COLLEGE DOCTORAL BRETAGNE — MATHS, TELECOMS, INFORMATIQUE, SIGNAL, SYSTEMES, ELECTRONIQUE

Université de Rennes

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE RENNES

ÉCOLE DOCTORALE N° 601
Mathématiques, Télécommunications, Informatique, Signal, Systèmes, Électronique
Spécialité : *Mathématiques et leurs interactions*

Par
Thomas CAVALLAZZI

Équations différentielles stochastiques dirigées par des bruits de Lévy : systèmes de particules en interaction de type champ moyen et processus de McKean-Vlasov

Thèse présentée et soutenue à Rennes, le 23 juin 2023
Unité de recherche : UMR CNRS 6625 Institut de Recherche Mathématique de Rennes (IRMAR)

Rapporteurs avant soutenance :

Benjamin	JOURDAIN	Professeur, École Nationale des Ponts et Chaussées
Arturo	KOHATSU-HIGA	Professeur, Ritsumeikan University

Composition du Jury :

Examinateurs :	François	BOLLEY	Professeur, ENS Rennes
	Arnaud	GUILLIN	Professeur, Université Clermont-Auvergne
	Benjamin	JOURDAIN	Professeur, École Nationale des Ponts et Chaussées
	Eva	LOCHERBACH	Professeure, Université Paris 1
	Milica	TOMASEVIC	Chargée de recherches CNRS, École Polytechnique
Dir. de thèse :	Mihai	GRADINARU	Professeur, Université de Rennes
Co-enc. de thèse :	Paul-Eric	CHAUDRU DE RAYNAL	Maître de Conférences, Université de Nantes

Remerciements

Me voilà arrivé au bout de mes trois années de thèse : quelle aventure ! Ces années ont été bien sûr très enrichissantes du point de vue mathématique, mais pas seulement. Regarder le chemin parcouru jusque-là me donne le vertige, et j'aimerais remercier toutes les personnes qui ont contribué à cela.

Mes premiers remerciements sont naturellement adressés à mes deux directeurs de thèse : Paul-Eric Chaudru de Raynal et Mihai Gradinaru, sans qui rien de tout cela n'aurait été possible. Mihai, je t'ai d'abord connu en tant qu'enseignant. C'est en grande partie grâce à ton excellent cours sur les chaînes de Markov et les martingales que j'ai eu envie d'explorer le monde extraordinaire des probabilités, et je ne regrette pas une seconde ! Je te remercie d'avoir éveillé cela en moi et de m'avoir donné la possibilité de faire ma thèse à Rennes. Merci pour ta bonne humeur, ton sourire et tous tes conseils. Paul-Eric, on ne se connaissait pas avant la thèse. J'ai eu beaucoup de chance de t'avoir comme encadrant ! Merci pour ta présence constante et tous tes conseils. On a passé un certain nombre d'heures à discuter et à réfléchir ensemble, tout cela entrecoupé de pauses café (tu n'auras toujours pas réussi à me convertir, quel dommage !). Travailler avec toi a toujours été très productif et agréable, même quand "ça ne marche pas (ou plus)"... Eh oui ! C'est ça la recherche ! Merci pour ton optimisme et de m'avoir donné un peu plus confiance en mon travail quand je doutais. Pour votre soutien à tous les deux, et pour tout le reste, je vous remercie chaleureusement.

Je tiens à remercier vivement Benjamin Jourdain et Arturo Kohatsu-Higa d'avoir accepté de rapporter ma thèse. Je vous suis très reconnaissant pour le temps que vous avez consacré à la lecture de mes travaux, et pour toutes vos remarques. Merci également à François Bolley, Arnaud Guillin, Eva Löcherbach et Milica Tomašević, qui me font l'honneur de faire partie de mon jury de thèse.

Le mathématicien que je suis aujourd'hui ne s'est évidemment pas construit tout seul : rien n'aurait été possible sans les professeurs que j'ai eus tout au long de ma scolarité. J'aimerais remercier tout d'abord mes professeurs de maths au lycée : Stéphane Brovelli, Christian Olivier et Valérie Rajot. Valérie, j'ai eu la chance de t'avoir en terminale et c'est grâce à toi que j'ai su que je voulais continuer dans les maths. Tu es une enseignante exceptionnelle, toujours à l'écoute de tes élèves et enthousiaste ! Ta pédagogie reste un modèle pour moi encore aujourd'hui. Enfin, merci pour ton soutien, qui n'est plus celui d'une professeure mais de l'amie que tu es devenue pour moi. Je tiens également à remercier mes professeurs de physique et de SVT : Thibaut Barbeau, Cécile Cazelles, Lise Munshi et Eric Perez. J'aimerais ensuite remercier les deux professeurs de maths que j'ai eu la chance d'avoir en prépa : Roger Mansuy et Nicolas Tosel. Vous êtes deux enseignants exceptionnels et des modèles pour l'enseignant que je suis. Merci pour votre bienveillance et votre humanité. Roger, merci d'avoir rendu la transition entre le lycée et la prépa aussi agréable. Nicolas, merci pour la vision des maths et le recul que tu m'as transmis, cela m'influence encore beaucoup aujourd'hui. Enfin, je remercie les enseignants qui m'ont formé à Rennes. Pour n'en citer que quelques-uns, merci à Jürgen Angst, Karine Beauchard, Benjamin Boutin, Jean-Christophe Breton, Arnaud Debussche, Thibaut Deheuvels, Jérémy Leborgne, Karel Pravda-Starov et Matthieu Romagny.

J'ai eu la chance d'enseigner à l'ENS Rennes pendant ces trois années. Je remercie les membres de l'équipe de l'ENS que j'ai toujours eu plaisir à croiser : Karine Beauchard, François Bolley, Arnaud Debussche, Magalie Fromont, Baptiste Huguet, Jérémy Leborgne, Salim Rostam et Rozenn Texier-Picard. Karine, merci pour la confiance que tu m'as accordée en me proposant ce monitorat. Travailler avec toi pour le cours d'EVNCD, après l'avoir apprécié en tant qu'élève, a été un très grand plaisir. Arnaud, merci pour tes conseils pour le post-doctorat. Magalie, j'ai été ravi de faire jury pour l'option A avec toi et je te remercie pour tes conseils. François, merci pour tes conseils et les discussions, mathématiques ou non, qu'on a eues durant ces années. Enfin, j'aimerais remercier tous les élèves que j'ai eus au cours de ces trois années. Cela a été un immense plaisir pour moi et une grande source de motivation.

Mon quotidien à l'IRMAR pendant la thèse a été très agréable et je remercie pour ça l'ensemble des personnes qui y travaillent. Merci en particulier à Eric Duvignac, Hélène Rousseaux et Marie-Aude Verger pour leur aide. J'ai eu l'immense chance de faire partie du bureau 232, partagé en première année avec François, Mercedes et Rémi. Pierre nous a ensuite rejoint lorsque Mercedes a terminé sa thèse. Mercedes, un grand merci pour ton accueil et tes conseils quand je suis arrivé à Rennes et tout au long de mon parcours. J'ai beaucoup apprécié organiser le TFJM^2 de Rennes avec toi ! François, Pierre et Rémi, je reviendrai vers vous par la suite... Je remercie enfin tous les doctorants que j'ai eu le plaisir de côtoyer dans cette aventure. J'aimerais remercier particulièrement Alice, Emilien, Hugo, Jérôme, Lisa, Marie, Martin, Isabelle, Maxime, Paul et Théo, et également Antoine, Emeline, Fabrice, Florian, Jérémy, Louis, Lucile et Mégane, qui ne sont plus à Rennes. Teoéii rrré ql m r i nst i v ore est-frosri sueacsazueu bsaersieanmst

J'aimerais remercier dans ce paragraphe des amis de tous horizons. Alice et Paul, je pense qu'on peut dire que vous êtes des membres invités du bureau 232 ! Paul, merci pour toutes les discussions et débats qu'on a pu avoir avec toi pendant ces années, et pour tout ton humour ! Alice, merci pour tous les bons moments passés avec nous. Les séances de badminton vont me manquer ! Je garderai en mémoire notre prouesse à rollers : de vrais riders ! Merci à Antoine, Bastien, Clarence, David et Maxence de m'avoir fait découvrir le festival Court Métrange. Antoine, on partage beaucoup de choses depuis notre arrivée à l'ENS et je voudrais te remercier pour ça. Tu es un ami que j'estime beaucoup. Merci pour ton soutien. Merci à Théo et Thibault, mes amis et fidèles camarades de l'option A pendant la prépa agreg, c'était un plaisir de travailler avec vous et de chercher le diable dans les détails ! J'aimerais également remercier d'autres amis de l'ENS : Alain, Clémentine et Emilie. C'est toujours un plaisir de vous revoir à diverses occasions ! Jérôme et Maxence, je suis heureux de vous avoir rencontrés. Merci pour les discussions et nos marches aléatoires au Thabor : les prochaines se feront ailleurs ! Julien et Sonia, mes compagnons de prépa, j'aimerais vous remercier d'avoir été là pendant ces années. J'espère qu'on aura l'occasion de se revoir plus souvent à l'avenir ! Théo, tu es un ami qui compte énormément. On a partagé beaucoup de choses. Je te remercie pour tout ça et pour ton soutien. Léa, que dire de notre amitié depuis toutes ces années ! On s'est toujours suivi et soutenu depuis tout ce temps et ce n'est pas prêt de s'arrêter ! Mille merci pour tout. Camille : on se suit depuis le collège et je trouve ça génial qu'on continue encore aujourd'hui ! Quel chemin parcouru : on peut être fier ! Merci d'être là depuis tout ce temps. Il y a enfin d'autres amis, toujours chers à mes yeux, mais que j'ai un peu perdu de vue : Antoine, Inès, Marine et My-Linh. J'espère que nos routes vont vite se recroiser !

Mégane, notre rencontre a été incroyable. Merci pour tous les moments inoubliables qu'on a vécus ensemble.

Emeline, je tiens à te remercier tout d'abord pour ton accueil chaleureux au début de ma thèse et pour les conseils que tu m'as donnés tout au long de cette aventure. J'ai adoré nos très nombreuses discussions et séances de travail : c'est tellement mieux de faire des maths à plusieurs au tableau ! J'ai trouvé en toi une amie géniale ! Enfin, merci pour ton soutien.

François, Pierre et Rémi, je ne sais pas par où commencer tant votre présence pendant ces trois années a été précieuse. Au début de l'histoire, il y a notre arrivée à l'ENS, il y a maintenant sept ans. La thèse nous a permis de nous rapprocher encore plus. Partager avec vous le bureau 232, bureau comptant le plus d'animaux dont notre mascotte Pies, a été exceptionnel. Bien sûr, on y a beaucoup travaillé, bercés parfois par un peu de musique. Discuter avec vous de maths, que ce soient des questions sur nos thèses respectives, sur nos enseignements ou autre, était tellement motivant et agréable. Mais il n'y a pas eu que ça, loin de là ! Il y a eu toutes nos pauses du midi suivies souvent de petites séances de jeux : le pedantix et ses compagnons nous auront bien amusés. On en aura appris des choses ! Bien sûr, on ne peut pas oublier toutes les discussions sur des sujets aussi bien sérieux qu'improbables, et tous les moments de rire (il est où notre toboggan en cas d'effondrement ! ?). Qu'est-ce qu'on s'est amusé à faire le court métrage pour Sciences en cour[t]s : mention spéciale à Alice pour sa précision hors pair quand il faut viser (c'est essentiel la localisation ...), à François pour son implacableuhhh raisonnement par transitivité, à Pierre pour avoir supporté notre incompétence pour accomplir la mission, à Raphaël dont on mesure les vastes connaissances en géopolitique extérieure sur la tribu borélienne, à Rémi pour son habileté stochastique pour ouvrir l'énorme coffre-fort, à Silvère pour avoir supporté le scotch, et à Théo pour l'ouverture grinçante de la porte. Notre amitié ne se résume pas qu'au bureau 232 : il y a eu tous les super moments passés ensemble dehors, que ce soient les soirées chez Pierre et Lisa ou au *Phare*, les soirées pizza, les balades, le volley, les séances films, les séances de sport aux Gayeulles (le défi a été largement relevé !), le badminton, et tout le reste. J'ai passé trois belles années à vos côtés, et je suis sûr que ça continuera à l'avenir, d'une manière ou d'une autre. Un immense merci à vous trois pour votre soutien constant dans les moments difficiles qu'il y a pu avoir, c'était tellement précieux. Bref, merci pour votre amitié exceptionnelle, vous êtes des piliers.

Enfin, j'aimerais remercier ma famille qui m'a toujours soutenu : Baba, Caroline, Claudia, François, Jade, Jean-Mi, Maman, Mamie, Mathieu, Olga, Papa, Philippe, Sylvie et tous les autres. J'ai également une pensée nostalgique pour Mamita et Papi. Merci également à Madame Callé et Serge, qui font presque partie de la famille. Claudia, merci pour les bons moments et les fous rires qu'on a eus en vacances. Sylvie, je suis content que tu aies rejoint la famille. Mamie, Caroline et Jean-Mi, merci pour tous les bons moments passés avec vous en famille autour de bons repas, et en Lozère pendant les vacances. Jade et François, merci de me faire rêver avec toutes vos aventures. Jade, tu es comme une sœur pour moi. Olga, merci pour ta joie de vivre, ta présence et ton soutien depuis toutes ces années. Nós nos entendemos sem precisar de falar. Philippe, merci de pouvoir compter sur toi quoi qu'il arrive. Baba, merci pour ton dévouement de toujours pour nous tous : on te doit tellement. Mathieu, merci d'être mon frère. Et enfin, Papa et Maman, merci pour toutes les valeurs que vous m'avez transmises depuis toujours et pour tout votre amour. La personne que je suis aujourd'hui vous doit énormément.

Table des matières

Résumé et notations

Cette thèse porte en grande partie sur l'étude des Équations Différentielles Stochastiques (EDS) de type McKean-Vlasov, qui sont de la forme

$$\begin{cases} dX_t = b(t, X_t, \mu_t)\, dt + \sigma(t, X_{t^-}, \mu_t)\, dZ_t, \quad t \geq 0, \\ \mu_t := [X_t], \\ X_0 = \xi, \end{cases} \tag{1}$$

où $[X_t]$ est la loi de X_t, $Z = (Z_t)_{t\geq 0}$ est un processus de Lévy indépendant de la donnée initiale ξ, $b : \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ et $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$, $\mathcal{P}(\mathbb{R}^d)$ étant l'espace des mesures de probabilité sur $\mathbb{R}^d$. Le système (1) est étroitement lié au système de N particules en interaction de type champ moyen

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\, dt + \sigma(t, X_{t^-}^{i,N}, \overline{\mu}_t^N)\, dZ_t^i, \quad t \geq 0, \quad i \in \{1, \dots, N\}, \\ \overline{\mu}_t^N := \frac{1}{N} \sum_{k=1}^{N} \delta_{X_t^{k,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{2}$$

où $(\xi^i, Z^i)_{i\geq 1}$ est une suite i.i.d. de même loi que (ξ, Z). Plus précisément, l'EDS de McKean-Vlasov (1) décrit la dynamique d'une particule quelconque du système (2) lorsque le nombre de particules N tend vers l'infini. De plus, il est également attendu que celles-ci deviennent indépendantes asymptotiquement. Il s'agit du phénomène de *propagation du chaos.*

Ce manuscrit présente les 5 travaux [Cav23], [Cav22a], [Cav22b], [Cav21] et [CL23], qui sont brièvement résumés dans la suite. L'article [CL23] a été accepté pour publication dans *Discrete and Continuous Dynamical Systems*. Les travaux [Cav23] et [Cav21] sont en révision respectivement à *Electronic Communications in Probability* et à *ESAIM: Probability and Statistics*. Les autres articles ont été soumis pour publication.

Dans le travail [Cav23], on s'intéresse au caractère bien posé de (1) sous des hypothèses lipschitziennes sur les coefficients b et σ. Il est supposé que le processus de Lévy Z admet un moment fini d'ordre $\beta \in [1, 2]$. Dans le cas où σ est constant égal à l'identité, on prouve des estimations classiques de propagation du chaos au niveau des trajectoires par couplage pour le système de particules (2). On améliore les taux de convergence obtenus dans le cas particulier d'un système de processus d'Ornstein-Uhlenbeck en interaction de type champ moyen dirigés par des processus α-stables, avec $\alpha \in (1, 2)$.

L'objectif principal des travaux [Cav22a, Cav22b] est de prouver des estimations de propagation du chaos faible pour l'EDS de McKean-Vlasov (1), où Z est un processus α-stable rotationnellement invariant, avec $\alpha \in (1, 2)$. On suppose que σ est constant égal à l'identité et que la dérive b est höldérienne par rapport aux variables espace et mesure en un sens précis. Il s'agit plus précisément d'exhiber des taux de convergence vers 0 pour $\mathbb{E}|\phi(\overline{\mu}_t^N) - \phi(\mu_t)|$ et $|\mathbb{E}(\phi(\overline{\mu}_t^N) - \phi(\mu_t))|$, lorsque N tend vers l'infini, ϕ étant dans une classe assez large de fonctions définies sur $\mathcal{P}_\beta(\mathbb{R}^d)$, l'espace des mesures de probabilité admettant un moment fini d'ordre β. Pour cela, on étudie dans [Cav22b] les propriétés régularisantes du semi-groupe, qui agit sur les fonctions définies sur $\mathcal{P}_\beta(\mathbb{R}^d)$, associé à l'EDS de McKean-Vlasov (1). On décrit ensuite la dynamique du semi-groupe au moyen d'une Équation aux Dérivées Partielles (EDP). Cela repose en grande partie sur la formule d'Itô le long d'un flot de mesures de probabilité qui permet de décrire la dynamique de l'application $t \in \mathbb{R}^+ \mapsto \phi(\mu_t)$. Celle-ci est prouvée dans [Cav22a] pour des flots de lois marginales de processus à sauts définis avec une mesure aléatoire de Poisson et une mesure de Poisson compensée. On utilise ensuite ces résultats pour prouver des estimations de propagation du chaos faible dans [Cav22b]. Dans le cas où la dérive b n'est pas bornée, on prouve dans [Cav22a] le même type d'estimations pour un système de processus d'Ornstein-Uhlenbeck α-stables en interaction de type champ moyen, avec $\alpha \in (1,2)$. Dans [Cav22b], on contrôle également ponctuellement l'erreur entre la densité d'une particule et celle de la solution de l'EDS de McKean-Vlasov limite, ce qui quantifie l'approximation de la loi d'une particule par la loi de l'EDS de McKean-Vlasov par rapport à la distance en variation totale.

Le travail [Cav21] porte sur la formule d'Itô le long du flot du lois marginales d'un processus d'Itô général de la forme

$$X_t := \int_0^t b_s \, ds + \int_0^t \sigma_s \, dB_s, \quad t \geq 0,$$

où B est un mouvement brownien et où les processus b et σ sont bornés, avec de plus σ uniformément elliptique. Le but précis est d'affaiblir les hypothèses de régularité sur la fonction à laquelle on veut appliquer la formule d'Itô le long d'un flot de mesures, au prix d'hypothèses d'intégrabilité. Cela a été fait pour la formule d'Itô classique par Krylov dans [Kry09, Partie 2.10] en exploitant l'effet régularisant du mouvement brownien grâce à l'ellipticité du coefficient de diffusion σ. Cet effet se traduit ici par l'existence d'une densité pour la loi de X_t, pour presque tout $t \geq 0$, qui appartient à un certain espace de Lebesgue. Ainsi, la formule d'Itô-Krylov pour un flot de mesures qui est établie dans [Cav21] peut être appliquée à des fonctions appartenant à un espace de type Sobolev de fonctions définies sur l'espace des mesures de probabilité.

Enfin, dans le travail [CL23], réalisé un collaboration avec Émeline Luirard, on ne s'intéresse plus au comportement d'un système de particules quand le nombre de particules tend vers l'infini, mais au comportement en temps long d'une seule particule dont la dynamique est détaillée dans ce qui suit. On considère le système cinétique inhomogène en temps d'EDS défini, sur l'intervalle de temps $[t_0, +\infty)$, où $t_0 > 0$, par

$$\begin{cases} dV_t = -\operatorname{sgn}(V_t) \dfrac{|V_t|^\gamma}{t^\beta} \, dt - X_t \, dt + dZ_t, \\ dX_t = V_t \, dt, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0), \end{cases}$$

où $\beta, \gamma > 0$ et $Z = (Z_t)_{t \geq 0}$ est soit un mouvement brownien, soit un processus α-stable, avec $\alpha \in (1,2)$. Cela décrit la dynamique d'une particule évoluant dans un potentiel quadratique et soumise à une force de frottement inhomogène en temps, et qui tend vers 0 quand le temps tend vers l'infini. Différents comportements asymptotiques sont établis en fonction des paramètres β, γ et α. Ils apparaissent à travers la convergence du processus vitesse-position, correctement changé d'échelle en temps et en espace.

Le manuscrit est structuré de la manière suivante.

- La Partie I est une introduction générale aux travaux présentés dans ce manuscrit. Cette introduction s'articule autour de deux chapitres. Dans le Chapitre 1, on présente les processus de Lévy, notamment les processus α-stables, puis les EDS de type McKean-Vlasov, ainsi que leurs liens avec les systèmes de particules en interaction de type champ moyen. On introduit précisément la notion de propagation du chaos au sens faible et fort. Enfin, on décrit heuristiquement la méthode reposant sur le semi-groupe qui est utilisée dans les articles [Cav22a] et [Cav22b] et qui permet de quantifier la propagation du chaos au sens faible. Dans le Chapitre 2, on présente les outils essentiels pour mettre en œuvre la méthode mentionnée précédemment : la notion de dérivation pour des fonctions définies sur un espace de mesures, celle de projection empirique et enfin la formule d'Itô le long d'un flot de mesures.

- La Partie II est un résumé des résultats obtenus au cours de la thèse. On présente les résultats démontrés, on les commente et les compare aux travaux déjà existants et on donne également des idées de preuves tout en discutant des hypothèses. Dans le Chapitre 3, on présente les travaux [Cav23], [Cav22a] et [Cav22b], qui traitent des EDS de McKean-Vlasov dirigées par des bruits de Lévy, principalement des processus α-stables, avec $\alpha \in (1, 2)$. Dans les Chapitres 4 et 5, on présente respectivement les travaux [Cav21] et [CL23].

- La Partie III rassemble les articles [Cav23], [Cav22a], [Cav22b], [Cav21] et [CL23]. Ils correspondent respectivement aux Chapitres 6, 7, 8, 9 et 10.

Notations. On rassemble ici des notations utilisées dans l'ensemble du manuscrit.

- $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ désigne un espace probabilisé muni d'une filtration satisfaisant les conditions usuelles.
- $\mathcal{P}(\mathbb{R}^d)$ est l'espace des mesures de probabilité sur $\mathbb{R}^d$.
- d_{TV} est la distance en variation totale sur $\mathcal{P}(\mathbb{R}^d)$.
- $\mathcal{P}_\beta(\mathbb{R}^d)$ est l'ensemble des mesures $\mu \in \mathcal{P}(\mathbb{R}^d)$ telles que $\int_{\mathbb{R}^d} |x|^\beta \, d\mu(x) < +\infty$, pour $\beta \geq 0$.
- W_β est la distance de Wasserstein d'ordre β sur $\mathcal{P}_\beta(\mathbb{R}^d)$, pour $\beta > 0$.
- $[\xi]$ désigne la loi d'une variable aléatoire ξ.
- δ_x est la masse de Dirac en $x \in \mathbb{R}^d$.
- $\overline{\mu}^N_{\boldsymbol{x}} := \frac{1}{N}\sum_{k=1}^N \delta_{x_k}$ est la mesure empirique associée à un vecteur $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$.
- $\tilde{\boldsymbol{z}}_{\boldsymbol{k}} := (0, \ldots, z, \ldots, 0) \in (\mathbb{R}^d)^N$, pour $z \in \mathbb{R}^d$, où z apparaît en k-ème position.
- B_r est la boule ouverte de $\mathbb{R}^d$, pour la norme euclidienne, centrée en 0 et de rayon $r > 0$.
- B_r^c désigne le complémentaire de la boule B_r dans $\mathbb{R}^d$.
- $a \wedge b$ est le minimum entre deux réels a et b.
- $a \vee b$ est le maximum entre deux réels a et b.
- sgn désigne la fonction signe sur $\mathbb{R}$ avec par convention $\mathrm{sgn}(0) = 0$.

- p' est l'exposant conjugué de $p \in [1, +\infty]$, défini par $\frac{1}{p} + \frac{1}{p'} = 1$.
- $L^p_{\text{loc}}(\mathbb{R}^d)$ est l'espace des fonctions mesurables f telles que pour tout $R > 0$, $f \in L^p(B_R)$.
- $W^{m,k}(\mathcal{O})$ est l'espace de Sobolev des fonctions appartenant à $L^k(\mathcal{O})$ et admettant des dérivées au sens des distributions d'ordre 1 à m dans $L^k(\mathcal{O})$, où $\mathcal{O}$ est un ouvert de $\mathbb{R}^d$. Il est muni de la norme suivante, définie pour $u \in W^{m,k}(\mathcal{O})$, par

$$\|u\|_{W^{m,k}(\mathcal{O})} := \sum_{\alpha \in \mathbb{N}^d,\, |\alpha| \leq m} \|\partial^\alpha u\|_{L^k(\mathcal{O})}.$$

- $W^{m,k}_{\text{loc}}(\mathbb{R}^d)$ est l'ensemble des fonctions u telles que pour tout $R > 0$, u appartienne à $W^{m,k}(B_R)$.
- $\mathbf{1}_A$ désigne la fonction indicatrice d'un ensemble A.
- $*$ désigne le produit de convolution entre deux fonctions ou mesures de probabilité lorsqu'il est bien défini.
- $\mathbb{R}^{d \times d}$ désigne l'ensemble des matrices sur $\mathbb{R}$ de taille $d \times d$.
- A^* est la transposée de la matrice $A \in \mathbb{R}^{d \times d}$.
- $A \cdot B$ est le produit scalaire usuel entre deux matrices $A, B \in \mathbb{R}^{d \times d}$, défini par $A \cdot B := \mathrm{Tr}(A^* B)$.
- $(x, y)^T$ désigne le vecteur $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, pour $x, y \in \mathbb{R}$.
- C est une constante générique qui ne dépend que des paramètres fixes du problème considéré et qui peut changer d'une ligne à l'autre.

Première partie

Introduction : des concepts aux outils

Chapitre 1

Des processus de Lévy aux équations différentielles stochastiques de McKean-Vlasov

1.1 Processus de Lévy

Les *processus de Lévy*, et plus particulièrement les *processus stables*, sont au cœur de ce manuscrit. Cette section vise à les introduire et à en donner quelques propriétés utiles dans la suite.

1.1.1 Motivation et définition

Les processus de Lévy forment une classe de processus aléatoires en temps continu fondamentale dans le domaine des probabilités. Outre leur rôle important en mathématiques, ils sont largement utilisés dans d'autres disciplines, où on les connait généralement sous le nom de *Lévy flights*. En physique, ils sont fondamentalement liés aux diffusions anormales (super-diffusion), qui sont des types de diffusions plus rapides décrites par des lois de puissances reliant le déplacement moyen au temps. On pourra consulter par exemple [MJCB14] pour une vue d'ensemble sur des modèles physiques de diffusions anormales. Les processus de Lévy apparaissent naturellement en biologie pour décrire certains phénomènes, par exemple les mouvements exploratoires d'êtres vivants comme des animaux ou encore des cellules. Ils sont également largement utilisés en finance. Donnons la définition mathématique des processus de Lévy ainsi que quelques-unes de leurs propriétés essentielles.

Définition 1.1 (Processus de Lévy)**.** Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Un processus stochastique $Z = (Z_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^d$ est un *processus de Lévy* s'il satisfait les conditions suivantes.

- Les trajectoires de Z sont presque sûrement continues à droite avec des limites à gauche (càdlàg) et partent initialement de 0.
- Les accroissements de Z sont stationnaires, c'est-à-dire que pour tous réels $0 \leq s \leq t$, $Z_t - Z_s$ a la même loi que Z_{t-s}.
- Les accroissements de Z sont indépendants, c'est-à-dire que pour tous réels $0 \leq t_1 \leq \cdots \leq t_l$, les variables aléatoires $Z_{t_1}, Z_{t_2} - Z_{t_1}, \ldots, Z_{t_l} - Z_{t_{l-1}}$ sont mutuellement indépendantes.

On peut penser les processus de Lévy comme des marches aléatoires en temps continu puisque leurs accroissements sont indépendants et stationnaires. Ils incluent de nombreux processus stochastiques fondamentaux comme le mouvement brownien, les processus de Poisson et les processus stables, sur lesquels nous reviendrons dans la suite.

1.1.2 Sauts d'un processus de Lévy

Puisque les trajectoires des processus de Lévy ne sont pas nécessairement continues, des sauts peuvent avoir lieu à des instants aléatoires. La description détaillée des sauts d'un processus de Lévy se fait à travers la notion de mesure aléatoire de Poisson.

Définition 1.2 (Mesure aléatoire de Poisson)**.** Soit ν une mesure σ-finie sur $\mathbb{R}^d\backslash\{0\}$. Une mesure aléatoire ponctuelle $\mathcal{N}$ sur $\mathbb{R}^+ \times \mathbb{R}^d\backslash\{0\}$ est une *mesure aléatoire de Poisson* avec intensité $dt \otimes \nu$ si elle vérifie les propriétés suivantes.

- Pour tout $t \geq 0$ et pour tout borélien A de $\mathbb{R}^d\backslash\{0\}$ tel que $\nu(A) < +\infty$, la variable aléatoire $\mathcal{N}([0,t] \times A)$ suit une loi de Poisson de paramètre $t\nu(A)$.
- Pour toute famille de boréliens $(Q_i)_i$ de $\mathbb{R}^+ \times \mathbb{R}^d$, les variables aléatoires $(\mathcal{N}(Q_i))_i$ sont indépendantes.

Le lien entre un processus de Lévy $Z = (Z_t)_{t\geq 0}$ et la notion de mesure aléatoire de Poisson se fait à travers l'étude du processus de sauts défini, pour $t \geq 0$, par $\Delta Z_t := Z_t - Z_{t^-}$, où Z_{t-} désigne la limite à gauche de Z au point t. On définit alors, pour tout borélien A de $\mathbb{R}^d\backslash\{0\}$

$$\mathcal{N}([0,t] \times A) := \mathrm{Card}\left\{s \in [0,t],\, \Delta Z_s \in A\right\}.$$

Ainsi, $\mathcal{N}([0,t] \times A)$ compte le nombre de sauts de Z à valeurs dans A et ayant lieu avant l'instant t. En utilisant [App09, Théorème 2.3.5], on vérifie que la mesure aléatoire $\mathcal{N}$ ainsi définie est une mesure aléatoire de Poisson, d'intensité ν donnée, pour tout borélien A de $\mathbb{R}^d\backslash\{0\}$, par

$$\nu(A) := \mathbb{E}\mathcal{N}([0,1] \times A).$$

On peut également vérifier que la mesure ν est une *mesure de Lévy* au sens de la définition suivante.

Définition 1.3 (Mesure de Lévy)**.** Une mesure σ-finie ν sur $\mathbb{R}^d$ est une *mesure de Lévy* si elle vérifie :

$$\nu(\{0\}) = 0 \quad \text{et} \quad \int_{\mathbb{R}^d} 1 \wedge |z|^2\, d\nu(z) < +\infty.$$

La mesure aléatoire de Poisson $\mathcal{N}$ associée au processus de Lévy Z, ou de manière équivalente sa mesure de Lévy ν, encodent toutes les informations sur les sauts de Z. On peut montrer que si A est un borélien de $\mathbb{R}^d$ borné inférieurement, alors le processus $(\mathcal{N}([0,t] \times A))_{t\geq 0}$, qui compte les sauts de Z à valeurs dans A, est un processus de Poisson d'intensité $\nu(A)$. Ainsi plus A est chargé par ν, plus le processus a tendance à avoir des sauts à valeurs dans A. On peut facilement associer à ce processus une martingale donnée par la définition suivante.

Définition 1.4 (Mesure aléatoire de Poisson compensée)**.** Soit $\mathcal{N}$ une mesure aléatoire de Poisson sur $\mathbb{R}^+ \times \mathbb{R}^d\backslash\{0\}$ d'intensité $dt \otimes \nu$. Pour tout borélien A de $\mathbb{R}^d\backslash\{0\}$ borné inférieurement (i.e. inclus dans le complémentaire d'une boule de rayon strictement positif), on pose, pour $t \geq 0$,

$$\widetilde{\mathcal{N}}([0,t] \times A) := \mathcal{N}([0,t] \times A) - t\nu(A).$$

$\widetilde{\mathcal{N}}$ est appelée *mesure aléatoire de Poisson compensée.*

Le processus $(\widetilde{\mathcal{N}}([0,t] \times A))_{t\geq 0}$ est un processus de Poisson compensé et il s'agit d'une martingale. On dit que $\widetilde{\mathcal{N}}$ est une mesure à valeurs martingales. On renvoie à [App09, Section 4] pour la définition des intégrales stochastiques contre $\mathcal{N}$ et $\widetilde{\mathcal{N}}$. Mentionnons simplement que l'intégrale stochastique contre la mesure compensée $\widetilde{\mathcal{N}}$ est assez semblable à l'intégrale contre le mouvement brownien et définit une martingale locale.

1.1.3 Décomposition de Lévy-Itô et symbole

À partir de la mesure aléatoire $\mathcal{N}$ associée au processus de Lévy Z, on peut reconstruire entièrement ce dernier grâce à la décomposition de Lévy-Itô [App09, Théorème 2.4.16].

Théorème 1.5. *[Décomposition de Lévy-Itô] Soit $Z = (Z_t)_{t\geq 0}$ un processus de Lévy à valeurs dans $\mathbb{R}^d$. Alors, Z se décompose de la manière suivante :*

$$Z_t = tb + \sigma W_t + \int_0^t \int_{B_1} z\widetilde{\mathcal{N}}(ds, dz) + \int_0^t \int_{B_1^c} z\mathcal{N}(ds, dz), \tag{1.1}$$

où $b \in \mathbb{R}^d$, $\sigma \in \mathbb{R}^{d\times d}$ est une matrice symétrique positive et W est un mouvement brownien standard sur $\mathbb{R}^d$, indépendant de la mesure de Poisson $\mathcal{N}$. Le triplet (b, A, ν) est déterminé, de manière unique, par le processus de Lévy Z.

Réciproquement, tout processus de cette forme est un processus de Lévy.

Les deux premiers termes de (1.1) représentent la partie continue de Z, il s'agit d'un mouvement brownien avec une dérive déterministe. Les deux derniers termes sont des termes de sauts : le premier porte sur les petits sauts et le second sur les grands sauts. Il y a presque sûrement un nombre fini de grands sauts sur chaque intervalle de temps borné et ils déterminent la finitude des moments de Z. Quant aux deux termes centraux de (1.1), ils forment la partie martingale de Z.

Mentionnons également la forme particulière de la fonction caractéristique de Z_t qui se déduit de la décomposition de Lévy-Itô.

Proposition 1.6. *Pour tout $t \geq 0$, la fonction caractéristique de Z_t est de la forme $\xi \in \mathbb{R}^d \mapsto e^{t\psi(\xi)}$ avec, pour tout $\xi \in \mathbb{R}^d$,*

$$\psi(\xi) := ib \cdot \xi - \frac{1}{2}(\sigma\sigma^* \xi \cdot \xi) + \int_{\mathbb{R}^d} e^{iz\cdot\xi} - 1 - iz \cdot \xi \mathbf{1}_{|z|<1}\, d\nu(y),$$

où ν est la mesure de Lévy associée à Z. La fonction ψ est appelée exposant caractéristique ou encore symbole de Z.

1.1.4 Semi-groupe et générateur

Les processus de Lévy vérifient la propriété de Markov forte grâce à l'indépendance et la stationnarité de leurs accroissements. Cela permet de leur associer un semi-groupe d'opérateurs de Feller (voir [App09, Théorème 3.1.9]).

Proposition 1.7 (Semi-groupe)**.** *Soit $f \in \mathcal{C}_0^0(\mathbb{R}^d; \mathbb{R})$, i.e. une fonction continue qui tend vers 0 en $+\infty$. On définit alors, pour tout $t \geq 0$,*

$$T_t f : x \in \mathbb{R}^d \mapsto T_t f(x) := \mathbb{E} f(x + Z_t).$$

La famille d'opérateurs linéaires $(T_t)_{t\geq 0}$ définit un semi-groupe de Feller au sens suivant.

- *Pour tout $t \geq 0$, $\mathcal{C}_0^0(\mathbb{R}^d;\mathbb{R})$ est stable par T_t.*
- *La propriété de semi-groupe est vérifiée, i.e. pour tous $s,t \geq 0$, $T_t \circ T_s = T_{t+s}$.*
- *Le semi-groupe est fortement continu, au sens où pour tout $f \in \mathcal{C}_0^0(\mathbb{R}^d;\mathbb{R})$*

$$\|T_t f - f\|_\infty \underset{t\to 0}{\longrightarrow} 0.$$

Au semi-groupe ainsi construit est associé son générateur infinitésimal $\mathcal{L}$ qui décrit la dynamique du semi-groupe.

Définition 1.8 (Générateur infinitésimal)**.** On définit l'opérateur non-borné L dont le domaine est l'ensemble des fonctions $f \in \mathcal{C}_0(\mathbb{R}^d;\mathbb{R})$ telle que $(T_t f - f)/t$ admette une limite, quand $t \to 0^+$, dans $\mathcal{C}_0(\mathbb{R}^d;\mathbb{R})$ muni de $\|\cdot\|_\infty$, et on note cette limite Lf.

Pour toute fonction f appartenant au domaine de L, par la propriété de semi-groupe, $T_t f$ est solution du problème de Cauchy

$$\begin{cases} y'(t) = Ly(t), & t \geq 0, \\ y(0) = f. \end{cases}$$

La forme du générateur est explicitée grâce à la décomposition de Lévy-Itô (1.1).

Proposition 1.9 (Générateur d'un processus de Lévy)**.** *Toute fonction $f \in \mathcal{C}_0^2(\mathbb{R}^d;\mathbb{R})$, i.e. f, ∇f et $\nabla^2 f$ sont dans $\mathcal{C}_0^0(\mathbb{R}^d)$, appartient au domaine de L et on a, pour tout $x \in \mathbb{R}^d$,*

$$Lf(x) = b \cdot \nabla f(x) + \frac{1}{2}\sum_{i,j=1}^d [\sigma\sigma^*]_{ij}\partial_{x_i,x_j} f(x) + \int_{\mathbb{R}^d} f(x+z) - f(x) - \nabla f(x)\cdot z\mathbf{1}_{|z|<1}\, d\nu(z).$$

On renvoie à [Sat99, Théorème 31.5] pour la preuve de ce résultat.

1.1.5 Processus α-stables

Les processus α-stables sont une classe importante de processus de Lévy et sont au cœur de cette thèse.

Définition 1.10 (Processus α-stable)**.** Un processus de Lévy Z est un *processus α-stable*, avec $\alpha \in (0,2]$, si pour tout $c > 0$, $(c^{1/\alpha} Z_{t/c})_{t\geq 0}$ a la même loi que $(Z_t)_{t\geq 0}$.

Un processus α-stable est donc défini par sa propriété d'auto-similarité. Cette propriété est en particulier centrale dans le travail [CL23], où le comportement asymptotique d'un système cinétique amorti et perturbé par un processus α-stable est étudié. Cela se fait à travers des changements d'échelle en temps et en espace, qui sont liés en partie à la propriété d'auto-similarité du bruit.

Lorsque $\alpha = 2$, on reconnait la propriété d'auto-similarité vérifiée par le mouvement brownien et on peut montrer que Z est effectivement un mouvement brownien (voir [Sat99, Théorème 14.2]). Lorsque $\alpha \in (0,2)$, un processus α-stable possède des sauts. Comme on l'a vu, ceux-ci sont décrits par une mesure de Lévy qui possède, d'après [Sat99, Théorème 14.3], la forme générique décrite dans la proposition suivante.

Proposition 1.11. *Soit Z un processus de Lévy α-stable, avec $\alpha \in (0,2)$. Alors il existe une mesure finie non nulle λ sur la sphère unité $\mathbb{S}^{d-1}$ de $\mathbb{R}^d$, telle que ν s'écrive en coordonnées sphériques $z = (r,\theta) \in \mathbb{R}_*^+ \times \mathbb{S}^{d-1}$*

$$\nu(dz) = \lambda(d\theta)\frac{dr}{r^{1+\alpha}}.$$

La mesure λ décrit la répartition des sauts dans les différentes directions de l'espace.

Proposition 1.12. *Soit $Z = (Z_t)_{t\geq 0}$ un processus α-stable avec $\alpha \in (0,2)$.*

- *On a l'équivalence*

$$\forall t \geq 0,\, \mathbb{E}|Z_t|^\beta < +\infty \Leftrightarrow \mathbb{E}|Z_1|^\beta < +\infty \Leftrightarrow \int_{B_1^c} |z|^\beta \, d\nu(z) < +\infty \Leftrightarrow \beta \in [0,\alpha).$$

  De plus, on a pour tout $t \geq 0$

$$\mathbb{E}|Z_t|^\beta = t^{\frac{\beta}{\alpha}}\mathbb{E}|Z_1|^\beta. \tag{1.2}$$

- *Si $\alpha > 1$, alors, pour tout $t \geq 0$, on a*

$$Z_t = \int_0^t \int_{\mathbb{R}^d} z\,\widetilde{\mathcal{N}}(ds,dz).$$

  Il s'agit donc d'une martingale centrée.

Pour la preuve de l'équivalence du premier point, on renvoie à [Sat99, Théorème 25.3]. L'expression (1.2) découle immédiatement de la propriété d'auto-similarité définissant un processus α-stable. La preuve du second point découle essentiellement de la décomposition de Lévy-Itô (voir [Sat99, Remarque 14.6 et Théorème 14.7]).

Définition 1.13 (Processus α-stable rotationnellement invariant)**.** On fixe $\alpha \in (0,2)$. Un processus α-stable Z sur $\mathbb{R}^d$ est dit *rotationnellement invariant* s'il existe $C > 0$ tel que

$$\nu(dz) = C\frac{dz}{|z|^{d+\alpha}}.$$

Ainsi, les sauts d'un processus stable rotationnellement invariants se répartissent uniformément dans toutes les directions de l'espace. Dans la suite, on prendra $C = 1$ pour simplifier. Notons que le générateur du processus α-stable rotationnellement invariant est donné, pour toute fonction f suffisamment régulière, par

$$\Delta^{\frac{\alpha}{2}} f(x) := \int_{\mathbb{R}^d} f(x+z) - f(x) - \nabla f(x)\cdot z\,\frac{dz}{|z|^{d+\alpha}}, \quad x \in \mathbb{R}^d. \tag{1.3}$$

Il s'agit du Laplacien fractionnaire. On peut le définir pour $f \in \mathcal{C}_b^{1+\gamma}(\mathbb{R}^d;\mathbb{R})$ avec $\gamma \in (\alpha-1,1]$ (i.e. f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^d$ et ∇f est bornée γ-höldérienne). Cela vient du fait que

$$\int_{B_1} |z|^\gamma \, d\nu(z) < +\infty \Leftrightarrow \gamma \in (\alpha-1,1].$$

La proposition suivante garantit l'existence d'une densité pour un processus stable rotationnellement invariant et fournit des estimations de gradient sur celle-ci.

Proposition 1.14. *Soit Z un processus α-stable rotationnellement invariant sur $\mathbb{R}^d$. Alors pour tout $t > 0$, Z_t admet une densité, notée $q(t, \cdot)$, par rapport à la mesure de Lebesgue et qui appartient à $\mathcal{C}^\infty(\mathbb{R}^d; \mathbb{R}^+)$. Introduisons, pour $j \geq 0$, la fonction ρ^j, définie sur $(0, +\infty) \times \mathbb{R}^d$, par*

$$\rho^j(t, x) := t^{-\frac{d}{\alpha}}(1 + t^{-\frac{1}{\alpha}}|x|)^{-d-\alpha-j}, \quad t > 0,\, x \in \mathbb{R}^d.$$

Alors on a les estimations de gradient suivantes : pour tout $j \in \mathbb{N}$, il existe une constante $C_j > 0$ telle que pour tout $t > 0$ et $y \in \mathbb{R}^d$

$$|\partial_y^j q(t, y)| \leq C_j t^{-\frac{j}{\alpha}} \rho^j(t, y).$$

La preuve de cette proposition se trouve dans [MZ22, Lemme 2.8]. Ces estimations sont l'un des éléments clés du travail [Cav22b], comme on le verra dans la suite.

1.2 Équations différentielles stochastiques de McKean-Vlasov et système de particules en interaction champ moyen

La majeure partie de cette thèse est consacrée aux EDS de type McKean-Vlasov et aux systèmes de particules en interaction de type champ moyen associés. Nous présentons ces deux notions et leurs liens dans cette section.

1.2.1 Équations différentielles stochastiques de McKean-Vlasov

Une EDS de type McKean-Vlasov est de la forme générale suivante

$$\begin{cases} dX_t = b(t, X_t, \mu_t)\, dt + \sigma(t, X_{t-}, \mu_t)\, dZ_t, \quad t \geq 0, \\ \mu_t := [X_t], \\ X_0 = \xi, \end{cases} \tag{1.4}$$

où $[X_t]$ désigne la loi de X_t et $Z = (Z_t)_{t \geq 0}$ est le bruit directeur de l'équation. Les coefficients b et σ sont définis sur $\mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, et à valeurs respectivement dans $\mathbb{R}^d$ et $\mathbb{R}^{d \times d}$. L'étude de ces équations a commencé naturellement par le cas où Z est un mouvement brownien. Elles ont été étudiées pour la première fois par McKean [McK66]. Celui-ci s'est intéressé à l'équation de Vlasov, qui décrit la distribution des particules chargées dans un plasma. La différence entre une EDS classique et une EDS de McKean-Vlasov réside dans le fait que les coefficients b et σ de l'équation dépendent non seulement de la position de la solution, mais également de sa loi. Les EDS de McKean-Vlasov sont également appelées EDS non-linéaires. Cela s'explique par le fait que, contrairement à une EDS classique, le flot des lois marginales $(\mu_t)_{t \geq 0}$ de la solution de (1.4) est solution d'une équation de Fokker-Planck non-linéaire. En effet, dans le cas où Z est un mouvement brownien standard et où $\sigma = \mathrm{Id}$, l'équation de Fokker-Planck associée est

$$\partial_t \mu_t = \frac{1}{2}\Delta \mu_t - \nabla \cdot (\mu_t\, b(t, \cdot, \mu_t)), \quad t \geq 0, \tag{1.5}$$

qui décrit une diffusion non-linéaire.

La première question qu'on peut se poser concerne le caractère bien posé de (1.4). Par bien posé, on entend l'existence et l'unicité des solutions, qui peuvent être considérées au sens fort ou au sens faible.

Dans le cas des EDS linéaires, i.e. quand les coefficients b et σ ne dépendent pas de la loi, il est bien connu que le bruit permet de montrer le caractère bien posé, sous des hypothèses de régularité en espace plus faible que lipschitzienne. Il s'agit du phénomène de *régularisation par le bruit*. On renvoie à [SV79], [Zvo74], [Ver81] et [Fla11] pour plus de détails à ce sujet. La dépendance des coefficients par rapport à la loi dans (1.4) est plus complexe à traiter, comme illustré dans la suite. On sépare la discussion selon que le bruit Z est un mouvement brownien ou un processus de Lévy plus général, et on se concentre sur la variable mesure des coefficients.

Bruit brownien

Donnons quelques résultats sur le caractère bien posé de (1.4), dans le cas où Z est un mouvement brownien.

- **Sans bruit ou avec un bruit qui peut dégénérer.** Le choix de la distance utilisée sur l'espaces des mesures est crucial, comme illustré dans ce qui suit, puisque les distances ne sont pas équivalentes.

  Si les coefficients b et σ sont globalement lipschitziens sur $\mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d)$, avec $p \geq 2$, localement uniformément en temps, alors il existe une unique solution forte à (1.4). On renvoie à [Szn91, Theorème 1.1] et [CD18a, Theorème 4.21] pour des preuves dans le cas $p = 2$. Ici, on travaille avec la distance de Wasserstein W_p, qui est contrôlée par la distance L^p sur les trajectoires. C'est ce qui permet de prouver le caractère bien posé dans ce cadre.

  Le contre-exemple de Scheutzow [Sch87], détaillé dans ce qui suit, montre que si on considère cette fois la distance en variation totale sur $\mathcal{P}(\mathbb{R}^d)$, alors on perd le caractère bien posé sous des hypothèses lipschitziennes et sans bruit. On n'a donc pas de théorème de Cauchy-Lipschitz pour la distance en variation totale. On se place en dimension $d = 1$, avec $\sigma = 0$, et on pose, pour $t \in \mathbb{R}^+$, $x \in \mathbb{R}$, et $\mu \in \mathcal{P}(\mathbb{R}^d)$,
  $$b(t, x, \mu) := \int_{\mathbb{R}} \tilde{b}(y)\, d\mu(y),$$
  avec $\tilde{b} : \mathbb{R} \to \mathbb{R}$ bornée et localement lipschitzienne. Alors, l'EDS de McKean-Vlasov (1.4) peut admettre plusieurs solutions. Ici, la dérive b est bornée et lipschitzienne par rapport à la distance en variation totale. Cela montre que le choix de la distance sur $\mathcal{P}(\mathbb{R}^d)$ fait dans les hypothèses lipschitziennes est crucial car il influence le caractère bien posé de (1.4).

- **Avec un bruit qui ne dégénère pas.** On peut se demander ce qui se passe dans le même cadre que le contre-exemple précédent, mais cette fois en présence de bruit et avec $\sigma = 1$. En supposant $\tilde{b}$ seulement mesurable et bornée, Shiga et Tanaka ont prouvé dans [ST85] l'unicité trajectorielle pour (1.4). Le bruit permet donc de restaurer l'unicité qui fait défaut dans le contre-exemple de Scheutzow [Sch87]. Il s'agit là encore d'un phénomène de *régularisation par le bruit*. De plus, ce résultat a été étendu par Jourdain [Jou97] dans le cas où la dérive b, définie sur $\mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, est bornée, lipschitzienne en variation totale par rapport à la mesure, et où $\sigma = \mathrm{Id}$. Le bruit permet donc ici d'établir le caractère bien posé de (1.4), sous des hypothèses lipschitziennes relatives à une plus grande classe de distances utilisées, par rapport à la variable mesure.

  Mentionnons d'autres travaux plus récents qui tirent profit du phénomène de régularisation par le bruit pour des EDS de McKean-Vlasov. On peut citer par exemple Flandoli, Issoglio et Russo [FIR17], Mishura et Veretennikov [MV21], Röckner et Zhang [RZ21], Chaudru de Raynal et Frikha [CdRF22] et Chaudru de Raynal, Jabir et Menozzi [CdRJM23].

Contrairement à la variable spatiale, se passer du caractère lipschitzien par rapport à l'argument mesure, relativement à une certaine distance, semble très compliqué même en présence de bruit. On peut l'illustrer à travers le contre-exemple suivant, dû à François Delarue. Si on considère (1.4) avec

$$\xi = 0, \quad \sigma = \mathrm{Id}, \quad \text{et} \quad b(t,x,\mu) = \tilde{b}\left(\int_{\mathbb{R}^d} y\, d\mu(y)\right),$$

où $\tilde{b}$ est hölderienne et bornée. Alors, la fonction b est hölderienne par rapport à W_1, mais l'EDS (1.4) peut admettre plusieurs solutions. En effet, si X est solution, alors on a, pour tout $t \geq 0$,

$$X_t = \int_0^t \tilde{b}(\mathbb{E}X_s)\, ds + Z_t.$$

En prenant l'espérance, on remarque que l'application $t \in \mathbb{R}^+ \mapsto \mathbb{E}(X_t)$ est solution de $y' = \tilde{b}(y)$, qui peut avoir plusieurs solutions puisque $\tilde{b}$ est seulement höldérienne.

Bruit de Lévy

Intéressons-nous maintenant au caractère bien posé de l'EDS de McKean-Vlasov (1.4) lorsque Z est un processus de Lévy général.

- **Avec un bruit qui peut dégénérer.** On commence par s'intéresser au cas où Z admet un moment d'ordre 2 fini. Jourdain, Méléard et Woyczynski ont démontré dans [JMW07] que (1.4) admet une unique solution forte sous des hypothèses lipschitziennes sur b et σ, par rapport à la distance de Wasserstein W_2 pour l'argument mesure. On peut mentionner également, toujours dans le cadre L^2, le travail [NBK+20], où les auteurs affaiblissent les hypothèses sur b et σ pour assurer l'existence et l'unicité d'une solution forte. En particulier, leur croissance n'est plus nécessairement linéaire.

  Dans le cas où Z admet seulement un moment d'ordre 1 fini, toujours sous des hypothèses lipschitziennes, par rapport à W_1 ici, on peut renvoyer à Graham [Gra92b]. L'intensité des sauts du bruit $\nu(\mathbb{R}^d)$, où ν est la mesure de Lévy de Z, est supposée finie dans ce travail. Cette condition n'est pas vérifiée pour les processus α-stables. Dans [Gra92a], il n'est plus supposé que $\nu(\mathbb{R}^d)$ est finie. Cependant, un point technique utilisé dans la preuve ne parait pas immédiat à justifier (voir la Remarque 3.3). C'est pour cela qu'on s'intéresse, dans [Cav23] au caractère bien posé de (1.4), où Z est un processus de Lévy général admettant un moment d'ordre $\beta \in [1,2]$ fini, et où les coefficients b et σ sont lipschitziens sur $\mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$. Lorsque $\beta \in (0,1)$, on ne peut pas travailler avec la distance de Wasserstein W_β (voir la Remarque 3.4). Enfin, dans le cas où Z n'admet pas de moment fini, l'existence faible est démontrée dans [JMW07] à travers l'étude du problème de martingale non-linéaire associé, et sous une hypothèse lipschitzienne, par rapport à l'argument mesure, relativement à une distance de Wasserstein modifiée. L'unicité n'est cependant pas établie.

- **Avec un bruit qui ne dégénère pas.** Mentionnons deux travaux récents qui tirent profit du phénomène de régularisation stochastique pour des EDS de McKean-Vlasov dirigées par un bruit de Lévy. Dans le cas höldérien, on renvoie à Frikha, Konakov et Menozzi [FKM21]. Plus précisément, les coefficients sont lipschitziens par rapport à une distance sur $\mathcal{P}(\mathbb{R}^d)$, définie par dualité avec une classe de fonctions tests bornées et höldériennes, qui métrise la convergence étroite. On renvoie également à Chaudru de Raynal, Jabir et Menozzi [CdRJM23] dans le cadre d'une dérive sous la forme convolution avec un noyau singulier appartenant à un espace de Besov. Dans ces deux travaux, les bruits sont des processus α-stables.

1.2.2 Systèmes de particules en interaction et propagation du chaos

Une des motivations pour étudier l'EDS de McKean-Vlasov (1.4) réside dans son lien étroit avec le système de particules en interaction défini, pour $N \geq 1$, par

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\, dt + \sigma(t, X_{t^-}^{i,N}, \overline{\mu}_t^N)\, dZ_t^i, \quad t \geq 0, \quad i \in \{1, \dots, N\}, \\ \overline{\mu}_t^N := \frac{1}{N} \sum_{k=1}^{N} \delta_{X_t^{k,N}}, \\ X_0^{i,N} = \xi^i \end{cases} \tag{1.6}$$

où $(\xi^i, Z^i)_{i \geq 1}$ est une suite i.i.d. de même loi que le couple de variables indépendantes (ξ, Z). Il s'agit d'un système de N particules soumises à des bruits i.i.d. $(Z^i)_{i \geq 1}$ et interagissant via la mesure empirique du système $\overline{\mu}^N$. Ce type d'interaction est appelé *champ moyen* puisque chaque particule interagit avec la distribution statistique de tout le système. Les particules $(X^{i,N})_{i \leq N}$ sont identiquement distribuées grâce à la symétrie de l'interaction et le caractère i.i.d. des bruits et des données initiales. En revanche, elles ne sont pas indépendantes puisque chacune interagit avec les autres. Le lien entre (1.4) et (1.6) est que la dynamique d'une particule quelconque de (1.6) est décrite, lorsque le nombre de particules N tend vers l'infini, par l'EDS de McKean-Vlasov (1.4). Une propriété plus forte liant une équation de McKean-Vlasov à son système de particules en interaction est la propriété de *propagation du chaos*. Elle stipule que la dynamique d'un nombre quelconque k de particules est décrit par k copies indépendantes de l'EDS de McKean-Vlasov (1.4), quand le nombre de particules N tend vers l'infini. L'origine de la notion de propagation du chaos remonte à Boltzmann et l'hypothèse de chaos moléculaire à l'origine de la théorie cinétique. La justification mathématique de ses travaux et de cette hypothèse est restée longtemps sans réponse. C'est Kac [Kac56] qui donne le premier une définition rigoureuse du chaos et l'illustre sur un système simplifié. Il a ensuite été étudié par McKean [McK67], puis par Sznitman [Szn91], lorsque les bruits $(Z^i)_{i \geq 1}$ sont des mouvements browniens. Mentionnons que, dans la définition de la propagation du chaos donnée dans [Szn91], les données initiales $(\xi^i)_{i \geq 1}$ sont supposées être seulement *chaotiques* au lieu d'indépendantes et identiquement distribuées.

Ces systèmes de particules en interaction de type champ moyen ont de nombreuses applications. En physique, ils ont un rôle crucial en théorie cinétique, comme mentionné précédemment. Ils ont également de nombreuses applications en biologie, notamment pour étudier la dynamique d'une population de cellules, en neurosciences pour modéliser les interactions entre neurones, en sciences sociales pour décrire des mouvements d'auto-organisation et aussi dans la théorie des jeux à champ moyen.

Donnons des définitions plus précises de la propagation du chaos au sens faible et au sens fort. Pour une présentation exhaustive, on renvoie à [CD22a] et [CD22b].

Définition 1.15 (Propagation du chaos faible)**.** On dit qu'il y a *propagation du chaos faible* pour le système de particules (1.6) si, pour tout $k \geq 1$, et pour tout $t \geq 0$, la loi de $(X_t^{1,N}, \ldots, X_t^{k,N})$ converge, quand N tend vers l'infini, vers la mesure produit $\mu_t^{\otimes k}$, où $(\mu_t)_{t\geq 0}$ est le flot de lois marginales de l'EDS de McKean-Vlasov (1.4).

D'après [Szn91, Proposition 2.2], cette définition est équivalente à la propriété suivante.

Proposition 1.16 (Approximation de champ moyen)**.** *Il y a propagation du chaos faible pour le système de particules* (1.6) *si et seulement si, pour tout $t \geq 0$, la suite de mesures aléatoires $(\overline{\mu}_t^N)_{N\geq 1}$ converge en loi vers la variable aléatoire constante égale à μ_t.*

Puisque ces deux notions sont équivalentes, on s'autorisera dans la suite à parler de *propagation du chaos faible* dès lors qu'on prouve une des deux propriétés équivalentes de la Proposition 1.16.

On peut également s'intéresser à la propagation du chaos au niveau des trajectoires. Introduisons des copies i.i.d. $(X^{i,\infty})_{i\geq 1}$ de l'EDS de McKean-Vlasov (1.4), où les données initiales et les bruits sont identiques à ceux du système de particules, c'est-à-dire $(\xi^i, Z^i)_{i\geq 1}$. Il s'agit d'un couplage entre (1.4) et (1.6).

Définition 1.17 (Propagation du chaos forte (ou trajectorielle))**.** On dit qu'il y a *propagation du chaos forte* (ou trajectorielle) pour le système de particules (1.6) si, pour tout $t \geq 0$,

$$\frac{1}{N}\sum_{k=1}^{N} \mathbb{E}|X_t^{k,N} - X_t^{k,\infty}| \underset{N\to+\infty}{\longrightarrow} 0.$$

Mentionnons qu'on peut utiliser d'autres couplages que les $(X^{i,\infty})_{i\geq 1}$ définis ici avec les mêmes bruits que les particules (voir par exemple [Ebe16] et [GBM22]). La terminologie de propagation du chaos faible ou forte utilisée ici provient des propriétés correspondantes pour les schémas numériques d'EDS. De plus, la propagation du chaos forte entraîne la propagation du chaos faible. En effet, s'il y a propagation du chaos au sens fort, alors, pour tout $k \geq 1$, la loi de $(X_t^{1,N}, \ldots, X_t^{k,N})$ converge, lorsque N tend vers l'infini, en distance de Wasserstein W_1 sur $\mathcal{P}_1((\mathbb{R}^d)^k)$ et donc en loi, vers $\mu_t^{\otimes k}$.

Dans les Définitions 1.15 et 1.17, la propagation du chaos est dite *qualitative* car il s'agit de résultats de convergence sans vitesse explicite par rapport à N. Quand des taux de convergence explicites par rapport à N sont prouvés, on dit que la propagation du chaos est *quantitative*. Précisons ce que cela signifie pour la propagation du chaos faible et forte. On se place désormais sur un intervalle de temps borné de la forme $[0, T]$, où $T > 0$ est fixé. Mentionnons qu'on peut s'intéresser à quantifier la propagation du chaos uniformément en temps, mais ces questions ne sont pas abordées dans la suite.

Concernant la propagation du chaos forte quantitative, compte tenu de la Définition 1.17, on peut sans ambiguïté dire qu'il s'agit de trouver une suite $(\varepsilon_N)_{N\geq 1}$ explicite par rapport à N, qui tend vers 0, et telle que pour tout $N \geq 1$ et $t \in [0, T]$

$$\frac{1}{N}\sum_{k=1}^{N} \mathbb{E}|X_t^{k,N} - X_t^{k,\infty}| \leq \varepsilon_N.$$

Bien sûr, on peut utiliser une distance L^p au lieu de la distance L^1 utilisée dans cette définition. Dans le travail [Cav23], on s'intéresse à la propagation du chaos forte quantitative pour (1.4), avec un bruit de Lévy admettant un moment d'ordre $\beta \in [1,2]$ fini, et où $\sigma = \mathrm{Id}$ et la dérive b est lipschitzienne sur $\mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$.

En ce qui concerne la propagation du chaos faible quantitative, on peut envisager de la définir de plusieurs manières. En considérant la Définition 1.15, on peut dire que la propagation du chaos faible quantitative consiste à trouver un contrôle explicite par rapport à N, et éventuellement k, sur la distance entre la loi jointe de k particules, notée $\mu_t^{1:k,N}$, et la loi produit $\mu_t^{\otimes k}$, pour un choix de distance sur $\mathcal{P}((\mathbb{R}^d)^k)$, et ce pour tout $k \geq 1$. Dans la suite, nous ne chercherons pas à quantifier la propagation du chaos faible de cette manière. En tenant compte de la Proposition 1.16, on peut dire que la propagation du chaos faible quantitative a pour objectif de trouver un contrôle explicite par rapport à N sur la distance entre la loi de la mesure empirique $[\overline{\mu}_t^N]$, et la loi limite δ_{μ_t}, pour une certaine distance sur $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ (par exemple, la distance notée $\mathcal{W}_{D_1}$ dans [CD22a, Définition 3.5]). Ce n'est pas non plus la définition que nous allons adopter dans la suite même si celle-ci entraînera un contrôle sur $\mathcal{W}_{D_1}([\overline{\mu}_t^N], \delta_{\mu_t})$. On définit dans ce qui suit la propagation du chaos faible quantitative en utilisant des fonctions tests définies sur $\mathcal{P}(\mathbb{R}^d)$.

Définition 1.18 (Propagation du chaos faible quantitative)**.** Soit $\mathscr{C}$ une classe de fonctions définies sur $\mathcal{P}(\mathbb{R}^d)$ à valeurs réelles, qui contient l'ensemble des fonctions ϕ telles qu'il existe $\varphi : \mathbb{R}^d \to \mathbb{R}$ bornée par 1 et lipschitzienne de constante de Lipschitz inférieure à 1, tel que pour tout $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\phi(\mu) = \int_{\mathbb{R}^d} \varphi \, d\mu.$$

On dit qu'il y a *propagation du chaos faible quantitative* (relativement à la classe de fonctions $\mathscr{C}$) s'il existe une suite explicite $(\varepsilon_N)_{N\geq 1}$, qui tend vers 0, et telle que pour tout $N \geq 1$, $t \in [0,T]$ et $\phi \in \mathscr{C}$

$$\mathbb{E}|\phi(\overline{\mu}_t^N) - \phi(\mu_t)| \leq \varepsilon_N. \tag{1.7}$$

Il est également naturel de vouloir trouver un contrôle explicite par rapport à N sur

$$|\mathbb{E}\phi(\overline{\mu}_t^N) - \phi(\mu_t)|. \tag{1.8}$$

Le contrôle (1.7) permet, grâce à la définition de $\mathcal{W}_{D_1}$ donnée dans [CD22a, Définition 3.5 et (49)], de déduire que pour tout $N \geq 1$ et $t \in [0,T]$

$$\mathcal{W}_{D_1}([\overline{\mu}_t^N], \delta_{\mu_t}) \leq \varepsilon_N. \tag{1.9}$$

Comme mentionné précédemment, cela permet bien de quantifier la propagation du chaos faible à travers la convergence de la mesure empirique du système grâce à la Proposition 1.16. Concernant le contrôle de (1.8), il permet par exemple de quantifier la vitesse d'approximation de la loi d'une particule par la loi du processus de McKean-Vlasov limite (voir la Remarque 3.25 pour plus de détails).

Lorsque Z est un mouvement brownien, les EDS de McKean-Vlasov et les systèmes de particules ont été beaucoup étudiés. En plus de McKean [McK67] et Sznitman [Szn91], on peut mentionner par exemple Gärtner [Gä88], Méléard [Mé96], Malrieu [Mal03], Mischler et al. [MMW15], Jabin et Wang [JW18], Lacker [Lac18, Lac21], Tomašević [Tom20] et Jabir [Jab19]. Le développement récent des jeux à champ moyen a donné une nouvelle impulsion pour étudier ces systèmes et le phénomène de propagation du chaos. En particulier, cela a mis à disposition un nouveau formalisme ainsi que de nouveaux outils, comme la notion *d'équations maîtresses*, qui sont des EDP sur l'espace des mesures de probabilités. Cela a permis de revisiter ou de généraliser quelques uns des travaux précédents. Dans cette direction, on peut renvoyer au livre de Carmona et Delarue [CD18a], ainsi qu'à Chaudru de Raynal et Frikha [CdRF22, CdRF21], Chassagneux et al. [CST22], Delarue and Tse [DT21] et Jourdain et Tsé [JT21]. Ces nouveaux outils sont aux coeur de cette thèse, en particulier dans les travaux [Cav22a] et [Cav22b].

On présente maintenant quelques travaux sur la propagation du chaos dans le cadre d'EDS de McKean-Vlasov dirigées par des processus de Lévy. Le cas où le bruit est un processus de Lévy général ayant un moment d'ordre 2 fini est traité par Jourdain, Méléard et Woyczysnki [JMW07]. Les auteurs établissent des estimations de propagation du chaos forte dans L^2 sous des hypothèses lipschitziennes sur les coefficients de l'EDS, par rapport à W_2 pour l'argument mesure. Toujours dans le même cadre lipschitzien, on fait référence à Neelima et al. [NBK+20], où les auteurs relaxent les hypothèses de [JMW07].

Intéressons-nous maintenant au cas où le bruit n'admet plus un moment d'ordre 2 fini. Dans [Gra92a], en suivant l'approche de Sznitman [Szn91] dans le cas brownien, Graham établit un résultat de propagation du chaos faible qualitatif sous des hypothèses lipschitziennes sur les coefficients de l'EDS, par rapport à W_1 pour l'argument mesure. L'EDS est dirigée par une mesure aléatoire de Poisson et une mesure aléatoire de Poisson compensée, et sa solution admet un moment d'ordre 1 fini. Enfin, en dimension 1, on peut mentionner Frikha et Li [FL21], où les auteurs étudient une EDS de McKean-Vlasov dirigée par une mesure de Poisson compensée avec des sauts positifs. Ils prouvent des inégalités de propagation du chaos forte dans L^1, sous des hypothèses lipschitziennes "d'un côté" en espace sur les coefficients (car les sauts sont positifs), par rapport à W_1 pour l'argument mesure.

1.2.3 Propagation du chaos faible quantitative : méthode basée sur le semi-groupe

On présente ici une méthode permettant de prouver la propagation du chaos faible quantitative, au sens donné dans la Définition 1.18. Cette stratégie a été originellement décrite dans [CD18a, Chapitre 5, p. 506 − 508], inspirée par [CDLL19] et [MMW15]. Elle a été utilisée notamment dans [CST22, DT21, CdRF21]. Cette méthode a été développée uniquement dans le cas d'un bruit brownien et l'objectif des travaux [Cav22a, Cav22b] est de la mettre en œuvre lorsque Z est un processus α-stable, avec $\alpha \in (1, 2)$. Mentionnons qu'une autre méthode permettant de prouver la propagation du chaos faible quantitative se trouve dans [MMW15]. En particulier, les auteurs étudient la propagation du chaos pour un modèle de collisions inélastiques pour l'équation de Bolzmann, qui présente des sauts.

Semi-groupe associé à l'EDS de McKean-Vlasov (1.4)

Fixons un horizon de temps fini T et $\beta \in [1, 2]$. On suppose que l'EDS de McKean-Vlasov (1.4) est bien posée au sens faible, et que si la donnée initiale ξ admet un moment d'ordre β fini, alors c'est le cas pour la solution en tout temps. On note alors $[X_t^{s,\mu}] \in \mathcal{P}_\beta(\mathbb{R}^d)$ la loi de la solution de (1.4) au temps $t \in [s, T]$, initialisée au temps s par une variable aléatoire de loi $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. Soit $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ une fonction. On définit alors, pour $0 \leq s \leq t \leq T$ et $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$,

$$\mathcal{T}_{s,t}\phi(\mu) := \phi([X_t^{s,\mu}]). \tag{1.10}$$

Cela définit le semi-groupe à 2 paramètres associé à l'EDS de McKean-Vlasov (1.4), qui vérifie la relation fondamentale

$$\forall\, 0 \leq s \leq \tau \leq t \leq T, \quad \mathcal{T}_{s,t} = \mathcal{T}_{s,\tau} \circ \mathcal{T}_{\tau,t}.$$

Il est alors naturel de vouloir étudier le problème de Cauchy associé afin de décrire la dynamique du semi-groupe à travers son générateur. On renvoie à la Sous-section 1.1.4 pour la définition du semi-groupe, agissant sur les fonctions définies sur $\mathbb{R}^d$, associé à un processus de Lévy. Puisque le semi-groupe agit ici sur les fonctions définies sur $\mathcal{P}_\beta(\mathbb{R}^d)$ à valeurs réelles, on est amené à se poser la question de la différentiabilité de telles fonctions. Deux notions de différentiation sont présentées dans le Chapitre 2 : la dérivée au sens de Lions ou L-dérivée (quand $\beta = 2$) et la dérivée plate, qui est la notion avec laquelle nous avons travaillé dans cette thèse. Dans le cadre des EDS linéaires, la dynamique du semi-groupe est décrite à l'aide de la formule d'Itô. On est ainsi amené à s'intéresser à l'équivalent de la formule d'Itô classique dans le cadre présent, à savoir une formule d'Itô le long d'un flot de mesures de probabilité. Il s'agit plus précisément de décrire la dynamique de l'application $t \in [0,T] \mapsto \phi([X_t^{s,\mu}])$ sous des hypothèses de régularité sur ϕ. Dans le cadre brownien et en prenant $\beta = 2$, la formule d'Itô le long d'un flot de mesures de probabilité est présentée dans le Théorème 2.28 du Chapitre 2. Dans le cas d'un bruit de Lévy plus général et avec $\beta \in (0,2]$, cela a été établi dans [Cav22a] (voir le Théorème 3.9). Grâce à cette formule d'Itô, on peut établir que, sous des hypothèses de régularité, $\mathcal{T}_{s,t}\phi$ est solution de l'EDP de Kolmogorov rétrograde suivante

$$\partial_s \mathcal{T}_{s,t}\phi(\mu) + \mathscr{L}_s \mathcal{T}_{s,t}\phi(\mu) = 0, \quad \forall (s,\mu) \in [0,t) \times \mathcal{P}_\beta(\mathbb{R}^d), \tag{1.11}$$

où $\mathscr{L}_s$ est un opérateur différentiel qui agit sur les fonctions définies sur $\mathcal{P}_\beta(\mathbb{R}^d)$ suffisamment régulières, et est le générateur du semi-groupe.

Méthode pour prouver la propagation du chaos faible quantitative

Soit $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ une fonction suffisamment régulière. On considère la fonction U définie, pour $(t,\mu) \in [0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$, par

$$U(t,\mu) := \mathcal{T}_{t,T}\phi(\mu).$$

L'objectif pour prouver la propagation du chaos faible quantitative est de contrôler

$$\phi(\overline{\mu}_T^N) - \phi(\mu_T) = U(T,\overline{\mu}_T^N) - U(T,\mu_T).$$

On commence par calculer la dérivée de l'application $t \in [0,T) \mapsto U(t,\overline{\mu}_t^N)$, à condition que U soit suffisamment régulière. Pour ce faire, on doit utiliser la dynamique du système de particules. Il s'agit donc d'appliquer la formule d'Itô classique à la projection empirique de U, définie par $(t,x_1,\dots,x_N) \in [0,T] \times (\mathbb{R}^d)^N \mapsto U\left(t, \frac{1}{N}\sum_{k=1}^N \delta_{x_k}\right)$, et au système de particules (1.6). La notion de projection empirique, qui est donc au cœur de la méthode, est présentée dans le Chapitre 2. En particulier, on fait le lien entre les dérivées partielles spatiales de la projection empirique et la dérivée, par rapport à la mesure, de la fonction U qu'on projette via la mesure empirique.

Le point crucial est de remarquer que l'application $t \in [0,T] \mapsto U(t,\mu_t)$ est constante puisque l'EDS de McKean-Vlasov est supposée bien posée au sens faible. Il est donc naturel de s'attendre à ce que la dérivée en temps $\frac{d}{dt}U(t, \overline{\mu}_t^N)$ tende vers 0, lorsque N tend vers l'infini. Le calcul de cette dérivée fait apparaître l'EDP (1.11) vérifiée par U, et un terme d'erreur puisque le système de particules n'est qu'une approximation de la dynamique de McKean-Vlasov. Puisqu'on veut un résultat quantitatif, on doit estimer précisément ce terme d'erreur pour obtenir un taux de convergence explicite par rapport à N. Cela nécessite des contrôles sur U et sa dérivée par rapport à la mesure.

Chapitre 2

Boîte à outils : calcul différentiel sur les espaces de mesures de probabilité

Le calcul différentiel pour des fonctions définies sur un espace de mesures de probabilité, qui est présenté dans ce chapitre, est l'un des outils essentiels de cette thèse comme expliqué dans le chapitre précédent. On commence par définir les espaces usuels de mesures de probabilité sur lesquels on va travailler et les distances dont ils sont munis. Ensuite, on décrit deux notions de dérivation : la dérivée au sens de Lions et la dérivée plate. On donne des exemples et on mentionne le lien entre ces dernières. Dans cette thèse, on travaillera seulement avec la dérivée plate. La notion de projection empirique, au cœur d'un des résultats de la thèse est également présentée. Enfin, on introduit la formule d'Itô le long d'un flot de mesures qui est l'un des sujets et outils principaux de cette thèse.

2.1 Espaces de mesures de probabilité

On commence par introduire les espaces de mesures de probabilité classiques utilisés dans ce manuscrit. On trouvera les preuves des résultats de cette section dans [Vil09].

2.1.1 Définitions

Définition 2.1. On note $\mathcal{P}(\mathbb{R}^d)$ l'ensemble des mesures de probabilité sur $\mathbb{R}^d$. Il est muni de la topologie de la convergence étroite, ou convergence en loi.

On introduit également la distance en variation totale. La convergence pour cette dernière implique la convergence étroite.

Définition 2.2 (Distance en variation totale)**.** La distance en variation totale sur $\mathcal{P}(\mathbb{R}^d)$ est définie, pour $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, par

$$d_{TV}(\mu,\nu) := \frac{1}{2} \sup_{\|f\|_\infty \le 1} \left| \int_{\mathbb{R}^d} f \, d(\mu - \nu) \right|.$$

Lorsque μ et ν ont des densités par rapport à la mesure de Lebesgue, on a

$$d_{TV}(\mu,\nu) = \frac{1}{2} \left\| \frac{d\mu}{dx} - \frac{d\nu}{dx} \right\|_{L^1}.$$

On fixe $\beta \in [0 + \infty)$.

Définition 2.3. On définit

$$\mathcal{P}_\beta(\mathbb{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} |x|^\beta \, d\mu(x) < +\infty \right\}.$$

Si $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, on dira que μ admet un moment d'ordre β fini. Lorsque $\beta = 0$, $\mathcal{P}_0(\mathbb{R}^d) = \mathcal{P}(\mathbb{R}^d)$.

Soient $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. On note $\Pi(\mu, \nu)$ l'ensemble des mesures de probabilité $\pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ ayant comme lois marginales μ et ν. De telles mesures de probabilité sont appelées plans de transport entre μ sur ν. On rappelle la définition de distance de Wasserstein.

Définition 2.4 (Distances de Wasserstein)**.** Soit $\beta \in (0, +\infty)$. L'application W_β définie par

$$W_\beta : \left\{ \begin{array}{ccc} \mathcal{P}_\beta(\mathbb{R}^d) \times \mathcal{P}_\beta(\mathbb{R}^d) & \to & \mathbb{R}^+ \\ (\mu, \nu) & \mapsto & \left(\inf\limits_{\pi \in \Pi(\mu,\nu)} \displaystyle\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^\beta \, d\pi(x, y) \right)^{\frac{1}{\beta} \wedge 1} \end{array} \right.$$

est une distance sur $\mathcal{P}_\beta(\mathbb{R}^d)$, appelée distance de Wasserstein.

Remarque 2.5. Une réécriture plus concise de la distance de Wasserstein est la suivante

$$\forall \mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d), \quad W_\beta(\mu, \nu) = \left(\inf_{[U]=\mu,\, [V]=\nu} \mathbb{E}\, |U - V|^\beta \right)^{\frac{1}{\beta} \wedge 1},$$

où on note $[U]$ la loi de U.

L'inégalité de Hölder entraîne directement la comparaison des distances de Wasserstein de la proposition qui suit.

Proposition 2.6. *Soient $1 \le p \le q < +\infty$, alors $W_p \le W_q$.*

2.1.2 Convergence en distance de Wasserstein

Soit $\beta \in (0, +\infty)$. On s'intéresse maintenant au lien entre la convergence d'une suite pour la distance de Wasserstein W_β et la convergence étroite.

Définition 2.7. On définit

$$\mathcal{C}_{b,\beta}(\mathbb{R}^d) := \left\{ f \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{R}), (1 + |\cdot|)^{-\beta} f(\cdot) \in \mathcal{C}_b^0(\mathbb{R}^d; \mathbb{R}) \right\} \supset \mathcal{C}_b^0(\mathbb{R}^d; \mathbb{R}).$$

Théorème 2.8. *Soient $(\mu_n)_n \in \mathcal{P}_\beta(\mathbb{R}^d)^{\mathbb{N}}$ et $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. On note $\rightharpoonup$ la convergence étroite des mesures de probabilité. Alors, il y a équivalence entre les assertions suivantes.*

(1) $\mu_n \overset{W_\beta}{\to} \mu$.

(2) $\forall f \in \mathcal{C}_{b,\beta}(\mathbb{R}^d), \quad \displaystyle\int_{\mathbb{R}^d} f \, d\mu_n \underset{n \to +\infty}{\to} \int_{\mathbb{R}^d} f \, d\mu$.

(3) $\mu_n \rightharpoonup \mu$ *et* $\displaystyle\int_{\mathbb{R}^d} |x|^\beta \, d\mu_n(x) \underset{n \to +\infty}{\to} \int_{\mathbb{R}^d} |x|^\beta \, d\mu(x)$.

(4) $\mu_n \rightharpoonup \mu$ *et* $\displaystyle\lim_{R \to +\infty} \sup_n \int_{|x| \ge R} |x|^\beta \, d\mu_n(x) = 0$.

2.1.3 Inégalité de convexité

Soit $\beta \in [1, +\infty)$. L'inégalité suivante, dite de convexité, s'avère utile pour obtenir certaines majorations.

Proposition 2.9. *Soient $\mu_1, \mu_2, \nu_1, \nu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$ et $\alpha \in [0,1]$. Alors on a l'inégalité dite de convexité suivante*

$$W_\beta^\beta(\alpha\mu_1 + (1-\alpha)\mu_2, \alpha\nu_1 + (1-\alpha)\nu_2) \leq \alpha W_\beta^\beta(\mu_1, \nu_1) + (1-\alpha) W_\beta^\beta(\mu_2, \nu_2)$$

Preuve. On prend un plan de transport optimal π_1 envoyant μ_1 sur ν_1 et un plan optimal π_2 envoyant μ_2 sur ν_2. On considère la probabilité

$$\pi = \alpha\pi_1 + (1-\alpha)\pi_2 \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d).$$

On remarque que les lois marginales de π sont $\alpha\mu_1 + (1-\alpha)\mu_2$ et $\alpha\nu_1 + (1-\alpha)\nu_2$. Ainsi on a

$$\begin{aligned}
&W_\beta^\beta(\alpha\mu_1 + (1-\alpha)\mu_2, \alpha\nu_1 + (1-\alpha)\nu_2) \\
&\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\beta \, d\pi(x,y) \\
&= \alpha \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\beta \, d\pi_1(x,y) + (1-\alpha) \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\beta \, d\pi_2(x,y) \\
&= \alpha W_\beta^\beta(\mu_1, \nu_1) + (1-\alpha) W_\beta^\beta(\mu_2, \nu_2).
\end{aligned}$$

□

On a également l'inégalité suivante.

Proposition 2.10. *Soient $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$. Pour $t \in [0,1]$, on définit $m_t := \nu + t(\mu - \nu)$. Alors si $t \in [0,1]$ et h tel que $t + h \in [0,1]$, on a*

$$W_\beta^\beta(m_{t+h}, m_t) \leq |h| W_\beta^\beta(\mu, \nu).$$

Preuve. On suppose $h > 0$, l'autre cas étant symétrique. Pour cela on remarque qu'on peut écrire

$$m_{t+h} = h\mu + (1-h)\frac{(1-(t+h))\nu + t\mu}{1-h} \quad \text{et} \quad m_t = h\nu + (1-h)\frac{(1-(t+h))\nu + t\mu}{1-h},$$

et on applique l'inégalité précédente. □

Les résultats de cette sous-section se généralisent immédiatement au cas $\beta \in (0,1)$ en retirant les puissances β des distances W_β qui apparaissent.

2.2 Dérivée de Lions

On fixe $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$. Il existe plusieurs manières de dériver la fonction u. On commence par présenter la dérivée de Lions ou L-dérivée introduite par Lions dans son cours au Collège de France [Lio] sur les jeux à champ moyen. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé complet et sans atomes. Pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, il existe $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$ de loi μ. On définit alors

$$\tilde{u} \left\{ \begin{array}{rl} L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) & \to \mathbb{R} \\ X & \mapsto u([X]), \end{array} \right.$$

où $[X]$ désigne la loi de X. La fonction $\tilde{u}$ est appelée lifting de u.

Définition 2.11. On dit que u est L-différentiable si $\tilde{u}$ est Fréchet-différentiable et on note $D\tilde{u}$ sa différentielle, identifiée à son gradient. De même, on dit que u est continûment L-différentiable si $\tilde{u}$ est $\mathcal{C}^1$ au sens usuel.

Remarque 2.12. Si $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$, la différentielle de $\tilde{u}$ en X est notée $D\tilde{u}(X)$. Cette notation porte légèrement confusion car on peut avoir l'impression que $D\tilde{u}(X)$ est l'image de la variable X par une certaine application déterministe, ce qui n'est a priori pas le cas. La proposition suivante, tirée de [CD18a, Proposition 5.24 et 5.25], apporte une précision à cette remarque.

Proposition 2.13. *Si u est L-différentiable, la loi du couple $(X, D\tilde{u}(X))$ ne dépend que de la loi de X et pas de la variable X de loi donnée. De plus, la variable aléatoire $D\tilde{u}(X) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ est $\sigma(X)$-mesurable et pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, il existe une fonction mesurable $\partial_\mu u(\mu) : \mathbb{R}^d \to \mathbb{R}^d$ telle que pour toute variable aléatoire X de loi μ, on a*

$$D\tilde{u}(X) \overset{p.s.}{=} \partial_\mu u(\mu)(X).$$

Remarque 2.14. L'application $\partial_\mu u(\mu)$ appartient à $L^2(\mu)$, donc est définie seulement μ-presque-partout.

Mentionnons la propriété fondamentale de représentation de la L-dérivée.

Proposition 2.15. *Supposons que u est continûment L-différentiable. Alors pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, il existe une version de $\partial_\mu u(\mu) \in L^2(\mathbb{R}^d, \mu)$, telle que l'application*

$$(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_\mu u(\mu)(v),$$

soit mesurable, où on munit $\mathcal{P}_2(\mathbb{R}^d)$ de la tribu borélienne associée à la topologie induite par W_2. De plus, si $\partial_\mu u(\mu)$ possède une version continue, alors les deux versions coïncident sur le support de μ.

Pour la preuve, on pourra consulter [CD18a, Proposition 5.33].

2.3 Dérivée plate

Présentons maintenant une autre notion de dérivation pour les fonctions définies sur un espace de mesures de probabilité. Il s'agit de la notion de dérivée plate, qui est celle utilisée dans cette thèse. On travaille sur $\mathcal{P}_\beta(\mathbb{R}^d)$ avec $\beta \in [0, +\infty)$. Bien sûr il y a un lien fort entre les deux notions de dérivation lorsque $\beta = 2$ comme nous le verrons dans la suite.

Définition 2.16. Soit $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$. On dit que u admet une *dérivée plate* s'il existe une application

$$\frac{\delta}{\delta m} u : (\mu, v) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \frac{\delta}{\delta m} u(\mu)(v) \in \mathbb{R}$$

continue qui satisfait les propriétés suivantes.

(1) Pour tout $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$ compact, il existe $C_\mathcal{K} > 0$ telle que

$$\forall v \in \mathbb{R}^d, \sup_{\mu \in \mathcal{K}} \left| \frac{\delta}{\delta m} u(\mu)(v) \right| \leq C_\mathcal{K}(1 + |v|^\beta).$$

(2) Pour tout $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$

$$u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(t\mu + (1-t)\nu)(v)\, d(\mu - \nu)(v)\, dt.$$

Cette notion utilise la convexité de $\mathcal{P}_\beta(\mathbb{R}^d)$. En fait, il s'agit d'une dérivée au sens de Gateaux sur l'espace de mesures signées finies.

Remarque 2.17. Si u admet une dérivée plate, alors u est continue par théorème de convergence dominée.

Donnons une caractérisation de la dérivée plate qui peut être utile quand on veut la calculer en pratique.

Proposition 2.18. *Soit $\frac{\delta}{\delta m} u : \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ une application continue vérifiant le premier point de la Définition 2.16. Alors u admet $\frac{\delta}{\delta m} u$ comme dérivée plate si et seulement si pour tout $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, l'application $t \in [0,1] \mapsto u(t\mu + (1-t)\nu)$ est dérivable avec*

$$\frac{d}{dt} u(t\mu + (1-t)\nu) = \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(t\mu + (1-t)\nu)(v)\, d(\mu - \nu)(v).$$

Preuve. On fait la preuve lorsque $\beta \in [1, +\infty)$, le cas $\beta \in [0,1)$ se traitant de façon complètement analogue. Le sens indirect est immédiat puisque la dérivée est continue par rapport à t (par convergence dominée), on peut donc intégrer entre 0 et 1. Pour le sens direct, on fixe $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$. On note pour $t \in [0,1]$, $m_t := t\mu + (1-t)\nu$. On fixe $t \in [0,1]$ et $h > 0$ tel que $t + h \in [0,1]$, le cas $h < 0$ étant analogue. Par définition de la dérivée plate on a

$$\begin{aligned}
u(m_{t+h}) - u(m_t) &= \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(v m_{t+h} + (1-v)m_t)(v)\, h\, d(\mu - \nu)(v)\, dv \\
&= h \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(m_t)(v)\, d(\mu - \nu)(v) \\
&+ h \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} u(\lambda m_{t+h} + (1-\lambda)m_t)(v) - \frac{\delta}{\delta m} u(m_t)(v) \right) d(\mu - \nu)(v)\, d\lambda.
\end{aligned}$$

Il s'agit donc de montrer que le dernier terme est un $o(h)$. Soit $(h_n)_n$ une suite strictement positive qui tend vers 0. Alors, pour $(v, \lambda) \in \mathbb{R}^d \times [0,1]$ fixé, on a

$$\frac{\delta}{\delta m} u(\lambda m_{t+h_n} + (1-\lambda)m_t)(v) \underset{n \to +\infty}{\to} \frac{\delta}{\delta m} u(m_t)(v).$$

En effet, on a d'après les Propositions 2.9 et 2.10

$$\begin{aligned}
W_\beta^\beta(\lambda m_{t+h_n} + (1-\lambda)m_t, m_t) &\leq \lambda W_\beta^\beta(m_{t+h_n}, m_t) \\
&\leq W_\beta^\beta(m_{t+h_n}, m_t) \\
&\leq h_n W_\beta^\beta(\mu,\nu) \\
&\underset{n\to+\infty}{\to} 0.
\end{aligned}$$

De plus, puisque $m_{t+h_n} \overset{W_\beta}{\to} m_t$, on déduit que l'ensemble $\{\lambda m_{t+h_n} + (1-\lambda)m_t,\ \lambda \in [0,1],\ n \geq 1\}$ est relativement compact. Il existe donc $C > 0$ telle que

$$\forall \lambda \in [0,1],\ \forall n \geq 1,\ \forall v \in \mathbb{R}^d,\ \left|\frac{\delta}{\delta m}u(\lambda m_{t+h_n} + (1-\lambda)m_t)(v) - \frac{\delta}{\delta m}u(m_t)(v)\right| \leq 2C(1+|v|^\beta).$$

Le théorème de convergence dominée assure donc que

$$\int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}u(\lambda m_{t+h_n} + (1-\lambda)m_t)(v) - \frac{\delta}{\delta m}u(m_t)(v)\right) d(\mu-\nu)(v)\, d\lambda \to 0.$$

D'où le résultat. □

Pour des fonctions dépendant en plus du temps et de l'espace, on introduit l'espace suivant de fonctions $\mathcal{C}^1$ qui sera utile dans la suite.

Définition 2.19. On définit $\mathcal{C}^1([0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d))$ comme l'espace des fonctions continues $u : [0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ vérifiant les propriétés suivantes.

(1) Pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, l'application $u(\cdot,\cdot,\mu)$ appartient à $\mathcal{C}^1([0,T]\times\mathbb{R}^d)$ avec $\partial_t u$ et $\partial_x u$ continues sur $[0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)$.

(2) Pour tout $(t,x) \in [0,T]\times\mathbb{R}^d$, l'application $u(t,x,\cdot)$ admet une dérivée plate $(\mu,v) \in \mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d \mapsto \frac{\delta}{\delta m}u(t,x,\mu)(v)$ tel que $\frac{\delta}{\delta m}u$ est continue sur $[0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)$.

(3) Pour tout $(t,x,\mu) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)$, l'application $\frac{\delta}{\delta m}u(t,x,\mu)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^d$ et $\partial_v\frac{\delta}{\delta m}u$ est continue sur $[0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$.

2.4 Exemples

Donnons deux exemples de calculs de dérivées plates sur $\mathcal{P}_\beta(\mathbb{R}^d)$ pour $\beta \in [0,+\infty)$.

Exemple 2.20. On se donne $\phi \in \mathcal{C}^0(\mathbb{R}^d,\mathbb{R})$ telle qu'il existe $C > 0$ vérifiant

$$\forall v \in \mathbb{R}^d,\ |\phi(v)| \leq C(1+|v|^\beta).$$

On considère

$$u : \mu \in \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d}\phi\, d\mu.$$

Alors pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$

$$\frac{\delta}{\delta m}u(\mu) = \phi.$$

En effet, pour $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, on a

$$u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \phi(v)\, d(\mu - \nu)(v)\, dt.$$

Lemme 2.21. *Soit $f \in \mathcal{C}^0(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$ telle qu'il existe $C > 0$ vérifiant*

$$\forall x \in \mathbb{R}^d,\ |f(x,y)| \leq C(1 + |x|^\beta + |y|^\beta).$$

Alors l'application $(\mu, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} f(x,y) d\mu(y)$ est continue.

Preuve. Soit $(\mu_n)_n$ et $(x_n)_n$ telles que

$$\mu_n \overset{W_\beta}{\to} \mu \quad \text{et} \quad x_n \to x.$$

Soit $\epsilon > 0$, comme $\mu_n \overset{W_p}{\to} \mu$, le Théorème 2.8 assure que $\mu_n \rightharpoonup \mu$, donc la suite est tendue et

$$\lim_{R \to +\infty} \sup_{n \geq 1} \int_{|x| \geq R} |x|^\beta\, d\mu_n(x) = 0.$$

Il existe donc $R > 0$ tel que

$$\sup_{n \geq 1} \int_{|x| \geq R} d\mu_n(x) \leq \epsilon \quad \text{et} \quad \sup_{n \geq 1} \int_{|x| \geq R} |x|^\beta\, d\mu_n(x) \leq \epsilon.$$

On a donc

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} f(x_n, y) d\mu_n(y) - \int_{\mathbb{R}^d} f(x,y) d\mu(y) \right| \\
&\leq \left| \int_{\mathbb{R}^d} f(x_n, y) d\mu_n(y) - \int_{\mathbb{R}^d} f(x,y) d\mu_n(y) \right| + \left| \int_{\mathbb{R}^d} f(x,y) d\mu_n(y) - \int_{\mathbb{R}^d} f(x,y) d\mu(y) \right| \\
&\leq \int_{|y| \leq R} |f(x_n, y) - f(x,y)| d\mu_n(y) + \int_{|y| > R} |f(x_n, y) - f(x,y)| d\mu_n(y) \\
&\quad + \left| \int_{\mathbb{R}^d} f(x,y) d\mu_n(y) - \int_{\mathbb{R}^d} f(x,y) d\mu(y) \right|
\end{aligned} \tag{2.1}$$

Le théorème de Heine assure que pour n assez grand et pour tout $y \in B_R$, la boule centrée en 0 et de rayon R

$$|f(x_n, y) - f(x,y)| \leq \epsilon.$$

Donc pour n assez grand

$$\int_{|y| \leq R} |f(x_n, y) - f(x,y)| d\mu_n(y) \leq \epsilon.$$

Le second terme de du membre de droite de (2.1) est majoré par

$$\int_{|y| > R} C(2 + |x_n|^\beta + |x|^\beta + 2|y|^\beta)\, d\mu_n(y).$$

Puisque pour n assez grand, $|x_n| \leq |x| + 1$, on déduit que pour n assez grand

$$\int_{|y|>R} |f(x_n, y) - f(x, y)| d\mu_n(y) \leq C(2 + |x|^\beta + (|x| + 1)^\beta + 2)\epsilon.$$

Le dernier terme du membre de droite de (2.1) tend vers 0 d'après le Théorème 2.8 puisque $f(x, \cdot) \in \mathcal{C}_{b,\beta}(\mathbb{R}^d)$, d'où le résultat. □

Exemple 2.22. On se donne $h \in \mathcal{C}^0(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R})$ telle qu'il existe $C > 0$ vérifiant

$$\forall x, y \in \mathbb{R}^d,\ |h(x, y)| \leq C(1 + |x|^\beta + |y|^\beta),$$

et on considère

$$u : \mu \in \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^d} h(x, y)\, d\mu(x)\, d\mu(y).$$

Alors, on a pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ et $v \in \mathbb{R}^d$

$$\frac{\delta}{\delta m} u(\mu)(v) = \int_{\mathbb{R}^d} h(y, v) d\mu(y) + \int_{\mathbb{R}^d} h(v, y)\, d\mu(y).$$

Le candidat pour la dérivée plate est bien continue sur $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ par le lemme précédent. De plus, on a

$$\left| \int_{\mathbb{R}^d} h(y, v) d\mu(y) + \int_{\mathbb{R}^d} h(v, y)\, d\mu(y) \right| \leq 2C \left(1 + |v|^\beta + \int_{\mathbb{R}^d} |y|^\beta\, d\mu(y) \right),$$

ce qui montre la condition de croissance que doit vérifier la dérivée plate. Enfin, fixons $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$ et $t \in [0, 1]$. On a

$$u(t\mu + (1 - t)\nu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d(\nu + t(\mu - \nu))(x)\, d(\nu + t(\mu - \nu))(y).$$

En développant et en dérivant, on obtient

$$\begin{aligned}
&\frac{d}{dt} u(t\mu + (1 - t)\nu) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d\nu(x)\, d(\mu - \nu)(y) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d\nu(y)\, d(\mu - \nu)(x) \\
&+ 2t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d(\mu - \nu)(x)\, d(\mu - \nu)(y) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d(\nu + t(\mu - \nu))(x)\, d(\mu - \nu)(y) + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x, y)\, d(\nu + t(\mu - \nu))(y)\, d(\mu - \nu)(x) \\
&= \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(t\mu + (1 - t)\nu)(v)\, d(\mu - \nu)(v).
\end{aligned}$$

D'où le résultat.

2.5 Lien entre dérivée de Lions et dérivée plate

Les deux notions de dérivation pour une fonction u définie sur $\mathcal{P}_2(\mathbb{R}^d)$ introduites précédemment sont distinctes. On a vu que la dérivée plate consiste à dériver $t \in [0, 1] \mapsto u(t\mu + (1 - t)\nu)$. En revanche, la L-dérivée permet en particulier de dériver les applications de la forme $t \in [0, 1] \mapsto u([tX + (1 - t)Y])$, où $[X] = \mu$ et $[Y] = \nu$. On a évidemment pas de manière générale $[tX+(1-t)Y] = t[X]+(1-t)[Y]$. Une bonne manière de comprendre $[tX + (1-t)Y]$ est de voir cette mesure de probabilité comme la loi de la variable $\mathbf{1}_{U\leq t}X + \mathbf{1}_{U>t}Y$, où U est une variable indépendante de X et Y et distribuée uniformément sur $[0,1]$.

Commençons par donner un contre-exemple d'une fonction qui admet une dérivée plate mais qui n'est pas L-dérivable.

Contre-exemple 2.23. On considère l'application moment d'ordre 1 définie par

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \qquad \mu & \mapsto \displaystyle\int_{\mathbb{R}^d} |x| \, d\mu(x). \end{cases}$$

Cette application admet comme dérivée plate l'application constante égale à $|\cdot|$ d'après l'Exemple 2.20. Montrons que u n'est pas L-dérivable. Le lifting de u est l'application

$$\tilde{u} : X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \mapsto \mathbb{E}|X|.$$

Si u était L-dérivable alors pour tout $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$, l'application $t \in \mathbb{R} \mapsto \mathbb{E}|tX|$ serait dérivable par composition. Ce n'est pas le cas comme on le voit en prenant $X = 1$.

On fixe toujours $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ et on cherche maintenant comment passer d'une dérivée plate à une L-dérivée.

Proposition 2.24. *Supposons que u admet une dérivée plate $\frac{\delta}{\delta m}u$. On suppose de plus que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\frac{\delta}{\delta m}u(\mu)$ est différentiable et vérifie les propriétés suivantes.*

(1) L'application $(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_v \frac{\delta}{\delta m}u(\mu)(v) \in \mathbb{R}^d$ est continue.

(2) Pour tout compact $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$, il existe $C_{\mathcal{K}} > 0$ telle que pour tout $v \in \mathbb{R}^d$

$$\sup_{\mu\in\mathcal{K}} \left| \partial_v \frac{\delta}{\delta m} u(\mu)(v) \right| \leq C_{\mathcal{K}}(1 + |v|).$$

Alors, la fonction u est L-dérivable. De plus on a

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\ \partial_\mu u(\mu) = \partial_v \frac{\delta}{\delta m} u(\mu).$$

La preuve de ce résultat se trouve dans [CD18a, Proposition 5.48].

Donnons maintenant un résultat assurant qu'une fonction L-dérivable admette également une dérivée plate. Il provient de [CD18a, Propositon 5.51].

Proposition 2.25. *Supposons que u est L-dérivable et que sa L-dérivée est globalement lipschitzienne. On suppose également que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, on peut trouver une version de $\partial_\mu u(\mu)$ telle que*

$$(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_\mu u(\mu)(v),$$

est continue. Alors u admet une dérivée plate qui satisfait l'hypothèse (2) *de croissance sous-linéaire uniformément sur les compacts de $\mathcal{P}_2(\mathbb{R}^d)$ de la Proposition 2.24.*

2.6 Projection empirique

Comme on l'a vu dans la Sous-section 1.2.3, la notion de projection empirique, définie dans ce qui suit, est un objet crucial pour prouver la propagation du chaos faible quantitative. Fixons $\beta \in [0, +\infty)$.

Définition 2.26 (Projection empirique)**.** Soit $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$. Pour $N \geq 1$, la *projection empirique* d'ordre N de u, qu'on note u^N, est définie par

$$u^N : \boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N \mapsto u(\overline{\mu}^N_{\boldsymbol{x}}),$$

où $\overline{\mu}^N_{\boldsymbol{x}} := \frac{1}{N} \sum_{k=1}^N \delta_{x_k}$ est la mesure empirique du vecteur $\boldsymbol{x}$.

Le résultat suivant permet d'assurer que la projection empirique est de classe $\mathcal{C}^1$.

Lemme 2.27. *Soit $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ une fonction admettant une dérivée plate $\frac{\delta}{\delta m} u$ qui vérifie les propriétés suivantes.*

(1) Pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(\mu) \in \mathcal{C}^1(\mathbb{R}^d; \mathbb{R})$.

(2) L'application $(\mu, v) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \partial_v \frac{\delta}{\delta m} u(\mu)(v) \in \mathbb{R}^d$ est continue.

Alors, pour tout $N \geq 1$, la projection empirique u^N de u est de classe $\mathcal{C}^1$ sur $(\mathbb{R}^d)^N$. De plus, pour tout $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ et $i \in \{1, \dots, N\}$,

$$\partial_{x_i} u^N(x_1, \dots, x_N) = \frac{1}{N} \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i).$$

Preuve. On fait la preuve dans le cas $\beta \geq 1$, le cas $\beta \in [0, 1)$ étant analogue. Par définition de la dérivée plate, on a pour $\mathbf{x} = (x_1, \dots, x_N), \mathbf{h} = (h_1, \dots, h_N) \in (\mathbb{R}^d)^N$

$$\begin{aligned}
&u^N(\boldsymbol{x} + \boldsymbol{h}) - u^N(\boldsymbol{x}) \\
&= \frac{1}{N} \sum_{i=1}^N \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i) \cdot h_i \\
&\quad + \frac{1}{N} \sum_{i=1}^N \int_0^1 \int_0^1 \left[\partial_v \frac{\delta}{\delta m} u(w \overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{h}} + (1-w) \overline{\mu}^N_{\boldsymbol{x}})(x_i + t h_i) - \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i) \right] \cdot h_i \, dt \, dw.
\end{aligned}$$

On majore le dernier terme du membre de droite de l'égalité précédente par

$$|\boldsymbol{h}| \frac{1}{N} \sum_{i=1}^N \int_0^1 \int_0^1 \left| \partial_v \frac{\delta}{\delta m} u(w \overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{h}} + (1-w) \overline{\mu}^N_{\boldsymbol{x}})(x_i + t h_i) - \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i) \right| dt \, dw.$$

Pour conclure à la différentiabilité de u^N, il suffit de voir que pour $i \in \{1, \dots, N\}$

$$\int_0^1 \int_0^1 \left| \partial_v \frac{\delta}{\delta m} u(w \overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{h}} + (1-w) \overline{\mu}^N_{\boldsymbol{x}})(x_i + t h_i) - \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i) \right| dt \, dw \underset{h \to 0}{\to} 0 \, .$$

Or on a

$$\sup_{w \in [0,1]} W_\beta^\beta(w \overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{h}} + (1-w) \overline{\mu}^N_{\boldsymbol{x}}, \overline{\mu}^N_{\boldsymbol{x}}) \leq \frac{1}{N} \sum_{k=1}^N |h_k|^\beta \underset{h \to 0}{\to} 0 \, .$$

La continuité de l'intégrande par rapport à $(\mu, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ permet de conclure que u^N est $\mathcal{C}^1$ et que pour tout $i \in \{1, \dots, N\}$

$$\partial_{x_i} u^N(x_1, \dots, x_N) = \frac{1}{N} \partial_x \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i).$$

□

2.7 Formule d'Itô le long d'un flot de mesures de probabilité

On introduit enfin le dernier outil qui est nécessaire à la méthode décrite dans la Sous-section 1.2.3 : la *formule d'Itô le long d'un flot de mesures de probabilité.*

2.7.1 Le résultat

On considère un processus d'Itô de la forme

$$X_t = X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dB_s, \quad t \geq 0,$$

où $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$, b et σ sont des processus progressivement mesurables à valeurs dans $\mathbb{R}^d$ et $\mathbb{R}^{d\times d}$ respectivement, et où B est un mouvement brownien sur $\mathbb{R}^d$. On note également $\mu_t = [X_t]$ pour tout $t \in \mathbb{R}^+$ et on suppose dans dans un premier temps que les hypothèses sur b et σ assurent que $t \in \mathbb{R}^+ \mapsto \mu_t \in \mathcal{P}_2(\mathbb{R}^d)$ est continue. Fixons une fonction $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$. Le but de la formule d'Itô le long d'un flot de mesures de probabilité est de décrire la dynamique de $t \in \mathbb{R}^+ \mapsto u(\mu_t)$, comme le fait la formule d'Itô classique pour $t \in \mathbb{R}^+ \mapsto f(X_t)$, où $f : \mathbb{R}^d \to \mathbb{R}$. Dans le cadre classique, la formule d'Itô requiert de la régularité sur la fonction f. Il en est de même pour la formule d'Itô pour un flot de mesures qui est énoncée dans le résultat suivant. Celui-ci se déduit de [CD18a, Théorème 5.99], où il est énoncé avec la L-dérivée, et de la Proposition 2.24.

Théorème 2.28. *Soit $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ une fonction $\mathcal{C}^1$ au sens de la Définition 2.19 et telle que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\partial_v \frac{\delta}{\delta m} u(\mu)$ est $\mathcal{C}^1$ et $\partial_v^2 \frac{\delta}{\delta m} u$ est continue sur $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$. On suppose que u vérifie les hypothèses de la Proposition 2.24 et que pour tout compact $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$*

$$\sup_{\mu \in \mathcal{K}} \int_{\mathbb{R}^d} \left| \partial_v^2 \frac{\delta}{\delta m} u(\mu)(v) \right|^2 d\mu(v) < +\infty.$$

On note toujours $\mu_t = [X_t]$ et on suppose que

$$\forall T > 0, \, \mathbb{E} \int_0^T |b_s|^2 + |\sigma_s|^4 \, ds < +\infty.$$

Alors, pour tout $t \in \mathbb{R}^+$, on a

$$u(\mu_t) = u(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s \right) ds + \frac{1}{2} \int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot a_s \right) ds,$$

où $a_s = \sigma_s \sigma_s^*$ et où pour toutes matrices $A, B \in \mathbb{R}^{d\times d}$, $A.B := Tr(AB^*)$ désigne le produit scalaire canonique.

2.7.2 Motivations et applications

La formule d'Itô de long d'un flot de mesures de probabilité a été développée sous l'impulsion de la théorie des jeux à champ moyen et des EDS de McKean-Vlasov. Les jeux à champ moyen ont été introduits indépendamment par Caines, Huang et Malhame [CHM06] et par Lasry et Lions [LL07]. La notion de *Master equation* ou *équation maîtresse* a été introduite par Lions dans son cours au Collège de France [Lio] dans le but de décrire les jeux à champ moyen. Les équations maîtresses sont des EDP définies sur un espace de mesures de probabilité. La formule d'Itô le long d'un flot de mesures est l'un des outils qui permet d'établir ce genre d'équations. On renvoie au cours de Lions [Lio], aux notes écrites par Cardaliaguet [Car10], et les deux livres écrits par Carmona et Delarue [CD18a, CD18b] pour plus de détails sur les jeux à champ moyen et les équations maîtresses. Mentionnons également Bensoussan, Frehse et Yam [BFY15] et Carmona et Delarue [CD14], où des équations maîtresses sont établies, grâce à la formule d'Itô le long d'un flot de mesures dans [CD14]. La question de l'existence et l'unicité de solutions classiques à ce type d'équations a été étudiée notamment par Cardaliaguet, Delarue, Lasry et Lions [CDLL19] et par Chassagneux, Crisan et Delarue [CCD15].

Comme expliqué dans la Sous-section 1.2.3 (voir (1.11)), la formule d'Itô le long d'un flot de mesures permet d'associer naturellement une EDP sur l'espace des mesures de probabilité à une EDS de McKean-Vlasov, plus précisément à son semi-groupe qui agit sur les fonctions définies sur un espace de mesures. On peut l'appeler également équation maîtresse. Cela s'avère être un outil important pour étudier le flot de mesures associé à une EDS de McKean-Vlasov, comme expliqué dans [CD18a, Chapitre 5]. Le lien entre une EDS de McKean-Vlasov et son EDP sur l'espace des mesures est au cœur du travail de Buckdahn, Li, Peng et Rainer [BLPR17], où les auteurs montrent que l'EDP est bien posée au sens classique et que sa solution s'exprime grâce au flot de mesures associé à la solution de l'EDS de McKean-Vlasov. Parallèlement, Chassagneux, Crisan et Delarue adoptent une approche similaire dans [CCD15]. Ils étudient le flot généré par un système progressif-rétrograde d'EDS sous des hypothèses plus faibles sur les coefficients de l'équation. Ces deux travaux ont été motivés par les jeux à champ moyen et la formule d'Itô pour un flot de mesures y joue un rôle clé. De plus, dans [CM17], Crisan et McMurray s'intéressent à ces EDP sur l'espace des mesures avec une condition terminale irrégulière en utilisant le calcul de Malliavin. Ils mettent en lumière un effet régularisant en ce qui concerne la différentiabilité de la solution de l'EDS par rapport à la loi initiale, et ce malgré le fait que le bruit brownien dans l'équation est seulement de dimension finie. Ce type d'effet régularisant par rapport à la mesure est également au cœur du travail [Cav22b] et permet de prouver la propagation du chaos faible quantitative comme expliqué dans la Sous-section 1.2.3. Mentionnons également que l'équation maîtresse associée au semi-groupe a été utilisée récemment par Jourdain et Tse [JT21] où les auteurs prouvent un théorème central limite pour le système de particules en interaction. Enfin, la formule d'Itô le long d'un flot de mesures est un outil important quand on étudie des problèmes de contrôles d'EDS de McKean-Vlasov. Elle permet d'obtenir un principe de programmation dynamique qui décrit la fonction valeur du problème, comme présenté dans [CD18a, Chapitre 6].

Comme dit précédemment, la formule d'Itô le long d'un flot de mesures de probabilité est au cœur de cette thèse. Dans le travail [Cav21], on affaiblit les hypothèses de régularité du Théorème 2.28 sur la fonction u à laquelle on peut appliquer la formule d'Itô. En particulier, on ne demande plus que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, la fonction $\frac{\delta}{\delta m}u(\mu)$ soit $\mathcal{C}^2$ au sens classique mais seulement qu'elle soit $\mathcal{C}^1$ et que $\partial_v \frac{\delta}{\delta m}u(\mu)$ appartienne à un certain espace de Sobolev. Il s'agit là encore d'un effet régularisant du bruit. Dans le travail [Cav22a], la formule d'Itô est établie pour un flot de lois marginales d'un processus à sauts défini par des intégrales stochastiques contre une mesure de Poisson et une mesure de Poisson compensée.

Deuxième partie

Résumé des travaux de la thèse et heuristiques

Chapitre 3

Propagation du chaos quantitative pour des EDS de McKean-Vlasov dirigées par un processus α-stable

Dans ce chapitre, on présente les résultats issus des travaux [Cav23], [Cav22a] et [Cav22b].

3.1 Caractère bien posé et propagation du chaos forte pour une EDS de McKean-Vlasov dirigée par un processus de Lévy général : cadre lipschitzien

Dans cette section, on présente les résultats de [Cav23]. L'objectif premier de ce travail est d'établir le caractère bien posé d'une EDS de McKean-Vlasov dirigée par un bruit de Lévy général admettant un moment d'ordre $\beta \in [1, 2]$ fini et sous des hypothèses lipschitziennes classiques, par rapport à la distance de Wasserstein W_β pour la variable mesure. Ce résultat, bien que très naturel, ne semblait pas avoir déjà été prouvé dans un cadre aussi général (voir la Remarque 3.3). Le second objectif de ce travail est de prouver la propagation du chaos forte quantitative sous des hypothèses plus restrictives, typiquement avec un coefficient de diffusion devant le bruit constant égal à l'identité. On n'a pas besoin ici des outils introduits dans le Chapitre 2 puisqu'on s'intéresse à la propagation du chaos forte.

Rappelons que $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ est un espace probabilisé filtré. On considère $\mathcal{N}$ une mesure aléatoire de Poisson sur $\mathbb{R}^+ \times \mathbb{R}^d \backslash \{0\}$ avec intensité $dt \otimes \nu$, où ν est une mesure de Lévy. On introduit alors $Z = (Z_t)_{t\geq 0}$ un processus de Lévy sur $\mathbb{R}^d$ qui s'écrit, pour tout $t \geq 0$,

$$Z_t = \int_0^t \int_{B_1} z \widetilde{\mathcal{N}}(ds, dz) + \int_0^t \int_{B_1^c} z \, \mathcal{N}(ds, dz).$$

On suppose que pour un certain $\beta \in [1, 2]$, on a

$$\int_{B_1^c} |z|^\beta \, d\nu(z) < +\infty.$$

Cela équivaut à supposer que pour tout $t \in \mathbb{R}^+$, Z_t a un moment d'ordre β fini d'après [Sat99, Théorème 25.3].

3.1.1 Caractère bien posé

On s'intéresse dans un premier temps au caractère bien posé de l'EDS de McKean-Vlasov suivante

$$\begin{cases} dX_t = b_t(X_t, \mu_t)\, dt + \sigma_t(X_{t^-}, \mu_t)\, dZ_t, \quad t \in [0,T], \\ \mu_t := [X_t], \\ X_0 = \xi, \end{cases} \tag{3.1}$$

où T est un horizon de temps fixé, ξ une variable aléatoire $\mathcal{F}_0$-mesurable indépendante de Z admettant un moment d'ordre β fini, $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}^d$ et $\sigma : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ sont des fonctions vérifiant les hypothèses suivantes.

Hypothèse (H1). Il existe $C > 0$ telle que pour tout $t \in [0,T]$, $x, y \in \mathbb{R}^d$ et $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, on a

$$|b_t(x,\mu) - b_t(y,\nu)| + |\sigma_t(x,\mu) - \sigma_t(y,\nu)| \leq C(|x-y| + W_\beta(\mu,\nu)), \tag{3.2}$$

et

$$|b_t(x,\mu)| + |\sigma_t(x,\mu)| \leq C(1 + |x| + M_\beta(\mu)),$$

où $M_\beta(\mu) := \left(\int_{\mathbb{R}^d} |x|^\beta\, d\mu(x)\right)^{\frac{1}{\beta}}$ pour $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$.

Il s'agit des hypothèses lipschitziennes et de croissance sous-linéaire usuelles.

Théorème 3.1. *Sous l'Hypothèse (H1), il existe une unique solution forte $(X_t)_{t\in[0,T]}$ à l'EDS (3.1) pour toute donnée initiale $\xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d)$. De plus, le flot de lois marginales $(\mu_t)_{t\in[0,T]}$ appartient à $\mathcal{C}^0([0,T]; \mathcal{P}_\beta(\mathbb{R}^d))$ et on a*

$$\mathbb{E} \sup_{t\leq T} |X_t|^\beta < +\infty. \tag{3.3}$$

Remarque 3.2. On peut ajouter un terme de la forme $(Bt + \Sigma W_t)_{t\geq 0}$ à Z, où $B \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d\times d}$ est une matrice symétrique positive de taille $d \times d$, et W est un mouvement brownien standard sur $\mathbb{R}^d$. En utilisant la décomposition de Lévy-Itô du Théorème 1.5, on peut donc considérer un processus de Lévy général Z ayant un moment d'ordre $\beta \in [1,2]$ fini.

Remarque 3.3. Dans [Gra92a], Graham énonce un résultat qui inclut notre théorème pour $\beta = 1$. Cependant, la justification de l'inégalité de Burkholder-Davis-Gundy utilisée pour passer de l'équation (1.5) à (1.6) n'est pas claire car c'est le crochet oblique qui est utilisé au lieu du crochet droit qui apparaît habituellement dans cette inégalité. L'argument ne semble pas immédiat, c'est pour cela qu'on prouve d'une autre manière le Théorème 3.1.

Comparons ce résultat aux travaux déjà existants. Lorsque $\beta = 2$, le caractère bien posé de (3.1) a été prouvé par Jourdain, Méléard et Woyczynski [JMW07]. Dans ce travail, l'existence au sens faible d'une solution est prouvée, quand $\beta = 0$, à travers le problème de martingales non-linéaire. Cependant, l'unicité n'est pas montrée lorsque $\beta = 0$. Quand $\beta = 1$, un résultat similaire au Théorème 3.1 est prouvé par Graham [Gra92b, Théorème 2.2]. Les différences principales sont les suivantes. Premièrement, dans [Gra92b], il n'y a pas d'intégrale de Poisson compensée $\widetilde{\mathcal{N}}$ dans la définition de Z et $\nu(\mathbb{R}^d)$ est supposée finie. Deuxièmement, lorsque la dérive b n'est pas bornée, il est supposé dans [Gra92b] que X_0 a un moment fini d'ordre 2, ce qui n'est pas le cas dans le Théorème 3.1. Aussi, en conservant nos notations, il est supposé l'existence d'une constante $C > 0$ telle que pour tout $t \in [0,T]$, $x \in \mathbb{R}^d$ et $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, on

a
$$\left|\int_{B_1^c} \sigma_t(x,\mu) z \, d\nu(z)\right|^2 + \int_{B_1^c} |\sigma_t(x,\mu)z|^2 \, d\nu(z) \leq C(1+|x|^2). \tag{3.4}$$

Cela suggère que σ est bornée par rapport à la variable mesure. De plus, (3.4) suggère fortement que
$$\int_{B_1^c} |z|^2 \, d\nu(z) < +\infty.$$

C'est le cas par exemple si $\sigma = \mathrm{Id}$. Cette condition sur ν est équivalente au fait que pour tout $t \in \mathbb{R}^+$, Z_t a un moment d'ordre 2 fini, ce qui n'est pas supposé dans le Théorème 3.1 puisque $\beta \in [1,2]$. Dans le cas où σ est uniformément elliptique, i.e. $\sigma\sigma^*$ est définie positive uniformément par rapport à ses variables, et bornée, on renvoie à [FKM21]. Dans ce travail, Frikha, Konakov et Menozzi prouvent le caractère bien posé de (3.1) sous des hypothèses höldériennes sur les coefficients par rapport aux variables espace et mesure. Ce résultat peut être appliqué dans le cadre lipschitzien mais on ne suppose pas ici que σ est uniformément elliptique. De plus, une autre hypothèse faite dans [FKM21] est que pour tout $(t,x) \in [0,T] \times \mathbb{R}^d$, les fonctions $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto b_t(x,\mu)$ et $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto \sigma_t(x,\mu)$ ont des dérivées plates bornées sur $\mathcal{P}(\mathbb{R}^d)$. Cela assure, au moins lorsque les coefficients dépendent linéairement de la mesure, que les coefficients sont bornés par rapport à la variable mesure, ce qui n'est pas le cas dans le Théorème 3.1

Remarque 3.4. Lorsque $\beta \in (0,1)$, le résultat d'unicité du Théorème 3.1 est faux sans une hypothèse de non-dégénérescence sur σ. Donnons un contre-exemple simple en posant, pour $t \in [0,T]$, $x \in \mathbb{R}^d$ et $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$
$$b_t(x,\mu) := \int_{\mathbb{R}^d} |x|^\beta \, d\mu(x), \quad \sigma_t(x,\mu) := 0, \quad \text{et} \quad \xi := 0.$$

L'Hypothèse (H1) est clairement satisfaite. De plus, toute solution de l'EDS de McKean-Vlasov correspondante est déterministe puisque la donnée initiale l'est et qu'il n'y a pas de bruit. On remarque alors que le problème est équivalent à résoudre l'équation différentielle
$$\begin{cases} y'(t) = |y(t)|^\beta, & t \in [0,T], \\ y(0) = 0, \end{cases}$$

qui admet plusieurs solutions. Cependant, sous l'Hypothèse (H1), il existe au moins une solution forte à l'EDS de McKean-Vlasov (3.1). On renvoie à la Section 6.4 pour une preuve de ce résultat qui repose sur le théorème de point fixe de Schauder.

Idée de preuve du Théorème 3.1. On fixe $\mu = (\mu_t)_{t\in[0,T]} \in \mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d))$ et on considère l'unique solution de
$$\begin{cases} dX_t^\mu = b_t(X_t^\mu, \mu_t)\, dt + \sigma_t(X_{t^-}^\mu, \mu_t)\, dZ_t, & t \in [0,T], \\ X_0^\mu = \xi. \end{cases} \tag{3.5}$$

Les coefficients $(t,x) \in [0,T]\times\mathbb{R}^d \mapsto b_t(x,\mu_t)$ et $(t,x) \in [0,T]\times\mathbb{R}^d \mapsto \sigma_t(x,\mu_t)$ de l'EDS sont lipschitziens et à croissance au plus linéaire en espace, uniformément en temps. D'après [Fou13, Proposition 2], on a
$$\mathbb{E} \sup_{t\leq T} |X_t^\mu|^\beta < +\infty.$$

L'application

$$\phi : \begin{cases} \mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d)) & \to \mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d)) \\ \mu & \mapsto ([X_t^\mu])_{t\in[0,T]} \end{cases} \tag{3.6}$$

est donc bien définie, et le but est de montrer qu'elle admet un unique point fixe en utilisant le théorème de Banach. Pour cela, on doit estimer $\mathbb{E}\sup_{s\le t}|X_s^\mu - X_s^\nu|^\beta$, pour $t\in[0,T]$. Pour obtenir ce contrôle, on utilise la méthode utilisée par Fournier dans [Fou13, Proposition 2], qui a déjà été utilisée dans le contexte d'EDS de McKean-Vlasov par Frikha et Li dans [FL21] pour prouver l'estimation de moments (3.3). Il s'agit d'abord de contrôler $\mathbb{E}\sup_{s\le t}|X_s^\mu - X_s^\nu|^\beta$ entre les temps de sauts de taille plus grande que 1 du bruit Z, par des outils de calcul stochastique usuels (on se ramène au cadre L^2). Ensuite, on étudie ce qui se passe aux instants des sauts de taille plus grande que 1. On parvient alors à prouver qu'il existe une constante $C>0$ telle que pour tout $n\ge 1$

$$\sup_{0\le s\le T} W_\beta^\beta(\phi^n(\mu)_s,\phi^n(\nu)_s) \le C^n\left(\frac{T^n}{n!}\right)^{\frac{\beta}{2}} \sup_{0\le s\le T} W_\beta^\beta(\mu_s,\nu_s),$$

ce qui permet d'appliquer le théorème du point fixe de Banach. □

3.1.2 Propagation du chaos forte

On considère le système de N particules en interaction suivant, qui est associé à (3.1), et donné par

$$\begin{cases} dX_t^{i,N} = b_t(X_t^{i,N},\overline{\mu}_t^N)\,dt + \sigma_t(X_{t-}^{i,N},\overline{\mu}_t^N)\,dZ_t^i, \quad t\in[0,T], \quad i\in\{1,\dots,N\}, \\ \overline{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{3.7}$$

où $(Z^i,\xi^i)_{i\ge1}$ sont i.i.d. de même loi que (Z,ξ). Le système de particules est bien défini d'après [App09, Théorème 6.2.9]. On introduit alors des copies i.i.d. de l'EDS de McKean-Vlasov (3.1), notées $(X^{i,\infty})_{i\ge1}$, où les données initiales et les bruits sont respectivement $(\xi^i)_{i\ge1}$ et $(Z^i)_{i\ge1}$. On prouve alors le résultat suivant de propagation du chaos forte quantitative. On suppose que $\beta\in[1,2)$, le cas $\beta=2$ étant traité dans [JMW07] et [NBK+20].

Théorème 3.5. *L'Hypothèse (H1) est remplacée par l'hypothèse plus forte suivante. On suppose que l'Hypothèse (H1) est vérifiée avec W_1 au lieu de W_β dans l'hypothèse lipschitzienne* (3.2) *et que $\sigma = Id$. Alors, il existe une constante $C>0$ telle que pour tout $N\ge1$*

$$\sup_{i\le N}\mathbb{E}\sup_{t\le T}|X_t^{i,N}-X_t^{i,\infty}| \le C\begin{cases} N^{\frac{1}{\beta}-1}, & \textit{si} \quad d=1,2 \quad \textit{ou} \quad d\ge3 \textit{ et } \beta<\frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \textit{si} \quad d\ge3 \textit{ et } \beta>\frac{d}{d-1}, \end{cases} \tag{3.8}$$

et

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\overline{\mu}_t^N,\mu_t) \le C\begin{cases} N^{\frac{1}{\beta}-1}, & \textit{si} \quad d=1,2 \quad \textit{ou} \quad d\ge3 \textit{ et } \beta<\frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \textit{si} \quad d\ge3 \textit{ et } \beta>\frac{d}{d-1}. \end{cases} \tag{3.9}$$

Remarque 3.6. La méthode utilisée dans la preuve du Théorème 3.5 ne peut pas être appliquée avec un coefficient non-constant σ sous l'Hypothèse (H1) (voir la Remarque 3.7). Cela semble rester un problème ouvert.

Comparons ce résultat aux travaux déjà existants. Commençons par le cadre L^2. Dans [JMW07], Jourdain, Méléard et Woyczynski traitent le cas d'un bruit de Lévy général admettant un moment d'ordre 2 fini. Les auteurs exhibent des taux de convergence pour la propagation du chaos forte dans L^2 sous des hypothèses lipschitziennes sur les coefficients b et σ, par rapport à W_2 pour l'argument mesure, qui sont similaires à l'Hypothèse (H1). Toujours dans le cadre lipschitzien, par rapport à W_1 ici, on peut également mentionner Neelima et al. [NBK+20], où un résultat de propagation du chaos du même type est prouvé, relaxant les hypothèses de [JMW07]. Dans le cas où le bruit admet seulement un moment d'ordre 1, on peut citer [Gra92a], où Graham prouve la propagation du chaos faible qualitative sous des hypothèses lipschitziennes, par rapport à la distance de Wasserstein W_1 pour l'argument mesure. L'EDS de McKean-Vlasov est dirigée par une mesure aléatoire de Poisson et sa mesure de Poisson compensée. Dans le cas unidimensionnel $d = 1$, Frikha et Li [FL21] s'intéressent à une EDS de McKean-Vlasov SDE dirigée par une mesure de Poisson compensée et ayant des sauts positifs uniquement. Les auteurs prouvent une estimation de propagation du chaos forte dans L^1, sous des hypothèses lipschitziennes "d'un seul côté" en espace sur les coefficients, par rapport à W_1 pour l'argument mesure. On obtient la même vitesse de convergence dans notre résultat. Cela n'est pas étonnant car elle provient directement de [FG15].

Preuve du Théorème 3.5. Pour prouver (3.8), on écrit pour $t \in [0, T]$

$$X_t^{i,N} - X_t^{i,\infty} = \int_0^t b_s(X_s^{i,N}, \overline{\mu}_s^N) - b_s(X_s^{i,\infty}, \mu_s)\, ds.$$

En utilisant l'hypothèse lipschitzienne sur b, il existe $C > 0$ telle que pour tout $t \in [0, T]$

$$\begin{aligned}
&\sup_{i \le N} \mathbb{E} \sup_{r \le t} |X_r^{i,N} - X_r^{i,\infty}| \\
&\le C \int_0^t \sup_{i \le N} \mathbb{E}|X_s^{i,N} - X_s^{i,\infty}|\, ds + C \int_0^t \mathbb{E} W_1(\overline{\mu}_s^N, \mu_s)\, ds \\
&\le C \int_0^t \sup_{i \le N} \mathbb{E}|X_s^{i,N} - X_s^{i,\infty}|\, ds + C \int_0^t \mathbb{E} W_1(\overline{\mu}_s^N, \tilde{\mu}_s^N) + \mathbb{E} W_1(\tilde{\mu}_s^N, \mu_s)\, ds,
\end{aligned}$$

où $\tilde{\mu}_s^N := \frac{1}{N} \sum_{k=1}^N \delta_{X_s^{k,\infty}}$ est la mesure empirique associée aux $(X^{i,\infty})_{i \ge 1}$. Puisque

$$W_1(\overline{\mu}_s^N, \tilde{\mu}_s^N) \le \frac{1}{N} \sum_{k=1}^N |X_s^{k,N} - X_s^{k,\infty}|,$$

l'inégalité de Gronwall assure qu'il existe $C > 0$ telle que pour tout $N \ge 1$

$$\sup_{i \le N} \mathbb{E} \sup_{t \le T} |X_t^{i,N} - X_t^{i,\infty}| \le C \int_0^T \mathbb{E} W_1(\tilde{\mu}_s^N, \mu_s)\, ds. \tag{3.10}$$

On conclut alors en utilisant [FG15, Théorème 1] puisque $(X^{i,\infty})_{i \ge 1}$ est une suite i.i.d. et que

$$\sup_{i \ge 1} \sup_{t \in [0,T]} \mathbb{E}|X_t^{i,\infty}|^\beta < +\infty,$$

par l'inégalité de Gronwall. L'inégalité (3.9) découle de (3.8) et [FG15] en écrivant

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}W_1(\overline{\mu}_t^N,\mu_t) &\leq \sup_{t\in[0,T]} \mathbb{E}W_1(\overline{\mu}_t^N,\tilde{\mu}_t^N) + \sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N,\mu_t)\\ &\leq \sup_{t\in[0,T]}\sup_{i\leq N}\mathbb{E}|X_t^{i,N}-X_t^{i,\infty}| + \sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N,\mu_t).\end{aligned}$$

□

Remarque 3.7. La méthode ne semble pas se généraliser au cas d'un coefficient de diffusion σ non-constant sans hypothèse supplémentaire sur Z. En effet, d'après l'inégalité de BDG, on aurait

$$\begin{aligned}&\mathbb{E}\left|\int_0^t\int_{B_1}\sigma_s(X_s^{i,N},\overline{\mu}_s^N)z-\sigma_s(X_s^{i,\infty},\mu_s)z\widetilde{\mathcal{N}}(ds,dz)\right|\\ &\qquad\qquad\leq C\mathbb{E}\left(\int_0^t\int_{B_1}|\sigma_s(X_s^{i,N},\overline{\mu}_s^N)-\sigma_s(X_s^{i,\infty},\mu_s)|^2|z|^2\,\mathcal{N}(ds,dz)\right)^{\frac{1}{2}}.\end{aligned}$$

A priori, on ne peut pas contrôler cette quantité par

$$\int_0^t \sup_{i\leq N}\mathbb{E}|X_s^{i,N}-X_s^{i,\infty}|+\mathbb{E}W_1(\overline{\mu}_s^N,\mu_s)\,ds,$$

puisque $\int_{B_1}|z|\,d\nu(z)$ n'est pas nécessairement finie. Par l'inégalité de Jensen, on peut obtenir un contrôle par

$$\left[\int_0^t \sup_{i\leq N}\mathbb{E}|X_s^{i,N}-X_s^{i,\infty}|^2+\mathbb{E}W_1(\overline{\mu}_s^N,\mu_s)^2\,ds\right]^{\frac{1}{2}},$$

ce qui ne suffit pas pour conclure.

On étudie maintenant un cas particulier pour lequel on peut améliorer les taux de convergence obtenus dans le Théorème 3.5. Supposons que $Z=(Z_t)_{t\geq 0}$ est un processus α-stable sur $\mathbb{R}^d$, avec $\alpha\in(1,2)$. Soient $A, A', B\in\mathbb{R}^{d\times d}$ des matrices de taille $d\times d$. Le système de particules considéré est (3.7) avec

$$\xi\in L^\alpha(\Omega,\mathcal{F}_0;\mathbb{R}^d),\quad b_t(x,\mu):=Ax+A'\int_{\mathbb{R}^d}y\,d\mu(y)\quad\text{et}\quad\sigma_t(x,\mu):=B.$$

Cela correspond à un système de processus d'Ornstein-Uhlenbeck α-stables en interaction. En gardant les mêmes notation, que pour le Théorème 3.5, on a le résultat suivant de propagation du chaos forte quantitative.

Théorème 3.8. *Il existe une constante $C>0$ telle que pour tout $N\geq 1$*

$$\sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X_t^{i,N}-X_t^{i,\infty}|\leq C\begin{cases}(\ln(N))^{\frac{1}{\alpha}}N^{\frac{1}{\alpha}-1}, & si\quad d=1,2\quad ou\quad d\geq 3\ et\ \alpha<\frac{d}{d-1},\\ N^{-\frac{1}{d}}, & si\quad d\geq 3\ et\ \alpha>\frac{d}{d-1},\end{cases}\tag{3.11}$$

et

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\overline{\mu}_t^N,\mu_t)\leq C\begin{cases}(\ln(N))^{\frac{1}{\alpha}}N^{\frac{1}{\alpha}-1}, & si\quad d=1,2\quad ou\quad d\geq 3\ et\ \alpha<\frac{d}{d-1},\\ N^{-\frac{1}{d}}, & si\quad d\geq 3\ et\ \alpha>\frac{d}{d-1}.\end{cases}\tag{3.12}$$

Idée de preuve du Théorème 3.8. La preuve de (3.11) consiste à raffiner la preuve de (3.8). L'idée clé est de retirer les sauts de taille plus grande que le nombre de particules N de tous les bruits dans une première étape. On définit donc, pour $i \geq 1$ et $t \in [0,T]$

$$Z^i_{N,t} := \int_0^t \int_{B_N} z \widetilde{\mathcal{N}}^i(ds, dz),$$

où $\widetilde{\mathcal{N}}^i$ est la mesure de Poisson compensée associée à Z^i. On considère alors, pour $i \in \{1, \dots, N\}$, $X^{i,\infty}_N$ l'unique solution de

$$\begin{cases} dX^{i,\infty}_{N,t} &= AX^{i,\infty}_{N,t}\,dt + A'\mathbb{E}X^{i,\infty}_{N,t}\,dt + B\,dZ^i_{N,t}, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ \mu_{N,t} &:= [X^{i,\infty}_{N,t}], \\ X^{i,\infty}_{N,0} &= \xi^i. \end{cases} \tag{3.13}$$

Notons que les variables aléatoires $(X^{i,\infty}_N)_{i \leq N}$ sont i.i.d. pour $N \geq 1$ fixé. On fait de même avec le système de particules en introduisant $(X^{i,N}_N)_{i \leq N}$ l'unique solution de

$$\begin{cases} dX^{i,N}_{N,t} = AX^{i,N}_{N,t}\,dt + A'\frac{1}{N}\sum\limits_{k=1}^N X^{k,N}_{N,t}\,dt + B\,dZ^i_{N,t}, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ \overline{\mu}^N_{N,t} := \frac{1}{N}\sum\limits_{j=1}^N \delta_{X^{j,N}_{N,t}}, \\ X^{i,N}_{N,0} = \xi^i, \end{cases} \tag{3.14}$$

Le premier objectif est de contrôler l'erreur L^1 entre $X^{i,N}_N$ et $X^{i,N}$ et également entre $X^{i,\infty}_N$ et $X^{i,\infty}$ pour tout $i \in \{1,\dots,N\}$. On prouve que ces erreurs sont de l'ordre de $N^{1-\alpha}$, ce qui correspond au taux de décroissance de

$$\int_{|z|>N} |z|\,d\nu(z),$$

et on déduit que

$$\sup_{i \leq N} \mathbb{E} \sup_{t \leq T} |X^{i,N}_t - X^{i,\infty}_t| \leq CN^{1-\alpha} + \sup_{i \leq N} \mathbb{E} \sup_{t \leq T} |X^{i,N}_{N,t} - X^{i,\infty}_{N,t}|. \tag{3.15}$$

Le dernier terme du membre de droite est contrôlé grâce à [FG15]. Dans le cas où $d = 1$, $d = 2$ ou $d \geq 3$ et $\alpha < d/(d-1)$, on travaille dans $\mathcal{P}_\alpha(\mathbb{R}^d)$ et on montre que

$$\sup_{t \leq T} \mathbb{E}|X^{i,\infty}_{N,t}|^\alpha \leq C\ln(N).$$

Cela vient du fait que

$$\int_{1 \leq |z| \leq N} |z|^\alpha\,d\nu(z) = \underset{N \to +\infty}{O}(\ln(N)).$$

Dans les autres cas, on travaille dans $\mathcal{P}_\beta(\mathbb{R}^d)$, pour $\beta \in (1,\alpha)$ bien choisi.

□

3.2 Caractère bien posé et propagation du chaos faible pour une EDS de McKean-Vlasov dirigée par un processus α-stable : cadre höldérien

Dans cette section, on présente les résultats des travaux [Cav22a] et [Cav22b].

3.2.1 Présentation du problème

Commençons par introduire le problème général. Soit $Z = (Z_t)_{t\geq 0}$ un processus α-stable sur $\mathbb{R}^d$, avec $\alpha \in (1,2)$, qu'on suppose rotationnellement invariant. On s'intéresse à l'EDS de McKean-Vlasov suivante

$$\begin{cases} dX_t = b(t, X_t, [X_t])\, dt + dZ_t, \quad t \in [0,T], \\ \mu_t := [X_t], \\ X_0 = \xi, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d), \end{cases} \tag{3.16}$$

où $T > 0$ est fixé, $[\xi]$ est la loi de la variable aléatoire ξ qui est indépendante de Z et avec un coefficient de dérive général $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$. Les hypothèses précises sur b seront précisées dans la suite. L'objectif final est de prouver la propagation de chaos pour le système de particules en interaction de type champ moyen associé et défini, pour $N \geq 1$, par

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\, dt + dZ_t^i, \quad t \in [0,T], \quad i \in \{1, \dots, N\}, \\ \overline{\mu}_t^N := \frac{1}{N} \sum_{j=1}^{N} \delta_{X_t^{j,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{3.17}$$

où $(Z^i, \xi^i)_{i\geq 1}$ est une suite i.i.d. de même loi que (Z, ξ).

Comme évoqué dans la Sous-section 1.2.2, la propagation du chaos peut être abordée au sens faible, c'est-à-dire au niveau de la convergence en loi au travers la convergence de la mesure empirique $\overline{\mu}^N$, ou au sens fort, c'est-à-dire au niveau des trajectoires par couplage (voir les Définitions 1.15 et 1.17). La propagation du chaos faible quantitative consiste plus précisément dans notre cadre à trouver un taux de convergence explicite par rapport à N pour $\mathbb{E}|\phi(\overline{\mu}_t^N) - \phi(\mu_t)|$ et $|\mathbb{E}(\phi(\overline{\mu}_t^N) - \phi(\mu_t))|$, pour ϕ dans une classe suffisamment large de fonctions définies sur un espace de mesures (voir la Définition 1.18). Ce sont de tels contrôles qu'on prouve dans [Cav22a, Cav22b]. On travaille dans [Cav22b] sous les hypothèses suivantes.

Hypothèse (H2).

(1) Le coefficient de dérive b est continu et globalement borné sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$.

(2) Pour tout $(t, \mu) \in [0,T] \times \mathcal{P}(\mathbb{R}^d)$, l'application $b(t, \cdot, \mu)$ est η-höldérienne sur $\mathbb{R}^d$, pour un certain $\eta \in (0,1]$ fixé et uniformément par rapport à $(t, \mu) \in [0,T] \times \mathcal{P}(\mathbb{R}^d)$, c'est-à-dire qu'il existe $C > 0$ tel que pour tout $t \in [0,T]$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ et $x_1, x_2 \in \mathbb{R}^d$

$$|b(t, x_1, \mu) - b(t, x_2, \mu)| \leq C|x_1 - x_2|^\eta.$$

(3) Pour tout $(t, x) \in [0,T] \times \mathbb{R}^d$, l'application $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto b(t, x, \mu)$ admet une dérivée plate telle que $\frac{\delta}{\delta m} b(t, x, \mu)(\cdot)$ est η-höldérienne sur $\mathbb{R}^d$ uniformément par rapport à $(t, x, \mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ et $\frac{\delta}{\delta m} b$ est bornée sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$.

(4) Pour tout $(t, x, v) \in [0,T] \times (\mathbb{R}^d)^2$, l'application $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto \frac{\delta}{\delta m} b(t, x, \mu)(v)$ admet une dérivée plate $\frac{\delta^2}{\delta m^2} b(t, x, \mu)(v, \cdot)$ qui est η-höldérienne uniformément par rapport à $(t, x, \mu, v) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$ et $\frac{\delta^2}{\delta m^2} b$ est bornée sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times (\mathbb{R}^d)^2$.

Ces hypothèses sont analogues à celles faites dans [CdRF21] dans le cas brownien. Les deux principaux objectifs de [Cav22a] et [Cav22b] sont de montrer que la méthode présentée dans la Sous-section 1.2.3 se généralise dans le cas d'EDS à sauts, et d'observer quels sont les changements sur les taux de convergence pour la propagation du chaos faible par rapport au cas brownien traité dans [CdRF21]. Le difficultés principales sont les suivantes. Premièrement, la dépendance par rapport à la mesure du coefficient de dérive b est générale ce qui requiert le calcul différentiel sur l'espaces des mesures de probabilités présenté dans le Chapitre 2. Deuxièmement, comme b n'est pas lipschitzien, on doit bénéficier d'un effet de *régularisation par le bruit*. Troisièmement, comme Z n'admet pas un moment d'ordre 2 fini, une difficulté supplémentaire est l'impossibilité de travailler dans L^2. On ne peut donc pas utiliser les outils développés dans ce cadre pour les jeux à champ moyen. De plus, la présence de sauts induit des difficultés supplémentaires pour développer ces outils.

On décrit maintenant la structure de la présentation des travaux [Cav22a] et [Cav22b]. Comme on l'a vu dans la Sous-section 1.2.3 qui présente la méthode basée sur le semi-groupe pour établir la propagation du chaos faible quantitative, un des outils essentiels est la formule d'Itô le long d'un flot de mesures de probabilité. Dans un premier temps, on présente donc la formule d'Itô le long de flots de lois marginales de processus à sauts définis par des intégrales contre une mesure de Poisson et une mesure de Poisson compensée. Celle-ci est prouvée dans [Cav22a]. L'autre objet fondamental de cette méthode est évidemment le semi-groupe associé à (3.16), dont la dynamique est décrite à l'aide de la formule d'Itô. Ainsi, dans un second temps, on va donc s'intéresser au caractère bien posé de (3.16), ce qui est nécessaire pour établir la propagation du chaos et qui permet de définir le semi-groupe (voir (1.10)). Cela est prouvé dans [Cav22b] sous des hypothèses plus faibles que l'Hypothèse (H2). On a également vu dans la Sous-section 1.2.3 que la méthode requiert une certaine régularité sur le semi-groupe, des estimations sur celui-ci, et de connaître sa dynamique qui est décrite par une EDP de Kolmogorov rétrograde. Ces propriétés reposent sur l'étude de la densité de transition associée à (3.16), en particulier sur des estimations sur ses dérivées par rapport aux variables espace et mesure. La Proposition 1.14 donnant des estimations de gradient sur la densité de transition de Z s'avère cruciale à cette étape, et c'est pour cela qu'on se restreint ici au cas où Z est rotationnellement invariant. Ces propriétés sur la densité de transition de (3.16) seront présentées juste avant d'en déduire la régularité et les estimations qui en découlent sur le semi-groupe, ce qui a été prouvé dans [Cav22b] sous l'Hypothèse (H2). On présente ensuite les résultats de propagation du chaos faible quantitative obtenus dans [Cav22b] pour le système de particules (3.17) sous l'Hypothèse (H2), ainsi que dans [Cav22a] pour un système particulier de processus d'Ornstein-Uhlenbeck α-stables en interaction de type champ moyen, avec $\alpha \in (1, 2)$. Enfin, on termine par présenter un dernier résultat issu de [Cav22b]. Il s'agit de quantifier la vitesse d'approximation de la loi d'une particule de (3.17) par la loi du processus de McKean-Vlasov limite au niveau des densités. On prouve une estimation ponctuelle entre la densité d'une particule et la densité du processus limite, ce qui permet de quantifier la vitesse d'approximation au niveau des lois en variation totale. La méthode repose sur des idées analogues à celles présentées dans la Sous-section 1.2.3, mais en utilisant la densité de transition de l'EDS de McKean-Vlasov au lieu du semi-groupe.

3.2.2 Formule d'Itô le long d'un flot de mesures associé à des intégrales stochastiques de Poisson

On présente ici la formule d'Itô pour un flot de mesures associé à un processus à sauts, qui est démontrée dans [Cav22a]. Comme expliqué précédemment, il s'agit d'un des outils essentiels qu'on utilise dans [Cav22a] et [Cav22b] pour prouver des estimations de propagation du chaos faible.

On fixe $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ un espace probabilisé filtré. On introduit une mesure aléatoire de Poisson $\mathcal{N}$ sur $[0,T]\times\mathbb{R}^d\backslash\{0\}$ avec intensité $dt\otimes\nu$, où ν est une mesure de Lévy, et on note $\widetilde{\mathcal{N}}(ds,dz) := \mathcal{N}(ds,dz) - ds\,d\nu(z)$ la mesure compensée associée à $\mathcal{N}$. On s'intéresse ici au processus à sauts $X = (X_t)_{t\in[0,T]}$ défini par

$$\forall t \in [0,T],\; X_t := X_0 + \int_0^t b_s\,ds + \int_0^t \int_{B_1} H(s,z)\,\widetilde{\mathcal{N}}(ds,dz) + \int_0^t \int_{B_1^c} K(s,z)\,\mathcal{N}(ds,dz), \tag{3.18}$$

où $b : [0,T]\times\Omega \to \mathbb{R}^d$, $H : [0,T]\times B_1\times\Omega \to \mathbb{R}^d$, et $K : [0,T]\times B_1^c\times\Omega \to \mathbb{R}^d$ sont des processus prévisibles. La loi de X_t est notée μ_t.

Fixons les hypothèses sur le processus $(X_t)_{t\in[0,T]}$. Soient $\beta\in(0,2]$ et $\gamma\in[0,1]$ tel que $\beta\leq 1+\gamma$. On fait les hypothèses suivantes.

(M) La variable aléatoire X_0 appartient à $L^\beta(\Omega,\mathcal{F}_0;\mathbb{R}^d)$ et on a

$$\mathbb{E}\int_0^T |b_s|^{\beta\vee 1}\,ds < +\infty. \tag{3.19}$$

(J1) Il existe un processus prévisible $(\widetilde{H}_s)_{s\in[0,T]}$, supposé localement borné presque sûrement, et tel que

$$\text{p.s.}\; \forall s\in[0,T],\; \forall z\in B_1,\; |H(s,z)| \leq |\widetilde{H}_s||z| \quad \text{et} \quad \mathbb{E}\int_0^T \int_{B_1} (|\widetilde{H}_s||z|)^{1+\gamma}\,d\nu(z)\,ds < +\infty. \tag{3.20}$$

(J2) On a

$$\mathbb{E}\int_0^T \int_{B_1^c} |K(s,z)|^\beta\,d\nu(z)\,ds < +\infty. \tag{3.21}$$

Les Hypothèses **(M)**, **(J1)** et **(J2)** et la condition $\beta\leq 1+\gamma$ permettent d'assurer que $\mu_t\in\mathcal{P}_\beta(\mathbb{R}^d)$ pour tout $t\in[0,T]$. Comme on le verra dans la suite, l'Hypothèse **(J1)** est, de plus, liée à la régularité höldérienne en espace du gradient de la dérivée plate de la fonction pour laquelle on veut prouver la formule d'Itô.

On énonce maintenant la formule d'Itô le long du flot de mesures $(\mu_t)_{t\in[0,T]}$.

Théorème 3.9 (Formule d'Itô)**.** *On suppose que les Hypothèses (**M**), **(J1)**, et **(J2)** sont satisfaites. Soit $u : \mathcal{P}_\beta(\mathbb{R}^d)\to\mathbb{R}$ une fonction admettant une dérivée plate $\frac{\delta}{\delta m}u$ vérifiant les propriétés suivantes.*

(1) Pour tout $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, la fonction $\frac{\delta}{\delta m}u(\mu)\in\mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ et $\partial_v\frac{\delta}{\delta m}u$ est continue sur $\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$.

(2) Si $\gamma>0$, pour tout compact $\mathcal{K}\subset\mathcal{P}_\beta(\mathbb{R}^d)$, il existe $C_\mathcal{K}>0$ tel que

$$\forall \mu\in\mathcal{K},\; \forall x,y\in\mathbb{R}^d,\; \left|\partial_v\frac{\delta}{\delta m}u(\mu)(x) - \partial_v\frac{\delta}{\delta m}u(\mu)(y)\right| \leq C_\mathcal{K}|x-y|^\gamma.$$

(3) Pour tout compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, on a

$$\begin{cases} \sup_{\mu \in \mathcal{K}} \int_{\mathbb{R}^d} \left| \partial_v \frac{\delta}{\delta m} u(\mu)(v) \right|^{\beta'} d\mu(v) < +\infty & si\ \beta > 1, \\ \sup_{v \in \mathbb{R}^d} \sup_{\mu \in \mathcal{K}} \left| \partial_v \frac{\delta}{\delta m} u(\mu)(v) \right| < +\infty & si\ \beta \leq 1, \end{cases}$$

où β' est l'exposant conjugué de β. On a alors, pour tout $t \in [0,T]$,

$$\begin{aligned} & u(\mu_t) - u(\mu_0) \\ & = \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s\right) ds \\ & + \int_0^t \int_{B_1^c} \mathbb{E}\left[\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right] d\nu(z)\, ds \\ & + \int_0^t \int_{B_1} \mathbb{E}\left[\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}) - \partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}) \cdot H(s,z)\right] d\nu(z)\, ds. \end{aligned} \tag{3.22}$$

Les hypothèses (1) et (3) sont naturelles pour assurer que le premier terme du membre de droite de (3.22) ait bien un sens. Ces hypothèses sont analogues à celles de la formule d'Itô du Théorème 2.28 dans le cas brownien. L'hypothèse (2) est directement liée à l'Hypothèse **(J1)** sur le processus considéré. Ces deux hypothèses permettent de montrer que le dernier terme du membre de droite de (3.22) est bien défini.

Remarque 3.10. L'hypothèse (3) est impliquée par l'hypothèse plus forte : pour tout compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, il existe $C_\mathcal{K} > 0$ tel que

$$\forall v \in \mathbb{R}^d,\ \sup_{\mu \in \mathcal{K}} \left| \partial_v \frac{\delta}{\delta m} u(\mu)(v) \right| \leq C_\mathcal{K}(1 + |v|^{\beta - 1} \mathbf{1}_{\beta > 1}).$$

Lorsque $\beta > 1$, cela découle de l'inégalité de Hölder.

Remarque 3.11. Il existe déjà dans la littérature des formules d'Itô le long d'un flot de mesures associé à une semi-martingale générale à sauts. Cela a été établi indépendamment par Guo, Pham et Wei [GPW20] et par Talbi, Touzi et Zhang [TTZ21]. Un point commun entre ces deux travaux est que le processus considéré est supposé avoir un moment d'ordre 2 fini, ce qui n'est pas le cas dans le Théorème 3.9. Ce cadre L^2 n'est pas adapté quand la mesure de Poisson $\mathcal{N}$ est associée à un processus α-stable avec $\alpha \in (0,2)$ puisqu'il possède seulement des moments finis d'ordre $\beta < \alpha$. La formule d'Itô du Théorème 3.9 peut être utilisée dans ce cadre puisque $\beta \in (0,2]$. Une autre différence entre notre formule d'Itô et [GPW20] est que l'on ne suppose pas que $\partial_v \frac{\delta}{\delta m} u$ est bornée si $\beta > 1$. De plus, dans [TTZ21], il est supposé que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, la fonction $\frac{\delta}{\delta m} u(\mu)$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}^d$, ce qui n'est pas le cas ici. Cette hypothèse est remplacée dans notre cadre par le fait que $\partial_v \frac{\delta}{\delta m} u(\mu)$ est supposée γ-höldérienne uniformément par rapport à $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, où γ apparaît dans l'Hypothèse **(J1)** qui est liée à l'intégrabilité du processus sur les petits sauts.

Idée de preuve du Théorème 3.9. On commence par localiser le processus en introduisant la suite de temps d'arrêt $(T_n)_{n \geq 1}$ définie, pour $n \geq 1$, par

$$T_n := \inf\{t \in [0,T],\ |X_t| \geq n \quad \text{ou} \quad |\widetilde{H}_s| \geq n\} \wedge T,$$

où $\widetilde{H}$ a été introduit dans l'Hypothèse **(J1)**. On définit alors le processus localisé X^n et son flot de mesure μ^n, pour tout $t \in [0,T]$, par

$$X^n_t := X_{t\wedge T_n} \quad \text{et} \quad \mu^n_t := \mathcal{L}(X^n_t). \tag{3.23}$$

Comme X est càdlàg et $\widetilde{H}$ est presque sûrement borné, $T_n = T$ pour n plus grand qu'une constante aléatoire. Le processus X^n approche X au sens où presque sûrement $\sup_{t\leq T} |X^n_t - X_t| \underset{n\to+\infty}{\longrightarrow} 0$. De plus, pour tout $t \in [0,T]$, $\mu^n_t \overset{\mathcal{P}_\beta(\mathbb{R}^d)}{\longrightarrow} \mu_t$.

On établit alors la formule d'Itô le long du flot de mesures $(\mu^n_t)_t$. Pour cela, on subdivise l'intervalle $[0,t]$ en posant, pour $m \geq 1$ et $k \in \{0,\ldots,m\}$

$$t^m_k := \frac{k}{m} t.$$

En linéarisant la somme télescopique des accroissements de u le long de la subdivision $(t^m_k)_k$ avec la dérivée plate et en utilisant la formule d'Itô classique en espace appliquée à la dérivée plate, on a

$$\begin{aligned} u(\mu^n_t) - u(\mu^n_0) &= \sum_{k=0}^{m-1} u(\mu^n_{t_{k+1}}) - u(\mu^n_{t_k}) \\ &= I_1 + I_2 + I_3, \end{aligned}$$

où

$$\begin{aligned} I_1 &:= \sum_{k=0}^{m-1} \int_0^1 \mathbb{E} \int_{t^m_k\wedge T_n}^{t^m_{k+1}\wedge T_n} \partial_v \frac{\delta}{\delta m} u(M^k_r)(X^n_s) \cdot b_s \, ds \, dr, \\ I_2 &:= \sum_{k=0}^{m-1} \int_0^1 \mathbb{E} \int_{t^m_k\wedge T_n}^{t^m_{k+1}\wedge T_n} \int_{B_1^c} \left[\frac{\delta}{\delta m} u(M^k_r)(X^n_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(M^k_r)(X^n_{s^-}) \right] d\nu(z) \, ds \, dr, \\ I_3 &:= \sum_{k=0}^{m-1} \int_0^1 \mathbb{E} \int_{t^m_k\wedge T_n}^{t^m_{k+1}\wedge T_n} \int_{B_1} \left[\frac{\delta}{\delta m} u(M^k_r)(X^n_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(M^k_r)(X^n_{s^-}) \right. \\ &\qquad \left. - H(s,z) \cdot \partial_v \frac{\delta}{\delta m} u(M^k_r)(X^n_{s^-}) \right] d\nu(z) \, ds \, dr, \end{aligned}$$

où $M^k_r := r\mu^n_{t^m_{k+1}} + (1-r)\mu^n_{t^m_k}$. En laissant tendre m vers l'infini par continuité, on déduit que pour tout $n \geq 1$ et $t \in [0,T]$

$$\begin{aligned} &u(\mu^n_t) - u(\mu^n_0) \\ &= \mathbb{E} \int_0^{t\wedge T_n} \left(\partial_v \frac{\delta}{\delta m} u(\mu^n_s)(X^n_s) \cdot b_s \right) ds \\ &\quad + \mathbb{E} \int_0^{t\wedge T_n} \int_{B_1^c} \left[\frac{\delta}{\delta m} u(\mu^n_s)(X^n_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(\mu^n_s)(X^n_{s^-}) \right] d\nu(z) \, ds \\ &\quad + \mathbb{E} \int_0^{t\wedge T_n} \int_{B_1} \left[\frac{\delta}{\delta m} u(\mu^n_s)(X^n_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(\mu^n_s)(X^n_{s^-}) - \partial_v \frac{\delta}{\delta m} u(\mu^n_s)(X^n_{s^-}) \cdot H(s,z) \right] d\nu(z) \, ds. \end{aligned} \tag{3.24}$$

Il reste ensuite à faire tendre n vers l'infini, essentiellement par continuité et convergence dominée grâce aux résultats d'approximation par localisation du processus X. L'hypothèse höldérienne sur $\partial_v \frac{\delta}{\delta m} u$

intervient car on utilise, pour la domination, que presque sûrement pour tout $s \in [0,t]$, $z \in B_1$, $n \in \mathbb{N}$

$$\left|\frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n + H(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n) - \partial_v \frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n) \cdot H(s,z)\right| \leq C|H(s,z)|^{1+\gamma}.$$

Cela découle de la formule de Taylor avec reste intégral. On peut alors conclure grâce à l'Hypothèse **(J1)**. Les Hypothèses **(M)** et **(J2)** permettent de passer à la limite $n \to +\infty$ dans les deux premiers termes du membre de droite de (3.24).

□

On peut également étendre la formule d'Itô du Théorème 3.9 au cas où la fonction u dépend aussi du temps et de l'espace. On introduit une autre mesure aléatoire de Poisson $\mathcal{M}$ sur $[0,T] \times \mathbb{R}^d$ admettant π comme mesure de Lévy. On définit alors, pour $t \in [0,T]$

$$Y_t := Y_0 + \int_0^t \kappa_s^2 \, ds + \int_0^t \int_{B_1} I(s,z) \, \widetilde{\mathcal{M}}(ds,dz) + \int_0^t \int_{B_1^c} J(s,z) \, \mathcal{M}(ds,dz),$$

où Y_0 est $\mathcal{F}_0$-mesurable et κ, I et J sont des processus prévisibles avec

$$\int_0^T |\kappa_s| \, ds + \int_0^T \int_{B_1} |I(s,z)|^{1+\Gamma} + |J(s,z)|^2 \, d\pi(z) \, ds < +\infty \quad \text{p.s.}$$

où $\Gamma \in (0,1]$ est fixé.

Théorème 3.12 (Extension de la formule d'Itô)**.** *A la place de* (3.19), (3.20) *et* (3.21) *dans les Hypothèses **(M)**, **(J1)** et **(J2)**, on suppose que*

$$\mathbb{E}|X_0|^\beta + \mathbb{E}\sup_{t \leq T} |b_s|^{\beta \vee 1} + \int_{B_1^c} \mathbb{E}\sup_{t \leq T} |K(t,z)|^\beta \, d\nu(z) + \int_{B_1} \mathbb{E}\sup_{t \leq T} (|\widetilde{H}_t||z|)^{1+\gamma} \, d\nu(z) < +\infty, \tag{3.25}$$

avec $\beta \leq 1+\gamma$. De plus, on suppose que pour tout $t \in [0,T]$ et $z \in \mathbb{R}^d$, $K(\cdot,z)$ et $H(\cdot,z)$ sont presque sûrement continus en t.

Soit $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ une fonction continue qui vérifie les propriétés suivantes.

(1) Pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ $u(\cdot,x,\mu)$ est continument dérivable et $\partial_t u$ est continue sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$. De plus, pour tout $t \in [0,T]$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $u(t,\cdot,\mu)$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^d$ avec $\partial_x u$ continue sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$ et tel que $\partial_x u(t,\cdot,\mu)$ est Γ-höldérienne uniformément par rapport à t et μ.

(2) Pour tout $t \in [0,T]$, $x \in \mathbb{R}^d$, la fonction $u(t,x,\cdot)$ admet une dérivée plate $\frac{\delta}{\delta m}u(t,x,\cdot)(\cdot)$ telle que $\frac{\delta}{\delta m}u$ est continue sur $[0,T] \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ et pour tout compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, il existe $C_\mathcal{K} > 0$ tel que

$$\forall t \in [0,T], \, \forall (x,\mu) \in \mathcal{K}, \, \forall v \in \mathbb{R}^d, \, \left|\frac{\delta}{\delta m}u(t,x,\mu)(v)\right| \leq C_\mathcal{K}(1+|v|^\beta).$$

De plus, on suppose que pour tout $t \in [0,T]$, $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $\frac{\delta}{\delta m}u(t,x,\mu)$ est différentiable et que $\partial_v \frac{\delta}{\delta m}u$ est continue sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.

(3) Si $\gamma > 0$, pour tout compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, il existe $C_\mathcal{K} > 0$ tel que

$$\forall t \in [0,T], \, \forall (x,\mu) \in \mathcal{K}, \, \forall v,v' \in \mathbb{R}^d, \, \left|\partial_v \frac{\delta}{\delta m}u(t,x,\mu)(v) - \partial_v \frac{\delta}{\delta m}u(t,x,\mu)(v')\right| \leq C_\mathcal{K}|v-v'|^\gamma.$$

(4) Pour tout compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, on a

$$\begin{cases} \sup_{t\in[0,T]} \sup_{(x,\mu)\in\mathcal{K}} \int_{\mathbb{R}^d} \left|\partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v)\right|^{\beta'} d\mu(v) < +\infty \quad \text{si } \beta > 1, \\ \sup_{t\in[0,T]} \sup_{v\in\mathbb{R}^d} \sup_{(x,\mu)\in\mathcal{K}} \left|\partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v)\right| < +\infty \quad \text{si } \beta \le 1. \end{cases}$$

Alors, la fonction $(t,x) \in [0,T] \times \mathbb{R}^d \mapsto u(t,x,\mu_t)$ est de classe $\mathcal{C}^1$ et $\partial_x u(t,\cdot,\mu_t)$ Γ-höldérienne uniformément par rapport à t. De plus, on a presque sûrement pour tout $t \in [0,T]$

$$\begin{aligned} &u(t,Y_t,\mu_t) - u(0,Y_0,\mu_0) \\ &= \int_0^t \partial_t u(s,Y_s,\mu_s)\,ds + \int_0^t \overline{\mathbb{E}}\left(\partial_v \frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_s)\cdot \bar{b}_s\right) ds \\ &\quad + \int_0^t \int_{B_1^c} \overline{\mathbb{E}}\left[\frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_{s^-} + \overline{K}(s,z)) - \frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_{s^-})\right] d\nu(z)\,ds \\ &\quad + \int_0^t \int_{B_1} \overline{\mathbb{E}}\left[\frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_{s^-} + \overline{H}(s,z)) - \frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_{s^-}) \right. \\ &\qquad\qquad\qquad \left. - \partial_v \frac{\delta}{\delta m} u(s,Y_s,\mu_s)(\overline{X}_{s^-})\cdot \overline{H}(s,z)\right] d\nu(z)\,ds \qquad (3.26) \\ &\quad + \int_0^t \partial_x u(s,Y_s,\mu_s)\cdot \kappa_s\,ds + \int_0^t \int_{B_1^c} u(s,Y_{s^-} + J(s,z),\mu_s) - u(s,Y_{s^-},\mu_s)\,\mathcal{M}(ds,dz) \\ &\quad + \int_0^t \int_{B_1} u(s,Y_{s^-} + I(s,z),\mu_s) - u(s,Y_{s^-},\mu_s)\,\widetilde{\mathcal{M}}(ds,dz) \\ &\quad + \int_0^t \int_{B_1} u(s,Y_{s^-} + I(s,z),\mu_s) - u(s,Y_{s^-},\mu_s) - \partial_x u(s,Y_{s^-},\mu_s)\cdot I(s,z)\,d\pi(z)\,ds, \end{aligned}$$

où $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ est une copie indépendante de $(\Omega,\mathcal{F},\mathbb{P})$ et $(\bar{b},\overline{H},\overline{K},\overline{X})$est une copie de (b,H,K,X).

L'idée est ici de dériver $t \in [0,T] \mapsto u(t,x,\mu_t)$ et d'appliquer ensuite la formule d'Itô classique en temps et en espace pour le processus Y.

3.2.3 Caractère bien posé de l'EDS de McKean-Vlasov et convergence des itérées de Picard

Avant de s'intéresser à la propagation du chaos, il est nécessaire de vérifier que l'EDS de McKean-Vlasov (3.16) est bien posée. Comme expliqué au début du Chapitre 3, on aura besoin dans la suite d'étudier la densité de transition associée et donc d'initialiser le problème en un temps $s \in [0,T]$ fixé. On considère ainsi

$$\begin{cases} dX_t^{s,\xi} = b(t, X_t^{s,\xi}, [X_t^{s,\xi}])\,dt + dZ_t, \quad t \in [s,T], \\ X_s^{s,\xi} = \xi, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d), \end{cases} \qquad (3.27)$$

où ξ est indépendante de Z.

Dans cette sous-section seulement, on suppose que la dérive $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ vérifie l'hypothèse suivante au lieu de (H2), qui est plus faible.

Hypothèse (H3).

(1) La fonction b est mesurable et bornée sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$.

(2) Pour tout $(t,\mu) \in [0,T] \times \mathcal{P}(\mathbb{R}^d)$, l'application $b(t,\cdot,\mu)$ est η-höldérienne sur $\mathbb{R}^d$ uniformément par rapport à $(t,\mu) \in [0,T] \times \mathcal{P}(\mathbb{R}^d)$, avec $\eta \in (0,1]$, i.e. il existe $C>0$ tel que pour tout $t \in [0,T]$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ et $x_1, x_2 \in \mathbb{R}^d$
$$|b(t,x_1,\mu) - b(t,x_2,\mu)| \leq C|x_1 - x_2|^\eta.$$

(3) Pour tout $(t,x) \in [0,T] \times \mathbb{R}^d$, l'application $b(t,x,\cdot)$ est lipschitzienne par rapport à la distance en variation totale d_{TV}, uniformément par rapport à $(t,x) \in [0,T] \times \mathbb{R}^d$, i.e. il existe $C>0$ tel que pour tout $t \in [0,T]$, $x \in \mathbb{R}^d$ et $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$
$$|b(t,x,\mu_1) - b(t,x,\mu_2)| \leq C d_{TV}(\mu_1,\mu_2).$$

On s'intéresse ici au caractère bien posé au sens faible de (3.27), c'est-à-dire au niveau des lois, ce qui se fait à travers l'étude problème de martingales non-linéaire associé à (3.27), dont on rappelle la définition.

Définition 3.13. Pour $(s,\mu) \in [0,T) \times \mathcal{P}(\mathbb{R}^d)$ fixés, on dit qu'une mesure de probabilité $\mathbb{P}$ sur l'espace de Skorokhod des fonctions càdlàg $\mathcal{D}([s,T];\mathbb{R}^d)$, muni de sa filtration canonique $(\mathcal{F}_t)_{t\in[s,T]}$, avec lois marginales $(\mathbb{P}_t)_{t\in[s,T]} \in \mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ est solution du problème de martingales non-linéaire associé à (3.27) avec condition initiale μ au temps s si le processus canonique $(y_t)_{t\in[s,T]}$ vérifie les deux conditions suivantes.

(1) On a $\mathbb{P}_s = \mu$.

(2) Pour tout $\phi \in \mathcal{C}_b^{1,2}([s,T] \times \mathbb{R}^d)$, le processus défini, pour $t \in [s,T]$, par
$$\phi(t,y_t) - \phi(s,y_s) - \int_s^t \left(\partial_r + L_r^{\mathbb{P}}\right) \phi(r,y_r)\, dr$$
est une $(\mathcal{D}([s,T];\mathbb{R}^d), (\mathcal{F}_t)_t, \mathbb{P})$-martingale issue de 0 au temps $t=s$ et où
$$L_r^{\mathbb{P}} f(t,x) := b(r,x,\mathbb{P}_r) \cdot \partial_x f(t,x) + \Delta^{\frac{\alpha}{2}} f(t,\cdot)(x), \tag{3.28}$$
avec $\Delta^{\frac{\alpha}{2}}$ le générateur de Z défini par (1.3) (Laplacien fractionnaire).

On peut maintenant énoncer le résultat d'existence et d'unicité faible pour l'EDS (3.27) qui est prouvé dans [Cav22b].

Théorème 3.14 (Caractère bien posé)**.** *Sous l'Hypothèse (H3), le problème de martingales associé à (3.27) est bien posé pour toute donnée initiale $\mu \in \mathcal{P}(\mathbb{R}^d)$. En particulier, l'EDS (3.27) admet une unique solution faible. De plus, pour tout $P \in \mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ avec $P_s = \mu \in \mathcal{P}(\mathbb{R}^d)$, on peut définir par récurrence $\overline{X}^{(m)}$ comme l'unique solution faible de*
$$\begin{cases} d\overline{X}_t^{(m)} = b(t, \overline{X}_t^m, [\overline{X}_t^{(m-1)}])\, dt + dZ_t, & t \in [s,T], \\ \overline{X}_s^{(m)} = \xi, \end{cases} \tag{3.29}$$
avec $[\xi] = \mu$ et $([\overline{X}_t^{(0)}])_{t\in[s,T]} = P$. Alors, en notant P^* l'unique solution faible de (3.27) avec donnée initiale $\mu \in \mathcal{P}(\mathbb{R}^d)$ au temps s, on a
$$\sup_{t\in[s,T]} d_{TV}(P_t^*, [\overline{X}_t^{(m)}]) \underset{m\to+\infty}{\longrightarrow} 0. \tag{3.30}$$

Idée de preuve du Théorème 3.14. On considère l'espace $\mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ qui est complet pour la distance $d_{s,T}$ définie pour $P,Q \in \mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ par

$$d_{s,T}(P,Q) := \sup_{r\in[s,T]} d_{TV}(P_r,Q_r).$$

On introduit alors l'espace

$$\mathcal{A}_{s,T,\mu} := \left\{P \in \mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d)),\ P_s = \mu\right\},$$

qui est également complet pour $d_{s,T}$. Pour tout $P \in \mathcal{A}_{s,T,\mu}$, on considère l'unique solution faible de l'EDS linéaire suivante où le flot de lois est fixé égal à P

$$\begin{cases} d\overline{X}_t^{s,\xi,P} = b(t,\overline{X}_t^{s,\xi,P},P_t)\,dt + dZ_t, \quad t\in[s,T],\\ \overline{X}_s^{s,\xi,P} = \xi. \end{cases}$$

Cette EDS linéaire est bien posée d'après [MP14]. On peut ainsi définir une application $\mathcal{I} : \mathcal{A}_{s,T,\mu} \to \mathcal{A}_{s,T,\mu}$ telle que pour tout $P \in \mathcal{A}_{s,T,\mu}$ et $t\in[s,T]$, $\mathcal{I}(P)_t = [X_t^{s,\xi,P}]$. On remarque alors que les solutions du problème de martingales correspondent exactement aux points fixes de l'application $\mathcal{I}$. Il reste à montrer que $\mathcal{I}$ admet une itérée contractante sur $(\mathcal{A}_{s,T,\mu}, d_{s,T})$ pour appliquer le théorème du point fixe de Banach. On utilise pour cela la méthode parametrix, justifiée par les points (1) et (2) de l'Hypothèse (H3). Elle permet d'exprimer la densité de $\overline{X}_t^{s,\xi,P}$ en fonction de P notamment. Donnons-en l'idée principale de manière heuristique. On considère l'EDS linéaire suivante

$$\begin{cases} dX_t^{s,x} = b(t,X_t^{s,x})\,dt + dZ_t, \quad t\in[s,T],\\ X_s^{s,x} = x, \end{cases}$$

où b est mesurable borné. Sa densité de transition $p(s,t,x,\cdot)$ est solution (fondamentale) de l'EDP de Kolmogorov rétrograde

$$\begin{cases} \partial_s p(s,t,x,y) + L_s p(s,t,\cdot,y)(x) = 0, \quad \forall (s,x)\in[0,t)\times\mathbb{R}^d,\\ p(s,t,x,\cdot) \underset{s\to t^-}{\longrightarrow} \delta_x, \quad \text{au sens faible}, \end{cases}$$

où pour tout $(t,x)\in[0,T]\times\mathbb{R}^d$, $L_s f(t,x) := b(s,x)\cdot\partial_x f(t,x) + \Delta^{\frac{\alpha}{2}} f(t,\cdot)(x)$. On réécrit alors le problème comme

$$\begin{cases} \partial_s p(s,t,x,y) + \Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x) = (\Delta^{\frac{\alpha}{2}} - L_s)p(s,t,\cdot,y)(x), \quad \forall (s,x)\in[0,t)\times\mathbb{R}^d,\\ p(s,t,x,\cdot) \underset{s\to t^-}{\longrightarrow} \delta_x, \quad \text{au sens faible}. \end{cases}$$

Or la solution fondamentale associée à l'opérateur $\partial_s + \Delta^{\frac{\alpha}{2}}$ est donnée par la densité de transition de Z notée $\widehat{p}(s,t,x,\cdot) := q(t-s,y-x)$, où $q(t,\cdot)$ est la densité de Z_t donnée par la Proposition 1.14. Cela assure qu'on a la représentation suivante de la densité p

$$p(s,t,x,y) = \widehat{p}(s,t,x,y) + \widehat{p}\otimes\left[(L_s - \Delta^{\frac{\alpha}{2}})p\right](s,t,x,y),$$

où
$$f \otimes g(s,t,x,y) := \int_s^t \int_{\mathbb{R}^d} f(s,r,x,z) g(r,t,z,y)\, dz\, dr,$$
lorsque l'intégrale est bien définie. En itérant, on obtient alors

$$\begin{aligned} p(s,t,x,y) &= \widehat{p}(s,t,x,y) + \sum_{k=1}^{\infty} \widehat{p} \otimes \mathcal{H}^k(s,t,x,y) \\ &= \widehat{p}(s,t,x,y) + p \otimes \mathcal{H}(s,t,x,y), \quad \text{(représentation implicite)} \end{aligned} \tag{3.31}$$

où

$$\mathcal{H}(s,t,x,y) := b(s,x) \cdot \partial_x \widehat{p}(s,t,x,y)$$

est le noyau parametrix et où $\widehat{p}$ est appelé proxy (densité de référence). Les estimations de gradient sur $\widehat{p}$ de la Proposition 1.14 permettent de montrer que la série dans (3.31) converge, b étant borné. C'est pour avoir ces estimations et les développements en série parametrix qui en découlent qu'on se restreint au cas d'un bruit α-stable rotationnellement invariant. On pourra consulter [KK18] et [MZ22] pour plus de détails sur la méthode parametrix dans ce cadre-là, ou encore la Section 8.10 de ce manuscrit.

Grâce à ces développements, on peut contrôler la différence des densités associées à deux flots de mesures $P, Q \in \mathcal{A}_{s,T,\mu}$. Ainsi, en utilisant le point (3) de l'Hypothèse (H3), on contrôle la distance $d_{s,T}(\mathcal{I}(P), \mathcal{I}(Q))$ en fonction de $d_{s,T}(P,Q)$, ce qui permet d'appliquer le théorème du point fixe de Banach.

□

Remarque 3.15. Le caractère bien posé était déjà démontré dans [FKM21] avec un coefficient de diffusion non constant devant le bruit Z. Dans ce travail, le théorème du point fixe utilisé dans la preuve est seulement appliqué en temps petit, ce qui suffit pour établir le caractère bien posé par recollement des solutions. Cependant, on a besoin dans la suite d'avoir la convergence des itérées de Picard sur tout l'intervalle de temps $[s,T]$. C'est pour cela qu'on refait la preuve de ce résultat dans la version plus simple ici où $\sigma = \mathrm{Id}$. On utilise le développement en série parametrix (3.31) complet contrairement à la représentation implicite utilisée dans [FKM21], ce qui permet d'obtenir ce résultat global.

3.2.4 Régularité de la densité de transition et problème de Cauchy associé

Une fois le caractère bien posé établi, on s'intéresse dans [Cav22b] à la régularité de la densité de transition associée à l'EDS de McKean-Vlasov (3.27). Comme on l'a mentionné précédemment, il s'agit d'une étape cruciale pour étudier ensuite les propriétés régularisantes du semi-groupe dont on a besoin pour prouver la propagation du chaos avec la méthode présentée dans la Sous-section 1.2.3. On introduit donc, pour $x \in \mathbb{R}^d$, le flot découplé associé à (3.27) solution de

$$\begin{cases} dX_t^{s,x,\mu} = b(t, X_t^{s,x,\mu}, [X_t^{s,\mu}])\, dt + dZ_t, & t \in [s,T], \\ X_s^{s,x,\mu} = x \in \mathbb{R}^d, \end{cases} \tag{3.32}$$

où on note $[X_t^{s,\mu}] := [X_t^{s,\xi}]$ si ξ est de loi μ. La notation fait sens puisque $[X_t^{s,\xi}]$ ne dépend de ξ qu'à travers sa loi d'après le caractère bien posé du problème de martingales. La loi de $X_t^{s,x,\mu}$ admet une densité notée

$p(\mu, s, t, x, \cdot)$, de même que la loi de $X_t^{s,\mu}$ notée $p(\mu, s, t, \cdot)$. De plus, ces deux densités vérifient la relation suivante qui provient du caractère bien posé du problème de martingales du Théorème 3.14

$$p(\mu, s, t, y) = \int_{\mathbb{R}^d} p(\mu, s, t, x, y)\, d\mu(x). \tag{3.33}$$

On étudie alors la régularité de la densité découplée $p(\mu, s, t, x, y)$ par rapport aux données initiales $(s, x, \mu) \in [0, t) \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, avec $\beta \in (1, \alpha)$, et on exhibe l'EDP de Kolmogorov rétrograde dont elle est solution.

Théorème 3.16. *Soient $(t, y) \in (0, T] \times \mathbb{R}^d$ fixés. Sous l'Hypothèse (H2), l'application $(\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d \mapsto p(\mu, s, t, x, y)$ appartient à $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d)$ (voir la Définition (2.19)), et est solution de l'EDP de Kolmogorov rétrograde suivante*

$$\begin{cases} \partial_s p(\mu, s, t, x, y) + \mathcal{L}_s p(\cdot, s, t, \cdot, y)(\mu, x) = 0, & \forall (\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d, \\ p(\mu, s, t, x, \cdot) \underset{s \to t-}{\longrightarrow} \delta_x, & \textit{au sens faible,} \end{cases} \tag{3.34}$$

où $\mathcal{L}_s$ est défini, pour toute fonction suffisamment régulière h sur $\mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, par

$$\begin{aligned} \mathcal{L}_s h(\mu, x) &:= b(s, x, \mu) \cdot \partial_x h(\mu, x) + \int_{\mathbb{R}^d} [h(\mu, x+z) - h(\mu, x) - z \cdot \partial_x h(\mu, x)] \frac{dz}{|z|^{d+\alpha}} \\ &\quad + \int_{\mathbb{R}^d} b(s, v, \mu) \cdot \partial_v \frac{\delta}{\delta m} h(\mu, x)(v)\, d\mu(v) \\ &\quad + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} h(\mu, x)(v+z) - \frac{\delta}{\delta m} h(\mu, x)(v) - z \cdot \partial_v \frac{\delta}{\delta m} h(\mu, x)(v) \right] \frac{dz}{|z|^{d+\alpha}}\, d\mu(v). \end{aligned} \tag{3.35}$$

De plus, il existe une constante positive C telle que pour tous $j \in \{0, 1\}$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$ et $x, y, v \in \mathbb{R}^d$

$$|\partial_x^j p(\mu, s, t, x, y)| \le C(t-s)^{-\frac{j}{\alpha}} \rho^j(t-s, y-x), \tag{3.36}$$

$$|\Delta^{\frac{\alpha}{2}} p(\mu, s, t, \cdot, y)(x)| \le C(t-s)^{-1} \rho^0(t-s, y-x), \tag{3.37}$$

$$\left| \partial_v^j \frac{\delta}{\delta m} p(\mu, s, t, x, y)(v) \right| \le C(t-s)^{j\frac{\eta-1}{\alpha} + 1 - \frac{1}{\alpha}} \rho^0(t-s, y-x), \tag{3.38}$$

et

$$\left| \Delta^{\frac{\alpha}{2}} \left[\frac{\delta}{\delta m} p(\mu, s, t, x, y) \right](v) \right| \le C(t-s)^{-\frac{1}{\alpha}} \rho^0(t-s, y-x), \tag{3.39}$$

où ρ^j est définie pour $(t, x) \in (0, +\infty) \times \mathbb{R}^d$ par

$$\rho^j(t, x) := t^{-\frac{d}{\alpha}} (1 + t^{-\frac{1}{\alpha}} |x|)^{-d-\alpha-j}. \tag{3.40}$$

Remarquons que l'estimation de gradient (3.36) est la même que celle de la densité de transition $\widehat{p}$ du processus α-stable rotationnellement invariant Z (voir la Proposition 1.14). D'autres estimations, typiquement des contrôles höldériens, sont prouvées dans le Théorème 8.18. Ils servent non seulement à prouver le Théorème 3.16, mais également dans la suite pour établir des contrôles höldériens sur le gradient de la dérivée plate du semi-groupe, dont on expliquera l'importance dans la suite.

Remarque 3.17. Introduisons, pour $(s,\mu) \in [0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$, l'opérateur L_s^μ associé à l'EDS linéaire obtenue en gelant l'argument mesure des coefficients de l'EDS de McKean-Vlasov (3.27) égal à μ. Plus précisément, il est défini, pour toute fonction f sur $\mathbb{R}^d$ suffisamment régulière, par

$$\forall x \in \mathbb{R}^d, \quad L_s^\mu f(x) := b(s,x,\mu) \cdot \partial_x f(x) + \Delta^{\frac{\alpha}{2}} f(x).$$

Alors, pour toute fonction h sur $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ suffisamment régulière, on a

$$\mathcal{L}_s h(\mu,x) = L_s^\mu h(\mu,\cdot)(x) + \int_{\mathbb{R}^d} L_s^\mu \left[\frac{\delta}{\delta m} h(\mu,x)\right](v)\, d\mu(v).$$

Idée de preuve du Théorème 3.16. Comme la dépendance par rapport à μ de $[X_t^{s,\mu}]$ est abstraite, on considère les itérées de Picard définies récursivement, pour $m \geq 0$, par

$$\begin{cases} dX_t^{s,\xi,(m+1)} = b(t, X_t^{s,\xi,(m+1)}, [X_t^{s,\mu,(m)}])\, dt + dZ_t, \quad t \in [s,T], \\ X_s^{s,\xi,(m+1)} = \xi, \quad [\xi] = \mu \in \mathcal{P}_\beta(\mathbb{R}^d), \end{cases} \tag{3.41}$$

où $[X_t^{s,\mu,(m)}] := [X_t^{s,\xi,(m)}]$ et où pour tout $t \in [s,T]$, $[X_t^{s,\mu,(0)}] = \rho \in \mathcal{P}_\beta(\mathbb{R}^d)$ est une mesure de probabilité quelconque fixée. On considère également la version découplée

$$\begin{cases} dX_t^{s,x,\mu,(m+1)} = b(t, X_t^{s,x,\mu,(m+1)}, [X_t^{s,\mu,(m)}])\, dt + dZ_t, \quad t \in [s,T], \\ X_s^{s,x,\mu,(m+1)} = x \in \mathbb{R}^d. \end{cases} \tag{3.42}$$

Les lois $[X_t^{s,\mu,(m)}]$ et $[X_t^{s,x,\mu,(m)}]$ ont des densités $p_m(\mu,s,t,\cdot)$ et $p_m(\mu,s,t,x,\cdot)$ qui vérifient la relation (3.33). On montre que ces densités convergent ponctuellement, lorsque m tend vers l'infini, vers les densités limites $p(\mu,s,t,\cdot)$ et $p(\mu,s,t,x,\cdot)$ grâce aux développements en série parametrix et au Théorème 3.14. On établit également que les applications $(s,x,\mu) \in [0,t) \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto p_m(\mu,s,t,x,y)$ sont $\mathcal{C}^1$ pour tout $t \in (0,T]$ et $y \in \mathbb{R}^d$ par récurrence et en utilisant des développements en série parametrix des densités. On prouve également des bornes et des contrôles höldériens uniformes en m sur les dérivées qu'on veut contrôler. Une fois ces estimations démontrées, on conclut en utilisant le théorème d'Ascoli que les densités limites ont la régularité $\mathcal{C}^1$ voulue. De plus, en passant à la limite dans les contrôles précédemment montrés sur les densités p_m, qui sont uniformes en m, on déduit les inégalités correspondantes pour la densité limite p. Pour établir que la densité découplée est solution de l'EDP (3.34), on fixe $h \in [0,s]$. En utilisant le caractère bien posé du problème de martingales prouvé au Théorème 3.14 et la propriété de Markov qui en découle, on obtient

$$p(\mu, s-h, t, x, y) = \mathbb{E}p([X_s^{s-h,\mu}], s, t, X_s^{s-h,x,\mu}, y). \tag{3.43}$$

On peut alors appliquer la formule d'Itô du Théorème 3.12 à la fonction $(\mu,x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto p(\mu,s,t,x,y)$. En prenant l'espérance, on obtient

$$\mathbb{E}p([X_s^{s-h,\mu}], s, t, X_s^{s-h,x,\mu}, y) = p(\mu,s,t,x,y) + \int_{s-h}^s \mathcal{L}_r p(\cdot,s,t,\cdot,y)([X_r^{s-h,\mu}], X_r^{s-h,x,\mu})\, dr.$$

Il s'ensuit que

$$\frac{1}{h}(p(\mu, s-h, t, x, y) - p(\mu, s, t, x, y)) = \frac{1}{h}\int_{s-h}^{s} \mathbb{E}\mathcal{L}_r p(\cdot, s, t, \cdot, y)([X_r^{s-h,\mu}], X_r^{s-h,x,\mu})\, dr,$$

et on conclut donc que

$$\frac{1}{h}(p(\mu, s-h, t, x, y) - p(\mu, s, t, x, y)) \underset{h\to 0^+}{\longrightarrow} \mathcal{L}_s p(\cdot, s, t, \cdot, y)(\mu, x).$$

□

Concernant la densité non découplée p, on a le résultat suivant qui se prouve de manière analogue au résultat sur la densité découplée.

Théorème 3.18. *Soient $(t, y) \in (0, T] \times \mathbb{R}^d$ fixés. Sous l'Hypothèse (H2), l'application $(\mu, s) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \mapsto p(\mu, s, t, y)$ appartient à $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t))$ (voir la Définition 2.19), et est solution (dite fondamentale) de l'EDP de Kolmogorov rétrograde suivante*

$$\begin{cases} \partial_s p(\mu, s, t, y) + \mathscr{L}_s p(\cdot, s, t, y)(\mu) = 0, & \forall (\mu, s) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t), \\ p(\mu, s, t, \cdot) \underset{s\to t-}{\longrightarrow} \mu, & \textit{au sens faible,} \end{cases} \tag{3.44}$$

où $\mathscr{L}_s$ est défini, pour toute fonction h définie sur $\mathcal{P}_\beta(\mathbb{R}^d)$ suffisamment régulière, par

$$\begin{aligned} \mathscr{L}_s h(\mu) := &\int_{\mathbb{R}^d} b(s, v, \mu) \cdot \partial_v \frac{\delta}{\delta m} h(\mu)(v)\, d\mu(v) \\ &+ \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} h(\mu)(v+z) - \frac{\delta}{\delta m} h(\mu)(v) - z \cdot \partial_v \frac{\delta}{\delta m} h(\mu)(v)\right] \frac{dz}{|z|^{d+\alpha}}\, d\mu(v). \end{aligned} \tag{3.45}$$

3.2.5 Étude du semi-groupe et EDP de Kolmogorov rétrograde

Grâce aux résultats précédents sur la densité de transition de l'EDS de McKean-Vlasov (3.27), on s'intéresse dans [Cav22b] au semi-groupe associé à l'EDS de McKean-Vlasov (3.27) agissant sur les fonctions définies sur $\mathcal{P}_\beta(\mathbb{R}^d)$, où on rappelle que $\beta \in (1, \alpha)$ est fixé. Pour une fonction $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ fixée, l'action du semi-groupe sur ϕ est donnée par la fonction U définie par

$$U(t, \mu) := \phi([X_T^{t,\mu}]), \quad \forall (t, \mu) \in [0, T] \times \mathcal{P}_\beta(\mathbb{R}^d), \tag{3.46}$$

où $[X_T^{t,\mu}]$ est la loi de la solution de l'EDS de McKean-Vlasov (3.16) au temps T, où la loi initiale est μ au temps t. Plus précisément, notre but est d'étudier les propriétés régularisantes du semi-groupe, c'est-à-dire le gain de régularité entre ϕ et U par rapport à la variable mesure. Ces propriétés régularisantes proviennent directement des propriétés sur la densité de transition de l'EDS de McKean-Vlasov (3.27) présentées dans la sous-section précédente.

Définissons plus précisément l'espace de fonctions de type Hölder sur lequel on fait agir le semi-groupe.

Définition 3.19. Soit $\delta \in (0, 1]$. On définit l'espace $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ comme l'ensemble des applications continues $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ admettant une dérivée plate telle qu'il existe une constante positive C telle que pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v_1, v_2 \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m}\phi(\mu)(v_1) - \frac{\delta}{\delta m}\phi(\mu)(v_2)\right| \leq C|v_1 - v_2|^\delta.$$

Le théorème suivant précise les propriétés régularisantes du semi-groupe agissant sur $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ et décrit sa dynamique grâce à une EDP de Kolmogorov rétrograde.

Théorème 3.20 (EDP de Kolmogorov rétrograde)**.** *Soit $\phi \in \mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$. Alors, sous l'Hypothèse (H2), l'application U définie par (3.46) appartient à $\mathcal{C}^0([0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)) \cap \mathcal{C}^1([0,T) \times \mathcal{P}_\beta(\mathbb{R}^d))$ (voir la Définition 2.19) et satisfait les propriétés suivantes.*

- *Il existe une constante positive C telle que pour tout $t \in [0,T)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v \in \mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v)\right| \leq C(T-t)^{\frac{\delta-1}{\alpha}}. \tag{3.47}$$

- *Pour tout $\gamma \in (0,1] \cap (0, (2\alpha-2) \wedge (\eta+\alpha-1))$, il existe une constante positive C telle que pour tout $t \in [0,T)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v_1, v_2 \in \mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2)\right| \leq C(T-t)^{\frac{\delta-1-\gamma}{\alpha}}|v_1 - v_2|^\gamma. \tag{3.48}$$

De plus, U est solution de l'EDP de Kolmogorov rétrograde suivante

$$\begin{cases} \partial_t U(t,\mu) + \mathscr{L}_t U(t,\cdot)(\mu) = 0, \quad \forall (t,\mu) \in [0,T) \times \mathcal{P}_\beta(\mathbb{R}^d), \\ U(T,\mu) = \phi(\mu), \quad \forall \mu \in \mathcal{P}_\beta(\mathbb{R}^d), \end{cases} \tag{3.49}$$

où $\mathscr{L}_t$ a été défini par (3.45). Il s'agit de l'unique solution de (3.49) parmi toutes les fonctions appartenant à $\mathcal{C}^0([0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)) \cap \mathcal{C}^1([0,T) \times \mathcal{P}_\beta(\mathbb{R}^d))$ vérifiant (3.47) et (3.48).

Les contrôles (3.47) et (3.48) avec $\gamma > \alpha - 1$ permettent d'appliquer la formule d'Itô du Théorème 3.9 le long du flot de lois marginales associé à l'EDS de McKean-Vlasov (3.16). La condition $\gamma > \alpha - 1$ permet de vérifier l'Hypothèse **(J1)** requise dans le Théorème 3.9 puisque pour un tel γ, on a

$$\int_{B_1} |z|^{1+\gamma}\, d\nu(z) < +\infty,$$

où ν est la mesure de Lévy de Z. Cela sert notamment à prouver (3.49), mais également à prouver la propagation du chaos faible quantitative. En effet, comme présenté dans la Sous-section 1.2.3, on aura besoin d'appliquer la formule d'Itô classique à la projection empirique de U et au système de particules (1.6). D'après le Lemme 2.27, la projection empirique de U est seulement de classe $\mathcal{C}^1$ sur $[0,T) \times (\mathbb{R}^d)^N$. Ainsi, l'utilisation de la formule d'Itô classique pour la projection empirique de U et le système de particules est justifiée par le contrôle höldérien (3.48), avec $\gamma > \alpha - 1$. La justification est complètement analogue au fait que l'Hypothèse **(J1)**, vérifiée ici pour $\gamma > \alpha - 1$, permet d'assurer qu'on peut appliquer la formule d'Itô du Théorème 3.9, le long du flot de lois marginales de la solution de l'EDS de McKean-Vlasov (3.16).

Remarque 3.21. Dans [Cav22a], on décrit également la dynamique du semi-groupe pour l'EDS de McKean-Vlasov (3.16) avec un coefficient de diffusion σ. On travaille sous l'hypothèse que l'EDS est bien posée au sens faible et sous des hypothèses de régularité sur la fonction U pour pouvoir appliquer la formule d'Itô du Théorème 3.9. L'étude de la densité de transition faite dans [Cav22b] permet justement de prouver la régularité requise sur U qui est énoncée dans le théorème précédent.

Idée de preuve du Théorème 3.20. Pour la régularité par rapport à $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ et les inégalités (3.47) et (3.48), cela découle des contrôles sur la densité de transition mentionnés dans la sous-section précédente et prouvés dans [Cav22b] (voir le Théorème 8.18). Donnons l'argument qui permet d'établir que U est solution de l'EDP (3.49). Soient $t \in [0,T)$ et $h \in [0,t]$. D'après la propriété de Markov provenant du caractère bien posé du problème de martingales associé à (3.27), on déduit que

$$U(t-h,\mu) = \phi([X_T^{t-h,\mu}]) = \phi([X_T^{t,[X_t^{t-h,\mu}]}]) = U(t,[X_t^{t-h,\mu}]).$$

En appliquant la formule d'Itô pour un flot de mesures du Théorème 3.9 à la fonction $U(t,\cdot)$, on obtient

$$\begin{aligned}
U(t-h,\mu) &= U(t,\mu) + \int_{t-h}^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \cdot b(s, X_s^{t-h,\mu}, [X_s^{t-h,\mu}])\right) ds \\
&\quad + \int_{t-h}^t \mathbb{E}\int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu} + z) - \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \right.\\
&\qquad \left. - \partial_v \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \cdot z\right) \frac{dz}{|z|^{d+\alpha}}\, ds \\
&= U(t,\mu) + \int_{t-h}^t \mathscr{L}_s U(t,\cdot)([X_s^{t-h,\mu}])\, ds.
\end{aligned}$$

Par continuité, on montre que

$$\frac{1}{h}(U(t-h,\mu) - U(t,\mu)) \underset{h\to 0^+}{\longrightarrow} \mathscr{L}_t U(t,\cdot)(\mu),$$

et donc que pour tout $t \in [0,T)$ et $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$

$$\partial_t U(t,\mu) + \mathscr{L}_t U(t,\cdot)(\mu) = 0.$$

L'unicité découle également de la formule d'Itô pour un flot de mesures.

$\square$

3.2.6 Propagation du chaos faible quantitative

On présente maintenant les résultats de propagation du chaos faible quantitative (voir la Définition 1.18) obtenus pour le système de particules en interaction (3.17) associé à (3.16). Rappelons les notations. Soit $(Z^n)_{n\geq 1}$ une suite i.i.d. de processus α-stables de même loi que Z et $(\xi^n)_n$ une suite i.i.d. de variables aléatoires de loi commune $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, où $\beta \in (1,\alpha)$ est toujours fixé. Le système de particules est défini par

$$\begin{cases}
dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\, dt + dZ_t^i, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\
\overline{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}, \\
X_0^{i,N} = \xi^i.
\end{cases} \tag{3.50}$$

On écrit ce système sous forme vectorielle. On pose, pour $t \in [0,T]$ et $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$

$$\mathbf{Z}_t^N := \begin{pmatrix} Z_t^1 \\ \vdots \\ Z_t^N \end{pmatrix}, \quad \boldsymbol{b}^N(t, \boldsymbol{x}) := \begin{pmatrix} b(t, x_1, \overline{\mu}_{\boldsymbol{x}}^N) \\ \vdots \\ b(t, x_N, \overline{\mu}_{\boldsymbol{x}}^N) \end{pmatrix}, \quad \text{et} \quad \boldsymbol{X}_t^N := \begin{pmatrix} X_t^{1,N} \\ \vdots \\ X_t^{N,N} \end{pmatrix}. \tag{3.51}$$

Le système de particules (3.50) peut alors être réécrit comme

$$\begin{cases} d\boldsymbol{X}_t^N = \boldsymbol{b}^N(t, \boldsymbol{X}_t^N)\, dt + d\mathbf{Z}_t^N, \quad t \in [0,T], \\ \boldsymbol{X}_0^N = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^N \end{pmatrix}. \end{cases} \tag{3.52}$$

Comme les processus $(Z^n)_{n\geq 1}$ sont indépendants, le processus $(\mathbf{Z}_t^N)_{t\geq 0}$ définit un processus α-stable sur $(\mathbb{R}^d)^N$, qu'on qualifie de cylindrique. L'EDS linéaire (3.52) est bien posée au sens faible en utilisant [CSZ18, Corollaire 1.4 (iii)] puisque sa dérive $\boldsymbol{b}^N$ est höldérienne en espace, uniformément en temps. Cela se déduit de l'Hypothèse (H2) en linéarisant avec la dérivée plate pour étudier la partie projection empirique. Le processus limite est solution de l'EDS de McKean-Vlasov (3.16) initialisée au temps $s = 0$ par une variable aléatoire ξ de loi $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$. On note $(\mu_t)_{t\in[0,T]}$ le flot de lois marginales de la solution. Définissons maintenant l'espace de fonctions tests utilisé pour quantifier la propagation du chaos faible.

Définition 3.22. Pour $\delta \in (0,1]$ et $L > 0$, on définit l'espace $\mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ comme l'ensemble des fonctions continues $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ admettant des dérivées plates d'ordre 1 et 2 $\frac{\delta}{\delta m}\phi$ et $\frac{\delta^2}{\delta m^2}\phi$ telles que pour tout $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v_1, v_2, v_1', v_2' \in \mathbb{R}^d$

$$\left| \frac{\delta}{\delta m}\phi(\mu)(v_1) - \frac{\delta}{\delta m}\phi(\mu)(v_2) \right| \leq L|v_1 - v_2|^\delta,$$

et

$$\left| \frac{\delta^2}{\delta m^2}\phi(\mu)(v_1, v_1') - \frac{\delta^2}{\delta m^2}\phi(\mu)(v_2, v_2') \right| \leq L(|v_1 - v_2|^\delta + |v_1' - v_2'|^\delta).$$

Remarquons que $\mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ est contenu dans $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$, défini à la Définition 3.19. Ainsi, le Théorème 3.20 qui décrit la régularité et la dynamique du semi-groupe associé à (3.16) s'applique pour $\phi \in \mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$. On peut maintenant énoncer le résultat de propagation du chaos faible quantitative prouvé dans [Cav22b].

Théorème 3.23 (Propagation du chaos faible quantitative)**.**
Soit $\delta \in (0,1]$, $L > 0$ et $\gamma \in (0,1] \cap (0, (\delta + \alpha - 1) \wedge (2\alpha - 2) \wedge (\eta + \alpha - 1))$ fixés. Alors, sous l'Hypothèse (H2), il existe une constante positive $C = C(d, T, \alpha, \beta, (H2), \gamma, \delta, L)$, croissante par rapport à T, telle que pour tout $\phi \in \mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ et $N \geq 1$

$$\mathbb{E}|\phi(\overline{\mu}_T^N) - \phi(\mu_T)| \leq C\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + \frac{C}{N^{1-\frac{1}{\beta}}}, \tag{3.53}$$

et

$$|\mathbb{E}(\phi(\overline{\mu}_T^N) - \phi(\mu_T))| \leq \frac{C}{N^\gamma}. \tag{3.54}$$

Remarque 3.24. Tout d'abord, mentionnons que le terme de donnée initiale $\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0)$ peut être traité en utilisant [FG15], en particulier dans le cas où μ_0 admet un moment fini d'ordre supérieur à β. En effet, si $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$ avec $q \geq 1$, on a

$$\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0) \leq C \begin{cases} N^{-\frac{1}{2}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{si} \quad d=1 \quad \text{et} \quad q \neq 2, \\ N^{-\frac{1}{2}}\ln(1+N) + N^{-\left(1-\frac{1}{q}\right)}, & \text{si} \quad d=2 \quad \text{et} \quad q \neq 2, \\ N^{-\frac{1}{d}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{si} \quad d\geq 3 \quad \text{et} \quad q \neq \frac{d}{d-1}. \end{cases} \tag{3.55}$$

Remarque 3.25. Précisons l'intérêt de l'estimation (3.54). Il s'agit d'un résultat qui quantifie l'approximation du semi-groupe associé à l'EDS de McKean-Vlasov (3.16) par sa projection empirique évaluée en le système de particules (3.17). De plus, grâce à (3.54), on peut quantifier la vitesse d'approximation de la loi d'une particule par la loi du processus limite de McKean-Vlasov par rapport à la distance de Wasserstein W_1. En effet, notons $\|\varphi\|_{\text{Lip}} := \sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{|x-y|}$ pour $\varphi : \mathbb{R}^d \to \mathbb{R}$. L'ensemble

$$\mathscr{C} := \left\{\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R},\ \exists \varphi : \mathbb{R}^d \to \mathbb{R}, \text{ avec } \|\varphi\|_{\text{Lip}} \leq 1, \text{ et } \phi(\mu) = \int_{\mathbb{R}^d} \varphi\, d\mu,\ \forall \mu \in \mathcal{P}_\beta(\mathbb{R}^d)\right\}$$

est contenu dans $\mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$. Ainsi grâce à (3.54) et au théorème de Kantorovich-Rubinstein [Vil09, Remarque 6.5], on déduit

$$\begin{aligned} \sup_{t\in[0,T]} W_1([X_t^{1,N}],\mu_t) &= \sup_{t\in[0,T]} \sup_{\varphi,\, \|\varphi\|_{\text{Lip}}\leq 1} \left|\mathbb{E}\varphi(X_t^{1,N}) - \int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\ &= \sup_{t\in[0,T]} \sup_{\varphi,\, \|\varphi\|_{\text{Lip}}\leq 1} \left|\mathbb{E}\left(\frac{1}{N}\sum_{k=1}^N \varphi(X_t^{k,N})\right) - \int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\ &= \sup_{t\in[0,T]} \sup_{\varphi,\, \|\varphi\|_{\text{Lip}}\leq 1} \left|\mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\overline{\mu}_t^N - \mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\ &= \sup_{t\in[0,T]} \sup_{\phi\in\mathscr{C}} \left|\mathbb{E}\phi(\overline{\mu}_t^N) - \mathbb{E}\phi(\mu_t)\right| \\ &\leq \frac{C_T}{N^\gamma}, \end{aligned} \tag{3.56}$$

où $\gamma \in (0,1] \cap (0, (\delta+\alpha-1) \wedge (2\alpha-2) \wedge (\eta+\alpha-1))$, et en utilisant le fait que la constante C du Théorème 3.23 est croissante par rapport à T.

Prenons formellement $\alpha = 2$ dans le théorème précédent, qui correspond au cas brownien traité dans [CdRF21]. Alors, on peut prendre $\gamma = 1$ et $\beta = 2$ dans le Théorème 3.23. Les taux de convergence prouvés dans le Théorème 3.23 sont les mêmes que ceux prouvés dans [CdRF21, Théorème 3.6], i.e. $N^{-\frac{1}{2}}$ pour (3.53) et N^{-1} pour (3.54). De plus, en dimension $d = 1$, on retrouve avec (3.53) le même taux de convergence que celui obtenu dans [FL21] pour la propagation du chaos forte dans L^1, puisque $\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0) \leq CN^{\frac{1}{\beta}-1}$ par [FG15].

Dans [Cav22a], on étudie également la propagation du chaos faible quantitative, mais cette fois pour un système particulier de processus d'Ornstein-Uhlenbeck stables en interaction de type champ moyen qu'on présente dans ce qui suit. Le bruit $Z = (Z_t)_{t\geq 0}$ est un processus α-stable, avec $\alpha \in (1,2)$, nondégénéré au sens où sa mesure de Lévy

$$\nu(dy) = d\lambda(\theta)\frac{dr}{r^{1+\alpha}}$$

vérifie qu'il existe $\rho > 0$ tel que pour tout $x \in \mathbb{R}^d$

$$\rho|x|^2 \leq \int_{\mathbb{S}^{d-1}} |\langle x, \theta\rangle|^2 \, d\lambda(\theta). \tag{3.57}$$

On fixe $A, A', B \in \mathbb{R}^{d\times d}$ des matrices de taille $d \times d$ telles que B est inversible et A et A' commutent. Soit $\beta \in [1, \alpha)$. En gardant les mêmes notations que pour l'EDS de McKean-Vlasov (3.16) et le système de particules (3.50), notre but est d'établir des inégalités de propagation du chaos faible pour le système suivant

$$\begin{cases} dX_t^{i,N} = AX_t^{i,N}\, dt + A'\dfrac{1}{N}\displaystyle\sum_{j=1}^N X_t^{j,N}\, dt + B\, dZ_t^i, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{3.58}$$

vers l'EDS de McKean-Vlasov limite

$$\begin{cases} dX_t &= (AX_t + A'\mathbb{E}X_t)\, dt + B\, dZ_t, \quad t \in [0,T], \\ X_0 &= \xi. \end{cases} \tag{3.59}$$

Remarquons que le résultat du Théorème 3.23 ne s'applique pas pour ce système puisque le coefficient de dérive, défini pour $\mu \in \mathcal{P}_1(\mathbb{R}^d)$ et $x \in \mathbb{R}^d$, par

$$b(x,\mu) := Ax + A'\int_{\mathbb{R}^d} y\, d\mu(y),$$

n'est pas borné par rapport aux variables espace et mesure. On prouve néanmoins dans [Cav22a] le résultat suivant.

Théorème 3.26. *Il existe une constante positive $C = C_T$, indépendante de $\phi \in \mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$ et de $N \geq 1$ et croissante par rapport à T, telle que pour tout $\phi \in \mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$ et $N \geq 1$, on a*

$$\mathbb{E}\left|\phi(\overline{\mu}_T^N) - \phi(\mu_T)\right| \leq C\,\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}, \tag{3.60}$$

et

$$|\mathbb{E}(\phi(\overline{\mu}_T^N) - \phi(\mu_T))| \leq C\,\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + \frac{C}{N^{\alpha-1}}. \tag{3.61}$$

De plus, si μ_0 appartient à $\mathcal{P}_2(\mathbb{R}^d)$, on a pour tout $\phi \in \mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$ et $N \geq 1$

$$|\mathbb{E}(\phi(\overline{\mu}_T^N) - \phi(\mu_T))| \leq \frac{C}{N^{\alpha-1}}. \tag{3.62}$$

Avant de présenter les preuves des Théorèmes 3.23 et 3.26, comparons les vitesses de convergence qui y sont obtenues. L'inégalité (3.54) du Théorème 3.23 est meilleure que l'inégalité (3.62) du Théorème 3.26, ce qui est naturel puisque les hypothèses sont plus fortes dans le Théorème 3.23. En ce qui concerne les inégalités (3.60) et (3.53) de ces deux théorèmes, on peut remarquer qu'on a une meilleure vitesse dans le cas du système de processus d'Ornstein-Uhlenbeck malgré que le coefficient de dérive est non-borné par rapport aux variables espace et mesure dans ce cas-là. Cela ne semble pas naturel puisque les hypothèses sont plus fortes dans le Théorème 3.23. L'explication est que le système de processus d'Ornstein-Uhlenbeck en interaction est un système particulier dont les coefficients sont linéaires. On utilise la même idée que celle présentée dans la preuve du Théorème 3.8 qui établit la propagation du chaos forte qualitative pour ce système. Il s'agit de considérer le système de particules et l'EDS de McKean-Vlasov limite pour lesquels on a retiré des bruits les sauts de taille plus grande que le nombre de particules. Tout comme dans le Théorème 3.8, l'explication du facteur logarithmique dans (3.60) provient du fait que

$$\int_{1\leq|z|\leq N} |z|^\alpha \, d\nu(z) \underset{N\to+\infty}{\sim} \ln(N).$$

Comparons maintenant les vitesses de convergence obtenues dans le Théorème 3.26 à celles prouvées au niveau des trajectoires par la méthode de couplage dans le Théorème 3.8. On peut tout d'abord mentionner que dans le Théorème 3.8, on n'a pas besoin de faire l'hypothèse (3.57) de non-dégénérescence sur ν, ni de supposer que la matrice B est inversible. Plaçons-nous dans le cas où la loi initiale μ_0 admet un moment d'ordre 2 fini. L'inégalité (3.60) quantifie la propagation du chaos à travers la mesure empirique (voir (1.9)). Moralement, elle permet de quantifier $\mathbb{E}W_1(\overline{\mu}_T^N, \mu_T)$. On dit "moralement" car pour que ce soit le cas, il faudrait prendre la borne supérieure sur les fonctions ϕ utilisées sous l'espérance, ce qui n'est pas le cas. L'inégalité (3.60) conduit au même taux de convergence que celui obtenu dans (3.12) d'après (3.55). En revanche, si on s'intéresse uniquement à la vitesse d'approximation de la loi d'une particule par le processus limite de McKean-Vlasov, l'inégalité (3.62) permet de montrer, grâce à la Remarque 3.25, que

$$\sup_{t\in[0,T]} W_1([X_t^{1,N}], \mu_t) \leq \frac{C}{N^{\alpha-1}}. \tag{3.63}$$

A priori, on ne peut pas obtenir un taux de convergence aussi fort avec la méthode de couplage mise en place dans le Théorème 3.8. De plus, l'avantage d'avoir des estimations comme (3.60) ou encore (3.53) dans le Théorème 3.23 plutôt que des contrôles comme ceux du Théorème 3.8 obtenus par couplage est qu'on sépare le terme de donnée initiale d'un terme de reste générique. Cela peut être utile si on a une meilleure vitesse de convergence vers 0 de $\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0)$ que les taux de convergence généraux prouvés dans [FG15].

Les preuves des Théorèmes 3.23 et 3.26 reposent sur la méthode présentée heuristiquement dans la Sous-section 1.2.3. La différence majeure entre les preuves de ces deux résultats réside dans le fait que pour le système de processus d'Ornstein-Uhlenbeck traité dans le Théorème 3.26, il y a l'étape préliminaire, mentionnée précédemment, qui consiste à retirer les sauts de taille plus grande que le nombre de particules N des bruits du système de particules et de l'EDS de McKean-Vlasov.

Idée de preuve du Théorème 3.23. On note pour tout $i \in \{1,\dots,N\}$ et $t \in [0,T]$

$$Z_t^i = \int_0^t \int_{\mathbb{R}^d} z \widetilde{\mathcal{N}}^i(ds, dz),$$

où $\widetilde{\mathcal{N}}^i$ est la mesure de Poisson compensée associée à Z^i. On introduit également $\boldsymbol{\mathcal{N}}^N$, $\widetilde{\boldsymbol{\mathcal{N}}}^N$ et $\boldsymbol{\nu}^N$ qui désignent respectivement la mesure aléatoire de Poisson, la mesure aléatoire de Poisson compensée, et la mesure de Lévy associées au processus α-stable cylindrique $\mathbf{Z}^N$, défini par (3.51). On considère alors U la solution de l'EDP de Kolmogorov rétrograde (3.49), avec condition terminale ϕ au temps T, définie par (3.46).

L'idée importante est de remarquer que l'application $t \in [0,T] \mapsto U(t,\mu_t)$ est constante puisque l'EDS de McKean-Vlasov (3.16) est bien posée. Grâce à la régularité de U montrée au Théorème 3.20, on peut appliquer la formule d'Itô à la projection empirique $\boldsymbol{x} = (x_1,\dots,x_N) \in (\mathbb{R}^d)^N \mapsto U(t,\overline{\mu}^N_{\boldsymbol{x}})$ (voir le Lemme 2.27 pour la régularité et l'expression des dérivées partielles de la projection empirique) et au système de particule (3.52). Ainsi, on obtient pour tout $t \in [0,T)$

$$
\begin{aligned}
&U(t,\overline{\mu}^N_t) - U(t,\mu_t) - \Big(U(0,\overline{\mu}^N_0) - U(0,\mu_0)\Big) \\
&= \int_0^t \partial_t U(s,\overline{\mu}^N_s)\,ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \partial_v \frac{\delta}{\delta m} U(s,\overline{\mu}^N_s)(X^{i,N}_s)\cdot b(s,X^{i,N}_s,\overline{\mu}^N_s)\,ds \\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) - \partial_{\boldsymbol{x}} U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\cdot \boldsymbol{z}\Big]\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \qquad (3.64)\\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N\cap\{|\boldsymbol{z}|\geq 1\}} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N\cap\{|\boldsymbol{z}|< 1\}} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&=: I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned}
$$

Remarquons que I_3 se réécrit de la façon suivante, grâce à l'indépendance des processus stables $(Z^i)_i$,

$$
\begin{aligned}
I_3 = \sum_{i=1}^N \int_0^t \int_{(\mathbb{R}^d)^N} \Big[&U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \\
&- \frac{1}{N}\partial_v \frac{\delta}{\delta m} U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-})\cdot z\Big]\,d\nu(z)\,ds,
\end{aligned}
$$

où $\tilde{\boldsymbol{z}}_{\boldsymbol{i}} := (0,\dots,0,z,0,\dots,0) \in (\mathbb{R}^d)^N$ pour $z \in \mathbb{R}^d$ et où z apparaît sur la i-ème coordonnée. L'objectif est ensuite de faire apparaître l'EDP vérifiée par U qui va permettre de simplifier cette expression. Contrairement au cas brownien traité dans [CdRF21], elle n'apparaît pas directement et il faut la forcer à apparaître et contrôler ensuite le terme d'erreur que cela induit. Pour la faire apparaître, on linéarise I_3 avec la dérivée plate en écrivant

$$
\begin{aligned}
I_3 &= \int_0^t \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Big[\frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x+z) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x) - \partial_v\frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x)\cdot z\Big]\,d\nu(z)\,d\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}(x)\,ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d}\int_0^1 \Big[\frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}) \\
&\qquad\qquad + \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}+z)\Big]\,dw\,d\nu(z)\,ds \\
&=: I_{3,A} + I_{3,B},
\end{aligned}
$$

où $m^i_{s,z,w} := w\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}$. En utilisant l'EDP de Kolmogorov rétrograde (3.49) et (3.64), on obtient la décomposition suivante

$$
\begin{aligned}
& U(t,\overline{\mu}_t^N) - U(t,\mu_t) - \Big(U(0,\overline{\mu}_0^N) - U(0,\mu_0)\Big) \\
&= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d}\int_0^1 \Big[\frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}) \\
&\qquad\qquad + \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})(X^{i,N}_{s^-}+z)\Big]\, dw\, d\nu(z)\, ds \\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N\cap\{|\boldsymbol{z}|\geq 1\}} \Big[U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})\Big]\, \widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \qquad (3.65)\\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N\cap\{|\boldsymbol{z}|< 1\}} \Big[U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})\Big]\, \widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&= I_{3,B} + I_4 + I_5.
\end{aligned}
$$

Les termes I_4 et I_5 sont alors majorés grâce à l'inégalité de Burkholder-Davis-Gundy pour obtenir (3.53), ou en utilisant le fait que ce sont des martingales centrées pour (3.54). Dans tous les cas, on utilise des majorations et contrôles höldériens prouvés pour la fonction U, notamment ceux du Théorème 3.20. Typiquement pour I_5, l'inégalité de Burkholder-Davis-Gundy dans L^β et la borne (3.47) sur $\partial_v \frac{\delta}{\delta m}U$ assurent que

$$
\mathbb{E}|I_5| \leq (\mathbb{E}|I_5|^\beta)^{\frac{1}{\beta}} \leq \frac{C}{N^{1-\frac{1}{\beta}}}.
$$

Concernant le terme $I_{3,B}$, on doit distinguer les grands sauts des petits sauts en le réécrivant comme

$$
\begin{aligned}
I_{3,B} &= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1}\int_0^1\int_0^1 \Big(\partial_v \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+\lambda z) \\
&\qquad\qquad - \partial_v \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})(X^{i,N}_{s^-}+\lambda z)\Big)\cdot z\, d\lambda\, dw\, d\nu(z)\, ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1^c}\int_0^1 \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})(X^{i,N}_{s^-}+z) \\
&\qquad\qquad + \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\mathbf{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-})\, dw\, d\nu(z)\, ds.
\end{aligned}
$$

On parvient alors à borner $\mathbb{E}|I_{3,B}|$ grâce aux majorations et contrôles höldériens prouvés sur U, notamment ceux du Théorème 3.20. On s'intéresse enfin au terme de donnée initiale $U(0,\overline{\mu}_0^N) - U(0,\mu_0)$. Pour démontrer (3.53), on remarque que $U(0,\cdot)$ est lipschitzienne par rapport à W_1 en linéarisant avec la dérivée plate et en utilisant (3.47). Pour prouver (3.54), on développe ce terme en linéarisant avec la dérivée plate et on le contrôle en utilisant le caractère i.i.d. des données initiales et les estimations sur U. On conclut la preuve du Théorème 3.23 en laissant tendre t vers T par continuité puisque $U(T,\cdot) = \phi$. □

Idée de preuve du Théorème 3.26. Dans le cas du Théorème 3.26 pour le système de processus d'Ornstein-Uhlenbeck, on doit adapter la méthode. En effet, comme le coefficient de dérive est non-borné, on n'a pas les mêmes contrôles sur U que ceux Théorème 3.20. En particulier, le gradient de la dérivée plate n'est plus borné en espace, comme c'est le cas dans (3.47). Cela fait apparaître dans les calculs, notamment pour contrôler le terme d'erreur $I_{3,B}$ dans (3.65), le moment d'ordre 2 de la mesure de Lévy ν qui est infini. Pour résoudre ce problème, on tronque les sauts de taille plus grande que le nombre de particules N des bruits de l'EDS de McKean-Vlasov (3.16) et du système de particules (3.50), comme fait dans le Théorème 3.8. Dans le Théorème 3.8, cette étape de troncature des grands sauts permettait uniquement d'améliorer les taux de convergence obtenus dans le Théorème 3.5, pour le système particulier de processus d'Ornstein-Uhlenbeck. Ici, cette étape est nécessaire si on veut appliquer la méthode reposant sur le semi-groupe décrite précédemment pour prouver le Théorème 3.23. On considère ainsi, pour $t \in [0,T]$,

$$Z_{N,t} := \int_0^t \int_{B_N} z\,\widetilde{\mathcal{N}}(ds,dz),$$

et $(X_{N,t})_{t\in[0,T]}$, la solution de l'EDS de McKean-Vlasov

$$\begin{cases} dX_{N,t} &= b(X_{N,t},\mu_{N,t})\,dt + B\,dZ_{N,t}, \quad t\in[0,T],\\ \mu_{N,t} &= [X_{N,t}],\\ X_0 &= \xi, \end{cases} \tag{3.66}$$

où $b(x,\mu) := Ax + A'\int_{\mathbb{R}^d} y\,d\mu(y)$. On note U_N la fonction définie par (3.46) associée à (3.66). On considère également pour $i\in\{1,\dots,N\}$ et $t\in[0,T]$

$$Z^i_{N,t} := \int_0^t \int_{B_N} z\,\widetilde{\mathcal{N}}^i(ds,dz)$$

et le système de particules tronqué

$$\begin{cases} dX^{i,N}_{N,t} = b(X^{i,N}_{N,t},\overline{\mu}^N_{N,t})\,dt + BdZ^i_{N,t}, \quad t\in[0,T], \quad i\in\{1,\dots,N\},\\ \overline{\mu}^N_{N,t} = \dfrac{1}{N}\displaystyle\sum_{j=1}^N \delta_{X^{j,N}_{N,t}},\\ X^{i,N}_{N,0} = \xi^i. \end{cases}$$

En utilisant le lemme de Gronwall, on contrôle, comme dans la preuve du Théorème 3.8, les erreurs faites avec les EDS initiales par

$$\mathbb{E}(W_1(\overline{\mu}^N_t,\overline{\mu}^N_{N,t}) + W_1(\mu_{N,t},\mu_t)) \le \frac{C}{N^{\alpha-1}}.$$

Cela provient du fait que

$$\int_{B_N^c} |z|\,d\nu(z) \le \frac{C}{N^{\alpha-1}},$$

comme utilisé dans la preuve du Théorème 3.8. On utilise ensuite la méthode décrite précédemment pour prouver le Théorème 3.23 afin contrôler

$$\mathbb{E}|\phi(\overline{\mu}^N_{N,T}) - \phi(\mu_{N,T})| \quad \text{et} \quad |\mathbb{E}(\phi(\overline{\mu}^N_{N,T}) - \phi(\mu_{N,T}))|,$$

en utilisant cette fois la fonction U_N et l'EDP associée à (3.66). On prouve les contrôles dont on a besoin sur U_N en utilisant l'expression de la densité de transition découplée associée à (3.66), qui est ici explicite et donnée par

$$p_N(\mu,s,t,x,y) = q_N\left(t-s, y - e^{tA}x - K_t\int_{\mathbb{R}^d} y\,d\mu(y)\right),$$

où $q_N(t,\cdot)$ est la densité du processus Ornstein-Uhlenbeck stable tronqué $\int_0^t e^{(t-s)A}B\,dZ_{N,s}$ et $t\in[0,T]\mapsto K_t$ une fonction explicite. On n'a ainsi pas besoin de mettre en place la méthode basée sur les itérées de Picard et la méthode parametrix utilisées dans [Cav22b] pour prouver le Théorème 3.20. Pour prouver ici les contrôles dont on a besoin sur U_N, on utilise le résultat suivant démontré dans [Cav22a], qui établit des contrôles uniformes en N sur q_N.

Proposition 3.27. *Pour tout $N \geq 1$ et pour tout $t \in (0,T]$, $q_N(t,\cdot) \in \mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^+)$. De plus, pour tout $\beta \in \mathbb{N}^d$ et $\gamma \in [0,\alpha)$, il existe une constante $C > 0$ qui dépend seulement de $T, d, \alpha, \eta, \beta, \gamma$ telle que pour tout $N \in \mathbb{N}$*

$$\int_{\mathbb{R}^d} |x|^\gamma |\partial_x^\beta q_N(t,x)|\, dx \leq C t^{\frac{\gamma - |\beta|}{\alpha}}. \tag{3.67}$$

L'hypothèse (3.57) de non-dégénérescence sur la mesure de Lévy et l'inversibilité de la matrice B sont cruciales pour prouver cette proposition.

□

3.2.7 Vitesse d'approximation de la loi d'une particule par le processus de McKean-Vlasov limite au niveau des densités

On conserve les mêmes notations que dans la sous-section précédente pour présenter le résultat quantitatif sur l'approximation de la loi d'une particule du système (3.50) par la loi du processus de McKean-Vlasov au niveau des densités. On travaille sous l'hypothèse suivante qui assure l'existence d'une densité pour le système de particules.

Hypothèse (H4). On suppose que pour tout $t > 0$, les particules $\boldsymbol{X}_t^N = (X_t^{1,N}, \ldots, X_t^{N,N})$ définies par (3.50), avec condition initiale $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$ au temps 0, admettent une densité sur $(\mathbb{R}^d)^N$ notée $\boldsymbol{p}^N(\mu_0, 0, t, \cdot)$.

Remarque 3.28. Le résultat d'existence d'une densité ne peut pas s'obtenir directement en utilisant la méthode parametrix dont on a parlé précédemment dans l'heuristique de preuve du Théorème 3.14. En effet, le bruit du système de particules $\boldsymbol{Z}^N$ est un processus α-stable cylindrique sur $(\mathbb{R}^d)^N$, c'est-à-dire ayant des coordonnées d-dimensionnelles indépendantes. Cela implique que la mesure de Lévy est dégénérée au sens où elle est supportée uniquement sur les axes d-dimensionnels de $(\mathbb{R}^d)^N$.

Lorsque $d = 1$, l'Hypothèse (H4) est impliquée par l'Hypothèse (H2) en utilisant [CHZ20a, Théorème 1.1]. En effet, l'Hypothèse (H2) assure que la la dérive du système de particules $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N \mapsto (b(t, x_i, \overline{\mu}_{\boldsymbol{x}}^N))_{i \in \{1,\ldots,N\}}$ est höldérienne uniformément par rapport $t \in [0,T]$. Si on veut se passer de l'Hypothèse (H4) en dimension $d > 1$, il s'agirait de voir si [CHZ20a, Théorème 1.1] se généralise dans ce cadre-là.

Sous l'Hypothèse (H4), la première particule admet une densité notée $p^{1,N}(\mu_0, s, t, \cdot)$. Elle est donnée, pour tout $y_1 \in \mathbb{R}^d$, par

$$p^{1,N}(\mu_0, 0, t, y_1) := \int_{(\mathbb{R}^d)^{N-1}} \boldsymbol{p}^N(\mu_0, 0, t, y_1, y_2, \ldots, y_N)\, dy_2 \ldots dy_N.$$

Par échangeabilité en loi, toutes les particules ont la même loi. Rappelons enfin que $p(\mu_0, 0, t, \cdot)$, définie par (3.33), est la densité de μ_t, qui est la loi de la solution de l'EDS de McKean-Vlasov (3.16) au temps t avec donnée initiale $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$ au temps 0.

Théorème 3.29. *Soient $\gamma \in (0,1] \cap (\alpha-1, (2\alpha-2) \wedge (\eta+\alpha-1))$ et $\gamma' \in [\alpha-1,1]$ fixés. On définit, pour $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0,T]$ et $y, z \in \mathbb{R}^d$*

$$q_0(\mu,0,t,y) := \int_{\mathbb{R}^d} \rho^0(t,y-x)\, d\mu(x) \quad \text{et} \quad f(z) := |z|^{1+\gamma}\mathbf{1}_{B_1}(z) + |z|\mathbf{1}_{B_1^c}(z),$$

où ρ^0 est définie par

$$\rho^0(t,x) = t^{-\frac{d}{\alpha}}(1+t^{-\frac{1}{\alpha}}|x|)^{-d-\alpha}.$$

On pose également $\zeta := -\left(1-\frac{2+\gamma}{\alpha}\right) \in (0,1)$ et on note $\mathcal{B}$ la fonction Bêta définie, pour $x, y > 0$, par

$$\mathcal{B}(x,y) := \int_0^1 (1-t)^{-1+x} t^{-1+y}\, dt.$$

Alors, sous les Hypothèses (H2) et (H4), les propriétés suivantes sont satisfaites.

- ***(Borne sur la densité d'une particule).*** *Il existe une constante positive $C = C(d,T,\alpha,\beta,(H2),\gamma)$ telle que pour tout $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0,T]$, $y \in \mathbb{R}^d$ et $N \geq 1$*

$$p^{1,N}(\mu_0,0,t,y) \leq Cq_0(\mu_0,0,t,y) + \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} \frac{t^{k(1-\zeta)}}{k(1-\zeta)} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta),1-\zeta)\right) \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0\left(\mu_0,0,t,y-\sum_{i\in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j), \tag{3.68}$$

où P_k désigne l'ensemble des parties de l'ensemble $\{1,\dots,k\}$ et par convention $q_0\left(\mu_0,0,t,y-\sum_{i\in\emptyset} z_i\right) := q_0(\mu_0,0,t,y)$.

- ***(Estimation ponctuelle d'approximation de la densité d'une particule).*** *Il existe une constante positive $C = C(d,T,\alpha,\beta,(H2),\gamma,\gamma')$ telle que pour tout $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0,T]$, $y \in \mathbb{R}^d$ et $N \geq 1$*

$$\begin{aligned} &\left|p^{1,N}(\mu_0,0,t,y) - p(\mu_0,0,t,y)\right| \\ &\leq \frac{C}{N^{\gamma'}} t^{1-\frac{1+\gamma'}{\alpha}} (1+M_{\gamma'}(\mu_0)) \int_{\mathbb{R}^d} (1+|x|^{\gamma'})\rho^0(t,y-x)\, d\mu_0(x) \\ &\quad + \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} t^{k(1-\zeta)} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta),1-\zeta)\right) \mathcal{B}(1+k(1-\zeta),1-\zeta) \\ &\quad \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0\left(\mu_0,0,t,y-\sum_{i\in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j). \end{aligned} \tag{3.69}$$

- ***(Estimation sur l'approximation de la loi d'une particule en variation totale).*** *Il existe une constante positive $C = C(d,T,\alpha,\beta,(H2),\gamma)$ telle que pour tout $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0,T]$ et $N \geq 1$*

$$d_{TV}([X_t^{1,N}],\mu_t) \leq \frac{C}{N^{\gamma}} t^{1-\frac{1+\gamma}{\alpha}} (1+M_\gamma(\mu_0)), \tag{3.70}$$

et

$$\sup_{t\in[0,T]} d_{TV}([X_t^{1,N}],\mu_t) \leq \frac{C}{N^{\alpha-1}}(1+M_{\alpha-1}(\mu_0)). \tag{3.71}$$

Comparons ce résultat avec l'estimation (3.56) par rapport à la distance de Wasserstein W_1. On trouve le même taux de convergence $N^{-\gamma}$ avec $\gamma \in (0,1] \cap (0,(2\alpha-2)\wedge(\eta+\alpha-1))$. Cependant, dans (3.70), il y a une singularité en temps de même que dans [CdRF21, Théorème 3.5], et qui n'est pas présente dans (3.56). Cette singularité en temps peut être retirée si on veut seulement le taux de convergence moins bon $N^{1-\alpha}$, qui est le même que celui obtenu dans (3.63) pour le système de processus d'Ornstein-Uhlenbeck en interaction. De plus, on retrouve le taux de convergence N^{-1} prouvé dans [CdRF21, Théorème 3.5] dans le cas brownien en prenant formellement $\alpha = 2$ dans le résultat précédent .

La méthode utilisée pour prouver l'estimation ponctuelle (3.69) repose sur des idées analogues à celles présentées dans la sous-section précédente concernant le Théorème 3.23. Elle a été employée dans le cas brownien dans [CdRF21]. Au lieu de considérer le semi-groupe associé à l'EDS de McKean-Vlasov, on travaille directement avec sa densité non-découplée, qui est solution fondamentale de l'EDP de Kolmogorov rétrograde (3.34). L'idée est d'étudier la dynamique de l'application $s\in[0,t)\mapsto p(\overline{\mu}_s^N,s,t,y)$ pour tout $y\in\mathbb{R}^d$. Les deux ingrédients sont alors les suivants. D'un côté, on quantifie par rapport à N l'erreur entre $\mathbb{E}p(\overline{\mu}_s^N,s,t,y)$ et $p(\mu_s,s,t,y) = p(\mu_0,0,t,y)$. De l'autre côté, on prouve que lorsque s tend vers t, alors $\mathbb{E}p(\overline{\mu}_s^N,s,t,y)$ converge simplement vers la $p^{1,N}(\mu_0,0,t,y)$.

Idée de preuve du Théorème 3.29. On commence par donner l'idée de la preuve de la borne (3.68) sur la densité d'une particule. En appliquant la formule d'Itô à la fonction $(s,\boldsymbol{x}) = (s,x_1,\dots,x_N)\in[0,t)\times(\mathbb{R}^d)^N \mapsto p(\overline{\mu}_{\boldsymbol{x}}^N,s,t,y) = \frac{1}{N}\sum_{k=1}^N p(\overline{\mu}_{\boldsymbol{x}}^N,s,t,x_k,y)$ et en gardant les mêmes notations que dans la sous-section précédente, on a

$$\begin{aligned}
p(\overline{\mu}_s^N,s,t,y) - p(\overline{\mu}_0^N,0,t,y) = &\int_0^s \partial_r p(\overline{\mu}_r^N,r,t,y)\,dr\\
&+\frac{1}{N}\sum_{i=1}^N\int_0^s \partial_v\frac{\delta}{\delta m}p(\overline{\mu}_r^N,r,t,y)(X_{r^-}^{i,N})\cdot b(r,X_{r^-}^{i,N},\overline{\mu}_r^N)\,dr\\
&+\sum_{i=1}^N\int_0^s\int_{(\mathbb{R}^d)^N}\Big[p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}+\tilde{\boldsymbol{z}}_i},r,t,y) - p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}},r,t,y)\\
&\qquad -\frac{1}{N}\partial_v\frac{\delta}{\delta m}p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}},r,t,y)(X_{r^-}^{i,N})\cdot z\Big]\,d\nu(z)\,dr\\
&+\int_0^s\int_{(\mathbb{R}^d)^N} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}+\boldsymbol{z}},r,t,y) - p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}},r,t,y)\,\widetilde{\boldsymbol{\mathcal{N}}}^N(dr,d\boldsymbol{z})\\
=:&\, I_1+I_2+I_3+I_4.
\end{aligned} \tag{3.72}$$

En utilisant les contrôles prouvés sur les dérivées de p, on prouve que I_4 définit une vraie martingale et que

$$\mathbb{E}I_4 = 0. \tag{3.73}$$

En linéarisant avec la dérivée plate, on décompose I_3 de la manière suivante

$$\begin{aligned}
I_3 = \int_0^s \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Big[& \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(x+z) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(x) \\
& - \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(x) \cdot z \Big] \, d\nu(z) \, d\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}(x) \, dr \\
+ \frac{1}{N} \sum_{i=1}^N \int_0^s \int_{B_1^c} \int_0^1 \Big[& \frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-} + z) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-} + z) \Big] \\
& - \Big[\frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-}) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-}) \Big] \, dw \, d\nu(z) \, dr \\
+ \frac{1}{N} \sum_{i=1}^N \int_0^s \int_{B_1} \int_0^1 \Big[& \frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-} + z) - \frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-}) \\
& + \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-}) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-} + z) \Big] \, dw \, d\nu(z) \, dr \\
=: I_{3,A} + I_{3,B} + I_{3,C},
\end{aligned}$$

où $m^i_{r,z,w} := w \overline{\mu}^N_{\boldsymbol{X}^N_{r^-} + \bar{z}_i} + (1-w) \overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}$. En prenant l'espérance dans (3.72) et en utilisant l'EDP de Kolmogorov rétrograde vérifiée par p du Théorème 3.18, on déduit que

$$\mathbb{E} p(\overline{\mu}^N_s, s, t, y) = \mathbb{E} p(\overline{\mu}^N_0, 0, t, y) + \mathbb{E} I_{3,B} + \mathbb{E} I_{3,C}. \tag{3.74}$$

On parvient alors à montrer, en utilisant les estimations prouvées sur les dérivées de la densité p, que

$$\begin{aligned}
& |\mathbb{E} I_{3,B}| + |\mathbb{E} I_{3,C}| \\
& \quad \leq \frac{C}{N^\gamma} \int_0^s \int_{\mathbb{R}^d} (t-r)^{-\zeta} f(z) \mathbb{E}(\rho^0(t-r, y - X^{1,N}_{r^-}) + \rho^0(t-r, y - X^{1,N}_{r^-} - z)) \, d\nu(z) \, dr.
\end{aligned} \tag{3.75}$$

Par le même raisonnement que dans le preuve de [CdRF21, Théorème 3.5], on obtient

$$\lim_{s \to t^-} \mathbb{E} p(\overline{\mu}^N_s, s, t, y) = p^{1,N}(\mu_0, 0, t, y). \tag{3.76}$$

L'argument repose sur le développement en série parametrix de p et le fait que $\widehat{p}(s, t, \cdot, y) \underset{s \to t^-}{\to} \delta_y$, au sens faible. En utilisant que $p(\mu_0, 0, t, y)$ est majorée par $q_0(\mu_0, 0, t, y)$, on déduit de (3.74), (3.75) et (3.76) que pour tout $t \in (0, T]$, $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $y \in \mathbb{R}^d$ et $N \geq 1$

$$\begin{aligned}
p^{1,N}(\mu_0, 0, t, y) \leq C q_0(\mu_0, 0, t, y) + \frac{C}{N^\gamma} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} & (t-r)^{-\zeta} f(z) [\rho^0(t-r, y-w-z) \\
& + \rho^0(t-r, y-w)] p^{1,N}(\mu_0, 0, r, w) \, dw \, d\nu(z) \, dr.
\end{aligned}$$

En itérant cette relation implicite, on parvient à démontrer la borne (3.68) sur $p^{1,N}$.

Pour prouver l'estimation (3.69), on utilise le lemme suivant qui se démontre en linéarisant avec la dérivée plate et en utilisant les estimations démontrées sur la densité p et ses dérivées.

Lemme 3.30. *Pour tout $\gamma \in [\alpha - 1, 1]$, il existe une constante positive $C > 0$ telle que pour tout*

$\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0,T]$, $y \in \mathbb{R}^d$ *et* $N \geq 1$

$$\left|\mathbb{E}p(\overline{\mu}_0^N, 0, t, y) - p(\mu_0, 0, t, y)\right| \leq \frac{C}{N^\gamma} t^{1-\frac{1+\gamma}{\alpha}} (1 + M_\gamma(\mu_0)) \int_{\mathbb{R}^d} (1 + |x|^\gamma) \rho^0(t, y - x)\, d\mu_0(x).$$

Expliquons maintenant comment on déduit l'estimation (3.69). On revient à l'identité (3.74) à laquelle on soustrait $p(\mu_0, 0, t, y)$ de chaque côté, et on utilise (3.74) et (3.75) pour obtenir

$$|\mathbb{E}p(\overline{\mu}_s^N, s, t, y) - p(\mu_0, 0, t, y)| \leq |\mathbb{E}p(\overline{\mu}_0^N, 0, t, y) - p(\mu_0, 0, t, y)| + |\mathbb{E}I_{3,B}| + |\mathbb{E}I_{3,C}|. \tag{3.77}$$

On utilise alors la borne (3.68) sur $p^{1,N}$ obtenue précédemment et (3.75), qui assurent que

$$\begin{aligned} |\mathbb{E}I_{3,B}| + |\mathbb{E}I_{3,C}| \leq \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} t^{k(1-\zeta)} \left(\prod_{j=1}^{k-2} \mathcal{B}(j(1-\zeta), 1-\zeta) \right) \mathcal{B}(1 + (k-1)(1-\zeta), 1-\zeta) \\ \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0 \left(\mu_0, 0, t, y - \sum_{i \in I} z_i \right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j). \end{aligned}$$

En combinant l'inégalité précédente, le Lemme 3.30 avec $\gamma' \in [\alpha - 1, 1]$, et en prenant la limite $s \to t^-$ dans (3.77) grâce à (3.76), on conclut la preuve de (3.69).

Les estimations (3.70) et (3.71) découlent directement de (3.69). En effet, il suffit d'intégrer (3.69) pour $y \in \mathbb{R}^d$. En utilisant l'asymptotique de la fonction Bêta, on conclut la preuve de (3.70) (en prenant $\gamma' = \gamma$ dans (3.69)) et (3.71) (en prenant $\gamma' = \alpha - 1$ dans (3.69)).

□

3.3 Perspectives

3.3.1 Propagation du chaos faible avec un coefficient de diffusion non constant

Une première piste de recherche pour généraliser le travail [Cav22b] est d'étudier le cas où il y a un coefficient de diffusion $\sigma : (t, x, \mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^{d \times d}$ devant le bruit Z dans (3.16) qui n'est plus constant. Pour ce faire, il serait intéressant d'étudier si on peut se servir des outils utilisés dans [MZ22]. En particulier, on pourrait utiliser le fait que la densité de l'intégrale d'une fonction déterministe contre le processus α-stable rotationnellement invariant Z peut être exprimée à l'aide d'un mouvement brownien et d'un $\frac{\alpha}{2}$-subordinateur indépendant, ainsi que les développements en séries parametrix utilisés dans [MZ22]. On peut également envisager d'autoriser les coefficients b et σ à ne plus être bornés et de traiter une partie des cas critique et super-critique $\alpha \leq 1$ en utilisant les outils de [MZ22], c'est-à-dire en utilisant le flot rétrograde associé à l'équation différentielle ordinaire sous-jacente pour pouvoir écrire les développements en série parametrix.

3.3.2 Propagation du chaos au niveau des densités

Nous avons commencé à réfléchir aux perspectives présentée dans cette sous-section avec Paul-Éric Chaudru de Raynal.

Dans le Théorème 3.29, on prouve un résultat quantitatif sur l'approximation de la densité d'une particule par la densité du processus de McKean-Vlasov limite, qui est inspiré de [CdRF21, Théorème 3.5] dans le cas brownien. On note, pour $k \in \{1, \ldots, N\}$, $p^{k,N}(\mu, 0, t, \cdot)$ la densité sur $(\mathbb{R}^d)^k$ de k particules au temps t et initialisées au temps 0 par des variables i.i.d. de loi μ. On peut se demander si on peut procéder de même pour obtenir une estimation de propagation du chaos sur

$$\left| \prod_{j=1}^{k} p(\mu, 0, t, y_j) - p^{k,N}(\mu, 0, t, y_1, \ldots, y_k) \right|,$$

pour $t \in (0, T]$ et $y_1, \ldots, y_k \in \mathbb{R}^d$, afin d'exhiber une majoration, explicite par rapport à N et k. Cela permettrait de quantifier la vitesse de propagation du chaos vis à vis de la distance en variation totale. Au lieu de considérer la densité découplée $p(\mu, s, t, x, y)$, solution de l'EDP de Kolmogorov rétrograde (3.34), on est amené à considérer $\prod_{j=1}^{k} p(\mu, s, t, x_j, y_j)$, qui est la densité découplée de k copies i.i.d. de l'EDS de McKean-Vlasov, initialisées respectivement en les $x_j \in \mathbb{R}^d$. L'application $(\mu, s, x_1, \ldots, x_k) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times (\mathbb{R}^d)^k \mapsto \prod_{j=1}^{k} p(\mu, s, t, x_j, y_j)$ vérifie une EDP analogue à (3.34). Il s'agit alors de voir comment adapter le reste de la preuve de [CdRF21, Théorème 3.5] dans le cas brownien ou du Théorème 3.29 dans le cas stable. Cela permettrait de quantifier la propagation du chaos en contrôlant la distance en variation totale entre la loi de k particule $\mu_t^{1:k,N}$ et la loi $\mu_t^{\otimes k}$, où $(\mu_t)_{t\in[0,T]}$ est le flot de lois marginales de l'EDS de McKean-Vlasov limite. De plus, comme mentionné dans la Sous-section 1.2.3, à partir du résultat de propagation du chaos faible quantitative du Théorème 3.23, on peut quantifier la distance entre la loi de la mesure empirique $[\overline{\mu}_t^N] \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ et δ_{μ_t} (voir (1.9)). On peut se demander si on peut alors retrouver un contrôle entre $\mu_t^{1:k,N}$ et $\mu_t^{\otimes k}$ à partir de ce résultat, et pour une certaine distance sur $\mathcal{P}((\mathbb{R}^d)^k)$.

3.3.3 Propagation du chaos au niveau des trajectoires par couplage

On peut également se demander si on peut prouver la propagation du chaos au niveau des trajectoires par couplage pour (3.16), sous des hypothèses höldériennes sur la dérive b, qui sont analogues à l'Hypothèse (H2). Plus précisément, on introduit $(X^{i,\infty})_{i\geq 1}$ des copies i.i.d. de (3.16) où les bruits et les données initiales sont les mêmes que pour le système de particules (1.6). L'objectif est d'exhiber des taux de convergence, explicites par rapport à N, pour $\mathbb{E}W_1(\overline{\mu}_T^N, \mu_T)$ et $\sup_{1\leq k\leq N} \mathbb{E}|X_T^{k,N} - X_T^{k,\infty}|$. On pourra utiliser la transformation de Zvonkin, comme fait dans [CdRF21], qui permet de retirer la dérive irrégulière afin de rendre b lipschitzienne, au prix d'un coefficient de diffusion σ qui n'est plus constant mais lipschitzien. La première étape serait donc d'établir un résultat de propagation du chaos au niveau des trajectoires pour (3.1) dans le cas où b et σ sont lipschitziens, avec σ non constant, et Z est un processus α-stable. Le cas $\sigma = \mathrm{Id}$ a été traité au Théorème 3.5 et il s'agirait donc d'ajouter un coefficient lipschitzien σ général, ce qui ne semble pas trivial (voir la Remarque 3.7). La seconde étape consisterait à ajouter un terme source à l'EDP de Kolmogorov rétrograde (3.49), pour pouvoir utiliser la transformation de Zvonkin dans ce cadre McKean-Vlasov. Cela a été fait dans le cas brownien dans [CdRF22, CdRF21].

3.3.4 Problèmes de contrôle pour des EDS de McKean-Vlasov dirigées par des processus de Lévy

Grâce à la formule d'Itô le long d'un flot de mesures du Théorème 3.9, on peut envisager de s'intéresser à des problèmes de contrôle pour des EDS de McKean-Vlasov dirigées par des processus de Lévy. Il s'agit, grâce à la formule d'Itô, de mettre en œuvre une méthode basée sur un principe de programmation dynamique, comme présenté dans [CD18a, Chapitre 6]. Dans le cas d'un bruit à sauts L^2, cela a été par Guo, Pham et Wei dans [GPW20]. On peut donc envisager de considérer des bruits de Lévy ayant un moment d'ordre $\beta \in (0, 2]$ grâce au Théorème 3.9.

Chapitre 4

Formule d'Itô-Krylov le long d'un flot de mesures de probabilité

Dans ce chapitre, on présente le travail [Cav21].

4.1 Objectif

Fixons $T > 0$ un horizon de temps fini, $d, d_1 \in \mathbb{N}^*$ avec $d_1 \geq d$, et $(B_t)_{t\geq 0}$ un $(\mathcal{F}_t)_{t\geq 0}$-mouvement brownien de dimension d_1. On s'intéresse au processus d'Itô sur $\mathbb{R}^d$ défini, pour $t \in [0, T]$, par

$$X_t := X_0 + \int_0^t b_s \, ds + \int_0^t \sigma_s \, dB_s, \tag{4.1}$$

où $X_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{R}^d)$, $b : [0, T] \times \Omega \to \mathbb{R}^d$ et $\sigma : [0, T] \times \Omega \to \mathbb{R}^{d\times d_1}$ sont des processus progressivement mesurables. On notera, pour $t \in [0, T]$, $\mu_t := [X_t]$ la loi de X_t et a la matrice $\sigma\sigma^*$. On fait les hypothèses suivantes.

(A) Il existe $K > 0$ tel que presque sûrement

$$\forall t \in [0, T],\ |b_t| + |\sigma_t| \leq K.$$

(B) Il existe $\delta > 0$ tel que presque sûrement

$$\forall t \in [0, T],\ \forall \lambda \in \mathbb{R}^d,\ a_t\lambda \cdot \lambda \geq \delta|\lambda|^2.$$

Le travail [Cav21] repose sur [Kry09, Section 2.10], où Krylov prouve une formule d'Itô pour l'application $t \in [0, T] \mapsto f(X_t)$, avec f appartenant à un certain espace de Sobolev sur $\mathbb{R}^d$ (voir le Théorème 4.4). Le but de ce travail est d'établir, de façon analogue, une formule d'Itô le long du flot $(\mu_t)_{t\in[0,T]}$ avec des hypothèses de type Sobolev sur la fonction $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ à laquelle on veut appliquer la formule d'Itô. En particulier, l'objectif est d'affaiblir les hypothèses de régularité habituelles pour appliquer la formule d'Itô (voir le Théorème 2.28), au prix d'hypothèses d'intégrabilité. Plus précisément, on ne demande plus à ce que pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, l'application $\frac{\delta}{\delta m}u(\mu)$ soit $\mathcal{C}^2$, mais seulement que $\partial_v \frac{\delta}{\delta m}u(\mu)$ appartienne à un espace de Sobolev.

Avant de présenter les résultats de [Cav21], mentionnons d'autres extensions de la formule d'Itô pour un flot de mesures. Cette dernière a été établie pour des flots de mesures associés à des semi-martingales càdlàg L^2 générales. Cela a été fait indépendamment par Guo, Pham et Wei dans [GPW20], qui ont étudié des problèmes de contrôle d'EDS de McKean-Vlasov à sauts et par Talbi, Touzi et Zhang dans [TTZ21] qui ont travaillé sur un problème d'arrêt optimal en champ moyen. Dans ces deux travaux, des principes de programmation dynamique sont établis grâce à la formule d'Itô. Dans le cas de processus à sauts n'admettant pas un moment d'ordre 2 fini, on renvoie au Théorème 3.9 présenté dans le chapitre précédent. On peut également mentionner les formules d'Itô-Wentzell-Lions pour des fonctionnelles dépendant d'un champ aléatoire de type Itô et d'un flot de mesures qui ont été prouvées par Dos Reis et Platonov dans [dRP22].

4.2 Contexte : inégalité de Krylov et formule d'Itô-Krylov

Le résultat sur lequel repose ce travail est l'inégalité de Krylov [Kry09, Théorème 4, Section 2.3]. Elle est rappelée dans le théorème suivant.

Théorème 4.1 (Inégalité de Krylov)**.** *Soit $b : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ et $\sigma : \mathbb{R}^+ \times \Omega \to \mathbb{R}^{d\times d_1}$ deux fonctions progressivement mesurables. On suppose que $p, d_1 \geq d$. De plus, on suppose qu'il existe $K > 0$ et $\delta > 0$ tels que pour tout $(t, \omega) \in \mathbb{R}^+ \times \Omega$*

$$|b_t(\omega)| + |\sigma_t(\omega)| \leq K,$$

et

$$\forall \lambda \in \mathbb{R}^d,\ (\sigma\sigma^*)_t(\omega)\lambda \cdot \lambda \geq \delta|\lambda|^2$$

Pour X_0 une variable aléatoire à valeurs dans $\mathbb{R}^d$ et mesurable par rapport à $\mathcal{F}_0$, on définit le processus d'Itô $X = (X_t)_{t\geq 0}$ par

$$\forall t \geq 0,\ X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dB_s.$$

Soit $\lambda > 0$. Alors, il existe une constante $N = N(d, p, \lambda, \delta, K) > 0$, telle que pour toute fonction mesurable $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$

$$\mathbb{E} \int_0^\infty e^{-\lambda t} |f(t, X_t)|\, dt \leq N \|f\|_{L^{p+1}(\mathbb{R}^+\times\mathbb{R}^d)}.$$

On utilise plus précisément le corollaire suivant.

Corollaire 4.2. *Si b et σ satisfont les Hypothèses **(A)** et **(B)**, il existe $N_1 = N_1(d, p, \delta, K, T) > 0$, tel que pour toute fonction mesurable $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$, on a*

$$\mathbb{E} \int_0^T |f(s, X_s)|\, ds \leq N_1 \|f\|_{L^{p+1}([0,T]\times\mathbb{R}^d)}.$$

L'inégalité de Krylov garantit l'existence d'une densité par rapport à la mesure de Lebesgue pour μ_s, pour presque tout $s \in [0, T]$, ainsi qu'une propriété d'intégrabilité sur cette dernière. C'est précisément de ce résultat que provient l'effet régularisant utilisé dans la suite pour prouver la formule d'Itô.

Proposition 4.3. *Sous les Hypothèses **(A)** et **(B)**, il existe une fonction $p \in L^1([0, T] \times \mathbb{R}^d; \mathbb{R}^+) \cap L^{(d+1)'}([0, T] \times \mathbb{R}^d; \mathbb{R}^+)$ telle que pour toute fonction mesurable $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}^+$*

$$\int_0^T \mathbb{E} f(s, X_s)\, ds = \int_{[0,T]\times\mathbb{R}^d} f(s, x)p(s, x)\, dx\, ds. \tag{4.2}$$

Si τ est un temps d'arrêt tel que $(X_t)_{t\in[0,T]}$ appartient presque sûrement à B_R sur l'ensemble $\{\tau > 0\}$, alors

$$\mathbb{E} \int_0^{\tau\wedge T} f(s, X_s)\, ds \leq \int_{[0,T]\times B_R} f(s, x)p(s, x)\, dx\, ds. \tag{4.3}$$

De plus, pour presque tout $s \in [0,T]$, $\mu_s = [X_s]$ admet $p(s,\cdot)$ comme densité et $p(s,\cdot) \in L^{(d+1)'}(\mathbb{R}^d)$ pour presque tout $s \in [0,T]$ par Fubini-Tonelli.

Idée de preuve de la Proposition 4.3. On considère la mesure μ définie pour $A \in \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathbb{R}^d)$ par

$$\mu(A) = \int_0^T \mathbb{E}\mathbf{1}_A(s, X_s)\, ds.$$

D'après l'inégalité de Krylov, pour tout $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^+$ mesurable, on a

$$\int_0^T \mathbb{E} f(s, X_s)\, ds = \int_{[0,T]\times\mathbb{R}^d} f(s,x)\, d\mu(s,x) \leq C\|f\|_{L^{p+1}([0,T]\times\mathbb{R}^d)}. \tag{4.4}$$

Le théorème de Radon-Nikodym assure qu'il existe $p \in L^1([0,T] \times \mathbb{R}^d; \mathbb{R}^+)$ tel que pour toute fonction mesurable $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^+$

$$\int_0^T \mathbb{E} f(s, X_s)\, ds = \int_{[0,T]\times\mathbb{R}^d} f(s,x)p(s,x)\, dx\, ds. \tag{4.5}$$

L'application $f \in L^{d+1}([0,T] \times \mathbb{R}^d) \mapsto \int_{[0,T]\times\mathbb{R}^d} f(s,x)p(s,x)\, dx\, ds$ est une forme linéaire continue sur $L^{d+1}([0,T]\times\mathbb{R}^d)$ d'après (4.4). On déduit par dualité que p appartient à $L^{(d+1)'}([0,T]\times\mathbb{R}^d)$ et on conclut par un argument de classe monotone.

□

Énonçons enfin la formule d'Itô-Krylov [Kry09, Théorème 1, Section 2.10].

Théorème 4.4 (Formule d'Itô-Krylov)**.** *Soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction continue telle que sa dérivée au sens des distributions ∇g appartient à l'espace de Sobolev $W^{1,k}_{loc}(\mathbb{R}^d)$, pour $k \geq d+1$. Sous les Hypothèses* ***(A)*** *et* ***(B)*** *sur b et σ, on a presque sûrement pour tout $t \in [0,T]$*

$$f(X_t) = f(X_0) + \int_0^t \nabla f(X_s) \cdot b_s\, ds + \int_0^t \nabla f(X_s) \cdot (\sigma_s dB_s) + \frac{1}{2}\int_0^t \nabla^2 f(X_s) \cdot a_s\, ds.$$

4.3 Formule d'Itô-Krylov le long d'un flot de mesures et exemples

On présente ici la formule d'Itô-Krylov le long du flot de mesures $(\mu_t)_{t\in[0,T]}$, défini comme la famille de lois marginales de (4.1). Notre but est de tirer profit de l'effet régularisant du bruit, qui ne dégénère pas grâce à l'Hypothèse **(B)**. Cet effet provient plus précisément de l'existence des densités $p(t,\cdot)$ de la Proposition 4.3 et de leur propriété d'intégrabilité. Si on regarde la formule d'Itô (4.6), l'effet régularisant provient notamment des espérances qui moyennisent, par rapport à la variable d'espace, les dérivées de $\frac{\delta}{\delta m}u$ sur toutes les trajectoires de $(X_t)_{t\geq 0}$. En effet, la régularisation par le bruit intervient uniquement à travers la variable spatiale de la dérivée plate et pas à travers la variable mesure comme on le verra dans les hypothèses. Ce n'est pas surprenant car l'espace $\mathcal{P}_2(\mathbb{R}^d)$ est de dimension infinie alors que le bruit brownien est de dimension finie. On ne peut donc pas espérer un effet régularisant important par rapport à la variable mesure de $\frac{\delta}{\delta m}u$. Le fait qu'un bruit de dimension finie ne peut pas avoir un effet régularisant total sur $\mathcal{P}_2(\mathbb{R}^d)$ est illustré dans [Mar20] dans le contexte des EDS de McKean-Vlasov. Ce phénomène de régularisation, lié à l'existence d'une densité et de bonnes propriétés sur celle-ci, est comparable à celui présenté dans la Sous-section 3.2.5 pour le semi-groupe associé à une

EDS de McKean-Vlasov dirigée par un processus α-stable. Ici, on a seulement l'existence d'une densité pour presque tout temps et une propriété d'intégrabilité sur cette dernière alors que dans la Section 3.2, on a des bornes ponctuelles sur la densité et ses dérivées.

Introduisons tout d'abord l'espace de mesures sur lequel nous allons travailler. Celui-ci est adapté aux propriétés démontrées précédemment sur la densité p. Soit $(\rho_n)_{n\geq 1}$ une approximation de l'unité sur $\mathbb{R}^d$, c'est-à-dire une suite de fonctions positives $\mathcal{C}^\infty$ telles que pour tout $n \geq 1$, $\int_{\mathbb{R}^d} \rho_n(x)\, dx = 1$ et ρ_n est nulle en dehors de la boule $B_{1/n}$.

Définition 4.5. On définit $\mathscr{P}(\mathbb{R}^d)$ comme l'espace des mesures $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ admettant une densité $\frac{d\mu}{dx}$ par rapport à la mesure de Lebesgue qui appartient à $L^{(d+1)'}(\mathbb{R}^d)$. On munit $\mathscr{P}(\mathbb{R}^d)$ d'une distance $d_{\mathscr{P}}$ satisfaisant les conditions suivantes.

- **(H1)** Pour tout $n \geq 1$, $\mu \in (\mathcal{P}_2(\mathbb{R}^d), W_2) \mapsto \mu * \rho_n \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$ est continue.
- **(H2)** Pour tout $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\mu * \rho_n \underset{n\to+\infty}{\longrightarrow} \mu$ pour la distance $d_{\mathscr{P}}$.

Remarquons que pour tout $n \geq 1$ et $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu * \rho_n \in \mathscr{P}(\mathbb{R}^d)$. En effet, sa densité est donnée par $x \mapsto \rho_n * \mu(x) = \int_{\mathbb{R}^d} \rho_n(x-y)\, d\mu(y)$. L'inégalité de Jensen assure qu'elle appartient à $L^{(d+1)'}(\mathbb{R}^d)$. L'espace $\mathscr{P}(\mathbb{R}^d)$ apparaît naturellement puisque d'après la Proposition 4.3, il existe $p \in L^1([0,T] \times \mathbb{R}^d; \mathbb{R}^+) \cap L^{(d+1)'}([0,T] \times \mathbb{R}^d; \mathbb{R}^+)$ tel que pour presque tout $t \in [0,T]$, μ_t admet comme densité $p(t,\cdot)$ et appartient donc à $\mathscr{P}(\mathbb{R}^d)$. Donnons deux exemples de choix pour la distance $d_{\mathscr{P}}$.

Exemple 4.6. La distance de Wasserstein W_2 satisfait clairement les Hypothèses **(H1)** et **(H2)** de la Définition 4.5. D'autres distances possibles, qu'on note d_k, sont définies, pour $k \in [d+1, +\infty[$, $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$, par

$$d_k(\mu,\nu) = \left\| \frac{d\mu}{dx} - \frac{d\nu}{dx} \right\|_{L^{k'}(\mathbb{R}^d)}.$$

La distance d_k est bien définie puisque pour $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\frac{d\mu}{dx} \in L^1(\mathbb{R}^d) \cap L^{(d+1)'}(\mathbb{R}^d)$, qui est contenu dans $L^{k'}(\mathbb{R}^d)$ par interpolation.

Introduisons maintenant l'espace de fonctions définies sur $\mathcal{P}_2(\mathbb{R}^d)$ pour lequel nous allons prouver la formule d'Itô-Krylov.

Définition 4.7. On définit $\mathcal{W}_1(\mathbb{R}^d)$ comme l'espace des fonctions continues $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ admettant une dérivée plate $\frac{\delta}{\delta m}u$ telle que pour tout $\mu \in \mathscr{P}(\mathbb{R}^d)$, la fonction $\frac{\delta}{\delta m}u(\mu)$ admet des dérivées au sens des distributions d'ordre 1 et 2 dans $L^k(\mathbb{R}^d)$, pour un certain $k \geq d+1$, et vérifie les propriétés suivantes.

(1) L'application $\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \partial_v \frac{\delta}{\delta m}u(\mu) \in \left(W^{1,k}(\mathbb{R}^d)\right)^d$ est continue pour un choix de distance $d_{\mathscr{P}}$ vérifiant **(H1)** et **(H2)**.

(2) Il existe $\alpha \in \mathbb{N}$ tel que $k \geq (1+\alpha)d$ et pour tout compact $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ et $\mu \in \mathcal{K} \cap \mathscr{P}(\mathbb{R}^d)$

$$\left\| \partial_v \frac{\delta}{\delta m} u(\mu) \right\|_{W^{1,k}(\mathbb{R}^d)} \leq C_{\mathcal{K}} \left(1 + \left\| \frac{d\mu}{dx} \right\|^{\alpha}_{L^{k'}(\mathbb{R}^d)} \right).$$

Remarque 4.8. (1) Nous avons choisi de travailler avec la dérivée plate plutôt que la L-dérivée. On pourra voir la Remarque 4.12 pour une justification de ce choix.

(2) L'espace $\mathcal{W}_1(\mathbb{R}^d)$ contient les fonctions vérifiant le point (1) de la Définition 4.7 avec $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ au lieu de $(\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$. En effet, le second point est clairement vérifié avec $\alpha = 0$ comme $\mathcal{K}$ est compact.

(3) L'hypothèse (2) de la Définition 4.7 permet de contrôler la croissance de $\left\|\partial_v \frac{\delta}{\delta m} u(\mu)\right\|_{W^{1,k}(\mathbb{R}^d)}$ par rapport à la mesure μ. Elle permet de tirer profit à la fois de la continuité du flot $\mathcal{P}_2(\mathbb{R}^d)$, mais aussi des propriétés d'intégrabilité de sa densité.

(4) Le théorème d'injection de Sobolev [Bre10, Corollaire 9.14] assure que pour tout $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(\mu)$ est de classe $\mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ et que $\partial_v \frac{\delta}{\delta m} u(\mu)$ est bornée et γ-höldérienne, où $\gamma := 1 - \frac{d}{k}$. Appuyons sur le fait qu'on n'a pas besoin de supposer que $\frac{\delta}{\delta m} u(\mu) \in W^{2,k}(\mathbb{R}^d)$ car il n'y a pas besoin d'hypothèse d'intégrabilité sur la dérivée plate $\frac{\delta}{\delta m} u$ puisqu'elle n'apparaît pas dans la formule d'Itô le long d'un flot de mesure (voir le Théorème 2.28).

On peut maintenant énoncer la formule d'Itô-Krylov pour les fonctions dans $\mathcal{W}_1(\mathbb{R}^d)$ démontrée dans [Cav21].

Théorème 4.9 (Formule d'Itô-Krylov)**.** *On suppose toujours que les Hypothèses* ***(A)*** *et* ***(B)*** *sur b et σ sont satisfaites. Soit u une fonction de $\mathcal{W}_1(\mathbb{R}^d)$. Alors on a pour tout $t \in [0,T]$*

$$u(\mu_t) = u(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s\right) ds + \frac{1}{2}\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot a_s\right) ds. \tag{4.6}$$

Donnons maintenant des exemples de fonctions dans $\mathcal{W}_1(\mathbb{R}^d)$.

Exemple 4.10 (Fonctionnelle linéaire). Soit $g \in \mathcal{C}^0(\mathbb{R}^d;\mathbb{R})$ admettant une dérivée au sens des distributions $\nabla g \in (W^{1,k}(\mathbb{R}^d))^d$ pour un certain $k \geq d+1$. Alors, la fonction

$$u : \left\{ \begin{array}{ll} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \mu & \mapsto \displaystyle\int_{\mathbb{R}^d} g(x)\, d\mu(x), \end{array} \right.$$

appartient à $\mathcal{W}_1(\mathbb{R}^d)$.

En effet, le théorème d'injection de Sobolev [Bre10, Corollaire 9.14] assure que $\nabla g \in L^\infty(\mathbb{R}^d)$ puisque $k \geq d+1$. Ainsi g est à croissance au plus linéaire et donc pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(\mu) = g$ (voir l'Exemple 2.20), qui satisfait clairement les hypothèses (1) et (2) (avec $\alpha = 0$) de la Définition 4.7.

On s'intéresse maintenant au cas multilinéaire.

Exemple 4.11 (Polynômes sur l'espace de Wasserstein). Soient $N \geq 2$ et $g \in \mathcal{C}^0((\mathbb{R}^d)^N;\mathbb{R})$ vérifiant les propriétés suivantes.

(1) Il existe $C > 0$ tel que pour tout $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$

$$|g(\boldsymbol{x})| \leq C(1 + |x_1|^2 + \cdots + |x_N|^2).$$

(2) La dérivée au sens des distributions ∇g appartient à $(W^{1,k}((\mathbb{R}^d)^N))^{Nd}$ pour un certain $k \in [Nd, +\infty[$.

Alors la fonction

$$u : \left\{ \begin{array}{ll} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \mu & \mapsto \displaystyle\int_{(\mathbb{R}^d)^N} g(x_1, \dots, x_N)\, d\mu(x_1) \dots d\mu(x_N), \end{array} \right.$$

appartient à $\mathcal{W}_1(\mathbb{R}^d)$ pour $d_{\mathscr{P}} = d_k$.

Remarque 4.12. (1) Dans la Définition 4.7, les dérivées distributionnelles d'ordre 1 et 2 de $\frac{\delta}{\delta m}u(\mu)$ ne sont pas nécessairement des fonctions intégrables pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Bien sûr, c'est le cas dans l'Exemple 4.10 puisque la dérivée plate ne dépend pas de la mesure μ. Cependant, dans l'Exemple 4.11 avec $N = 2$, la dérivée plate est donnée par

$$\frac{\delta}{\delta m}u(\mu)(v) = \int_{\mathbb{R}^d} g(v,y)\,d\mu(y) + \int_{\mathbb{R}^d} g(y,v)\,d\mu(y), \tag{4.7}$$

d'après l'Exemple 2.22. Formellement, la dérivée distributionnelle par rapport à v de la première intégrale dans (4.7) est

$$\int_{\mathbb{R}^d} \partial_v g(v,y)\,d\mu(y).$$

Ce terme n'a aucune raison d'être bien défini pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ car on a seulement supposé que $\nabla g \in (W^{1,k}(\mathbb{R}^{2d}))^{2d}$ avec $k \geq 2d$. En effet, pour $k = 2d$, on sait seulement par injection de Sobolev que ∇g appartient à $(L^r(\mathbb{R}^{2d}))^{2d}$ avec $r \in [2d, +\infty[$. Cela justifie aussi notre choix de travailler avec la dérivée plate plutôt que la L-dérivée. En effet, la L-dérivée de u serait égale à $\partial_v \frac{\delta}{\delta m}u(\mu)$ d'après la Proposition 2.24, terme qui n'est pas bien défini pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ comme on l'a vu. Ainsi u n'est pas nécessairement L-dérivable.

(2) Nos hypothèses sur les dérivées spatiales de $\frac{\delta}{\delta m}u$ dans la Définition 4.7 sont faites seulement sur $\mathscr{P}(\mathbb{R}^d)$ plutôt que sur tout l'espace $\mathcal{P}_2(\mathbb{R}^d)$ essentiellement car dans la formule d'Itô (4.6), ces dérivées apparaissent seulement dans une intégrale et le long du flot $(\mu_s)_{s\in[0,T]}$ qui appartient à $\mathscr{P}(\mathbb{R}^d)$ pour presque tout $s \in [0,T]$. Cependant, on suppose que u est continue sur $\mathcal{P}_2(\mathbb{R}^d)$ car le flot $s \in [0,T] \mapsto \mu_s \in \mathcal{P}_2(\mathbb{R}^d)$ est continu mais μ_t n'a aucune raison d'appartenir à $\mathscr{P}(\mathbb{R}^d)$ pour tout $t \in [0,T]$ et encore moins d'être continu pour $d_{\mathscr{P}}$.

L'exemple suivant illustre le théorème pour une fonctionnelle non-linéaire en la mesure.

Exemple 4.13. Soient $F \in \mathcal{C}^1(\mathbb{R};\mathbb{R})$ et $g \in \mathcal{C}^0(\mathbb{R}^d;\mathbb{R})$ admettant une dérivée distributionnelle ∇g appartenant à $(W^{1,k}(\mathbb{R}^d))^d$ pour un certain $k \geq d+1$. Alors

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \mu & \mapsto F\left(\int_{\mathbb{R}^d} g\,d\mu\right) \end{cases}$$

appartient à $\mathcal{W}_1(\mathbb{R}^d)$ pour $d_{\mathscr{P}} = W_2$.

4.4 Stratégie de preuve

On commence par montrer que les termes qui apparaissent dans la formule d'Itô (4.9) sont bien définis. Pour cela, on utilise le caractère borné de b et σ, ainsi que les propriétés d'intégrabilité de la densité $p(t,\cdot)$ de μ_t (voir la Proposition 4.3) et enfin le contrôle de $\left\|\partial_v \frac{\delta}{\delta m}u(\mu)\right\|_{W^{1,k}(\mathbb{R}^d)}$ fait dans les hypothèses. On régularise ensuite u par convolution en posant, pour $n \geq 1$,

$$u^n : \mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto u(\mu * \rho_n).$$

Cela régularise la dérivée plate spatialement puisqu'on a, pour tout $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ et $v \in \mathbb{R}^d$,

$$\frac{\delta}{\delta m} u^n(\mu)(v) = \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(\mu * \rho_n)(x) \rho_n(v-x)\, dx = \frac{\delta}{\delta m} u(\mu * \rho_n) * \rho_n(v).$$

On vérifie qu'on peut appliquer la formule d'Itô pour un flot de mesures classique rappelée au Théorème 2.28. On obtient que, pour tout $n \geq 1$ et $t \in [0,T]$,

$$u^n(\mu_t) = u^n(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} u^n(\mu_s)(X_s) \cdot b_s\right) ds + \frac{1}{2}\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta}{\delta m} u^n(\mu_s)(X_s) \cdot a_s\right) ds.$$

Il reste ensuite à laisser tendre n vers $+\infty$ en utilisant notamment la continuité de

$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \partial_v \frac{\delta}{\delta m} u(\mu) \in W^{1,k}(\mathbb{R}^d).$$

Le contrôle de la croissance par rapport à μ de $\left\| \partial_v \frac{\delta}{\delta m} u(\mu) \right\|_{W^{1,k}(\mathbb{R}^d)}$ permet de justifier l'utilisation du théorème de convergence dominée.

4.5 Extension au cas où la fonction dépend aussi du temps et de l'espace

Dans [Cav21], on démontre aussi la formule d'Itô-Krylov pour une fonction $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$. On introduit l'espace de fonctions généralisant $\mathcal{W}_1(\mathbb{R}^d)$.

Définition 4.14. On définit $\mathcal{W}_2(\mathbb{R}^d)$ comme l'espace des fonctions continues $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ vérifiant les propriétés suivantes pour une distance $d_{\mathscr{P}}$ qui satisfait les Hypothèses **(H1)** et **(H2)** de la Définition 4.5.

(1) Pour tout $(x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $u(\cdot, x, \mu) \in \mathcal{C}^1$ et $\partial_t u$ est continue sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$.

(2) Il existe $k_1 \geq d+1$ tel que pour tout $(t,\mu) \in [0,T] \times \mathscr{P}(\mathbb{R}^d)$, $u(t,\cdot,\mu) \in W^{2,k_1}_{\mathrm{loc}}(\mathbb{R}^d)$ et pour tout $t \in [0,T]$ et $R > 0$

$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \partial_x u(t,\cdot,\mu) \in \left(W^{1,k_1}(B_R)\right)^d,$$

est continue et $\partial_x u$ et $\partial_x^2 u$ sont mesurables par rapport à $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d)$.

(3) Pour tout $(t,x) \in [0,T] \times \mathbb{R}^d$, $u(t,x,\cdot)$ admet une dérivée plate $\frac{\delta}{\delta m} u(t,x,\cdot)(\cdot)$ qui est continue sur $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, et tel que pour tout compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ et $t \in [0,T]$, il existe $C > 0$ tel que pour tout $v \in \mathbb{R}^d$

$$\sup_{(x,\mu) \in \mathcal{K}} \left| \frac{\delta}{\delta m} u(t,x,\mu)(v) \right| dx \leq C(1 + |v|^2).$$

(4) Il existe $k_2 \geq 2d$ tel que pour tout $(t,\mu) \in [0,T] \times \mathscr{P}(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(t,\cdot,\mu)(\cdot)$ admet une dérivée distributionnelle par rapport à v d'ordre 1 et 2 et telle que pour tout $t \in [0,T]$ et $R > 0$

$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \left(\partial_v \frac{\delta}{\delta m} u(t,\cdot,\mu)(\cdot)\, ,\, \partial_v^2 \frac{\delta}{\delta m} u(t,\cdot,\mu)(\cdot)\right) \in (L^{k_2}(B_R \times \mathbb{R}^d))^d \times (L^{k_2}(B_R \times \mathbb{R}^d))^{d \times d},$$

est continue et mesurable par rapport à $(t,x,\mu,v) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d$.

(5) Il existe $\alpha_1, \alpha_2 \in \mathbb{N}$ avec $k_1 \geq (2\alpha_1 + 1)d$, $k_2 \geq (\alpha_2 + 2)d$ tel que pour tout compact $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ et pour tout $R > 0$, il existe $C_{\mathcal{K},R} > 0$ tel que pour tout $\mu \in \mathcal{K} \cap \mathscr{P}(\mathbb{R}^d)$

$$\begin{cases} \sup_{t \leq T} \left\{ \|\partial_x u(t, \cdot, \mu)\|_{L^{k_1}(B_R)} + \left\|\partial_x^2 u(t, \cdot, \mu)\right\|_{L^{k_1}(B_R)} \right\} \leq C_{\mathcal{K},R} \left(1 + \left\|\frac{d\mu}{dx}\right\|_{L^{k_1'}(\mathbb{R}^d)}^{\alpha_1}\right) \\ \sup_{t \leq T} \left\{ \left\|\partial_v \frac{\delta}{\delta m} u(t, \cdot, \mu)(\cdot)\right\|_{L^{k_2}(B_R \times \mathbb{R}^d)} + \left\|\partial_v^2 \frac{\delta}{\delta m} u(t, \cdot, \mu)(\cdot)\right\|_{L^{k_2}(B_R \times \mathbb{R}^d)} \right\} \leq C_{\mathcal{K},R} \left(1 + \left\|\frac{d\mu}{dx}\right\|_{L^{k_2'}(\mathbb{R}^d)}^{\alpha_2}\right). \end{cases}$$

Remarque 4.15. Le contrôle de l'hypothèse (3) est assez naturel. Si la borne supérieure était prise seulement sur un compact de $\mathcal{P}_2(\mathbb{R}^d)$, ce serait assuré par la définition de la dérivée plate. Or nous avons besoin de contrôler également $\frac{\delta}{\delta m} u$ localement uniformément par rapport à la variable $x \in \mathbb{R}^d$ du fait de la convolution en espace et en mesure qui est utilisée naturellement dans la preuve. Les hypothèses (2), (4) et (5) sont des généralisations de celles présentes dans la Définition 4.7, adaptées au cadre présent. Dans l'hypothèse (5), la condition sur k_2 et α_2 change un peu en comparaison à la condition analogue dans la Définition 4.7. C'est essentiellement dû au fait qu'on à affaire à des fonctions définies sur $\mathbb{R}^{2d}$ au lieu de $\mathbb{R}^d$ à cause de la variable spatiale de la fonction u.

On énonce maintenant l'extension de la formule d'Itô-Krylov. On considère toujours le flot de lois marginales $(\mu_t)_{t \in [0,T]}$ du processus X défini par (4.1). Soit également $(\eta_s)_{s \in [0,T]}$ et $(\gamma_s)_{s \in [0,T]}$ deux processus progressivement mesurables à valeurs respectivement dans $\mathbb{R}^d$ et $\mathbb{R}^{d \times d_1}$ et vérifiant les Hypothèses **(A)** et **(B)**. On pose, pour $t \leq T$,

$$\xi_t := \xi_0 + \int_0^t \eta_s \, ds + \int_0^t \gamma_s \, dB_s,$$

où ξ_0 est une variable aléatoire $\mathcal{F}_0$-mesurable à valeurs dans $\mathbb{R}^d$.

Théorème 4.16 (Extension de la formule d'Itô-Krylov)**.** *Soit u une fonction dans $\mathcal{W}_2(\mathbb{R}^d)$. On a presque sûrement pour tout $t \in [0, T]$*

$$\begin{aligned} u(t, \xi_t, \mu_t) = u(0, \xi_0, \mu_0) &+ \int_0^t (\partial_t u(s, \xi_s, \mu_s) + \partial_x u(s, \xi_s, \mu_s) \cdot \eta_s) \, ds + \frac{1}{2} \int_0^t \partial_x^2 u(s, \xi_s, \mu_s) \cdot \gamma_s \gamma_s^* \, ds \\ &+ \int_0^t \overline{\mathbb{E}} \left(\partial_v \frac{\delta}{\delta m} u(s, \xi_s, \mu_s)(\overline{X_s}) \cdot \overline{b_s} \right) ds + \frac{1}{2} \int_0^t \overline{\mathbb{E}} \left(\partial_v^2 \frac{\delta}{\delta m} u(s, \xi_s, \mu_s)(\overline{X_s}) \cdot \overline{a_s} \right) ds \\ &+ \int_0^t \partial_x u(s, \xi_s, \mu_s) \cdot (\gamma_s \, dB_s), \end{aligned} \tag{4.8}$$

où $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ est une copie de $(\Omega, \mathcal{F}, \mathbb{P})$ et $(\overline{X}, \overline{b}, \overline{\sigma})$ une copie indépendante de (X, b, σ).

La preuve de ce résultat repose sur la même idée que celle du Théorème 4.9. Elle est cependant plus technique.

4.6 Perspectives

La perspective présentée dans cette section vient d'être abordée en collaboration avec Paul-Éric Chaudru de Raynal et Stéphane Menozzi.

Le but est d'exhiber des taux de convergence pour la propagation du chaos du système de particules associé à l'EDS de McKean-Vlasov suivante

$$\begin{cases} dX_t = \displaystyle\int_{\mathbb{R}^d} b(t, X_t - y)\, d\mu_t(y)\, dt + dB_t, \quad t \in [0,T], \\ \mu_t := [X_t], \end{cases} \tag{4.9}$$

où B est un mouvement brownien et $b \in L^q((0,T); L^p(\mathbb{R}^d))$ avec $p, q \in [1, \infty)$ vérifiant la condition de Krylov-Röckner $\frac{d}{p} + \frac{2}{q} < 1$. La propagation du chaos qualitative, i.e. sans taux de convergence, a déjà été prouvée dans [Tom20] et également dans [HHMT20], où un principe de grandes déviations est établi. Pour un résultat de propagation du chaos quantitative dans le cas d'interactions singulières, on peut mentionner [Han22], où la propagation du chaos est quantifiée en terme d'entropie. Notre objectif est d'obtenir des estimations de propagation du chaos analogues à celles du Théorème 3.23. Pour ce faire, on peut appliquer les résultats de [CdRF21] avec une dérive b régularisée par convolution. Ensuite, le but est de prouver des estimations sur la solution de l'EDP de Kolmogorov rétrograde associée à (4.9) avec une dérive régularisée, qui doivent être uniformes par rapport au paramètre de régularisation. Cela permettrait de quantifier la propagation du chaos d'une manière différente qu'en terme d'entropie et on n'aurait besoin d'aucune hypothèse structurelle sur la dérive b contrairement à [Han22], où il est supposé que b ne dépend pas du temps et que $\nabla \cdot b = 0$. On aurait également besoin de l'inégalité de Krylov pour gérer l'approximation par une dérive régulière.

Comportement asymptotique d'un système cinétique inhomogène en temps dans un potentiel quadratique

Dans ce chapitre, on présente le travail [CL23], réalisé en collaboration avec Émeline Luirard.

5.1 Présentation générale du contexte

On commence par présenter le contexte général dans lequel s'inscrit l'article [CL23], ainsi que les travaux auxquels il fait suite.

On s'intéresse au comportement asymptotique d'un système modélisant une particule avec vitesse $V \in \mathbb{R}$ et position $X \in \mathbb{R}$. La particule évolue dans le potentiel quadratique $\mathcal{U} : x \mapsto \frac{x^2}{2}$, avec une force de frottement inhomogène en temps b, et est soumise à une force aléatoire représentant l'interaction avec l'environnement. La dynamique de la particule est décrite par le système Hamiltonien amorti suivant, perturbé par un processus α-stable L avec $\alpha \in (1, 2]$,

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}L_t - b(t, V_t)\,\mathrm{d}t - \nabla\mathcal{U}(X_t)\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0). \end{cases} \tag{5.1}$$

Les systèmes Hamiltoniens stochastiques, c'est-à-dire lorsque $b = 0$, ont été largement étudiés. Comprendre leur comportement asymptotique est l'un des enjeux principaux. Le processus Hamiltonien, ou énergie, associé au système (5.1) lorsque $b = 0$ est défini par $H_t := \frac{1}{2}|V_t|^2 + \mathcal{U}(X_t)$. Par exemple, le comportement en temps long du processus Hamiltonien correctement changé d'échelle a été étudié dans [AK94]. Le cas d'un système aléatoire Hamiltonien amorti a été abordé dans [Wu01] (voir aussi les références dans cet article).

Le comportement en temps long d'une particule libre, c'est-à-dire sans potentiel $\mathcal{U}$, a également déjà été étudié. On mentionne notamment [GO13], [FT21], [GL21a], [GL21b]. Même dans le cadre homogène en temps, divers comportements asymptotiques peuvent être observés. Par exemple, lorsque L est un mouvement brownien, qu'on notera B dans la suite, l'EDS suivante de type Langevin a été étudiée dans [FT21]

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}B_t - \dfrac{\rho}{2}\dfrac{V_t}{1+V_t^2}\,\mathrm{d}t \\ \mathrm{d}X_t = V_t\,\mathrm{d}t. \end{cases}$$

Dans ce système, la force de frottement se comporte comme $-\dfrac{\rho}{v}$, lorsque $|v|$ est grand, ce qui implique heuristiquement que le processus vitesse se comporte, loin de zéro, comme un processus (signé) de Bessel de dimension $1-\rho$. Différents comportements asymptotiques pour le processus position apparaissent alors en fonction de la valeur de ρ. Plus précisément, lorsque $\rho \geq 5$, en utilisant un changement d'échelle approprié, les auteurs prouvent que le processus position se comporte asymptotiquement comme un mouvement brownien. Un processus α-stable apparaît comme dynamique limite pour la position lorsque $\rho \in [1,5)$. L'indice d'auto-similarité α est alors une fonction de ρ qui interpole les puissances des changements d'échelle $\frac{1}{2}$ (mouvement brownien) et $\frac{3}{2}$ (processus de Bessel intégré). Ce dernier processus apparaît asymptotiquement lorsque $\rho \in (0,1)$. Cependant, les outils utilisés dans ce travail, notamment la notion de mesure invariante, fonction d'échelle et mesure vitesse sont limités à des coefficients homogènes en temps.

Dans [GO13], [GL21a] et [GL21b], toujours dans le cas où $\mathcal{U}=0$, la force de frottement b a une forme particulière et dépend du temps. Plus précisément, le système considéré est le suivant

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}B_t - \rho\dfrac{\operatorname{sgn}(V_t)\,|V_t|^\gamma}{t^\beta}\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t. \end{cases} \tag{5.2}$$

La force de frottement est inhomogène en temps et dépend de paramètres positifs β, γ et ρ. Lorsque la particule se déplace lentement, la mécanique classique assure que la force de frottement est linéaire, i.e. $\gamma=1$. Lorsque la particule évolue dans un régime turbulent, la mécanique des fluides assure que $\gamma=2$. C'est pourquoi dans un cadre général, on suppose que la force de frottement s'écrit comme $v \mapsto -\rho\operatorname{sgn}(v)\,|v|^\gamma$. De plus, l'intensité de la force de frottement peut dépendre du temps à travers le coefficient de friction $t \mapsto \rho_t$. Pour une particule évoluant dans un fluide, c'est le cas par exemple si la viscosité du fluide ou encore la géométrie de la particule évoluent avec le temps. C'est pour cette raison qu'une dépendance en temps est ajoutée à la force de frottement dans [GO13], [GL21a] et [GL21b]. Dans ces travaux, il est supposé que $\rho_t = \frac{\rho}{t^\beta}$. L'objectif principal derrière l'étude de ce modèle est de comprendre la compétition entre la force de frottement, qui tend à immobiliser la particule mais de moins en moins efficacement au fur et à mesure que le temps passe, et le bruit qui la perturbe constamment. La Proposition 1.12 assure, par propriété d'auto-similarité de L, que $\mathbb{E}\left[|L_t|\right]$ est proportionnelle à $t^{\frac{1}{\alpha}}$. Cela montre que L_t agit à une échelle typique de $t^{\frac{1}{\alpha}}$ et donc que, lorsque α décroit, cela perturbe la vitesse avec des valeurs typiques plus élevées. L'intérêt des travaux mentionnés précédemment est d'étudier le comportement asymptotique du système en fonction des différents paramètres et sous les prisme de la compétition entre les deux actions opposées mentionnées précédemment.

Donnons deux exemples importants dans le cas brownien avant d'expliquer plus en détails les résultats de [GO13, GL21a, GL21b]. Lorsque $\beta=0$, le coefficient de frottement ne décroît pas avec le temps. Par un argument d'ergodicité, le processus vitesse converge vers son unique mesure invariante et donc, le processus position changé d'échelle $(\varepsilon^{\frac{1}{2}}X_{t/\varepsilon})_{t>0}$ se comporte comme un mouvement brownien lorsque ε tend vers 0. Lorsque "$\beta=+\infty$", i.e. lorsqu'il n'y a pas de force de frottement, le processus vitesse-position changé d'échelle $(\varepsilon^{\frac{1}{2}} V_{t/\varepsilon}, \varepsilon^{\frac{3}{2}} X_{t/\varepsilon})_{t>0}$ converge en loi vers $\left(B_t, \int_0^t B_s \, \mathrm{d}s\right)_{t>0}$. Lorsque $\beta > 0$, la force de frottement disparaît asymptotiquement : elle ralentit le système de moins en moins efficacement et on s'attend à une transition entre les deux cas extrêmes mentionnés précédemment, tant sur le processus limite que sur le changement d'échelle utilisé.

Dans [GO13], les auteurs étudient la convergence en loi, lorsque t tend vers l'infini, de $r_t V_t$, pour une certaine vitesse de convergence r_t. Dans [GL21a], les résultats obtenus dans [GO13] sont étendus au processus vitesse-position. Plus précisément, les auteurs étudient la convergence en loi du processus changé d'échelle $(r_{\varepsilon,V} V_{t/\varepsilon}, r_{\varepsilon,X} X_{t/\varepsilon})_t$ pour deux taux de convergence $r_{\varepsilon,V}$ et $r_{\varepsilon,X}$. Ces résultats ont été enfin généralisés dans [GL21b] pour un bruit directeur α-stable. Pour être plus précis, trois régimes asymptotiques sont mis en lumière dans ces travaux en fonction des paramètres β, γ et α, l'indice d'autosimilarité de L.

- Lorsque la force de frottement b décroit suffisamment vite vers 0 avec le temps, c'est-à-dire lorsque β est assez grand relativement à γ et α, le processus vitesse-position changé d'échelle se comporte comme s'il n'y avait pas de frottements et converge donc en loi vers le processus de Kolmogorov $(L, \int_0^\cdot L)$, comme dans le cas extrême "$\beta = +\infty$" mentionné précédemment dans le cas brownien et avec le même changement d'échelle $(r_{\varepsilon,V}, r_{\varepsilon,X}) := (\varepsilon^{\frac{1}{\alpha}}, \varepsilon^{1+\frac{1}{\alpha}})$.
- Lorsque les paramètres β, γ et α s'équilibrent, le processus limite est toujours de la forme cinétique $(\mathcal{V}, \int_0^\cdot \mathcal{V})$, mais le processus $\mathcal{V}$ n'a plus la même loi que le bruit et est ergodique.
- Lorsque la force de frottement ne décroit pas suffisamment vite pour être négligée, le processus limite n'est plus sous forme cinétique. Le processus vitesse changé d'échelle converge en lois finie-dimensionnelles vers un bruit blanc. Dans ce cas, le comportement asymptotique est en quelque sorte une interpolation entre les deux cas extrêmes $\beta = 0$ et "$\beta = +\infty$", ce qui est expliqué par la décroissance plus lente de l'intensité de la force de frottement avec le temps.

5.2 Système étudié et comportement en temps long

Dans [CL23], on s'intéresse au comportement asymptotique de la solution du système cinétique d'EDS suivant, défini sur l'intervalle de temps $[t_0, +\infty)$, où $t_0 > 0$

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}L_t - \mathrm{sgn}(V_t)\dfrac{|V_t|^\gamma}{t^\beta}\,\mathrm{d}t - X_t\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0). \end{cases} \tag{SKE}$$

La force de frottement est inhomogène en temps et dépend de deux paramètres positifs β et γ. Le processus directeur L est soit un mouvement brownien, soit un processus α-stable symétrique, i.e. rotationnellement invariant en dimension 1, avec $\alpha \in (1, 2)$. Il s'agit du même modèle que (5.2), avec une force de rappel supplémentaire qui provient du potentiel quadratique. Plus précisément, notre but est d'étudier le comportement, lorsque $\varepsilon \to 0$, du processus position-vitesse changé d'échelle $(Z_t^{(\varepsilon)})_t := \left(r_\varepsilon (X_{t/\varepsilon}, V_{t/\varepsilon})^T\right)_{t>0}$, pour un taux de convergence r_ε approprié. L'une de nos motivations est d'analyser comment la présence du potentiel quadratique influence les résultats obtenus dans [GL21a, GL21b] à travers l'effet confinant qu'a le potentiel quadratique sur la position X. En effet, l'effet confinant se fait sur la position de la particule et ne disparaît pas asymptotiquement contrairement à la force de frottement. Il s'agit ici d'étudier une compétition entre la force de frottement et le potentiel quadratique, qui freinent et confinent la particule, contre le bruit qui perturbe sa vitesse.

Le système sous-jacent sans bruit et force de frottement est l'oscillateur harmonique classique

$$\begin{cases} v_t' = -x_t, \\ x_t' = v_t. \end{cases}$$

Le comportement intrinsèquement oscillatoire induit par le potentiel quadratique empêche le processus changé d'échelle $Z^{(\varepsilon)}$ de converger en loi en tant que processus. Cependant, on parvient à prouver que ses lois marginales de dimension 1 convergent. Afin d'obtenir la convergence de tout le processus, l'idée cruciale est de retirer, en un certain sens, les oscillations présentes dans le système. Plus précisément, on pose $Y_t := \Theta_t^{-1}(X_t, V_t)^T$, où Θ_t est la rotation sur $\mathbb{R}^2$ d'angle $-t$, définie par

$$\Theta_t := \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

On énonce maintenant les principaux résultats démontrés dans [CL23]. Le théorème suivant traite le cas où le bruit L est un mouvement brownien ($\alpha = 2$). Il porte sur la convergence en loi dans l'espace des fonctions continues $\mathcal{C}((0, +\infty), \mathbb{R}^2)$ muni de la topologie associée à la convergence uniforme sur tout compact.

Théorème 5.1 (Cas brownien, i.e. $\alpha = 2$)**.** *On définit $q := \frac{\beta}{\gamma+1}$, $r_\varepsilon := \varepsilon^{q \wedge \frac{1}{2}}$ et on pose $(Y_t^{(\varepsilon)})_{t \geq \varepsilon t_0} := \left(r_\varepsilon \Theta_{t/\varepsilon}^{-1}(X_{t/\varepsilon}, V_{t/\varepsilon})^T\right)_{t \geq \varepsilon t_0}$. Soit $\mathcal{B}$ un mouvement brownien standard sur $\mathbb{R}^2$.*

- *(Régime sur-critique i.e. $2q > 1$). Le processus $Y^{(\varepsilon)}$ converge en loi vers $\left(\mathcal{B}_{\frac{t}{2}}\right)_{t>0}$.*
- *(Régime critique i.e. $2q = 1$). Supposons qu'on est dans le cas linéaire $\gamma = 1$. Le processus $Y^{(\varepsilon)}$ converge en loi vers $\left(\frac{1}{\sqrt{2t}} \int_0^t \sqrt{s}\, \mathrm{d}\mathcal{B}_s\right)_{t>0}$, qui est le processus Gaussien centré dont le noyau de covariance est donné par $K(s,t) := \frac{(s \wedge t)^2}{4\sqrt{st}} I_2$.*
- *(Régime sous-critique i.e. $2q < 1$). Supposons qu'on est dans le cas linéaire $\gamma = 1$ et que $\beta \in \left(\frac{1}{2}, 1\right)$. Le processus $Y^{(\varepsilon)}$ converge en lois finie-dimensionnelles vers le processus Gaussien centré dont le noyau de covariance est donné par $K(s,t) := \frac{1}{2} s^\beta \mathbf{1}_{\{s=t\}} I_2$.*

Supposons maintenant que $\alpha \in (1, 2)$. On note ψ l'exposant caractéristique (symbole) du processus α-stable symétrique L. Il découle de [Sat99, Théorème 14.15 p. 86] qu'il existe $a > 0$ tel que pour tout $\xi \in \mathbb{R}$,

$$\psi(\xi) = -a|\xi|^\alpha. \tag{5.3}$$

Dans le théorème suivant, la convergence en loi se fait dans l'espace des fonctions càdlàg $\mathcal{D}((0, +\infty), \mathbb{R}^2)$, c'est-à-dire les fonctions définies sur $(0, +\infty)$ qui sont càdlàg sur tout segment contenu dans $(0, +\infty)$, qui est muni de la distance de Shorokhod.

Théorème 5.2 (Cas stable, i.e. $\alpha \in (1, 2)$)**.** *Supposons que $\gamma \in (0, \alpha)$. On définit $q := \frac{\beta}{\gamma+\alpha-1}$, $r_\varepsilon := \varepsilon^{q \wedge \frac{1}{\alpha}}$ et on pose $(Y_t^{(\varepsilon)})_{t \geq \varepsilon t_0} := \left(r_\varepsilon \Theta_{t/\varepsilon}^{-1}(X_{t/\varepsilon}, V_{t/\varepsilon})^T\right)_{t \geq \varepsilon t_0}$. Soit $\mathcal{L}$ le processus α-stable rotationnellement* *invariant sur $\mathbb{R}^2$, dont l'exposant caractéristique est donné par*

$$\xi \in \mathbb{R}^2 \mapsto -\widetilde{C} \, \|\xi\|^\alpha \,, \quad \textit{avec } \widetilde{C} := \frac{a}{2\pi} \int_0^{2\pi} |\cos(x)|^\alpha \, \mathrm{d}x.$$

- *(Régime sur-critique i.e. $\alpha q > 1$). Le processus $Y^{(\varepsilon)}$ converge en loi vers $(\mathcal{L}_t)_{t>0}$.*
- *(Régime critique i.e. $\alpha q = 1$). Supposons qu'on est dans le cas linéaire $\gamma = 1$. Le processus $Y^{(\varepsilon)}$ converge en loi vers le processus de type-Lévy $\left(\frac{1}{\sqrt{t}} \int_0^t \sqrt{s} \, \mathrm{d}\mathcal{L}_s\right)_{t>0}$.*
- *(Régime sous-critique i.e. $\alpha q < 1$). Supposons qu'on est dans le cas linéaire $\gamma = 1$ et que $\beta \in \left(\frac{1}{2}, 1\right)$. Alors, pour tout $(t_1, \cdots, t_d) \in (0, +\infty)^d$, $\left(Y_{t_1}^{(\varepsilon)}, \cdots, Y_{t_d}^{(\varepsilon)}\right)$ converge en loi vers la mesure produit $\mu_{t_1} \otimes \cdots \otimes \mu_{t_d}$, où μ_t est la mesure de probabilité dont la fonction caractéristique est donnée par*

$$\xi \in \mathbb{R}^2 \mapsto \exp\left(-\frac{2}{\alpha} \widetilde{C} \, \|\xi\|^\alpha \, t^\beta\right).$$

Remarque 5.3. La symétrie de L est seulement utilisée pour assurer le caractère bien posé de (SKE) lorsque $\gamma < 1$.

Remarque 5.4. Le processus Hamiltonien associé à notre système est donné par

$$H_t := \frac{1}{2} |V_t|^2 + \frac{1}{2} |X_t|^2 = \frac{1}{2} \|Z_t\|^2 = \frac{1}{2} \|Y_t\|^2 \,.$$

En combinant les résultats des théorèmes précédents avec le *continuous mapping theorem*, on déduit la convergence du processus Hamiltonien changé d'échelle $(H_t^{(\varepsilon)})_{t>0} := (r_\varepsilon^2 H_{t/\varepsilon})_{t>0}$ lorsque $\varepsilon \to 0$, soit en tant que processus dans les cas critique et sur-critique, soit en lois finie-dimensionnelles dans le cas sous-critique.

Par exemple dans le cas brownien sur-critique, le processus Hamiltonien limite est donné par $(H_t^0)_{t\geq 0} := (\frac{1}{2}\|\mathcal{B}_{\frac{t}{2}}\|^2)_{t\geq 0}$. C'est un carré de Bessel, qui est solution de

$$\mathrm{d}H_t^0 = \sqrt{H_t^0} \, \mathrm{d}B_t + \frac{1}{2} \, \mathrm{d}t, \quad H_0^0 = 0,$$

où B est un mouvement brownien standard sur $\mathbb{R}$. On remarque qu'on retrouve le même processus Hamiltonien limite que dans [AK94, Théorème 2.1] pour le système non-amorti. L'explication est la suivante : si la force de frottement décroît suffisamment vite avec le temps, c'est-à-dire si β est assez grand, alors le processus Hamiltonien changé d'échelle se comporte asymptotiquement comme si il n'y avait pas de frottements.

Comme corollaire de nos deux théorèmes, on obtient la convergence en loi, lorsque t tend vers l'infini, de $t^{-q\wedge\frac{1}{\alpha}}(X_t, V_t)^T$.

Corollaire 5.5. *On pose $(Z_t^{(\varepsilon)})_{t\geq\varepsilon t_0} := (r_\varepsilon(X_{t/\varepsilon}, V_{t/\varepsilon})^T)_{t\geq\varepsilon t_0}$, où $r_\varepsilon := \varepsilon^{q\wedge\frac{1}{\alpha}}$. Le processus $Z^{(\varepsilon)}$ ne converge pas en loi, cependant sous les mêmes hypothèses que dans les Théorèmes 5.1 et 5.2, on déduit la convergence en loi $r_{1/t}(X_t, V_t)^T$ lorsque $t \to +\infty$ vers des limites explicites.*

Dans le cas brownien, la limite est soit $\mathcal{N}(0, \frac{1}{2}I_2)$ dans les régimes sur-critique et sous-critique, ou $\mathcal{N}(0, \frac{1}{4}I_2)$ dans le régime critique.

Dans le cas stable, en gardant les mêmes notations que dans le Théorème 5.2, la fonction caractéristique de la limite est donnée, pour tout $\xi \in \mathbb{R}^2$, par

- $\exp\left(-\widetilde{C}\,\|\xi\|^\alpha\right)$ *dans le régime sur-critique,*
- $\exp\left(-\left(1+\frac{\alpha}{2}\right)^{-1}\widetilde{C}\,\|\xi\|^\alpha\right)$ *dans le régime critique,*
- $\exp\left(-\frac{2}{\alpha}\widetilde{C}\,\|\xi\|^\alpha\right)$ *dans le régime sous-critique.*

Ici, le changement entre les différents régimes apparaît à travers les différents paramètres d'échelle pour les lois limites. Mentionnons également que dans le cas brownien, la position et la vitesse deviennent indépendantes en temps grand puisque la matrice de covariance de la loi limite gaussienne est diagonale. Cependant, ce n'est plus vrai dans le cas α-stable avec $\alpha \in (1,2)$. En effet, la loi limite est une loi stable rotationnellement invariante sur $\mathbb{R}^2$, qui ne peut pas avoir des coordonnées indépendantes.

De façon similaire à [GL21a, GL21b], on met en lumière trois régimes asymptotiques déterminés par la même relation entre les différents paramètres du problème. Cependant, le taux de convergence de la position X est différent de celui prouvé dans [GL21a, GL21b], lorsque $\mathcal{U} = 0$. L'explication heuristique est que la présence du potentiel quadratique permet au bruit de propager plus facilement de l'équation sur la vitesse à celle sur la position, qui est donc plus diffusive en quelque sorte. On pourra consulter [FFPV17] pour plus de détails à ce sujet. Remarquons également que le processus position tend à croître moins vite dans notre cas que dans le cas $\mathcal{U} = 0$. Par exemple, dans le régime sur-critique brownien, X_t se comporte en loi asymptotiquement comme $\mathcal{N}(0, \frac{t}{2})$ dans notre cas, mais comme $\mathcal{N}(0, \frac{t^3}{3})$ dans le cas $\mathcal{U} = 0$ traité dans [GL21a]. La différence des variances s'interprète comme le fait que le potentiel quadratique confine la particule à travers la force de rappel qu'il induit.

5.3 Stratégie de preuve

Dans notre modèle, les équations sur la vitesse et la position ne peuvent pas être traitées de façon indépendante comme dans [GL21a, GL21b] puisqu'elles sont couplées. L'idée est d'écrire le système sous forme vectorielle, et d'utiliser la méthode de variation de la constante qui permet de se ramener à un système sans potentiel en dimension 2.

Régime sur-critique. La formule d'Itô pour un produit assure que, pour $t \geq t_0$,

$$\Theta_t^{-1}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = \Theta_{t_0}^{-1}\begin{pmatrix} X_{t_0} \\ V_{t_0} \end{pmatrix} + \int_{t_0}^t \Theta_s^{-1}\begin{pmatrix} 0 \\ \operatorname{sgn}(V_s)\frac{|V_s|^\gamma}{s^\beta} \end{pmatrix} \mathrm{d}s + \int_{t_0}^t \begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix} \mathrm{d}L_s, \tag{5.4}$$

où on rappelle que Θ_t est la rotation d'angle $-t$. On a alors, pour tout $t \geq \varepsilon t_0$,

$$Y_t^{(\varepsilon)} = \varepsilon^{\frac{1}{\alpha}}\Theta_{t_0}^{-1}\begin{pmatrix} x_0 \\ v_0 \end{pmatrix} + \varepsilon^{\frac{1}{\alpha}}\int_{t_0}^{t/\varepsilon} \Theta_s^{-1}\frac{|V_s|^\gamma}{s^\beta}\,\mathrm{d}s + M_t^{(\varepsilon)}, \tag{5.5}$$

où

$$M_t^{(\varepsilon)} := \varepsilon^{\frac{1}{\alpha}}\int_{t_0}^{t/\varepsilon} \begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix} \mathrm{d}L_s.$$

Pour traiter la convergence en loi du processus $M^{(\varepsilon)}$, lorsque $\varepsilon \to 0$ et dans le cas où le bruit est brownien, on tire profit de la théorie des processus Gaussiens. La convergence est donc caractérisée par la moyenne et la fonction covariance. Dans le cas où le bruit n'est pas brownien, on doit étudier la convergence en loi d'une intégrale de Wiener-Lévy, c'est-à-dire l'intégrale d'une fonction déterministe contre le processus stable L. Le point clé est que de tels processus ont des accroissements indépendants même si ce ne sont pas forcément des processus de Lévy. On utilise alors la fonction caractéristique pour étudier la convergence des lois finie-dimensionnelles.

Les deux premiers termes du membre de droite de (5.5) peuvent être négligés car tendent vers 0 dans $L^1(\Omega, \mathbb{P})$ grâce à l'estimation de moments suivante qu'on démontre.

Proposition 5.6. *Pour tout $\kappa \geq 0$ dans le cas brownien ou pour tout $\kappa \in [0, \alpha)$ si $\alpha < 2$, il existe une constante $C_{\kappa,t_0} > 0$ telle que*

$$\forall t \geq t_0, \ \mathbb{E}\left[\left\|(V_t, X_t)^T\right\|^\kappa\right] \leq C_{\kappa,t_0} t^{\frac{\kappa}{\alpha}}. \tag{5.6}$$

C'est à cette étape qu'apparaît la distinction entre le régime sur-critique et les deux autres. En effet, dans les régimes critique et sous-critique, le second terme du membre de droite de (5.5) ne tend pas vers 0 dans $L^1(\Omega, \mathbb{P})$, lorsque $\varepsilon \to 0$. De plus, la tension du processus changé d'échelle $(Y_t^{(\varepsilon)})_{t>0}$, lorsque $\varepsilon \to 0$, est prouvée grâce à la Proposition 5.6 dans les régimes critiques et sur-critiques. Remarquons également que l'estimation (5.6) est la même que celle prouvée seulement pour le processus vitesse dans [GL21a, GL21b].

Régimes critique et sous-critique. Dans ces deux régimes, on se restreint au cas d'une force de frottement linéaire en espace, i.e. $\gamma = 1$. La méthode de changement de temps utilisée dans [GL21a] et [GO13] ne semble pas s'adapter ici car les équations sur la vitesse et la position sont couplées. Le cas linéaire nous permet d'utiliser l'Équation Différentielle Ordinaire (EDO) linéaire sous-jacente, c'est-à-dire lorsqu'il n'y a pas de bruit, qui est

$$x''(t) + \frac{x'(t)}{t^\beta} + x(t) = 0, \quad t \geq t_0. \tag{5.7}$$

On établit notamment un développement asymptotique à l'ordre 1 des solutions de cette EDO et de leur dérivée, lorsque $\beta > \frac{1}{2}$. C'est la seule raison qui induit cette restriction sur β dans nos résultats. De plus, on considère la résolvante $(R_t)_{t \geq t_0}$ de cette EDO, qui est solution, à valeurs dans les matrices de taille 2×2, de

$$R_t' = \begin{pmatrix} 0 & 1 \\ -1 & -\frac{1}{t^\beta} \end{pmatrix} R_t, \quad t \geq t_0,$$

avec une donnée initiale R_{t_0} au temps t_0 qu'on ne précise pas. Soit f la fonction définie, pour $t > 0$, par

$$f(t) := \begin{cases} \frac{1}{\sqrt{t}} & \text{si } \beta = 1, \\ \exp\left(-\frac{t^{1-\beta}}{2(1-\beta)}\right) & \text{sinon.} \end{cases}$$

Grâce aux développements asymptotiques mentionnés précédemment, on montre que lorsque t tend vers l'infini, R_t se comporte comme $f(t)\Theta_t$, pour certain choix de donnée initiale R_{t_0}. On écrit alors, grâce à la formule d'Itô, pour $t \geq t_0$

$$R_t^{-1}\begin{pmatrix}X_t\\V_t\end{pmatrix} = R_{t_0}^{-1}\begin{pmatrix}X_{t_0}\\V_{t_0}\end{pmatrix} + \int_{t_0}^{t} R_s^{-1}\,\mathrm{d}\begin{pmatrix}0\\L_s\end{pmatrix}.$$

On multiplie cette équation par $e^{-tA}R_t$ et on fait le changement d'échelle décrit dans les Théorèmes 5.1 et 5.2. On obtient alors, pour $t \geq \varepsilon t_0$,

$$Y_t^{(\varepsilon)} = \varepsilon^q f\left(\frac{t}{\varepsilon}\right)\Phi_{t/\varepsilon}R_{t_0}^{-1}\begin{pmatrix}x_0\\v_0\end{pmatrix} + \Phi_{t/\varepsilon}\widetilde{M}_t^{(\varepsilon)}, \tag{5.8}$$

où

$$\Phi_t := \frac{e^{-tA}R_t}{f(t)}, \quad \forall t > 0, \quad \text{et} \quad \widetilde{M}_t^{(\varepsilon)} := \varepsilon^q f\left(\frac{t}{\varepsilon}\right)\int_{t_0}^{t/\varepsilon} R_s^{-1}\,\mathrm{d}\begin{pmatrix}0\\L_s\end{pmatrix}, \quad \forall t \geq \varepsilon t_0.$$

Le comportement asymptotique de R_t assure que Φ_t converge vers I_2 lorsque t tend vers l'infini. Le premier terme du membre de droite de (5.8) tend donc vers 0 quand $\varepsilon \to 0$. Il suffit alors d'étudier la convergence du processus $\widetilde{M}^{(\varepsilon)}$, lorsque $\varepsilon \to 0$. La convergence en loi ou en lois finie-dimensionnelles de ce dernier est traitée de façon analogue au régime sur-critique, mais repose cette fois sur le développement asymptotique de la résolvante R_t.

5.4 Perspectives

De nombreuses questions restent ouvertes. Dans le cas linéaire $\gamma = 1$, le comportement asymptotique, lorsque $\beta \in (0, 1/2]$, n'a pas été établi. On s'attend à la convergence, en lois finie-dimensionnelles, vers le même type de processus que lorsque $\beta \in (1/2, 1)$, c'est-à-dire un processus ayant des marginales indépendantes (bruit blanc). On est tenté de dire que les lois limites sont les mêmes que lorsque $\beta \in (1/2, 1)$, puisque leur expression a toujours un sens lorsque $\beta \in (0, 1/2]$. Le point technique qui manque pour traiter ces cas est le développement asymptotique des solutions de l'EDO (5.7).

Dans le cas non-linéaire $\gamma \neq 1$ et dans les régimes critiques et sous-critiques (i.e. quand la force de frottement ne peut pas être négligée), on n'a aucune idée des processus limites. La seule conjecture qu'on puisse faire est que, dans le régime sous-critique, le processus changé d'échelle converge encore, en lois finie-dimensionnelles, vers un bruit blanc, comme dans le cas linéaire.

On pourrait également envisager de considérer un système de particules en interaction de type champ moyen. Chaque particules évoluerait selon (SKE), avec des bruits i.i.d. et avec un terme d'interaction de type champ moyen supplémentaire. Cela ouvrirait à des questions de temps long sur les systèmes de particules, ou encore sur l'EDS de McKean-Vlasov correspondante.

Troisième partie

Travaux et preuves

Chapter 6

Well-posedness and propagation of chaos for Lévy-driven McKean-Vlasov SDEs under Lipschitz assumptions

This chapter corresponds to the article [Cav23]. It is in revision in *Electronic Communications in Probability*.

Abstract. The first goal of this note is to prove the strong well-posedness of McKean-Vlasov SDEs driven by Lévy processes on $\mathbb{R}^d$ having a finite moment of order $\beta \in [1, 2]$ and under standard Lipschitz assumptions on the coefficients. Then, we prove a quantitative propagation of chaos result at the level of paths for the associated interacting particle system, with constant diffusion coefficient. Finally, we improve the rates of convergence obtained for a particular mean-field system of interacting stable-driven Ornstein-Uhlenbeck processes.

6.1 Introduction and results

Let us fix $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space and $\mathcal{N}$ a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}^d \backslash \{0\}$ with intensity $dt \otimes \nu$, where ν is a Lévy measure, i.e.

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} 1 \wedge |z|^2 \, d\nu(z) < +\infty,$$

where $a \wedge b$ denotes the minimum between to real numbers a and b. We denote by $\widetilde{\mathcal{N}}(dt, dz) := \mathcal{N}(dt, dz) - dt \otimes d\nu(z)$ the associated compensated Poisson random measure. We consider $Z = (Z_t)_{t\geq 0}$ a Lévy process on $\mathbb{R}^d$ written, for all $t \geq 0$, as

$$Z_t = \int_0^t \int_{B_1} z \, \widetilde{\mathcal{N}}(ds, dz) + \int_0^t \int_{B_1^c} z \, \mathcal{N}(ds, dz),$$

where B_1 is the open ball of $\mathbb{R}^d$ centered at 0 and of radius 1 and B_1^c is its complementary in $\mathbb{R}^d$.

We assume that there exists $\beta \in [1, 2]$ such that the Lévy measure ν satisfies

$$\int_{B_1^c} |z|^\beta \, d\nu(z) < +\infty.$$

This is equivalent to assume that for any $t \in \mathbb{R}^+$, Z_t has a finite moment of order β by [Sat99, Theorem 25.3]. Let us denote by $\mathcal{P}_\beta(\mathbb{R}^d)$ the space of probability measures on $\mathbb{R}^d$ having a finite moment of order β, which is endowed with the Wasserstein metric W_β. We are interested in the well-posedness of the following Lévy-driven McKean-Vlasov SDE

$$\begin{cases} dX_t = b_t(X_t, \mu_t)\, dt + \sigma_t(X_{t-}, \mu_t)\, dZ_t, \quad t \in [0,T], \\ \mu_t := [X_t], \\ X_0 = \xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d), \end{cases} \tag{6.1}$$

where T is a fixed finite horizon of time, $[X_t]$ denotes the distribution of X_t and $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathcal{M}_d(\mathbb{R})$ are measurable maps, $\mathcal{M}_d(\mathbb{R})$ being the space of matrices of size $d \times d$ on $\mathbb{R}$. The first motivation to study (6.1) lies into its connexion with the following mean-field interacting particle system

$$\begin{cases} dX_t^{i,N} = b_t(X_t^{i,N}, \overline{\mu}_t^N)\, dt + \sigma_t(X_{t-}^{i,N}, \overline{\mu}_t^N)\, dZ_t^i, \quad t \in [0,T], \quad i \in \{1, \dots, N\}, \\ \overline{\mu}_t^N := \frac{1}{N} \sum\limits_{j=1}^N \delta_{X_t^{j,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{6.2}$$

where $(Z^i, \xi^i)_{i \geq 1}$ are i.i.d. with same distribution as (Z, ξ). The link between (6.1) and (6.2) is that for any $k \geq 1$, the dynamics of k particles is expected to be described by k independent copies of(6.1) when the total number of particles N tends to infinity. This is the so-called propagation of chaos phenomenon. It was originally studied by McKean [McK67] and then investigated by Sznitman [Szn91] when Z is a Brownian motion. For a detailed review on the topic of propagation of chaos, we refer the reader to [CD22a, CD22b].

We are going to work under the following Lipschitz assumptions.

Assumption (H1). There exists a constant $C > 0$ such that for all $t \in [0,T]$, $x, y \in \mathbb{R}^d$ and $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, we have

$$|b_t(x,\mu) - b_t(y,\nu)| + |\sigma_t(x,\mu) - \sigma_t(y,\nu)| \leq C(|x-y| + W_\beta(\mu,\nu)), \tag{6.3}$$

and

$$|b_t(x,\mu)| + |\sigma_t(x,\mu)| \leq C(1 + |x| + M_\beta(\mu)),$$

where $M_\beta(\mu) = \left(\int_{\mathbb{R}^d} |x|^\beta\, d\mu(x)\right)^{\frac{1}{\beta} \wedge 1}$ for $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$.

6.1.1 Well-posedness of the McKean-Vlasov SDE (6.1)

Theorem 6.1. *Under Assumption (H1), there exists a unique strong solution $(X_t)_{t\in[0,T]}$ to (6.1) for all initial datum $\xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d)$. Moreover, the flow of marginal distributions $(\mu_t)_{t\in[0,T]}$ belongs to $\mathcal{C}^0([0,T]; \mathcal{P}_\beta(\mathbb{R}^d))$ and we have*

$$\mathbb{E} \sup_{t \leq T} |X_t|^\beta < +\infty. \tag{6.4}$$

Remark 6.2. We can easily add a term of the form $(Bt + \Sigma W_t)_{t\geq 0}$ to Z, where $B \in \mathbb{R}^d$, $\Sigma \in \mathcal{M}_d(\mathbb{R})$ is a symmetric positive semidefinite matrix of size $d \times d$ and W is a standard Brownian motion on $\mathbb{R}^d$.

Using the Lévy-Itô decomposition given in [App09, Theorem 2.4.16], we can thus consider a general Lévy process Z having a finite moment of order $\beta \in [1, 2]$.

Remark 6.3. In [Gra92a], Graham proves a theorem including our result for $\beta = 1$. However, the justification of the Burkholder-Davis-Gundy (BDG) inequality used to justify the passage from (1.5) to (1.6) is not clear. Indeed, it is the angle bracket that is used instead of the sharp bracket, which appears in the standard BDG inequality. The argument does not seem immediate. That is why we prove it differently in Theorem 6.1.

Let us compare our result with the existing literature. When $\beta = 2$, the well-posedness of (6.1) was proved by Jourdain, Méléard and Woyczynski [JMW07]. In this work, the weak existence is also proved when $\beta = 0$ through the relative nonlinear martingale problem. However, uniqueness is not shown when $\beta = 0$. When $\beta = 1$, a result similar to Theorem 6.1 is proved by Graham in [Gra92b, Theorem 2.2]. The main differences are the following. Firstly, in [Gra92b], there is no integral with respect to the compensated Poisson random measure $\widetilde{\mathcal{N}}$ in the definition of Z. Secondly, in the case where the drift b is unbounded, it is supposed in [Gra92b] that X_0 has a finite moment of order 2, which is not the case in Theorem 6.1, and also that, keeping our notations, there exists $C > 0$ such that for all $t \in [0, T]$, $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$, we have

$$\left| \int_{B_1^c} \sigma_t(x, \mu) z \, d\nu(z) \right|^2 + \int_{B_1^c} |\sigma_t(x, \mu) z|^2 \, d\nu(z) \leq C(1 + |x|^2). \tag{6.5}$$

It suggests that σ is bounded with respect to its measure variable, which is not the case in Theorem 6.1. Moreover, (6.5) strongly suggests that

$$\int_{B_1^c} |z|^2 \, d\nu(z) < +\infty.$$

It is the case when $\sigma = \mathrm{Id}$ for example. However, this condition on ν is equivalent to the fact that for any $t \in \mathbb{R}^+$, Z_t has a finite moment of order 2, which is not supposed in Theorem 6.1 since $\beta \in [1, 2]$. In the non-degenerate case, i.e. when σ is uniformly elliptic, we refer to [FKM21]. In this work, Frikha, Menozzi and Konakov prove the well-posedness of (6.1) under Hölder assumptions on the coefficients with respect to both space and measure variables. Of course, this result can be applied to Lipschitz continuous coefficients but in Theorem 6.1, we do not assume that the diffusion coefficient σ is uniformly elliptic. Moreover, another assumption made in [FKM21] is that for all $(t, x) \in [0, T] \times \mathbb{R}^d$, the maps $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto b_t(x, \mu)$ and $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto \sigma_t(x, \mu)$ have bounded linear derivatives, where $\mathcal{P}(\mathbb{R}^d)$ is the space of probability measures on $\mathbb{R}^d$. Note that, at least when the coefficients depend linearly on the measure, this assumption implies the boundedness of the coefficients with respect to the measure variable. This is not the case here.

Remark 6.4. Notice that when $\beta \in (0, 1)$, the uniqueness result of Theorem 6.1 is false without a non-degeneracy assumption on the diffusion coefficient σ. Let us give a simple counterexample by setting, for $t \in [0, T]$, $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$

$$b_t(x, \mu) := \int_{\mathbb{R}^d} |x|^\beta \, d\mu(x), \quad \sigma_t(x, \mu) := 0, \quad \text{and} \quad \xi := 0.$$

Assumption (H1) is clearly satisfied. Moreover, the solution to the corresponding McKean-Vlasov SDE is deterministic since there is no noise and the initial distribution is deterministic. We easily remark that the problem is equivalent to solve the ordinary differential equation

$$\begin{cases} y'(t) = |y(t)|^{\beta}, & t \in [0,T], \\ y(0) = 0. \end{cases}$$

It is well-known that there exists several solutions to this problem. However, under Assumption (H1), there exists at least one strong solution to the McKean-Vlasov SDE (6.1). We refer to Appendix 6.4 for a proof of this result.

6.1.2 Propagation of chaos for the interacting particle system (6.2)

We now focus on the propagation of chaos for the interacting particle system (6.2). Under Assumption (H1), the SDE (6.2) admits a unique strong solution by [App09, Theorem 6.2.9]. Propagation of chaos can be understood in the weak sense, i.e. in distribution through the convergence of the empirical measure $\overline{\mu}^N$, or in the strong sense, i.e. at the level of paths by coupling. Our aim is to prove quantitative strong propagation of chaos. Let us introduce the i.i.d. copies of the limiting McKean-Vlasov SDE (6.1), which are denoted by $(X^{i,\infty})_{i\geq 1}$, where the initial data and the noises are respectively $(\xi^i)_{i\geq 1}$ and $(Z^i)_{i\geq 1}$. We assume that $\beta \in [1,2)$ since the case $\beta = 2$ is treated in [JMW07] and [NBK+20].

Theorem 6.5. *Assume that Assumption (H1) holds true with W_1 instead of W_β in the Lipschitz control (6.3) and with $\sigma = Id$. Then, there exists a constant $C > 0$ such that for all $N \geq 1$*

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_t^{i,N} - X_t^{i,\infty}| \leq C \begin{cases} N^{\frac{1}{\beta}-1}, & \text{if } \ d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \beta < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if } \ d \geq 3 \text{ and } \beta > \frac{d}{d-1}, \end{cases} \tag{6.6}$$

and

$$\sup_{t\in[0,T]} \mathbb{E} W_1(\overline{\mu}_t^N, \mu_t) \leq C \begin{cases} N^{\frac{1}{\beta}-1}, & \text{if } \ d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \beta < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if } \ d \geq 3 \text{ and } \beta > \frac{d}{d-1}. \end{cases} \tag{6.7}$$

Remark 6.6. The method used in the proof of Theorem 6.5 cannot be applied to prove quantitative strong propagation of chaos with a general non-constant diffusion coefficient σ under Assumption (H1). It remains, to the best of our knowledge, an open problem.

We now compare our result with the existing literature. In [Gra92a], Graham proves qualitative weak propagation of chaos, i.e. without rate of convergence, under Lipschitz assumptions for an interacting particle system driven by a Poisson random measure and its compensated measure. It is supposed that the Poisson random measure is associated with a Poisson process having a finite moment of order 1 and that the set of jumps is discrete. Jourdain, Méléard and Woyczynski treat in [JMW07] the case of a general Lévy noise having a finite moment of order 2. The authors exhibit rates of convergence for the strong propagation of chaos in L^2 under standard Lipschitz assumptions on the drift and diffusion coefficients b and σ which are similar to Assumption (H1). Still in the Lipschitz framework, we mention Neelima et al. [NBK+20], where quantitative strong propagation of chaos is proved in L^2, relaxing the assumptions of [JMW07]. In the one-dimensional case, Frikha and Li [FL21] study a McKean-Vlasov SDE driven by a compensated Poisson random measure with positive jumps. They give a rate of convergence for the strong propagation of chaos in L^1 under one-sided Lipschitz assumptions on the coefficients.

Let us now study a particular example for which we can improve the rates of convergence obtained in Theorem 6.5. Assume that $Z = (Z_t)_{t\geq 0}$ is an α-stable process on $\mathbb{R}^d$ with $\alpha \in (1, 2)$. Let us fix also $A, A', B \in \mathcal{M}_d(\mathbb{R})$ matrices of size $d \times d$. We are interested in the interacting particle system (6.2) and the limiting McKean-Vlasov SDE (6.1) with

$$\xi \in L^\alpha(\Omega, \mathcal{F}_0; \mathbb{R}^d), \quad b_t(x, \mu) := Ax + A' \int_{\mathbb{R}^d} y \, d\mu(y) \quad \text{and} \quad \sigma_t(x, \mu) := \mathrm{Id}.$$

This corresponds to a system of interacting stable-driven Ornstein-Uhlenbeck processes. Keeping the same notations as in Theorem 6.5, we have the following quantitative propagation of chaos result.

Theorem 6.7. *There exists a positive constant C such that for all $N \geq 1$*

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_t^{i,N} - X_t^{i,\infty}| \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1, 2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}, \end{cases} \tag{6.8}$$

and

$$\sup_{t\in[0,T]} \mathbb{E} W_1(\overline{\mu}_t^N, \mu_t) \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1, 2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}. \end{cases} \tag{6.9}$$

6.2 Proof of Theorem 6.1

Let us fix $\mu = (\mu_t)_{t\in[0,T]} \in \mathcal{C}^0([0, T]; \mathcal{P}_\beta(\mathbb{R}^d))$. By using [App09, Theorem 6.2.9], we deduce that the SDE

$$\begin{cases} dX_t^\mu = b_t(X_t^\mu, \mu_t) \, dt + \sigma_t(X_{t-}^\mu, \mu_t) \, dZ_t, \quad t \in [0, T], \\ X_0^\mu = \xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d), \end{cases} \tag{6.10}$$

admits a unique strong solution X^μ. Moreover, note that the coefficients of this standard SDE $(t, x) \in [0, T] \times \mathbb{R}^d \mapsto b_t(x, \mu_t)$ and $(t, x) \in [0, T] \times \mathbb{R}^d \mapsto \sigma_t(x, \mu_t)$ are at most of linear growth with respect to the space variable x, uniformly with respect to $t \in [0, T]$. By using Proposition 2 in Fournier [Fou13], we get that

$$\mathbb{E} \sup_{t\leq T} |X_t^\mu|^\beta < +\infty.$$

The map

$$\phi : \begin{cases} \mathcal{C}^0([0, T]; \mathcal{P}_\beta(\mathbb{R}^d)) & \to \mathcal{C}^0([0, T]; \mathcal{P}_\beta(\mathbb{R}^d)) \\ \mu & \mapsto ([X_t^\mu])_{t\in[0,T]} \end{cases} \tag{6.11}$$

is thus well-defined. The goal is now to prove that ϕ has a unique fixed point thanks to the Banach fixed point theorem. This is enough to prove the strong well-posedness of (6.1). The space $\mathcal{C}^0([0, T]; \mathcal{P}_\beta(\mathbb{R}^d))$ is endowed with the uniform metric associated with W_β. We fix $\mu, \nu \in \mathcal{C}^0([0, T]; \mathcal{P}_\beta(\mathbb{R}^d))$ and we aim at estimating $\mathbb{E} \sup_{s\leq t} |X_s^\mu - X_s^\nu|^\beta$, for $t \in [0, T]$. We employ the method used by Fournier in the proof of [Fou13, Proposition 2], which was used in the context of McKean-Vlasov SDEs by Frikha and Li in [FL21] to prove the moment estimation (6.4). The first step is to consider the SDE (6.10) without the big jumps term. Namely, we assume that for $\xi_1, \xi_2 \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d)$, and for all $t \in [0, T]$

$$X_t^\mu = \xi_1 + \int_0^t b_s(X_s^\mu, \mu_s) \, ds + \int_0^t \int_{B_1} \sigma_s(X_{s-}^\mu, \mu_s) z \widetilde{\mathcal{N}}(ds, dz), \tag{6.12}$$

and

$$X^{\nu}_t = \xi_2 + \int_0^t b_s(X^{\nu}_s, \nu_s)\, ds + \int_0^t \int_{B_1} \sigma_s(X^{\nu}_{s-}, \nu_s) z \widetilde{\mathcal{N}}(ds, dz). \tag{6.13}$$

Note that by definition of ϕ, ξ_1 is equal to ξ_2, however in the next step of the proof, we need to take different initial data for the SDE. Using the Lipschitz assumption on the coefficients, Jensen and the BDG inequalities, we obtain that for a constant $C = C_T$ depending only on T and which can change from line to line, we have for all $t \in [0, T]$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{s\le t} |X^{\mu}_s - X^{\nu}_s|^2 \mid \xi_1, \xi_2\right) \\
&\qquad \le C\left[|\xi_1 - \xi_2|^2 + \int_0^t \mathbb{E}\left(|X^{\mu}_s - X^{\nu}_s|^2 \mid \xi_1, \xi_2\right) ds + \int_0^t W^2_{\beta}(\mu_s, \nu_s)\, ds\right].
\end{aligned}$$

Gronwall's lemma ensures that

$$\mathbb{E}\left(\sup_{s\le t} |X^{\mu}_s - X^{\nu}_s|^2 \mid \xi_1, \xi_2\right) \le C|\xi_1 - \xi_2|^2 + C\int_0^t W^2_{\beta}(\mu_s, \nu_s)\, ds.$$

It follows from Jensen's inequality that

$$\begin{aligned}
\mathbb{E}\left(\sup_{s\le t} |X^{\mu}_s - X^{\nu}_s|^{\beta} \mid \xi_1, \xi_2\right) &\le \left(\mathbb{E}\left(\sup_{s\le t} |X^{\mu}_s - X^{\nu}_s|^2 \mid \xi_1, \xi_2\right)\right)^{\frac{\beta}{2}} \\
&\le C|\xi_1 - \xi_2|^{\beta} + C\left(\int_0^t W^2_{\beta}(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}}.
\end{aligned}$$

Taking the expectation yields for all $t \in [0, T]$

$$\mathbb{E}\left(\sup_{s\le t} |X^{\mu}_s - X^{\nu}_s|^{\beta}\right) \le C\mathbb{E}|\xi_1 - \xi_2|^{\beta} + C\left(\int_0^t W^2_{\beta}(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}}. \tag{6.14}$$

Let us now add the big jumps. We denote by $(T_n)_{n\ge1}$ the sequence of jumping times of $(Z_t)_{t\ge0}$ having a size greater than 1, and by $(\Delta Z_n)_{n\ge1}$ the associated sequence of jumps, which is an i.i.d. sequence of random variables with common distribution $\frac{\nu_{|B_1^c}}{\nu(B_1^c)}$ and independent of $(T_n)_{n\ge1}$. We can write the restriction of the Poisson random measure $\mathcal{N}$ on $\mathbb{R}^+ \times B_1^c$ as

$$\sum_{n\ge1} \delta_{(T_n, \Delta Z_n)},$$

which is independent on the restriction of $\mathcal{N}$ on $\mathbb{R}^+ \times B_1\backslash\{0\}$. Let us denote by $\mathcal{G}$ the σ-algebra generated by $(T_n)_{n\ge1}$. Notice that on the time interval $[0, T_1)$, X^{μ} and X^{ν} defined in (6.10) are respectively solutions to (6.12) and (6.13) with $\xi_1 = \xi_2$. Thus, using (6.14) with the conditional expectation with respect to $\mathcal{G}$ instead of the expectation, we deduce that

$$\mathbb{E}\left(\sup_{s<t\wedge T_1} |X^{\mu}_s - X^{\nu}_s|^{\beta} \mid \mathcal{G}\right) \le C\left(\int_0^{t\wedge T_1} W^2_{\beta}(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}}. \tag{6.15}$$

Let us now deal with the first big jump of Z, which occurs at time T_1. We have

$$X^\mu_{T_1} - X^\nu_{T_1} = X^\mu_{T_1^-} - X^\nu_{T_1^-} + \Big(\sigma_{T_1}(X^\mu_{T_1^-}, \mu_{T_1}) - \sigma_{T_1}(X^\nu_{T_1^-}, \nu_{T_1})\Big)\, \Delta Z_1.$$

It follows from the Lipschitz assumption on σ that

$$|X^\mu_{T_1} - X^\nu_{T_1}|^\beta \leq C|X^\mu_{T_1^-} - X^\nu_{T_1^-}|^\beta(1 + |\Delta Z_1|^\beta) + W^\beta_\beta(\mu_{T_1}, \nu_{T_1})|\Delta Z_1|^\beta.$$

Since ΔZ_1 is independent of $\mathcal{G}$ and $\mathbb{E}|\Delta Z_1|^\beta < +\infty$, we deduce by (6.15) that almost surely on the set $\{T_1 \leq T\}$

$$\mathbb{E}\left(|X^\mu_{T_1} - X^\nu_{T_1}|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1\leq t} \leq C\left[\left(\int_0^{t\wedge T_1} W^2_\beta(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}} + W^\beta_\beta(\mu_{T_1}, \nu_{T_1})\right].$$

We thus have by the preceding inequality and (6.15)

$$\begin{aligned}
&\mathbb{E}\left(|X^\mu_{t\wedge T_1} - X^\nu_{t\wedge T_1}|^\beta \mid \mathcal{G}\right)\\
&= \mathbb{E}\left(|X^\mu_{T_1} - X^\nu_{T_1}|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1\leq t} + \mathbb{E}\left(|X^\mu_{t} - X^\nu_{t}|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1> t}\\
&\leq C\left[\left(\int_0^{t\wedge T_1} W^2_\beta(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}} + W^\beta_\beta(\mu_{T_1}, \nu_{T_1})\right].
\end{aligned}$$

Following the same lines and using (6.14), we prove that for any $n \geq 1$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\leq s< t\wedge T_{n+1}} |X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\\
&\qquad\leq C\left[\mathbb{E}\left(|X^\mu_{t\wedge T_n} - X^\nu_{t\wedge T_n}|^\beta \mid \mathcal{G}\right) + \left(\int_{t\wedge T_n}^{t\wedge T_{n+1}} W^2_\beta(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}}\right], \qquad (6.16)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\leq s\leq t\wedge T_{n+1}} |X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right) \qquad (6.17)\\
&\leq C\left[\mathbb{E}\left(|X^\mu_{t\wedge T_n} - X^\nu_{t\wedge T_n}|^\beta \mid \mathcal{G}\right) + \left(\int_{t\wedge T_n}^{t\wedge T_{n+1}} W^2_\beta(\mu_s, \nu_s)\, ds\right)^{\frac{\beta}{2}} + W^2_\beta(\mu_{T_{n+1}}, \nu_{T_{n+1}})\right],
\end{aligned}$$

where C is independent of n. Reasoning by induction, we deduce that for a constant $C > 1$ depending only on T, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\le s<t\wedge T_{n+1}} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\\
&\le C^{n+1}\left[\sum_{k=0}^{n}\left(\int_{t\wedge T_k}^{t\wedge T_{k+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W_\beta^\beta(\mu_{T_k},\nu_{T_k})\right]\\
&\le C^{n+1}\left[(n+1)^{1-\frac{\beta}{2}}\left(\int_0^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W_\beta^\beta(\mu_{T_k},\nu_{T_k})\right],
\end{aligned}$$

by Jensen's inequality and with the convention that $T_0 = 0$. Thus, for a certain constant $K > C$, one has for all $j \le n$

$$\mathbb{E}\left(\sup_{t\wedge T_j\le s<t\wedge T_{j+1}} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right) \le K^{j+2}\left[\left(\int_0^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W_\beta^\beta(\mu_{T_k},\nu_{T_k})\right].$$

Summing the preceding inequality over $j \in \{0,\dots,n\}$, we deduce that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s<t\wedge T_{n+1}} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\\
&\qquad\le \frac{K^{n+3}}{1-K}\left[\left(\int_0^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W_\beta^\beta(\mu_{T_k},\nu_{T_k})\right]. \qquad (6.18)
\end{aligned}$$

Let us denote by $(N_t)_{t\ge 0}$ the Poisson process associated with the jumping times $(T_n)_{n\ge 1}$ which has an intensity $\lambda = \nu(B_1^c)$. One has

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{N_t=n}\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{t<T_1}\right) + \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_n\le t<T_{n+1}}\right)\\
&\le \mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s<t\wedge T_1} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{t<T_1}\right)\\
&\quad + \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s<t\wedge T_{n+1}} |X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_n\le t<T_{n+1}}\right)
\end{aligned}$$

Using (6.15) and (6.18), we obtain that for any $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s^\mu - X_s^\nu|^\beta\right) \\
&\leq C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\right)^{\frac{\beta}{2}} \\
&\quad + \sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\mathbb{E}\left(\left[\left(\int_0^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^n W_\beta^\beta(\mu_{T_k},\nu_{T_k})\right]\mathbf{1}_{T_n\leq t<T_{n+1}}\right) \\
&\leq C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\right)^{\frac{\beta}{2}} \\
&\quad + \sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\mathbb{P}(N_t=n)\left(\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \mathbb{E}\left(\sum_{k=1}^n W_\beta^\beta(\mu_{T_k},\nu_{T_k}) \mid N_t=n\right)\right).
\end{aligned} \tag{6.19}$$

Let us recall that the conditional distribution of $(T_1,\dots,T_n)$ given $N_t=n$ admits the following density with respect to the Lebesgue measure

$$(t_1,\dots,t_n)\in[0,t]^n \mapsto \frac{n!}{t^n}\mathbf{1}_{t_1<\cdots<t_n}.$$

This yields

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^n W_\beta^\beta(\mu_{T_k},\nu_{T_k}) \mid N_t=n\right) &= \int_{[0,t]^n}\frac{n!}{t^n}\mathbf{1}_{t_1<\cdots<t_n}\sum_{k=1}^n W_\beta^\beta(\mu_{t_k},\nu_{t_k})\,dt_1\dots dt_n \\
&= \int_{[0,t]^n}\frac{1}{t^n}\sum_{k=1}^n W_\beta^\beta(\mu_{t_k},\nu_{t_k})\,dt_1\dots dt_n \\
&= \sum_{k=1}^n \frac{1}{t^n}t^{n-1}\int_0^t W_\beta^\beta(\mu_s,\nu_s)\,ds \\
&= \frac{n}{t}\int_0^t W_\beta^\beta(\mu_s,\nu_s)\,ds.
\end{aligned}$$

Injecting this equality in (6.19), we get

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s^\mu - X_s^\nu|^\beta\right) \\
&\leq \sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\frac{(\lambda t)^n}{n!}e^{-\lambda t}\left(\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \frac{n}{t}\int_0^t W_\beta^\beta(\mu_s,\nu_s)\,ds\right) \\
&\quad + C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\right)^{\frac{\beta}{2}}.
\end{aligned}$$

This proves the existence of a constant $C>0$ depending only on T such that for all $t\in[0,T]$

$$\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s^\mu - X_s^\nu|^\beta\right) \leq C\left[\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \int_0^t W_\beta^\beta(\mu_s,\nu_s)\,ds\right]. \tag{6.20}$$

Note that (6.20) is true if $\beta\in(0,1)$ since we have only used that $0<\beta\leq 2$. Changing again the constant

C, Hölder's inequality yields for all $t \in [0,T]$

$$\sup_{0\le s\le t} W_\beta^\beta(\phi(\mu)_s, \phi(\nu)_s) \le \mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta\right) \le C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}. \tag{6.21}$$

Raising the preceding inequality to the power $\frac{2}{\beta}$ and reasoning by induction, we prove that for any $n \ge 1$ and for any $t \in [0,T]$

$$\sup_{0\le s\le t} W_\beta^2(\phi^n(\mu)_s, \phi^n(\nu)_s) \le \frac{C^{\frac{2n}{\beta}} t^n}{n!} \sup_{0\le s\le t} W_\beta^2(\mu_s,\nu_s),$$

which yields

$$\sup_{0\le s\le T} W_\beta^\beta(\phi^n(\mu)_s, \phi^n(\nu)_s) \le C^n\left(\frac{T^n}{n!}\right)^{\frac{\beta}{2}} \sup_{0\le s\le T} W_\beta^\beta(\mu_s,\nu_s).$$

This proves that for n large enough, ϕ^n is a contraction on $\mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d))$. The function ϕ has thus a unique fixed point by the Banach fixed point theorem, which concludes the proof.

6.3 Proof of Theorem 6.5 and Theorem 6.7

Proof of Theorem 6.5. To prove (6.6), we write for all $t \in [0,T]$

$$X_t^{i,N} - X_t^{i,\infty} = \int_0^t b_t(X_s^{i,N}, \overline{\mu}_s^N) - b_t(X_s^{i,\infty}, \mu_s)\,ds.$$

Using the Lipschitz assumption on b, there exists $C > 0$ such that for all $t \in [0,T]$

$$\begin{aligned}
&\sup_{i\le N}\mathbb{E}\sup_{r\le t}|X_r^{i,N} - X_r^{i,\infty}| \\
&\le C\int_0^t \sup_{i\le N}\mathbb{E}|X_s^{i,N} - X_s^{i,\infty}|\,ds + C\int_0^t \mathbb{E}W_1(\overline{\mu}_s^N, \mu_s)\,ds \\
&\le C\int_0^t \sup_{i\le N}\mathbb{E}|X_s^{i,N} - X_s^{i,\infty}|\,ds + C\int_0^t \mathbb{E}W_1(\overline{\mu}_s^N, \tilde{\mu}_s^N) + \mathbb{E}W_1(\tilde{\mu}_s^N, \mu_s)\,ds,
\end{aligned}$$

where $\tilde{\mu}_s^N := \frac{1}{N}\sum_{k=1}^N \delta_{X_s^{k,\infty}}$ is the empirical measure associated with $(X^{i,\infty})_{i\ge 1}$. Using that

$$W_1(\overline{\mu}_s^N, \tilde{\mu}_s^N) \le \frac{1}{N}\sum_{k=1}^N |X_s^{k,N} - X_s^{k,\infty}|$$

and Gronwall's inequality, there exists $C > 0$ such that for all $N \ge 1$

$$\sup_{i\le N}\mathbb{E}\sup_{t\le T}|X_t^{i,N} - X_t^{i,\infty}| \le C\int_0^T \mathbb{E}W_1(\tilde{\mu}_s^N, \mu_s)\,ds. \tag{6.22}$$

We conclude using [FG15, Theorem 1] since $(X^{i,\infty})_{i\ge 1}$ are i.i.d. and

$$\sup_{i\ge 1}\sup_{t\in[0,T]} \mathbb{E}|X_t^{i,\infty}|^\beta < +\infty$$

by Gronwall's inequality. The inequality (6.7) follows from (6.6) and [FG15] because

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}W_1(\overline{\mu}_t^N,\mu_t) &\le \sup_{t\in[0,T]} \mathbb{E}W_1(\overline{\mu}_t^N,\tilde{\mu}_t^N)+\sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N,\mu_t)\\ &\le \sup_{t\in[0,T]}\sup_{i\le N}\mathbb{E}|X_t^{i,N}-X_t^{i,\infty}|+\sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N,\mu_t).\end{aligned}$$

Proof of Theorem 6.7. Let us prove (6.8). As a first step, we remove the jumps of size larger than the number of particles N from all the noises. We thus define, for $i\ge 1$ and $t\in[0,T]$

$$Z_{N,t}^i := \int_0^t\int_{B_N} z\widetilde{\mathcal{N}}^i(ds,dz),$$

where $\mathcal{N}^i$ is the Poisson random measure associated with Z^i and $\widetilde{\mathcal{N}}^i$ its compensated Poisson random measure. We define, for all $i\in\{1,\dots,N\}$, $X_N^{i,\infty}$ as the unique solution to

$$\begin{cases} dX_{N,t}^{i,\infty} &= AX_{N,t}^{i,\infty}\,dt + A'\mathbb{E}X_{N,t}^{i,\infty}\,dt + B\,dZ_{N,t}^i, \quad t\in[0,T], \quad i\in\{1,\dots,N\},\\ \mu_{N,t} &:= [X_{N,t}^{i,\infty}],\\ X_{N,0}^{i,\infty} &= \xi^i. \end{cases} \tag{6.23}$$

For any $N\ge 1$ fixed, the random variables $(X_N^{i,\infty})_{i\le N}$ are i.i.d. We proceed similarly for the particle system by defining $(X_N^{i,N})_{i\le N}$ as the unique solution to

$$\begin{cases} dX_{N,t}^{i,N} = AX_{N,t}^{i,N}\,dt + A'\frac{1}{N}\sum\limits_{k=1}^N X_{N,t}^{k,N}\,dt + B\,dZ_{N,t}^i, \quad t\in[0,T], \quad i\in\{1,\dots,N\},\\ \overline{\mu}_{N,t}^N := \frac{1}{N}\sum\limits_{j=1}^N \delta_{X_{N,t}^{j,N}},\\ X_{N,0}^{i,N} = \xi^i, \end{cases} \tag{6.24}$$

The first objective is to control the L^1-error respectively between $X_N^{i,N}$ and $X^{i,N}$ and between $X_N^{i,\infty}$ and $X^{i,\infty}$ for all $i\in\{1,\dots,N\}$. We write for all $t\in[0,T]$

$$X_{N,t}^{i,N}-X_t^{i,N} = \int_0^t (b(X_{N,s}^{i,N},\overline{\mu}_{N,s}^N)-b(X_s^{i,N},\overline{\mu}_s^N))\,ds + B\int_0^t\int_{B_N^c} z\widetilde{\mathcal{N}}^i(ds,dz).$$

Using the fact that b is Lipschitz continuous on $\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R}^d)$ and BDG's inequality, there exists $C>0$ independent on $N\ge 1$ and $t\in[0,T]$, which can change from line to line, such that for all $t\in[0,T]$

$$\begin{aligned}\sup_{i\le N}\mathbb{E}\sup_{r\le t}|X_{N,r}^{i,N}-X_r^{i,N}| \le C\Bigg[&\int_0^t \sup_{i\le N}\mathbb{E}|X_{N,s}^{i,N}-X_s^{i,N}|\,ds + \int_0^t \frac{1}{N}\sum_{j=1}^N \mathbb{E}|X_{N,s}^{j,N}-X_s^{j,N}|\,ds\\ &+\sup_{i\le N}\mathbb{E}\left(\int_0^t\int_{B_N^c}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac12}\Bigg].\end{aligned}$$

Using the subadditivity of the square root and the fact that the integral with respect to $\mathcal{N}^i$ is a discrete sum, one has for all $t \in [0,T]$

$$\left(\int_0^t \int_{B_N^c} |z|^2 \mathcal{N}^i(ds,dz)\right)^{\frac{1}{2}} \leq \int_0^t \int_{B_N^c} |z| \mathcal{N}^i(ds,dz).$$

It follows that

$$\begin{aligned}
&\sup_{i\leq N} \mathbb{E} \sup_{r\leq t} |X_{N,r}^{i,N} - X_r^{i,N}| \\
&\leq C \left[\int_0^t \sup_{i\leq N} \mathbb{E}|X_{N,s}^{i,N} - X_s^{i,N}|\, ds + \mathbb{E}\int_0^t \int_{B_N^c} |z| \mathcal{N}^i(ds,dz)\right] \\
&= C \left[\int_0^t \sup_{i\leq N} \mathbb{E}|X_{N,s}^{i,N} - X_s^{i,N}|\, ds + \int_0^t \int_{B_N^c} |z|\, d\nu(z)\, ds\right] \\
&\leq C \left[\int_0^t \sup_{i\leq N} \mathbb{E}|X_{N,s}^{i,N} - X_s^{i,N}|\, ds + \int_N^\infty r\, \frac{dr}{r^{1+\alpha}}\right].
\end{aligned}$$

Gronwall's inequality ensures that

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_{N,t}^{i,N} - X_t^{i,N}| \leq CN^{1-\alpha}. \tag{6.25}$$

We similarly get that for some constant $C > 0$ independent of N

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_{N,t}^{i,\infty} - X_t^{i,\infty}| \leq CN^{1-\alpha}. \tag{6.26}$$

The triangle inequality, (6.25) and (6.26) yield

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_t^{i,N} - X_t^{i,\infty}| \leq CN^{1-\alpha} + \sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_{N,t}^{i,N} - X_{N,t}^{i,\infty}|. \tag{6.27}$$

The second term in the right hand-side of (6.27) is controlled as in the proof of Theorem 6.5. We get that there exists $C > 0$ such that for all $N \geq 1$

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_{N,t}^{i,N} - X_{N,t}^{i,\infty}| \leq C \int_0^T \mathbb{E} W_1(\tilde{\mu}_{N,s}^N, \mu_{N,s})\, ds. \tag{6.28}$$

As the random variables $(X_N^{i,\infty})_{i\leq N}$ are i.i.d. when N is fixed, we are going to use [FG15, Theorem 1] to control $\mathbb{E}W_1(\tilde{\mu}_{N,s}^N, \mu_{N,s})$ uniformly with respect to $s \in [0,T]$. We start by controlling the moments of $X_N^{i,\infty}$. Let us fix $\beta \in [1,\alpha]$. Gronwall's inequality ensures that there exists $C > 0$ such that for any $t \in [0,T]$

$$\mathbb{E}|X_{N,t}^{i,\infty}|^\beta \leq C \sup_{s\leq t} \mathbb{E}|Z_{N,s}^i|^\beta.$$

If $\beta < \alpha$ and since Z admits a finite moment of order β, it is clear that

$$\sup_{N\geq 1} \sup_{s\leq T} \mathbb{E}|X_{N,s}^{i,\infty}|^\beta \leq \sup_{N\geq 1} \sup_{s\leq T} \mathbb{E}|Z_{N,s}^i|^\beta < +\infty. \tag{6.29}$$

If $\beta = \alpha$, BDG's and Jensen's inequalities yield

$$\begin{aligned}
\mathbb{E}|Z^i_{N,t}|^\alpha &\le C\mathbb{E}\left(\int_0^t\int_{B_N}|z|^2\,\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\\
&\le C\left[\left(\mathbb{E}\int_0^t\int_{B_1}|z|^2\,\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}} + \mathbb{E}\left(\int_0^t\int_{B_N\setminus B_1}|z|^2\,\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\right]\\
&\le C\left[\left(t\int_{B_1}|z|^2\,d\nu(z)\right)^{\frac{\alpha}{2}} + \mathbb{E}\left(\int_0^t\int_{B_N\setminus B_1}|z|^2\,\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\right] \qquad (6.30)\\
&\le C\left[1+\int_1^N r^\alpha\,\frac{dr}{r^{1+\alpha}}\right]\\
&\le C\ln(N).
\end{aligned}$$

If $d=1$, $d=2$ or $d\ge 3$ and $\alpha<\frac{d}{d-1}$, it follows from (6.30) and [FG15] that for all $N\ge 1$

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_{N,t},\mu_{N,t}) \le C(\ln(N))^{\frac{1}{\alpha}}\begin{cases}(N^{-\frac{1}{2}}+N^{\frac{1}{\alpha}-1}), & \text{if}\quad d=1,\\ (N^{-\frac{1}{2}}\ln(N+1)+N^{\frac{1}{\alpha}-1}), & \text{if}\quad d=2,\\ (N^{-\frac{1}{d}}+N^{\frac{1}{\alpha}-1}), & \text{if}\quad d\ge 3 \text{ and } \alpha\neq\frac{d}{d-1}.\end{cases}$$

Since $\alpha<2$ and if $d\ge 3$, we have assumed that $\alpha<\frac{d}{d-1}$ and thus $\frac{1}{d}>1-\frac{1}{\alpha}$, we deduce that for all $N\ge 1$

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_{N,s},\mu_{N,s}) \le C(\ln(N))^{\frac{1}{\alpha}}N^{\frac{1}{\alpha}-1}. \qquad (6.31)$$

In the case where $d\ge 3$ and $\alpha>\frac{d}{d-1}$, let us introduce $\beta\in\left(\frac{d}{d-1},\alpha\right)$. By (6.29) with this choice of β and [FG15], for all $N\ge 1$

$$\sup_{t\in[0,T]}W_1(\tilde{\mu}^N_{N,t},\mu_{N,t}) \le CN^{-\frac{1}{d}}. \qquad (6.32)$$

This ends the proof of (6.8) thanks to (6.28) and (6.27) since $\frac{\alpha-1}{\alpha}<\alpha-1$ because in the case where $d\ge 3$ and $\alpha>\frac{d}{d-1}$, then $\alpha-1\ge\frac{1}{d}$.

For the proof of (6.9), keeping the same notations as in the proof of Theorem 6.5, we have

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\overline{\mu}^N_t,\mu_t) \le \sup_{t\in[0,T]}\sup_{i\le N}\mathbb{E}|X^{i,N}_t - X^{i,\infty}_t| + \sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_t,\mu_t).$$

The first term in the right hand-side of the preceding inequality is controlled by (6.8). For the second one, we use the following decomposition

$$\begin{aligned}
\sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_t,\mu_t) &\le \sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_t,\tilde{\mu}^N_{N,t}) + \sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_{N,t},\mu_{N,t}) + \sup_{t\in[0,T]}\mathbb{E}W_1(\mu_{N,t},\mu_t)\\
&=: I_1+I_2+I_3.
\end{aligned}$$

Using (6.25) and (6.26), we get that

$$I_1+I_3 \le CN^{1-\alpha}. \qquad (6.33)$$

For I_2, the inequalities (6.32) and (6.31) prove that for all $N \geq 1$

$$I_2 \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}. \end{cases} \tag{6.34}$$

Gathering (6.33) and (6.34) ends the proof of (6.9) as previously.

6.4 Appendix: Existence of a solution to (6.1) when $\beta \in (0,1)$

Let us fix $\beta \in (0,1)$. We have seen in Remark 6.4 that in this case, uniqueness for (6.1) fails to be true under Assumption (H1). However, the existence of solutions to (6.1) is given in the following proposition.

Proposition 6.8. *We assume that Assumption (H1) is satisfied and that there exists $\delta > 0$ such that*

$$\int_{B_1^c} |z|^{\beta+\delta} \, d\nu(z) < +\infty.$$

Then, there exists a strong solution $(X_t)_{t\in[0,T]}$ to (6.1) for all $\xi \in L^{\beta+\delta}(\Omega, \mathcal{F}_0; \mathbb{R}^d)$. Moreover, we have

$$\mathbb{E} \sup_{t\leq T} |X_t|^{\beta+\delta} < +\infty. \tag{6.35}$$

Proof. The strategy relies on a compactness argument. Let us denote by $D_T := D([0,T]; \mathbb{R}^d)$ the Skorokhod space, i.e. the space of càdlàg $\mathbb{R}^d$-valued functions defined on $[0,T]$. We endow D_T with the Skorokhod metric d, which makes it Polish (see [Bas11, Section 34]). By definition of d, we have for any $f, g \in D_T$

$$d(f,g) \leq \|f - g\|_\infty := \sup_{t\in[0,T]} |f_t - g_t|.$$

The previous inequality becomes an equality if $g = 0$. We also denote by $\mathcal{P}_\beta(D_T)$ the space of probability measures $\mu \in \mathcal{P}(D_T)$ such that

$$\int_{D_T} d(f,0)^\beta \, d\mu(f) = \int_{D_T} \|f\|_\infty^\beta \, d\mu(f) < +\infty.$$

It is endowed with the Wasserstein metric of order β defined, for any $\mu, \nu \in \mathcal{P}_\beta(D_T)$, by

$$\mathcal{W}_\beta(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \int_{D_T\times D_T} d^\beta(f,g) \, d\pi(f,g),$$

where $\Pi(\mu,\nu)$ denotes the set of probability measure on $D_T \times D_T$ having μ and ν as marginal distributions. For any fixed $t \in [0,T]$, we define the projection $\pi_t : f \in D_T \mapsto f_t \in \mathbb{R}^d$. It is a measurable function so that if μ belongs to $\mathcal{P}(D_T)$, we can define $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ as the push-forward measure of μ by π_t. Notice that if $\mu \in \mathcal{P}_\beta(D_T)$, the function $t \in [0,T] \mapsto \mu_t$ belongs to $D([0,T]; \mathcal{P}_\beta(\mathbb{R}^d))$. Let us fix $\mu \in \mathcal{P}_\beta(D_T)$. Reasoning as in the proof of Proposition 6.1, the standard SDE

$$\begin{cases} dX_t^\mu = b_t(X_t^\mu, \mu_t) \, dt + \sigma_t(X_{t-}^\mu, \mu_t) \, dZ_t, \quad t \in [0,T], \\ X_0^\mu = \xi \in L^\beta(\Omega, \mathcal{F}_0) \end{cases} \tag{6.36}$$

admits a unique strong solution X^μ such that

$$\mathbb{E} \sup_{t \le T} |X_t^\mu|^\beta < +\infty.$$

The following function is thus well-defined

$$\phi : \left\{ \begin{array}{rl} \mathcal{P}_\beta(D_T) & \to \mathcal{P}_\beta(D_T) \\ \mu & \mapsto [(X_t^\mu)_{t\in[0,T]}]. \end{array} \right. \tag{6.37}$$

The goal is to prove that ϕ has at least one fixed point using Schauder's fixed point theorem. By the estimation (6.20) obtained in the proof of Theorem 6.1, we have

$$\mathcal{W}_\beta(\phi(\mu), \phi(\nu)) \le \mathbb{E}\left(\sup_{0\le s\le T} |X_s^\mu - X_s^\nu|^\beta \right) \le C \left(\int_0^T W_\beta^2(\mu_s, \nu_s)\, ds \right)^{\frac{\beta}{2}}. \tag{6.38}$$

Let us show that this implies the continuity of ϕ. Consider $(\mu^n)_n \in \mathcal{P}_\beta(D_T)$ a sequence which converges towards $\mu \in \mathcal{P}_\beta(D_T)$ with respect to $\mathcal{W}_\beta$. For almost all $t \in [0,T]$, the sequence $(\mu_t^n)_n$ converges in distribution to μ_t (see [Bil99, Section 13]). Let us fix such a $t \in [0,T]$. We aim at proving that the previous convergence holds true with respect to W_β. By [Vil09, Definition 6.8], it is enough to prove that

$$\lim_{R\to+\infty} \sup_n \int_{|x|\ge R} |x|^\beta d\mu_t^n(x) = 0.$$

But since

$$\begin{aligned} \int_{|x|\ge R} |x|^\beta d\mu_t^n(x) &= \int_{|f_t|\ge R} |f_t|^\beta d\mu^n(f) \\ &\le \int_{d(f,0)\ge R} d(f,0)^\beta d\mu^n(f), \end{aligned}$$

we conclude using that $\mu^n \xrightarrow{\mathcal{W}_\beta} \mu$. Thus, for almost all $t \in [0,T]$, $(\mu_t^n)_n$ converges to μ_t with respect to W_β. Coming back to (6.38), and using the dominated convergence theorem justified since

$$\sup_{n\ge 1} \int_{D_T} \|f\|_\infty^\beta \, d\mu^n(f) < +\infty,$$

we conclude that ϕ is continuous. Following the same lines as to prove (6.20), we show that for some constant $C = C_T > 0$

$$\mathbb{E} \sup_{t\le T} |X_t^\mu|^\beta \le C \left[1 + \left(\int_0^T M_\beta^2(\mu_s)\, ds \right)^{\frac{\beta}{2}} \right]. \tag{6.39}$$

Let us define, for $R > 0$,

$$\overline{\mathcal{B}}_R := \left\{ \mu \in \mathcal{P}_\beta(D_T), \int_{D_T} \|f\|_\infty^\beta \, d\mu(f) \le R \right\}.$$

This is a closed and convex subset of $\mathcal{P}_\beta(D_T)$, which is stable by ϕ for R large enough owing to (6.39) and since $\beta \in (0,1)$. In the following, we fix $R > 0$ such that $\phi(\overline{\mathcal{B}}_R) \subset \overline{\mathcal{B}}_R$. It remains to prove that $\phi(\overline{\mathcal{B}}_R)$ is relatively compact in $\mathcal{P}_\beta(D_T)$ to conclude that ϕ admits a fixed point by Schauder's theorem. Let us fix $(\mu^n)_n$ a sequence of $\overline{\mathcal{B}}_R$. In a first step, we prove with Aldou's criterion (see [Bas11, Theorem

34.8]) that $([(X_t^{\mu^n})_{t\in[0,T]}])_n$ is tight, and thus relatively compact in $\mathcal{P}(D_T)$. For $t \in [0,T]$ and $A > 0$, we have

$$\begin{aligned}\mathbb{P}(|X_t^{\mu^n}| \geq A) &\leq \frac{1}{A^\beta}\mathbb{E}|X_t^{\mu^n}|^\beta \\ &\leq \frac{R}{A^\beta}.\end{aligned}$$

This yields for all $t \in [0,T]$

$$\lim_{A\to+\infty} \sup_n \mathbb{P}(|X_t^{\mu^n}| \geq A) = 0. \tag{6.40}$$

Let $(\tau_n)_n$ be a sequence of stopping times and $(\delta_n)_n$ a sequence of real numbers converging to 0. We assume that $\tau_n \leq T$ and $0 \leq \tau_n + \delta_n \leq T$ for any $n \geq 1$. It remains to prove that for any fixed $\epsilon > 0$

$$\mathbb{P}(|X_{\tau_n+\delta_n}^{\mu^n} - X_{\tau_n}^{\mu^n}| \geq \epsilon) \underset{n\to+\infty}{\longrightarrow} 0.$$

For $A > 0$ which will be chosen latter, we set

$$T_A^n := \inf\left\{t \leq T,\, |X_t^{\mu^n}| \geq A\right\}.$$

Markov's inequality yields

$$\mathbb{P}(|X_{\tau_n+\delta_n}^{\mu^n} - X_{\tau_n}^{\mu^n}| \geq \epsilon) \leq \mathbb{P}(T_A^n \leq T) + \frac{1}{\epsilon^\beta}\mathbb{E}(|X_{\tau_n+\delta_n}^{\mu^n} - X_{\tau_n}^{\mu^n}|^\beta \mathbf{1}_{T_A^n \geq T}). \tag{6.41}$$

Notice first that

$$\begin{aligned}\mathbb{P}(T_A^n \leq T) &\leq \mathbb{P}(\sup_{t\leq T}|X_t^{\mu^n}|^\beta \geq A^\beta) \\ &\leq \frac{R}{A^\beta}.\end{aligned} \tag{6.42}$$

Then, by the triangle inequality, we have

$$\begin{aligned}\mathbb{E}(|X_{\tau_n+\delta_n}^{\mu^n} - X_{\tau_n}^{\mu^n}|^\beta \mathbf{1}_{T_A^n \geq T}) &\leq \mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n} b_s(X_s^{\mu^n}, \mu_s^n)\, ds\right|^\beta \mathbf{1}_{T_A^n \geq T}\right) \\ &\quad + \mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n}\int_{B_1} \sigma_s(X_{s^-}^{\mu^n}, \mu_s^n) z \widetilde{\mathcal{N}}(ds,dz)\right|^\beta \mathbf{1}_{T_A^n \geq T}\right) \\ &\quad + \mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n}\int_{B_1^c} \sigma_s(X_{s^-}^{\mu^n}, \mu_s^n) z \mathcal{N}(ds,dz)\right|^\beta \mathbf{1}_{T_A^n \geq T}\right) \\ &=: I_1 + I_2 + I_3\end{aligned}$$

We now estimate I_1, I_2 and I_3. Using the linear growth assumption on b, we have for a constant $C > 0$ independent of n

$$\begin{aligned}I_1 &\leq \mathbb{E}\left|\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n} C(1 + |X_s^{\mu^n}| + M_\beta(\mu_s^n))\, ds\right|^\beta \\ &\leq C|\delta_n|^\beta(1 + A^\beta + R^\beta).\end{aligned}$$

Thanks to BDG's and Jensen's inequalities, we obtain that

$$
\begin{aligned}
I_2 &\leq \mathbb{E}\left(\left|\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1}\sigma_s(X_{s-}^{\mu^n},\mu_s^n)z\widetilde{\mathcal{N}}(ds,dz)\right|^\beta\right)\\
&\leq C\mathbb{E}\left(\left|\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1}|\sigma_s(X_{s-}^{\mu^n},\mu_s^n)|^2|z|^2\mathcal{N}(ds,dz)\right|^{\frac{\beta}{2}}\right)\\
&\leq C(1+A^2+R^2)^{\frac{\beta}{2}}\left(\mathbb{E}\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1}|z|^2\,d\nu(z)\,ds\right)^{\frac{\beta}{2}}\\
&\leq C(1+A^\beta+R^\beta)|\delta_n|^{\frac{\beta}{2}}.
\end{aligned}
$$

Since $\beta<1$, the subadditivity of the map $|\cdot|^\beta$ yields

$$
\begin{aligned}
I_3 &\leq \mathbb{E}\left(\left|\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1^c}\sigma_s(X_{s-}^{\mu^n},\mu_s^n)z\mathcal{N}(ds,dz)\right|^\beta\right)\\
&\leq \mathbb{E}\left(\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1^c}|\sigma_s(X_{s-}^{\mu^n},\mu_s^n)|^\beta|z|^\beta\mathcal{N}(ds,dz)\right)\\
&\leq C(1+A^\beta+R^\beta)\left(\mathbb{E}\int_{\tau_n\wedge T_A^n}^{(\tau_n+\delta_n)\wedge T_A^n}\int_{B_1^c}|z|^\beta\,d\nu(z)\,ds\right)\\
&\leq C(1+A^\beta+R^\beta)|\delta_n|.
\end{aligned}
$$

Using (6.41), (6.42), and the upper-bounds obtained previously for I_1, I_2 and I_3, we deduce that

$$
\mathbb{P}(|X_{\tau_n+\delta_n}^{\mu^n}-X_{\tau_n}^{\mu^n}|\geq\epsilon)\leq\frac{R}{A^\beta}+\frac{C}{\epsilon^\beta}(1+A^\beta+R^\beta)(|\delta_n|+|\delta_n|^\beta+|\delta_n|^{\frac{\beta}{2}}).
$$

Since R is fixed, we can choose A large enough and then let n tend to $+\infty$ to obtain that $(X_{\tau_n+\delta_n}^{\mu^n}-X_{\tau_n}^{\mu^n})_n$ converges in probability to 0. Thus, $([(X_t^{\mu^n})_{t\in[0,T]}])_n$ is relatively compact in $\mathcal{P}(D_T)$. The relative compactness in $\mathcal{P}_\beta(D_T)$ follows from the fact that

$$
\sup_{\mu\in\overline{\mathcal{B}}_R}\mathbb{E}\sup_{t\leq T}|X_t^\mu|^{\beta+\delta}<+\infty.
$$

Indeed, this is a consequence of [Fou13, Proposition 2], since for all $t\in[0,T]$, $\mu\in\overline{\mathcal{B}}_R$ and $x\in\mathbb{R}^d$

$$
|b_t(x,\mu_t)|+|\sigma_t(x,\mu_t)|\leq C(1+|x|+R),
$$

and

$$
\int_{B_1^c}|z|^{\beta+\delta}\,d\nu(z)<+\infty.
$$

We have proved that $\phi(\overline{\mathcal{B}}_R)$ is relatively compact in $\mathcal{P}_\beta(D_T)$. Thus, Schauder's fixed point theorem yields the existence of a solution to (6.1). The moment estimation (6.35) directly follows from [Fou13, Proposition 2]. □

Itô's formula for the flow of measures of Poisson stochastic integrals and applications

This chapter corresponds to the article [Cav22a]. It has been submitted for publication.

Abstract. We prove Itô's formula for the flow of measures associated with a jump process defined by a drift, an integral with respect to a Poisson random measure and with respect to the associated compensated Poisson random measure. We work in $\mathcal{P}_\beta(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$ having a finite moment of order $\beta \in (0,2]$. As an application, we exhibit the backward Kolmogorov partial differential equation stated on $[0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$ associated with a McKean-Vlasov stochastic differential equation driven by a Poisson random measure. It describes the dynamics of the semigroup acting on functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$ associated with the McKean-Vlasov stochastic differential equation, under regularity assumptions on it. Finally, we use the semigroup and the backward Kolmogorov equation to prove new quantitative weak propagation of chaos results for a mean-field system of interacting Ornstein-Uhlenbeck processes driven by i.i.d. α-stable processes with $\alpha \in (1,2)$.

7.1 Introduction

Let us fix $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space satisfying the usual conditions and $\mathcal{N}$ a Poisson random measure on $[0,T] \times \mathbb{R}^d$, for some finite horizon of time $T > 0$, with intensity measure $dt \otimes \nu$, where ν is a Lévy measure, i.e. $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{R}^d} \min(|z|^2, 1)\, d\nu(z) < +\infty.$$

We consider the jump process $X = (X_t)_{t\in[0,T]}$ with values in $\mathbb{R}^d$ defined, for all $t \in [0,T]$, by

$$X_t := X_0 + \int_0^t b_s\, ds + \int_0^t \int_{\{|z|<1\}} H(s,z)\, \widetilde{\mathcal{N}}(ds,dz) + \int_0^t \int_{\{|z|\geq 1\}} K(s,z)\, \mathcal{N}(ds,dz), \tag{7.1}$$

where $\widetilde{\mathcal{N}}(ds,dz) := \mathcal{N}(ds,dz) - ds\, d\nu(z)$ is the associated compensated Poisson random measure, b, H, K are predictable processes and X_0 is $\mathcal{F}_0$-measurable. The distribution of X_t is denoted by μ_t. The assumptions made on X_0, ν, b, H and K (see Section 7.2) ensure that for all $t \in [0,T]$, μ_t belongs to $\mathcal{P}_\beta(\mathbb{R}^d)$, the space of probability measures having a finite moment of order β, for some $\beta \in (0,2]$.

In this work, we are interested in Itô's formula for the flow of probability measures $(\mu_t)_{t\in[0,T]}$. This formula describes the dynamics of the map $t \in [0,T] \mapsto u(\mu_t)$, for some function $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$. Itô's formula for a flow of probability measures has been developed over the last decade to deal with mean-field games, which were initiated independently by Caines, Huang and Malhame in [CHM06] and by Lasry and Lions in [LL07]. An important object introduced by Lions in his lectures at Collège de France [Lio] is the master equation related to a mean-field game, which is a Partial Differential Equation (PDE) stated on the space of probability measures. We refer to [Lio, Car10, CD18a, CD18b] for more details on mean-field games and associated master equations. As for the standard Itô formula, Itô's formula along a flow of probability measures allows to associate a PDE stated on the space of probability measures to a McKean-Vlasov Stochastic Differential Equation (SDE), which is the related master equation in this context. The well-posedness of these equations was studied for example in [CCD15, BLPR17, CDLL19, CM17, CdRF22]. Moreover, the problem of propagation of chaos for a mean-field interacting particle system towards the corresponding McKean-Vlasov SDE can also be tackled using the associated PDE on the space of measures. It turns out to be an efficient tool to prove the quantitative weak propagation of chaos of the empirical measure $(\overline{\mu}_t^N)_t$ of the particle system towards the flow of marginal distributions of the solution to the McKean-Vlasov SDE $(\mu_t)_t$. More precisely, a quantitative weak propagation result consists in finding an explicit rate of convergence with respect to N for $\mathbb{E}|u(\overline{\mu}_T^N) - u(\mu_T)|$ and $|\mathbb{E}u(\overline{\mu}_T^N) - \mathbb{E}u(\mu_T)|$ and for u belonging to a certain class of functions defined on the space of probability measures. This strategy was originally described in Chapter 5 of [CD18a]. It was adopted for example by Chaudru de Raynal and Frikha in [CdRF22, CdRF21], by Chassagneux, Szpruch and Tse in [CST22] and by Delarue and Tse in [DT21]. Let us also mention that this PDE on the space of probability measures has recently been used in [JT21] to prove a central limit theorem for interacting particle systems. Finally, Itô's formula for a flow of measures is also a key tool to tackle McKean-Vlasov control problems. Indeed, it induces a dynamic programming principle which describes the value function of the problem as presented in [CD15] or in Chapter 6 of [CD18a] (see also the references therein).

In the first part of this work, we focus on Itô's formula for the flow of measures $(\mu_t)_t$ associated with (7.1) and for a function $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ with $\beta \in (0,2]$. It states that if u is regular enough, we have for all $t \in [0,T]$

$$
\begin{aligned}
&u(\mu_t) - u(\mu_0)\\
&= \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s)\cdot b_s\right) ds\\
&\quad + \int_0^t \int_{|z|\geq 1} \mathbb{E}\left[\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right] d\nu(z)\, ds\\
&\quad + \int_0^t \int_{|z|<1} \mathbb{E}\left[\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}) - \partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\cdot H(s,z)\right] d\nu(z)\, ds,
\end{aligned}
$$

where $\frac{\delta}{\delta m} u$ denotes the linear (functional) derivative of u (see Definition 7.1). The precise assumptions are given in Section 7.2 (see Theorem 7.2). In the Brownian setting, Itô's formula for a flow of measures is proved for example in [BLPR17] (see Theorem 6.1) or in Section 3 of [CCD15] and Chapter 5 of [CD18a] (see Theorem 5.99) under less restrictive assumptions. Let us also mention [dRP22, Cav21] for some recent extensions of Itô's formula for a flow of measures in the Brownian case. Concerning jump processes, Itô's formula has recently been extended to flows of measures generated by càdlàg semi-martingales. It was established independently by Guo, Pham and Wei in [GPW20], who studied McKean-Vlasov control problems with jumps and by Talbi, Touzi and Zhang in [TTZ21] who worked on mean-field optimal stopping problems. A common point between these two works is that the jump process considered has a finite moment of order 2. However, this framework is not adapted when the Poisson random measure $\mathcal{N}$ stems from an α-stable Lévy process with $\alpha \in (0,2)$ since it only has finite moments of order $\beta < \alpha$. Itô's formula given in Theorem 7.2 can be applied for these processes since $\beta \in (0,2]$. Another difference between this work and [GPW20] is that we do not assume $\partial_v \frac{\delta}{\delta m} u$ to be uniformly bounded if $\beta > 1$. Moreover, the authors of [TTZ21] assume that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the function $\frac{\delta}{\delta m} u(\mu)$ is of class $\mathcal{C}^2$ on $\mathbb{R}^d$, which is not the case in this paper. It is replaced here by the γ-Hölder continuity of $\partial_v \frac{\delta}{\delta m} u(\mu)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, where γ is such that $x \mapsto |x|^{1+\gamma} \in L^1(\{|z|<1\}, \nu)$. The strategy of the proof is the following. First, we localize $X = (X_t)_t$ using stopping times and we precise in Proposition 7.13 in which sense the sequence of stopped processes $(X^n)_n$ approximates X. Then, we establish Itô's formula for the flow of measures associated with the stopped process $(X^n_t)_t$ thanks to a standard method of time discretization. We finally let n tend to infinity essentially by using the approximation results of Proposition 7.13. Moreover, we extend Itô's formula for functions depending also on the time and space variables in Theorem 7.8.

Then, we consider a general Lévy-driven McKean-Vlasov SDE, which is assumed to be well-posed in the weak sense, of the form

$$\begin{cases} dX_t = b_t(X_t, \mu_t)\, dt + \sigma_t(X_{t^-}, \mu_t)\, dZ_t, \quad t \in [0,T], \\ \mu_t = [X_t], \\ X_0 = \xi \in L^\beta(\Omega, \mathcal{F}_0), \end{cases} \tag{7.2}$$

where $[X_t]$ denotes the distribution of the random variable X_t, $\beta \in (0,2]$. The coefficients b and σ are assumed to be at most of linear growth when $\beta \geq 1$ or uniformly bounded when $\beta < 1$ and $Z = (Z_t)_t$ is a Lévy process on $\mathbb{R}^d$ having the following Lévy-Itô decomposition

$$Z_t = \int_0^t \int_{\{|z|<1\}} z \widetilde{\mathcal{N}}(ds, dz) + \int_0^t \int_{\{|z|\geq 1\}} z \mathcal{N}(ds, dz),$$

and a finite moment of order β. Our aim is to describe the dynamics of the semigroup acting on functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$ associated with the McKean-Vlasov SDE (7.2). For a fixed function $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$, the action of the semigroup on u is given by the map ϕ_u defined by

$$\phi_u : \begin{cases} [0,T] \times \mathcal{P}_\beta(\mathbb{R}^d) & \to \mathbb{R} \\ (t, \mu) & \mapsto u([X_T^{T-t,\mu}]), \end{cases} \tag{7.3}$$

where $[X_T^{T-t,\mu}]$ denotes the distribution of the solution to (7.2) at time T starting at time $T-t$ from a random variable ξ with distribution μ. As in [CD18a], we prove using Itô's formula for a flow of measures that, under regularity assumptions on ϕ_u, the function $(t,\mu) \in [0,T] \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \phi_u(T-t, \mu)$ is a classical solution to a backward Kolmogorov PDE on $[0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$ (see Theorem 7.16).

Finally, we prove new quantitative weak propagation of chaos results, in the sense previously defined, for a mean-field system of interacting stable Ornstein-Uhlenbeck processes defined as follows. We assume that the driving process $Z = (Z_t)_t$ is an α-stable process on $\mathbb{R}^d$ with $\alpha \in (1,2)$ and with a non degenerate

Lévy measure (see (7.50)). More precisely, we introduce $(Z^n)_n$ an i.i.d. sequence of stable processes having the same distribution as $Z = (Z_t)_t$, A, A', B matrices of size $d \times d$ such that B is invertible and $(X_0^n)_n$ an i.i.d. sequence of random variables with common distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, for some $\beta \in [1, \alpha)$. The particle system is the unique solution to the following classical SDE on $(\mathbb{R}^d)^N$, for $N \geq 1$

$$\begin{cases} dX_t^{i,N} = AX_t^{i,N}\, dt + A' \dfrac{1}{N} \displaystyle\sum_{j=1}^N X_t^{j,N}\, dt + B\, dZ_t^i, \quad t \in [0,T], \quad i \in \{1, \ldots, N\}, \\ X_0^{i,N} = X_0^i. \end{cases} \tag{7.4}$$

We denote by $X = (X_t)_t$ the solution to the corresponding McKean-Vlasov SDE

$$\begin{cases} dX_t = (AX_t + A'\mathbb{E}X_t)\, dt + B\, dZ_t, \quad t \in [0,T], \\ X_0 = \xi, \end{cases} \tag{7.5}$$

where the distribution of ξ is μ_0.

The propagation of chaos phenomenon was originally studied by McKean in [McK67] and later by Sznitman in [Szn91]. It states that for any $k \geq 1$, the limiting behaviour of $(X^{1,N}, \ldots, X^{k,N})$, when N tends to infinity, is expected to be described by k independent copies of the McKean-Vlasov process (7.5). It can be shown in the weak sense i.e. in distribution or in the strong sense i.e. at the level of paths by coupling. It has been of course widely studied in the Brownian case. Let us focus on some works dealing with propagation of chaos for mean-field systems with jumps. In [Gra92a], following the approach of Sznitman [Szn91] in the Brownian case, Graham proves weak propagation of chaos under Lipschitz assumptions on the coefficients for a mean-field system driven by a Poisson random measure and its compensated measure. He works in the L^1 framework i.e. the Poisson random measure is associated with a Poisson process having a finite moment of order 1. The set of jumps of the noise is also assumed to be discrete, which is not the case for stable processes for example. In the case of a McKean-Vlasov SDE driven by a general Lévy process having a finite moment of order 2, we refer to Jourdain, Méléard and Woyczynski [JMW07]. The authors prove rates of convergence for the strong propagation of chaos in L^2 under standard Lipschitz assumptions on the coefficients. Moreover, it has also been established in [NBK+20] by Neelima et al. under relaxed assumptions. We also mention [MMW15] where Mischler, Mouhot and Wennberg exhibit conditions yielding rates of convergence for the propagation of chaos. As an application, they study an inelastic Boltzmann collision jump process. Finally, let us also refer to [FL21], where Frikha and Li study a one-dimensional McKean-Vlasov SDE driven by a compensated Poisson random measure with positive jumps. They prove explicit rates of convergence for the strong propagation of chaos in L^1 under one-sided Lipschitz assumptions.

Concerning our mean-field system of interacting stable Ornstein-Uhlenbeck processes (7.4), we are interested in quantitative weak propagation of chaos. Let us denote by $\overline{\mu}_t^N$ the empirical measure of the particle system (7.4) and by μ_t the distribution of X_t. We prove in Theorem 7.20 that there exists a constant $C > 0$ such that for any $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ in a certain class of regular functions described in Theorem 7.20 and $N \geq 1$, one has

$$\mathbb{E}\left|u(\overline{\mu}_T^N) - u(\mu_T)\right| \leq C\, \mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + C \ln(N)^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, \tag{7.6}$$

where W_1 is the Wasserstein metric of order 1. We also show that if $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, one has

$$|\mathbb{E}(u(\overline{\mu}_T^N) - u(\mu_T))| \leq CN^{1-\alpha}. \tag{7.7}$$

As a consequence of (7.7), if $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, one has the following mean-field limit estimate

$$\sup_{t\in[0,T]} W_1([X_t^{1,N}], \mu_t) \leq \frac{C_T}{N^{\alpha-1}}. \tag{7.8}$$

Of course, the rate of convergence is better at the level of semigroup i.e. in (7.7) than in (7.6). Notice that if we formally take $\alpha = 2$, which corresponds to the Brownian case treated in [CdRF22, CdRF21], we recover the same rates of convergence, up to the factor $\ln(N)$ in (7.6), as in Theorem 3.6 of [CdRF21], even though the drift is unbounded here. In dimension $d = 1$, we recover with (7.6) the same rate of convergence proved in [FL21], for the strong propagation of chaos in L^1, since $\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) \leq CN^{\frac{1}{\beta}-1}$ by [FG15]. Notice that if μ_0 belongs to $\mathcal{P}_2(\mathbb{R})$, then using [FG15], we have a better rate of propagation of chaos since (7.6) becomes

$$\mathbb{E}\left|u(\overline{\mu}_T^N) - u(\mu_T)\right| \leq C\,\ln(N)^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}.$$

In [Cav23], quantitative strong propagation of chaos is proved for (7.4). We refer to Remark 7.22 for a precise comparison of the rates of convergence. In particular, the estimate (7.8) is better than the one that can be deduced from [Cav23, Theorem 3].

The method that we use relies on regularity properties and estimates on the solution to the backward Kolmogorov PDE ϕ_u defined in (7.3) (see Proposition 7.24). This strategy was originally described in Chapter 5 of [CD18a] (pages 506 – 508), inspired by [CDLL19] and [MMW15], and was employed for example in [CST22, DT21, CdRF21], as mentioned above. Let us describe the main ideas. We begin by computing the time derivative of the map $t \in [0,T) \mapsto \phi_u(T-t, \overline{\mu}_t^N)$ by applying the standard Itô's formula for the empirical projection $(t, x_1, \dots, x_N) \in [0,T] \times (\mathbb{R}^d)^N \mapsto \phi_u\left(T-t, \frac{1}{N}\sum_{k=1}^N \delta_{x_k}\right)$ and for the particle system. Noting that the map $t \in [0,T] \mapsto \phi_u(T-t, \mu_t)$ is constant, we naturally expect that the time derivative previously computed tends to 0 as N converges to infinity. This convergence is shown with an explicit rate of convergence using the PDE satisfied by ϕ_u and some estimates on ϕ_u. Finally, we express $u(\overline{\mu}_T^N) - u(\mu_T) = \phi_u(0, \overline{\mu}_T^N) - \phi_u(0, \mu_T)$ as the sum of $\phi_u(T, \overline{\mu}_0^N) - \phi_u(T, \mu_0)$ plus a remainder term related to the time derivative previously estimated. Since the initial data are i.i.d., the first term is controlled by standard estimates, for example those in [FG15].

An important difference between the jump case in comparison with the Brownian case is that the Kolmogorov PDE satisfied by ϕ_u does not directly appear when we apply Itô's formula for the empirical projection of ϕ_u and for the particle system. In order to use the backward Kolmogorov PDE, we thus have to control the corresponding error term. To do this and because of the unboundedness of the drift, we need to remove the big jumps of the driving processes in a first step. The key point is to consider the solutions to (7.4) and (7.5) driven by truncated noises for which we remove the jumps bigger than the number of particles N. We need to perform this truncation procedure to control the error term mentioned above. This is essentially because the moment of order 2 of the Lévy measure ν appears in our computations due to the unboundedness of the drift. We refer to Remark 7.23 for more details. The presence of the factor $\ln(N)$ in (7.6) comes from this preliminary step.

Let us eventually emphasize that this mean-field system of interacting Ornstein-Uhlenbeck processes is a first application. We aim at showing that this method, already used in the Brownian setting, and leading to quantitative weak propagation of chaos results, can be generalized in the context of jump processes. Our next goal is to follow the same strategy to establish quantitative weak propagation of chaos results for a general class of Lévy-driven McKean-Vlasov SDEs, under mild regularity assumptions on the coefficients, as done in [CdRF21].

The paper is organized as follows. In Section 7.2, we state and prove Itô's formula in Theorem 7.2 and Theorem 7.8 for the flow of measure of Poisson stochastic integrals. In Section 7.3, and more precisely in Theorem 7.16, we use our Itô's formula to derive the backward Kolmogorov PDE on the space of measures describing the dynamics of the semigroup associated with a general Lévy-driven McKean-Vlasov SDE. Then, in Section 7.4, we study the mean-field system of interacting stable Ornstein-Uhlenbeck processes and we prove in Theorem 7.20 quantitative weak propagation of chaos. Appendix 7.5 is devoted to the proof of moment estimates, which are uniform with respect to the truncation of the big jumps, on the density and its derivatives of a stable Ornstein-Uhlenbeck process. Finally, in Appendix 7.6, we prove the regularity properties and the controls required in Section 7.4 on the solution to the backward Kolmogorov PDE associated with (7.5) driven by the truncated stable process mentioned above.

Let us finally introduce some notations used several times in the article.

Notations

- $\mathcal{P}_\beta(\mathbb{R}^d)$ denotes the set of probability measures μ on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |x|^\beta \, d\mu(x) < +\infty$, for $\beta > 0$. It is equipped with the Wasserstein metric of order β denoted by W_β, which makes it complete. Denoting by $\Pi(\mu,\nu)$ the set of couplings between two probability measures $\mu,\nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the metric W_β is defined by

$$W_\beta(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^\beta \, d\pi(x,y) \right)^{\min\left(\frac{1}{\beta},1\right)}.$$

- $\overline{\mu}^N_{\boldsymbol{x}} := \frac{1}{N}\sum_{k=1}^N \delta_{x_k}$ denotes the empirical measure, for $\boldsymbol{x} = (x_1,\dots,x_N) \in (\mathbb{R}^d)^N$.
- $\tilde{\boldsymbol{z}}_{\boldsymbol{k}} := (0,\dots,z,\dots,0) \in (\mathbb{R}^d)^N$ for $z \in \mathbb{R}^d$, where z appears in the k-th position.
- B_r is the open ball in $\mathbb{R}^d$ centered at 0 and of radius r for the euclidean norm.
- B_r^c denotes the complementary of B_r.
- $\mathbf{1}_A$ denotes the indicator function of some measurable set A.
- p' is the conjugate exponent of $p \in [1,\infty]$ defined by $\frac{1}{p} + \frac{1}{p'} = 1$.
- $a \wedge b$ denotes the minimum between a and b.
- $a \vee b$ denotes the maximum between a and b.
- C is a generic constant that may depend only on the fixed parameters of the problem and which may change from line to line.

7.2 Itô's formula for the flow of measures of Poisson stochastic integrals

7.2.1 Assumptions and Itô's formula for a flow of measures

Let ν be a Lévy measure on $\mathbb{R}^d$, i.e. $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{R}^d} |z|^2 \wedge 1 \, d\nu(z) < +\infty.$$

We introduce a $(\mathcal{F}_t)_{t\in[0,T]}$-Poisson random measure $\mathcal{N}$ on $[0,T] \times \mathbb{R}^d \backslash \{0\}$ with intensity measure $dt \otimes \nu$, and we denote by $\widetilde{\mathcal{N}}(ds, dz) := \mathcal{N}(ds, dz) - ds\, d\nu(z)$ the compensated Poisson random measure associated with $\mathcal{N}$. Let us define the jump process $X = (X_t)_{t\in[0,T]}$ by

$$\forall t \in [0,T],\; X_t := X_0 + \int_0^t b_s \, ds + \int_0^t \int_{B_1} H(s,z)\, \widetilde{\mathcal{N}}(ds, dz) + \int_0^t \int_{B_1^c} K(s,z)\, \mathcal{N}(ds, dz), \tag{7.9}$$

where $b : [0,T] \times \Omega \to \mathbb{R}^d$, $H : [0,T] \times B_1 \times \Omega \to \mathbb{R}^d$, and $K : [0,T] \times B_1^c \times \Omega \to \mathbb{R}^d$ are predictable processes. The distribution of X_t is denoted by μ_t.

Let us fix our assumptions on $(X_t)_t$ in order to establish Itô's formula for the flow $(\mu_t)_t$. We assume that there exists two constants $\beta \in (0,2]$ and $\gamma \in [0,1]$ such that $\beta \leq 1 + \gamma$ and satisfying the following properties.

(M) The random variable X_0 belongs to $L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d)$ and we have

$$\mathbb{E} \int_0^T |b_s|^{\beta \vee 1} \, ds < +\infty. \tag{7.10}$$

(J1) There exists a predictable process $(\widetilde{H}_s)_{s\in[0,T]}$ which is assumed to be almost surely locally bounded and to satisfy

$$\text{a.s.}\, \forall s \in [0,T],\, \forall z \in B_1,\, |H(s,z)| \leq |\widetilde{H}_s||z|, \quad \text{and} \quad \mathbb{E} \int_0^T \int_{B_1} (|\widetilde{H}_s||z|)^{1+\gamma} \, d\nu(z)\, ds < +\infty. \tag{7.11}$$

(J2) We have

$$\mathbb{E} \int_0^T \int_{B_1^c} |K(s,z)|^\beta \, d\nu(z)\, ds < +\infty. \tag{7.12}$$

Definition 7.1 (Linear derivative). A function $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ is said to have a linear derivative if there exists a function $\frac{\delta}{\delta m} u \in \mathcal{C}^0(\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d; \mathbb{R})$ satisfying the following properties.

1. For all compact subset $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists a constant $C_\mathcal{K} > 0$ such that

$$\forall \mu \in \mathcal{K},\, \forall v \in \mathbb{R}^d,\, \left| \frac{\delta}{\delta m} u(\mu)(v) \right| \leq C_\mathcal{K}(1 + |v|^\beta).$$

2. For all $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, we have

$$u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(t\mu + (1-t)\nu)(v)\, d(\mu - \nu)(v)\, dt.$$

We now state Itô's formula along the flow of measures $(\mu_t)_{t\in[0,T]}$.

Theorem 7.2 (Itô's formula)**.** *Suppose that Assumptions* $(\boldsymbol{M})$, ***(J1)****, and* ***(J2)*** *are satisfied. Let* $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ *be a function having a linear derivative* $\frac{\delta}{\delta m}u$ *which satisfies the following properties.*

1. *For any* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, *the function* $\frac{\delta}{\delta m}u(\mu) \in \mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ *and* $\partial_v \frac{\delta}{\delta m}u$ *is continuous on* $\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$.
2. *If* $\gamma > 0$, *for any compact* $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, *there exists* $C_\mathcal{K} > 0$ *such that*
$$\forall \mu \in \mathcal{K},\, \forall x, y \in \mathbb{R}^d,\, \left|\partial_v \frac{\delta}{\delta m}u(\mu)(x) - \partial_v \frac{\delta}{\delta m}u(\mu)(y)\right| \le C_\mathcal{K}|x-y|^\gamma.$$
3. *For any compact* $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, *we have*
$$\begin{cases} \displaystyle\sup_{\mu\in\mathcal{K}} \int_{\mathbb{R}^d} \left|\partial_v \frac{\delta}{\delta m}u(\mu)(v)\right|^{\beta'} d\mu(v) < +\infty, & \text{if } \beta > 1, \\ \displaystyle\sup_{v\in\mathbb{R}^d}\sup_{\mu\in\mathcal{K}} \left|\partial_v \frac{\delta}{\delta m}u(\mu)(v)\right| < +\infty, & \text{if } \beta \le 1, \end{cases}$$

where β' *is the conjugate exponent of* β *when* $\beta \ge 1$. *Then, we have for all* $t \in [0,T]$

$$\begin{aligned} &u(\mu_t) - u(\mu_0) \\ &= \int_0^t \mathbb{E}\left(\partial_v \frac{\delta}{\delta m}u(\mu_s)(X_s)\cdot b_s\right) ds \qquad (7.13)\\ &+ \int_0^t \int_{B_1^c} \mathbb{E}\left[\frac{\delta}{\delta m}u(\mu_s)(X_{s^-} + K(s,z)) - \frac{\delta}{\delta m}u(\mu_s)(X_{s^-})\right] d\nu(z)\, ds \\ &+ \int_0^t \int_{B_1} \mathbb{E}\left[\frac{\delta}{\delta m}u(\mu_s)(X_{s^-} + H(s,z)) - \frac{\delta}{\delta m}u(\mu_s)(X_{s^-}) - \partial_v \frac{\delta}{\delta m}u(\mu_s)(X_{s^-})\cdot H(s,z)\right] d\nu(z)\, ds. \end{aligned}$$

Remark 7.3. Assumption (3) is implied by the following stronger assumption: for all compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that

$$\forall v \in \mathbb{R}^d,\, \sup_{\mu\in\mathcal{K}} \left|\partial_v \frac{\delta}{\delta m}u(\mu)(v)\right| \le C_\mathcal{K}(1 + |v|^{\beta-1}\mathbf{1}_{\beta>1}).$$

When $\beta > 1$, it follows from the fact that $\beta'(\beta-1) = \beta$.

We specify in the three following corollaries the particular case where the Poisson random measure is associated with an α-stable Lévy process $Z = (Z_t)_t$ with $\alpha \in (0,2)$. Consider $\sigma : [0,T]\times\Omega \to \mathbb{R}^{d\times d}$ a predictable process such that

$$\mathbb{E}\int_0^T |\sigma_s|^{1+\gamma}\, ds < +\infty,$$

for some $\gamma \in [0,1]$ which will be specified. Fix also X_0 and b satisfying Assumption **(M)** for some $\beta \le 1+\gamma$. In this example, the process X considered is defined, for all $t \in [0,T]$, by

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dZ_s.$$

Corollary 7.4 (Sub-critical case)**.** *Assume that* $\alpha \in (1,2)$, $\beta \in (1,\alpha)$ *and* $\gamma \in (\alpha-1, 1]$. *Then, the process* $X = (X_t)_t$ *satisfies Assumptions* ***(M)****,* ***(J1)*** *and* ***(J2)*** *with* $K(s,z) = H(s,z) = \sigma_s z$. *Let* $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ *be a function having a linear derivative* $\frac{\delta}{\delta m}u$ *which satisfies the following properties.*

1. *For any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the function $\frac{\delta}{\delta m}u(\mu) \in \mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ and $\partial_v \frac{\delta}{\delta m}u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.*
2. *For any compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*

$$\forall \mu \in \mathcal{K},\, \forall x, y \in \mathbb{R}^d,\, \left| \partial_v \frac{\delta}{\delta m}u(\mu)(x) - \partial_v \frac{\delta}{\delta m}u(\mu)(y) \right| \leq C_\mathcal{K}|x-y|^\gamma.$$

3. *For any compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, we have*

$$\sup_{\mu \in \mathcal{K}} \int_{\mathbb{R}^d} \left| \partial_v \frac{\delta}{\delta m}u(\mu)(v) \right|^{\beta'} d\mu(v) < +\infty.$$

Then, Itô's formula (7.13) holds true for X and u.

Corollary 7.5 (Super-critical case)**.** *Assume that $\alpha \in (0,1)$, $\beta \in (0,\alpha)$ and $\gamma = 0$. Then, the process $X = (X_t)_t$ satisfies Assumptions **(M)**, **(J1)** and **(J2)** with $K(s,z) = H(s,z) = \sigma_s z$. Let $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a function having a linear derivative $\frac{\delta}{\delta m}u$ which satisfies the following properties.*

1. *For any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the function $\frac{\delta}{\delta m}u(\mu) \in \mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ and $\partial_v \frac{\delta}{\delta m}u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.*
2. *For any compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, we have*

$$\forall v \in \mathbb{R}^d,\, \sup_{v \in \mathbb{R}^d} \sup_{\mu \in \mathcal{K}} \left| \partial_v \frac{\delta}{\delta m}u(\mu)(v) \right| < +\infty.$$

Then, Itô's formula (7.13) holds true for X and u.

Corollary 7.6 (Critical case)**.** *Assume that $\alpha = 1$, $\beta \in (0,1)$ and $\gamma \in (0,1]$. Then, the process $X = (X_t)_t$ satisfies Assumptions **(M)**, **(J1)** and **(J2)** with $K(s,z) = H(s,z) = \sigma_s z$. Let $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a function having a linear derivative $\frac{\delta}{\delta m}u$ which satisfies the following properties.*

1. *For any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the function $\frac{\delta}{\delta m}u(\mu) \in \mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ and $\partial_v \frac{\delta}{\delta m}u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.*
2. *For any compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*

$$\forall \mu \in \mathcal{K},\, \forall x, y \in \mathbb{R}^d,\, \left| \partial_v \frac{\delta}{\delta m}u(\mu)(x) - \partial_v \frac{\delta}{\delta m}u(\mu)(y) \right| \leq C_\mathcal{K}|x-y|^\gamma.$$

3. *For any compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, we have*

$$\forall v \in \mathbb{R}^d,\, \sup_{v \in \mathbb{R}^d} \sup_{\mu \in \mathcal{K}} \left| \partial_v \frac{\delta}{\delta m}u(\mu)(v) \right| < +\infty.$$

Then, Itô's formula (7.13) holds true for X and u.

Let us give an example of a function u which satisfies the assumptions of the three preceding corollaries. We take $d = 1$ to simplify the computations.

Example 7.7. Let $\beta \in (0,2)$. For $\varepsilon > 0$, consider a function χ_ε of class $\mathcal{C}^2$ on $\mathbb{R}$ such that

- $\forall x \in \mathbb{R}$, $\chi_\varepsilon(x) \in [0,1]$,
- χ_ε is equal to 1 on $[-2\varepsilon, 2\varepsilon]^c$ and to 0 on $[-\varepsilon, \varepsilon]$.

Define the function u by

$$\forall \mu \in \mathcal{P}_\beta(\mathbb{R}),\ u(\mu) := \int_{\mathbb{R}} |x|^\beta \chi_\varepsilon(x)\, d\mu(x).$$

Then, u satisfies the assumptions of Corollaries 7.4, 7.5 and 7.6 with $\gamma = 1$.

Proof. An easy computation shows that u has a linear derivative given, for all $\mu \in \mathcal{P}_\beta(\mathbb{R})$ and $v \in \mathbb{R}$, by

$$\frac{\delta}{\delta m} u(\mu)(v) = |v|^\beta \chi_\varepsilon(v),$$

which is clearly of class $\mathcal{C}^2$. Moreover, one has for all $\mu \in \mathcal{P}_\beta(\mathbb{R})$ and $v \in \mathbb{R}$

$$\forall \mu \in \mathcal{P}_\beta(\mathbb{R}),\ \forall v \in \mathbb{R},\ \partial_v \frac{\delta}{\delta m} u(\mu)(v) = \beta \mathrm{sgn}(v)|v|^{\beta-1}\chi_\varepsilon(v) + |v|^\beta \chi_\varepsilon'(v).$$

Thus, the condition in Remark 7.3 is satisfied, i.e. for all $\mu \in \mathcal{P}_\beta(\mathbb{R})$ and $v \in \mathbb{R}$

$$\left|\partial_v \frac{\delta}{\delta m} u(\mu)(v)\right| \leq C(1 + |v|^{\beta-1}\mathbf{1}_{\beta>1}).$$

Moreover, we have for all $\mu \in \mathcal{P}_\beta(\mathbb{R})$ and $v \in \mathbb{R}$

$$\partial_v^2 \frac{\delta}{\delta m} u(\mu)(v) = \beta(\beta-1)\mathrm{sgn}(v)|v|^{\beta-2}\chi_\varepsilon(v) + |v|^\beta \chi_\varepsilon''(v) + 2\beta \mathrm{sgn}(v)|v|^{\beta-1}\chi_\varepsilon'(v).$$

The right-hand side term is clearly bounded on $\mathbb{R}$, and thus Assumption (2) in Theorem 7.2 is satisfied with $\gamma = 1$ by the mean value theorem. $\square$

We now extend Itô's formula in the case where the function u depends also on the time and space variables. Let us introduce $\mathcal{M}$ another Poisson random measure on $[0,T] \times \mathbb{R}^d \backslash \{0\}$ with Lévy measure π. Then we define, for $t \in [0,T]$

$$Y_t := Y_0 + \int_0^t \kappa_s\, ds + \int_0^t \int_{B_1} I(s,z)\, \widetilde{\mathcal{M}}(ds,dz) + \int_0^t \int_{B_1^c} J(s,z)\, \mathcal{M}(ds,dz),$$

where Y_0 is $\mathcal{F}_0$-measurable, κ, I and J are predictable processes with

$$\int_0^T |\kappa_s|\, ds + \int_0^T \int_{B_1} |I(s,z)|^{1+\Gamma} + |I(s,z)|^2\, d\pi(z)\, ds < +\infty \quad \text{a.s.}$$

for some $\Gamma \in (0,1]$.

Theorem 7.8 (Extension of Itô's formula)**.** *Instead of* (7.10), (7.11) *and* (7.12) *in Assumptions **(M)**, **(J1)** and **(J2)**, we assume that*

$$\mathbb{E}|X_0|^\beta + \mathbb{E}\sup_{t\leq T}|b_s|^{\beta\vee 1} + \int_{B_1^c} \mathbb{E}\sup_{t\leq T}|K(t,z)|^\beta\, d\nu(z) + \int_{B_1} \mathbb{E}\sup_{t\leq T}(|\widetilde{H}_t||z|)^{1+\gamma}\, d\nu(z) < +\infty, \tag{7.14}$$

with $\beta \leq 1+\gamma$. Moreover, we assume that for all $t \in [0,T]$ and $z \in \mathbb{R}^d$, $K(\cdot,z)$ and $H(\cdot,z)$ are almost surely continuous at t.

Let $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a continuous function satisfying the following properties.

1. *For any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ $u(\cdot, x, \mu)$ is continuously differentiable and $\partial_t u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$. Moreover for any $t \in [0,T]$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $u(t, \cdot, \mu)$ is of class $\mathcal{C}^1$ on $\mathbb{R}^d$ with $\partial_x u$ continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$ and such that $\partial_x u(t, \cdot, \mu)$ is Γ-Hölder continuous uniformly with respect to $t \in [0,T]$ and $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$.*
2. *For any $t \in [0,T]$, $x \in \mathbb{R}^d$, the function $u(t,x,\cdot)$ admits a linear derivative $\frac{\delta}{\delta m} u(t,x,\cdot)(\cdot)$ such that $\frac{\delta}{\delta m} u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ and for all compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*
$$\forall t \in [0,T],\ \forall (x,\mu) \in \mathcal{K},\ \forall v \in \mathbb{R}^d,\ \left| \frac{\delta}{\delta m} u(t,x,\mu)(v) \right| \leq C_\mathcal{K}(1 + |v|^\beta).$$
Moreover, we assume that for any $t \in [0,T]$, $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(t,x,\mu)$ is differentiable and that $\partial_v \frac{\delta}{\delta m} u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.
3. *If $\gamma > 0$, for any compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*
$$\forall t \in [0,T],\ \forall (x,\mu) \in \mathcal{K},\ \forall v, v' \in \mathbb{R}^d,\ \left| \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v) - \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v') \right| \leq C_\mathcal{K} |v - v'|^\gamma.$$
4. *For any compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, we have*
$$\begin{cases} \sup\limits_{t \in [0,T]} \sup\limits_{(x,\mu) \in \mathcal{K}} \displaystyle\int_{\mathbb{R}^d} \left| \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v) \right|^{\beta'} d\mu(v) < +\infty, & \text{if } \beta > 1, \\ \sup\limits_{t \in [0,T]} \sup\limits_{v \in \mathbb{R}^d} \sup\limits_{(x,\mu) \in \mathcal{K}} \left| \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v) \right| < +\infty, & \text{if } \beta \leq 1. \end{cases}$$

Then, the function $(t,x) \in [0,T] \times \mathbb{R}^d \mapsto u(t,x,\mu_t)$ is of class $\mathcal{C}^1$, with $\partial_x u(t, \cdot, \mu_t)$ Γ-Hölder continuous uniformly with respect to t. Moreover, we have almost surely for all $t \in [0,T]$

$$\begin{aligned}
&u(t, Y_t, \mu_t) - u(0, Y_0, \mu_0) \\
&= \int_0^t \partial_t u(s, Y_s, \mu_s)\, ds + \int_0^t \overline{\mathbb{E}} \left(\partial_v \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_s) \cdot \overline{b}_s \right) ds \\
&\quad + \int_0^t \int_{B_1^c} \overline{\mathbb{E}} \left[\frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-} + \overline{K}(s,z)) - \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-}) \right] d\nu(z)\, ds \\
&\quad + \int_0^t \int_{B_1} \overline{\mathbb{E}} \Big[\frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-} + \overline{H}(s,z)) - \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-}) \\
&\qquad\qquad - \partial_v \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-}) \cdot \overline{H}(s,z) \Big] d\nu(z)\, ds \qquad (7.15) \\
&\quad + \int_0^t \partial_x u(s, Y_s, \mu_s) \cdot \kappa_s\, ds + \int_0^t \int_{B_1^c} u(s, Y_{s^-} + J(s,z), \mu_s) - u(s, Y_{s^-}, \mu_s)\, \mathcal{M}(ds, dz) \\
&\quad + \int_0^t \int_{B_1} u(s, Y_{s^-} + I(s,z), \mu_s) - u(s, Y_{s^-}, \mu_s)\, \widetilde{\mathcal{M}}(ds, dz) \\
&\quad + \int_0^t \int_{B_1} u(s, Y_{s^-} + I(s,z), \mu_s) - u(s, Y_{s^-}, \mu_s) - \partial_x u(s, Y_{s^-}, \mu_s) \cdot I(s,z)\, d\pi(z)\, ds,
\end{aligned}$$

where $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ is an independent copy of $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\overline{b}, \overline{H}, \overline{K}, \overline{X})$ is a copy of (b, H, K, X).

Before proving Itô's formulas of Theorem 7.2 and Theorem 7.8, we gather in the next section some useful properties of the process X and on its flow of marginal distributions $(\mu_t)_t$. We also introduce the localization of X that is at the core of our proof.

7.2.2 Preliminary study of the process and localization

We introduce the driving Lévy process Y defined, for $t \in [0,T]$, by

$$Y_t := \int_0^t \int_{B_1} z\,\widetilde{\mathcal{N}}(ds,dz) + \int_0^t \int_{B_1^c} z\,\mathcal{N}(ds,dz). \tag{7.16}$$

The jumping times of Y associated with jumps of size greater than 1 are denoted by $(\tilde{T}_n)_n$ and the associated jumps $(Z_n)_n$ are defined, for all $n \geq 1$, by

$$Z_n := \delta Y_{\tilde{T}_n} := Y_{\tilde{T}_n} - Y_{\tilde{T}_n^-}.$$

The sequence $(\tilde{T}_{n+1} - \tilde{T}_n)_n$ is i.i.d. with common distribution exponential of parameter $\nu(B_1^c)$ and $(Z_n)_n$ is independent of $(\tilde{T}_n)_n$ with common distribution $\frac{\nu_{|B_1^c}}{\nu(B_1^c)}$. We can write, for $t \in [0,T]$

$$\int_0^t \int_{B_1^c} z\,\mathcal{N}(ds,dz) = \sum_{n=1}^{\infty} Z_n \mathbf{1}_{\tilde{T}_n \leq t} = \sum_{n=1}^{N_t} Z_n,$$

where $N_t := \sum_{k=1}^{\infty} \mathbf{1}_{\tilde{T}_k \leq t}$ is the associated Poisson process. Let us begin with the continuity of the flow $(\mu_t)_t$

Proposition 7.9. *There exists a constant $C > 0$ such that for all stopping time τ and for all $0 \leq s \leq t \leq T$, one has*

$$\begin{aligned}&\mathbb{E}|X_{t\wedge\tau} - X_{s\wedge\tau}|^\beta \\ &\leq C\left[\left(\mathbb{E}\int_s^t |b_u|^{\beta\vee 1}\,du\right)^{\frac{\beta}{\beta\vee 1}} + \left(\mathbb{E}\int_s^t \int_{B_1} |H(r,z)|^{1+\gamma}\,d\nu(z)\,dr\right)^{\frac{\beta}{1+\gamma}} + \mathbb{E}\int_s^t \int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr\right].\end{aligned} \tag{7.17}$$

Thus, the map $t \in [0,T] \mapsto \mu_t \in \mathcal{P}_\beta(\mathbb{R}^d)$ is continuous. Moreover, we have the following moment estimate

$$\mathbb{E}\sup_{t\leq T} |X_t|^\beta < +\infty. \tag{7.18}$$

Proof. For the drift term, we work in $L^{\beta\vee 1}(\Omega)$. In fact, Jensen's inequality ensures that

$$\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau} b_u\,du\right|^{\beta\vee 1} \leq C\,\mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau} |b_u|^{\beta\vee 1}\,du.$$

Then, a discussion on the relative position of τ with respect to s and t yields

$$\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau} b_u\,du\right|^{\beta\vee 1} \leq C\,\mathbb{E}\int_s^t |b_u|^{\beta\vee 1}\,du.$$

The compensated Poisson integral is treated in $L^{1+\gamma}(\Omega)$ since $\beta \leq 1+\gamma$. The Burkholder-Davis-Gundy

(BDG) inequality together with the fact that $1+\gamma \leq 2$ imply that

$$
\begin{aligned}
&\mathbb{E}\left|\int_0^{t\wedge\tau}\int_{B_1} H(r,z)\,\widetilde{\mathcal{N}}(dr,dz) - \int_0^{s\wedge\tau}\int_{B_1} H(r,z)\,\widetilde{\mathcal{N}}(dr,dz)\right|^{1+\gamma} \\
&\leq C\,\mathbb{E}\left(\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1} |H(r,z)|^2\,\mathcal{N}(dr,dz)\right)^{\frac{1+\gamma}{2}} \qquad (7.19)\\
&= C\,\mathbb{E}\left(\sum_{s\wedge\tau<r\leq t\wedge\tau} |H(r,\delta Y_r)|^2\right)^{\frac{1+\gamma}{2}} \\
&\leq C\,\mathbb{E}\sum_{s\wedge\tau<r\leq t\wedge\tau} |H(r,\delta Y_r)|^{1+\gamma} \qquad (7.20)\\
&= C\,\mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1} |H(r,z)|^{1+\gamma}\,\mathcal{N}(dr,dz) \\
&= C\,\mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1} |H(r,z)|^{1+\gamma}\,d\nu(z)\,dr \\
&\leq C\,\mathbb{E}\int_{s}^{t}\int_{B_1} |H(r,z)|^{1+\gamma}\,d\nu(z)\,dr.
\end{aligned}
$$

Finally, for the Poisson integral, we work in $L^\beta(\Omega)$ and we begin with the case $\beta \leq 1$. In this case, the same computations as done for the compensated Poisson random integral yield

$$
\begin{aligned}
\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} K(r,z)\,\mathcal{N}(dr,dz)\right|^\beta &= \mathbb{E}\left|\sum_{n=1}^{\infty} K(\tilde{T}_n, Z_n)\mathbf{1}_{s\wedge\tau<\tilde{T}_n\leq t\wedge\tau}\right|^\beta \\
&\leq C\,\mathbb{E}\int_s^t\int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr. \qquad (7.21)
\end{aligned}
$$

Let us now deal with the case $\beta \in (1,2]$. Writing artificially the Poisson integral as a compensated Poisson integral and using BDG's and Jensen's inequalities, we deduce that

$$
\begin{aligned}
&\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} K(r,z)\,\mathcal{N}(dr,dz)\right|^\beta \\
&\leq C\left[\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} K(r,z)\,\widetilde{\mathcal{N}}(dr,dz)\right|^\beta + \mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} K(r,z)\,d\nu(z)\,dr\right|^\beta\right] \\
&\leq C\left[\mathbb{E}\left|\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} |K(r,z)|^2\,\mathcal{N}(dr,dz)\right|^{\frac{\beta}{2}} + \mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr\right] \qquad (7.22)\\
&\leq C\left[\mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} |K(r,z)|^\beta\,\mathcal{N}(dr,dz) + \mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr\right] \\
&\leq C\,\mathbb{E}\int_{s\wedge\tau}^{t\wedge\tau}\int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr \\
&\leq C\,\mathbb{E}\int_{s}^{t}\int_{B_1^c} |K(r,z)|^\beta\,d\nu(z)\,dr.
\end{aligned}
$$

We have thus proved that (7.17) holds true. From this, we deduce that the map $t \in [0,T] \mapsto \mu_t \in \mathcal{P}_\beta(\mathbb{R}^d)$ is continuous using the dominated convergence theorem and Assumptions **(M)**, **(J1)** and **(J2)**.

For the moment estimate (7.18), Jensen's inequality ensures that

$$\sup_{t\le T}\left|\int_0^t b_s\,ds\right|^{\beta\vee 1}\le C\,\mathbb{E}\int_0^T |b_s|^{\beta\vee 1}\,ds.$$

The upper-bound is finite owing to Assumption **(M)**. For the compensated Poisson integral, we work in $L^{1+\gamma}(\Omega)$ since $\beta\le 1+\gamma$. Reasoning as in (7.19), it follows from BDG's inequality that

$$\begin{aligned}\mathbb{E}\sup_{t\le T}\left|\int_0^t\int_{B_1} H(s,z)\,\widetilde{\mathcal{N}}(ds,dz)\right|^{1+\gamma} &\le C\,\mathbb{E}\left(\int_0^T\int_{B_1}|H(s,z)|^2\,\mathcal{N}(ds,dz)\right)^{\frac{1+\gamma}{2}}\\ &\le C\,\mathbb{E}\int_0^T\int_{B_1}|H(s,z)|^{1+\gamma}\,d\nu(z)\,ds,\end{aligned}$$

which is finite thanks to Assumption **(J1)**. Finally, for the Poisson integral, we have

$$\mathbb{E}\sup_{t\le T}\left|\int_0^t\int_{B_1^c} K(s,z)\,\mathcal{N}(ds,dz)\right|^{\beta}\le \mathbb{E}\left(\int_0^T\int_{B_1^c}|K(s,z)|\,\mathcal{N}(ds,dz)\right)^{\beta}.$$

Reasoning as in (7.21) if $\beta\le 1$ and as in (7.22) if $\beta>1$, one deduces that

$$\mathbb{E}\sup_{t\le T}\left|\int_0^t\int_{B_1^c} K(s,z)\,\mathcal{N}(ds,dz)\right|^{\beta}\le C\,\mathbb{E}\int_0^T\int_{B_1^c}|K(s,z)|^{\beta}\,d\nu(z)\,ds.$$

The upper-bound is finite by Assumption **(J2)**.

□

We now introduce the localization of process X.

Definition 7.10 (Localized process and approximate flow)**.** Let us introduce the sequence of localizing stopping times $(T_n)_n$ defined, for all $n\ge 1$, by

$$T_n := \inf\{t\in[0,T],\ |X_t|\ge n \quad\text{or}\quad |\widetilde{H}_s|\ge n\}\wedge T,$$

where $\widetilde{H}$ was introduced in Assumption **(J2)**. Then, we define the localized process X^n and its flow of marginal distributions, for all $t\in[0,T]$, by

$$X_t^n := X_{t\wedge T_n}\quad\text{and}\quad \mu_t^n := [X_t^n]. \tag{7.23}$$

Remark 7.11. Since X is a càdlàg process and $\widetilde{H}$ is almost surely locally bounded, it is immediate that almost surely $T_n = T$, for n bigger than some random constant.

Let us start with the continuity of the sequence of approximate flows $(\mu^n)_n$ and a uniform moment estimate as in Proposition 7.9.

Proposition 7.12. *We have*

$$\sup_n \mathbb{E}|X_t^n - X_s^n|^{\beta}\xrightarrow[s\to t]{} 0, \tag{7.24}$$

which yields the continuity of the map $t\in[0,T]\mapsto\mu_t^n\in\mathcal{P}_\beta(\mathbb{R}^d)$ for all $n\ge 1$. Moreover, the following *uniform moment estimate holds true*

$$\mathbb{E} \sup_{n\geq 1} \sup_{t\leq T} |X_t^n|^\beta < +\infty. \tag{7.25}$$

Proof. It follows immediately from Proposition 7.9. □

We can now state the main approximation result that we will use to prove Theorem 7.2.

Proposition 7.13 (Approximation)**.** *We have almost surely*

$$\sup_{t\leq T} |X_t^n - X_t| \underset{n\to+\infty}{\longrightarrow} 0.$$

Moreover, for all $t \in [0,T]$, $\mu_t^n \overset{\mathcal{P}_\beta(\mathbb{R}^d)}{\longrightarrow} \mu_t$.

Proof. The first point is a direct consequence of the fact that $T_n = T$ a.s. for n large enough (see Remark 7.11). For the convergence in $\mathcal{P}_\beta(\mathbb{R}^d)$, we show that $\mathbb{E}|X_t^n - X_t|^\beta \to 0$, for any $t \in [0,T]$. This is a consequence of the first point and of the dominated convergence theorem by (7.25). □

We end this preliminary section with a compactness result on the sequence of approximate flows $(\mu^n)_n$.

Proposition 7.14. *The set* $\{\mu_t^n,\, n \geq 1,\, t \in [0,T]\}$ *is relatively compact in* $\mathcal{P}_\beta(\mathbb{R}^d)$.

Proof. Let $(\mu_{s_k}^{n_k})_k$ be a subsequence of $\{\mu_t^n,\, n \geq 1,\, t \in [0,T]\}$. Up to extracting a subsequence, we can assume that $s_k \to s \in [0,T]$. If there exists k_0 such that $n_k = n_{k_0}$ for infinitely many $k \geq 1$. The continuity of $t \in [0,T] \mapsto \mu_t^{n_{k_0}}$ allows to conclude that along this subsequence

$$\mu_{s_k}^{n_k} \overset{\mathcal{P}_\beta(\mathbb{R}^d)}{\longrightarrow} \mu_s^{n_{k_0}}.$$

Otherwise, we can assume that $n_k \to +\infty$. The triangle inequality yields

$$W_\beta(\mu_{s_k}^{n_k}, \mu_s) \leq \sup_{n\geq 1} W_\beta(\mu_{s_k}^n, \mu_s^n) + W_\beta(\mu_s^{n_k}, \mu_s).$$

The first term of the right-hand side of the preceding inequality converges to 0 owing to Proposition 7.12, as well as the second one by Proposition 7.13. □

7.2.3 Proof of Theorem 7.2 and Theorem 7.8

Proof of Theorem 7.2. In Itô's formula (7.13), there are three types of integrals: the drift term, the small jumps term coming from the compensated Poisson integral and the big jumps term stemming from the Poisson integral. They will be always treated separately in the proof. Let us fix $t \in (0,T]$ and prove Itô's formula at time t.

Step 1: All the terms in Itô's formula (7.13) are well-defined.

If $\beta > 1$, the drift term is well-defined since Hölder's inequality implies that

$$\int_0^T \mathbb{E}\left|\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s\right| ds \le \left(\mathbb{E}\int_0^T |b_s|^\beta\, ds\right)^{1/\beta} \left(\mathbb{E}\int_0^T \left|\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s)\right|^{\beta'} ds\right)^{1/\beta'}.$$

The right-hand side term is finite thanks to Assumption **(M)** and Assumption (3) in Theorem 7.2 since $(\mu_t)_{t\in[0,T]}$ is compact in $\mathcal{P}_\beta(\mathbb{R}^d)$ by Proposition 7.9. If $\beta \le 1$, Assumption (3) in Theorem 7.2 ensures that there exists $C > 0$ such that

$$\int_0^T \mathbb{E}\left|\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s\right| ds \le C\,\mathbb{E}\int_0^T |b_s|\, ds,$$

which is finite thanks to Assumption **(M)**. We consider now the big jumps term in (7.13). The growth condition in the definition of the linear derivative (see Definition 7.1) and Assumption **(J2)** ensure that

$$\mathbb{E}\left|\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right| \le C\left(1 + \mathbb{E}\sup_{s\le T}|X_s|^\beta + |K(s,z)|^\beta\right).$$

The right-hand side term belongs to $L^1([0,T]\times B_1^c, ds\otimes\nu)$ because of the moment estimate (7.18) in Proposition 7.9 and Assumption **(J2)**. We finally focus on the small jumps term. Note that if $\gamma = 0$ and since we have assumed that $\beta \le \gamma + 1$, then $\beta \le 1$ and Assumption (3) in Theorem 7.2 ensures that for all compact $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that for all $v \in \mathbb{R}^d$, we have

$$\sup_{\mu\in\mathcal{K}}\left|\partial_v \frac{\delta}{\delta m} u(\mu)(v)\right| \le C_\mathcal{K}.$$

Thus, the mean value theorem yields for all $s \in [0,T]$ and for all $z \in B_1$

$$\begin{aligned}\left|\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}) - H(s,z)\cdot\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right| &\le C|H(s,z)| \\ &= C|H(s,z)|^{1+\gamma}.\end{aligned} \tag{7.26}$$

If $\gamma > 0$, then (7.26) holds true. Indeed, it follows from Taylor's formula and Assumption (2) in Theorem 7.2 that for all $s \in [0,T]$ and $z \in B_1$

$$\begin{aligned}&\left|\frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + H(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}) - H(s,z)\cdot\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right| \\ &= \left|\int_0^1 \left[\partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + rH(s,z)) - \partial_v \frac{\delta}{\delta m} u(\mu_s)(X_{s^-})\right]\cdot H(s,z)\, dr\right| \\ &\le C|H(s,z)|^{1+\gamma}.\end{aligned} \tag{7.27}$$

Assumption **(J1)** allows us to conclude that the small jumps term is well-defined in both cases.

Step 2: Itô's formula for μ^n.

We fix $n \ge 1$ and we aim at proving Itô's formula (7.13) for the localized process X^n defined, for all

$t \in [0,T]$, by $X^n_t = X_{t\wedge T_n}$. For $m \geq 1$, we define the following subdivision of $[0,t]$ by

$$\forall k \in \{0,\dots,m\},\, t^m_k := \frac{k}{m}t.$$

We will omit the index m in the sequel and denote by $(t_k)_k$ this subdivision. By definition of the linear derivative, one has for all $k \in \{0,\dots,m\}$

$$u(\mu^n_{t_{k+1}}) - u(\mu^n_{t_k}) = \int_0^1 \mathbb{E}\left(\frac{\delta}{\delta m}u(M^k_r)(X^n_{t_{k+1}}) - \frac{\delta}{\delta m}u(M^k_r)(X^n_{t_k})\right) dr, \tag{7.28}$$

where $M^k_r := r\mu^n_{t_{k+1}} + (1-r)\mu^n_{t_k}$, omitting again the dependence in n and m. Let us denote by F the function $\frac{\delta}{\delta m}u(M^k_r)$. Note that F is of class $\mathcal{C}^1$ et ∇F is γ-Hölder continuous. We can apply the standard Itô formula for F and X^n by Assumption **(J1)**, which yields almost surely for all $t \in [0,T]$

$$\begin{aligned} F(X^n_t) = F(X^n_0) &+ \int_0^{t\wedge T_n} \nabla F(X^n_s)\cdot b_s\, ds \\ &+ \int_0^{t\wedge T_n}\int_{B_1^c} [F(X^n_{s^-} + K(s,z)) - F(X^n_{s^-})]\, \mathcal{N}(ds,dz) \\ &+ \int_0^{t\wedge T_n}\int_{B_1} [F(X^n_{s^-} + H(s,z)) - F(X^n_{s^-})]\, \widetilde{\mathcal{N}}(ds,dz) \\ &+ \int_0^{t\wedge T_n}\int_{B_1} [F(X^n_{s^-} + H(s,z)) - F(X^n_{s^-}) - H(s,z)\cdot\nabla F(X^n_{s^-})]\, d\nu(z)\, ds. \end{aligned} \tag{7.29}$$

Now, we aim at taking the expectation in the previous formula. The compensated Poisson integral is centered. Indeed by definition of T_n, we have almost surely for all $s \in [0, t\wedge T_n)$ and for all $z \in B_1$

$$|X^n_{s^-}| + |H(s,z)| \leq 2n.$$

Thus, the mean value theorem ensures that for some constant $C_n > 0$ depending only on n and F, we have almost surely for all $s \in [0,t\wedge T_n)$ and for all $z \in B_1$

$$|F(X^n_{s^-} + H(s,z)) - F(X^n_{s^-})|^2 \leq C_n |H(s,z)|^2.$$

Thus, by definition of T_n and by Assumption **(J1)**, one has almost surely for all $s \in [0,t\wedge T_n)$ and for all $z \in B_1$

$$|F(X^n_{s^-} + H(s,z)) - F(X^n_{s^-})|^2 \leq n^2 C_n |z|^2.$$

From this inequality, we deduce that

$$\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1} |F(X^n_{s^-} + H(s,z)) - F(X^n_{s^-})|^2\, d\nu(z)\, ds < +\infty.$$

It follows that the compensated Poisson integral in (7.29) is a true centered martingale. For the Poisson integral in (7.29), we use the growth property of F coming from the definition of the linear derivative. It yields

$$|F(X^n_{s^-} + K(s,z)) - F(X^n_{s^-})| \leq C\left(1 + \sup_{s\leq T}|X_s|^\beta + |K(s,z)|^\beta\right).$$

Thus, Assumption **(J2)** and the moment estimate (7.18) in Proposition 7.9 prove that

$$\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1^c}|F(X_{s-}^n+K(s,z))-F(X_{s-}^n)|\,d\nu(z)\,ds<+\infty.$$

This yields

$$\begin{aligned}&\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1^c}[F(X_{s-}^n+K(s,z))-F(X_{s-}^n)]\,\mathcal{N}(ds,dz)\\&=\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1^c}[F(X_{s-}^n+K(s,z))-F(X_{s-}^n)]\,d\nu(z)\,ds.\end{aligned}$$

Taking the expectation in the other term of Itô's formula (7.29) can be done arguing as in Step 1. It follows that

$$\begin{aligned}\mathbb{E}F(X_t^n)=\mathbb{E}F(X_0^n)&+\mathbb{E}\int_0^{t\wedge T_n}\partial_v F(X_s^n)\cdot b_s\,ds\\&+\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1^c}[F(X_{s-}^n+K(s,z))-F(X_{s-}^n)]\,d\nu(z)\,ds\\&+\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1}[F(X_{s-}^n+H(s,z))-F(X_{s-}^n)-H(s,z)\cdot\partial_v F(X_{s-}^n)]\,d\nu(z)\,ds.\end{aligned}\tag{7.30}$$

Recall that F is equal to $\frac{\delta}{\delta m}u(M_r^k)$. Thus, by using (7.28) and (7.30), one has

$$\begin{aligned}u(\mu_t^n)-u(\mu_0^n)&=\sum_{k=0}^{m-1}u(\mu_{t_{k+1}}^n)-u(\mu_{t_k}^n)\\&=I_1+I_2+I_3,\end{aligned}\tag{7.31}$$

where

$$\begin{aligned}I_1&:=\sum_{k=0}^{m-1}\int_0^1\mathbb{E}\int_{t_k\wedge T_n}^{t_{k+1}\wedge T_n}\partial_v\frac{\delta}{\delta m}u(M_r^k)(X_s^n)\cdot b_s\,ds\,dr,\\I_2&:=\sum_{k=0}^{m-1}\int_0^1\mathbb{E}\int_{t_k\wedge T_n}^{t_{k+1}\wedge T_n}\int_{B_1^c}\left[\frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n+K(s,z))-\frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n)\right]d\nu(z)\,ds\,dr,\\I_3&:=\sum_{k=0}^{m-1}\int_0^1\mathbb{E}\int_{t_k\wedge T_n}^{t_{k+1}\wedge T_n}\int_{B_1}\left[\frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n+H(s,z))-\frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n)\right.\\&\qquad\left.-H(s,z)\cdot\partial_v\frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n)\right]d\nu(z)\,ds\,dr.\end{aligned}$$

Our goal is now to let m tend to infinity in (7.31). For the small jumps term I_3, it can be rewritten as

$$\begin{aligned}I_3=\int_0^1\mathbb{E}\int_0^{t\wedge T_n}\int_{B_1}&\left[\frac{\delta}{\delta m}u(M_r^{\lfloor ms/t\rfloor})(X_{s-}^n+H(s,z))-\frac{\delta}{\delta m}u(M_r^{\lfloor ms/t\rfloor})(X_{s-}^n)\right.\\&\left.-H(s,z)\cdot\partial_v\frac{\delta}{\delta m}u(M_r^{\lfloor ms/t\rfloor})(X_{s-}^n)\right]d\nu(z)\,ds\,dr,\end{aligned}$$

where $\lfloor\cdot\rfloor$ is the integer part function and since $\lfloor ms/t\rfloor$ is equal to k if and only if $s\in[t_k,t_{k+1})$. The continuity of the flow $s \in [0,T] \mapsto \mu_s^n \in \mathcal{P}_\beta(\mathbb{R}^d)$ ensures that for all $r \in [0,1]$ and $s \in [0,t]$

$$M_r^{\lfloor ms/t \rfloor} \underset{m\to+\infty}{\longrightarrow} \mu_s^n,$$

for the Wasserstein metric W_β. The continuity of $\frac{\delta}{\delta m}u$ and $\partial_v \frac{\delta}{\delta m}u$ on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$ implies that

$$\begin{aligned}
&\frac{\delta}{\delta m}u(M_r^{\lfloor ms/t \rfloor})(X_{s-}^n + H(s,z)) - \frac{\delta}{\delta m}u(M_r^{\lfloor ms/t \rfloor})(X_{s-}^n) - H(s,z) \cdot \partial_v \frac{\delta}{\delta m}u(M_r^{\lfloor ms/t \rfloor})(X_{s-}^n) \\
&\underset{m\to+\infty}{\longrightarrow} \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n + H(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) - H(s,z) \cdot \partial_v \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n).
\end{aligned}$$

Note that the set $\{v\mu_s^n + (1-v)\mu_t^n,\ s,t \in [0,T],\ v \in [0,1]\}$ is compact in $\mathcal{P}_\beta(\mathbb{R}^d)$ thanks to Proposition 7.14. As done at the end of Step 1 in (7.26) and (7.27), Assumption (2) in Theorem 7.2 guarantees that the integrand in I_3 is bounded, up to some multiplicative constant, by $|\widetilde{H}_s|^{1+\gamma}|z|^{1+\gamma}$. Assumption **(J1)** proves that this quantity belongs to $L^1([0,1] \times [0,t] \times B_1 \times \Omega, dr \otimes ds \otimes \nu \otimes \mathbb{P})$. Thus, the dominated convergence theorem ensures that I_3 converges, when $m \to +\infty$, towards

$$\mathbb{E} \int_0^{t\wedge T_n} \int_{B_1} \left[\frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n + H(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) - H(s,z) \cdot \partial_v \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) \right] d\nu(z)\, ds.$$

Using the same arguments, we get that

$$I_2 \underset{m\to+\infty}{\longrightarrow} \mathbb{E} \int_0^t \int_{B_1^c} \left[\frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n + K(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) \right] d\nu(z)\, ds.$$

From the growth condition in the definition of the linear derivative, we deduce that almost surely for all $s \in [0,T]$, $r \in [0,1]$, $z \in B_1$, $n \in \mathbb{N}$

$$\left| \frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n + K(s,z)) - \frac{\delta}{\delta m}u(M_r^k)(X_{s-}^n) \right| \le C \left(1 + \sup_{s\le T} |X_s|^\beta + |K(s,z)|^\beta \right).$$

The upper-bound belongs to $L^1([0,1] \times [0,t] \times B_1^c \times \Omega, dr \otimes ds \otimes \nu \otimes \mathbb{P})$ by Assumption **(J2)**. This justifies the use of the dominated convergence theorem, as we did for I_3. Finally, I_1 can be handled in the same way, the domination being an immediate consequence of the localization of the process and the continuity of $\partial_v \frac{\delta}{\delta m}u$. Letting m tend to infinity in (7.31), one has for all $t \in [0,T]$

$$\begin{aligned}
&u(\mu_t^n) - u(\mu_0^n) \\
&= \mathbb{E} \int_0^{t\wedge T_n} \left(\partial_v \frac{\delta}{\delta m}u(\mu_s^n)(X_s^n) \cdot b_s \right) ds \\
&\quad + \mathbb{E} \int_0^{t\wedge T_n} \int_{B_1^c} \left[\frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n + K(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) \right] d\nu(z)\, ds \\
&\quad + \mathbb{E} \int_0^{t\wedge T_n} \int_{B_1} \left[\frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n + H(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) - \partial_v \frac{\delta}{\delta m}u(\mu_s^n)(X_{s-}^n) \cdot H(s,z) \right] d\nu(z)\, ds \\
&=: A_1 + A_2 + A_3.
\end{aligned} \tag{7.32}$$

Step 3: Itô's formula for X.

Our goal is now to let n tend to infinity in Itô's formula (7.32) for the localized process. For the left-hand side term of (7.32), we have seen in Proposition 7.13 that for all $t \in [0,T]$, when $n \to \infty$

$$\mu_t^n \overset{\mathcal{P}_\beta(\mathbb{R}^d)}{\longrightarrow} \mu_t.$$

The continuity of u on $\mathcal{P}_\beta(\mathbb{R}^d)$ ensures that $u(\mu_t^n) \to u(\mu_t)$. For the right-hand side of (7.32), let us start with the limit of A_1 when n tends to $+\infty$. It follows from the continuity of $\partial_v \frac{\delta}{\delta m} u$ and from Proposition 7.13 that almost surely for all $s \le t$

$$\partial_v \frac{\delta}{\delta m} u(\mu_s^n)(X_s^n) \cdot b_s \mathbf{1}_{s \le T_n} \underset{n \to +\infty}{\longrightarrow} \partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s.$$

For $\beta \le 1$, we use the dominated convergence theorem justified by Assumption (3) in Theorem 7.2 and Assumption **(M)**. For $\beta > 1$, we conclude with a uniform integrability argument in the space $L^1([0,t] \times \Omega, ds \otimes \mathbb{P})$. Indeed, notice that $b \in L^\beta([0,t] \times \Omega, ds \otimes d\mathbb{P})$ and

$$\sup_{n \ge 1} \mathbb{E} \int_0^t \left| \partial_v \frac{\delta}{\delta m} u(\mu_s^n)(X_s^n) \mathbf{1}_{s \le T_n} \right|^{\beta'} ds < +\infty,$$

thanks to Assumption (3) of Theorem 7.2 and the relative compactness of $\{\mu_s^n,\, n \ge 1,\, 0 \le s \le T\}$ in $\mathcal{P}_\beta(\mathbb{R}^d)$ of Proposition 7.14. Thus, the sequence $\left((\partial_v \frac{\delta}{\delta m} u(\mu_s^n)(X_s^n) \cdot b_s \mathbf{1}_{s \le T_n})_{s \in [0,t]} \right)_n$ is uniformly integrable in $L^1([0,t] \times \Omega, ds \otimes \mathbb{P})$. This comes from the fact that if $(X_n)_n$ is a sequence of $\mathbb{R}^d$-valued random variables bounded in $L^{\beta'}$ and $Y \in L^\beta$, then $(X_n \cdot Y)_n$ is uniformly integrable. It is clearly bounded in L^1 by Hölder's inequality. Moreover, if A is a measurable set, then we have for all $n \ge 1$

$$\mathbb{E}(|X_n \cdot Y| \mathbf{1}_A) \le \sup_{n \ge 1} (\mathbb{E}|X_n|^{\beta'})^{\frac{1}{\beta'}} (\mathbb{E}(|Y|^\beta \mathbf{1}_A))^{\frac{1}{\beta}}.$$

The conclusion follows from the fact that Y is uniformly integrable. We have thus proved that

$$A_1 \underset{n \to +\infty}{\longrightarrow} \mathbb{E} \int_0^t \partial_v \frac{\delta}{\delta m} u(\mu_s)(X_s) \cdot b_s \, ds.$$

We now focus on the big jumps term A_2. The continuity of $\frac{\delta}{\delta m} u$ and Proposition 7.13 prove that almost surely for all $s \in [0,t]$, $z \in B_1^c$

$$\frac{\delta}{\delta m} u(\mu_s^n)(X_{s^-}^n + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s^n)(X_{s^-}^n) \underset{n \to +\infty}{\longrightarrow} \frac{\delta}{\delta m} u(\mu_s)(X_{s^-} + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s)(X_{s^-}).$$

The growth assumption in the definition of the linear derivative together with Assumption **(J2)** ensure that for all $s \in [0,t]$, $z \in B_1^c$, $n \in \mathbb{N}$

$$\left| \frac{\delta}{\delta m} u(\mu_s^n)(X_{s^-}^n + K(s,z)) - \frac{\delta}{\delta m} u(\mu_s^n)(X_{s^-}^n) \right| \le C \left(1 + \sup_{s \le T} |X_s|^\beta + |K(s,z)|^\beta \right).$$

Using Assumption **(J2)** and the moment estimate (7.18), one has

$$\mathbb{E} \int_0^t \int_{B_1^c} \left(1 + \sup_{s \le T} |X_s|^\beta + |K(s,z)|^\beta \right) d\nu(z) \, ds < +\infty.$$

Thus, the dominated convergence theorem yields

$$A_2 \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\int_0^t \int_{B_1^c} \left[\frac{\delta}{\delta m}u(\mu_s)(X_{s^-}+K(s,z)) - \frac{\delta}{\delta m}u(\mu_s)(X_{s^-})\right] d\nu(z)\, ds.$$

Eventually, we let n tend to infinity in the small jumps term A_3 thanks to the dominated convergence theorem. For the domination, we use the relative compactness of $\{\mu_s^n,\, n\geq 1,\, 0\leq s\leq T\}$ in $\mathcal{P}_\beta(\mathbb{R}^d)$ given by Proposition 7.14 and Assumption (2) in Theorem 7.2. Reasoning as in the end of Step 1 in (7.26) and (7.27), we have almost surely, for all $s\in[0,t]$, $z\in B_1$, $n\geq 1$

$$\left|\frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n+H(s,z)) - \frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n) - \partial_v\frac{\delta}{\delta m}u(\mu_s^n)(X_{s^-}^n)\cdot H(s,z)\right| \leq C|H(s,z)|^{1+\gamma}.$$

It follows from Assumption **(J1)** that the dominated convergence theorem can be applied. It yields

$$A_3 \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\int_0^{t\wedge T_n} \int_{B_1} \left[\frac{\delta}{\delta m}u(\mu_s)(X_{s^-}+H(s,z)) - \frac{\delta}{\delta m}u(\mu_s)(X_{s^-}) - \partial_v\frac{\delta}{\delta m}u(\mu_s)(X_{s^-})\cdot H(s,z)\right] d\nu(z)\, ds.$$

We conclude the proof by taking the limit $n\to+\infty$ in (7.32). □

Proof of Theorem 7.8. It is enough to prove that the map $(t,x)\in[0,T]\times\mathbb{R}^d\mapsto u(t,x,\mu_t)$ is of class $\mathcal{C}^1$. Then, we conclude by applying the standard Itô's formula for this function and for the process $(Y_t)_t$ since $x\in\mathbb{R}^d\mapsto\partial_x u(t,x,\mu_t)$ is Γ-Hölder continuous uniformly in time. Let us fix $\mathcal{K}\subset\mathbb{R}^d$ a compact subset, $x\in\mathcal{K}$, $t\in[0,T]$ and $h\in\mathbb{R}$ such that $t+h\in[0,T]$. We have

$$u(t+h,x,\mu_{t+h}) - u(t,x,\mu_t) = [u(t+h,x,\mu_{t+h}) - u(t,x,\mu_{t+h})] + [u(t,x,\mu_{t+h}) - u(t,x,\mu_t)]\,.$$

The continuity of $\partial_t u$ on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)$ ensures that

$$\begin{aligned}\frac{1}{h}u(t+h,x,\mu_{t+h}) - u(t,x,\mu_{t+h}) &= \frac{1}{h}\int_t^{t+h}\partial_t u(s,x,\mu_{t+h})\, ds\\ &\underset{h\to0}{\longrightarrow} \partial_t u(t,x,\mu_t).\end{aligned}$$

Then, using Itô's formula of Theorem 7.2, we obtain that

$$\begin{aligned}&u(t,x,\mu_{t+h}) - u(t,x,\mu_t)\\ &= \int_t^{t+h}\mathbb{E}\left(\partial_v\frac{\delta}{\delta m}u(t,x,\mu_s)(X_s)\cdot b_s\right) ds \\ &\quad+\int_t^{t+h}\int_{B_1^c}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-}+K(s,z)) - \frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\right] d\nu(z)\, ds\\ &\quad+\int_t^{t+h}\int_{B_1}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-}+H(s,z))\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\right.\\ &\qquad\qquad\left.-\partial_v\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\cdot H(s,z)\right] d\nu(z)\, ds.\end{aligned} \tag{7.33}$$

First, note that the function $(s,t,x)\in[0,T]^2\times\mathcal{K}\mapsto\mathbb{E}\left(\partial_v\frac{\delta}{\delta m}u(t,x,\mu_s)(X_s)\cdot b_s\right)$ is continuous. Indeed, it easily follows from the continuity of $\partial_v\frac{\delta}{\delta m}u$ and b, the fact that for all $t\in[0,T]$, $X_{t^-}=X_t$ almost surely, and from a uniform integrability argument. The uniform integrability of the family

$\left(\partial_v \frac{\delta}{\delta m} u(t,x,\mu_s)(X_s)\cdot b_s\right)_{(s,t,x)\in[0,T]^2\times\mathcal{K}}$ comes from the integrability assumption on b and from Assumption (4) in Theorem 7.8. It follows from the dominated convergence theorem that the function

$$(s,t,x)\in[0,T]^2\times\mathcal{K}\mapsto \int_{B_1^c}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-}+K(s,z))-\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\right]d\nu(z)$$

is continuous. Indeed, the domination is easily deduced from Assumption (2) in Theorem 7.8, the assumption on K and from (7.18). Finally, the function

$$\begin{aligned}&(s,t,x)\in[0,T]^2\times\mathcal{K}\mapsto\\ &\int_{B_1}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-}+H(s,z))-\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})-\partial_v\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\cdot H(s,z)\right]d\nu(z)\end{aligned}$$

is also continuous using the dominated convergence theorem. Indeed by Assumption (3), we find that almost surely for all $x\in\mathcal{K}$, $z\in B_1$, $s,t\in[0,T]$

$$\left|\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-}+H(s,z))-\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})-\partial_v\frac{\delta}{\delta m}u(t,x,\mu_s)(X_{s^-})\cdot H(s,z)\right|\leq C\sup_{s\leq T}(|\widetilde{H}_s||z|)^{1+\gamma}.$$

This shows that

$$\begin{aligned}&\frac{1}{h}(u(t,x,\mu_{t+h})-u(t,x,\mu_t))\\ &\underset{h\to 0}{\longrightarrow}\mathbb{E}\left(\partial_v\frac{\delta}{\delta m}u(t,x,\mu_t)(X_t)\cdot b_t\right)\\ &\quad+\int_{B_1^c}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_t)(X_{t^-}+K(t,z))-\frac{\delta}{\delta m}u(t,x,\mu_t)(X_{t^-})\right]d\nu(z)\\ &\quad+\int_{B_1}\mathbb{E}\left[\frac{\delta}{\delta m}u(t,x,\mu_t)(X_{t^-}+H(t,z))-\frac{\delta}{\delta m}u(t,x,\mu_t)(X_{t^-})-\partial_v\frac{\delta}{\delta m}u(t,x,\mu_t)(X_{t^-})\cdot H(t,z)\right]d\nu(z).\end{aligned}$$

We have thus proved that $t\in[0,T]\mapsto u(t,x,\mu_t)$ is differentiable with $\frac{d}{dt}u(t,\cdot,\mu_t)$ continuous on $[0,T]\times\mathbb{R}^d$ and that Itô's formula (7.15) holds true.

□

7.3 Backward Kolmogorov PDE on the space of measures and empirical projection

In this section, we use Itô's formula to derive the backward Kolmogorov PDE on the space of measures associated with a general Lévy-driven McKean-Vlasov SDE. More precisely, it describes the dynamics of the associated semigroup acting on functions defined on the space of measures under regularity assumptions on it. In the Brownian case, it has been done in Chapter 5 of [CD18a] and in [CdRF22]. Let us introduce $Z=(Z_t)_t$ a Lévy process on $\mathbb{R}^d$ which has the following Lévy-Itô decomposition

$$Z_t=\int_0^t\int_{B_1}z\widetilde{\mathcal{N}}(ds,dz)+\int_0^t\int_{B_1^c}z\mathcal{N}(ds,dz),\tag{7.34}$$

where $\mathcal{N}$ is the associated Poisson random measure with Lévy measure ν. We fix our assumptions on the Lévy process Z. We assume that there exists $\beta\in(0,2]$ and $\gamma\in[0,1]$ with $\beta\leq 1+\gamma$ and such that the Lévy measure ν satisfies the following properties.

(J1') We have

$$\int_{B_1} |z|^{1+\gamma}\, d\nu(z) < +\infty. \tag{7.35}$$

(J2') For all $t \in [0,T]$, Z_t has a finite moment of order β i.e.

$$\int_{B_1^c} |z|^{\beta}\, d\nu(z) < +\infty. \tag{7.36}$$

Note that we can always take $\gamma = 1$ by definition of a Lévy measure. We are interested in the following McKean-Vlasov SDE

$$\begin{cases} dX_t = b_t(X_t,\mu_t)\,dt + \sigma_t(X_{t-},\mu_t)\,dZ_t, \quad t\in[0,T],\\ \mu_t = [X_t],\\ X_0 = \xi \in L^{\beta}(\Omega,\mathcal{F}_0), \end{cases} \tag{7.37}$$

where $b : [0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)\to\mathbb{R}^{d\times d}$ are measurable. We make the following assumptions on the McKean-Vlasov SDE (7.37).

(H1) There is weak existence and uniqueness for (7.37), for any initial time $s\in[0,T)$ and for any initial distribution μ_0 in $\mathcal{P}_\beta(\mathbb{R}^d)$.

(H2) If $\beta \geq 1$, there exists $C>0$ such that for all $t\in[0,T]$, $x\in\mathbb{R}^d$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$

$$\begin{aligned} |b_t(x,\mu)| &\leq C\left(1+|x|+M_\beta(\mu)\right),\\ |\sigma_t(x,\mu)| &\leq C, \end{aligned} \tag{7.38}$$

where $M_\beta(\mu) = \left(\int_{\mathbb{R}^d}|x|^\beta\,d\mu(x)\right)^{\frac{1}{\beta}}$.

If $\beta<1$, there exists $C>0$ such that for all $t\in[0,T]$, $x\in\mathbb{R}^d$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$

$$|b_t(x,\mu)| + |\sigma_t(x,\mu)| \leq C. \tag{7.39}$$

(H3) The coefficients b and σ are continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d)$.

We refer to [Cav23] for a well-posedness result, in the strong sense, in a Lipschitz framework. Notice that under these assumptions, an immediate adaptation of Lemma 4.1 in [FL21] ensures that $(\mu_t)_t$ belongs to $\mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d))$ and that

$$\mathbb{E}\sup_{t\in[0,T]}|X_t|^\beta < +\infty. \tag{7.40}$$

The action of the semigroup associated with (7.37) is given in the following definition.

Definition 7.15 (Semigroup)**.** For a function $u:\mathcal{P}_\beta(\mathbb{R}^d)\to\mathbb{R}$, the action of the semigroup associated with the McKean-Vlasov SDE (7.37) on the map u is given by the function ϕ_u defined by

$$\phi_u : \begin{cases} [0,T]\times\mathcal{P}_\beta(\mathbb{R}^d) & \to \mathbb{R}\\ \qquad (t,\mu) & \mapsto u([X_T^{T-t,\mu}]), \end{cases} \tag{7.41}$$

where $[X_T^{T-t,\mu}]$ denotes the distribution of any solution to (7.37) at time T starting at time $T-t$ from any random variable ξ with distribution μ. This makes sense by the well-posedness assumption **(H1)** for (7.37).

Theorem 7.16 (Backward Kolmogorov PDE)**.** *Assume that **(H1)**, **(H2)**, **(H3)** hold and that the function ϕ_u satisfies the following properties.*

1. *The function ϕ_u belongs to $\mathcal{C}^0([0,T]\times\mathcal{P}_\beta(\mathbb{R}^d);\mathbb{R})$, $\partial_t\phi_u$ exists and is continuous on $(0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)$, and $\frac{\delta}{\delta m}\phi_u$ and $\partial_v\frac{\delta}{\delta m}\phi_u$ exist and are continuous on $(0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$.*

2. *If $\gamma>0$, for any compact $\mathcal{K}\subset\mathcal{P}_\beta(\mathbb{R}^d)$ and $a>0$, there exists $C>0$ such that for all $\mu\in\mathcal{K}$, $x,y\in\mathbb{R}^d$ and $t\in[a,T]$*
$$\left|\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(x)-\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(y)\right|\le C|x-y|^\gamma.$$

3. *For any compact $\mathcal{K}\subset\mathcal{P}_\beta(\mathbb{R}^d)$ and $a>0$, we have*
$$\begin{cases}\displaystyle\sup_{t\in[a,T]}\sup_{\mu\in\mathcal{K}}\int_{\mathbb{R}^d}\left|\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(v)\right|^{\beta'}d\mu(v)<+\infty & \text{if } \beta>1,\\ \displaystyle\sup_{t\in[a,T]}\sup_{\mu\in\mathcal{K}}\sup_{v\in\mathbb{R}^d}\left|\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(v)\right|<+\infty & \text{if } \beta\le 1.\end{cases}$$

Then, the function ϕ_u satisfies
$$\begin{cases}\partial_t\phi_u(T-t,\mu)=\mathscr{L}_t\phi_u(T-t,\mu), & \forall(t,\mu)\in[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d),\\ \phi_u(T-t,\cdot)_{|t=T}=u(\cdot),\end{cases}\tag{7.42}$$

where the operator $\mathscr{L}_t$ is defined, for any $s\in(0,T]$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, by
$$\begin{aligned}&\mathscr{L}_t\phi_u(s,\mu)\\ &:=\int_{\mathbb{R}^d}\partial_v\frac{\delta}{\delta m}\phi_u(s,\mu)(v)\,b_t(v,\mu)\,d\mu(v)\\ &\quad+\int_{\mathbb{R}^d}\int_{B_1^c}\left[\frac{\delta}{\delta m}\phi_u(s,\mu)(v+\sigma_t(v,\mu)z)-\frac{\delta}{\delta m}\phi_u(s,\mu)(v)\right]d\nu(z)\,d\mu(v)\\ &\quad+\int_{\mathbb{R}^d}\int_{B_1}\left[\frac{\delta}{\delta m}\phi_u(s,\mu)(v+\sigma_t(v,\mu)z)-\frac{\delta}{\delta m}\phi_u(s,\mu)(v)-\partial_v\frac{\delta}{\delta m}\phi_u(s,\mu)(v)\cdot(\sigma_t(v,\mu)z)\right]d\nu(z)\,d\mu(v).\end{aligned}\tag{7.43}$$

Proof. First, note that by definition of ϕ_u and thanks to the well-posedness assumption **(H1)**, the function $t\in[0,T]\mapsto\phi_u(T-t,\mu_t)$ is constant. Differentiating with respect to t the function $t\mapsto\phi(T-t,\mu_t)$ thanks to Itô's formula of Theorem 7.8 and the regularity assumptions made of ϕ_u, one has for all $t \in [0,T)$

$$\begin{aligned}
0 &= -\partial_t \phi_u(T-t,\mu_t) \\
&\quad + \int_{\mathbb{R}^d} \partial_v \frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v)\, b_t(v,\mu)\, d\mu_t(v) \\
&\quad + \int_{\mathbb{R}^d}\int_{B_1^c} \left[\frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v+\sigma_t(v,\mu_t)z) - \frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v)\right] d\nu(z)\, d\mu_t(v) \\
&\quad + \int_{\mathbb{R}^d}\int_{B_1} \left[\frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v+\sigma_t(v,\mu_t)z) - \frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v)\right. \\
&\qquad\qquad \left. -\partial_v \frac{\delta}{\delta m}\phi_u(T-t,\mu_t)(v)\cdot(\sigma_t(v,\mu_t)z)\right] d\nu(z)\, d\mu_t(v).
\end{aligned}$$

We conclude the proof by noting that the flow of measures $(\mu_t)_t$ associated with the McKean-Vlasov SDE (7.37) can be initialised at time t with any measure $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. □

Let us recall the definition of empirical projection.

Definition 7.17 (Empirical projection)**.** Fix $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$. For all $N \geq 1$, the empirical projection u^N of u is defined for all $\boldsymbol{x} = (x_1,\dots,x_N) \in (\mathbb{R}^d)^N$ by

$$u^N(\boldsymbol{x}) := u(\overline{\mu}^N_{\boldsymbol{x}}),$$

where $\overline{\mu}^N_{\boldsymbol{x}} = \frac{1}{N}\sum_{j=1}^N \delta_{x_j}$ is the empirical measure.

We also state the following Lemma on the link between the regularity of u and u^N. The proof is completely analogous to [CD18a, Proposition 5.91].

Lemma 7.18. *Let $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a function admitting a linear derivative $\frac{\delta}{\delta m}u$ satisfying the following properties.*

1. *For all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $\frac{\delta}{\delta m}u(\mu) \in \mathcal{C}^2(\mathbb{R}^d;\mathbb{R})$.*
2. *The functions $\partial_v \frac{\delta}{\delta m}u$ and $\partial_v^2 \frac{\delta}{\delta m}u$ are continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$.*
3. *For all $v \in \mathbb{R}^d$, the function $\mu \in \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \partial_v \frac{\delta}{\delta m}u(\mu)(v)$ has a linear derivative*

$$(\mu, v') \mapsto \frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}u(\mu)(v,v') \in \mathbb{R}^d,$$

which is $\mathcal{C}^1$ with respect to v' and such that $\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d\times\mathbb{R}^d$.

Then, for all $N \geq 1$, the empirical projection u^N of u is of class $\mathcal{C}^2$. Moreover, we have for all $\boldsymbol{x} = (x_1,\dots,x_N) \in (\mathbb{R}^d)^N$ and $i,j \in \{1,\dots N\}$

$$\partial_{x_i} u^N(x_1,\dots,x_N) = \frac{1}{N}\partial_v \frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_i),$$

and

$$\partial_{x_j}\partial_{x_i} u^N(x_1,\dots,x_N) = \frac{1}{N^2}\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_i,x_j) + \mathbf{1}_{i=j}\frac{1}{N}\partial_v^2\frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_j).$$

We consider now the mean-field interacting particle system associated with the McKean-Vlasov SDE (7.37). The goal is to approximate in some sense $\mathscr{L}_t u(\overline{\mu}^N_{\boldsymbol{x}})$ by the generator of the particle system applied to the empirical projection u^N of u. This will be crucial to prove quantitative weak propagation of chaos using the Kolmogorov backward PDE associated with (7.37), as explained in Chapter 5 of [CD18a].

Let us introduce the particle system associated with (7.37) and some notations used in the sequel. We denote by $(Z^n)_n$ an i.i.d. sequence of Lévy processes having the same distribution as $Z = (Z_t)_t$ defined in (7.34), and by $(X_0^n)_n$ an i.i.d. sequence of random variables with common distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$. For a fixed integer $N \geq 1$, the system of N particles associated with (7.37) is defined as the solution to the following classical SDE on $(\mathbb{R}^d)^N$, provided that it is well-posed in the weak sense,

$$\begin{cases} dX_t^{i,N} = b_t(X_t^{i,N}, \overline{\mu}_t^N)\, dt + \sigma_t(X_{t-}^{i,N}, \overline{\mu}_{t-}^N)\, dZ_t^i, \quad t \in [0,T], \quad i \leq N, \\ \overline{\mu}_t^N := \frac{1}{N} \sum_{j=1}^N \delta_{X_t^{j,N}}, \\ X_0^{i,N} = X_0^i. \end{cases} \tag{7.44}$$

We can write for all $i \in \{1, \dots, N\}$ and for all $t \in [0,T]$

$$Z_t^i = \int_0^t \int_{B_1} z\, \widetilde{\mathcal{N}}^i(ds, dz) + \int_0^t \int_{B_1^c} z\, \mathcal{N}^i(ds, dz),$$

where $\mathcal{N}^i$ is the Poisson random measure associated with Z^i. Then, we set for all $t \in [0,T]$

$$\mathbf{Z}_t^N := \begin{pmatrix} Z_t^1 \\ \vdots \\ Z_t^N \end{pmatrix} \in (\mathbb{R}^d)^N.$$

As the Lévy processes $(Z^n)_n$ are independent, the process $(\mathbf{Z}_t^N)_t$ is a Lévy process in $(\mathbb{R}^d)^N$. Its Poisson random measure $\boldsymbol{\mathcal{N}}^N$ and its Lévy measure $\boldsymbol{\nu}^N$ are defined as follows. For all $\phi : [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}^+$, one has

$$\begin{aligned} \int_0^T \int_{(\mathbb{R}^d)^N} \phi(s, \boldsymbol{x})\, \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{x}) &= \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \phi(s, 0, \dots, 0, x_i, 0, \dots, 0)\, \mathcal{N}^i(ds, dx_i) \\ &= \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \phi(s, \tilde{\boldsymbol{x}}_{\boldsymbol{i}})\, \mathcal{N}^i(ds, dx), \end{aligned} \tag{7.45}$$

where $\tilde{\boldsymbol{x}}_{\boldsymbol{i}} := (0, \dots, x, \dots, 0) \in (\mathbb{R}^d)^N$ for $x \in \mathbb{R}^d$, where x appears in the i-th position. For all $\phi : (\mathbb{R}^d)^N \to \mathbb{R}^+$, one has

$$\int_{(\mathbb{R}^d)^N} \phi(\boldsymbol{x})\, d\boldsymbol{\nu}^N(\boldsymbol{x}) = \sum_{i=1}^N \int_{\mathbb{R}^d} \phi(\tilde{\boldsymbol{x}}_{\boldsymbol{i}})\, d\nu(x). \tag{7.46}$$

Note that since the processes $(Z^n)_n$ are independent, for all $t \in [0,T]$, the support of the random measure $\boldsymbol{\mathcal{N}}^N(t, d\boldsymbol{x})$ is contained in

$$\bigcup_{i=0}^{N-1} \{0_{\mathbb{R}^d}\}^i \times \mathbb{R}^d \times \{0_{\mathbb{R}^d}\}^{N-1-i} \subset (\mathbb{R}^d)^N.$$

Let us define for all $t \in [0,T]$, $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$

$$\mathbf{b}_t^N(\boldsymbol{x}) := \begin{pmatrix} b_t(x_1, \overline{\mu}_{\boldsymbol{x}}^N) \\ \vdots \\ b_t(x_N, \overline{\mu}_{\boldsymbol{x}}^N) \end{pmatrix} \in (\mathbb{R}^d)^N \quad \text{and} \quad \boldsymbol{\sigma}_t^N(\boldsymbol{x}) := \begin{pmatrix} \sigma_t(x_1, \overline{\mu}_{\boldsymbol{x}}^N) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_t(x_N, \overline{\mu}_{\boldsymbol{x}}^N) \end{pmatrix} \in \mathbb{R}^{Nd \times Nd}.$$

Thus, writing $\boldsymbol{X}_t^N = \begin{pmatrix} X_t^{1,N} \\ \vdots \\ X_t^{N,N} \end{pmatrix}$, the SDE (7.44) defining the particle system can be rewritten as

$$\begin{cases} d\boldsymbol{X}_t^N = \mathbf{b}_t^N(\boldsymbol{X}_t^N)\, dt + \boldsymbol{\sigma}_t^N(\boldsymbol{X}_{t^-}^N)\, d\mathbf{Z}_t^N, \quad t \in [0,T], \\ \boldsymbol{X}_0^N = \begin{pmatrix} X_0^1 \\ \vdots \\ X_0^N \end{pmatrix}. \end{cases} \tag{7.47}$$

Let us finally recall the definition of the family of operators $(L_t)_t$ associated with the particle system (7.44). For any $f : (\mathbb{R}^d)^N \to \mathbb{R}$ of class $\mathcal{C}^2$ and for any $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$, we define $L_t f$ by

$$L_t^N f(\boldsymbol{x}) := \boldsymbol{b}_t^N(\boldsymbol{x}) \cdot \nabla f(\boldsymbol{x}) + \int_{(\mathbb{R}^d)^N} f(\boldsymbol{x} + \boldsymbol{\sigma}_t^N(\boldsymbol{x})\boldsymbol{z}) - f(\boldsymbol{x}) - \nabla f(\boldsymbol{x}) \cdot (\boldsymbol{\sigma}_t^N(\boldsymbol{x})\boldsymbol{z}\mathbf{1}_{|\boldsymbol{z}|<1})\, d\boldsymbol{\nu}^N(\boldsymbol{z}). \tag{7.48}$$

Proposition 7.19. *Let $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a function satisfying the assumptions of Theorem 7.2 with β and γ satisfying (7.35) and (7.36), and Lemma 7.18. Assume moreover that $\frac{\delta}{\delta m}\frac{\delta}{\delta m}u$ and $\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u$ exist and are uniformly bounded on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$. Then, we have for all $t \in [0,T]$ and $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$*

$$\begin{aligned} L_t^N u^N(x_1, \dots, x_N) &= \mathscr{L}_t u(\overline{\mu}_{\boldsymbol{x}}^N) + \frac{1}{N}\sum_{k=1}^N \int_{\mathbb{R}^d}\int_0^1 \frac{\delta}{\delta m}u(m_{t,z,w}^k)(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N)z) - \frac{\delta}{\delta m}u(m_{t,z,w}^k)(x_k) \\ &\qquad - \left[\frac{\delta}{\delta m}u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N)z) - \frac{\delta}{\delta m}u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k)\right] dw\, d\nu(z) \\ &= \mathscr{L}_t u(\overline{\mu}_{\boldsymbol{x}}^N) + O\left(\frac{1}{N}\right), \end{aligned}$$

where $m_{t,z,w}^k := w\overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{\sigma}_t(\boldsymbol{x},\overline{\mu}_{\boldsymbol{x}}^N)\tilde{\boldsymbol{z}}_{\boldsymbol{k}}} + (1-w)\overline{\mu}_{\boldsymbol{x}}^N$, for $w \in [0,1]$.

The last term is interpreted as an error term and we will need to control it to prove our propagation of chaos estimates in Section 7.4.

Proof. We easily check that we can apply the operator L_t to the empirical projection u^N thanks to the regularity assumptions. It yields for all $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$

$$\begin{aligned}
L_t^N u^N(\boldsymbol{x}) &= \frac{1}{N}\sum_{k=1}^N b_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N)\partial_v \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k) \\
&\quad + \sum_{k=1}^N \int_{\mathbb{R}^d} u(\overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{\sigma}_t(\boldsymbol{x})\tilde{\boldsymbol{z_k}}}) - u(\overline{\mu}_{\boldsymbol{x}}^N) - \frac{1}{N}\partial_v \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k) \cdot (\sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z \mathbf{1}_{|z|\le 1})\, d\nu(z) \\
&= \frac{1}{N}\sum_{k=1}^N b_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N)\partial_v \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k) \\
&\quad + \frac{1}{N}\sum_{k=1}^N \int_{\mathbb{R}^d}\int_0^1 \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k) \\
&\qquad\qquad - \partial_v \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k) \cdot (\sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z \mathbf{1}_{|z|\le 1})\, dw\, d\nu(z) \\
&= \mathscr{L}_t u(\overline{\mu}_{\boldsymbol{x}}^N) + \frac{1}{N}\sum_{k=1}^N \int_{\mathbb{R}^d}\int_0^1 \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k) \\
&\qquad\qquad - \left[\frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k)\right] dw\, d\nu(z).
\end{aligned}$$

It remains to show the bound on the error term between $L_t^N u^N(\boldsymbol{x})$ and $\mathscr{L}_t u(\overline{\mu}_{\boldsymbol{x}}^N)$. We split the domain of integration $\mathbb{R}^d$ of the last integral into B_1 and B_1^c and we first consider the integral over B_1^c. Defining $m^k_{t,z,w,r} := r m^k_{t,z,w} + (1-r)\overline{\mu}_{\boldsymbol{x}}^N = \overline{\mu}_{\boldsymbol{x}}^N + rw\overline{\mu}^N_{\boldsymbol{x}+\boldsymbol{\sigma}_t(\boldsymbol{x})\tilde{\boldsymbol{z_k}}}$, we have

$$\begin{aligned}
&\int_{B_1^c}\int_0^1 \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k) \\
&\qquad - \left[\frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k)\right] dw\, d\nu(z) \\
&= \frac{1}{N}\int_{B_1^c}\int_{[0,1]^2} w\left[\frac{\delta}{\delta m}\frac{\delta}{\delta m} u(m^k_{t,z,w,r})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z, x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z)\right. \\
&\qquad \left. - \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(m^k_{t,z,w,r})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z, x_k)\right] dr\, dw\, d\nu(z) \\
&\quad + \frac{1}{N}\int_{B_1^c}\int_{[0,1]^2} w\left[\frac{\delta}{\delta m}\frac{\delta}{\delta m} u(m^k_{t,z,w,r})(x_k, x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(m^k_{t,z,w,r})(x_k, x_k)\right] dr\, dw\, d\nu(z).
\end{aligned} \tag{7.49}$$

Since $\frac{\delta}{\delta m}\frac{\delta}{\delta m} u$ is uniformly bounded and $\nu(B_1^c) < +\infty$, we deduce that

$$\begin{aligned}
&\left|\int_{B_1^c}\int_0^1 \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(m^k_{t,z,w})(x_k)\right. \\
&\qquad \left. - \left[\frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k + \sigma_t(x_k, \overline{\mu}_{\boldsymbol{x}}^N) z) - \frac{\delta}{\delta m} u(\overline{\mu}_{\boldsymbol{x}}^N)(x_k)\right] dw\, d\nu(z)\right| \\
&\le \frac{C}{N}.
\end{aligned}$$

For the integral over B_1, one has

$$
\begin{aligned}
&\int_{B_1}\int_0^1 \frac{\delta}{\delta m}u(m^k_{t,z,w})(x_k+\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z) - \frac{\delta}{\delta m}u(m^k_{t,z,w})(x_k) \\
&\qquad\qquad - \left[\frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_k+\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z) - \frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_k)\right] dw\, d\nu(z) \\
&= \int_{B_1}\int_{[0,1]^2} \left[\partial_v \frac{\delta}{\delta m}u(m^k_{t,z,w})(x_k+h\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z) - \partial_v\frac{\delta}{\delta m}u(\overline{\mu}^N_{\boldsymbol{x}})(x_k+h\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\right]\cdot(\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\, dh\, dw\, d\nu(z) \\
&= \frac{1}{N}\int_{B_1}\int_{[0,1]^3} \left[\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u(m^k_{t,z,w,r})(x_k+h\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z, x_k+\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\right. \\
&\qquad\qquad \left. - \frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u(m^k_{t,z,w,r})(x_k+h\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z, x_k)\right]\cdot(\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\, dr\, dh\, dw\, d\nu(z) \\
&= \frac{1}{N}\int_{B_1}\int_{[0,1]^4} \left[\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u(m^k_{t,z,w,r})(x_k+h\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z, x_k+l\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)(\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad \cdot(\sigma_t(x_k,\overline{\mu}^N_{\boldsymbol{x}})z)\, dl\, dr\, dh\, dw\, d\nu(z).
\end{aligned}
$$

We conclude the proof using that σ and $\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}u$ are bounded and that $\int_{B_1}|z|^2\,d\nu(z)$ is finite.

□

7.4 Propagation of chaos for a mean-field system of interacting stable Ornstein-Uhlenbeck processes

We now study a nonlinear stable Ornstein-Uhlenbeck process driven by an α-stable process with $\alpha\in(1,2)$, which can be written, by [Sat99, Remark 14.6 and Theorem 14.7], for $t\geq 0$

$$Z_t = \int_0^t\int_{\mathbb{R}^d} z\,\widetilde{\mathcal{N}}(ds,dz).$$

We assume that Z is non degenerate in the following sense. Writing $y = r\theta\in\mathbb{R}^d\backslash\{0\}$ with $r\in\mathbb{R}^+_*$ and $\theta\in\mathbb{S}^{d-1}$, where $\mathbb{S}^{d-1}$ denotes the unit sphere in $\mathbb{R}^d$, the Lévy measure ν of Z decomposes as

$$\nu(dy) = d\mu(\theta)\frac{dr}{r^{1+\alpha}},$$

where μ is a non-zero finite measure on $\mathbb{S}^{d-1}$. We assume that ν satisfies the following non-degeneracy assumption.

(ND) There exists $\eta>0$ such that for all $\lambda\in\mathbb{R}^d$

$$\eta|\lambda|^2 \leq \int_{\mathbb{S}^{d-1}} |\langle\lambda,\theta\rangle|^2\,d\mu(\theta). \tag{7.50}$$

Let $A, A', B\in\mathcal{M}_d(\mathbb{R})$ be matrices of size $d\times d$ such that B is invertible and such that A and A' commute. Let us also fix $\beta\in[1,\alpha)$, and $\xi\in L^\beta(\Omega,\mathcal{F}_0)$ with distribution $\mu_0\in\mathcal{P}_\beta(\mathbb{R}^d)$. We study the following nonlinear SDE

$$\begin{cases} dX_t = (AX_t + A'\mathbb{E}X_t)\,dt + B\,dZ_t, \quad t\in[0,T], \\ X_0 = \xi. \end{cases} \tag{7.51}$$

Note that weak uniqueness holds for SDE (7.51) i.e. the distribution of $(X_t)_t$ depends only on the distribution μ_0 of ξ and not on the random variable ξ chosen.

We consider the mean-field system of interacting Ornstein-Uhlenbeck processes associated with (7.51). Let us denote by $(Z^n)_n$ an i.i.d. sequence of α-stable processes having the same distribution as $Z = (Z_t)_t$, and by $(X_0^n)_n$ an i.i.d. sequence of random variables with common distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$. For a fixed integer $N \geq 1$, the system of N particles associated with (7.37) is defined as the unique solution to the following classical SDE on $(\mathbb{R}^d)^N$

$$\begin{cases} dX_t^{i,N} = AX_t^{i,N}\,dt + A'\dfrac{1}{N}\displaystyle\sum_{j=1}^N X_t^{j,N}\,dt + B\,dZ_t^i, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ X_0^{i,N} = X_0^i. \end{cases} \tag{7.52}$$

We now state our quantitative weak propagation of chaos result in the next theorem.

Theorem 7.20. *Let $\mathscr{C}$ be the class of continuous functions $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ admitting two linear derivatives $\frac{\delta}{\delta m}u$ and $\frac{\delta}{\delta m}\frac{\delta}{\delta m}u$ such that $\frac{\delta}{\delta m}u(\mu)$ and $\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\mu)$ are Lipschitz continuous with Lipschitz constant smaller than 1. Then, there exists a constant $C = C_T$ independent of $u \in \mathscr{C}$ and $N \geq 1$ and non-decreasing with respect to T such that for all $u \in \mathscr{C}$ and $N \geq 1$, we have*

$$\mathbb{E}\left|u(\overline{\mu}_T^N) - u(\mu_T)\right| \leq C\,\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}, \tag{7.53}$$

and

$$|\mathbb{E}(u(\overline{\mu}_T^N) - u(\mu_T))| \leq C\,\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + \frac{C}{N^{\alpha-1}}. \tag{7.54}$$

Note that in the particular case where the initial distribution μ_0 belongs to $\mathcal{P}_2(\mathbb{R}^d)$, we have

$$|\mathbb{E}(u(\overline{\mu}_T^N) - u(\mu_T))| \leq \frac{C}{N^{\alpha-1}}. \tag{7.55}$$

Remark 7.21.

- The initial data terms can be handled using [FG15], in particular in the case where μ_0 has more moments than β. Indeed, one has if $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$

$$\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) \leq C \begin{cases} N^{-\frac{1}{2}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d=1 \quad \text{and} \quad q \neq 2, \\ N^{-\frac{1}{2}}\ln(1+N) + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d=2 \quad \text{and} \quad q \neq 2, \\ N^{-\frac{1}{d}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d\geq 3 \quad \text{and} \quad q \neq \frac{d}{d-1}. \end{cases} \tag{7.56}$$

- Let us denote by $\|\varphi\|_{\mathrm{Lip}} := \sup_{x\neq y}\frac{|\varphi(x)-\varphi(y)|}{|x-y|}$ for $\varphi : \mathbb{R}^d \to \mathbb{R}$. The set

$$\left\{\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R},\ \exists\varphi : \mathbb{R}^d \to \mathbb{R},\ \text{with}\ \|\varphi\|_{\mathrm{Lip}} \leq 1,\ \text{and}\ \phi(\mu) = \int_{\mathbb{R}^d}\varphi\,d\mu,\ \forall\mu \in \mathcal{P}_\beta(\mathbb{R}^d)\right\}$$

  is contained in $\mathscr{C}$. This allows quantify the approximation of the distribution of one particle by the limiting McKean-Vlasov process with respect to W_1. Indeed, we have by the Kantorovich-

Rubinstein theorem [Vil09, Remark 6.5] and if $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$

$$
\begin{aligned}
\sup_{t\in[0,T]} W_1([X_t^{1,N}],\mu_t) &= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left| \mathbb{E}\varphi(X_t^{1,N}) - \int_{\mathbb{R}^d} \varphi\, d\mu_t \right| \\
&= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left| \mathbb{E}\left(\frac{1}{N}\sum_{k=1}^N \varphi(X_t^{k,N}) \right) - \int_{\mathbb{R}^d} \varphi\, d\mu_t \right| \\
&= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left| \mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\overline{\mu}_t^N - \mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\mu_t \right| \\
&= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left| \mathbb{E}\phi(\overline{\mu}_t^N) - \mathbb{E}\phi(\mu_t) \right| \\
&\le \frac{C}{N^{\alpha-1}},
\end{aligned}
\tag{7.57}
$$

since the constant C in Theorem 7.20 is non-decreasing with respect to T.

Remark 7.22. Let us now compare the rates of convergence obtained in Theorem 7.20 with those proved in [Cav23, Theorem 3] at the level of paths by coupling. We assume that the initial distribution μ_0 belongs to $\mathcal{P}_2(\mathbb{R}^d)$. The estimate (7.53) quantifies the propagation of chaos through the convergence of the empirical measure. It allows morally to control $\mathbb{E}W_1(\overline{\mu}_T^N,\mu_T)$. We say "morally" because it would be true if the we could take the supremum over all functions $u \in \mathscr{C}$ inside the expectation, which is not the case in Theorem 7.20. The rate of convergence in (7.53) is the same as in [Cav23, Theorem 3], if we use (7.56) to control the initial data term in (7.53). However, if we have a better rate of convergence for $\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0)$, then the rate of convergence in (7.53) can be better than the one in [Cav23, Theorem 3], in particular when $d \ge 3$. Moreover, concerning the convergence of the distribution of one particle towards the distribution of the limiting McKean-Vlasov process, the estimate (7.57) is better than the one deduced from [Cav23, Theorem 3].

Remark 7.23. In order to obtain our propagation of chaos estimates, we follow the method presented in [CD18a, Chapter 5, pages 506 − 508]. We need to apply the operator L_t^N associated with the particle system, defined in (7.48), for the empirical projection $\boldsymbol{x} \in (\mathbb{R}^d)^N \mapsto \phi_u(t,\overline{\mu}_{\boldsymbol{x}}^N)$ of ϕ_u, which was defined in Definition 7.15. In order to use the backward Kolmogorov PDE satisfied by ϕ_u, we need to control the difference between $L_t^N\phi_u(t,\overline{\mu}_{\boldsymbol{x}}^N)$ and $\mathscr{L}_t\phi_u(t,\overline{\mu}_{\boldsymbol{x}}^N)$, as done in Proposition 7.19. However, the assumptions on $\frac{\delta}{\delta m}\frac{\delta}{\delta m}\phi_u$ made in this result are not satisfied. Indeed, for this interacting Ornstein-Uhlenbeck system, we can only show (see Proposition 7.24) that for any fixed $t \in (0,T]$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, and $v, v' \in \mathbb{R}^d$

$$
\left| \frac{\delta}{\delta m}\frac{\delta}{\delta m}\phi_u(t,\mu)(v,v') \right| \le C(1+|v|)(1+|v'|). \tag{7.58}
$$

This is due to the unboundedness of drift which is only at most of linear growth with respect to $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. The error term in (7.49), in the proof of Proposition 7.19, is not directly well-defined since $\int_{\mathbb{R}^d} |z|^2\, d\nu(z)$ is infinite. To get around this difficulty, we will consider, in a first step, truncated versions of the operators L_t^N and $\mathscr{L}_t$ by removing the big jumps of the noises. Namely we replace ν by the restriction of ν on the ball B_N, where N is also the number of particles of the system. For example, in the proof of Theorem 7.20, the terms (7.71), (7.72) and (7.74) are not well-defined without removing the big jumps.

We thus set for $t \in [0,T]$

$$
Z_{N,t} := \int_0^t \int_{B_N} z\,\widetilde{\mathcal{N}}(ds,dz). \tag{7.59}
$$

Then, we consider the McKean-Vlasov SDE (7.51) driven by Z_N. Its solution is denoted by $(X_{N,t})_t$ and satisfies

$$\begin{cases} dX_{N,t} = (AX_{N,t} + A'\mathbb{E}X_{N,t})\,dt + B\,dZ_{N,t}, \quad t \in [0,T], \\ X_0 = \xi \in L^\beta(\Omega, \mathcal{F}_0). \end{cases} \tag{7.60}$$

We denote by $\mu_{N,t}$ the distribution of $X_{N,t}$ and by $\phi_{N,u}$ the solution to the backward Kolmogorov PDE associated with the McKean-Vlasov SDE (7.60) defined in Definition 7.15.

Before proving the preceding theorem, we state in the following proposition regularity properties and bounds on $\phi_{N,u}$ that we will use.

Proposition 7.24. *For all $u \in \mathscr{C}$ and $N \geq 1$, the function $\phi_{N,u}$ satisfies the following properties.*

1. *The function $\phi_{N,u}$ belongs to $\mathcal{C}^0([0,T] \times \mathcal{P}_\beta(\mathbb{R}^d); \mathbb{R})$, $\partial_t \phi_{N,u}$ exists and is continuous on $(0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$. Moreover $\frac{\delta}{\delta m}\phi_{N,u}$, $\partial_v \frac{\delta}{\delta m}\phi_{N,u}$ and $\partial_v^2 \frac{\delta}{\delta m}\phi_{N,u}$ exist and are continuous on $(0,T] \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.*
2. *The functions $\frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}$ and $\partial_{v'} \frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}$ exist and are continuous on $(0,T] \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$.*
3. *There exists three continuous functions on $(0,T]$ denoted by g_1, g_2 and g_3 such that g_1 is globally bounded on $[0,T]$ and $g_2 \in L^1(0,T)$ which satisfy the following properties. For all $u \in \mathcal{C}$, $N \geq 1$, $t \in (0,T]$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $v, v' \in \mathbb{R}^d$, one has*

$$\begin{aligned} \left| \partial_v \frac{\delta}{\delta m}\phi_{N,u}(t,\mu)(v) \right| &\leq g_1(t), \\ \left| \partial_{v'} \frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}(t,\mu)(v,v') \right| &\leq g_2(t), \\ \left| \partial_v^2 \frac{\delta}{\delta m}\phi_{N,u}(t,\mu)(v) \right| &\leq g_3(t). \end{aligned} \tag{7.61}$$

The proof is postponed to the Appendix (Section 7.6).

We can now prove the propagation of chaos estimates of Theorem 7.20 for the mean-field system of interacting Ornstein-Uhlenbeck processes (7.52).

Proof of Theorem 7.20. We assume that the matrix B equal to the identity I_d since it does not change anything to the proof. Let us fix $N \geq 1$. As we did in (7.59) for $Z = (Z_t)_t$, we remove the big jumps of the i.i.d. copies $(Z^i)_i$ of Z by defining for all $i \in \{1, \dots, N\}$ and $t \in [0,T]$

$$Z^i_{N,t} := \int_0^t \int_{B_N} z\,\widetilde{\mathcal{N}}^i(ds,dz).$$

For the sake of clarity, we set for $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_1(\mathbb{R}^d)$

$$b(x,\mu) := Ax + A' \int_{\mathbb{R}^d} x\,d\mu(x).$$

Then, we consider $(X_{N,t})_t$ the nonlinear Ornstein-Uhlenbeck process driven by Z_N together with $\boldsymbol{X}^N_N :=$

$(X^{1,N}_{N,t},\dots,X^{N,N}_{N,t})_t$ the associated interacting particle system, i.e. the solutions to

$$\begin{cases} dX_{N,t} = b(X_{N,t},\mu_{N,t})\,dt + dZ_{N,t}, \quad t\in[0,T],\\ \mu_{N,t} = [X_{N,t}],\\ X_{N,0} = \xi \in L^\beta(\Omega,\mathcal{F}_0), \end{cases} \tag{7.62}$$

and

$$\begin{cases} dX^{i,N}_{N,t} = b(X^{i,N}_{N,t},\overline{\mu}^N_{N,t})\,dt + dZ^i_{N,t}, \quad t\in[0,T], \quad i\in\{1,\dots,N\},\\ \overline{\mu}^N_{N,t} = \dfrac{1}{N}\displaystyle\sum_{j=1}^N \delta_{X^{j,N}_{N,t}},\\ X^{i,N}_{N,0} = X^i_0. \end{cases} \tag{7.63}$$

Note that the particles $(X^{1,N}_N,\dots,X^{N,N}_N)$ have the same distribution. Let us also emphasize that we remove the jumps that are bigger than the number of particles N. Recall that the operator $\mathscr{L}_N$ associated with (7.62) defined in Theorem 7.16 is given, for all $s\in[0,T]$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, by

$$\begin{aligned} \mathscr{L}_N\phi_{N,u}(s,\mu) &:= \int_{\mathbb{R}^d}\partial_v\frac{\delta}{\delta m}\phi_{N,u}(s,\mu)(v)\,b(v,\mu)\,d\mu(v) \\ &\quad + \int_{\mathbb{R}^d}\int_{B_N}\left[\frac{\delta}{\delta m}\phi_{N,u}(s,\mu)(v+z) - \frac{\delta}{\delta m}\phi_{N,u}(s,\mu)(v) - \partial_v\frac{\delta}{\delta m}\phi_{N,u}(s,\mu)(v)\cdot z\right]d\nu(z)\,d\mu(v). \end{aligned} \tag{7.64}$$

In order to establish (7.53), our aim is to control $\mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_t) - \phi_u(T-t,\mu_t)|$ with respect to $t\in[0,T)$ and uniformly with respect to N. We decompose it in the following way

$$\begin{aligned} &\mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_t) - \phi_{N,u}(T-t,\mu_t)| \\ &\le \mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_t) - \phi_{N,u}(T-t,\overline{\mu}^N_{N,t})| + \mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t,\mu_{N,t})| \\ &\quad + \mathbb{E}|\phi_{N,u}(T-t,\mu_{N,t}) - \phi_{N,u}(T-t,\mu_t)|. \end{aligned}$$

Since $\partial_v\frac{\delta}{\delta m}\phi_{N,u}$ is bounded on $[0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d$ uniformly with respect to N by Proposition 7.24, we deduce by the Kantorovich-Rubinstein theorem that for some constant $C>0$, we have for all $N\ge1$, $t\in[0,T)$, $\mu,\nu\in\mathcal{P}_\beta(\mathbb{R}^d)$

$$|\phi_{N,u}(T-t,\mu) - \phi_{N,u}(T-t,\nu)| \le CW_1(\mu,\nu).$$

Thus, we obtain

$$\begin{aligned} \mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_t) - \phi_{N,u}(T-t,\mu_t)| &\le C\,\mathbb{E}W_1(\overline{\mu}^N_t,\overline{\mu}^N_{N,t}) + \mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t,\mu_{N,t})| \\ &\quad + C\,\mathbb{E}W_1(\mu_{N,t},\mu_t) \\ &=: A_1 + A_2 + A_3. \end{aligned} \tag{7.65}$$

First, we deal with A_1. Note that we can write, for $t\in[0,T]$,

$$X^{k,N}_{N,t} - X^{k,N}_t = \int_0^t (b(X^{k,N}_{N,s},\overline{\mu}^N_{N,s}) - b(X^{k,N}_s,\overline{\mu}^N_s)\,ds - \int_0^t\int_{B^c_N} z\,\widetilde{\mathcal{N}}^k(ds,dz).$$

Using the fact that b is Lipschitz continuous on $\mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$, one has for all $t \in [0,T]$

$$\begin{aligned}\mathbb{E}|X^{k,N}_{N,t} - X^{k,N}_t| &\leq C\left[\int_0^t \mathbb{E}|X^{k,N}_{N,s} - X^{k,N}_s|\,ds + \int_0^t \frac{1}{N}\sum_{j=1}^N \mathbb{E}|X^{j,N}_{N,s} - X^{j,N}_s|\,ds + \int_0^t \int_{B^c_N} |z|\,d\nu(z)\,ds\right]\\ &\leq C\left[\int_0^t \mathbb{E}|X^{k,N}_{N,s} - X^{k,N}_s|\,ds + \int_0^t \frac{1}{N}\sum_{j=1}^N \mathbb{E}|X^{j,N}_{N,s} - X^{j,N}_s|\,ds + \frac{1}{N^{\alpha-1}}\right].\end{aligned}$$

Gronwall's inequality ensures that there exists a constant $C > 0$ depending only on T and such that for any $t \in [0,T]$

$$\mathbb{E}W_1(\overline{\mu}^N_t, \overline{\mu}^N_{N,t}) \leq \frac{1}{N}\sum_{k=1}^N \mathbb{E}|X^{k,N}_{N,t} - X^{k,N}_t| \leq \frac{C}{N^{\alpha-1}}. \tag{7.66}$$

We follow the same lines of reasoning to treat A_3. Indeed, writing for all $t \in [0,T]$

$$X_{N,t} - X_t = \int_0^t (b(X_{N,s}, \mu_{N,s}) - b(X_s, \mu_s)\,ds + \int_0^t \int_{B^c_N} z\,\widetilde{\mathcal{N}}(ds, dz),$$

we deduce as previously that there exists $C > 0$ such that for any $t \in [0,T]$

$$\mathbb{E}W_1(\mu_{N,t}, \mu_t) \leq \frac{C}{N^{\alpha-1}}. \tag{7.67}$$

It remains to treat the term A_2 in (7.65). Using Lemma 7.18 and the regularity properties satisfied by $\phi_{N,u}$ stated in Proposition 7.24, we obtain that the empirical projection $(t, \boldsymbol{x}) \in [0,T) \times (\mathbb{R}^d)^N \mapsto \phi_{N,u}(T-t, \overline{\mu}^N_{\boldsymbol{x}})$ belongs to $\mathcal{C}^{1,2}([0,T) \times (\mathbb{R}^d)^N)$. Moreover, for all $t \in [0,T)$ and $\boldsymbol{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$, we have

$$\partial_{\boldsymbol{x}} \phi_{N,u}(T-t, \overline{\mu}^N_{\boldsymbol{x}}) = \frac{1}{N}\begin{pmatrix} \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-t, \overline{\mu}^N_{\boldsymbol{x}})(x_1) \\ \vdots \\ \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-t, \overline{\mu}^N_{\boldsymbol{x}})(x_N)\end{pmatrix}.$$

We denote by $\boldsymbol{\mathcal{N}}^N$ the Poisson random measure associated with $\boldsymbol{Z}^N$ defined in (7.45) and by $\boldsymbol{\nu}^N$ its associated Lévy measure defined in (7.46). Applying the standard Itô formula for this function and the $(\mathbb{R}^d)^N$-valued process $(\boldsymbol{X}^N_{N,t})_t$ and noticing that the map $t \in [0,T] \mapsto \phi_{N,u}(T-t, \mu_{N,t})$ is constant, we obtain that for all $t \in [0,T)$

$$
\begin{aligned}
&\phi_{N,u}(T-t,\overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t,\mu_{N,t}) - \Big(\phi_{N,u}(T,\overline{\mu}^N_0) - \phi_{N,u}(T,\mu_0)\Big) \\
&= -\int_0^t \partial_t \phi_{N,u}(T-s,\overline{\mu}^N_{N,s})\,ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{N,s})(X^{i,N}_{N,s}) \cdot b(X^{i,N}_{N,s},\overline{\mu}^N_{N,s})\,ds \\
&\quad + \int_0^t \int_{(B_N)^N} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}}) - \partial_{\boldsymbol{x}}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\cdot \boldsymbol{z}\Big]\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \\
&\quad + \int_0^t \int_{(B_N)^N\cap\{|\boldsymbol{z}|<1\}} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \qquad (7.68)\\
&\quad + \int_0^t \int_{(B_N)^N\cap\{|\boldsymbol{z}|\geq 1\}} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&=: I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned}
$$

Note that

$$
\int_0^t \int_{(B_N)^N\cap\{|\boldsymbol{z}|\geq 1\}} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big]\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds
$$

is well-defined. Indeed, recalling that for $z\in\mathbb{R}^d$, $\tilde{\boldsymbol{z}}_{\boldsymbol{i}} = (0,\dots,0,z,0,\dots,0)\in(\mathbb{R}^d)^N$, where z appears in the i-th coordinate, one has by definition of $\boldsymbol{\nu}^N$ in (7.46)

$$
\begin{aligned}
&\int_0^t \int_{(B_N)^N\cap\{|\boldsymbol{z}|\geq 1\}} \Big|\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big|\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \\
&= \int_0^t \int_{(\mathbb{R}^d)^N} \mathbf{1}_{(B_N)^N\cap\{|\boldsymbol{z}|\geq 1\}}(\boldsymbol{z}) \Big|\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big|\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \\
&= \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \mathbf{1}_{(B_N)^N\cap\{|\boldsymbol{z}|\geq 1\}}(\tilde{\boldsymbol{z}}_{\boldsymbol{i}}) \Big|\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big|\,d\nu(z)\,ds \\
&= \sum_{i=1}^N \int_0^t \int_{B_N\setminus B_1} \Big|\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}})\Big|\,d\nu(z)\,ds \\
&\leq \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_N\setminus B_1} \int_0^1 \Big|\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s,m^i_{s,z,w})(X^{i,N}_{N,s^-}+hz)\Big|\,|z|\,dw\,d\nu(z)\,ds \\
&\leq C\int_0^t g_1(T-s)\,ds \int_{B_N\setminus B_1} |z|\,d\nu(z),
\end{aligned}
$$

where $m^i_{s,z,w} := w\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{N,s-}}$. We conclude since g_1 is bounded by Proposition 7.24.

Using the form of the Lévy measure of the particle system in (7.46), we can write

$$
\begin{aligned}
I_3 &= \sum_{i=1}^N \int_0^t \int_{(B_N)^N} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\tilde{\boldsymbol{z}}_i}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}}) \\
&\qquad\qquad - \frac{1}{N}\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-})\cdot z\Big]\, d\nu(z)\, ds \\
&= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_N}\int_0^1 \Big[\frac{\delta}{\delta m}\phi_{N,u}(T-s,m^i_{s,z,w})(X^{i,N}_{N,s^-}+z) \\
&\qquad - \frac{\delta}{\delta m}\phi_{N,u}(T-s,m^i_{s,z,w})(X^{i,N}_{N,s^-}) - \frac{1}{N}\partial_v\frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-})\cdot z\Big]\, dw\, d\nu(z)\, ds,
\end{aligned}
$$

where $m^i_{s,z,w} := w\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\tilde{\boldsymbol{z}}_i} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}}$. In order to make appear the backward Kolmogorov PDE (7.42), we decompose I_3 in the following way

$$
\begin{aligned}
I_3 &= \int_0^t \int_{\mathbb{R}^d}\int_{B_N} \Big[\frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(x+z) - \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(x) \\
&\qquad - \partial_v\frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(x)\cdot z\Big]\, d\nu(z)\, d\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}}(x)\, ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N\int_0^t\int_{B_N}\int_0^1 \Big[\frac{\delta}{\delta m}\phi_{N,u}(T-s,m^i_{s,z,w})(X^{i,N}_{N,s^-}+z) - \frac{\delta}{\delta m}\phi_{N,u}(T-s,m^i_{s,z,w})(X^{i,N}_{N,s^-}) \\
&\qquad + \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-}) - \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-}+z)\Big]\, dw\, d\nu(z)\, ds \\
&=: I_{3,A} + I_{3,B}.
\end{aligned}
$$

Since $(\boldsymbol{X}^N_{N,s})_{s\in[0,T]}$ is càdlàg, we deduce that almost surely, for almost all $s\in[0,t]$,

$$
\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}} = \overline{\mu}^N_{\boldsymbol{X}^N_{N,s}} = \overline{\mu}^N_{N,s}.
$$

Thanks to the backward Kolmogorov PDE (7.42) in Theorem 7.16 applied with $\beta\in[1,\alpha)$ and $\gamma=1$ and justified by Proposition 7.24, one has

$$
\begin{aligned}
I_1 + I_2 + I_{3,A} &= \int_0^t -\partial_t\phi_{N,u}(T-s,\overline{\mu}^N_{N,s}) + \mathscr{L}_N\phi_{N,u}(T-s,\overline{\mu}^N_{N,s})\, ds \\
&= 0.
\end{aligned}
$$

Thus, we obtain the following decomposition

$$
\begin{aligned}
&\phi_{N,u}(T-t,\overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t,\mu_{N,t}) - \Big(\phi_{N,u}(T,\overline{\mu}^N_0) - \phi_{N,u}(T,\mu_0)\Big) \\
&= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_N} \int_0^1 \Big[\frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w})(X^{i,N}_{N,s^-}+z) - \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w})(X^{i,N}_{N,s^-}) \\
&\qquad + \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-}) - \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-}+z)\Big]\, dw\, d\nu(z)\, ds \\
&\quad + \int_0^t \int_{B_N\setminus B_1} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})\Big]\, \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \\
&\quad + \int_0^t \int_{B_1} \Big[\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})\Big]\, \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \\
&= I_{3,B} + I_4 + I_5.
\end{aligned}
\tag{7.69}
$$

It follows that for all $t \in [0,T)$

$$
\mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t,\mu_{N,t})| \le \mathbb{E}|\phi_{N,u}(T,\overline{\mu}^N_0) - \phi_{N,u}(T,\mu_0)| + \mathbb{E}(|I_{3,B}| + |I_4| + |I_5|). \tag{7.70}
$$

We treat each term separately. For $I_{3,B}$, we write

$$
\begin{aligned}
&I_{3,B} \\
&= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_N} \int_{[0,1]^2} \Big[\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w})(X^{i,N}_{N,s^-}+hz) \\
&\qquad - \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s,\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}})(X^{i,N}_{N,s^-}+hz)\Big] \cdot z\, dh\, dw\, d\nu(z)\, ds \\
&= \frac{1}{N^2}\sum_{i=1}^N \int_0^t \int_{B_N} \int_{[0,1]^3} \Big[\frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w,r})(X^{i,N}_{N,s^-}+hz, X^{i,N}_{N,s^-}+z) \\
&\qquad - \frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w,r})(X^{i,N}_{N,s^-}+hz, X^{i,N}_{N,s^-})\Big] \cdot z\, dr\, dh\, dw\, d\nu(z)\, ds \\
&= \frac{1}{N^2}\sum_{i=1}^N \int_0^t \int_{B_N} \int_{[0,1]^4} \Big[\partial_{v'} \frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w,r})(X^{i,N}_{N,s^-}+hz, X^{i,N}_{N,s^-}+kz)z\Big] \\
&\qquad \cdot z\, dk\, dr\, dh\, dw\, d\nu(z)\, ds,
\end{aligned}
\tag{7.71}
$$

where $m^i_{s,z,w,r} := r m^i_{s,z,w} + (1-r)\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}} = \overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}} + rw\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\tilde{\boldsymbol{z}}_i}$. We deduce that there exists a constant $C > 0$ independent of N such that for all $t \in (0,T]$

$$
\begin{aligned}
\mathbb{E}|I_{3,B}| &\le \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_N} g_2(T-s)|z|^2\, d\nu(z)\, ds \\
&\le \frac{C}{N^{\alpha-1}}.
\end{aligned}
\tag{7.72}
$$

Let us recall that $m^i_{s,z,w} = w\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}+\tilde{\boldsymbol{z}}_i} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}}$. It follows from the Cauchy-Schwarz inequality and the L^2-isometry of compensated Poisson random integrals that there exists a constant $C > 0$

independent of N such that we have for all $t \in (0, T]$

$$\begin{aligned}
\mathbb{E}|I_4| &\leq \left(\mathbb{E}\left(\int_0^t \int_{(B_N)^N \cap \{|\boldsymbol{z}|<1\}} \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}}) \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z})\right)^2\right)^{\frac{1}{2}} \\
&= \left(\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_1} \left|\phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}+\tilde{\boldsymbol{z}}_i}) - \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}})\right|^2 d\nu(z)\, ds\right)^{\frac{1}{2}} \qquad (7.73) \\
&= \left(\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_1} \left|\frac{1}{N}\int_{[0,1]^2} \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w})(X^{i,N}_{N,s^-} + hz)\cdot z\, dh\, dw\right|^2 d\nu(z)\, ds\right)^{\frac{1}{2}} \\
&\leq \left(\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_1} \frac{1}{N^2}\, |g_1(T-s)|^2\, C^2 |z|^2\, d\nu(z)\, ds\right)^{\frac{1}{2}} \\
&\leq \frac{C}{\sqrt{N}}.
\end{aligned}$$

Finally, for I_5, BDG's inequality and the fact that $\alpha \in (1,2)$ yield, for all $t \in [0, T)$,

$$\begin{aligned}
\mathbb{E}|I_5| &\leq \left(\mathbb{E}\left(\int_0^t \int_{(B_N)^N \cap \{|\boldsymbol{z}|\geq 1\}} \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}}) \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z})\right)^\alpha\right)^{\frac{1}{\alpha}} \\
&\leq C\left(\mathbb{E}\left[\int_0^t \int_{(B_N)^N \cap \{|\boldsymbol{z}|\geq 1\}} \left|\phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}})\right|^2 \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z})\right]^{\frac{\alpha}{2}}\right)^{\frac{1}{\alpha}} \\
&\leq C\left(\mathbb{E}\int_0^t \int_{(B_N)^N \cap \{|\boldsymbol{z}|\geq 1\}} \left|\phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}+\boldsymbol{z}}) - \phi_{N,u}(T-s, \overline{\mu}^N_{\mathbf{X}^N_{N,s^-}})\right|^\alpha \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z})\right)^{\frac{1}{\alpha}} \qquad (7.74) \\
&= C\left(\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_N\setminus B_1} \left|\frac{1}{N}\int_{[0,1]^2} \partial_v \frac{\delta}{\delta m}\phi_{N,u}(T-s, m^i_{s,z,w})(X^{i,N}_{N,s^-} + hz)\cdot z\, dh\, dw\right|^\alpha d\nu(z)\, ds\right)^{\frac{1}{\alpha}} \\
&\leq C\left(\sum_{i=1}^N \mathbb{E}\int_0^t \int_{B_N\setminus B_1} \left|\frac{1}{N} g_1(T-s)\right|^\alpha |z|^\alpha\, d\nu(z)\, ds\right)^{\frac{1}{\alpha}} \\
&\leq C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}.
\end{aligned}$$

As a consequence of (7.69), (7.72), (7.73), (7.74), and the fact that $(\alpha-1)\wedge\frac{1}{2} > 1-\frac{1}{\alpha}$, there exists $C > 0$ such that for all $N \geq 1$ and $t \in [0, T)$

$$\mathbb{E}|\phi_{N,u}(T-t, \overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t, \mu_{N,t})| \leq \mathbb{E}|\phi_{N,u}(T, \overline{\mu}^N_0) - \phi_{N,u}(T, \mu_0)| + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}. \qquad (7.75)$$

Note that our assumptions ensure that the function $\phi_{N,u}(T, \cdot)$ is Lipschitz continuous with respect to the Wasserstein metric W_1 uniformly with respect to N. It is a consequence of the Kantorovich-Rubinstein theorem since $\frac{\delta}{\delta m}\phi_{N,u}(T, \mu)$ is Lipschitz continuous uniformly with respect to μ and N by Proposition 7.24. We thus obtain for all $N \geq 1$, $t \in [0, T)$

$$\mathbb{E}|\phi_{N,u}(T-t, \overline{\mu}^N_{N,t}) - \phi_{N,u}(T-t, \mu_{N,t})| \leq C\,\mathbb{E}W^1(\overline{\mu}^N_0, \mu_0) + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}. \qquad (7.76)$$

Using (7.65), (7.66), (7.67) and (7.76), we deduce that for all $t \in [0,T)$

$$\mathbb{E}|\phi_{N,u}(T-t,\overline{\mu}_t^N) - \phi_{N,u}(T-t,\mu_t)| \leq C\,\mathbb{E}W^1(\overline{\mu}_0^N,\mu_0) + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}.$$

We conclude the proof of (7.53) by letting t tend to T thanks to the continuity of $\phi_{N,u}$. The estimate (7.54) is also proved. Indeed, coming back to (7.69), we see with the previous computations that I_4 and I_5 are centered because of the martingale property of compensated Poisson random integrals. We thus obtain that

$$|\mathbb{E}(\phi_{N,u}(T-t,\overline{\mu}_{N,t}^N) - \phi_{N,u}(T-t,\mu_{N,t}))| \leq |\mathbb{E}(\phi_{N,u}(T,\overline{\mu}_0^N) - \phi_{N,u}(T,\mu_0))| + \mathbb{E}|I_{3,B}|.$$

The initial data term is handled in the same way as previously and the control of $\mathbb{E}|I_{3,B}|$ has already been done in (7.72). As a consequence, we obtain that for all $t \in [0,T)$

$$|\mathbb{E}(\phi_{N,u}(T-t,\overline{\mu}_t^N) - \phi_{N,u}(T-t,\mu_t))| \leq C\,\mathbb{E}W^1(\overline{\mu}_0^N,\mu_0) + \frac{C}{N^{\alpha-1}},$$

and we conclude by letting t tend to T. Finally, to prove (7.55) in the case where the initial condition belongs to $\mathcal{P}_2(\mathbb{R}^d)$, we use the same argument as at the end of the proof of [CdRF21, Theorem 3.6 page 44]. To do this, we remark that, thanks to Proposition 7.24, for a constant $C>0$ independent of N, we have for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and for all $v, v' \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m}\frac{\delta}{\delta m}\phi_{N,u}(T,\mu)(v,v')\right| \leq C(1+|v|)(1+|v'|).$$

It yields

$$|\mathbb{E}(\phi_{N,u}(T,\overline{\mu}_0^N) - \phi_{N,u}(T,\mu_0))| \leq \frac{C}{N},$$

which ends the proof.

□

7.5 Appendix: Estimates on the density of a truncated stable Ornstein-Uhlenbeck process

Fix $Z := (Z_t)_t$ an α-stable process on $\mathbb{R}^d$, with $\alpha \in (1,2)$. Let us denote by $\mathcal{N}$ the Poisson random measure associated with Z and by $\widetilde{\mathcal{N}}$ its compensated Poisson random measure. We can write, for any $t \in [0,T]$,

$$Z_t = \int_0^t \int_{\mathbb{R}^d} z\,\widetilde{\mathcal{N}}(ds,dz).$$

Writing $y = r\theta \in \mathbb{R}^d\backslash\{0\}$ with $r \in \mathbb{R}_*^+$ and $\theta \in \mathbb{S}^{d-1}$, where $\mathbb{S}^{d-1}$ denotes the unit sphere in $\mathbb{R}^d$, the Lévy measure of Z decomposes as

$$\nu(dy) = d\mu(\theta)\frac{dr}{r^{1+\alpha}},$$

for μ a non-zero finite measure on $\mathbb{S}^{d-1}$. We assume that ν satisfies the following non-degeneracy assumption

(ND) There exists $\eta > 0$ such that for all $\lambda \in \mathbb{R}^d$

$$\eta|\lambda|^2 \leq \int_{\mathbb{S}^{d-1}} |\langle \lambda, \theta \rangle|^2 \, d\mu(\theta).$$

In Appendix 7.5, we denote by $\langle \cdot, \cdot \rangle$ the usual scalar product on $\mathbb{R}^d$ for the sake of clarity. The Lévy symbol associated with Z is given, for $\lambda \in \mathbb{R}^d$, by

$$\psi(\lambda) := \int_{\mathbb{R}^d} \left[e^{i\langle \lambda, y \rangle} - 1 - i\langle \lambda, y \rangle \right] \, d\nu(y).$$

We introduce, for $\delta \in (0, +\infty]$, the truncated symbol ψ_δ defined, for $\lambda \in \mathbb{R}^d$, by

$$\psi_\delta(\lambda) := \int_{|y|<\delta} \left[e^{i\langle \lambda, y \rangle} - 1 - i\langle \lambda, y \rangle \right] \, d\nu(y). \tag{7.77}$$

It corresponds to remove the jump of size bigger than δ from Z.

We now remove the jumps of Z_t having a size bigger than $t^{\frac{1}{\alpha}}$, which is the typical scale of the α-stable process. Let us fix $t \in (0, T]$. We define $\tilde{Z}^1 = (\tilde{Z}^1_s)_s$ by

$$\tilde{Z}^1_s := \int_0^s \int_{|z| \geq t^{\frac{1}{\alpha}}} z \widetilde{\mathcal{N}}(du, dz),$$

and $\tilde{Z}^2_s := Z_s - \tilde{Z}^1_s$. Thus, one has

$$Z_t = \tilde{Z}^1_t + \tilde{Z}^2_t,$$

where $\tilde{Z}^1$ and $\tilde{Z}^2$ are independent Lévy processes with Lévy symbols respectively given, for all $\lambda \in \mathbb{R}^d$, by

$$\psi^1(\lambda) = \psi_\infty(\lambda) - \psi_{t^{\frac{1}{\alpha}}}(\lambda) \quad \text{and} \quad \psi^2(\lambda) = \psi_{t^{\frac{1}{\alpha}}}(\lambda).$$

We also introduce, for $N \in \mathbb{N}$ such that $N \geq T^{\frac{1}{\alpha}}$, the truncated stable process defined, for $s \in [0, T]$, by

$$Z_{N,s} := \int_0^s \int_{|z| \leq N} z \widetilde{\mathcal{N}}(du, dz). \tag{7.78}$$

We set for $s \in [0, T]$

$$\tilde{Z}^3_{N,s} := \int_0^s \int_{t^{\frac{1}{\alpha}} \leq |z| < N} z \mathcal{N}(du, dz),$$

and

$$\tilde{Z}^4_{N,s} = -\int_0^s \int_{t^{\frac{1}{\alpha}} \leq |z| < N} z \, d\nu(z) \, du.$$

We have the following decomposition for $Z_{N,t}$

$$Z_{N,t} = \tilde{Z}^2_t + \tilde{Z}^3_{N,t} + \tilde{Z}^4_{N,t},$$

where the three random variables in the right-hand side term of the preceding equality are mutually independent.

Let $A, B \in \mathcal{M}_d(\mathbb{R})$ be two matrices such that B is invertible and $T > 0$ a finite horizon of time. The Ornstein-Uhlenbeck process $Y = (Y_t)_t$ associated with Z is defined as the solution to the following well-posed SDE

$$dY_t = AY_t dt + B\,dZ_t, \quad t \in [0,T],$$

starting at $t=0$ from 0. Using Itô's formula, one can see that Y can be expressed as

$$Y_t = \int_0^t e^{(t-s)A} B\,dZ_s, \quad t \in [0,T]. \tag{7.79}$$

We also consider the truncated Ornstein-Uhlenbeck process $(Y_{N,t})_t$ defined by (7.79) with Z replaced by the truncated stable process Z_N. It thus writes, for $t \in [0,T]$

$$\begin{aligned} Y_{N,t} &= \int_0^t e^{(t-s)A} B\,d\tilde{Z}^2_s + \int_0^t e^{(t-s)A} B\,d\tilde{Z}^3_{N,s} + \int_0^t e^{(t-s)A} B\,d\tilde{Z}^4_{N,s} \\ &=: \tilde{Y}^2_t + \tilde{Y}^3_{N,t} + \tilde{Y}^4_{N,t}. \end{aligned} \tag{7.80}$$

We begin with the existence of a density for the small jumps process $\tilde{Y}^2$ of the truncated stable Ornstein-Uhlenbeck process Y_N, and on some decay and moment estimates on it.

Proposition 7.25. *For all $t \in (0,T]$ the distribution of $\tilde{Y}^2_t$ has a density with respect to the Lebesgue measure $\tilde{p}^2(t,\cdot)$ which belongs to the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ of regular and rapidly decreasing functions. Moreover, for any $m \geq 0$, and for any multi-index $\beta \in \mathbb{N}^d$, there exists a constant $C_{T,m,\beta} > 0$ such that for all $t \in (0,T]$ and $x \in \mathbb{R}^d$, one has*

$$\left|\partial_x^\beta \tilde{p}^2(t,x)\right| \leq \frac{C_{T,m,\beta}}{t^{\frac{d+|\beta|}{\alpha}}} \left(1 + \frac{|x|}{t^{\frac{1}{\alpha}}}\right)^{-m}. \tag{7.81}$$

Moreover, for all $\gamma \geq 0$, there exists a constant $C > 0$ depending only on $T, d, \alpha, \gamma, \beta, \eta$ such that for all $t \in (0,T]$

$$\int_{\mathbb{R}^d} |x|^\gamma |\partial_x^\beta \tilde{p}^2(t,x)|\,dx \leq C t^{\frac{\gamma - |\beta|}{\alpha}}. \tag{7.82}$$

Proof. We fix $t > 0$. Reasoning as in [Sat99, page 105], the characteristic function of $\tilde{Y}^2_t$ is given by the function

$$p \in \mathbb{R}^d \mapsto \phi_t(p) := \exp\left(\int_0^t \psi^2(B^* e^{sA^*} p)\,ds\right),$$

where A^* denotes the transpose matrix of a matrix A. Changing variables in $v := \frac{s}{t}$ and $\rho := \frac{r}{t^{\frac{1}{\alpha}}}$, one has for $p \in \mathbb{R}^d$

$$\begin{aligned} \phi_t(p) &= \exp\left(\int_0^t \int_0^{t^{\frac{1}{\alpha}}} \int_{\mathbb{S}^{d-1}} e^{i\langle B^* e^{sA^*} p, r\theta\rangle} - 1 - i\langle B^* e^{sA^*} p, r\theta\rangle\,d\mu(\theta)\,\frac{dr}{r^{1+\alpha}}\,ds\right) \\ &= \exp\left(\int_0^1 \int_0^1 \int_{\mathbb{S}^{d-1}} e^{i\langle B^* e^{tvA^*} p, t^{\frac{1}{\alpha}}\rho\theta\rangle} - 1 - i\langle B^* e^{tvA^*} p, t^{\frac{1}{\alpha}}\rho\theta\rangle\,d\mu(\theta)\,\frac{d\rho}{\rho^{1+\alpha}}\,dv\right) \\ &= \exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} t^{\frac{1}{\alpha}} p)\,dv\right). \end{aligned}$$

Reasoning as in [CHZ20b, Lemma 3.2], there exists $\eta > 0$ such that for any $q \in \mathbb{R}^d$

$$\mathrm{Re}(\psi_1(q)) \leq -\eta(|q|^\alpha \wedge |q|^2). \tag{7.83}$$

We have thus for any $p \in \mathbb{R}^d$

$$\left|\exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} t^{\frac{1}{\alpha}} p)\, dv\right)\right| = \exp\left(\int_0^1 \mathrm{Re}(\psi_1(B^* e^{tvA^*} t^{\frac{1}{\alpha}} p))\, dv\right)$$
$$\leq \exp\left(-\eta \int_0^1 |B^* e^{tvA^*} t^{\frac{1}{\alpha}} p|^\alpha \wedge |B^* e^{tvA^*} t^{\frac{1}{\alpha}} p|^2 dv\right).$$

We now prove that there exists $\eta_T > 0$ such that for all $s \in [0,T]$ and for any $q \in \mathbb{R}^d$

$$|B^* e^{sA^*} q| \geq \eta_T |q| \tag{7.84}$$

First, note that

$$\begin{aligned}|B^* e^{sA^*} q|^2 &= \langle BB^* e^{sA^*} q, e^{sA^*} q\rangle \\ &\geq \eta_1 |e^{sA^*} q|^2,\end{aligned}$$

for $\eta_1 > 0$ since BB^* is positive-definite because B is invertible. We prove that there exists $\eta_2 > 0$ such that for any $q \in \mathbb{R}^d$ and $s \in [0,T]$

$$|e^{sA^*} q|^2 \geq \eta_2 |q|^2,$$

which will conclude the proof of (7.84). Reasoning by contradiction, there exists $(s_n)_n \in [0,T]^{\mathbb{N}}$ and $(q_n)_n \in (\mathbb{R}^d)^{\mathbb{N}}$ with $|q_n| = 1$ for all n and such that for all n

$$\langle e^{s_n A^*} e^{s_n A} q_n, q_n\rangle \leq \frac{1}{n}. \tag{7.85}$$

By compactness, one can assume that $(s_n)_n$ converges to $s \in [0,T]$ and that $(q_n)_n$ converges to $q \in \mathbb{R}^d$, with $|q| = 1$. Letting n tend to infinity in (7.85), we obtain that $q = 0$ since e^{sA} is invertible. This is a contradiction.

Thus, there exists a constant $\eta_T > 0$ such that for any $q \in \mathbb{R}^d$ and $t \in [0,T]$

$$\left|\exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} q)\, dv\right)\right| \leq \exp\left(-\eta_T |q|^\alpha \wedge |q|^2\right). \tag{7.86}$$

This proves that the characteristic function of $\tilde{Y}_t^2$ belongs to $L^1(\mathbb{R}^d)$. It follows that the distribution of $\tilde{Y}_t^2$ has a density with respect to the Lebesgue measure denoted by $\tilde{p}^2(t,\cdot)$ and given, for all $x \in \mathbb{R}^d$, by

$$\tilde{p}^2(t,x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle x,p\rangle} \exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} t^{\frac{1}{\alpha}} p)\, dv\right) dp$$

Changing variables in $q := t^{\frac{1}{\alpha}} p$ yields

$$\tilde{p}^2(t,x) = \frac{t^{-\frac{d}{\alpha}}}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\left\langle q, \frac{x}{t^{\frac{1}{\alpha}}}\right\rangle} \exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} q)\, dv\right) dq.$$

Let us define for all $t > 0$ and $q \in \mathbb{R}^d$

$$g_t(q) := \exp\left(\int_0^1 \psi_1(B^* e^{tvA^*} q)\, dv\right).$$

We have proved in (7.86) that for any $q \in \mathbb{R}^d$

$$|g_t(q)| \leq e^{-\eta_T(|q|^\alpha \wedge |q|^2)}. \tag{7.87}$$

It follows that for any $t > 0$, $\tilde{p}^2(t, \cdot)$ belongs to $\mathcal{C}^\infty(\mathbb{R}^d)$ and that for any multi-index $\beta \in \mathbb{N}^d$

$$\partial^\beta \tilde{p}^2(t,x) = \frac{t^{-\frac{d+|\beta|}{\alpha}}}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\left\langle q, \frac{x}{t^{\frac{1}{\alpha}}} \right\rangle} (-iq)^\beta g_t(q)\, dq. \tag{7.88}$$

Now, we prove that for all $t \in (0,T]$ $g_t \in \mathcal{C}^\infty(\mathbb{R}^d)$. For the derivatives of order 1, we differentiate under the integral to obtain that for $j \in \{1, \dots, d\}$

$$\partial_{q_j} g_t(q) = \left(\int_0^1 \int_0^1 \int_{\mathbb{S}^{d-1}} i\rho \left[e^{tvA} B\theta \right]_j \left(e^{i\langle B^* e^{tvA^*} q, \rho\theta \rangle} - 1 \right) d\mu(\xi) \, \frac{d\rho}{\rho^{1+\alpha}}\, dv \right) g_t(q),$$

where $[v]_j$ denotes the $j-$th component of a vector $v \in \mathbb{R}^d$. This term is well-defined since the mean value theorem and the continuity of $t \in [0,T] \mapsto e^{tA}$ ensure that there exists $C_T > 0$ such that for all $q \in \mathbb{R}^d$, $v \in [0,1]$, $\theta \in \mathbb{S}^{d-1}$

$$\left| e^{i\langle q, \rho e^{tvA} B\theta \rangle} - 1 \right| \leq C_T |q| \rho. \tag{7.89}$$

We can easily prove by induction that for any $b \in \mathbb{N}^d$, $\partial^b g_t(q)$ exists and can be expressed as a linear combination of terms of the form

$$\left(\prod_{l=1}^N \left[\int_0^1 \int_0^1 \int_{\mathbb{S}^{d-1}} \prod_{j=1}^{M_l} \left[\rho e^{tvA} B\theta \right]_{k_j} \left(e^{i\langle q, \rho e^{tvA} B\theta \rangle} - \mathbf{1}_{M_l = 1} \right) d\mu(\theta) \, \frac{d\rho}{\rho^{1+\alpha}}\, dv \right] \right) g_t(q),$$

where $N \geq 1$, $M_l \geq 1$ and $k_l \in \{1, \dots, d\}$.

Using (7.87), we deduce that there exists two constants $C > 0$ and $D \geq 1$ depending on T and b such that for all $t \in (0,T]$ and $q \in \mathbb{R}^d$

$$\left| \prod_{l=1}^N \int_0^1 \int_0^1 \int_{\mathbb{S}^{d-1}} \prod_{j=1}^{M_l} \left[\rho e^{tvA} B\theta \right]_{k_l} \left(e^{i\langle q, \rho e^{tvA} B\theta \rangle} - \mathbf{1}_{M_l = 1} \right) d\mu(\theta) \, \frac{d\rho}{\rho^{1+\alpha}}\, dv \right| |g_t(q)| \leq C(1+|q|^D) e^{-\eta_T(|q|^\alpha \wedge |q|^2)}.$$

Indeed, the only difficulty appears when $M_i = 1$ since $\int_0^1 \rho \, \frac{d\rho}{\rho^{1+\alpha}}$ is equal to infinity. However in this case, we can use (7.89) and the fact that

$$\int_0^1 \rho^2 \, \frac{d\rho}{\rho^{1+\alpha}} < +\infty.$$

This shows that for all multi-indices $a, b \in \mathbb{N}^d$

$$\sup_{t \leq T} \| q^a \partial^b g_t(q) \|_{L^1_q(\mathbb{R}^d)} < +\infty. \tag{7.90}$$

We easily see that the same estimates hold replacing $q \in \mathbb{R}^d \mapsto g_t(q)$ by $q \in \mathbb{R}^d \mapsto (-iq)^\beta g_t(q)$. Thus, one has for any $m \geq 0$

$$\sup_{t\leq T}\sup_{x\in\mathbb{R}^d}\left\{(1+|x|)^m\left|\mathcal{F}_q((-iq)^\beta g_t(q))(x)\right|\right\} \leq C\sum_{\gamma\in\mathbb{N}^d,\,|\gamma|\leq m}\sup_{t\leq T}\left\|\partial_q^\gamma((-iq)^\beta g_t(q))\right\|_{L^1_q(\mathbb{R}^d)}$$
$$< +\infty,$$

where $\mathcal{F}_q$ denotes the Fourier transform with respect to the variable $q \in \mathbb{R}^d$. Using (7.88), it follows that for a constant $C = C_{T,m,\beta} > 0$, we have for all $t \in (0,T]$ and for all $x \in \mathbb{R}^d$

$$\left|\partial^\beta \tilde{p}^2(t,x)\right| \leq \frac{C}{t^{\frac{d+|\beta|}{\alpha}}}\left(1+\frac{|x|}{t^{\frac{1}{\alpha}}}\right)^{-m}.$$

This ends the proof of (7.81). The moment estimate (7.82) directly follows from (7.81) by choosing m large enough and from a change of variables.

□

We can now study the density of the truncated Ornstein-Uhlenbeck process Y_N defined in (7.80).

Proposition 7.26. *For all $N \geq 1 \vee T^{\frac{1}{\alpha}}$ and for all $t \in (0,T]$ the distribution of $Y_{N,t}$ has a density with respect to the Lebesgue measure denoted by $p^N(t,\cdot) \in \mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^+)$. Moreover, for any $\beta \in \mathbb{N}^d$ and $\gamma \in [0,\alpha)$ there exists a constant $C > 0$ depending only on $T, d, \alpha, \eta, \beta, \gamma$ such that for any $N \in \mathbb{N}$ and $t \in (0,T]$*

$$\int_{\mathbb{R}^d}|x|^\gamma|\partial_x^\beta p^N(t,x)|\,dx \leq Ct^{\frac{\gamma-|\beta|}{\alpha}}. \tag{7.91}$$

The same results hold for $(Y_t)_t$, defined in (7.79), which has a density $p(t,\cdot)$ satisfying the moment estimate (7.91).

Moreover, for any $N \geq 1\vee T^{\frac{1}{\alpha}}$ and $t_0 \in (0,T]$, p_N belongs to $\mathcal{C}^{1,\infty}([t_0,T]\times\mathbb{R}^d)$ and $p^N(t,\cdot)$, $\partial_t p^N(t,\cdot)$ belong to $\mathcal{S}(\mathbb{R}^d)$ uniformly in time on $[t_0,T]$. More precisely, for all $m \geq 0$, and for any multi-index $\beta \in \mathbb{N}^d$, there exists a constant $C_{t_0,T,m,\beta,N} > 0$ such that for all $x \in \mathbb{R}^d$, one has

$$\sup_{t\in[t_0,T]}\left|\partial_x^\beta p^N(t,x)\right| \leq C_{t_0,T,m,\beta,N}\,(1+|x|)^{-m}, \tag{7.92}$$

and

$$\sup_{t\in[t_0,T]}\left|\partial_x^\beta\partial_t p^N(t,x)\right| \leq C_{t_0,T,m,\beta,N}\,(1+|x|)^{-m}, \tag{7.93}$$

.

Proof. **Proof of (7.91).** Let us fix $t \in (0,T]$. We recall that we have decomposed $Y_{N,t}$ as the following sum of three independent random variables

$$\begin{aligned}Y_{N,t} &= \tilde{Y}_t^2 + \tilde{Y}_{N,t}^3 + \tilde{Y}_{N,t}^4 \\ &= \tilde{Y}_t^2 + \int_0^t\int_{t^{\frac{1}{\alpha}}\leq|z|<N} e^{(t-s)A}Bz\,\mathcal{N}(ds,dz) - \int_0^t\int_{t^{\frac{1}{\alpha}}\leq|z|<N} e^{(t-s)A}Bz\,d\nu(z)\,ds.\end{aligned}$$

The density of $Y_{N,t}$ is thus given for all $x \in \mathbb{R}^d$ by

$$p^N(t,x) = \mathbb{E}(\tilde{p}^2(t,x-\tilde{Y}_{N,t}^3-\tilde{Y}_{N,t}^4)).$$

The regularity of $p^N(t,\cdot)$ follows from Proposition 7.25 by differentiation under the integral. We now fix $\gamma \in [0,\alpha)$. Then, we write by Fubini-Tonelli's theorem

$$\begin{aligned}\int_{\mathbb{R}^d} |x|^\gamma |\partial_x^\beta p^N(t,x)|\,dx &= \int_{\mathbb{R}^d} |x|^\gamma |\mathbb{E}(\partial_x^\beta \tilde{p}^2(t, x - \tilde{Y}^3_{N,t} - \tilde{Y}^4_{N,t}))|\,dx \\ &\leq C \int_{\mathbb{R}^d} \left(|x|^\gamma + \mathbb{E}|\tilde{Y}^3_{N,t}|^\gamma + \mathbb{E}|\tilde{Y}^4_{N,t}|^\gamma\right) |\partial_x^\beta \tilde{p}^2(t,x)|\,dx.\end{aligned}$$

Using the moment estimate (7.82), it remains to prove that for a constant $C > 0$ independent of N, one has for all $t \in [0,T]$

$$\mathbb{E}|\tilde{Y}^3_{N,t}|^\gamma + \mathbb{E}|\tilde{Y}^4_{N,t}|^\gamma \leq C t^{\frac{\gamma}{\alpha}}.$$

For $\tilde{Y}^3_{N,t}$, note that $\int_0^t \int_{t^{\frac{1}{\alpha}} \leq |z|} |z|\,\mathcal{N}(ds,dz)$ has the same distribution as $t^{\frac{1}{\alpha}} \int_0^1 \int_{1 \leq |z|} |z|\,\mathcal{N}(ds,dz)$. This can be seen easily using the characteristic function. Moreover, since $\gamma < \alpha$, we have

$$\mathbb{E}\left|\int_0^1 \int_{1 \leq |z|} |z|\,\mathcal{N}(ds,dz)\right|^\gamma < +\infty.$$

We thus obtain that for all $t \in [0,T]$

$$\begin{aligned}\mathbb{E}|\tilde{Y}^3_{N,t}|^\gamma &= \mathbb{E}\left|\int_0^t \int_{t^{\frac{1}{\alpha}} \leq |z| < N} e^{(t-s)A} Bz\,\mathcal{N}(ds,dz)\right|^\gamma \\ &\leq C_T \mathbb{E}\left|\int_0^t \int_{t^{\frac{1}{\alpha}} \leq |z|} |z|\,\mathcal{N}(ds,dz)\right|^\gamma \\ &\leq C_T t^{\frac{\gamma}{\alpha}}.\end{aligned}$$

Finally, for $\tilde{Y}^4_{N,t}$, one has for all $t \in [0,T]$

$$\begin{aligned}\mathbb{E}|\tilde{Y}^4_{N,t}|^\gamma &= \mathbb{E}\left|\int_0^t \int_{t^{\frac{1}{\alpha}} \leq r < N} \int_{\mathbb{S}^{d-1}} e^{(t-s)A} Br\theta \frac{dr}{r^{1+\alpha}}\,d\mu(\theta)\,ds\right|^\gamma \\ &\leq \mathbb{E}\left|\frac{t^{1/\alpha - 1}}{\alpha - 1} \int_0^t \int_{\mathbb{S}^{d-1}} e^{(t-s)A} B\theta\,d\mu(\theta)\,ds\right|^\gamma \\ &\leq C_T t^{\frac{\gamma}{\alpha}}.\end{aligned}$$

This concludes the proof of (7.91). The same reasoning holds for the moment estimates on $(Y_t)_t$.

Proof of (7.92) and (7.93). As in the proof of Proposition 7.25, we obtain the following expression for the density of Y_t^N

$$p^N(t,x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle x,p\rangle} \exp\left(\int_0^t \psi_N(B^* e^{sA^*} p)\,ds\right) dp, \quad x \in \mathbb{R}^d,$$

where ψ_N was defined in (7.77). We follow the same lines as in the proof of Proposition 7.25. Indeed, reasoning as in Lemma 3.2 in [CHZ20b], there exists $\eta > 0$ such that for any $q \in \mathbb{R}^d$

$$\mathrm{Re}(\psi_N(q)) \leq -\eta(|q|^\alpha \wedge |q|^2), \tag{7.94}$$

It follows that for some constant $\eta_T > 0$, one has for all $p \in \mathbb{R}^d$, $t \in [0,T]$.

$$\exp\left(\int_0^t \psi_N(B^* e^{sA^*} p)\, ds\right) \le \exp(-t\eta_T(|q|^\alpha \wedge |q|^2)).$$

Moreover, we prove that there exists $C > 0$ such that for any $q \in \mathbb{R}^d$

$$|\psi_N(q)| \le C|q|^\alpha, \tag{7.95}$$

Indeed, using Taylor formula, one has

$$\begin{aligned}
|\psi_N(q)| &\le \int_{|y|\le|q|^{-1}} |e^{i\langle q,y\rangle} - 1 - i\langle q,y\rangle|\, d\nu(y) + \int_{|q|^{-1}\le|y|\le N} |e^{i\langle q,y\rangle} - 1 - i\langle q,y\rangle|\, d\nu(y)\\
&\le \int_{|y|\le|q|^{-1}} C|y|^2|q|^2\, d\nu(y) + \int_{|q|^{-1}\le|y|\le N} 2 + |q||y|\, d\nu(y)\\
&\le C|q|^2(|q|^{-1})^{2-\alpha} + |q|^\alpha + C|q|(|q|^{-1})^{\alpha-1}\\
&\le C|q|^\alpha.
\end{aligned}$$

Thanks to (7.95), we prove by differentiation under the integral that $p^N \in \mathcal{C}^{1,\infty}((0,T]\times\mathbb{R}^d)$ and that for all $t \in (0,T]$ and $x \in \mathbb{R}^d$

$$\partial_t p^N(t,x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-i\langle x,p\rangle}\psi_N(B^* e^{tA^*}p)\exp\left(\int_0^t \psi_N(B^* e^{sA^*}p)\, ds\right) dp.$$

As for (7.90), we prove that for all multi-indices $a, b \in \mathbb{N}^d$, we have

$$\begin{aligned}
&\sup_{t\in[t_0,T]} \left\| p^a \partial_p^b \left(\psi_N(B^* e^{tA^*}p)\exp\left(\int_0^t \psi_N(B^* e^{sA^*}p)\, ds\right)\right)\right\|_{L^1_p(\mathbb{R}^d)}\\
&=: C_{t_0,T,a,b,N} < +\infty.
\end{aligned}$$

The same reasoning used in the proof of Proposition 7.25 allows to deduce (7.92) and (7.93).

□

7.6 Appendix: Proof of Proposition 7.24

Proof of Proposition 7.24. We recall that X_N denotes the solution to SDE (7.60) driven by the truncated stable process $(Z_{N,t})_t$ defined in (7.78). We fix $(t,\mu) \in (0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$ and X_0 a random variable with distribution μ. Applying Itô's formula for the function $f : (t,x) \in [0,+\infty) \times \mathbb{R}^d \mapsto e^{-tA}x$, we obtain that for all $t \ge 0$

$$X_{N,t} = e^{tA}X_0 + \int_0^t e^{(t-s)A}A'\mathbb{E}X_{N,s}\, ds + \int_0^t e^{(t-s)A}B\, dZ_{N,s}.$$

By differentiating the map $t \in \mathbb{R}^+ \mapsto e^{-tA}\mathbb{E}X_{N,t}$, we deduce, since A and A' commute, that for all $t \ge 0$

$$\mathbb{E}X_{N,t} = e^{t(A+A')}\mathbb{E}X_0,$$

since Z_N is a centered process. It yields for all $t \geq 0$

$$X_{N,t} = e^{tA}X_0 + \left(\int_0^t e^{(t-s)A}A' e^{s(A+A')}\,ds\right)\mathbb{E}X_0 + \int_0^t e^{(t-s)A}B\,dZ_{N,s}$$
$$=: e^{tA}X_0 + K_t\mathbb{E}X_0 + Y_{N,t}. \tag{7.96}$$

The process $(Y_{N,t})_t$ has been studied in Appendix 7.5. By Proposition 7.26, for $t > 0$, it admits a density $p^N(t,\cdot)$ satisfying the moment estimates (7.91) and the Schwartz estimates (7.92) and (7.93). Since the moment estimates (7.91) are uniform with respect to N, we denote X_N by X, for the sake of clarity, and we omit all the indices N in the following. By Proposition 7.26 and (7.96), we deduce that X_t has a density with respect to the Lebesgue measure depending on the initial data $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and denoted by $q(\mu,t,\cdot)$. Moreover, using (7.96), we have for all $y \in \mathbb{R}^d$

$$q(\mu,t,y) = \int_{\mathbb{R}^d} p(t, y - e^{tA}x - K_t M(\mu))\,d\mu(x), \tag{7.97}$$

where $p = p^N$ is given by Proposition 7.26, and $M(\mu) := \int_{\mathbb{R}^d} x\,d\mu(x)$, for $\mu \in \mathcal{P}_1(\mathbb{R}^d)$.

We deduce, using Proposition 7.26, that $q(\mu,\cdot,y)$ is of class $\mathcal{C}^1$ on $(0,T]$ and that $q(\cdot,t,y)$ admits a linear derivative given for all $v \in \mathbb{R}^d$ by

$$\frac{\delta}{\delta m}q(\mu,t,y)(v) = p(t, y - e^{tA}v - K_tM(\mu)) - \int_{\mathbb{R}^d} \partial_y p(t, y - e^{tA}x - K_tM(\mu))\cdot(K_tv)\,d\mu(x).$$

It is easy to see by differentiation under the integral that for all $t \in (0,T]$ and $y \in \mathbb{R}^d$, $\frac{\delta}{\delta m}q(\mu,t,y)$ is of class $\mathcal{C}^1$ and that one has for all $v \in \mathbb{R}^d$

$$\partial_v \frac{\delta}{\delta m}q(\mu,t,y)(v) = -e^{tA^*}\partial_y p(t, y - e^{tA}v - K_tM(\mu)) - \int_{\mathbb{R}^d} K_t^*\partial_y p(t, y - e^{tA}x - K_tM(\mu))\,d\mu(x).$$

Using Proposition 7.26, we see that $\partial_t q$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times(0,T]\times\mathbb{R}^d$ and that $\frac{\delta}{\delta m}q$ and $\partial_v\frac{\delta}{\delta m}q$ are continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times(0,T]\times\mathbb{R}^d\times\mathbb{R}^d$.

Let us fix $u \in \mathcal{C}$. Since SDE (7.51) is time-homogeneous, the associated function ϕ_u is thus given, for any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and for any $t \in (0,T]$, by $\phi_u(t,\mu) = u(\theta(\mu,t))$, where $\theta(\mu,t) := q(\mu,t,y)\,dy$ is the distribution of any solution to SDE (7.51) starting as $t = 0$ from the distribution $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. Moreover, reasoning exactly as in the proof on Proposition 2.3 of [CdRF21], we prove using Proposition 7.26 that ϕ_u satisfies the regularity properties (1) and (2) of Proposition 7.24. For the time derivative, it ensures that

$$\partial_t\phi_u(t,\mu) = \int_{\mathbb{R}^d}\frac{\delta}{\delta m}u(\theta(\mu,t))(y)\partial_t q(\mu,t,y)\,dy.$$

Moreover, one has for all $v \in \mathbb{R}^d$

$$\frac{\delta}{\delta m}\phi_u(t,\mu)(v) = \int_{\mathbb{R}^d}\frac{\delta}{\delta m}u(\theta(\mu,t))(y)\frac{\delta q}{\delta m}(\mu,t,y)(v)\,dy,$$

and that

$$\partial_v \frac{\delta}{\delta m}\phi_u(t,\mu)(v) = -\int_{\mathbb{R}^d} \frac{\delta}{\delta m}u(\theta(\mu,t))(y)e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\,dy \\ -\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{\delta}{\delta m}u(\theta(\mu,t))(y)K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))\,d\mu(x)\,dy. \tag{7.98}$$

Using Proposition 7.26, we obtain that for all $x\in\mathbb{R}^d$

$$\int_{\mathbb{R}^d}\partial_y p(t,y-e^{tA}x-K_tM(\mu))\,dy=0. \tag{7.99}$$

We can thus write (7.98) in the following centered form

$$\begin{aligned}&\partial_v \frac{\delta}{\delta m}\phi_u(t,\mu)(v)\\ &=-\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v+K_tM(\mu))\right)e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\,dy\\ &\quad-\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}x+K_tM(\mu))\right)K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))\,d\mu(x)\,dy.\end{aligned}$$

Using the Lipschitz assumption on $\frac{\delta}{\delta m}u$ as well as the moment estimate in Proposition 7.26, we deduce that there exists $C>0$ depending only on T, which may change from line to line, such that for any $v\in\mathbb{R}^d$

$$\begin{aligned}\left|\partial_v \frac{\delta}{\delta m}\phi_u(t,\mu)(v)\right| &\le C\int_{\mathbb{R}^d}|y||\partial_y p(t,y)|\,dy\\ &\le C.\end{aligned} \tag{7.100}$$

Similarly, one has for all $v\in\mathbb{R}^d$

$$\begin{aligned}&\partial_v^2 \frac{\delta}{\delta m}\phi_u(t,\mu)(v)\\ &=-\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v+K_tM(\mu))\right)e^{tA^*}\partial_y^2 p(t,y-e^{tA}v-K_tM(\mu))e^{tA}\,dy.\end{aligned}$$

Using the Lipschitz assumption on $\frac{\delta}{\delta m}u$ as well as the moment estimate in Proposition 7.26, we deduce that there exists $C>0$ depending only on T, such that for any $v\in\mathbb{R}^d$

$$\begin{aligned}\left|\partial_v^2 \frac{\delta}{\delta m}\phi_u(t,\mu)(v)\right| &\le C\int_{\mathbb{R}^d}|y||\partial_y^2 p(t,y)|\,dy\\ &\le Ct^{-\frac{1}{\alpha}}.\end{aligned} \tag{7.101}$$

Following again the same lines as in Proposition 2.3 of [CdRF21], we obtain that for all $v\in\mathbb{R}^d$, the function $\partial_v\frac{\delta}{\delta m}\phi_u(t,\cdot)(v)$ admits a linear derivative given for all $v'\in\mathbb{R}^d$ by

$$
\begin{aligned}
&\frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_u(t,\mu)(v,v') \\
&= -\int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy' \\
&\quad + \int_{\mathbb{R}^{3d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu)) \\
&\qquad\qquad \partial_y p(t,y'-e^{tA}x-K_tM(\mu))\cdot(K_tv')\,dy\,dy'\,d\mu(x) \\
&\quad + \int_{\mathbb{R}^{d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) e^{tA^*}\partial_y^2 p(t,y-e^{tA}v-K_tM(\mu))K_tv'\,dy \\
&\quad - \int_{\mathbb{R}^{d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) K_t^*\partial_y p(t,y-e^{tA}v'-K_tM(\mu))\,dy \\
&\quad - \int_{\mathbb{R}^{3d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy'\,d\mu(x) \\
&\quad + \int_{\mathbb{R}^{4d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu)) \\
&\qquad\qquad \Big(\partial_y p(t,y'-e^{tA}x'-K_tM(\mu))\cdot K_tv'\Big)\,dy\,dy'\,d\mu(x')\,d\mu(x) \\
&\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) K_t^*\partial_y^2 p(t,y-e^{tA}x-K_tM(\mu))K_tv'\,dy\,d\mu(x).
\end{aligned}
$$

By differentiation under the integral, we deduce that for all $v, v' \in \mathbb{R}^d$

$$
\begin{aligned}
&\partial_{v'}\frac{\delta}{\delta m}\partial_v \frac{\delta}{\delta m}\phi_u(t,\mu)(v,v') \\
&= \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\otimes e^{tA^*}\partial_y p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy' \\
&\quad + \int_{\mathbb{R}^{3d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\otimes K_t^*\partial_y p(t,y'-e^{tA}x-K_tM(\mu))\,dy\,dy'\,d\mu(x) \\
&\quad + \int_{\mathbb{R}^{d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) e^{tA^*}\partial_y^2 p(t,y-e^{tA}v-K_tM(\mu))K_t\,dy \\
&\quad + \int_{\mathbb{R}^{d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) K_t^*\partial_y^2 p(t,y-e^{tA}v'-K_tM(\mu))e^{tA}\,dy \\
&\quad + \int_{\mathbb{R}^{3d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))\otimes e^{tA^*}\partial_y p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy'\,d\mu(x) \\
&\quad + \int_{\mathbb{R}^{4d}} \frac{\delta}{\delta m}\frac{\delta}{\delta m} u(\theta(\mu,t))(y,y') K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu)) \\
&\qquad\qquad \otimes\Big(K_t^*\partial_y p(t,y'-e^{tA}x'-K_tM(\mu))\Big)\,dy\,dy'\,d\mu(x)\,d\mu(x') \\
&\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} u(\theta(\mu,t))(y) K_t^*\partial_y^2 p(t,y-e^{tA}x-K_tM(\mu))K_t\,dy\,d\mu(x).
\end{aligned}
$$

It follows from (7.99) that we can write

$$
\begin{aligned}
&\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(v,v')\\
&=\int_{\mathbb{R}^{2d}}\Big(\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(y,y')-\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v+K_tM(\mu),e^{tA}v'+K_tM(\mu))\Big)\\
&\qquad\qquad e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\otimes e^{tA^*}\partial_y p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy'\\
&\quad+\int_{\mathbb{R}^{3d}}\Big(\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(y,y')-\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v+K_tM(\mu),e^{tA}x+K_tM(\mu))\Big)\\
&\qquad\qquad e^{tA^*}\partial_y p(t,y-e^{tA}v-K_tM(\mu))\otimes K_t^*\partial_y p(t,y'-e^{tA}x-K_tM(\mu))\,dy\,dy'\,d\mu(x)\\
&\quad+\int_{\mathbb{R}^{d}}\Big(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v+K_tM(\mu))\Big)\,e^{tA^*}\partial_y^2 p(t,y-e^{tA}v-K_tM(\mu))K_t\,dy\\
&\quad+\int_{\mathbb{R}^{d}}\Big(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}v'+K_tM(\mu))\Big)\,K_t^*\partial_y^2 p(t,y-e^{tA}v'-K_tM(\mu))e^{tA}\,dy\\
&\quad+\int_{\mathbb{R}^{3d}}\Big(\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(y,y')-\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}x+K_tM(\mu),e^{tA}v'+K_tM(\mu))\Big)\\
&\qquad\qquad K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))\otimes e^{tA^*}\partial_y p(t,y'-e^{tA}v'-K_tM(\mu))\,dy\,dy'\,d\mu(x)\\
&\quad+\int_{\mathbb{R}^{4d}}\Big(\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(y,y')-\frac{\delta}{\delta m}\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}x+K_tM(\mu),e^{tA}x'+K_tM(\mu))\Big)\\
&\qquad K_t^*\partial_y p(t,y-e^{tA}x-K_tM(\mu))\otimes\Big(K_t^*\partial_y p(t,y'-e^{tA}x'-K_tM(\mu))\Big)\,dy\,dy'\,d\mu(x)\,d\mu(x')\\
&\quad+\int_{\mathbb{R}^{2d}}\Big(\frac{\delta}{\delta m}u(\theta(\mu,t))(y)-\frac{\delta}{\delta m}u(\theta(\mu,t))(e^{tA}x+K_tM(\mu))\Big)\,K_t^*\partial_y^2 p(t,y-e^{tA}x-K_tM(\mu))K_t\,dy\,d\mu(x).
\end{aligned}
$$

Using the Lipschitz assumption on $\frac{\delta}{\delta m}u$ and $\frac{\delta}{\delta m}\frac{\delta}{\delta m}u$, as well as the moment estimate in Proposition 7.26, we deduce that there exists $C>0$ depending only on T such that for any $v,v'\in\mathbb{R}^d$

$$
\begin{aligned}
\Big|\partial_{v'}\frac{\delta}{\delta m}\partial_v\frac{\delta}{\delta m}\phi_u(t,\mu)(v,v')\Big| &\le C\Big(\int_{\mathbb{R}^d}|y||\partial_y^2 p(t,y)|\,dy+\int_{\mathbb{R}^d}|y||\partial_y p(t,y)|\,dy\int_{\mathbb{R}^d}|\partial_y p(t,y)|\,dy\Big)\\
&\le Ct^{-\frac{1}{\alpha}}. \qquad (7.102)
\end{aligned}
$$

Note that the right-hand side term belongs to $L^1(0,T)$ since $\alpha\in(1,2)$. Thus, (7.100), (7.101) and (7.102) ensure that the point (3) is Proposition 7.24 is satisfied. It concludes the proof. □

Chapter 8

Quantitative weak propagation of chaos for stable-driven McKean-Vlasov SDEs

This chapter corresponds to the article [Cav22b]. It has been submitted for publication.

Abstract. We consider a general McKean-Vlasov stochastic differential equation driven by a rotationally invariant α-stable process on $\mathbb{R}^d$, with $\alpha \in (1,2)$. We assume that the diffusion coefficient is the identity matrix and that the drift is bounded and Hölder continuous with respect to both space and measure variables, in some precise sense. The main goal of this work is to prove new weak propagation of chaos estimates for the associated mean-field interacting particle system. We also establish a pointwise control on the difference between the density of one particle and the density of the limiting McKean-Vlasov SDE. Our study relies on the regularizing properties and the dynamics of the semigroup associated with the McKean-Vlasov stochastic differential equation, which acts on functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$, the space of probability measures on $\mathbb{R}^d$ having a finite moment of order $\beta \in (1,\alpha)$. More precisely, the dynamics of the semigroup is described by a backward Kolmogorov partial differential equation defined on the strip $[0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$.

8.1 Introduction

In this work, we are interested in the following McKean-Vlasov Stochastic Differential Equation (SDE) driven by $Z = (Z_t)_t$ a rotationally invariant α-stable process on $\mathbb{R}^d$, with $\alpha \in (1,2)$,

$$\begin{cases} dX_t = b(t, X_t, [X_t])\, dt + dZ_t, \quad t \in [0,T], \\ \mu_t := [X_t], \\ X_0 = \xi, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d), \end{cases} \tag{8.1}$$

where $T > 0$ is fixed, $[\xi]$ denotes the distribution of the random variable ξ which is independent of Z, $\mathcal{P}(\mathbb{R}^d)$ is the space of probability measures on $\mathbb{R}^d$ and where $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ is the drift coefficient. This kind of SDEs are also called nonlinear SDEs since the associated Fokker-Planck equation solved by the flow of marginal distributions $(\mu_t)_t$ is nonlinear. We also focus on the associated mean-field interacting particle system defined, for all $N \geq 1$, by

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\, dt + dZ_t^i, \quad t \in [0,T], \quad i \in \{1, \dots, N\}, \\ \overline{\mu}_t^N := \frac{1}{N} \sum_{j=1}^N \delta_{X_t^{j,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{8.2}$$

where $(Z^i, \xi^i)_{i\geq 1}$ are i.i.d. copies of (Z, ξ). The connection between (8.1) and (8.2) is that the McKean-Vlasov SDE (8.1) describes the dynamics of one particle of the system (8.2), when the number of particles N tends to infinity. A stronger property that is also expected to hold true in general is the so-called propagation of chaos. It states that for all $k \geq 1$, the dynamics of k particles is described when N tends to infinity by k independent copies of the limiting McKean-Vlasov SDE (8.1). This was originally studied by McKean [McK67] and then investigated by Sznitman [Szn91], when Z is a Brownian motion. These mean-field systems have many applications for example in physics (kinetic theory), in biology to study the motion of a cell population for example, in neuroscience to model the interactions between neurons, in social sciences to describe self-organization behaviors and also in the Mean-Field Games theory.

Propagation of chaos can be considered in the weak sense, i.e. in distribution through the convergence of the empirical measure $\overline{\mu}^N$ (see [Szn91]), or in the strong sense, i.e. at the level of paths by coupling. The terminology of weak and strong propagation of chaos used here is based of the corresponding properties for numerical schemes for SDEs. It can be qualitative or quantitative when a rate of convergence is shown. By quantitative weak propagation of chaos, we precisely mean to find a rate of convergence of $\mathbb{E}|\phi(\overline{\mu}_t^N) - \phi(\mu_t)|$ and $|\mathbb{E}(\phi(\overline{\mu}_t^N) - \phi(\mu_t))|$, for ϕ in a large class of functions defined on the space of probability measures.

When Z is a Brownian motion, McKean-Vlasov SDEs and mean-field systems have been widely studied. In addition to McKean [McK67] and Sznitman [Szn91], one can mention for example Gärtner [Gä88], Méléard [Mé96], Malrieu [Mal03], Mischler et al. [MMW15], Jabin and Wang [JW18], Lacker [Lac18, Lac21], Tomašević [Tom20] and Jabir [Jab19]. For a detailed review on the topic of propagation of chaos, we refer to [CD22a, CD22b]. The recent development of Mean-Field Games has given a new impulsion to study these systems and the associated propagation of chaos. In particular, it provides a new formalism and new tools, such as the notion of master equations, which are Partial Differential Equations (PDEs) on the space of measures, that allow to revisit or generalize some of the preceding works. In this direction, one can mention for example the book of Carmona and Delarue [CD18a], Chaudru de Raynal and Frikha [CdRF22, CdRF21], Chassagneux et al. [CST22], Delarue and Tse [DT21] and Jourdain and Tsé [JT21]. Such new tools are at the core of this work.

It is natural to consider other types of noise such as Lévy processes, which are also largely used to model physical systems (Lévy flights and anomalous diffusion), see e.g. [MJW01] for the physical point of view and [JMW05] for the mathematical point of view. The propagation of chaos phenomenon has also been studied for McKean-Vlasov SDEs driven by Lévy processes. In [Gra92a], following the approach of Sznitman [Szn91], Graham proves qualitative weak propagation of chaos under Lipschitz assumptions (with respect to the 1- Wasserstein metric W_1 for the measure variable) for a mean-field system driven by a Poisson random measure and its compensated measure. He works in the L^1 framework, i.e. the Poisson random measure is associated with a Poisson process having a finite moment of order 1. When the driving noise is a Lévy process having a finite moment of order 2, we refer to Jourdain et al. [JMW07]. The authors prove quantitative strong propagation of chaos in L^2 under standard Lipschitz assumptions on the drift and diffusion coefficients (with respect to the 2-Wasserstein metric W_2). We also refer to Neelima et al. [NBK+20] where quantitative strong propagation of chaos is proved in L^2, relaxing the assumptions of [JMW07]. We also mention [MMW15], where Mischler, Mouhot and Wennberg exhibit conditions leading to weak propagation of chaos estimates. As an application, they study an inelastic Boltzmann collision jump process. In the one-dimensional case, Frikha and Li [FL21] study a McKean-Vlasov SDE driven by a compensated Poisson random measure with positive jumps. They prove quantitative strong propagation of chaos in L^1 under one-sided Lipschitz assumptions on the coefficients (with respect to W_1 for the measure variable).

The aim of this work is to prove quantitative weak propagation of chaos for (8.1), assuming that the drift b is bounded and Hölder continuous in some precise sense with respect to both space and measure variables (see Assumption (H3) for the precise assumptions). The method that we used relies on the semigroup acting on functions defined on $\mathcal{P}(\mathbb{R}^d)$, associated with the McKean-Vlasov SDE. We will give more details below. The study in the Brownian case has been made by Chaudru de Raynal and Frikha [CdRF22, CdRF21], under similar assumptions. The main difficulties here are the following. Firstly, the dependence of the drift b with respect to the measure variable is general, so that differential calculus for functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$ is needed. We use the notion of linear (functional) derivative (see Definition 8.35). Secondly, since the coefficients are not Lipschitz continuous, one needs to benefit from a regularization by noise phenomenon, in particular to prove the well-posedness of (8.1). Thirdly, since Z is a Lévy process having an infinite moment of order 2, another difficulty here is the impossibility of working in L^2. It is thus impossible to rely on the tools, mentioned above, developed for Mean-Field Games in the L^2 framework. Moreover, the presence of jumps induces supplementary technical difficulties to develop such tools.

In the first part of this work, the weak well-posedness of (8.1) is established through the related nonlinear martingale problem by using the Banach fixed point theorem applied on a suitable complete metric space (see Assumption (H2) and Theorem 8.2). Notice that the well-posedness of (8.1) was already shown in [FKM21] by using the same fixed point argument. The fixed point theorem is only proved to hold in small time in [FKM21], which is enough for the well-posedness. However, in the following, we need to apply the Banach fixed point theorem on the whole time interval $[0, T]$ in order to use Picard iterations to approximate the solution to (8.1). That is why we do the proof a bit differently to obtain this result.

We then study the regularity of the transition density associated with (8.1), in particular the differentiability with respect to the initial distribution $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, with $\beta \in (1, \alpha)$, where $\mathcal{P}_\beta(\mathbb{R}^d)$ is the space of probability measures μ on $\mathbb{R}^d$ having a finite moment of order β. Let us introduce, for $x \in \mathbb{R}^d$ and $s \in [0, T)$, the following decoupled stochastic flow associated to SDE (8.1)

$$\begin{cases} dX_t^{s,x,\mu} = b(t, X_t^{s,x,\mu}, [X_t^{s,\mu}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,x,\mu} = x, \end{cases} \tag{8.3}$$

where $[X_t^{s,\mu}]$ denotes the distribution of the solution to (8.1) at time t, starting at time s by any random variable ξ with distribution $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. The notation $[X_t^{s,\mu}]$ makes sense since by weak well-posedness, the distribution of the solution to (8.1) depends on the initial condition only through its distribution μ. The distributions of $X_t^{s,x,\mu}$ and $X_t^{s,\mu}$ admit densities respectively denoted by $p(\mu, s, t, x, \cdot)$ and $p(\mu, s, t, \cdot)$. Moreover, they satisfy the following relation stemming from the well-posedness of the martingale problem

$$p(\mu, s, t, y) = \int_{\mathbb{R}^d} p(\mu, s, t, x, y)\, d\mu(x). \tag{8.4}$$

Let us fix $t \in (0, T]$ and $y \in \mathbb{R}^d$. We study the regularity of the map $(s, x, \mu) \in [0, t) \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto p(\mu, s, t, x, y)$ and we prove in Theorem 8.4 that p is solution to the following backward Kolmogorov PDE

$$\begin{cases} \partial_s p(\mu, s, t, x, y) + \mathcal{L}_s p(\cdot, s, t, \cdot, y)(\mu, x) = 0, \quad \forall (\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d, \\ p(\mu, s, t, x, \cdot) \underset{s \to t-}{\longrightarrow} \delta_x, \quad \text{in the weak sense,} \end{cases} \tag{8.5}$$

where $\mathcal{L}_s$ is defined, for smooth enough functions h on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$, by

$$\begin{aligned} \mathcal{L}_s h(\mu, x) := {} & b(s, x, \mu) \cdot \partial_x h(\mu, x) + \int_{\mathbb{R}^d} [h(\mu, x+z) - h(\mu, x) - z \cdot \partial_x h(\mu, x)] \, \frac{dz}{|z|^{d+\alpha}} \\ & + \int_{\mathbb{R}^d} b(s, v, \mu) \cdot \partial_v \frac{\delta}{\delta m} h(\mu, x)(v)\, d\mu(v) \\ & + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} h(\mu, x)(v+z) - \frac{\delta}{\delta m} h(\mu, x)(v) - z \cdot \partial_v \frac{\delta}{\delta m} h(\mu, x)(v) \right] \frac{dz}{|z|^{d+\alpha}}\, d\mu(v), \end{aligned} \tag{8.6}$$

where $\frac{\delta}{\delta m}$ denotes the linear derivative (see Definition 8.35). Estimates and Hölder controls on the derivatives of p with respect to s, x and μ are proved in Theorem 8.18 under Assumption (H3) by using Picard iterations and the parametrix method.

Once that the regularity of the transition density has been studied, we focus on the regularizing properties of the semigroup, acting on functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$, associated with (8.1). It is at the core of the method to prove the propagation of chaos. For a function $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$, the action of the McKean-Vlasov semigroup on ϕ is given by the function U defined by

$$U : (t, \mu) \in [0, T] \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \phi([X_T^{t,\mu}]), \tag{8.7}$$

where $[X_T^{t,\mu}]$ denotes the distribution of the solution to (8.1) at time T and starting at time t with a random variable ξ with distribution $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. More precisely, we prove that if $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ has a linear derivative that is uniformly Hölder continuous, then the map U defined by (8.7) is more regular. Moreover, estimates and Hölder controls on the derivative of U with respect to the measure variable are proved. Inspired by [CD18a] (see (5.128)), [CDLL19] and [CdRF22], we describe the dynamics of the semigroup in Theorem 8.9 by showing that U is the unique classical solution to the backward Kolmogorov PDE

$$\begin{cases} \partial_t U(t, \mu) + \mathscr{L}_t U(t, \cdot)(\mu) = 0, \quad \forall (t, \mu) \in [0, T) \times \mathcal{P}_\beta(\mathbb{R}^d), \\ U(T, \mu) = \phi(\mu), \quad \forall \mu \in \mathcal{P}_\beta(\mathbb{R}^d), \end{cases} \tag{8.8}$$

where $\mathscr{L}_t$ is defined, for smooth enough functions h on $\mathcal{P}_\beta(\mathbb{R}^d)$, by

$$\begin{aligned}\mathscr{L}_t h(\mu) := &\int_{\mathbb{R}^d} b(t,v,\mu)\cdot\partial_v \frac{\delta}{\delta m}h(\mu)(v)\,d\mu(v) \\ &+\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left[\frac{\delta}{\delta m}h(\mu)(v+z)-\frac{\delta}{\delta m}h(\mu)(v)-z\cdot\partial_v\frac{\delta}{\delta m}h(\mu)(v)\right]\frac{dz}{|z|^{d+\alpha}}\,d\mu(v).\end{aligned}\tag{8.9}$$

Then, we use the preceding results to prove quantitative propagation of chaos for the particle system (8.2) towards the limiting McKean-Vlasov SDE (8.1).

Firstly, we prove in Theorem 8.12 weak propagation of chaos estimates for the particle system (8.2), as defined above. We refer the reader to Remark 8.14 for a comparison with the existing literature. The method that we use relies on the solution to the backward Kolmogorov PDE (8.8) U defined in (8.7), and thus on the McKean-Vlasov semigroup. This strategy was originally described in [CD18a, pages 506 – 508], inspired by [CDLL19] and [MMW15], and was employed for example in [CST22, DT21, CdRF21]. Let us describe the main ideas. We begin by computing the time derivative of the map $t\in[0,T)\mapsto U(t,\overline{\mu}^N_t)$ by applying the standard Itô formula for the empirical projection $(t,x_1,\dots,x_N)\in[0,T]\times(\mathbb{R}^d)^N\mapsto U\left(t,\frac{1}{N}\sum_{k=1}^N\delta_{x_k}\right)$ and for the particle system. Noting that $t\in[0,T]\mapsto U(t,\mu_t)$ is constant, we naturally expect that the time derivative previously computed tends to 0 as N converges to infinity. This convergence has to be shown with an explicit rate of convergence using the PDE satisfied by U and some estimates on U. Finally, we express $\phi(\overline{\mu}^N_T)-\phi(\mu_T)=U(T,\overline{\mu}^N_T)-U(T,\mu_T)$ as the sum of $U(0,\overline{\mu}^N_0)-U(0,\mu_0)$ plus a remainder term related to the time derivative which was previously estimated. Since the initial data are i.i.d., the first term is controlled by standard estimates, for example those in [FG15].

Secondly, we focus on the approximation of the distribution of one particle of the system (8.2) by the limiting McKean-Vlasov process at the level of densities in Theorem 8.16. We prove a pointwise error bound between the density of one particle and the density of the limiting McKean-Vlasov SDE. This allows to quantify the decrease with respect to N of the total variation distance between the distributions of one particle and the solution to the McKean-Vlasov SDE. The method used to prove this result relies on similar ideas than presented in the preceding paragraph. It was used in the Brownian case in [CdRF21] (see also references therein). Instead of considering the semigroup associated with the McKean-Vlasov SDE, we work with the not decoupled density of the solution to the McKean-Vlasov SDE (8.1), which is somehow a fundamental solution to the Backward Kolmogorov PDE (8.8) (see Theorem 8.7). Denoting by $p(\mu,s,t,\cdot)$ the density of the McKean-Vlasov SDE (8.1) at time t and starting at time s from a random variable with distribution $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, the idea is to study the dynamics of the map $s\in[0,t)\mapsto p(\overline{\mu}^N_s,s,t,y)$, for any $y\in\mathbb{R}^d$. The two ingredients are the following. On the one hand, we quantify the error with respect to N between $\mathbb{E}p(\overline{\mu}^N_s,s,t,y)$ and $p(\mu_s,s,t,y)=p(\mu_0,0,t,y)$. On the other hand, we prove that when s tends to t, then $\mathbb{E}p(\overline{\mu}^N_s,s,t,\cdot)$ converges pointwise to the density of one particle at time t and starting at time 0 from μ_0.

The paper is organized as follows. In Section 8.2, we present our results and we comment them. The proofs are given in the next sections. Section 8.3 is dedicated to prove the weak well-posedness of the McKean-Vlasov SDE (8.1) (Theorem 8.2). Then, we study the regularity of its transition density in Section 8.4 (Theorem 8.18). Section 8.5 is devoted to the proof of the regularizing properties of the semigroup and the backward Kolmogorov PDE that describes its dynamics (Theorem 8.9). In Section 8.6, we prove the quantitative weak propagation of chaos result (Theorem 8.12) and in Section 8.7 the error bound for the approximation of the distribution of one particle by the limiting McKean-Vlasov SDE at the level of densities (Theorem 8.16). We prove in Section 8.8 the technical proposition leading to the regularity of the transition density and the related estimates stated in Theorem 8.18. In Appendix 8.9, we gather some definitions and propositions related to differential calculus for functions defined on $\mathcal{P}(\mathbb{R}^d)$ or $\mathcal{P}_\beta(\mathbb{R}^d)$. Finally, Appendix 8.10 aims at presenting the parametrix method and estimates on the density of the solution to a linear stable-driven SDE, which are the core of the proof of Theorem 8.18.

Let us finally introduce some notations used several times in the article.

Notations

- $\mathcal{P}(\mathbb{R}^d)$ denotes the set of probability measures on $\mathbb{R}^d$.
- d_{TV} is the total variation metric on $\mathcal{P}(\mathbb{R}^d)$.
- $\mathcal{P}_\beta(\mathbb{R}^d)$ denotes the set of probability measures μ on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |x|^\beta \, d\mu(x) < +\infty$, for $\beta > 0$. It is equipped with the Wasserstein metric of order β denoted by W_β, which makes it complete. Denoting by $\Pi(\mu,\nu)$ the set of couplings between two probability measures $\mu,\nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the metric W_β is defined by

$$W_\beta(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^\beta \, d\pi(x,y)\right)^{\frac{1}{\beta}\wedge 1}.$$

- $M_\beta(\mu)$ denotes the moment of order β of $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ defined by

$$M_\beta(\mu) := \left(\int_{\mathbb{R}^d} |x|^\beta \, d\mu(x)\right)^{\frac{1}{\beta}\wedge 1}.$$

- $[\xi]$ denotes the distribution of the random variable ξ.
- $\overline{\mu}^N_{\boldsymbol{x}} := \frac{1}{N}\sum_{k=1}^N \delta_{x_k}$ denotes the empirical measure, for $\boldsymbol{x} = (x_1,\dots,x_N) \in (\mathbb{R}^d)^N$.
- $\tilde{\boldsymbol{z}}_{\boldsymbol{k}} := (0,\dots,z,\dots,0) \in (\mathbb{R}^d)^N$ for $z \in \mathbb{R}^d$, where z appears in the k-th position.
- B_r is the open ball in $\mathbb{R}^d$ centered at 0 and of radius r for the euclidean norm.
- B_r^c denotes the complementary of B_r.
- $a \wedge b$ denotes the minimum between a and b.
- $a \vee b$ denotes the maximum between a and b.
- $\mathcal{B}$ denotes the Beta function defined, for all $x, y > 0$, by

$$\mathcal{B}(x,y) := \int_0^1 (1-t)^{-1+x} t^{-1+y} \, dt.$$

- C is a generic constant that may depend only on the fixed parameters of the problem and which may change from line to line.

8.2 Overview on the main results and comments

Let us fix $Z = (Z_t)_t$ a rotationally invariant α-stable process on $\mathbb{R}^d$ with $\alpha \in (1,2)$. Its associated Poisson random measure is denoted by $\mathcal{N}$ and the compensated Poisson random measure by $\widetilde{\mathcal{N}}$. Since $\alpha \in (1,2)$, by [Sat99, Remark 14.6 and Theorem 14.7], we can write for all $t \geq 0$

$$Z_t = \int_0^t \int_{\mathbb{R}^d} z \, \widetilde{\mathcal{N}}(ds, dz).$$

The density of Z_t is denoted by $q(t,\cdot)$ and the Lévy measure ν of Z is given by

$$d\nu(z) := \frac{dz}{|z|^{d+\alpha}}.$$

The generator of Z is the fractional Laplacian $\Delta^{\frac{\alpha}{2}}$ which is defined for all $f \in \mathcal{C}_b^{1+\gamma}(\mathbb{R}^d;\mathbb{R})$, with $\gamma > \alpha - 1$ (i.e. f belongs to $\mathcal{C}_b^1(\mathbb{R}^d;\mathbb{R})$ and ∇f is γ-Hölder continuous) and for all $x \in \mathbb{R}^d$, by

$$\Delta^{\frac{\alpha}{2}} f(x) := \int_{\mathbb{R}^d} (f(x+z) - f(x) - \nabla f(x) \cdot z) \, d\nu(z). \tag{8.10}$$

We are interested in the following stable-driven McKean-Vlasov SDE

$$\begin{cases} dX_t^{s,\xi} = b(t, X_t^{s,\xi}, [X_t^{s,\xi}]) \, dt + dZ_t, \quad t \in [s,T], \\ X_s^{s,\xi} = \xi, \quad [\xi] = \mu \in \mathcal{P}(\mathbb{R}^d), \end{cases} \tag{8.11}$$

where $[\xi]$ denotes the distribution of the random variable ξ and $s \in [0,T)$.

Let us define, for $k > -\alpha$ the function ρ^k on $(0,+\infty) \times \mathbb{R}^d$ by

$$\forall t > 0, \, x \in \mathbb{R}^d, \, \rho^k(t,x) := t^{-\frac{d}{\alpha}} (1 + t^{-\frac{1}{\alpha}} |x|)^{-d-\alpha-k}. \tag{8.12}$$

These functions are related to gradient estimates on the transition density q of the stable process Z (see Lemma 8.43). Some useful properties of these functions are given in Lemma 8.42.

8.2.1 Well-posedness of the nonlinear martingale problem and Picard iterations

The first point is to prove the existence and uniqueness, in the weak sense, of the solution to (8.16). Let us first recall the definition of the nonlinear martingale problem associated with (8.16).

Definition 8.1. Let us fix $(s,\mu) \in [0,T) \times \mathcal{P}(\mathbb{R}^d)$. We say that a probability measure $\mathbb{P}$ on the Skorokhod space $\mathcal{D}([s,T];\mathbb{R}^d)$, endowed with its canonical filtration $(\mathcal{F}_t)_t$, with time marginal distributions $(\mathbb{P}_t)_{t\in[s,T]} \in \mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ solves the nonlinear martingale problem associated to the SDE (8.11) with initial distribution μ at time s if the canonical process $(y_t)_{t\in[s,T]}$ satisfies the two following conditions.

1. We have $\mathbb{P}_s = \mu$.
2. For any $\phi \in \mathcal{C}_b^{1,2}([s,T] \times \mathbb{R}^d)$, the process defined for $t \in [s,T]$ by

$$\phi(t, y_t) - \phi(s, y_s) - \int_s^t \left(\partial_r + L_r^{\mathbb{P}}\right) \phi(r, y_r) \, dr$$

is a $(\mathcal{D}([s,T];\mathbb{R}^d),(\mathcal{F}_t)_t,\mathbb{P})$-martingale starting from 0 at time $t=s$ and where

$$L_r^{\mathbb{P}} f(t,x) := b(r,x,\mathbb{P}_r)\cdot\partial_x f(t,x) + \Delta^{\frac{\alpha}{2}} f(t,\cdot)(x). \tag{8.13}$$

We assume that the drift $b:[0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}^d$ satisfies the following properties.

Assumption (H2).

1. The drift b is measurable and globally bounded on $[0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)$.
2. For any $(t,\mu)\in[0,T]\times\mathcal{P}(\mathbb{R}^d)$, the map $b(t,\cdot,\mu)$ is η-Hölder continuous on $\mathbb{R}^d$ uniformly with respect to $(t,\mu)\in[0,T]\times\mathcal{P}(\mathbb{R}^d)$, with $\eta\in(0,1]$, i.e. there exists $C>0$ such that for all $t\in[0,T]$, $\mu\in\mathcal{P}(\mathbb{R}^d)$ and $x_1,x_2\in\mathbb{R}^d$
$$|b(t,x_1,\mu)-b(t,x_2,\mu)|\le C|x_1-x_2|^\eta.$$
3. For any $(t,x)\in[0,T]\times\mathbb{R}^d$, the map $b(t,x,\cdot)$ is Lipschitz continuous with respect to the total variation metric d_{TV} uniformly with respect to $(t,x)\in[0,T]\times\mathbb{R}^d$, i.e. there exists $C>0$ such that for all $t\in[0,T]$, $x\in\mathbb{R}^d$ and $\mu_1,\mu_2\in\mathcal{P}(\mathbb{R}^d)$
$$|b(t,x,\mu_1)-b(t,x,\mu_2)|\le C d_{TV}(\mu_1,\mu_2).$$

We can now state the weak well-posedness result.

Theorem 8.2 (Weak well-posedness)**.** *Under Assumption (H2), the martingale problem associated with the McKean-Vlasov SDE* (8.11) *is well-posed for all initial distribution $\mu\in\mathcal{P}(\mathbb{R}^d)$. In particular, the SDE* (8.11) *is well-posed in the weak sense. Moreover, for any $P\in\mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d))$ with $P_s=\mu\in\mathcal{P}(\mathbb{R}^d)$, we can define recursively $\overline{X}^{(m)}$ as the unique weak solution to*

$$\begin{cases} d\overline{X}_t^{(m)} = b(t,\overline{X}_t^m,[\overline{X}_t^{(m-1)}])\,dt + dZ_t, & t\in[s,T],\\ \overline{X}_s^{(m)} = \xi, \end{cases} \tag{8.14}$$

with $[\xi]=\mu$ and $([\overline{X}_t^{(0)}])_{t\in[s,T]}=P$. Then, denoting by P^* the unique solution to the martingale problem associated with (8.11)*, with initial condition $\mu\in\mathcal{P}(\mathbb{R}^d)$ at time s, we have*

$$\sup_{t\in[s,T]} d_{TV}(P_t^*,[\overline{X}_t^{(m)}]) \underset{m\to+\infty}{\longrightarrow} 0, \tag{8.15}$$

where d_{TV} is the total variation metric.

This result is proved in Section 8.3.

Remark 8.3. The well-posedness of the nonlinear martingale problem was already proved in [FKM21]. Therein, the fixed point theorem is only shown to hold in small time since it is enough to prove the well-posedness. However, we will need in the following the convergence of the sequence of Picard iterations on the whole interval $[s,T]$. That is why we do the proof a bit differently. Instead of using the implicit parametrix expansion, we write the complete parametrix expansion, which yields this global result.

8.2.2 Cauchy problem for the transition density associated with the McKean-Vlasov SDE

Now that the well-posedness has been established, we focus on the regularity of the transition density associated with (8.16). In particular, our aim is to study its regularity with respect to the initial distribution μ. To do this, we fix $s \in [0, T)$, $\beta \in (1, \alpha)$ and we consider the following stable-driven McKean-Vlasov SDE

$$\begin{cases} dX_t^{s,\xi} = b(t, X_t^{s,\xi}, [X_t^{s,\xi}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,\xi} = \xi, \quad [\xi] = \mu \in \mathcal{P}_\beta(\mathbb{R}^d). \end{cases} \tag{8.16}$$

The associated martingale problem is well-posed using Theorem 8.2 and there is weak existence and uniqueness for the SDE (8.16). Moreover the distribution of $X_t^{s,\xi}$ depends only on μ and not on the choice of the random variable ξ such that $[\xi] = \mu$ and we denote it by

$$[X_t^{s,\mu}] := [X_t^{s,\xi}]. \tag{8.17}$$

We introduce, for $x \in \mathbb{R}^d$, the following decoupled stochastic flow associated to SDE (8.16)

$$\begin{cases} dX_t^{s,x,\mu} = b(t, X_t^{s,x,\mu}, [X_t^{s,\mu}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,x,\mu} = x \in \mathbb{R}^d. \end{cases} \tag{8.18}$$

The distribution of $X_t^{s,x,\mu}$ admits a density denoted by $p(\mu, s, t, x, \cdot)$ by using the parametrix expansion given in Theorem 8.41. The distribution of $X_t^{s,\mu}$ has also a density denoted by $p(\mu, s, t, \cdot)$. Moreover, it satisfies the following relation stemming from the well-posedness of the martingale problem

$$p(\mu, s, t, y) = \int_{\mathbb{R}^d} p(\mu, s, t, x, y)\, d\mu(x). \tag{8.19}$$

We introduce the following assumption, which is stronger than Assumption (H2).

Assumption (H3).

1. The drift b is jointly continuous and globally bounded on $[0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$.
2. For any $(t, \mu) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$, the map $b(t, \cdot, \mu)$ is η-Hölder continuous on $\mathbb{R}^d$, for some fixed $\eta \in (0, 1]$, uniformly with respect to $(t, \mu) \in [0, T] \times \mathcal{P}(\mathbb{R}^d)$, i.e. there exists $C > 0$ such that for all $t \in [0, T]$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $x_1, x_2 \in \mathbb{R}^d$
$$|b(t, x_1, \mu) - b(t, x_2, \mu)| \leq C|x_1 - x_2|^\eta.$$
3. For any $(t, x) \in [0, T] \times \mathbb{R}^d$, the map $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto b(t, x, \mu)$ has a linear derivative such that $\frac{\delta}{\delta m} b(t, x, \mu)(\cdot)$ is η-Hölder continuous on $\mathbb{R}^d$ uniformly with respect to $(t, x, \mu) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$ and $\frac{\delta}{\delta m} b$ is bounded on $[0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$.
4. For any $(t, x, v) \in [0, T] \times (\mathbb{R}^d)^2$, the map $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto \frac{\delta}{\delta m} b(t, x, \mu)(v)$ has a linear derivative such that $\frac{\delta^2}{\delta m^2} b(t, x, \mu)(v, \cdot)$ is η-Hölder continuous uniformly with respect to $(t, x, \mu, v) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d$ and $\frac{\delta^2}{\delta m^2} b$ is bounded on $[0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \times (\mathbb{R}^d)^2$.

We need this stronger assumptions in order to study the regularity with respect to the initial distribution of the transition density p.

Let us now state in the next theorem the backward Kolmogorov PDE satisfied by the decoupled transition density p as well as some important gradient estimates on p.

Theorem 8.4. *Let us fix $(t,y) \in (0,T] \times \mathbb{R}^d$. Under Assumption (H3), the mapping $(\mu,s,x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0,t) \times \mathbb{R}^d \mapsto p(\mu,s,t,x,y)$ belongs to $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0,t) \times \mathbb{R}^d)$ (see Definition 8.36), and is solution to the following backward Kolmogorov PDE*

$$\begin{cases} \partial_s p(\mu,s,t,x,y) + \mathcal{L}_s p(\cdot,s,t,\cdot,y)(\mu,x) = 0, \quad \forall (\mu,s,x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0,t) \times \mathbb{R}^d, \\ p(\mu,s,t,x,\cdot) \underset{s \to t-}{\longrightarrow} \delta_x, \quad \textit{in the weak sense,} \end{cases} \tag{8.20}$$

where $\mathcal{L}_s$ is defined, for smooth enough function h on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$, by

$$\begin{aligned} \mathcal{L}_s h(\mu,x) := {} & b(s,x,\mu) \cdot \partial_x h(\mu,x) + \int_{\mathbb{R}^d} [h(\mu,x+z) - h(\mu,x) - z \cdot \partial_x h(\mu,x)] \, \frac{dz}{|z|^{d+\alpha}} \\ & + \int_{\mathbb{R}^d} b(s,v,\mu) \cdot \partial_v \frac{\delta}{\delta m} h(\mu,x)(v) \, d\mu(v) \\ & + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} h(\mu,x)(v+z) - \frac{\delta}{\delta m} h(\mu,x)(v) - z \cdot \partial_v \frac{\delta}{\delta m} h(\mu,x)(v) \right] \frac{dz}{|z|^{d+\alpha}} \, d\mu(v). \end{aligned} \tag{8.21}$$

Moreover, there exists a positive constant $C > 0$ such that for all $j \in \{0,1\}$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$ and $x,y,v \in \mathbb{R}^d$

$$|\partial_x^j p(\mu,s,t,x,y)| \le C(t-s)^{-\frac{j}{\alpha}} \rho^j(t-s,y-x), \tag{8.22}$$

$$|\Delta^{\frac{\alpha}{2}} p(\mu,s,t,\cdot,y)(x)| \le C(t-s)^{-1} \rho^0(t-s,y-x), \tag{8.23}$$

$$\left| \partial_v^j \frac{\delta}{\delta m} p(\mu,s,t,x,y)(v) \right| \le C(t-s)^{j\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} \rho^0(t-s,y-x), \tag{8.24}$$

and

$$\left| \Delta^{\frac{\alpha}{2}} \left[\frac{\delta}{\delta m} p(\mu,s,t,x,y) \right](v) \right| \le C(t-s)^{-\frac{1}{\alpha}} \rho^0(t-s,y-x), \tag{8.25}$$

where ρ^j was defined by (8.12).

The proof of this theorem is postponed to Section 8.4. More precisely, certain other estimates (see Theorem 8.18), in particular Hölder controls, are needed not only to prove the preceding theorem, but also to exhibit the regularizing properties of the semigroup associated with (8.16) which are presented in the next subsection.

Remark 8.5. Contrary to the Brownian case ($\alpha = 2$) studied in [CdRF22, CdRF21], we do not need to prove $\mathcal{C}^2$ regularity with respect to $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. This would impose stronger assumptions.

Remark 8.6. Let us introduce, for $(s,\mu) \in [0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$, the operator L_s^μ associated with the linear SDE related to the McKean-Vlasov SDE (8.16), i.e. where the measure argument is frozen equal to μ. It is defined, for smooth enough function f on $\mathbb{R}^d$ by

$$\forall x \in \mathbb{R}^d, \quad L_s^\mu f(x) := b(s,x,\mu) \cdot \partial_x f(x) + \Delta^{\frac{\alpha}{2}} f(x).$$

Then, we have for all smooth enough function h on $\mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$

$$\mathcal{L}_s h(\mu, x) = L_s^\mu h(\mu, \cdot)(x) + \int_{\mathbb{R}^d} L_s^\mu \left[\frac{\delta}{\delta m} h(\mu, x)\right](v)\, d\mu(v).$$

Let us now deal with the not decoupled density of the McKean-Vlasov SDE (8.16).

Theorem 8.7. *Let us fix $(t, y) \in (0, T] \times \mathbb{R}^d$. Under Assumption (H3), the mapping $(\mu, s) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \mapsto p(\mu, s, t, y)$ belongs to $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t))$ (see Definition 8.36), and is solution to the following backward Kolmogorov PDE*

$$\begin{cases} \partial_s p(\mu, s, t, y) + \mathscr{L}_s p(\cdot, s, t, y)(\mu) = 0, \quad \forall (\mu, s) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t), \\ p(\mu, s, t, \cdot) \underset{s \to t-}{\longrightarrow} \mu, \quad \textit{in the weak sense}, \end{cases} \tag{8.26}$$

where $\mathscr{L}_s$ is defined, for smooth enough functions h on $\mathcal{P}_\beta(\mathbb{R}^d)$, by

$$\begin{aligned} \mathscr{L}_s h(\mu) := &\int_{\mathbb{R}^d} b(s, v, \mu) \cdot \partial_v \frac{\delta}{\delta m} h(\mu)(v)\, d\mu(v) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} h(\mu)(v + z) - \frac{\delta}{\delta m} h(\mu)(v) - z \cdot \partial_v \frac{\delta}{\delta m} h(\mu)(v)\right] \frac{dz}{|z|^{d+\alpha}}\, d\mu(v). \end{aligned} \tag{8.27}$$

We do not give the proof of this result since it follows from the same reasoning as in the proof of (8.20) in Theorem 8.4 presented in Section 8.4 (see also [CdRF21, Theorem 3.3] for the proof in the Brownian case).

8.2.3 Backward Kolmogorov PDE on the space of measures

We can now focus on the study of the semigroup associated with (8.16), acting on functions defined on $\mathcal{P}_\beta(\mathbb{R}^d)$. Let us recall that $\beta \in (1, \alpha)$ is fixed. For a fixed function $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$, the action of the semigroup on ϕ is given by the map U defined by

$$U(t, \mu) := \phi([X_T^{t,\mu}]), \quad \forall (t, \mu) \in [0, T] \times \mathcal{P}_\beta(\mathbb{R}^d), \tag{8.28}$$

where $[X_T^{t,\mu}]$ is the flow of marginal distributions of (8.16), where the initial distribution is equal to μ at time t and was defined in (8.17). We aim at studying the regularizing properties of the semigroup, i.e. the gain of regularity between ϕ and U with respect to the measure variable. This will be crucial to prove the propagation of chaos. The regularization of ϕ by a smooth flow of probability measures is presented in Proposition 8.39.

We define the space of functions on which the semigroup acts.

Definition 8.8. Let us fix $\delta \in (0, 1]$. The space $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ is defined as the set of continuous functions $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ admitting a linear derivative such that there exists a positive constant C such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v_1, v_2 \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m} \phi(\mu)(v_1) - \frac{\delta}{\delta m} \phi(\mu)(v_2)\right| \le C|v_1 - v_2|^\delta.$$

We state in the next theorem the regularizing properties of the semigroup acting on $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ and we describe its dynamics through the backward Kolmogorov PDE that it satisfies.

Theorem 8.9 (Backward Kolmogorov PDE)**.** *Let us fix $\phi \in \mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$. Then, under Assumption (H3), the function U defined in (8.28) belongs to $\mathcal{C}^0([0,T]\times\mathcal{P}_\beta(\mathbb{R}^d))\cap\mathcal{C}^1([0,T)\times\mathcal{P}_\beta(\mathbb{R}^d))$ (see Definition 8.36) and satisfies the following properties.*

- *There exists a positive constant C such that for all $t\in[0,T)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $v\in\mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v)\right| \leq C(T-t)^{\frac{\delta-1}{\alpha}}. \tag{8.29}$$

- *For all $\gamma\in(0,1]\cap(0,(2\alpha-2)\wedge(\eta+\alpha-1))$, there exists a positive constant C such that for all $t\in[0,T)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $v_1,v_2\in\mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2)\right| \leq C(T-t)^{\frac{\delta-1-\gamma}{\alpha}} |v_1-v_2|^\gamma. \tag{8.30}$$

Moreover, U is solution to the following backward Kolmogorov PDE

$$\begin{cases}\partial_t U(t,\mu) + \mathscr{L}_t U(t,\cdot)(\mu) = 0, \quad \forall (t,\mu)\in[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d),\\ U(T,\mu)=\phi(\mu), \quad \forall \mu\in\mathcal{P}_\beta(\mathbb{R}^d),\end{cases} \tag{8.31}$$

where $\mathscr{L}_t$ was defined in (8.27). It is the unique solution to (8.31) among all functions in $\mathcal{C}^0([0,T]\times\mathcal{P}_\beta(\mathbb{R}^d))\cap\mathcal{C}^1([0,T)\times\mathcal{P}_\beta(\mathbb{R}^d))$ satisfying (8.29) and (8.30).

We prove this result in Section 8.5.

8.2.4 Quantitative weak propagation of chaos

We are now going to use the regularizing properties and the dynamics of the semigroup given in Theorem 8.9 to prove quantitative weak propagation of chaos for the mean-field interacting particle system associated with (8.16). Let us introduce $(Z^n)_n$ an i.i.d. sequence of α-stable processes having the same distribution as Z and $(X_0^n)_n$ an i.i.d. sequence of random variables with common distribution $\mu_0\in\mathcal{P}_\beta(\mathbb{R}^d)$, where $\beta\in(1,\alpha)$ is still fixed. For any $N\geq 1$, the system of N particles associated with (8.16) is defined as the unique solution to the following classical SDE

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \overline{\mu}_t^N)\,dt + dZ_t^i, \quad t\in[0,T], \quad i\in\{1,\dots,N\},\\ \overline{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}},\\ X_0^{i,N} = X_0^i. \end{cases} \tag{8.32}$$

This linear SDE on $(\mathbb{R}^d)^N$ is well-posed in the weak sense using [CSZ18, Corollary 1.4 (iii)] since its drift coefficient is Hölder continuous in space uniformly in time. The limiting McKean-Vlasov SDE is (8.16) starting at time $s=0$ from any random variable ξ with distribution $\mu_0\in\mathcal{P}_\beta(\mathbb{R}^d)$. We denote by $(\mu_t)_{t\in[0,T]}$ the flow of marginal distributions of its solution. Let us define the space of test functions that we use to quantify the weak propagation of chaos of the particle system (8.32) towards the McKean-Vlasov SDE (8.16).

Definition 8.10. For $\delta \in (0,1]$ and $L > 0$, we define the space $\mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ as the set of continuous functions $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ admitting two linear derivatives $\frac{\delta}{\delta m}\phi$ and $\frac{\delta^2}{\delta m^2}\phi$ such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $v_1, v_2, v_1', v_2' \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m}\phi(\mu)(v_1) - \frac{\delta}{\delta m}\phi(\mu)(v_2)\right| \leq L|v_1 - v_2|^\delta,$$

and

$$\left|\frac{\delta^2}{\delta m^2}\phi(\mu)(v_1, v_1') - \frac{\delta^2}{\delta m^2}\phi(\mu)(v_2, v_2')\right| \leq L(|v_1 - v_2|^\delta + |v_1' - v_2'|^\delta).$$

Remark 8.11. Note that $\mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ is a subspace of $\mathcal{C}^{(1,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ defined in Definition 8.8.

We now state our quantitative propagation of chaos result.

Theorem 8.12 (Quantitative weak propagation of chaos)**.**
Let us fix $\delta \in (0,1]$, $L > 0$ and $\gamma \in (0,1] \cap (0, (\delta+\alpha-1) \wedge (2\alpha-2) \wedge (\eta+\alpha-1))$. Then, under Assumption (H3), there exists a positive constant $C = C(d, T, \alpha, \beta, (H3), \gamma, \delta, L)$, non-decreasing with respect to T, such that for all $\phi \in \mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ and $N \geq 1$, it holds

$$\mathbb{E}|\phi(\overline{\mu}_T^N) - \phi(\mu_T)| \leq C\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) + \frac{C}{N^{1-\frac{1}{\beta}}}. \tag{8.33}$$

Moreover, there exists a positive constant $C = C(d, T, \alpha, \beta, (H3), \gamma, \delta, L, M_1(\mu_0))$, non-decreasing with respect to T, such that for all $\phi \in \mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ and $N \geq 1$, it holds

$$|\mathbb{E}(\phi(\overline{\mu}_T^N) - \phi(\mu_T))| \leq \frac{C}{N^\gamma}. \tag{8.34}$$

The proof is given in Section 8.6.

Remark 8.13.

1. The initial data term in (8.33) can be handled using Fournier and Guillin [FG15], in particular in the case where μ_0 has more moments than β. Indeed, one has if $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$ with $q \geq 1$

$$\mathbb{E}W_1(\overline{\mu}_0^N, \mu_0) \leq C \begin{cases} N^{-\frac{1}{2}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d = 1 \quad \text{and} \quad q \neq 2, \\ N^{-\frac{1}{2}}\ln(1+N) + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d = 2 \quad \text{and} \quad q \neq 2, \\ N^{-\frac{1}{d}} + N^{-\left(1-\frac{1}{q}\right)}, & \text{if} \quad d \geq 3 \quad \text{and} \quad q \neq \frac{d}{d-1}. \end{cases} \tag{8.35}$$

2. Let us justify why the estimate (8.34) is interesting. Firstly, this result quantifies the approximation of semigroup associated with the McKean-Vlasov SDE by its empirical projection defined in Definition 8.37 applied to the particle system. Secondly, it allows to quantify the approximation of the distribution of one particle by the distribution of the limiting McKean-Vlasov process with respect to W_1. Indeed, denoting by $\|\varphi\|_{\text{Lip}} := \sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{|x-y|}$ for $\varphi : \mathbb{R}^d \to \mathbb{R}$, the set

$$\mathscr{C} := \left\{\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R},\, \exists \varphi : \mathbb{R}^d \to \mathbb{R}, \text{ with } \|\varphi\|_{\text{Lip}} \leq 1, \text{ and } \phi(\mu) = \int_{\mathbb{R}^d} \varphi\, d\mu,\, \forall \mu \in \mathcal{P}_\beta(\mathbb{R}^d)\right\}$$

is contained in $\mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$. Thus, the estimate (8.34) and Kantorovich-Rubinstein's theorem

[Vil09, Remark 6.5] ensure that

$$\begin{aligned}
\sup_{t\in[0,T]} W_1([X_t^{1,N}],\mu_t) &= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left|\mathbb{E}\varphi(X_t^{1,N}) - \int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\
&= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left|\mathbb{E}\left(\frac{1}{N}\sum_{k=1}^N \varphi(X_t^{k,N})\right) - \int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\
&= \sup_{t\in[0,T]} \sup_{\varphi,\,\|\varphi\|_{\mathrm{Lip}}\le 1} \left|\mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\overline{\mu}_t^N - \mathbb{E}\int_{\mathbb{R}^d} \varphi\, d\mu_t\right| \\
&= \sup_{t\in[0,T]} \sup_{\phi\in\mathscr{C}} \left|\mathbb{E}\phi(\overline{\mu}_t^N) - \mathbb{E}\phi(\mu_t)\right| \\
&\le \frac{C_T}{N^\gamma},
\end{aligned} \tag{8.36}$$

where $\gamma \in (0,1]\cap(0,(2\alpha-2)\wedge(\eta+\alpha-1))$ and since the constant C in Theorem 8.12 is non-decreasing with respect to T.

Let us now compare our result with the existing literature.

Remark 8.14. 1. Let us formally take $\alpha = 2$, which corresponds to the Brownian case treated in [CdRF21]. Then, we can take $\gamma = 1$ and $\beta = 2$ in Theorem 8.12. The rates of convergence proved in our theorem are precisely those proved in [CdRF21, Theorem 3.6], i.e. $N^{-\frac{1}{2}}$ for (8.33) and N^{-1} for (8.34).

2. In dimension $d = 1$, we recover with (8.33) the same rate of convergence obtained in [FL21], for the strong propagation of chaos in L^1, since $\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0) \le CN^{\frac{1}{\beta}-1}$ by (8.35).

3. In [Cav22a], the example of a nonlinear Ornstein-Uhlenbeck process is treated using the same method. It corresponds to take

$$b(t,x,\mu) := x + \int_{\mathbb{R}^d} y\, d\mu(y).$$

However, there is preliminary step in the proof, which consists in removing the jumps larger than the number of particles N from all the noises. This is due to the unboundedness of b with respect to both space and measures variables in this case, which yields weaker estimates on the semigroup. It is proved in [Cav22a] that there exists a positive constant C such that for any $\phi \in \mathcal{C}_1^{(2,1)}(\mathcal{P}_\beta(\mathbb{R}^d))$

$$\mathbb{E}\left|\phi(\overline{\mu}_T^N) - \phi(\mu_T)\right| \le C\,\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0) + C\frac{\ln(N)^{\frac{1}{\alpha}}}{N^{1-\frac{1}{\alpha}}}, \tag{8.37}$$

which is better than (8.33) in spite of our stronger assumption of boundedness on b. Indeed, by removing the large jumps in a first step, we can take $\beta = \alpha$ in (8.33), up to the logarithmic factor present in (8.37). This factor precisely comes from the fact that $\int_{1\le|z|\le N} |z|^\alpha\, d\nu(z) \underset{N\to+\infty}{\sim} \ln(N)$, which is the price to pay to take $\beta = \alpha$ in (8.33). The estimate (8.34) is better in our framework since the rate of the corresponding estimate in [Cav22a] is $N^{1-\alpha}$ and we can take $\gamma > \alpha - 1$ in Theorem 8.12. This is natural since the drift is unbounded in [Cav22a].

8.2.5 Quantitative approximation of the distribution of one particle by the limiting McKean-Vlasov process at the level of densities

We keep the same notations as in the preceding subsection. We present here the quantitative propagation of chaos result at the level of densities for the particle system (8.32). Let us first introduce the following assumption which deals with the existence of a density for (8.32).

Assumption (H4). We assume that for any $t > 0$, the particles $(X_t^{1,N}, \dots, X_t^{N,N})$ defined by (8.32) with initial distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$ at time 0 have a density on $(\mathbb{R}^d)^N$ denoted by $\boldsymbol{p}^N(\mu_0, 0, t, \cdot)$.

Remark 8.15. When $d = 1$, notice that Assumption (H4) is implied by Assumption (H3) using [CHZ20a, Theorem 1.1]. Indeed, Assumption (H3) ensures that the map $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N \mapsto (b(t, x_i, \overline{\mu}_{\boldsymbol{x}}^N))_{i \in \{1,\dots,N\}}$ is Hölder continuous uniformly with respect to $t \in [0, T]$. If we want to avoid Assumption (H4) in dimension $d > 1$, we need to study if we can extend [CHZ20a, Theorem 1.1] to this framework.

Under Assumption (H4), the density of the first particle exists and is denoted by $p^{1,N}(\mu_0, s, t, \cdot)$. It is given, for all $y_1 \in \mathbb{R}^d$, by

$$p^{1,N}(\mu_0, 0, t, y_1) = \int_{(\mathbb{R}^d)^{N-1}} \boldsymbol{p}^N(\mu_0, 0, t, y_1, y_2, \dots, y_N)\, dy_2 \dots dy_N.$$

Note that by exchangeability, all the particles have the same distribution. Let us recall that $p(\mu_0, 0, t, \cdot)$, defined in (8.19), is the density on $\mathbb{R}^d$ of μ_t, which is the distribution of the solution to the McKean-Vlasov SDE (8.16) at time t with initial distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$ at time 0.

Theorem 8.16. *Let us fix $\gamma \in (0, 1] \cap (\alpha - 1, (2\alpha - 2) \wedge (\eta + \alpha - 1))$ and $\gamma' \in [\alpha - 1, 1]$. We define for $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0, T]$ and $y, z \in \mathbb{R}^d$*

$$q_0(\mu, 0, t, y) := \int_{\mathbb{R}^d} \rho^0(t, y - x)\, d\mu(x), \quad \text{and} \quad f(z) := |z|^{1+\gamma} \mathbf{1}_{B_1}(z) + |z| \mathbf{1}_{B_1^c}(z),$$

where ρ^0 was defined by (8.12). We also set $\zeta := -\left(1 - \frac{2+\gamma}{\alpha}\right) \in (0, 1)$ since $\gamma \in (0, 2\alpha - 2)$, and we denote by $\mathcal{B}$ the Beta function defined, for all $x, y > 0$, by

$$\mathcal{B}(x, y) := \int_0^1 (1 - t)^{-1+x} t^{-1+y}\, dt.$$

Then, under Assumptions (H3) and (H4), the following properties are satisfied.

- ***(Upper-bound for the density of one particle).*** *There exists a positive constant $C = C(d, T, \alpha, \beta, (H3), \gamma)$ such that for all $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0, T]$, $y \in \mathbb{R}^d$ and $N \geq 1$*

$$p^{1,N}(\mu_0, 0, t, y) \leq C q_0(\mu_0, 0, t, y) + \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} \frac{t^{k(1-\zeta)}}{k(1-\zeta)} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta), 1-\zeta) \right) \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0\left(\mu_0, 0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j), \tag{8.38}$$

where P_k denotes the set of all subsets of $\{1, \dots, k\}$ and by convention $q_0\left(\mu_0, 0, t, y - \sum_{i \in \emptyset} z_i\right) := q_0\left(\mu_0, 0, t, y\right)$.

- ***(Pointwise estimate for the approximation of the density of one particle).*** *There exists a positive constant $C = C(d, T, \alpha, \beta, (H3), \gamma, \gamma')$ such that for all $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0, T]$, $y \in \mathbb{R}^d$ and $N \geq 1$*

$$\Big|p^{1,N}(\mu_0, 0, t, y) - p(\mu_0, 0, t, y)\Big| \leq \frac{C}{N^{\gamma'}} t^{1-\frac{1+\gamma'}{\alpha}}(1 + M_{\gamma'}(\mu_0)) \int_{\mathbb{R}^d} (1 + |x|^{\gamma'})\rho^0(t, y - x)\, d\mu_0(x)$$
$$+ \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} t^{k(1-\zeta)} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta), 1-\zeta)\right) \mathcal{B}(1 + k(1-\zeta), 1-\zeta) \tag{8.39}$$
$$\sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0\left(\mu_0, 0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j).$$

- ***(Estimate for the approximation of the distribution of one particle in total variation).*** *There exists a positive constant $C = C(d, T, \alpha, \beta, (H3), \gamma)$ such that for all initial distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t \in (0, T]$ and $N \geq 1$*

$$d_{TV}([X_t^{1,N}], \mu_t) \leq \frac{C}{N^\gamma} t^{1-\frac{1+\gamma}{\alpha}}(1 + M_\gamma(\mu_0)), \tag{8.40}$$

and

$$\sup_{t\in[0,T]} d_{TV}([X_t^{1,N}], \mu_t) \leq \frac{C}{N^{\alpha-1}}(1 + M_{\alpha-1}(\mu_0)). \tag{8.41}$$

The proof is given in Section 8.7.

Remark 8.17. Let us compare this result with the estimate (8.36) quantified with the Wasserstein metric W_1. We find the same rate of convergence $N^{-\gamma}$ with $\gamma \in (0, 1] \cap (0, (2\alpha - 2) \wedge (\eta + \alpha - 1))$. However in (8.40), there is a time-integrable singularity, similarly to [CdRF21, Theorem 3.5], which is not present in (8.36). This time singularity can be removed with the slower rate of convergence $N^{1-\alpha}$. We recover the rate of convergence N^{-1} shown in [CdRF21, Theorem 3.5] in the Brownian case by taking formally $\alpha = 2$ in the preceding result.

8.3 Well-posedness of the nonlinear martingale problem and Picard iterations

This section is dedicated to the proof of Theorem 8.2. The proof is based on the Banach fixed point theorem on a suitable complete metric space.

Introduction of the complete space and parametrix expansion. Let us consider the space $\mathcal{C}^0([s, T]; \mathcal{P}(\mathbb{R}^d))$ which is complete under the uniform metric $d_{s,T}$ associated to the total variation metric d_{TV} defined, for $P, Q \in \mathcal{C}^0([s, T]; \mathcal{P}(\mathbb{R}^d))$ by

$$d_{s,T}(P, Q) := \sup_{r\in[s,T]} d_{TV}(P_r, Q_r).$$

We introduce the space

$$\mathcal{A}_{s,T,\mu} := \left\{P \in \mathcal{C}^0([s, T]; \mathcal{P}(\mathbb{R}^d)),\ P_s = \mu\right\}.$$

Note that it is a closed subspace of $(\mathcal{C}^0([s,T];\mathcal{P}(\mathbb{R}^d)), d_{s,T})$ and thus $(\mathcal{A}_{s,T,\mu}, d_{s,T})$ is complete. For any $P \in \mathcal{A}_{s,T,\mu}$, we consider the following linear time-inhomogeneous SDE

$$\begin{cases} d\overline{X}_t^{s,\xi,P} = b(t, \overline{X}_t^{s,\xi,P}, P_t)\,dt + dZ_t, \quad t \in [s,T], \\ \overline{X}_s^{s,\xi,P} = \xi. \end{cases}$$

Notice that this SDE is well-posed in the weak sense since it is the case for the related linear martingale problem by [MP14]. Its flow of marginal distributions $([X_t^{s,\xi,P}])_{t\in[s,T]}$ belongs to $\mathcal{A}_{s,T,\mu}$. We can thus define a map $\mathcal{I} : \mathcal{A}_{s,T,\mu} \to \mathcal{A}_{s,T,\mu}$ such that for any $P \in \mathcal{A}_{s,T,\mu}$, $\mathcal{I}(P)_t = [X_t^{s,\xi,P}]$. We remark that a probability measure $\mathbb{P}$ on the Skorokhod space $\mathcal{D}([s,T];\mathbb{R}^d)$ solves the martingale problem related to the McKean-Vlasov SDE (8.11) if and only if its flow of marginal distributions $(\mathbb{P}_t)_{t\in[s,T]}$ is a fixed point of $\mathcal{I}$. Our goal is thus to prove that for some $m \geq 1$, the m-th iterate $\mathcal{I}^{(m)}$ is a contraction on $(\mathcal{A}_{s,T,\mu}, d_{s,T})$. We fix $P^1, P^2 \in \mathcal{A}_{s,T,\mu}$ and we define recursively for all $m \geq 1$, $\overline{X}^{1,(m)}$ and $\overline{X}^{2,(m)}$ as the unique weak solutions to

$$\begin{cases} d\overline{X}_t^{i,(m)} = b(t, \overline{X}_t^{i,m}, [\overline{X}_t^{i,(m-1)}])\,dt + dZ_t, \quad t \in [s,T], \quad i \in \{1,2\}, \\ \overline{X}_s^{i,(m)} = \xi, \end{cases} \tag{8.42}$$

with $([\overline{X}_t^{i,(0)}])_{t\in[s,T]} = P^i$. We also introduce the associated decoupling fields. Namely they are the weak solutions to

$$\begin{cases} d\overline{X}_t^{x,i,(m)} = b(t, \overline{X}_t^{x,i,m}, [\overline{X}_t^{i,(m-1)}])\,dt + dZ_t, \quad t \in [s,T], \quad i \in \{1,2\}, \\ \overline{X}_s^{x,i,(m)} = x. \end{cases} \tag{8.43}$$

Thanks to Theorem 8.41, the distribution of $\overline{X}_t^{x,i,(m)}$ has a density with respect to the Lebesgue measure denoted by $p_{i,m}(\mu, s, t, x, \cdot)$. Remark that the notation makes sense since, by weak well-posedness, the distribution of $\overline{X}_t^{x,i,(m)}$ depends on the initial condition ξ only through its distribution μ. Moreover, by weak well-posedness of SDE (8.42), $\overline{X}_t^{i,(m)}$ has a density $p_{i,m}(\mu, s, t, \cdot)$ such that for all $y \in \mathbb{R}^d$

$$p_{i,m}(\mu, s, t, y) = \int_{\mathbb{R}^d} p_{i,m}(\mu, s, t, x, y)\,d\mu(x).$$

Let us give the implicit parametrix representation of $p_{i,m}(\mu, s, t, x, y)$, which is given in Appendix 8.10. We define for all $0 \leq s \leq r < t \leq T$ and $x, y \in \mathbb{R}^d$

$$\begin{aligned} \widehat{p}(r,t,x,y) &:= q(t-r, y-x), \\ \mathcal{H}_{i,m}(\mu, r, t, x, y) &:= b(r, x, [\overline{X}_r^{i,(m-1)}]) \cdot \partial_x \widehat{p}(r,t,x,y), \end{aligned} \tag{8.44}$$

where $q(t,\cdot)$ is the density of Z_t. The space-time convolution between to functions f and g is given by

$$f \otimes g(\mu, r, t, x, y) := \int_r^t \int_{\mathbb{R}^d} f(\mu, r, r', x, z) g(\mu, r', t, z, y)\,dz\,dr', \tag{8.45}$$

when it is well-defined. The convolution iterates of $\mathcal{H}_{i,m}$ are defined recursively by $\mathcal{H}_{i,m}^{k+1} = \mathcal{H}_{i,m} \otimes \mathcal{H}_{i,m}^k$. By convention $\mathcal{H}_{i,m}^0 = \mathrm{Id}$. By Assumption (H2), we can apply Theorem 8.41 which ensures that

$$p_{i,m}(\mu, s, t, x, y) = \widehat{p}(s,t,x,y) + p_{i,m} \otimes \mathcal{H}_{i,m}(\mu, s, t, x, y). \tag{8.46}$$

Using Theorem 8.41 and Proposition 8.44, we deduce that there exists a positive constant K such that for all $i \in \{1,2\}$, $m \geq 1$, $k \geq 1$, $0 \leq s \leq r < t \leq T$ and $x, y \in \mathbb{R}^d$

$$\begin{cases} |p_{i,m}(\mu,s,t,x,y)| \leq K\rho^0(t-s,y-x), \\ \left|\mathcal{H}^k_{i,m}(\mu,r,t,x,y)\right| \leq K^k(t-r)^{-\frac{1}{\alpha}+(k-1)(1-\frac{1}{\alpha})} \displaystyle\prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x), \end{cases} \tag{8.47}$$

where the functions ρ^j were defined in (8.12) and $\mathcal{B}$ is the Beta function defined, for all $x, y > 0$ by

$$\mathcal{B}(x,y) := \int_0^1 (1-t)^{-1+x}t^{-1+y}\,dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

Γ being the Gamma function. For any $0 \leq s \leq r < t \leq T$, $x, y \in \mathbb{R}^d$, we define

$$\begin{aligned} \Delta p_m(\mu,s,t,x,y) &:= p_{1,m}(\mu,s,t,x,y) - p_{2,m}(\mu,s,t,x,y), \\ \Delta\mathcal{H}_m(\mu,r,t,x,y) &:= \mathcal{H}_{1,m}(\mu,r,t,x,y) - \mathcal{H}_{2,m}(\mu,r,t,x,y). \end{aligned}$$

It follows from the definition of $\mathcal{H}_{i,m}$, the Lipschitz continuity of $b(s,x,\cdot)$ with respect to the total variation metric (and its expression as the L^1 norm of the difference between the densities) and (8.235), that for some constant $C > 0$, one has for all $m \geq 1$

$$\begin{aligned} |\Delta\mathcal{H}_{m+1}(\mu,r,t,x,y)| &\leq Cd_{TV}([\overline{X}^{1,(m)}_r],[\overline{X}^{2,(m)}_r])(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x) \\ &\leq C\int_{\mathbb{R}^{2d}} |\Delta p_m(\mu,s,r,x',y')|\,d\mu(x')\,dy'\,(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x). \end{aligned} \tag{8.48}$$

We similarly obtain for $m = 0$ that

$$|\Delta\mathcal{H}_1(\mu,r,t,x,y)| \leq Cd_{s,r}(P^1,P^2)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x). \tag{8.49}$$

Let us prove that for all $m \geq 1$, the following representation formula holds true

$$\Delta p_m(\mu,s,t,x,y) = \sum_{k=0}^{\infty} p_{2,m} \otimes \Delta\mathcal{H}_m \otimes \mathcal{H}^k_{1,m}(\mu,s,t,x,y), \tag{8.50}$$

the series being absolutely convergent. Starting from the implicit parametrix representation formula (8.46), we have

$$\Delta p_m(\mu,s,t,x,y) = p_{2,m} \otimes \Delta\mathcal{H}_m(\mu,s,t,x,y) = \Delta p_m \otimes \mathcal{H}_{1,m}(\mu,s,t,x,y).$$

Iterating this procedure, we easily prove by induction that for all $N \geq 1$

$$\Delta p_m(\mu,s,t,x,y) = \sum_{k=0}^{N} p_{2,m} \otimes \Delta\mathcal{H}_m \otimes \mathcal{H}^k_{1,m}(\mu,s,t,x,y) + \Delta p_m \otimes \mathcal{H}^{N+1}_{1,m}(\mu,s,t,x,y). \tag{8.51}$$

We want to take the limit $N \to +\infty$ in (8.51). By using (8.47) and the convolution inequality (8.234), we obtain that for some positive constant C

$$
\begin{aligned}
&\left|\Delta p_m \otimes \mathcal{H}_{1,m}^{N+1}(\mu,s,t,x,y)\right| \\
&\leq \int_s^t \int_{\mathbb{R}^d} C\rho^0(r-s,z-x) C^{N+1}(t-r)^{-\frac{1}{\alpha}+N\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{N} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-z)\,dz\,dr \\
&\leq C^{N+2}(t-s)^{(N+1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{N} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^0(t-s,y-x).
\end{aligned}
$$

The upper-bound converges to 0 as N tends to infinity thanks to the asymptotic behavior of the Beta function. Following the same lines, we prove using again (8.47) that the series appearing in (8.51) is absolutely convergent. Letting N tend to infinity in (8.51) yields the representation formula (8.50).

We are now going to prove by induction that there exists a positive constant C such that for all $m \geq 1$, $t \in (s,T]$, $x,y \in \mathbb{R}^d$

$$
|\Delta p_m(\mu,s,t,x,y)| \leq C^m (t-s)^{m\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{m-1} \mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right) d_{s,t}(P^1,P^2)\rho^0(t-s,y-x). \tag{8.52}
$$

Base case $m = 1$. It follows from (8.47), (8.49) and the convolution inequality (8.234) that

$$
\begin{aligned}
|p_{2,1} \otimes \Delta\mathcal{H}_1(\mu,s,t,x,y)| &\leq C\int_s^t \int_{\mathbb{R}^d} \rho^0(r-s,z-x)(t-r)^{-\frac{1}{\alpha}} d_{s,r}(P^1,P^2)\rho^1(t-r,y-z)\,dz\,dr \\
&\leq C(t-s)^{1-\frac{1}{\alpha}} d_{s,t}(P^1,P^2)\rho^0(t-s,y-x),
\end{aligned} \tag{8.53}
$$

since $d_{s,r}(P^1,P^2) \leq d_{s,t}(P^1,P^2)$ for $r \in [s,t]$. Following the same lines and using the bound (8.47) on $\mathcal{H}_{1,1}^k$, we show that for some constant $C > 0$, one has for all $k \geq 1$

$$
\begin{aligned}
&\left|p_{2,1} \otimes \Delta\mathcal{H}_1 \otimes \mathcal{H}_{1,1}^k(\mu,s,t,x,y)\right| \\
&\leq C(t-s)^{1-\frac{1}{\alpha}} d_{s,t}(P^1,P^2)\rho^0(t-s,y-x) C^k (t-s)^{k\left(1-\frac{1}{\alpha}\right)} \prod_{l=1}^{k-1} \mathcal{B}\left(l\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

By summing over $k \geq 1$, we deduce using the asymptotic behavior of the Beta function that for some constant $C > 0$, we have

$$
\begin{aligned}
&\sum_{k=1}^{\infty} \left|p_{2,1} \otimes \Delta\mathcal{H}_1 \otimes \mathcal{H}_{1,1}^k(\mu,s,t,x,y)\right| \\
&\leq C(t-s)^{1-\frac{1}{\alpha}} d_{s,t}(P^1,P^2)\rho^0(t-s,y-x).
\end{aligned} \tag{8.54}
$$

Plugging (8.53) and (8.54) into the representation formula (8.50) for $m = 1$ concludes the proof of the base case.

Induction step. We assume that (8.52) holds true at step m for a certain constant C that will be chosen at the end of the induction step to ensure that (8.52) is verified at step $m+1$. We denote by K

any constant independent of m and C appearing in the induction step. By using (8.48), one has

$$|\Delta\mathcal{H}_{m+1}(\mu,r,t,x,y)| \leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)$$
$$C^m(r-s)^{m\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{m-1}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,r}(P^1,P^2).$$

This inequality combined with (8.47) yields

$$\begin{aligned}
&|p_{2,m+1}\otimes\Delta\mathcal{H}_{m+1}(\mu,s,t,x,y)| \\
&\leq K\int_s^t\int_{\mathbb{R}^d}\rho^0(r-s,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)C^m(r-s)^{m\left(1-\frac{1}{\alpha}\right)} \\
&\qquad\prod_{j=1}^{m-1}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,r}(P^1,P^2)\,dz\,dr \\
&\leq KC^m(t-s)^{(m+1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{m}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,t}(P^1,P^2)\rho^0(t-s,y-x).
\end{aligned}\tag{8.55}$$

Following the same lines and using the bound (8.47) on $\mathcal{H}^k_{1,m+1}$, we show that for some constant $K>0$, one has for all $k\geq 1$

$$\begin{aligned}
&\left|p_{2,m+1}\otimes\Delta\mathcal{H}_{m+1}\otimes\mathcal{H}^k_{1,m+1}(\mu,s,t,x,y)\right| \\
&\leq KC^m(t-s)^{(m+1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{m}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,t}(P^1,P^2)\rho^0(t-s,y-x) \\
&\qquad K^k(t-s)^{k\left(1-\frac{1}{\alpha}\right)}\prod_{l=1}^{k-1}\mathcal{B}\left(l\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}$$

By summing over $k\geq 1$, we deduce, thanks to the asymptotic behavior of the Beta function, that for some constant $K>0$, we have

$$\begin{aligned}
&\sum_{k=1}^{\infty}\left|p_{2,m+1}\otimes\Delta\mathcal{H}_{m+1}\otimes\mathcal{H}^k_{1,m+1}(\mu,s,t,x,y)\right| \\
&\leq KC^m(t-s)^{(m+1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{m}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,t}(P^1,P^2)\rho^0(t-s,y-x).
\end{aligned}\tag{8.56}$$

Finally, plugging (8.55) and (8.56) into the representation formula (8.50), we obtain

$$|\Delta p_{m+1}(\mu,s,t,x,y)| \leq KC^m(t-s)^{(m+1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{m}\mathcal{B}\left(1+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)d_{s,t}(P^1,P^2)\rho^0(t-s,y-x).$$

This ends the proof of the induction step provided that we choose $C\geq K$ in (8.52), which is possible since K does not depend on m.

Conclusion of the proof of Theorem 8.2. Using the asymptotic behavior of the Beta function, we get that for m large enough, we have for all $t \in [s,T]$, $x, y \in \mathbb{R}^d$

$$|\Delta p_m(\mu, s, t, x, y)| \leq \varepsilon d_{s,t}(P^1, P^2)\rho^0(t-s, y-x),$$

where $\varepsilon > 0$ is such that

$$\varepsilon \int_{\mathbb{R}^d} \rho^0(t-s, y)\, dy = \varepsilon \int_{\mathbb{R}^d} \rho^0(1, y)\, dy = \frac{1}{2}.$$

Finally, we have

$$\begin{aligned} d_{s,T}(\mathcal{I}^{(m)}(P^1), \mathcal{I}^{(m)}(P^2)) &= \sup_{t \in (s,T]} \sup_{h,\, \|h\|_\infty \leq 1} \left| \int_{\mathbb{R}^d} h(y)(p_{1,m}(\mu, s, t, y) - p_{2,m}(\mu, s, t, y))\, dy \right| \\ &\leq \sup_{t \in (s,T]} \int_{\mathbb{R}^{2d}} |\Delta p_m(\mu, s, t, x, y)|\, d\mu(x)\, dy \\ &\leq \frac{1}{2} d_{s,T}(P^1, P^2). \end{aligned}$$

The Banach fixed point theorem ensures that $\mathcal{I}$ has a unique fixed point in $\mathcal{A}_{s,T,\mu}$. Thus, the martingale problem associated to the McKean-Vlasov SDE (8.11) is well-posed. Moreover, we know that for any initial data $P \in \mathcal{A}_{s,T,\mu}$, the sequence $(\mathcal{I}^{(m)}(P))_{m \geq 1}$ converges towards the solution to the martingale problem with respect to the metric $d_{s,T}$. This proves (8.15).

8.4 Properties of the transition density

This section is devoted to prove Theorem 8.4. It will be a direct consequence of the following result.

Theorem 8.18 (Regularity estimates on the decoupled transition density)**.** *Let us fix $0 \leq s < t \leq T$ and $y \in \mathbb{R}^d$. Under Assumption (H3), the mapping $(\mu, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto p(\mu, s, t, x, y)$ belongs to $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d)$ (see Definition (8.36)). Moreover, it satisfies the following properties.*

- *There exists $C > 0$ such that for all $j \in \{0, 1\}$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\partial_x^j p(\mu, s, t, x, y)| \leq C(t-s)^{-\frac{j}{\alpha}} \rho^j(t-s, y-x). \tag{8.57}$$

- *For all $j \in \{0, 1\}$ and $\gamma \in (0, 1]$ with $\gamma \in (0, (2\alpha - 2) \wedge (\eta + \alpha - 1))$ if $j = 1$, there exists $C > 0$ such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$|\partial_x^j p(\mu, s, t, x_1, y) - \partial_x^j p(\mu, s, t, x_2, y)| \leq C(t-s)^{-\frac{j+\gamma}{\alpha}} |x_1 - x_2|^\gamma \left[\rho^j(t-s, y-x_1) + \rho^j(t-s, y-x_2)\right]. \tag{8.58}$$

- *There exists $C > 0$ such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v \in \mathbb{R}^d$*

$$\left| \frac{\delta}{\delta m} p(\mu, s, t, x, y)(v) \right| \leq C(t-s)^{1-\frac{1}{\alpha}} \rho^0(t-s, y-x). \tag{8.59}$$

- *There exists $C > 0$ such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v \in \mathbb{R}^d$*

$$\left| \partial_v \frac{\delta}{\delta m} p(\mu, s, t, x, y)(v) \right| \leq C(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} \rho^0(t-s, y-x). \tag{8.60}$$

- *For all* $\gamma \in (0,1] \cap (0,(2\alpha-2)\wedge(\eta+\alpha-1))$*, there exists* $C>0$ *such that for all* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $0 \le s < t \le T$*,* $x, y, v_1, v_2 \in \mathbb{R}^d$

$$\left|\partial_v \frac{\delta}{\delta m} p(\mu,s,t,x,y)(v_1) - \partial_v \frac{\delta}{\delta m} p(\mu,s,t,x,y)(v_2)\right| \le C(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}} |v_1-v_2|^\gamma \rho^0(t-s,y-x). \tag{8.61}$$

- *For all* $\gamma \in (0,1]$*, there exists* $C>0$ *such that for all ,* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $0 \le s < t \le T$*,* $x, y, v_1, v_2 \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m} p(\mu,s,t,x,y)(v_1) - \frac{\delta}{\delta m} p(\mu,s,t,x,y)(v_2)\right| \le C(t-s)^{1-\frac{1+\gamma}{\alpha}} |v_1-v_2|^\gamma \rho^0(t-s,y-x). \tag{8.62}$$

- *For all* $\gamma \in (0,1]$*, there exists* $C>0$ *such that for all* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $0 \le s < t \le T$*,* $x_1, x_2, y, v \in \mathbb{R}^d$

$$\begin{aligned}\left|\frac{\delta}{\delta m} p(\mu,s,t,x_1,y)(v) - \frac{\delta}{\delta m} p(\mu,s,t,x_2,y)(v)\right| &\le C(t-s)^{1-\frac{1+\gamma}{\alpha}} |x_1-x_2|^\gamma \\ &\quad \left[\rho^0(t-s,y-x_1) + \rho^0(t-s,y-x_2)\right].\end{aligned} \tag{8.63}$$

- *For all* $\gamma \in (0,1] \cap (0,\eta+\alpha-1)$*, there exists* $C>0$ *such that for all* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $0 \le s < t \le T$*,* $x_1, x_2, y, v \in \mathbb{R}^d$

$$\begin{aligned}\left|\partial_v \frac{\delta}{\delta m} p(\mu,s,t,x_1,y)(v) - \partial_v \frac{\delta}{\delta m} p(\mu,s,t,x_2,y)(v)\right| &\le C(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}} |x_1-x_2|^\gamma \\ &\quad \left[\rho^0(t-s,y-x_1) + \rho^0(t-s,y-x_2)\right].\end{aligned} \tag{8.64}$$

- *For all* $j \in \{0,1\}$ *and* $\gamma \in (0,1]$ *with* $\gamma \in (0,\alpha-1+\eta)$ *if* $j=1$*, there exists* $C>0$ *such that for all* $0 \le s < t \le T$*,* $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $x, y \in \mathbb{R}^d$

$$\left|\partial_x^j p(\mu_1,s,t,x,y) - \partial_x^j p(\mu_2,s,t,x,y)\right| \le C(t-s)^{1-\frac{1+\gamma+j}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \rho^j(t-s,y-x). \tag{8.65}$$

- *For all* $\gamma \in (0,1]$*, there exists* $C>0$ *such that for all* $0 \le s < t \le T$*,* $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $x, y, v \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m} p(\mu_1,s,t,x,y)(v) - \frac{\delta}{\delta m} p(\mu_2,s,t,x,y)(v)\right| \le C(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \rho^0(t-s,y-x). \tag{8.66}$$

- *For all* $\gamma \in (0,1]$*, there exists* $C>0$ *such that for all* $0 \le s < t \le T$*,* $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $x, y, v \in \mathbb{R}^d$

$$\left|\partial_v \frac{\delta}{\delta m} p(\mu_1,s,t,x,y)(v) - \partial_v \frac{\delta}{\delta m} p(\mu_2,s,t,x,y)(v)\right| \le C(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \rho^0(t-s,y-x). \tag{8.67}$$

- *For all* $j \in \{0,1\}$*,* $\gamma \in \left(0, 1-\frac{j}{\alpha}\right)$*, there exists a constant* $C>0$ *such that for all* $t \in (0,T]$*,* $s_1, s_2 \in [0,t)$*,* $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$*,* $x, y \in \mathbb{R}^d$

$$|\partial_x^j p(\mu, s_1, t, x, y) - \partial_x^j p(\mu, s_2, t, x, y)| \\ \leq C\left[\frac{|s_1 - s_2|^\gamma}{(t-s_1)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1, y-x) + \frac{|s_1 - s_2|^\gamma}{(t-s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_2, y-x)\right]. \quad (8.68)$$

- *For all $\gamma \in (0,1)$, there exists a constant $C > 0$ such that for all $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\left|\frac{\delta}{\delta m}p(\mu, s_1, t, x, y)(v) - \frac{\delta}{\delta m}p(\mu, s_2, t, x, y)(v)\right| \\ \leq C\left[\frac{|s_1 - s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1\vee s_2, y-x) + \frac{|s_1 - s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1\wedge s_2, y-x)\right]. \quad (8.69)$$

- *For all $\gamma \in \left(0, 1+\frac{\eta-1}{\alpha}\right)$, there exists a constant $C > 0$ such that for all $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\left|\partial_v\frac{\delta}{\delta m}p(\mu, s_1, t, x, y)(v) - \partial_v\frac{\delta}{\delta m}p(\mu, s_2, t, x, y)(v)\right| \quad (8.70) \\ \leq C\left[\frac{|s_1 - s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\vee s_2, y-x) + \frac{|s_1 - s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\wedge s_2, y-x)\right].$$

- *For all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, v \in \mathbb{R}^d$, we have*

$$\int_{\mathbb{R}^d}\frac{\delta}{\delta m}p(\mu, s, t, x, y)(v)\,dy = \int_{\mathbb{R}^d}\partial_v\frac{\delta}{\delta m}p(\mu, s, t, x, y)(v)\,dy = 0. \quad (8.71)$$

Before proving this result, let us introduce the parametrix method that is at the core of the proof (see Appendix 8.10 for more details). We denote by $q(t,\cdot)$ the density of Z_t. We define for all $0 \leq s \leq r < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$

$$\widehat{p}(\mu, s, r, t, x, y) = \widehat{p}(r, t, x, y) := q(t-r, y-x), \quad (8.72) \\ \mathcal{H}(\mu, s, r, t, x, y) := b(r, x, [X_r^{s,\mu}]) \cdot \partial_x\widehat{p}(s, r, t, x, y).$$

Note that the proxy $\widehat{p}(s, r, t, x, \cdot)$ does not depend on μ and s and is the density at time $t > r$ of the solution to

$$\begin{cases} d\widehat{X}_t^{r,x} = dZ_t, \\ \widehat{X}_r^{r,x} = x \in \mathbb{R}^d, \end{cases} \quad (8.73)$$

and $\mathcal{H}$ is the associated parametrix kernel. We also define the space-time convolution operator between to functions f and g by

$$f \otimes g(\mu, s, r, t, x, y) := \int_r^t\int_{\mathbb{R}^d} f(\mu, s, r, r', x, z)g(\mu, s, r', t, z, y)\,dz\,dr', \quad (8.74)$$

when it is well-defined. The space-time convolution iterates $\mathcal{H}^k$ of $\mathcal{H}$ are defined recursively by $\mathcal{H}^1 := \mathcal{H}$ and $\mathcal{H}^{k+1} := \mathcal{H} \otimes \mathcal{H}^k$. By convention $f \otimes \mathcal{H}^0$ is equal to f. In order to simplify a bit the notations, we will write $f \otimes g(\mu, s, t, x, y) := f \otimes g(\mu, s, s, t, x, y)$, $\mathcal{H}(\mu, s, t, x, y) := \mathcal{H}(\mu, s, s, t, x, y)$, and the same for other maps. Finally, we denote by Φ the solution to the following Volterra integral equation

$$\Phi(\mu, s, r, t, x, y) = \mathcal{H}(\mu, s, r, t, x, y) + \mathcal{H} \otimes \Phi(\mu, s, r, t, x, y),$$

which is given by the uniform convergent series

$$\Phi(\mu, s, r, t, x, y) = \sum_{k=1}^{\infty} \mathcal{H}^k(\mu, s, r, t, x, y). \tag{8.75}$$

Then, we have using Theorem 8.41, that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$, and $x, y \in \mathbb{R}^d$

$$p(\mu, s, t, x, y) = \widehat{p}(s, t, x, y) + \sum_{k=0}^{\infty} \widehat{p} \otimes \mathcal{H}^k(\mu, s, t, x, y). \tag{8.76}$$

Let us now prove Theorem 8.18.

Step 1: Properties of the Picard approximation of the transition density associated to (8.18). In order to study the regularity with respect to μ of $p(\mu, s, t, x, y)$, we consider an approximation sequence based on Picard iteration. We fix a measure $\nu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $s \in [0, T)$ and we start by considering the following stable-driven McKean-Vlasov SDE

$$\begin{cases} dX_t^{s,\xi,(1)} = b(t, X_t^{s,\xi,(1)}, \nu)\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,\xi,(1)} = \xi, \quad [\xi] = \mu \in \mathcal{P}_\beta(\mathbb{R}^d). \end{cases} \tag{8.77}$$

The associated martingale problem is well-posed by [MP14] and there is weak existence and uniqueness for SDE (8.77). As previously, the distribution of $X_t^{s,\xi,(1)}$ is denoted by $[X_t^{s,\mu,(1)}]$. We also introduce, for $x \in \mathbb{R}^d$, the following decoupled stochastic flow associated to SDE (8.77)

$$\begin{cases} dX_t^{s,x,\mu,(1)} = b(t, X_t^{s,x,\mu,(1)}, [X_t^{s,\mu,(1)}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,x,\mu,(1)} = x \in \mathbb{R}^d. \end{cases} \tag{8.78}$$

Then, for all $m \ge 1$, we define recursively

$$\begin{cases} dX_t^{s,\xi,(m+1)} = b(t, X_t^{s,\xi,(m+1)}, [X_t^{s,\mu,(m)}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,\xi,(m+1)} = \xi, \quad [\xi] = \mu \in \mathcal{P}_\beta(\mathbb{R}^d). \end{cases} \tag{8.79}$$

and

$$\begin{cases} dX_t^{s,x,\mu,(m+1)} = b(t, X_t^{s,x,\mu,(m+1)}, [X_t^{s,\mu,(m)}])\, dt + dZ_t, \quad t \in [s, T], \\ X_s^{s,x,\mu,(m+1)} = x \in \mathbb{R}^d. \end{cases} \tag{8.80}$$

Note that these are not McKean-Vlasov SDEs but linear SDEs. The densities of $[X_t^{s,\mu,(m)}]$ and

$[X_t^{s,x,\mu,(m)}]$ exist by 8.41 and are denoted by $p_m(\mu, s, t, \cdot)$ and $p_m(\mu, s, t, x, \cdot)$ and satisfy

$$p_m(\mu, s, t, y) = \int_{\mathbb{R}^d} p_m(\mu, s, t, x, y)\, d\mu(x). \tag{8.81}$$

Let us now give their parametrix expansions. We define for all $0 \leq s < t \leq T$, $r \in [s, t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $x, y \in \mathbb{R}^d$

$$\begin{aligned} \widehat{p}_m(s, r, t, x, y) &:= \widehat{p}(r, t, x, y) = q(t - r, y - x), \\ \mathcal{H}_m(\mu, s, r, t, x, y) &:= b(r, x, [X_r^{s,\mu,(m-1)}]) \cdot \partial_x \widehat{p}(s, r, t, x, y). \end{aligned} \tag{8.82}$$

Note that the proxy $\widehat{p}_m(s, r, t, x, \cdot)$ does not depend on m, μ and s. We denote by Φ_m the solution to the following Volterra integral equation

$$\Phi_m(\mu, s, r, t, x, y) = \mathcal{H}_m(\mu, s, r, t, x, y) + \mathcal{H}_m \otimes \Phi_m(\mu, s, r, t, x, y), \tag{8.83}$$

which is given by the uniform convergent series

$$\Phi_m(\mu, s, r, t, x, y) = \sum_{k=1}^{\infty} \mathcal{H}_m^k(\mu, s, r, t, x, y). \tag{8.84}$$

Let us recall that the Beta function $\mathcal{B}$ is defined, for all $x, y > 0$ by

$$\mathcal{B}(x, y) := \int_0^1 (1 - t)^{-1+x} t^{-1+y}\, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)},$$

where Γ is the Gamma function.

Applying Proposition 8.44 and Theorem 8.41, since all the controls are uniform with respect to the measure argument and thus on $m \geq 1$ too, we deduce the following two propositions.

Proposition 8.19.

- *There exists $C > 0$ such that for all $k \geq 1$, $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\mathcal{H}_m^k(\mu, s, r, t, x, y)| \leq C^k (t - r)^{-\frac{1}{\alpha} + (k-1)\left(1 - \frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \rho^1(t - r, y - x). \tag{8.85}$$

- *For $\gamma \in (0, \eta]$ such that $\gamma < \alpha - 1$, there exists $C > 0$ depending on γ such that for all $k \geq 1$, $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$\begin{aligned} |\mathcal{H}_m^k(\mu, s, r, t, x_1, y) - \mathcal{H}_m^k(\mu, s, r, t, x_2, y)| &\leq C^k (t - r)^{-\frac{\gamma+1}{\alpha} + (k-1)\left(1 - \frac{1}{\alpha}\right)} |x_1 - x_2|^\gamma \\ &\prod_{j=1}^{k-1} \mathcal{B}\left(-\frac{\gamma}{\alpha} + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \left[\rho^1(t - r, y - x_1) + \rho^1(t - r, y - x_2)\right]. \end{aligned} \tag{8.86}$$

- *The series (8.84) defining Φ_m is absolutely convergent and there exists $C > 0$ such that for all*

$m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$ and $x, y \in \mathbb{R}^d$

$$|\Phi_m(\mu, s, r, t, x, y)| \leq C(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r, y-x). \tag{8.87}$$

- *For $\gamma \in (0, \eta]$ such that $\gamma < \alpha - 1$, there exists $C > 0$ depending on γ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$|\Phi_m(\mu, s, r, t, x_1, y) - \Phi_m(\mu, s, r, t, x_2, y)| \leq C(t-r)^{-\frac{\gamma+1}{\alpha}}|x_1 - x_2|^\gamma \left[\rho^1(t-r, y-x_1) + \rho^1(t-r, y-x_2)\right]. \tag{8.88}$$

Proposition 8.20. *For any $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $x \in \mathbb{R}^d$, the distribution of $X_t^{s,x,\mu,(m)}$ has a density with respect to the Lebesgue measure denoted by $p_m(\mu, s, t, x, \cdot)$ and given by the absolutely convergent parametrix series*

$$\begin{aligned} p_m(\mu, s, t, x, y) &= \widehat{p}(s, t, x, y) + \sum_{k=1}^{\infty} \widehat{p} \otimes \mathcal{H}_m^k(\mu, s, t, x, y) \\ &= \widehat{p}(s, t, x, y) + \widehat{p} \otimes \Phi_m(s, t, x, y). \end{aligned} \tag{8.89}$$

For any $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $y \in \mathbb{R}^d$, $p_m(\mu, s, t, \cdot, y)$ is of class $\mathcal{C}^1$ on $\mathbb{R}^d$. Moreover, the following properties hold true.

- *There exists $C > 0$ such that for all $j \in \{0, 1\}$, $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\partial_x^j p_m(\mu, s, t, x, y)| \leq C(t-s)^{-\frac{j}{\alpha}}\rho^j(t-s, y-x). \tag{8.90}$$

- *For all $j \in \{0, 1\}$ and $\gamma \in (0, 1]$ with $\gamma \in (0, (2\alpha - 2) \wedge (\eta + \alpha - 1))$ if $j = 1$, there exists $C > 0$ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$|\partial_x^j p_m(\mu, s, t, x_1, y) - \partial_x^j p_m(\mu, s, t, x_2, y)| \leq C(t-s)^{-\frac{j+\gamma}{\alpha}}|x_1 - x_2|^\gamma \left[\rho^j(t-s, y-x_1) + \rho^j(t-s, y-x_2)\right]. \tag{8.91}$$

We state in the following proposition all the properties satisfied by the transition densities p_m which are used to prove Theorem 8.18. As the proof is rather long and technical, it is postponed to Section 8.8.

Proposition 8.21. *For any $m \geq 1$, $t \in (0, T]$, $y \in \mathbb{R}^d$, the map $(\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d \mapsto p_m(\cdot, \cdot, t, \cdot, y)$ belongs to $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d)$ and satisfies the following properties.*

- *There exists $C > 0$ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v \in \mathbb{R}^d$*

$$\begin{aligned} \left|\frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v)\right| &\leq (t-s)^{1-\frac{1}{\alpha}}\rho^0(t-s, y-x) \\ &\quad \left(\sum_{k=1}^{m} C^k (t-s)^{(k-1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right). \end{aligned} \tag{8.92}$$

- *There exists $C > 0$ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v \in \mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v)\right| \leq (t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} \rho^0(t-s, y-x)$$
$$\left(\sum_{k=1}^m C^k (t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\right). \quad (8.93)$$

- *For all $\tilde{\eta} \in (0, \eta \wedge (\alpha-1))$ there exists a constant $C > 0$ such that for all $m \geq 1$, $0 \leq s < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$*

$$|\partial_s p_m(\mu, s, t, x, y)| \leq (t-s)^{-1} \rho^{-\tilde{\eta}}(t-s, y-x) \sum_{k=1}^m C^k (t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}$$
$$\prod_{j=1}^{k-1} \mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right), 1-\frac{1}{\alpha}\right). \quad (8.94)$$

- *For all $\gamma \in (0,1] \cap (0, (2\alpha-2) \wedge (\eta+\alpha-1))$, there exists $C > 0$ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v_1, v_2 \in \mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v_1) - \partial_v \frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v_2)\right| \leq (t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}} |v_1 - v_2|^\gamma \rho^0(t-s, y-x)$$
$$\left(\sum_{k=1}^m C^k (t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\right). \quad (8.95)$$

- *There exists $C > 0$ such that for all $\gamma \in (0,1]$, $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v_1, v_2 \in \mathbb{R}^d$*

$$\left|\frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v_1) - \frac{\delta}{\delta m} p_m(\mu, s, t, x, y)(v_2)\right| \leq C(t-s)^{1-\frac{1+\gamma}{\alpha}} |v_1 - v_2|^\gamma \rho^0(t-s, y-x). \quad (8.96)$$

- *There exists $C > 0$ such that for all $\gamma \in (0,1]$, $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x_1, x_2, y, v \in \mathbb{R}^d$*

$$\left|\frac{\delta}{\delta m} p_m(\mu, s, t, x_1, y)(v) - \frac{\delta}{\delta m} p_m(\mu, s, t, x_2, y)(v)\right| \leq C(t-s)^{1-\frac{1+\gamma}{\alpha}} |x_1 - x_2|^\gamma$$
$$\left[\rho^0(t-s, y-x_1) + \rho^0(t-s, y-x_2)\right]. \quad (8.97)$$

- *For all $\gamma \in (0,1] \cap (0, \eta+\alpha-1)$, there exists $C > 0$ such that for all $m \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x_1, x_2, y, v \in \mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m} p_m(\mu, s, t, x_1, y)(v) - \partial_v \frac{\delta}{\delta m} p_m(\mu, s, t, x_2, y)(v)\right| \leq C(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}} |x_1 - x_2|^\gamma$$
$$\left[\rho^0(t-s, y-x_1) + \rho^0(t-s, y-x_2)\right]. \quad (8.98)$$

- *There exists $C > 0$ such that for all $\gamma \in (0,1]$, $m \geq 1$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y \in \mathbb{R}^d$*

$$|p_m(\mu_1, s, t, x, y) - p_m(\mu_2, s, t, x, y)| \leq C(t-s)^{1-\frac{1+\gamma}{\alpha}} W_1^\gamma(\mu_1, \mu_2) \rho^0(t-s, y-x). \quad (8.99)$$

- *For all $j \in \{0,1\}$ and $\gamma \in (0,1]$ with $\gamma \in (0, \alpha - 1 + \eta)$ if $j = 1$, there exists $C > 0$ such that for all $m \geq 1$, $0 \leq s < t \leq T$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$*

$$\left|\partial_x^j p_m(\mu_1, s, t, x, y) - \partial_x^j p_m(\mu_2, s, t, x, y)\right| \leq C(t-s)^{1-\frac{1+\gamma+j}{\alpha}} W_1^\gamma(\mu_1, \mu_2)\rho^j(t-s, y-x). \quad (8.100)$$

- *For all $\gamma \in (0,1]$, there exists $C > 0$ such that for all $m \geq 1$, $0 \leq s < t \leq T$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\begin{aligned}\left|\frac{\delta}{\delta m} p_m(\mu_1, s, t, x, y)(v) - \frac{\delta}{\delta m} p_m(\mu_2, s, t, x, y)(v)\right| &\leq (t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1, \mu_2)\rho^0(t-s, y-x) \\ &\sum_{k=1}^m C^k (t-s)^{(k-1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 - \frac{\gamma}{\alpha} + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right). \quad (8.101)\end{aligned}$$

- *For all $\gamma \in (0,1]$, there exists $C > 0$ such that for all $m \geq 1$, $0 \leq s < t \leq T$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\begin{aligned}\left|\partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x, y)(v) - \partial_v \frac{\delta}{\delta m} p_m(\mu_2, s, t, x, y)(v)\right| &\leq (t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1, \mu_2)\rho^0(t-s, y-x) \\ &\sum_{k=1}^m C^k (t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + \frac{\eta-\gamma}{\alpha} + j\left(1 + \frac{\eta-1}{\alpha}\right), 1 - \frac{1}{\alpha}\right). \quad (8.102)\end{aligned}$$

- *For all $j \in \{0,1\}$, $\gamma \in \left(0, 1 - \frac{j}{\alpha}\right)$, there exists a constant $C > 0$ such that for all $m \geq 1$, $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$*

$$\begin{aligned}&|\partial_x^j p_m(\mu, s_1, t, x, y) - \partial_x^j p_m(\mu, s_2, t, x, y)| \\ &\qquad \leq C\left[\frac{|s_1 - s_2|^\gamma}{(t-s_1)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1, y-x) + \frac{|s_1 - s_2|^\gamma}{(t-s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_2, y-x)\right]. \quad (8.103)\end{aligned}$$

- *For all $\gamma \in (0,1)$, there exists a constant $C > 0$ such that for all $m \geq 1$, $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\begin{aligned}&\left|\frac{\delta}{\delta m} p_m(\mu, s_1, t, x, y)(v) - \frac{\delta}{\delta m} p_m(\mu, s_2, t, x, y)(v)\right| \\ &\leq \left[\frac{|s_1 - s_2|^\gamma}{(t-s_1 \vee s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1 \vee s_2, y-x) + \frac{|s_1 - s_2|^\gamma}{(t-s_1 \wedge s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1 \wedge s_2, y-x)\right] \quad (8.104) \\ &\qquad \sum_{k=1}^m C^k (t-s_1 \vee s_2)^{(k-1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right).\end{aligned}$$

- *For all $\gamma \in \left(0, 1 + \frac{\eta-1}{\alpha}\right)$, there exists a constant $C > 0$ such that for all $m \geq 1$, $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$
\begin{aligned}
&\left|\partial_v \frac{\delta}{\delta m} p_m(\mu, s_1, t, x, y)(v) - \partial_v \frac{\delta}{\delta m} p_m(\mu, s_2, t, x, y)(v)\right| \\
&\leq \left[\frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} \rho^0(t - s_1 \vee s_2, y - x) + \frac{|s_1 - s_2|^\gamma}{(t - s_1 \wedge s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} \rho^0(t - s_1 \wedge s_2, y - x)\right] \\
&\quad \sum_{k=1}^{m} C^k (t - s_1 \vee s_2)^{(k-1)(1+\frac{\eta - 1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(\left[2\left(1 + \frac{\eta - 1}{\alpha}\right) - \gamma\right] \wedge 1 + (j-1)\left(1 + \frac{\eta - 1}{\alpha}\right), 1 - \frac{1}{\alpha}\right).
\end{aligned}
\tag{8.105}
$$

Remark 8.22. Note that by the asymptotic behavior of the Beta function, we get that all the series appearing in the right-hand side members of the preceding inequalities are convergent. This provides controls which are uniform with respect to m.

Step 2: Passage to the limit in the previous estimates. We are going to take the limit $m \to +\infty$ in all the estimates proved in Proposition 8.21 to deduce that the transition density $p(\mu, s, t, x, y)$ of the McKean-Vlasov SDE (8.16) satisfies the regularity properties and estimates of Theorem 8.18. This will be done along a converging subsequence given by the Arzelà-Ascoli theorem. Notice that all the partial sums of the series appearing in the all the upper-bounds of Proposition 8.21 have a limit when $m \to +\infty$ using the asymptotic behavior of the Beta function.

First of all, note that Theorem 8.2 yields

$$
\sup_{r \in [s,t]} d_{TV}([X_r^{s,\mu,(m)}], [X_r^{s,\mu}]) \underset{m \to +\infty}{\longrightarrow} 0.
$$

It follows that for all $k \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$, $x, y \in \mathbb{R}^d$

$$
\mathcal{H}_m^k(\mu, s, r, t, x, y) \underset{m \to +\infty}{\longrightarrow} \mathcal{H}^k(\mu, s, r, t, x, y),
$$

where $\mathcal{H}(\mu, s, t, x, y)$ was defined in (8.72). We can thus let m tend to infinity in the parametrix series (8.89) which yields the following pointwise convergence

$$
p_m(\mu, s, t, x, y) \underset{m \to +\infty}{\longrightarrow} p(\mu, s, t, x, y) \tag{8.106}
$$

thanks to the parametrix expansion (8.76) of p. Let us fix $(t, y) \in (0, T] \times \mathbb{R}^d$ and $\mathcal{K}$ a compact subset of $\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d$. Using (8.91), (8.100) and (8.103), we deduce that the sequence of functions $(p_m(\cdot, \cdot, t, \cdot, y))_m \in \mathcal{C}^0(\mathcal{K})^\mathbb{N}$ is uniformly equi-continuous on $\mathcal{K}$. It is also uniformly bounded by (8.90). The Arzelà-Ascoli theorem ensures that we can extract a subsequence of $(p_m(\cdot, \cdot, t, \cdot, y))_m$ which converges uniformly on $\mathcal{K}$, necessarily towards $p(\cdot, \cdot, t, \cdot, y)$ by (8.106). This yields the continuity of $p(\cdot, \cdot, t, \cdot, y)$ on $\mathcal{K}$ and thus, since $\mathcal{K}$ is arbitrary, on $\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d$. Moreover, passing to the limit in (8.90), (8.91), (8.100) and (8.103) for $j = 1$, along the converging subsequence of $(p_m(\cdot, \cdot, t, \cdot, y))_m$ previously obtained, we get that (8.57), (8.58), (8.65) and (8.68) hold true for $j = 0$. We now prove that $p(\mu, s, t, \cdot, y)$ is of class $\mathcal{C}^1$ on $\mathbb{R}^d$ for any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $s \in [0, t)$. To do this, we fix $R > 0$. By (8.90) and (8.91), we can apply the Arzelà-Ascoli theorem to the sequence $(\partial_x p_m(\mu, s, t, \cdot, y))_m \in \mathcal{C}^0(B_R)^\mathbb{N}$, where B_R denotes the open ball of $\mathbb{R}^d$ with radius R. Since R is arbitrary, we can construct, using a diagonal extraction procedure, a continuous function on $\mathbb{R}^d$ which is the limit, uniformly on each compact subset of $\mathbb{R}^d$, of a subsequence of $(\partial_x p_m(\mu,s,t,\cdot,y))_m$. This proves that $p_m(\mu,s,t,\cdot,y)$ is of class $\mathcal{C}^1$. By (8.90), (8.91), (8.100) and (8.103), the continuity of $\partial_x p(\cdot,\cdot,t,\cdot,y)$ on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times\mathbb{R}^d$ is again a consequence of the Arzelà-Ascoli theorem applied to $(\partial_x p_m(\cdot,\cdot,t,\cdot,y))_m \in (\mathcal{C}^0(\mathcal{K}))^{\mathbb{N}}$, where $\mathcal{K}$ is an arbitrary compact subset of $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times\mathbb{R}^d$. Taking the limit $m\to+\infty$ along a converging subsequence in (8.90), (8.91), (8.100) and (8.103), we deduce that (8.57), (8.58), (8.65) and (8.68) are satisfied.

Let us now focus on the existence of the linear derivative of p. We fix $\tau<t$ and $\mathcal{K}$ a closed and bounded subset of $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,\tau]\times(\mathbb{R}^d)^2$. Note that since $\beta>1$, $\mathcal{K}$ is relatively compact in $\mathcal{P}_1(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$ for the metric d defined for all $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $s_1,s_2\in[0,\tau]$, $x_1,x_2,v_1,v_2\in\mathbb{R}^d$ by

$$d((\mu_1,s_1,x_1,v_1),(\mu_2,s_2,x_2,v_2)) := W_1(\mu_1,\mu_2)+|s_1-s_2|+|x_1-x_2|+|v_1-v_2|.$$

Using (8.96), (8.97), (8.101), (8.104), we deduce that the sequence of functions $\left(\frac{\delta}{\delta m}p_m(\cdot,\cdot,t,\cdot,y)(\cdot)\right)_m$ is uniformly equi-continuous on $\mathcal{K}$ with respect to the metric d. Moreover (8.92) ensures that

$$\sup_{m\geq 1}\ \sup_{(\mu,s,x,v)\in\mathcal{P}_\beta(\mathbb{R}^d)\times[0,\tau]\times(\mathbb{R}^d)^2}\left|\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\right|<+\infty. \tag{8.107}$$

Then, we apply the Arzelà-Ascoli theorem, which gives the existence of a subsequence of $\left(\frac{\delta}{\delta m}p_m(\cdot,\cdot,t,\cdot,y)(\cdot)\right)_m$ which converges uniformly on $\mathcal{K}$ with respect to d. Since this is true for all $\tau<t$ and for every bounded subset $\mathcal{K}$ of $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,\tau]\times(\mathbb{R}^d)^2$, we can use a diagonal extraction procedure. This yields the existence of function g continuous with respect to d on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$ such that, up to an extraction, $\left(\frac{\delta}{\delta m}p_m(\cdot,\cdot,t,\cdot,y)(\cdot)\right)_m$ converges towards g uniformly on each compact subset of $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$. Note that g is also continuous with respect to the usual metric on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$ and that (8.107) implies that for each $\tau\in[0,t)$, we have

$$\sup_{(\mu,s,x,v)\in\mathcal{P}_\beta(\mathbb{R}^d)\times[0,\tau]\times(\mathbb{R}^d)^2}|g(\mu,s,x,v)|<+\infty.$$

We now prove that $p(\cdot,s,t,x,y)$ admits a linear derivative given, for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $s\in[0,t)$, $x,v\in\mathbb{R}^d$ by

$$\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)=g(\mu,s,x,v), \tag{8.108}$$

which is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$. For all $\mu,\nu\in\mathcal{P}_\beta(\mathbb{R}^d)$, one has

$$p_m(\mu,s,t,x,y)-p_m(\nu,s,t,x,y)=\int_0^1\int_{\mathbb{R}^d}\frac{\delta}{\delta m}p_m(\lambda\mu+(1-\lambda)\nu,s,t,x,y)(v)\,d(\mu-\nu)(v)\,d\lambda.$$

We take the limit $m\to+\infty$ along the subsequence of $\left(\frac{\delta}{\delta m}p_m(\cdot,\cdot,t,\cdot,y)(\cdot)\right)_m$ converging towards g that we have obtained above. By the dominated convergence theorem justified by (8.107) and since $(p_m)_m$ converges pointwise towards p, we obtain that

$$p(\mu,s,t,x,y)-p(\nu,s,t,x,y)=\int_0^1\int_{\mathbb{R}^d}g(\lambda\mu+(1-\lambda)\nu,s,x,v)\,d(\mu-\nu)(v)\,d\lambda.$$

This proves (8.108). Moreover, taking the limit $m\to+\infty$ in (8.92), (8.96), (8.97), (8.101) and (8.104) along the converging subsequence yields (8.59), (8.62), (8.63), (8.66) and (8.69). Using again the Arzelà-Ascoli theorem, we prove that for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $s\in[0,t)$, $x,y\in\mathbb{R}^d$, the map $\frac{\delta}{\delta m}p(\mu,s,t,x,y)(\cdot)$ is of class $\mathcal{C}^1$ on $\mathbb{R}^d$, that $\partial_v \frac{\delta}{\delta m} p(\cdot,\cdot,t,\cdot,y)(\cdot)$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^2$ and that it satisfies (8.60), (8.61), (8.64), (8.67) and (8.70).

Remark 8.23. We have made the proof of the extraction of converging subsequences for a fixed $y\in\mathbb{R}^d$. However, following exactly the same lines as for the estimates of Hölder continuity with respect to x (8.91), (8.97) and (8.98), we can prove similar estimates with respect to y. This ensures that the converging subsequences can also be assumed to converge uniformly on each compact subset of $\mathbb{R}^d$ with respect to y.

Let us prove (8.71). We fix $\varepsilon>0$. Thanks to (8.92) and (8.93), we can find a compact subset $\mathcal{K}$ of $\mathbb{R}^d$ such that

$$\int_{\mathcal{K}^c}\sup_{m\geq 1}\left\{\left|\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\right|+\left|\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\right|\right\}dy\leq\varepsilon$$

By Remark (8.23), up to extracting a subsequence, we can assume that the functions $\left(\frac{\delta}{\delta m}p_m(\mu,s,t,x,\cdot)(v)\right)_m$ and $\left(\partial_v\frac{\delta}{\delta m}p_m(\mu,s,t,x,\cdot)(v)\right)_m$ converge uniformly on $\mathcal{K}$ towards $\frac{\delta}{\delta m}p(\mu,s,t,x,\cdot)(v)$ and $\partial_v\frac{\delta}{\delta m}p(\mu,s,t,x,\cdot)(v)$. Noticing that (8.71) is true for p_m by (8.174), (8.164), (8.175) and (8.165). We can thus write

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d}\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)\,dy\right|\\
&=\left|\int_{\mathbb{R}^d}\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)-\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\,dy\right|\\
&\leq\int_{\mathcal{K}}\left|\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)-\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\right|dy+\varepsilon.
\end{aligned}$$

We conclude by letting m tend to infinity and with a similar reasoning for $\partial_v\frac{\delta}{\delta m}p$. This ends the proof of Theorem 8.18.

Proof of Theorem 8.4. We first need to prove that for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $t\in(0,T]$, $x,y\in\mathbb{R}^d$, the map $s\in[0,t)\mapsto p(\mu,\cdot,t,x,y)$ is of class $\mathcal{C}^1$ on $[0,t)$, that $\partial_s p(\cdot,\cdot,t,\cdot,y)$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times\mathbb{R}^d$ and satisfies (8.20). Let us fix $h\in[0,s]$. By the well-posedness of the nonlinear martingale problem proved in Theorem 8.2, we deduce that the transition density satisfies the following Markov property

$$p(\mu,s-h,t,x,y)=\mathbb{E}p([X_s^{s-h,\mu}],s,t,X_s^{s-h,x,\mu},y).\tag{8.109}$$

By (8.57), (8.59), (8.60), (8.58) and (8.61) with $\gamma>\alpha-1$, we can apply Itô's formula of Proposition 8.40 for the function $(\mu,x)\in\mathcal{P}_\beta(\mathbb{R}^d)\times\mathbb{R}^d\mapsto p(\mu,s,t,x,y)$. Taking the expectation in Itô's formula, we obtain that

$$\mathbb{E}p([X_s^{s-h,\mu}],s,t,X_s^{s-h,x,\mu},y)=p(\mu,s,t,x,y)+\int_{s-h}^{s}\mathcal{L}_r p(\cdot,s,t,\cdot,y)([X_r^{s-h,\mu}],X_r^{s-h,x,\mu})\,dr,$$

where $\mathcal{L}_r$ was defined in (8.21). We thus have

$$\frac{1}{h}(p(\mu,s-h,t,x,y)-p(\mu,s,t,x,y))=\frac{1}{h}\int_{s-h}^{s}\mathbb{E}\mathcal{L}_r p(\cdot,s,t,\cdot,y)([X_r^{s-h,\mu}],X_r^{s-h,x,\mu})\,dr.$$

Using the continuity and the boundedness of b as well as the Hölder continuity and the bounds on

$p(\cdot,\cdot,t,\cdot,y)$, $\partial_x p(\cdot,\cdot,t,\cdot,y)$, $\frac{\delta}{\delta m}p(\cdot,\cdot,t,\cdot,y)(\cdot)$ and $\partial_v \frac{\delta}{\delta m}p(\cdot,\cdot,t,\cdot,y)(\cdot)$ proved above, we find that

$$\frac{1}{h}(p(\mu,s-h,t,x,y)-p(\mu,s,t,x,y)) \underset{h\to 0^+}{\longrightarrow} \mathcal{L}_s p(\cdot,s,t,\cdot,y)(\mu,x).$$

The map $s\in[0,t)\mapsto p(\mu,s,t,x,y)$ is thus left-differentiable on $[0,t)$. It also follows that the map $(\mu,s,x)\in\mathcal{P}_\beta(\mathbb{R}^d)\times[0,t)\times\mathbb{R}^d\mapsto \mathcal{L}_s p(\cdot,s,t,\cdot,y)(\mu,x)$ is continuous. This proves that $p(\mu,\cdot,t,x,y)$ is $\mathcal{C}^1$ on $[0,t)$ and that it satisfies for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $s\in[0,t)$, $x,y\in\mathbb{R}^d$

$$\partial_s p(\mu,s,t,x,y) = -\mathcal{L}_s p(\cdot,s,t,\cdot,y)(\mu,x).$$

Let us now fix $f:\mathbb{R}^d\to\mathbb{R}$ a bounded and uniformly continuous function. We fix $\varepsilon>0$. There exists $\delta>0$ such that for all $x,y\in\mathbb{R}^d$ with $|x-y|\le\delta$, we have $|f(x)-f(y)|\le\varepsilon$. Using (8.57), we obtain that

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} f(y)p(\mu,s,t,x,y)\,dy - f(x)\right| &= \sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d}(f(y)-f(x))p(\mu,s,t,x,y)\,dy\right| \\
&\le \varepsilon + C\|f\|_\infty \int_{|y|>\delta}\rho^0(t-s,y)\,dy \\
&\le \varepsilon + C\|f\|_\infty \int_{|y|>\delta}(t-s)^{-\frac{d}{\alpha}}(1+(t-s)^{-\frac{1}{\alpha}}|y|)^{-d-\alpha}\,dy \\
&= \varepsilon + C\|f\|_\infty \int_{|z|>(t-s)^{-\frac{1}{\alpha}}\delta}(1+|z|)^{-d-\alpha}\,dz.
\end{aligned}$$

We conclude taking the lim sup when $s\to t$ in the preceding inequality that $p(\mu,s,t,x,\cdot)\underset{s\to t^-}{\longrightarrow}\delta_x$ in the weak sense.

The estimates (8.22) and (8.24) have been proved in Theorem 8.18 and the estimate (8.23) is proved in Theorem 8.41. It thus remains to prove (8.25). Using (8.59), (8.60), (8.61) and the same reasoning as used in the proof of (8.229) in Theorem 8.41, we find that there exists a positive constant C such that for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\le s<t\le T$, $x,y,v\in\mathbb{R}^d$

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d}\left[\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v+z)-\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)-z\cdot\partial_v\frac{\delta}{\delta m}p(\mu,s,t,x,y)(v)\right]\frac{dz}{|z|^{d+\alpha}}\right| \\
&\qquad\le C(t-s)^{-1+1-\frac{1}{\alpha}}\rho^0(t-s,y-x).
\end{aligned}$$

This concludes the proof of (8.25) and thus ends the proof of Theorem 8.4.

□

8.5 Backward Kolmogorov PDE on the space of measures

We prove Theorem 8.9 in this section.

Step 1: Continuity of U on $[0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)$. Reasoning exactly as in the proof of [CdRF22, Proposition 6.1], the continuity of U on $[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d)$ follows from the continuity of the map $(\mu,s,x)\in\mathcal{P}_\beta(\mathbb{R}^d)\times[0,T)\times\mathbb{R}^d\mapsto p(\mu,s,t,x,y)$ and (8.57). Let us prove it on $[0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)$. Let $(t_n)_n\in[0,T)^{\mathbb{N}}$, $(\mu_n)_n\in\mathcal{P}_\beta(\mathbb{R}^d)^{\mathbb{N}}$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$ such that $|t_n-T|\underset{n\to+\infty}{\longrightarrow}0$ and $W_\beta(\mu_n,\mu)\underset{n\to+\infty}{\longrightarrow}0$. Since ϕ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d)$, it is enough to prove that $W_\beta([X_T^{t_n,\mu_n}],\mu)\underset{n\to+\infty}{\longrightarrow}0$. We start with the weak convergence. Let us fix $f:\mathbb{R}^d\to\mathbb{R}$ a bounded and uniformly continuous function. We write

$$\begin{aligned}\int_{\mathbb{R}^{2d}} f(y)p(\mu_n,t_n,T,x,y)\,dy\,d\mu_n(x)-\int_{\mathbb{R}^d} f(x)d\mu(x) &= \int_{\mathbb{R}^{2d}}(f(y)-f(x))p(\mu_n,t_n,T,x,y)\,dy\,d\mu_n(x)\\ &\quad+\int_{\mathbb{R}^d} f(x)d(\mu_n-\mu)(x)\\ &=: I_1+I_2.\end{aligned}$$

Since $W_\beta(\mu_n,\mu)\underset{n\to+\infty}{\longrightarrow}0$, it is clear, by definition, that $I_2\underset{n\to+\infty}{\longrightarrow}0$. Let us now deal with I_1. We fix $\varepsilon>0$. There exists $\delta>0$ such that for all $x,y\in\mathbb{R}^d$ with $|x-y|\le\delta$, we have $|f(x)-f(y)|\le\varepsilon$. Using (8.57), we obtain that

$$\begin{aligned}|I_1| &\le \int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}(f(y)-f(x))p(\mu_n,t_n,T,x,y)\,dy\right|\,d\mu_n(x)\\ &\le \varepsilon + C\|f\|_\infty\int_{|y|>\delta}\rho^0(T-t_n,y)\,dy\\ &\le \varepsilon + C\|f\|_\infty\int_{|y|>\delta}(T-t_n)^{-\frac{d}{\alpha}}(1+(T-t_n)^{-\frac{1}{\alpha}}|y|)^{-d-\alpha}\,dy\\ &= \varepsilon + C\|f\|_\infty\int_{|z|>(T-t_n)^{-\frac{1}{\alpha}}\delta}(1+|z|)^{-d-\alpha}\,dz.\end{aligned}$$

We conclude that $I_1\underset{n\to+\infty}{\longrightarrow}0$. It remains to show that

$$\int_{\mathbb{R}^{2d}}|y|^\beta p(\mu_n,t_n,T,x,y)\,d\mu_n(x)\,dy\underset{n\to+\infty}{\longrightarrow}\int_{\mathbb{R}^d}|x|^\beta d\mu(x).$$

To see this, we write

$$\begin{aligned}\int_{\mathbb{R}^{2d}}|y|^\beta p(\mu_n,t_n,T,x,y)\,dy\,d\mu_n(x)-\int_{\mathbb{R}^d}|x|^\beta d\mu(x) &= \int_{\mathbb{R}^{2d}}(|y|^\beta-|x|^\beta)p(\mu_n,t_n,T,x,y)\,dy\,d\mu_n(x)\\ &\quad+\int_{\mathbb{R}^d}|x|^\beta d(\mu_n-\mu)(x)\\ &=: J_1+J_2.\end{aligned}$$

Since $W_\beta(\mu_n,\mu)\underset{n\to+\infty}{\longrightarrow}0$, we deduce that $J_2\underset{n\to+\infty}{\longrightarrow}0$. For J_1, by the mean-value theorem, there exists a positive constant C such that for all $x,y\in\mathbb{R}^d$

$$||y|^\beta-|x|^\beta|\le C|y-x|(|y|^{\beta-1}+|x|^{\beta-1}).$$

We obtain by (8.57) and the space-time inequality (8.231) that

$$
\begin{aligned}
|J_1| &\le C(T-t_n)^{\frac{1}{\alpha}} \int_{\mathbb{R}^{2d}} (|y|^{\beta-1} + |x|^{\beta-1})\rho^{-1}(T-t_n, y-x)\, dy\, d\mu_n(x) \\
&\le C(T-t_n)^{\frac{1}{\alpha}} \int_{\mathbb{R}^{2d}} \rho^{-\beta}(T-t_n, y-x) + |x|^{\beta-1}\rho^{-1}(T-t_n, y-x)\, dy\, d\mu_n(x) \\
&\le C(T-t_n)^{\frac{1}{\alpha}} \left(1 + \int_{\mathbb{R}^d} |x|^{\beta-1}\, d\mu_n(x)\right) \\
&\le C(T-t_n)^{\frac{1}{\alpha}},
\end{aligned}
$$

since $\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} |x|^{\beta-1}\, d\mu_n(x) < +\infty$. We have thus proved that $W_\beta([X_T^{t_n,\mu_n}], \mu) \underset{n\to+\infty}{\longrightarrow} 0$, which concludes the first step.

Step 2: Estimate (8.29) **and** (8.30)**.** By Proposition 8.39 and Proposition 8.21, we know that the map U belongs to $\mathcal{C}^1([0,T) \times \mathcal{P}_\beta(\mathbb{R}^d))$. Moreover, for any $t \in [0,T)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $v \in \mathbb{R}^d$, we have

$$
\begin{aligned}
\frac{\delta}{\delta m} U(t,\mu)(v) &= \int_{\mathbb{R}^d} \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) p(\mu,t,T,v,y)\, dy \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(x)\right) \frac{\delta}{\delta m} p(\mu,t,T,x,y)(v)\, dy\, d\mu(x),
\end{aligned} \tag{8.110}
$$

and

$$
\begin{aligned}
\partial_v \frac{\delta}{\delta m} U(t,\mu)(v) &= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(v)\right) \partial_x p(\mu,t,T,v,y)\, dy \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(x)\right) \partial_v \frac{\delta}{\delta m} p(\mu,t,T,x,y)(v)\, dy\, d\mu(x).
\end{aligned} \tag{8.111}
$$

By (8.57), (8.60), we similarly get that there exists a positive constant C such that for all $t \in [0,T)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $v \in \mathbb{R}^d$

$$
\begin{aligned}
\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v)\right| &\le C(T-t)^{\frac{\delta-1}{\alpha}} + C(T-t)^{\frac{\delta-1}{\alpha}+1+\frac{\eta-1}{\alpha}} \\
&\le C(T-t)^{\frac{\delta-1}{\alpha}}.
\end{aligned}
$$

We now prove (8.30). Let us first assume that $|v_1 - v_2| \ge (T-t)^{\frac{1}{\alpha}}$. In this case, it follows from (8.29) that

$$
\begin{aligned}
\left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2)\right| &\le \left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1)\right| + \left|\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2)\right| \\
&\le C(T-t)^{\frac{\delta-1}{\alpha}} \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}} |v_1 - v_2|^\gamma.
\end{aligned}
$$

Assume now that $|v_1 - v_2| < (T-t)^{\frac{1}{\alpha}}$. Using (8.111) and the fact that $\int_{\mathbb{R}^d} \partial_x p(\mu,t,T,v_1,y)\, dy = 0$, one has

$$
\begin{aligned}
&\partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2) \\
&= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(v_2) \right) (\partial_x p(\mu,t,T,v_1,y) - \partial_x p(\mu,t,T,v_2,y))\, dy \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu}])(x) \right) (\partial_v \frac{\delta}{\delta m} p(\mu,t,T,x,y)(v_1) - \partial_v \frac{\delta}{\delta m} p(\mu,t,T,x,y)(v_2))\, dy\, d\mu(x).
\end{aligned}
$$

Thanks to the uniform δ-Hölder continuity of $\frac{\delta}{\delta m}\phi(\mu)(\cdot)$, (8.58), (8.61) since $\gamma \in (0,(2\alpha-2)\wedge(\eta+\alpha-1))$ and (8.232) since $|v_1 - v_2| < (T-t)^{\frac{1}{\alpha}}$, we obtain that

$$
\begin{aligned}
&\left| \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2) \right| \\
&\leq C \int_{\mathbb{R}^d} |y - v_2|^{\delta} (T-t)^{-\frac{\gamma+1}{\alpha}} |v_1 - v_2|^{\gamma} \rho^1(T-t, y - v_2)\, dy \\
&\quad + C \int_{\mathbb{R}^{2d}} |y - x|^{\delta} (T-t)^{-\frac{\gamma+1}{\alpha}+1+\frac{\eta-1}{\alpha}} |v_1 - v_2|^{\gamma} \rho^0(T-t, y-x)\, dy\, d\mu(x).
\end{aligned}
$$

Using the space-time inequality (8.231), we deduce that

$$
\begin{aligned}
&\left| \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_1) - \partial_v \frac{\delta}{\delta m} U(t,\mu)(v_2) \right| \\
&\leq C\ (T-t)^{\frac{\delta-\gamma-1}{\alpha}} |v_1 - v_2|^{\gamma}.
\end{aligned}
$$

Step 3: Time-derivative and backward Kolmogorov PDE (8.31)**.** Let us fix $t \in [0,T)$ and $h \in [0,t]$ such that $t-h \in [0,T)$. From the Markov property stemming from the well-posedness of the associated martingale problem related to (8.16), we obtain that

$$
U(t-h,\mu) = \phi([X_T^{t-h,\mu}]) = \phi([X_T^{t,[X_t^{t-h,\mu}]}]) = U(t,[X_t^{t-h,\mu}]).
$$

Since $U \in \mathcal{C}^1([0,T)\times\mathcal{P}_\beta(\mathbb{R}^d))$ and thanks to (8.29) and (8.30), we can apply Itô's formula of Proposition 8.40 for the function $U(t,\cdot)$. It yields

$$
\begin{aligned}
U(t-h,\mu) &= U(t,\mu) + \int_{t-h}^{t} \mathbb{E}\left(\partial_v \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \cdot b(s, X_s^{t-h,\mu}, [X_s^{t-h,\mu}]) \right) ds \\
&\quad + \int_{t-h}^{t} \mathbb{E} \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu} + z) - \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \right. \\
&\qquad \left. - \partial_v \frac{\delta}{\delta m} U(t,[X_s^{t-h,\mu}])(X_s^{t-h,\mu}) \cdot z \right) \frac{dz}{|z|^{d+\alpha}}\, ds \\
&= U(t,\mu) + \int_{t-h}^{t} \mathscr{L}_s U(t,\cdot)([X_s^{t-h,\mu}])\, ds,
\end{aligned}
$$

where $\mathscr{L}_s$ was defined in (8.27). Using the continuity and the boundedness of b as well as the regularity properties of the map $\mu \mapsto U(t,\mu)$ obtained above, we find that

$$\frac{1}{h}(U(t-h,\mu)-U(t,\mu)) \underset{h\to 0^+}{\longrightarrow} \mathscr{L}_t U(t,\cdot)(\mu).$$

This proves the left-differentiability of the map $t\in[0,T)\mapsto U(t,\mu)$. It also follows from the regularity of the drift b and of U on $[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d)$ that the map $(t,\mu)\in[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d)\mapsto \mathscr{L}_t U(t,\cdot)(\mu)$ is continuous. Thus, the map $t\in[0,T)\mapsto U(t,\mu)$ is $\mathcal{C}^1$ and satisfies for all $t\in[0,T)$ and $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$

$$\partial_t U(t,\mu)+\mathscr{L}_t U(t,\cdot)(\mu)=0.$$

This shows that U is solution to the backward Kolmogorov PDE (8.31).

Step 4: Uniqueness of the solution to (8.31). Let us consider $V\in\mathcal{C}^0([0,T]\times\mathcal{P}_\beta(\mathbb{R}^d))\cap\mathcal{C}^1([0,T)\times\mathcal{P}_\beta(\mathbb{R}^d))$ another solution to (8.31) satisfying (8.29) and (8.30). Let us fix $(t,\mu)\in[0,T)\times\mathcal{P}_\beta(\mathbb{R}^d)$. For any $\tau\in[t,T)$, we can apply Itô's formula of Proposition 8.40 for the map $(s,\mu)\in[t,\tau]\times\mathcal{P}_\beta(\mathbb{R}^d)\mapsto V(s,\mu)$ which yields

$$V(\tau,[X_\tau^{t,\mu}])=V(t,\mu)+\int_t^\tau \partial_s V(s,[X_s^{t,\mu}])\,ds+\int_t^\tau \mathscr{L}_s V(s,\cdot)([X_s^{t,\mu}])\,ds.$$

Since V solves (8.31), we obtain that $V(\tau,[X_\tau^{t,\mu}])=V(t,\mu)$. We then use the continuity of the maps $(t,\mu)\in[0,T]\times\mathcal{P}_\beta(\mathbb{R}^d)\mapsto V(t,\mu)$ and $\tau\in[t,T]\mapsto[X_\tau^{t,\mu}]\in\mathcal{P}_\beta(\mathbb{R}^d)$ to let τ tends to T. This yields

$$\phi([X_T^{t,\mu}])=U(t,\mu)=V(t,\mu).$$

8.6 Quantitative weak propagation of chaos

This section is dedicated to prove Theorem 8.12. We first need to establish additional regularity properties on the solution to the PDE (8.31) to prove our weak propagation of chaos result.

Proposition 8.24. *Let us fix $\delta\in(0,1]$, $L>0$ and $\gamma\in(0,1]\cap(0,(2\alpha-2)\wedge(\eta+\alpha-1))$. Then, under Assumption (H3), there exists positive constant $C=C(d,T,\alpha,\beta,(H3),\delta,L,\gamma)$ such that for all $\phi\in\mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ (defined in Definition 8.10), the solution U of the backward Kolmogorov PDE (8.31) with terminal condition ϕ at time T satisfies the following properties.*

- *For all $t\in[0,T)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $v\in\mathbb{R}^d$*

$$\left|\partial_v\frac{\delta}{\delta m}U(t,\mu)(v)\right|\le C(T-t)^{\frac{\delta-1}{\alpha}}. \tag{8.112}$$

- *For all $t\in[0,T)$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $v_1,v_2\in\mathbb{R}^d$*

$$\left|\partial_v\frac{\delta}{\delta m}U(t,\mu_1)(v_1)-\partial_v\frac{\delta}{\delta m}U(t,\mu_2)(v_2)\right|\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}}\left(|v_1-v_2|^\gamma+W_1^\gamma(\mu_1,\mu_2)\right). \tag{8.113}$$

- *For all $t\in[0,T)$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $v\in\mathbb{R}^d$*

$$\left|\frac{\delta}{\delta m}U(t,\mu_1)(v)-\frac{\delta}{\delta m}U(t,\mu_2)(v)\right|\le C(T-t)^{\frac{\delta-1}{\alpha}}W_1(\mu_1,\mu_2). \tag{8.114}$$

Proof of Proposition 8.24. First, note that (8.112) have been proved in Theorem 8.9.

Proof of (8.113). We write

$$\begin{aligned}\partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_1) - \partial_v \frac{\delta}{\delta m}U(t,\mu_2)(v_2) &= \partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_1) - \partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_2)\\ &\quad + \partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_2) - \partial_v \frac{\delta}{\delta m}U(t,\mu_2)(v_2)\\ &=: I_1 + I_2.\end{aligned} \tag{8.115}$$

Then, using (8.30), we obtain that

$$|I_1| \le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}}|v_1-v_2|^{\gamma}. \tag{8.116}$$

We now focus on I_2. Using (8.111), we get

$$\begin{aligned}&\partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_2) - \partial_v \frac{\delta}{\delta m}U(t,\mu_2)(v_2)\\ &= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(v_2)\right)\partial_x p(\mu_1,t,T,v_2,y)\,dy\\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x)\right)\partial_v \frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v_2)\,dy\,d\mu_1(x)\\ &\quad - \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(v_2)\right)\partial_x p(\mu_2,t,T,v_2,y)\,dy\\ &\quad - \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right)\partial_v \frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\,dy\,d\mu_2(x).\end{aligned}$$

We decompose it in the following way

$$\begin{aligned}&\partial_v \frac{\delta}{\delta m}U(t,\mu_1)(v_2) - \partial_v \frac{\delta}{\delta m}U(t,\mu_2)(v_2)\\ &= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(v_2) - \left[\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(v_2)\right]\right)\partial_x p(\mu_1,t,T,v_2,y)\,dy\\ &\quad + \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(v_2)\right)\left(\partial_x p(\mu_1,t,T,v_2,y) - \partial_x p(\mu_2,t,T,v_2,y)\right)\,dy\\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x) - \left[\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right]\right)\\ &\qquad\qquad \partial_v \frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v_2)\,dy\,d\mu_1(x)\\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right)\\ &\qquad\qquad \left(\partial_v \frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v_2) - \partial_v \frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\right)\,dy\,d\mu_1(x)\\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right)\partial_v \frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\,dy\,d(\mu_1-\mu_2)(x)\\ &=: J_1 + J_2 + J_3 + J_4 + J_5.\end{aligned} \tag{8.117}$$

Before estimating each term of the preceding decomposition, let us prove that for all $\gamma \in (0,1]$, there exists a positive constant C such that for any $\phi \in \mathcal{C}_L^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$, $t \in [0,T)$, $\mu_1, \mu_2, \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$

$$\left|\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x) - \left[\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right]\right| \\ \leq C(T-t)^{-\frac{\gamma}{\alpha}}|x-y|^\delta W_1^\gamma(\mu_1,\mu_2). \quad (8.118)$$

To prove this, we write

$$\begin{aligned}
&\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x) - \left[\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right] \\
&= \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,v) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(x,v)\right)(p(\mu_1,t,T,v) - p(\mu_2,s,T,v))\,dv\,d\lambda \\
&= \int_0^1 \int_{\mathbb{R}^{2d}} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,v) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(x,v)\right)p(\mu_1,t,T,x',v)\,d(\mu_1-\mu_2)(x')\,dv\,d\lambda \\
&\quad + \int_0^1 \int_{\mathbb{R}^{2d}} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,v) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(x,v)\right)\left(p(\mu_1,t,T,x',v) - p(\mu_2,t,T,x',v)\right)\,d\mu_2(x')\,dv\,d\lambda \\
&=: K_1 + K_2,
\end{aligned}$$

where $m_\lambda := \lambda[X_T^{t,\mu_1}] + (1-\lambda)[X_T^{t,\mu_2}]$, for $\lambda \in [0,1]$. It follows from (8.65) and the δ-Hölder continuity of $\frac{\delta^2}{\delta m^2}\phi(\mu)(\cdot,v)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $v \in \mathbb{R}^d$ that

$$|K_2| \leq C(T-t)^{1-\frac{\gamma+1}{\alpha}}|y-x|^\delta W_1^\gamma(\mu_1,\mu_2). \quad (8.119)$$

Concerning K_1, we set for $x' \in \mathbb{R}^d$

$$F(x') := \int_{\mathbb{R}^d} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,v) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(x,v)\right)p(\mu_1,t,T,x',v)\,dv.$$

Let us fix $x', x'' \in \mathbb{R}^d$. Thanks to (8.58) and the δ-Hölder continuity of $\frac{\delta^2}{\delta m^2}\phi(\mu)(\cdot,v)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $v \in \mathbb{R}^d$, one gets that

$$|F(x') - F(x'')| \leq C|x-y|^\delta (T-t)^{-\frac{\gamma}{\alpha}}|x'-x''|^\gamma.$$

Taking an optimal coupling between μ_1 and μ_1 for W_1 directly yields, with Jensen's inequality,

$$|K_1| \leq C|x-y|^\delta (T-t)^{-\frac{\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2). \quad (8.120)$$

Combining (8.119) and (8.120) concludes the proof of (8.118) since $\alpha \in (1,2)$. We can now turn to estimate each term of (8.117). Using (8.118), (8.57) and the space-time inequality (8.231), one obtains that

$$\begin{aligned}
|J_1| &\leq C\int_{\mathbb{R}^d} (T-t)^{-\frac{\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2)|y-v_2|^\delta (T-t)^{-\frac{1}{\alpha}}\rho^1(T-t,y-v_2)\,dy \\
&\leq C(T-t)^{\frac{\delta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\int_{\mathbb{R}^d} \rho^{1-\delta}(T-t,y-v_2)\,dy \\
&\leq C(T-t)^{\frac{\delta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned} \quad (8.121)$$

For J_2, it follows from the δ-Hölder continuity of $\frac{\delta}{\delta m}\phi(\mu)(\cdot)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, (8.65) since $\gamma < \eta + \alpha - 1$, and the space-time inequality (8.231) that

$$\begin{aligned}
|J_2| &\le C \int_{\mathbb{R}^d} |y - v_2|^\delta (T-t)^{1-\frac{2+\gamma}{\alpha}} W_1^\gamma(\mu_1, \mu_2) \rho^1(T-t, y - v_2)\, dy \\
&\le C(T-t)^{1+\frac{\delta-2-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2),
\end{aligned} \tag{8.122}$$

since $\alpha \in (1,2)$. For J_3, (8.118), (8.60) and the space-time inequality (8.231) yield

$$\begin{aligned}
|J_3| &\le C \int_{\mathbb{R}^{2d}} (T-t)^{-\frac{\gamma}{\alpha}} |y-x|^\delta W_1^\gamma(\mu_1,\mu_2)(T-t)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} \rho^0(T-t, y-x)\, dy\, d\mu_1(x) \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned} \tag{8.123}$$

It follows from the δ-Hölder continuity of $\frac{\delta}{\delta m}\phi(\mu)(\cdot)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, (8.67) and the space-time inequality (8.231) that

$$\begin{aligned}
|J_4| &\le C \int_{\mathbb{R}^{2d}} |y-x|^\delta (T-t)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \rho^0(T-t, y-x)\, dy\, d\mu_1(x) \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned} \tag{8.124}$$

We finally deal with J_5. We set for $x \in \mathbb{R}^d$

$$H(x) := \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x) \right) \partial_v \frac{\delta}{\delta m} p(\mu_2, t, T, x, y)(v_2)\, dy.$$

Let us prove that for any $x, x' \in \mathbb{R}^d$

$$|H(x) - H(x')| \le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}} |x-x'|^\gamma. \tag{8.125}$$

Assume first that $|x - x'| > (T-t)^{\frac{1}{\alpha}}$. The δ-Hölder continuity of $\frac{\delta}{\delta m}\phi(\mu)(\cdot)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, (8.60) and the space-time inequality (8.231) ensure that

$$\begin{aligned}
|H(x) - H(x')| &\le |H(x)| + |H(x')| \\
&\le C \int_{\mathbb{R}^d} |y-x|^\delta (T-t)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} \rho^0(T-t, y-x)\, dy \\
&\le C(T-t)^{\frac{\delta-1}{\alpha}+1+\frac{\eta-1}{\alpha}} \\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}} |x-x'|^\gamma.
\end{aligned}$$

In the case where $|x-x'| \le (T-t)^{\frac{1}{\alpha}}$, thanks to (8.71), we write

$$
\begin{aligned}
&H(x)-H(x')\\
&=\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x')-\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right)\partial_v\frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\,dy\\
&\quad+\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)-\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x')\right)\left(\partial_v\frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)-\partial_v\frac{\delta}{\delta m}p(\mu_2,t,T,x',y)(v_2)\right)dy\\
&=\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)-\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x')\right)\left(\partial_v\frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)-\partial_v\frac{\delta}{\delta m}p(\mu_2,t,T,x',y)(v_2)\right)dy.
\end{aligned}
$$

Thanks to the δ-Hölder continuity of $\frac{\delta}{\delta m}\phi(\mu)(\cdot)$ uniformly with respect to $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$ and (8.64) since $\gamma<\eta+\alpha-1$, we directly get that

$$
|H(x)-H(x')|\le C\int_{\mathbb{R}^d}|y-x'|^{\delta}(T-t)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|x-x'|^{\gamma}\left[\rho^0(T-t,y-x')+\rho^0(T-t,y-x)\right]dy.
$$

Since $|x-x'|\le (T-t)^{\frac{1}{\alpha}}$, it follows from (8.232) and the space-time inequality (8.231) that

$$
\begin{aligned}
|H(x)-H(x')|&\le C\int_{\mathbb{R}^d}|y-x'|^{\delta}(T-t)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|x-x'|^{\gamma}\rho^0(T-t,y-x')\,dy\\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}}|x-x'|^{\gamma}.
\end{aligned}
$$

Then, it follows from (8.125) that

$$
\begin{aligned}
|J_5|&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\\
&\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2).
\end{aligned}\tag{8.126}
$$

Coming back to (8.115) and using (8.117), (8.121), (8.122), (8.123), (8.124), (8.126), we have shown that

$$
|I_2|\le C(T-t)^{\frac{\delta-1-\gamma}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2).
$$

This estimate together with (8.116) ends the proof of (8.113).

Proof of (8.114)**.** It follows from (8.111) that

$$
\begin{aligned}
&\frac{\delta}{\delta m}U(t,\mu_1)(v)-\frac{\delta}{\delta m}U(t,\mu_2)(v)\\
&=\int_{\mathbb{R}^d}\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y)p(\mu_1,t,T,v,y)\,dy\\
&\quad+\int_{\mathbb{R}^{2d}}\left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y)-\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x)\right)\frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v)\,dy\,d\mu_1(x)\\
&\quad-\int_{\mathbb{R}^d}\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)p(\mu_2,t,T,v,y)\,dy\\
&\quad-\int_{\mathbb{R}^{2d}}\left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)-\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right)\frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v)\,dy\,d\mu_2(x).
\end{aligned}
$$

We decompose it in the following way

$$
\begin{aligned}
&\frac{\delta}{\delta m}U(t,\mu_1)(v) - \frac{\delta}{\delta m}U(t,\mu_2)(v) \\
&= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)\right) p(\mu_1,t,T,v,y)\,dy \\
&\quad + \int_{\mathbb{R}^d} \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)\,(p(\mu_1,t,T,v,y) - p(\mu_2,t,T,v,y))\,dy \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(x) - \left[\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right]\right) \\
&\qquad\qquad \frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v_2)\,dy\,d\mu_1(x) \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right) \\
&\qquad\qquad \left(\frac{\delta}{\delta m}p(\mu_1,t,T,x,y)(v_2) - \frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\right)\,dy\,d\mu_1(x) \\
&\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(x)\right) \frac{\delta}{\delta m}p(\mu_2,t,T,x,y)(v_2)\,dy\,d(\mu_1-\mu_2)(x) \\
&=: J_1 + J_2 + J_3 + J_4 + J_5.
\end{aligned} \tag{8.127}
$$

We first focus of J_1. Let us prove that there exists a positive constant C such that for all $t \in [0,T)$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $y \in \mathbb{R}^d$

$$
\left|\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y)\right| C(T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2). \tag{8.128}
$$

To prove this, we write

$$
\begin{aligned}
&\frac{\delta}{\delta m}\phi([X_T^{t,\mu_1}])(y) - \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) \\
&= \int_0^1 \int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z)(p(\mu_1,t,T,z) - p(\mu_2,t,T,z))\,dz\,d\lambda \\
&= \int_0^1 \int_{\mathbb{R}^{2d}} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z)(p(\mu_1,t,T,x,z) - p(\mu_2,t,T,x,z))\,d\mu_1(x)\,dz\,d\lambda \\
&\quad + \int_0^1 \int_{\mathbb{R}^{2d}} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z)p(\mu_2,t,T,x,z)\,d(\mu_1-\mu_2)(x)\,dz\,d\lambda \\
&=: K_1 + K_2,
\end{aligned}
$$

where $m_\lambda := \lambda[X_T^{t,\mu_1}] + (1-\lambda)[X_T^{t,\mu_2}]$. We rewrite K_1 as

$$
K_1 = \int_0^1 \int_{\mathbb{R}^{2d}} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m}p(M_r,t,T,x,z)(w)\,d(\mu_1-\mu_2)(w)\,dr\,d\mu_1(x)\,dz\,d\lambda,
$$

where $M_r = r\mu_1 + (1-r)\mu_2$. We set for $w \in \mathbb{R}^d$

$$
F(w) := \int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z)\frac{\delta}{\delta m}p(M_r,t,T,x,z)(w)\,dz,
$$

where λ, r and x are fixed. For $w_1, w_2 \in \mathbb{R}^d$, one has

$$
\begin{aligned}
&F(w_1) - F(w_2)\\
&= \int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) \int_0^1 \partial_v \frac{\delta}{\delta m} p(M_r,t,T,x,z)(sw_1 + (1-s)w_2)\cdot(w_1-w_2)\, ds\, dz\\
&= \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,x)\right) \partial_v \frac{\delta}{\delta m} p(M_r,t,T,x,z)(sw_1 + (1-s)w_2)\cdot(w_1-w_2)\, dz\, ds.
\end{aligned}
$$

It follows from the δ-Hölder continuity of $\frac{\delta^2}{\delta m^2}\phi(\mu)(y,\cdot)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$, (8.60) and the space-time inequality (8.231) that

$$
|F(w_1) - F(w_2)| \le C(T-t)^{\frac{\delta-1}{\alpha}+1+\frac{\eta-1}{\alpha}} |w_1 - w_2|.
$$

This yields

$$
\begin{aligned}
|K_1| &\le C(T-t)^{\frac{\delta-1}{\alpha}+1+\frac{\eta-1}{\alpha}} W_1(\mu_1,\mu_2)\\
&\le C(T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2). \qquad (8.129)
\end{aligned}
$$

We control K_2 as for K_1 by studying the regularity with respect to x of the function G given by

$$
G(x) := \int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) p(\mu_2,t,T,x,z)\, dz.
$$

For $x_1, x_2 \in \mathbb{R}^d$, one has

$$
\begin{aligned}
G(x_1) - G(x_2) &= \int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) \int_0^1 \partial_x p(\mu_2,t,T,rx_1+(1-r)x_2,z)\cdot(x_1-x_2)\, dr\, dz\\
&= \int_0^1 \int_{\mathbb{R}^d} \left(\frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,z) - \frac{\delta^2}{\delta m^2}\phi(m_\lambda)(y,rx_1+(1-r)x_2)\right)\\
&\qquad \partial_x p(\mu_2,t,T,rx_1+(1-r)x_2,z)\cdot(x_1-x_2)\, dz\, dr.
\end{aligned}
$$

Using δ-Hölder continuity of $\frac{\delta^2}{\delta m^2}\phi(\mu)(y,\cdot)$ uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$ and $y \in \mathbb{R}^d$, (8.57) and the space-time inequality (8.231), we obtain that

$$
|G(x_1) - G(x_2)| \le C(T-t)^{\frac{\delta-1}{\alpha}} |x_1 - x_2|.
$$

It proves that

$$
|K_2| \le C(T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2). \qquad (8.130)
$$

Combining (8.129) and (8.130) concludes the proof of (8.128). We can now turn to estimate J_1 in (8.127). Using (8.128), one gets

$$
\begin{aligned}
|J_1| &\le C\int_{\mathbb{R}^d} (T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2) p(\mu_2,t,T,v,y)\, dy\\
&\le C(T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2). \qquad (8.131)
\end{aligned}
$$

For J_2, we rewrite it as

$$J_2 = \int_{\mathbb{R}^d} \frac{\delta}{\delta m}\phi([X_T^{t,\mu_2}])(y) \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} p(M_r, t, T, v, y)(z)\, d(\mu_1 - \mu_2)(z)\, dr\, dy,$$

where $M_r := r\mu_1 + (1-r)\mu_2$. Following the same lines as for the proof of the above estimates on K_1 (8.129), we obtain that

$$|J_2| \le C(T-t)^{\frac{\delta-1}{\alpha}} W_1(\mu_1,\mu_2). \tag{8.132}$$

Following same lines as in the proof of (8.123), (8.124) and (8.126) by using (8.59), (8.66) and (8.63), we can prove the following estimates

$$|J_3| + |J_4| + |J_5| \le C(T-t)^{\frac{\delta-1}{\alpha}+1-\frac{1}{\alpha}} W_1(\mu_1,\mu_2). \tag{8.133}$$

Gathering (8.127) (8.131), (8.132) and (8.133), we have proved (8.114).

□

Proof of Theorem 8.12. Let us first introduce some notations. We can write for all $i \in \{1,\dots,N\}$ and for all $t \in [0,T]$

$$Z_t^i = \int_0^t \int_{\mathbb{R}^d} z\, \widetilde{\mathcal{N}}^i(ds,dz),$$

where $\mathcal{N}^i$ is the Poisson random measure associated with Z^i, and $\widetilde{\mathcal{N}}^i$ is the associated compensated Poisson random measure. Then, we set for all $t \in [0,T]$

$$\mathbf{Z}_t^N := \begin{pmatrix} Z_t^1 \\ \vdots \\ Z_t^N \end{pmatrix} \in (\mathbb{R}^d)^N.$$

As the Lévy processes $(Z^n)_n$ are independent, the process $(\mathbf{Z}_t^N)_t$ is a Lévy process in $(\mathbb{R}^d)^N$. Il is a cylindrical α-stable process. Its Poisson random measure $\boldsymbol{\mathcal{N}}^N$ and its Lévy measure $\boldsymbol{\nu}^N$ are defined as follows. For all $\varphi : [0,T] \times (\mathbb{R}^d)^N \to \mathbb{R}^+$, one has

$$\int_0^T \int_{\mathbb{R}^{Nd}} \varphi(s,\boldsymbol{x})\, \boldsymbol{\mathcal{N}}^N(ds,d\boldsymbol{x}) = \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \varphi(s,0,\dots,0,x_i,0,\dots,0)\, \mathcal{N}^i(ds,dx_i). \tag{8.134}$$

For all $\phi : \mathbb{R}^{Nd} \to \mathbb{R}^+$, one has

$$\int_{\mathbb{R}^{Nd}} \phi(\boldsymbol{x})\, d\boldsymbol{\nu}^N(\boldsymbol{x}) = \sum_{i=1}^N \int_{\mathbb{R}^d} \phi(0,\dots,0,x_i,0,\dots,0)\, d\nu(x_i). \tag{8.135}$$

Note that since $(Z^n)_n$ are independent processes, for all $t \in [0,T]$, the support of the random measure $\boldsymbol{\mathcal{N}}^N(t,d\boldsymbol{x})$ is contained in

$$\bigcup_{i=0}^{N-1} \{0\}_{\mathbb{R}^d}^i \times \mathbb{R}^d \times \{0\}_{\mathbb{R}^d}^{N-1-i} \subset (\mathbb{R}^d)^N.$$

Let us define for all $t \in [0,T]$, $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$

$$\boldsymbol{b}^N(t, \boldsymbol{x}) := \begin{pmatrix} b(t, x_1, \overline{\mu}^N_{\boldsymbol{x}}) \\ \vdots \\ b(t, x_N, \overline{\mu}^N_{\boldsymbol{x}}) \end{pmatrix} \in (\mathbb{R}^d)^N,$$

where $\overline{\mu}^N_{\boldsymbol{x}} = \frac{1}{N}\sum_{j=1}^N \delta_{x_j}$. Thus, writing for $t \in [0,T]$ and $N \geq 1$

$$\boldsymbol{X}^N_t = \begin{pmatrix} X^{1,N}_t \\ \vdots \\ X^{N,N}_t \end{pmatrix},$$

the SDE (8.32) defining the particle system can be rewritten as

$$\begin{cases} d\boldsymbol{X}^N_t = \boldsymbol{b}^N(t, \boldsymbol{X}^N_t)\,dt + d\mathbf{Z}^N_t, \quad t \in [0,T], \\ \boldsymbol{X}^N_0 = \begin{pmatrix} X^1_0 \\ \vdots \\ X^N_0 \end{pmatrix}. \end{cases} \tag{8.136}$$

Proof of (8.33)**.** Let us consider U the solution to the backward Kolmogorov PDE (8.31) with terminal condition ϕ at time T given in Theorem 8.9. Using Lemma 8.38 and the fact that $U \in \mathcal{C}^1([0,T) \times \mathcal{P}_\beta(\mathbb{R}^d))$, we obtain that the function $(t, \boldsymbol{x}) \in [0,T) \times (\mathbb{R}^d)^N \mapsto U(t, \overline{\mu}^N_{\boldsymbol{x}})$ belongs to $\mathcal{C}^1([0,T) \times (\mathbb{R}^d)^N)$. Moreover, for all $t \in [0,T)$ and $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$, we have by Proposition 8.38

$$\partial_{\boldsymbol{x}} U(t, \overline{\mu}^N_{\boldsymbol{x}}) = \frac{1}{N} \begin{pmatrix} \partial_v \frac{\delta}{\delta m} U(t, \overline{\mu}^N_{\boldsymbol{x}})(x_1) \\ \vdots \\ \partial_v \frac{\delta}{\delta m} U(t, \overline{\mu}^N_{\boldsymbol{x}})(x_N) \end{pmatrix}.$$

Let us fix $\gamma \in (0,1] \cap (\alpha - 1, (\delta + \alpha - 1) \wedge (2\alpha - 2) \wedge (\eta + \alpha - 1))$. We easily see using (8.113) that the map $\boldsymbol{x} \mapsto \partial_{\boldsymbol{x}} U(t, \overline{\mu}^N_{\boldsymbol{x}})$ is γ-Hölder continuous locally uniformly with respect to $t \in [0,T)$. This comes from the fact that for all $\boldsymbol{x} = (x_1, \dots, x_N), \boldsymbol{y} = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N$,

$$W_1^\gamma(\overline{\mu}^N_{\boldsymbol{x}}, \overline{\mu}^N_{\boldsymbol{y}}) \leq \frac{1}{N^\gamma} \sum_{k=1}^N |x_k - y_k|^\gamma.$$

We denote by $\boldsymbol{\mathcal{N}}^N$ the Poisson random measure associated with $\boldsymbol{Z}^N = (Z^1, \dots, Z^N)$ defined in (8.134) and by $\boldsymbol{\nu}^N$ its associated Lévy measure defined in (8.135). Since $\gamma > \alpha - 1$, we can apply the standard Itô formula for this function and the $(\mathbb{R}^d)^N$-valued process $(\boldsymbol{X}^N_t)_t$. Noticing that $t \in [0,T] \mapsto U(t, \mu_t)$ is constant by the definition of U and the well-posedness of the McKean-Vlasov SDE, we obtain that for all $t \in [0,T)$

$$
\begin{aligned}
&U(t,\overline{\mu}^N_t) - U(t,\mu_t) - \Big(U(0,\overline{\mu}^N_0) - U(0,\mu_0)\Big) \\
&= \int_0^t \partial_t U(s,\overline{\mu}^N_s)\,ds \\
&\quad + \frac{1}{N}\sum_{i=1}^N \int_0^t \partial_v \frac{\delta}{\delta m} U(s,\overline{\mu}^N_s)(X^{i,N}_s)\cdot b(s,X^{i,N}_s,\overline{\mu}^N_s)\,ds \\
&\quad + \int_0^t \int_{(\mathbb{R}^d)^N} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) - \partial_{\boldsymbol{x}} U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\cdot \boldsymbol{z}\Big]\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \qquad (8.137)\\
&\quad + \int_0^t \int_{\{|\boldsymbol{z}|\geq 1\}} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&\quad + \int_0^t \int_{\{|\boldsymbol{z}|< 1\}} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big]\,\widetilde{\boldsymbol{\mathcal{N}}}^N(ds,d\boldsymbol{z}) \\
&=: I_1 + I_2 + I_3 + I_4 + I_5.
\end{aligned}
$$

Note that the term

$$
\int_0^t \int_{\{|\boldsymbol{z}|\geq 1\}} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big]\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds
$$

is well-defined. Indeed, if we set, for $h \in \mathbb{R}^d$, $\tilde{\boldsymbol{h}}_{\boldsymbol{j}} := (0,\ldots,0,h,0,\ldots,0) \in (\mathbb{R}^d)^N$, where h appears in the j-th coordinate, one has using (8.112)

$$
\begin{aligned}
&\int_0^t \int_{\{|\boldsymbol{z}|\geq 1\}} \Big|U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big|\,d\boldsymbol{\nu}^N(\boldsymbol{z})\,ds \\
&= \sum_{i=1}^N \int_0^t \int_{B_1^c} \Big|U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})\Big|\,d\nu(z)\,ds \\
&\leq \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1^c}\int_0^1 \Big|\partial_v \frac{\delta}{\delta m} U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+hz)\Big|\,|z|\,dw\,d\nu(z)\,ds \\
&\leq C\int_0^t (T-s)^{\frac{\delta-1}{\alpha}}\,ds \int_{B_1^c} |z|\,d\nu(z),
\end{aligned}
$$

where $m^i_{s,z,w} := w\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{N,s^-}}$. We conclude since $\alpha \in (1,2)$.

By (8.135), we write

$$
\begin{aligned}
I_3 &= \sum_{i=1}^N \int_0^t \int_{(\mathbb{R}^d)^N} \Big[U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}) - U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \\
&\qquad\qquad - \frac{1}{N}\partial_v \frac{\delta}{\delta m} U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-})\cdot z\Big]\,d\nu(z)\,ds \\
&= \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d}\int_0^1 \Big[\frac{\delta}{\delta m} U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) \\
&\qquad\qquad - \frac{\delta}{\delta m} U(s,m^i_{s,z,w})(X^{i,N}_{s^-}) - \partial_v \frac{\delta}{\delta m} U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-})\cdot z\Big]\,dw\,d\nu(z)\,ds,
\end{aligned}
$$

where $m^i_{s,z,w} = w\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}$. In order to make appear the backward Kolmogorov PDE

(8.31), we decompose I_3 in the following way

$$\begin{aligned}
I_3 &= \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x+z) - \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x) - \partial_v \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(x) \cdot z \right] d\nu(z)\, d\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}(x)\, ds \\
&\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \int_0^1 \left[\frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-} + z) - \frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-}) \right. \\
&\qquad\qquad \left. + \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-} + z) \right] dw\, d\nu(z)\, ds \\
&=: I_{3,A} + I_{3,B}.
\end{aligned}$$

Since $(\boldsymbol{X}^N_s)_{s\in[0,T]}$ is càd-làg, we deduce that almost surely for almost all $s \in [0,t]$ we have $\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}} = \overline{\mu}^N_{\boldsymbol{X}^N_s} = \overline{\mu}^N_s$. Owing to the backward Kolmogorov PDE (8.31) in Theorem 8.9, one has

$$\begin{aligned}
I_1 + I_2 + I_{3,A} &= \int_0^t \partial_s U(s, \overline{\mu}^N_{N,s}) + \mathscr{L}_s U(s,\cdot)(\overline{\mu}^N_s)\, ds \\
&= 0.
\end{aligned}$$

Thus, we obtain the following decomposition, for any $t \in [0,T)$,

$$\begin{aligned}
&U(t, \overline{\mu}^N_t) - U(t, \mu_t) - \Big(U(0, \overline{\mu}^N_0) - U(0, \mu_0) \Big) \\
&= \frac{1}{N} \sum_{i=1}^N \int_0^t \int_{\mathbb{R}^d} \int_0^1 \left[\frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-} + z) - \frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-}) \right. \\
&\qquad\qquad \left. + \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m} U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-} + z) \right] dw\, d\nu(z)\, ds \\
&\quad + \int_0^t \int_{\{|\boldsymbol{z}|\geq 1\}} \Big[U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \Big] \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \qquad (8.138) \\
&\quad + \int_0^t \int_{\{|\boldsymbol{z}|< 1\}} \Big[U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \Big] \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \\
&= I_{3,B} + I_4 + I_5.
\end{aligned}$$

It follows that for all $t \in [0,T)$

$$\mathbb{E}|U(t, \overline{\mu}^N_t) - U(t, \mu_t)| \leq \mathbb{E}|U(0, \overline{\mu}^N_0) - U(0, \mu_0)| + \mathbb{E}(|I_{3,B}| + |I_4| + |I_5|). \qquad (8.139)$$

We now treat each term separately. For $I_{3,B}$, one has

$$
\begin{aligned}
\mathbb{E}|I_{3,B}| \leq{}& \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1}\int_0^1 \Big| \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}) \\
&\qquad + \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}+z)\Big| \, dw\, d\nu(z)\, ds \\
&+ \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1^c}\int_0^1 \Big| \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}) \\
&\qquad + \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}+z)\Big| \, dw\, d\nu(z)\, ds \\
\leq{}& \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1}\int_0^1\int_0^1 \Big| \partial_v \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+\lambda z) \\
&\qquad - \partial_v \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}+\lambda z)\Big| \, |z|\, d\lambda\, dw\, d\nu(z)\, ds \\
&+ \frac{1}{N}\sum_{i=1}^N \int_0^t \int_{B_1^c}\int_0^1 \Big| \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-}+z) - \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}+z)\Big| \\
&\qquad + \Big| \frac{\delta}{\delta m}U(s,\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}})(X^{i,N}_{s^-}) - \frac{\delta}{\delta m}U(s,m^i_{s,z,w})(X^{i,N}_{s^-})\Big| \, dw\, d\nu(z)\, ds.
\end{aligned}
$$

Recall that $m^i_{s,z,w} = w\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}$. Moreover, one has

$$
\begin{aligned}
W_1(m^i_{s,z,w}, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) &\leq W_1(\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \\
&\leq \frac{|z|}{N}.
\end{aligned}
$$

Then, by (8.113) with $\gamma \in (0,1] \cap (\alpha-1, (\delta+\alpha-1)\wedge(2\alpha-2)\wedge(\eta+\alpha-1))$ and (8.114), we deduce that

$$
\begin{aligned}
\mathbb{E}|I_{3,B}| \leq{}& \frac{C}{N}\sum_{i=1}^N \int_0^t \int_{B_1} (T-s)^{\frac{\delta-1-\gamma}{\alpha}} \frac{|z|^{1+\gamma}}{N^\gamma}\, d\nu(z)\, ds \\
&+ \frac{C}{N}\sum_{i=1}^N \int_0^t \int_{B_1^c} (T-s)^{\frac{\delta-1}{\alpha}} \frac{|z|}{N} d\nu(z)\, ds.
\end{aligned}
$$

Since $\gamma < \delta+\alpha-1$, we have $\frac{\delta-1-\gamma}{\alpha} > -1$ and therefore the map $s \in [0,T) \mapsto (T-s)^{\frac{\delta-1-\gamma}{\alpha}}$ is integrable and $z \in B_1 \mapsto |z|^{1+\gamma} \in L^1(B_1,\nu)$. It yields

$$
\mathbb{E}|I_{3,B}| \leq \frac{C}{N^\gamma}. \tag{8.140}
$$

Let us now focus on I_4. Recall that $m^i_{s,z,w} = w\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}$. We introduce $\Gamma \in (1,2]$ such that $\Gamma > \alpha$ and $\delta > 1-\alpha/\Gamma$, which is possible since $\delta > 0$. Thanks to BDG's inequalities and the subadditivity of the map $|\cdot|^{\Gamma/2}$, there exists a constant $C>0$ independent of N such that we have for all $t \in (0, T]$

$$
\begin{aligned}
\mathbb{E}|I_4| &\leq \left(\mathbb{E}\left| \int_0^t \int_{\{|\boldsymbol{z}|<1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right] \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \right|^\Gamma \right)^{\frac{1}{\Gamma}} \\
&\leq C \left(\mathbb{E}\left[\int_0^t \int_{\{|\boldsymbol{z}|<1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right|^2 \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z}) \right]^{\frac{\Gamma}{2}} \right)^{\frac{1}{\Gamma}} \\
&\leq C \left(\mathbb{E} \int_0^t \int_{\{|\boldsymbol{z}|<1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right|^\Gamma \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z}) \right)^{\frac{1}{\Gamma}} \qquad (8.141) \\
&= C \left(\sum_{i=1}^N \mathbb{E} \int_0^t \int_{B_1} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}_i}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right|^\Gamma d\nu(z)\, ds \right)^{\frac{1}{\Gamma}} \qquad (8.142) \\
&= C \left(\sum_{i=1}^N \mathbb{E} \int_0^t \int_{B_1} \left| \frac{1}{N} \int_{[0,1]^2} \partial_v \frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-} + hz) \cdot z\, dh\, dw \right|^\Gamma d\nu(z)\, ds \right)^{\frac{1}{\Gamma}} \\
&\leq C \left(\sum_{i=1}^N \mathbb{E} \int_0^t \int_{B_1} \frac{1}{N^\Gamma} (T-s)^{\frac{\Gamma(\delta-1)}{\alpha}} |z|^\Gamma d\nu(z)\, ds \right)^{\frac{1}{\Gamma}} \\
&\leq \frac{C}{N^{1-\frac{1}{\Gamma}}}.
\end{aligned}
$$

Indeed, the time integral is finite since $\Gamma(\delta-1)/\alpha > -1$ since $\delta > 1 - \alpha/\Gamma$ by choice of Γ and $z \in B_1 \mapsto |z|^\Gamma$ is integrable with respect to ν since $\gamma > \alpha$. Finally for I_5, BDG's inequalities and the fact that $1 < \beta < 2$ yields, for all $t \in [0, T)$,

$$
\begin{aligned}
\mathbb{E}|I_5| &\leq \left(\mathbb{E}\left(\int_0^t \int_{\{|\boldsymbol{z}|>1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right] \widetilde{\boldsymbol{\mathcal{N}}}^N(ds, d\boldsymbol{z}) \right)^\beta \right)^{\frac{1}{\beta}} \\
&\leq C \left(\mathbb{E}\left[\int_0^t \int_{\{|\boldsymbol{z}|>1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right|^2 \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z}) \right]^{\frac{\beta}{2}} \right)^{\frac{1}{\beta}} \\
&\leq C \left(\mathbb{E} \int_0^t \int_{\{|\boldsymbol{z}|>1\}} \left| U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}+\boldsymbol{z}}) - U(s, \overline{\mu}^N_{\boldsymbol{X}^N_{s^-}}) \right|^\beta \boldsymbol{\mathcal{N}}^N(ds, d\boldsymbol{z}) \right)^{\frac{1}{\beta}} \qquad (8.143) \\
&= C \left(\sum_{i=1}^N \mathbb{E} \int_0^t \int_{B_1^c} \left| \frac{1}{N} \int_{[0,1]^2} \partial_v \frac{\delta}{\delta m} U(s, m^i_{s,z,w})(X^{i,N}_{s^-} + hz) \cdot z\, dh\, dw \right|^\beta d\nu(z)\, ds \right)^{\frac{1}{\beta}} \\
&\leq C \left(\sum_{i=1}^N \mathbb{E} \int_0^t \int_{B_1^c} \left| \frac{1}{N} (T-s)^{\frac{\delta-1}{\alpha}} \right|^\beta |z|^\beta d\nu(z)\, ds \right)^{\frac{1}{\beta}} \\
&\leq \frac{C}{N^{1-\frac{1}{\beta}}}.
\end{aligned}
$$

Indeed, the time integral is finite since $1 < \beta < \alpha$ and thus $\beta(\delta-1)/\alpha > -1$, and the map $z \in B_1^c \mapsto |z|^\beta$ is integrable with respect to ν since $\beta < \alpha$. As a consequence of (8.138), (8.140), (8.141), (8.143), and the fact that $\gamma > 1 - \frac{1}{\beta}$ since $1 < \beta < \alpha < 1 + \gamma$ and $1 - \frac{1}{\Gamma} > 1 - \frac{1}{\beta}$ since $\beta < \Gamma$, we have, for a constant $C > 0$, that for all $N \geq 1$ and $t \in [0,T)$

$$\mathbb{E}|U(t,\overline{\mu}_t^N) - U(t,\mu_t)| \leq \mathbb{E}|U(0,\overline{\mu}_0^N) - U(0,\mu_0)| + \frac{C}{N^{1-\frac{1}{\beta}}}. \tag{8.144}$$

It follows from (8.112) that the map $\frac{\delta}{\delta m}U(0,\mu)(\cdot)$ is Lipschitz uniformly with respect to $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$. The Kantorovich-Rubinstein theorem ensures that

$$\mathbb{E}|U(0,\overline{\mu}_0^N) - U(0,\mu_0)| \leq C\mathbb{E}W_1(\overline{\mu}_0^N,\mu_0).$$

By Fatou's lemma, the continuity of U on $[0,T] \times \mathcal{P}_\beta(\mathbb{R}^d)$ and since $\overline{\mu}_{T-}^N = \overline{\mu}_T^N$ almost surely, we can let t tend to T in (8.144), which concludes the proof of (8.33).

Proof of (8.34)**.** Coming back to (8.138) and using (8.141) and (8.143), we see that I_4 an I_5 are centered random variables because of the martingale property of compensated Poisson random integrals. We thus obtain that for any $t \in [0,T)$

$$|\mathbb{E}(U(t,\overline{\mu}_t^N) - U(t,\mu_t))| \leq |\mathbb{E}(U(0,\overline{\mu}_0^N) - U(0,\mu_0))| + \mathbb{E}|I_{3,B}|.$$

The control of $\mathbb{E}|I_{3,B}|$ has already been done in (8.140). It follows that there exists a positive constant C such that for all $t \in [0,T)$

$$|\mathbb{E}(U(t,\overline{\mu}_t^N) - U(t,\mu_t))| \leq C\,|\mathbb{E}(U(0,\overline{\mu}_0^N) - U(0,\mu_0))| + \frac{C}{N^\gamma}. \tag{8.145}$$

Reasoning as in [CdRF21], after (5.24), we can write using the exchangeability in law of the initial data $(\xi^i)_{i\in\{1,\dots,N\}}$

$$\mathbb{E}(U(0,\overline{\mu}_0^N) - U(0,\mu_0)) = \int_0^1 \mathbb{E}\left(\frac{\delta}{\delta m}U(0,\tilde{\mu}_0^{N,\lambda_1})(\tilde{\xi}) - \frac{\delta}{\delta m}U(0,\mu_0^{N,\lambda_1})(\tilde{\xi})\right) d\lambda_1,$$

where $\mu_0^{N,\lambda_1} := \lambda_1\overline{\mu}_0^N + (1-\lambda_1)\mu_0$, $\tilde{\mu}_0^{N,\lambda_1} := \lambda_1\tilde{\mu}_0^N + (1-\lambda_1)\mu_0$, $\tilde{\mu}_0^N := \overline{\mu}_0^N + \frac{1}{N}(\delta_{\tilde{\xi}} - \delta_{\xi^1})$, $(\xi^i)_i$ being an i.i.d. sequence of random variable with common distribution μ_0 and $\tilde{\xi}$ a random variable independent of $(\xi^i)_i$ with distribution μ_0. This ensures, using (8.114), that there exists a positive constant C such that for all $\phi \in \mathcal{C}^{(2,\delta)}(\mathcal{P}_\beta(\mathbb{R}^d))$ and $N \geq 1$

$$\begin{aligned}
|\mathbb{E}(U(0,\overline{\mu}_0^N) - U(0,\mu_0))| &\leq \int_0^1 C\mathbb{E}W_1(\tilde{\mu}_0^{N,\lambda_1},\overline{\mu}_0^{N,\lambda_1})\,d\lambda_1 \\
&\leq \int_0^1 C\mathbb{E}W_1(\tilde{\mu}_0^N,\overline{\mu}_0^N)\,d\lambda_1 \\
&\leq \int_0^1 \frac{C}{N}\mathbb{E}|\tilde{\xi} - \xi_1|\,d\lambda_1 \\
&\leq \frac{C}{N},
\end{aligned}$$

the constant C depending on $M_1(\mu_0)$ in (8.34). We conclude by letting t tend to T in (8.145) that (8.34) holds true.

□

8.7 Quantitative approximation of the distribution of one particle by the limiting McKean-Vlasov process at the level of densities

This section is devoted to the proof of Theorem 8.16. We start to show that the density bound (8.38) holds. Let us fix $t \in (0, T]$. As in the proof of Theorem 8.12, we easily see using Theorem 8.18 that we can apply Itô's formula for the function $(s, \boldsymbol{x}) \in [0, t) \times (\mathbb{R}^d)^N \mapsto p(\overline{\mu}^N_{\boldsymbol{x}}, s, t, y) = \frac{1}{N}\sum_{k=1}^N p(\overline{\mu}^N_{\boldsymbol{x}}, s, t, x_k, y)$. Keeping the same notations as in the proof of Theorem 8.12, we get that

$$\begin{aligned}
p(\overline{\mu}^N_s, s, t, y) - p(\overline{\mu}^N_0, 0, t, y) = {}& \int_0^s \partial_r p(\overline{\mu}^N_r, r, t, y)\, dr \\
&+ \frac{1}{N}\sum_{i=1}^N \int_0^s \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_r, r, t, y)(X^{i,N}_{r^-}) \cdot b(r, X^{i,N}_{r^-}, \overline{\mu}^N_r)\, dr \\
&+ \sum_{i=1}^N \int_0^s \int_{(\mathbb{R}^d)^N} \Big[p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}}, r, t, y) - p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y) \\
&\qquad - \frac{1}{N}\partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-}) \cdot z \Big]\, d\nu(z)\, dr \\
&+ \int_0^s \int_{(\mathbb{R}^d)^N} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}+\boldsymbol{z}}, r, t, y) - p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)\, \widetilde{\boldsymbol{\mathcal{N}}}^N(dr, d\boldsymbol{z}) \\
=: {}& I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{8.146}$$

First, notice that reasoning as in the proof of Theorem 8.12, by using (8.66) for the big jumps, and (8.60) for the small jumps, we prove that I_4 is a true martingale and that

$$\mathbb{E} I_4 = 0. \tag{8.147}$$

By definition of the linear derivative, we can write

$$\begin{aligned}
I_3 = \frac{1}{N}\sum_{i=1}^N \int_0^s \int_{\mathbb{R}^d} \int_0^1 \Big[& \frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-} + z) \\
& - \frac{\delta}{\delta m} p(m^i_{r,z,w}, r, t, y)(X^{i,N}_{r^-}) - \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}, r, t, y)(X^{i,N}_{r^-}) \cdot z \Big]\, dw\, d\nu(z)\, dr,
\end{aligned}$$

where $m^i_{r,z,w} := w\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}+\tilde{\boldsymbol{z}}_{\boldsymbol{i}}} + (1-w)\overline{\mu}^N_{\boldsymbol{X}^N_{r^-}}$. We can now decompose I_3 in the following way

$$
\begin{aligned}
I_3 = \int_0^s \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Big[&\frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(x+z) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(x) \\
&\qquad -\partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(x)\cdot z\Big]\, d\nu(z)\, d\overline{\mu}^N_{\mathbf{X}^N_{r^-}}(x)\, dr \\
+ \frac{1}{N}\sum_{i=1}^N \int_0^s \int_{B_1^c}\int_0^1 \Big[&\frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(X^{i,N}_{r^-}+z) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(X^{i,N}_{r^-}+z)\Big] \\
&- \Big[\frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(X^{i,N}_{r^-}) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(X^{i,N}_{r^-})\Big]\, dw\, d\nu(z)\, dr \\
+ \frac{1}{N}\sum_{i=1}^N \int_0^s \int_{B_1}\int_0^1 \Big[&\frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(X^{i,N}_{r^-}+z) - \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(X^{i,N}_{r^-}) \\
&+ \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(X^{i,N}_{r^-}) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(X^{i,N}_{r^-}+z)\Big]\, dw\, d\nu(z)\, dr \\
=: I_{3,A} + I_{3,B} + I_{3,C}.
\end{aligned}
$$

By taking the expectation in (8.146), we deduce thanks to (8.147) and the backward Kolmogorov PDE (8.26) satisfied by p (see Theorem 8.7) that

$$
\mathbb{E}p(\overline{\mu}^N_s, s,t,y) = \mathbb{E}p(\overline{\mu}^N_0, 0,t,y) + \mathbb{E}I_{3,B} + \mathbb{E}I_{3,C}. \tag{8.148}
$$

We first aim at controlling $I_{3,B}$. To do this, we write for $v \in \mathbb{R}^d$

$$
\begin{aligned}
\frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(v) &- \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(v) \\
&= p(m^i_{r,z,w}, r,t,v,y) - p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,v,y) \\
&\quad + \int_{\mathbb{R}^d} \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,x,y)(v)\, w d(\overline{\mu}^N_{\mathbf{X}^N_{r^-}+\tilde{\boldsymbol{z}}_i} - \overline{\mu}^N_{\mathbf{X}^N_{r^-}})(x) \\
&\quad + \int_{\mathbb{R}^d} \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,x,y)(v) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,x,y)(v)\, d\overline{\mu}^N_{\mathbf{X}^N_{r^-}}(x).
\end{aligned}
$$

Using (8.65), (8.63) and (8.66) with $\gamma = 1$, we obtain

$$
\begin{aligned}
&\left|\frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(v) - \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(v)\right| \\
&\le \frac{C}{N}(t-r)^{-\zeta}|z|\left[\rho^0(t-r,y-v) + \rho^0(t-r,y-X^{i,N}_{r^-}-z) + \rho^0(t-r,y-X^{i,N}_{r^-}) + \frac{1}{N}\sum_{k=1}^N \rho^0(t-r,y-X^{k,N}_{r^-})\right],
\end{aligned}
$$

where we recall that $\zeta := -\left(1-\frac{2+\gamma}{\alpha}\right)$. It follows that

$$
\mathbb{E}|I_{3,B}| \le \frac{C}{N}\int_0^s \int_{B_1^c} (t-r)^{-\zeta}|z|\mathbb{E}(\rho^0(t-r,y-X^{i,N}_{r^-}) + \rho^0(t-r,y-X^{i,N}_{r^-}-z))\, d\nu(z)\, dr. \tag{8.149}
$$

Let us now focus on $I_{3,C}$ which can be written in the following way

$$I_{3,C} = \frac{1}{N}\sum_{i=1}^N \int_0^s \int_{B_1}\int_0^1\int_0^1 \Big[\partial_v \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(X^{i,N}_{r^-}+\lambda z) \\ -\partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(X^{i,N}_{r^-}+\lambda z)\Big]\cdot z\, d\lambda\, dw\, d\nu(z)\, dr.$$

As previously, we need to control for $v\in\mathbb{R}^d$

$$\begin{aligned}\partial_v \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(v) - \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(v)& \\ = \partial_x p(m^i_{r,z,w}, r,t,v,y) - \partial_x p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,v,y)& \\ + \int_{\mathbb{R}^d} \partial_v \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,x,y)(v)\, w d(\overline{\mu}^N_{\mathbf{X}^N_{r^-}+\tilde{\boldsymbol{z}}_i} - \overline{\mu}^N_{\mathbf{X}^N_{r^-}})(x)& \\ + \int_{\mathbb{R}^d} \partial_v \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,x,y)(v) - \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,x,y)(v)\, d\overline{\mu}^N_{\mathbf{X}^N_{r^-}}(x).&\end{aligned}$$

Using (8.65), (8.64) and (8.67) with γ, we obtain

$$\left|\partial_v \frac{\delta}{\delta m} p(m^i_{r,z,w}, r,t,y)(v) - \partial_v \frac{\delta}{\delta m} p(\overline{\mu}^N_{\mathbf{X}^N_{r^-}}, r,t,y)(v)\right| \le \frac{C}{N^\gamma}(t-r)^{1-\frac{2+\gamma}{\alpha}}|z|^\gamma \\ \left[\rho^0(t-r,y-v)+\rho^0(t-r,y-X^{i,N}_{r^-}-z)+\rho^0(t-r,y-X^{i,N}_{r^-})+\frac{1}{N}\sum_{k=1}^N \rho^0(t-r,y-X^{k,N}_{r^-})\right].$$

This yields, by definition of ζ,

$$\mathbb{E}|I_{3,C}| \le \frac{C}{N^\gamma}\int_0^s\int_{B_1}(t-r)^{-\zeta}|z|^{1+\gamma}\mathbb{E}(\rho^0(t-r,y-X^{i,N}_{r^-})+\rho^0(t-r,y-X^{i,N}_{r^-}-z))\, d\nu(z)\, dr. \tag{8.150}$$

Gathering (8.148), (8.149) and (8.150), we deduce that for any $s\in[0,t)$, $\mu_0\in\mathcal{P}_\beta(\mathbb{R}^d)$, $y\in\mathbb{R}^d$ and $N\ge 1$

$$\begin{aligned}\mathbb{E}p(\overline{\mu}^N_s, s,t,y) &\le \mathbb{E}p(\overline{\mu}^N_0, 0,t,y) \\ &\quad + \frac{C}{N^\gamma}\int_0^s\int_{\mathbb{R}^d}(t-r)^{-\zeta}f(z)\mathbb{E}(\rho^0(t-r,y-X^{i,N}_{r^-})+\rho^0(t-r,y-X^{i,N}_{r^-}-z))\, d\nu(z)\, dr.\end{aligned} \tag{8.151}$$

By (8.57), one has

$$\begin{aligned}\mathbb{E}p(\overline{\mu}^N_0, 0,t,y) &= \mathbb{E}\frac{1}{N}\sum_{i=1}^N\Big(p(\overline{\mu}^N_0,0,t,\xi^i,y)\Big) \\ &\le Cq_0(\mu_0,0,t,y),\end{aligned} \tag{8.152}$$

where we recall that

$$q_0(\mu_0,0,t,y) := \int_{\mathbb{R}^d}\rho^0(t,y-x)\, d\mu_0(x).$$

The same reasoning as in the proof of [CdRF21, Theorem 3.5] (see (5.13) and below) shows that

$$\lim_{s\to t^-} \mathbb{E}p(\overline{\mu}_s^N, s, t, y) = p^{1,N}(\mu_0, 0, t, y). \tag{8.153}$$

This is justified by the parametrix expansion of the transition density p of the McKean-Vlasov SDE (see (8.76)). Indeed, it allows to write, for $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$, $x, y \in \mathbb{R}^d$, $p(\mu_0, s, t, x, y) = \widehat{p}(s, t, x, y) + R(\mu_0, s, t, x, y)$ with

$$|R(\mu_0, s, t, x, y)| \le C(t-s)^{-\frac{1}{\alpha}} \rho^1(t-s, y-x).$$

Using (8.151), (8.152) and (8.153), we deduce that for any $t \in (0, T]$, $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $y \in \mathbb{R}^d$ and $N \ge 1$

$$\begin{aligned} p^{1,N}(\mu_0, 0, t, y) \le C q_0(\mu_0, 0, t, y) + \frac{C}{N^\gamma} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (t-r)^{-\zeta} f(z) [\rho^0(t-r, y-w-z) \\ + \rho^0(t-r, y-w)] p^{1,N}(\mu_0, 0, r, w) \, dw \, d\nu(z) \, dr. \end{aligned}$$

Denoting, for $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$ and $y \in \mathbb{R}^d$

$$g(\mu, s, t, y) := (t-s)^{-\zeta} \rho^0(t-s, y), \tag{8.154}$$

the previous inequality can be rewritten as

$$p^{1,N}(\mu_0, 0, t, y) \le C q_0(\mu_0, 0, t, y) + \frac{C}{N^\gamma} \int_{\mathbb{R}^d} [p^{1,N} \otimes g(\mu_0, 0, t, y-z) + p^{1,N} \otimes g(\mu_0, 0, t, y)] f(z) \, d\nu(z),$$

where we recall that

$$p^{1,N} \otimes g(\mu, 0, t, y) := \int_0^t \int_{\mathbb{R}^d} p^{1,N}(\mu, 0, r, w) g(\mu, r, t, y-w) \, dw \, dr.$$

Notice that g yields a time-integrable singularity since $\zeta \in (0,1)$. Then, we can easily prove by induction that for any $M \ge 1$

$$\begin{aligned} p^{1,N}(\mu_0, 0, t, y) \le C q_0(\mu_0, 0, t, y) + \sum_{k=1}^M \frac{C^{k+1}}{N^{k\gamma}} \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0 \otimes g^k \left(\mu_0, 0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^k f(z_j) \, d\nu(z_j) \\ + \frac{C^{M+1}}{N^{(M+1)\gamma}} \sum_{I \in P_{M+1}} \int_{(\mathbb{R}^d)^{M+1}} p^{1,N} \otimes g^{M+1} \left(\mu_0, 0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^{M+1} f(z_j) \, d\nu(z_j), \end{aligned} \tag{8.155}$$

where P_k is the set of all subsets of $\{1, \dots, k\}$ and g^k is defined by the recursive relation $g^{k+1} = g^k \otimes g$. As in the proof of (8.241), we prove by induction that for any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$ and $y \in \mathbb{R}^d$

$$|g^k(\mu, s, t, y)| \le C^{k-1} (t-s)^{-\zeta + (k-1)(1-\zeta)} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta), 1-\zeta) \right) \rho^0(t-s, y). \tag{8.156}$$

Thus, sending M to infinity in (8.155) and using (8.156), we deduce that

$$p^{1,N}(\mu_0,0,t,y) \leq Cq_0(\mu_0,0,t,y) + \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} \sum_{I\in P_k} \int_{(\mathbb{R}^d)^k} q_0 \otimes g^k \left(\mu_0,0,t,y-\sum_{i\in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j).$$

By injecting (8.156) in the previous inequality and using the convolution inequality (8.234), we get

$$p^{1,N}(\mu_0,0,t,y) \leq Cq_0(\mu_0,0,t,y) + \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta),1-\zeta)\right) \int_0^t (t-r)^{-\zeta+(k-1)(1-\zeta)}\, dr$$
$$\sum_{I\in P_k} \int_{(\mathbb{R}^d)^k} q_0 \left(\mu_0,0,t,y-\sum_{i\in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j).$$

This concludes the proof of (8.38).

Before proving (8.39), we state and show the following Lemma.

Lemma 8.25. *For any $\gamma \in [\alpha-1,1]$, there exists a constant $C>0$ such that for any $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $t\in(0,T]$, $y\in\mathbb{R}^d$ and $N\geq 1$*

$$\left|\mathbb{E}p(\overline{\mu}_0^N,0,t,y) - p(\mu_0,0,t,y)\right| \leq \frac{C}{N^\gamma} t^{1-\frac{1+\gamma}{\alpha}} (1+M_\gamma(\mu_0)) \int_{\mathbb{R}^d} (1+|x|^\gamma)\rho^0(t,y-x)\, d\mu_0(x).$$

Proof of Lemma 8.25. Recall that $(\xi^i)_{i\geq1}$ is an i.i.d. sequence with common distribution $\mu_0 \in \mathcal{P}_\beta(\mathbb{R}^d)$ and let us introduce $\tilde{\xi}$ a random variable with distribution μ_0 independent of $(\xi^i)_{i\geq1}$. Reasoning as in Lemma 5.1 in [CdRF21] (see (5.4) and below), we obtain that

$$\begin{aligned}
\mathbb{E}p(\overline{\mu}_0^N,0,t,y) - p(\mu_0,0,t,y) &= \frac{1}{N}\int_0^1 \mathbb{E}\left(\frac{\delta}{\delta m}p(\mu_0^{\lambda_1,N},0,t,\xi^1,y)(\xi^1) - \frac{\delta}{\delta m}p(\mu_0^{\lambda_1,N},0,t,\xi^1,y)(\tilde{\xi})\right) d\lambda_1 \\
&\quad + \frac{N-1}{N}\int_0^1 \mathbb{E}\left(\frac{\delta}{\delta m}p(\tilde{\mu}_0^{\lambda_1,N},0,t,\xi^1,y)(\tilde{\xi}) - \frac{\delta}{\delta m}p(\mu_0^{\lambda_1,N},0,t,\xi^1,y)(\tilde{\xi})\right) d\lambda_1 \\
&=: I_1 + I_2,
\end{aligned}$$

where $\mu_0^{\lambda_1,N} := \lambda_1\overline{\mu}_0^N + (1-\lambda_1)\mu_0$ and $\tilde{\mu}_0^{\lambda_1,N} := \lambda_1\tilde{\mu}_0^N + (1-\lambda_1)\mu_0$ with $\tilde{\mu}_0^N := \overline{\mu}_0^N + \frac{1}{N}(\delta_{\tilde{\xi}} - \delta_{\xi^2})$. Using (8.62), we obtain that for some constant $C>0$ depending on γ

$$\begin{aligned}
|I_1| &\leq \frac{C}{N}\int_0^1 t^{1-\frac{1+\gamma}{\alpha}} \mathbb{E}(|\xi^1-\tilde{\xi}|^\gamma \rho^0(t,y-\xi^1))\, d\lambda_1 \\
&\leq \frac{C}{N} t^{1-\frac{1+\gamma}{\alpha}} \left[\mathbb{E}(|\xi^1|^\gamma\rho^0(t,y-\xi^1)) + \mathbb{E}(|\tilde{\xi}|^\gamma\rho^0(t,y-\xi^1))\right] \\
&\leq \frac{C}{N} t^{1-\frac{1+\gamma}{\alpha}} (1+M_\gamma(\mu_0)) \int_{\mathbb{R}^d} (1+|x|^\gamma)\rho^0(t,y-x)\, d\mu_0(x).
\end{aligned}$$

Using now (8.101), one has

$$|I_2| \leq C\frac{N-1}{N} t^{1-\frac{1+\gamma}{\alpha}} \int_0^1 \mathbb{E}(W_1^\gamma(\tilde{\mu}_0^{\lambda_1,N},\mu_0^{\lambda_1,N})\rho^0(t,y-\xi^1))\, d\lambda_1.$$

Since we have

$$\begin{aligned} W_1(\tilde{\mu}_0^{\lambda_1,N}, \mu_0^{\lambda_1,N}) &\leq W_1(\tilde{\mu}_0^N, \overline{\mu}_0^N) \\ &\leq \frac{1}{N}|\xi^2 - \tilde{\xi}|, \end{aligned}$$

we deduce that

$$\begin{aligned} |I_2| &\leq \frac{C}{N^\gamma} t^{1-\frac{1+\gamma}{\alpha}} \mathbb{E}(|\xi^2 - \tilde{\xi}|\rho^0(t, y - \xi^1)) \\ &\leq \frac{C}{N^\gamma} t^{1-\frac{1+\gamma}{\alpha}} (1 + M_\gamma(\mu_0)) \int_{\mathbb{R}^d} \rho^0(t, y - x)\, d\mu_0(x). \end{aligned}$$

This concludes the proof of Lemma 8.25. □

Let us now prove the estimate (8.39). We come back to the identity (8.148) subtracting $p(\mu_0, 0, t, y)$ from its both sides and use (8.149) and (8.150) to obtain that

$$|\mathbb{E}p(\overline{\mu}_s^N, s, t, y) - p(\mu_0, 0, t, y)| \leq |\mathbb{E}p(\overline{\mu}_0^N, 0, t, y) - p(\mu_0, 0, t, y)| + R^N(\mu_0, s, t, y). \tag{8.157}$$

where

$$\begin{aligned} R^N(\mu_0, s, t, y) := \frac{C}{N^\gamma} \int_0^s \int_{(\mathbb{R}^d)^2} (t-r)^{-\zeta} f(z) \\ \Big[\rho^0(t-r, y-w-z) + \rho^0(t-r, y-w)\Big] p^{1,N}(\mu_0, 0, r, w)\, dw\, d\nu(z)\, dr \end{aligned}$$

We can now use (8.38) to get that

$$\begin{aligned} R^N(\mu_0, s, t, y) \leq{}& \frac{C^2}{N^\gamma} \int_0^s \int_{(\mathbb{R}^d)^3} (t-r)^{-\zeta} f(z) \Big[\rho^0(t-r, y-w-z) + \rho^0(t-r, y-w)\Big] \\ &\qquad \rho^0(r, w-x)\, d\mu_0(x)\, dw\, d\nu(z)\, dr \\ &+ \sum_{k=1}^{\infty} \frac{C^{k+2}}{N^{(k+1)\gamma}} \left(\prod_{j=1}^{k-1} \mathcal{B}(j(1-\zeta), 1-\zeta)\right) \\ &\int_0^s \int_{(\mathbb{R}^d)^3} (t-r)^{-\zeta} \frac{r^{k(1-\zeta)}}{k(1-\zeta)} f(z_{k+1}) \Big[\rho^0(t-r, y-w-z_{k+1}) + \rho^0(t-r, y-w)\Big] \\ &\left[\sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} \rho^0\left(r, w - x - \sum_{i \in I} z_i\right) \prod_{j=1}^k f(z_j)\, d\nu(z_j)\right] d\mu_0(x)\, dw\, d\nu(z_{k+1})\, dr. \end{aligned}$$

Thanks to the convolution inequality (8.234), we obtain that

$$\begin{aligned} R^N(\mu, s, t, y) \leq \sum_{k=1}^{\infty} \frac{C^{k+1}}{N^{k\gamma}} t^{k(1-\zeta)} \left(\prod_{j=1}^{k-2} \mathcal{B}(j(1-\zeta), 1-\zeta)\right) \mathcal{B}(1 + (k-1)(1-\zeta), 1-\zeta) \\ \sum_{I \in P_k} \int_{(\mathbb{R}^d)^k} q_0\left(\mu_0, 0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^k f(z_j)\, d\nu(z_j). \end{aligned}$$

Using the preceding inequality, Lemma 8.25 with $\gamma' \in [\alpha - 1, 1]$, and taking the limit $s \to t^-$ in (8.157) thanks to (8.153) concludes the proof of (8.39).

The estimates (8.40) and (8.41) are direct consequences of (8.39). Indeed, we integrate (8.39) over $y \in \mathbb{R}^d$ noticing that

$$\int_{\mathbb{R}^d} \int_{(\mathbb{R}^d)^k} \sum_{I \in P_k} q_0 \left(\mu_0, t, y - \sum_{i \in I} z_i\right) \prod_{j=1}^{k} f(z_j)\, d\nu(z_j)\, dy \leq 2^k \int_{\mathbb{R}^d} \rho^0(1, y)\, dy \left(\int_{\mathbb{R}^d} f(z)\, d\nu(z)\right)^k,$$

which is finite since $f \in L^1(\mathbb{R}^d, \nu)$. Using the asymptotics of the Beta function, we conclude the proof of (8.40) (by taking $\gamma' = \gamma$ in (8.39)) and (8.41) (by taking $\gamma' = \alpha - 1$ in (8.39)).

8.8 Proof of Proposition 8.21

This section is dedicated to prove Proposition 8.21, which is rather long and technical. We proceed by induction on $m \geq 1$. The base case $m = 1$ is immediate. It is enough to apply Theorem 8.41 since $p_1(\mu, s, t, x, y)$ does not depend on μ. For the induction step, we assume that all the estimates of Proposition 8.21 are satisfied for p_m. For the sake of clarity, we denote by K a positive constant depending only on (d, α, b, T) appearing in the induction step and independent of the induction assumption, which may change from line to line and which will determine the choice of the constant C appearing in Proposition 8.21. Let us introduce a notation used in this section. If f is a function defined on $\mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times (\mathbb{R}^d)^2$, we define $\Delta_{\mu_1, \mu_2} f(\cdot, s, x, v) := f(\mu_1, s, x, v) - f(\mu_2, s, x, v)$ for $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$. The same notations holds with respect to the variables x and v. For the time variable, if $s_1, s_2 \in [0, t)$, we define $\Delta_{s_1, s_2} f(\mu, \cdot, x, v) := f(\mu, s_1 \vee s_2, x, v) - f(\mu, s_1 \wedge s_2, x, v)$.

8.8.1 Preparatory technical results

Lemma 8.26.
- *There exists $K > 0$ such that for all $0 \leq s \leq r \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, v \in \mathbb{R}^d$, one has*

$$\left|\frac{\delta}{\delta m}\left[b(r, x, [X_r^{s,\mu,(m)}])\right](v)\right| \leq K\left[1 + \sum_{k=1}^{m} C^k (r-s)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right]. \tag{8.158}$$

- *There exists $K > 0$ such that for all $0 \leq s \leq r \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, v \in \mathbb{R}^d$, one has*

$$\begin{aligned}&\left|\partial_v \frac{\delta}{\delta m}\left[b(r, x, [X_r^{s,\mu,(m)}])\right](v)\right| \\ &\leq K(r-s)^{\frac{\eta-1}{\alpha}}\left[1 + \sum_{k=1}^{m} C^k (r-s)^{k(1+\frac{\eta-1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + \frac{\eta-1}{\alpha} + j\left(1 + \frac{\eta-1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].\end{aligned} \tag{8.159}$$

- *There exists $K > 0$ such that for all $0 \leq s \leq r < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$, one has*

$$\begin{aligned}\left|\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu, s, r, t, x, y)(v)\right| &\leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r, y-x) \\ &\quad \left[1 + \sum_{k=1}^{m} C^k (r-s)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].\end{aligned} \tag{8.160}$$

- *There exists $K > 0$ such that for all $0 \leq s \leq r < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$, one has*

$$\left|\partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu, s, r, t, x, y)(v)\right| \leq K(r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r, y-x)$$
$$\left[1 + \sum_{k=1}^{m} C^k (r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + \frac{\eta-1}{\alpha} + j\left(1 + \frac{\eta-1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right]. \quad (8.161)$$

Proof of Lemma 8.26. Using the induction assumption, we deduce that for all $y \in \mathbb{R}^d$, the map $\mu \in \mathcal{P}_\beta(\mathbb{R}^d) \mapsto p_m(\mu, s, r, y)$ has a linear derivative given, for all $v \in \mathbb{R}^d$, by

$$\frac{\delta}{\delta m}p_m(\mu, s, t, y)(v) = p_m(\mu, s, r, v, y) + \int_{\mathbb{R}^d} \frac{\delta}{\delta m}p_m(\mu, s, r, x', y)(v)\, d\mu(x').$$

Thus, the map $\mu \in \mathcal{P}_\beta(\mathbb{R}^d) \mapsto b(r, x, [X_r^{s,\mu,(m)}])$ has a linear derivative given by

$$\frac{\delta}{\delta m}\left[b(r, x, [X_r^{s,\mu,(m)}])\right](v) = \int_{\mathbb{R}^d} \frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(z) p_m(\mu, s, r, v, z)\, dz \quad (8.162)$$
$$+ \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(z) \frac{\delta}{\delta m}p_m(\mu, s, r, x', z)(v)\, dz\, d\mu(x').$$

Using the boundedness of $\frac{\delta}{\delta m}b$ on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$, (8.90), and the induction assumption, we deduce that

$$\begin{aligned}
&\left|\frac{\delta}{\delta m}\left[b(r, x, [X_r^{s,\mu,(m)}])\right](v)\right| \\
&\leq K + K\int_{\mathbb{R}^{2d}} \left|\frac{\delta}{\delta m}p_m(\mu, s, r, x', z)(v)\right| dz\, d\mu(x') \\
&\leq K + K\int_{\mathbb{R}^{2d}} \left(\sum_{k=1}^{m} C^k (r-s)^{(k-1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right) \\
&\qquad\qquad (r-s)^{1-\frac{1}{\alpha}}\rho^0(r-s, z-x')\, dz\, d\mu(x') \\
&\leq K + K\sum_{k=1}^{m} C^k (r-s)^{k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right).
\end{aligned}$$

This proves (8.158). Let us now focus on (8.159). Note that for all $v, x' \in \mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d} \partial_x p_m(\mu, s, t, v, z)\, dz = \int_{\mathbb{R}^d} \partial_v \frac{\delta}{\delta m}p_m(\mu, s, t, x', z)(v)\, dz = 0.$$

Thus, using the induction assumption and differentiating (8.162) with respect to v, one has for all $v \in \mathbb{R}^d$

$$\partial_v \frac{\delta}{\delta m}\left[b(r, x, [X_r^{s,\mu,(m)}])\right](v) \quad (8.163)$$
$$= \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(z) - \frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(v)\right] \partial_x p_m(\mu, s, r, v, z)\, dz$$
$$+ \int_{\mathbb{R}^{2d}} \left[\frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(z) - \frac{\delta}{\delta m}b(r, x, [X_r^{s,\mu,(m)}])(x')\right] \partial_v \frac{\delta}{\delta m}p_m(\mu, s, r, x', z)(v)\, dz\, d\mu(x').$$

Using the η-Hölder continuity of $\frac{\delta}{\delta m}b$ with respect to v, which is uniform with respect to the other variables, (8.90), and the induction assumption, we deduce that

$$\begin{aligned}
&\left|\partial_v \frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](v)\right| \\
&\leq K(r-s)^{\frac{\eta-1}{\alpha}} + K\int_{\mathbb{R}^{2d}} |z-x'|^\eta \left|\partial_v \frac{\delta}{\delta m} p_m(\mu,s,r,x',z)(v)\right| dz\, d\mu(x') \\
&\leq K(r-s)^{\frac{\eta-1}{\alpha}} + K\int_{\mathbb{R}^{2d}} \left(\sum_{k=1}^m C^k (r-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\right) \\
&\qquad\qquad (r-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} |z-x'|^\eta \rho^0(r-s,z-x')\, dz\, d\mu(x') \\
&\leq K(r-s)^{\frac{\eta-1}{\alpha}} + K(r-s)^{\frac{\eta-1}{\alpha}} \sum_{k=1}^m C^k (r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right), 1-\frac{1}{\alpha}\right).
\end{aligned}$$

It ends the proof of (8.159). Using (8.158), it is clear that $\mathcal{H}_{m+1}$ has a linear derivative given for all $v \in \mathbb{R}^d$ by

$$\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,r,t,x,y)(v) = \left(\frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](v)\right)\cdot \partial_x \widehat{p}(r,t,x,y), \tag{8.164}$$

and that

$$\partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,r,t,x,y)(v) = \left(\partial_v\frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](v)\right)\cdot \partial_x \widehat{p}(r,t,x,y). \tag{8.165}$$

Thus, (8.160) and (8.161) follow directly from (8.158), (8.159), and (8.235). □

Lemma 8.27. *For any $k \geq 1$, $0 \leq s \leq r < t \leq T$, $x, y \in \mathbb{R}^d$, $\mathcal{H}_{m+1}^k(\cdot,s,r,t,x,y)$ has a linear derivative which is $\mathcal{C}^1$ and satisfies the following properties.*

- *There exists a positive constant K_m depending on m and a positive constant C independent of k and m such that for all $k \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$, $x, y, v \in \mathbb{R}^d$, one has*

$$\begin{aligned}\left|\frac{\delta}{\delta m}\mathcal{H}_{m+1}^k(\mu,s,r,t,x,y)(v)\right| &\leq K_m k C^{k-1}(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x),\end{aligned} \tag{8.166}$$

and

$$\begin{aligned}\left|\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}^k(\mu,s,r,t,x,y)(v)\right| &\leq K_m k C^{k-1}(r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x).\end{aligned} \tag{8.167}$$

- *There exists a positive constant K_m depending on m and a positive constant C independent of k*

and m such that for all $k \geq 1$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s < t \leq T$, $x, y, v \in \mathbb{R}^d$, one has

$$\left|\widehat{p} \otimes \frac{\delta}{\delta m}\mathcal{H}^k_{m+1}(\mu, s, t, x, y)(v)\right| \leq K_m k C^{k-1}(t-s)^{1-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right), 1\right) \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^0(t-s, y-x), \quad (8.168)$$

and

$$\left|\widehat{p} \otimes \partial_v \frac{\delta}{\delta m}\mathcal{H}^k_{m+1}(\mu, s, t, x, y)(v)\right| \leq K_m k C^{k-1}(t-s)^{\frac{\eta-1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right), 1+\frac{\eta-1}{\alpha}\right) \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^0(t-s, y-x). \quad (8.169)$$

Proof of Lemma 8.27. We proceed by induction on k to prove (8.166) for all $m \geq 1$. The base case $k = 1$ is a direct consequence of (8.160). Assume now that (8.166) holds for $\mathcal{H}^k_{m+1}$ and let us prove it for $\mathcal{H}^{k+1}_{m+1}$. By definition, we have

$$\mathcal{H}^{k+1}_{m+1}(\mu, s, r, t, x, y) = \int_r^t \int_{\mathbb{R}^d} \mathcal{H}_{m+1}(\mu, s, r, r', x, z)\mathcal{H}^k_{m+1}(\mu, s, r', t, z, y)\, dz\, dr'.$$

Using Lemma 8.26 and the induction assumption, we deduce that $\mathcal{H}^{k+1}_{m+1}(\cdot, s, r, t, x, y)$ has a linear derivative given by

$$\frac{\delta}{\delta m}\mathcal{H}^{k+1}_{m+1} = \frac{\delta}{\delta m}\mathcal{H}_{m+1} \otimes \mathcal{H}^k_{m+1} + \mathcal{H}_{m+1} \otimes \frac{\delta}{\delta m}\mathcal{H}^k_{m+1}.$$

Using (8.85), the induction assumption (8.166) and the convolution inequality (8.234), we deduce that

$$\begin{aligned}
&\left|\frac{\delta}{\delta m}\mathcal{H}^{k+1}_{m+1}(\mu, s, r, t, x, y)(v)\right| \\
&\leq K_m \int_r^t \int_{\mathbb{R}^d} (r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r, z-x) C^k (t-r')^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r', y-z)\, dz\, dr' \\
&\quad + \int_r^t \int_{\mathbb{R}^d} K(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r, z-x) K_m k C^{k-1}(t-r')^{(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)(t-r')^{-\frac{1}{\alpha}}\rho^1(t-r', y-z)\, dz\, dr' \\
&\leq K_m C^k (t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r, y-x) \\
&\quad + K K_m k C^{k-1}(t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r, y-x) \\
&\leq K_m (k+1) C^k (t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right)\rho^1(t-r, y-x),
\end{aligned}$$

if we choose $C \geq K$. Following the same lines, we prove that for any $k \geq 1$, $\frac{\delta}{\delta m}\mathcal{H}^k_{m+1}$ is $\mathcal{C}^1$ with respect to $v \in \mathbb{R}^d$ and that it satisfies

$$\partial_v \frac{\delta}{\delta m}\mathcal{H}^{k+1}_{m+1} = \partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1} \otimes \mathcal{H}^k_{m+1} + \mathcal{H}_{m+1} \otimes \partial_v \frac{\delta}{\delta m}\mathcal{H}^k_{m+1}.$$

Using (8.85), the induction assumption (8.167) and the convolution inequality (8.234), we deduce that

$$\begin{aligned}
&\left|\partial_v \frac{\delta}{\delta m}\mathcal{H}^{k+1}_{m+1}(\mu,s,r,t,x,y)(v)\right| \\
&\leq K_m \int_r^t \int_{\mathbb{R}^d} (r-s)^{\frac{\eta-1}{\alpha}}(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r,z-x) C^k (t-r')^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r',y-z)\,dz\,dr' \\
&\quad + \int_r^t \int_{\mathbb{R}^d} K(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r,z-x) K_m k C^{k-1}(r'-s)^{\frac{\eta-1}{\alpha}}(t-r')^{(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)(t-r')^{-\frac{1}{\alpha}}\rho^1(t-r',y-z)\,dz\,dr \\
&\leq K_m C^k (r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x) \\
&\quad + K K_m k C^{k-1}(r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x) \\
&\leq K_m (k+1) C^k (r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-x),
\end{aligned}$$

if we chose $C \geq K$. It proves (8.167). The estimates (8.168) and (8.169) follow immediately from (8.166), (8.167), Lemma 8.43 and the convolution inequality (8.234).

□

Lemma 8.28.
- *For any $0 \leq s < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, we have*

$$\left|\partial_s \left[b(t,x,[X_t^{s,\mu,(m)}])\right]\right| \leq K \int_{\mathbb{R}^{2d}} (1 \wedge |x'-y|^\eta)|\partial_s p_m(\mu,s,t,x',y)|\,dy\,d\mu(x'). \tag{8.170}$$

- *For any $0 \leq s < r < t \leq T$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$, we have*

$$|\partial_s \mathcal{H}_{m+1}(\mu,s,r,t,x,y)| \leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\int_{\mathbb{R}^{2d}}(1 \wedge |x'-y|^\eta)|\partial_s p_m(\mu,s,r,x',y)|\,dy\,d\mu(x'). \tag{8.171}$$

Proof of Lemma 8.28. **Proof of** (8.170)**.**

By the induction assumption at step m, we see by the dominated convergence theorem that the map $s \in [0,t) \mapsto b(t,x,[X_t^{s,\mu,(m)}])$ is differentiable and that

$$\partial_s \left[b(t,x,[X_t^{s,\mu,(m)}])\right] = \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu,(m)}])(y) - \frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu,(m)}])(x')\right)\partial_s p_m(\mu,s,t,x',y)\,d\mu(x')\,dy.$$

The boundedness of $\frac{\delta}{\delta m}b$ and the η-Hölder continuity of $\frac{\delta}{\delta m}b(t,x,\mu)(\cdot)$ allows to conclude.

Proof of (8.171).

Because of the expression (8.82) of $\mathcal{H}_{m+1}$ and since $\widehat{p}(\mu,s,r,t,x,y)$ does not depend on s, (8.171) follows from (8.170) and (8.235). $\square$

Lemma 8.29. • *There exists a positive constant K_m depending on m and a positive constant C such that for all $k\geq 1$, $m\geq 1$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s<r<t\leq T$, $x,y\in\mathbb{R}^d$*

$$\left|\partial_s\mathcal{H}^k_{m+1}(\mu,s,r,t,x,y)\right|\leq K_m k C^{k-1}(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\rho^1(t-r,y-x)\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right). \tag{8.172}$$

• *There exists a positive constant K_m depending on m such that for all $k\geq 1$, $m\geq 1$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s<r<t\leq T$, $x,y\in\mathbb{R}^d$*

$$|\partial_s\Phi_{m+1}(\mu,s,r,t,x,y)|\leq K_m(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x). \tag{8.173}$$

Proof of Lemma 8.29.. **Proof of (8.172).** We proceed by induction on $k\geq 1$. The base case $k=1$ comes from (8.171) and the induction assumption (8.94) for $\partial_s p_m$. Indeed, the induction assumption and the space-time inequality (8.231) ensure that

$$\int_{\mathbb{R}^{2d}}(1\wedge|x'-y|^\eta)|\partial_s p_m(\mu,s,r,x',y)|\,dy\,d\mu(x')\leq K_m(r-s)^{\frac{\eta}{\alpha}-1}\int_{\mathbb{R}^{2d}}\rho^{-\tilde{\eta}-\eta}(r-s,y-x')\,dy\;d\mu(x').$$

We conclude since $\eta+\tilde{\eta}<\alpha$, the map $\rho^{-\tilde{\eta}-\eta}(r-s,\cdot)$ belongs to $L^1(\mathbb{R}^d)$ and $\int_{\mathbb{R}^d}\rho^{-\tilde{\eta}-\eta}(t-s,y)\,dy$ is equal to a constant independent of s and t. For the induction step, we assume that (8.172) holds at step k. By definition, we have

$$\mathcal{H}^{k+1}_{m+1}(\mu,s,r,t,x,y)=\int_r^t\int_{\mathbb{R}^d}\mathcal{H}_{m+1}(\mu,s,r,r',x,z)\mathcal{H}^k_{m+1}(\mu,s,r',t,z,y)\,dz\,dr'.$$

It follows from the induction assumption and the dominated convergence theorem that the map $s\in[0,t)\mapsto\mathcal{H}^{k+1}_{m+1}(\mu,s,r,t,x,y)$ is differentiable and that

$$\begin{aligned}\partial_s\mathcal{H}^{k+1}_{m+1}(\mu,s,r,t,x,y)&=\int_r^t\int_{\mathbb{R}^d}\partial_s\mathcal{H}_{m+1}(\mu,s,r,r',x,z)\mathcal{H}^k_{m+1}(\mu,s,r',t,z,y)\,dz\,dr'\\&\quad+\int_r^t\int_{\mathbb{R}^d}\mathcal{H}_{m+1}(\mu,s,r,r',x,z)\partial_s\mathcal{H}^k_{m+1}(\mu,s,r',t,z,y)\,dz\,dr'\\&=:I_1+I_2.\end{aligned}$$

Using the base case $k=1$ and (8.85), we deduce that

$$
\begin{aligned}
|I_1| \leq \int_r^t \int_{\mathbb{R}^d} & K_m (r-s)^{\frac{\eta}{\alpha}-1} (r'-r)^{-\frac{1}{\alpha}} \rho^1(r'-r, z-x) C^k (t-r')^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
& \prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^1(t-r', y-z)\, dr'\, dz \\
\leq K_m C^k & (r-s)^{\frac{\eta}{\alpha}-1} (t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^1(t-r, y-x).
\end{aligned}
$$

For I_2, from the induction assumption (8.172), (8.85), the convolution inequality (8.234) and since $\frac{\eta}{\alpha} < 1$, we obtain that

$$
\begin{aligned}
|I_2| \leq \int_r^t \int_{\mathbb{R}^d} & K (r'-r)^{-\frac{1}{\alpha}} \rho^1(r'-r, z-x) K_m k C^{k-1} (r'-s)^{\frac{\eta}{\alpha}-1} (t-r')^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
& \prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^1(t-r', y-z)\, dz\, dr' \\
\leq (r-s)^{\frac{\eta}{\alpha}-1} & \int_r^t \int_{\mathbb{R}^d} K (r'-r)^{-\frac{1}{\alpha}} \rho^1(r'-r, z-x) K_m k C^{k-1} (t-r')^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
& \prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^1(t-r', y-z)\, dz\, dr' \\
\leq K_m k C^k & (r-s)^{\frac{\eta}{\alpha}-1} (t-r)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)} \rho^1(t-r, y-x) \prod_{j=1}^{k} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right),
\end{aligned}
$$

provided that we choose $C \geq K$ in (8.172). This concludes the induction step for (8.172).

Proof of (8.173). Using the definition of Φ_{m+1} (8.84) and (8.172), we obtain by the dominated convergence theorem that $s \in [0, r) \mapsto \phi_{m+1}(\mu, s, r, t, x, y)$ is continuously differentiable with

$$
\partial_s \Phi_{m+1}(\mu, s, r, t, x, y) = \sum_{k=1}^{\infty} \partial_s \mathcal{H}_{m+1}^k(\mu, s, r, t, x, y).
$$

Then, (8.173) follows immediately from (8.172).

□

8.8.2 First part of the proof of the induction step

We split the proof of the induction step into different parts for the sake of clarity. We start by proving the estimates (8.92), (8.93) and (8.94).

Proof of (8.92) and (8.93). We start by showing that for all $0 \leq s < t \leq T$, $x, y \in \mathbb{R}^d$, the map $p_{m+1}(\cdot, s, t, x, y)$ admits a linear derivative given which is $\mathcal{C}^1$ with respect to $v \in \mathbb{R}^d$ and such that for all $v \in \mathbb{R}^d$

$$
\frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v) = \sum_{k=1}^{\infty} \widehat{p} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}^k(\mu, s, t, x, y)(v), \tag{8.174}
$$

and

$$\partial_v \frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v) = \sum_{k=1}^{\infty} \widehat{p} \otimes \partial_v \frac{\delta}{\delta m} \mathcal{H}^k_{m+1}(\mu, s, t, x, y)(v). \tag{8.175}$$

where the series are absolutely convergent. Moreover, we also have the following representation formulas

$$\frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v) = \sum_{k=0}^{\infty} p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}^k_{m+1}(\mu, s, t, x, y)(v), \tag{8.176}$$

and

$$\partial_v \frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v) = \sum_{k=0}^{\infty} p_{m+1} \otimes \partial_v \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}^k_{m+1}(\mu, s, t, x, y)(v). \tag{8.177}$$

To prove (8.174), we fix $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$ and we write thanks to Lemma 8.27 and Fubini's theorem

$$\begin{aligned}
&p_{m+1}(\mu_1, s, t, x, y) - p_{m+1}(\mu_2, s, t, x, y)\\
&= \sum_{k=1}^{\infty} \widehat{p} \otimes \mathcal{H}^k_{m+1}(\mu_1, s, t, x, y) - \widehat{p} \otimes \mathcal{H}^k_{m+1}(\mu_2, s, t, x, y)\\
&= \sum_{k=1}^{\infty} \int_s^t \int_{\mathbb{R}^d} \widehat{p}(s, r, x, z)(\mathcal{H}^k_{m+1}(\mu_1, s, r, t, z, y) - \mathcal{H}^k_{m+1}(\mu_2, s, r, t, z, y))\, dz\, dr\\
&= \sum_{k=1}^{\infty} \int_s^t \int_{\mathbb{R}^d} \widehat{p}(s, r, x, z) \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} \mathcal{H}^k_{m+1}(\lambda\mu_1 + (1-\lambda)\mu_2, s, r, t, z, y)(v)\, d(\mu_1 - \mu_2)(v)\, d\lambda\, dz\, dr\\
&= \int_0^1 \int_{\mathbb{R}^d} \sum_{k=1}^{\infty} \int_s^t \int_{\mathbb{R}^d} \widehat{p}(s, r, x, z) \frac{\delta}{\delta m} \mathcal{H}^k_{m+1}(\lambda\mu_1 + (1-\lambda)\mu_2, s, r, t, z, y)(v)\\
&\qquad\qquad d(\mu_1 - \mu_2)(v)\, d\lambda\, dz\, dr\, d(\mu_1 - \mu_2)(v)\, d\lambda\\
&= \int_0^1 \int_{\mathbb{R}^d} \sum_{k=1}^{\infty} \widehat{p} \otimes \frac{\delta}{\delta m} \mathcal{H}^k_{m+1}(\lambda\mu_1 + (1-\lambda)\mu_2, s, t, x, y)(v)\, d(\mu_1 - \mu_2)(v)\, d\lambda.
\end{aligned}$$

Note that owing to (8.168) the series (8.174) is absolutely convergent, locally uniformly with respect to $(\mu, s, x, v) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, t) \times \mathbb{R}^d \times \mathbb{R}^d$. This concludes the proof of (8.174). Moreover, we have proved that the map $(\mu, s, x, v) \mapsto \frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v)$ is continuous. By differentiation under the integral using (8.167), we obtain (8.175). The dominated convergence theorem yields the continuity of the map $(\mu, s, x, v) \mapsto \partial_v \frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v)$. Let us now focus on the representation formula (8.176). Using the parametrix expansion (8.89) of p_{m+1}, one has for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \le s < t \le T$, $x, y \in \mathbb{R}^d$

$$p_{m+1}(\mu, s, t, x, y) = \widehat{p}(s, t, x, y) + p_{m+1} \otimes \mathcal{H}_{m+1}(\mu, s, t, x, y).$$

We can easily see by induction thanks to (8.166) that for any $k \ge 1$, one has

$$\frac{\delta}{\delta m} \mathcal{H}^k_{m+1}(\mu, s, r, t, x, y)(v) = \sum_{j=1}^{k} \mathcal{H}^{k-j}_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}^{j-1}_{m+1}(\mu, s, r, t, x, y)(v).$$

We plug this expression into (8.174), and since the series is absolutely convergent, we obtain setting

$l = k - j$ and $i = j - 1$ and by Fubini's theorem that

$$\begin{aligned}\frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v) &= \sum_{k=1}^{\infty} \sum_{j=1}^{k} \widehat{p} \otimes \mathcal{H}_{m+1}^{k-j} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}_{m+1}^{j-1}(\mu, s, t, x, y)(v) \\ &= \sum_{l=0}^{\infty} \sum_{i=0}^{\infty} \widehat{p} \otimes \mathcal{H}_{m+1}^{l} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}_{m+1}^{i}(\mu, s, t, x, y)(v) \\ &= \sum_{i=0}^{\infty} p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}_{m+1}^{i}(\mu, s, t, x, y)(v).\end{aligned}$$

This is exactly (8.176). Let us now prove that the estimate (8.92) is still true at step $m+1$. It follows from (8.160), (8.90) and the convolution inequality (8.234) that for some positive constant K independent of m one has

$$\begin{aligned}&\left|\widehat{p} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}(\mu, s, t, x, y)(v)\right| \\ &\leq K \int_s^t \int_{\mathbb{R}^d} \left[1 + \sum_{k=1}^{m} C^k (r-s)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right] \\ &\qquad\qquad \rho^0(r-s, z-x)(t-r)^{-\frac{1}{\alpha}} \rho^1(t-r, y-z)\, dz\, dr \\ &\leq K(t-s)^{1-\frac{1}{\alpha}} + K \sum_{k=1}^{m} C^k (t-s)^{(k+1)(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \rho^0(t-s, y-x).\end{aligned}$$

Using this inequality, the bound (8.85) for $\mathcal{H}_{m+1}^k$, the convolution inequality (8.234) and summing over $k \geq 0$, we find that there exists a constant K independent of m such that

$$\begin{aligned}&\left|\frac{\delta}{\delta m} p_{m+1}(\mu, s, t, x, y)(v)\right| \\ &\leq \sum_{k=0}^{\infty} \left|p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1} \otimes \mathcal{H}_{m+1}^{k}(\mu, s, t, x, y)(v)\right| \\ &\leq K(t-s)^{1-\frac{1}{\alpha}} + K \sum_{k=1}^{m} C^k (t-s)^{(k+1)(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \rho^0(t-s, y-x) \\ &\leq (t-s)^{1-\frac{1}{\alpha}} \rho^0(t-s, y-x) \left(\sum_{k=1}^{m+1} C^{k+1} (t-s)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + j\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right),\end{aligned}$$

provided that we choose $C \geq K$ in (8.92). It ends the proof of the induction step for (8.92).

Notice that by differentiating under the integral (8.176) using (8.90), (8.161) and (8.85), we obtain the representation formula (8.177) for $\partial_v \frac{\delta}{\delta m} p_{m+1}$. Let us now prove that the estimate (8.93) is verified at step $m+1$. It follows from (8.161), (8.90) and the convolution inequality (8.234) that for some positive constant K independent of m one has

$$\begin{aligned}
&\left|\widehat{p}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,t,x,y)(v)\right|\\
&\leq K\int_s^t\int_{\mathbb{R}^d}(r-s)^{\frac{\eta-1}{\alpha}}\left[1+\sum_{k=1}^m C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\
&\qquad\qquad\rho^0(r-s,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\
&\leq K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k}\mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\rho^0(t-s,y-x).
\end{aligned}$$

Using this inequality, the bound (8.85) for $\mathcal{H}^k_{m+1}$, the convolution inequality (8.234) and summing over $k\geq 0$, we find that there exists a constant K independent of m such that

$$\begin{aligned}
&\left|\partial_v\frac{\delta}{\delta m}p_{m+1}(\mu,s,t,x,y)(v)\right|\\
&\leq\sum_{k=0}^{\infty}\left|p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\otimes\mathcal{H}^k_{m+1}(\mu,s,t,x,y)(v)\right|\\
&\leq K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k}\mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\
&\qquad\qquad(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s,y-x)\\
&\leq(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s,y-x)\left(\sum_{k=1}^{m+1}C^k(t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right),
\end{aligned}$$

provided that we choose $C\geq K$ in (8.93). It ends the proof of the induction step for (8.93).

Proof of (8.94)**.** Let us first prove the following representation formula for $\partial_s\Phi_{m+1}$

$$\begin{aligned}
\partial_s\Phi_{m+1}(\mu,s,r,t,x,y)&=\left[\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right](\mu,s,r,t,x,y)\\
&\quad+\Phi_{m+1}\otimes\left[\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right](\mu,s,r,t,x,y).
\end{aligned}\tag{8.178}$$

By differentiating the relation $\Phi_{m+1}=\mathcal{H}_{m+1}+\mathcal{H}_{m+1}\otimes\Phi_{m+1}$, we obtain that

$$\begin{aligned}
\partial_s\Phi_{m+1}(\mu,s,r,t,x,y)&=\partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,y)+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}(\mu,s,r,t,x,y)\\
&\quad+\mathcal{H}_{m+1}\otimes\partial_s\Phi_{m+1}(\mu,s,r,t,x,y).
\end{aligned}$$

Notice that by (8.172) for $k=1$ and (8.87), we get that

$$\left|\left[\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right](\mu,s,r,t,x,y)\right|\leq K_m(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x).$$

The kernel $\left[\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right](\mu,s,r,t,x,y)$ yields a time-integrable singularity of order $(r-$

$t)^{-\frac{1}{\alpha}}$. We can thus iterate the previous relation to obtain that

$$\begin{aligned}
&\partial_s \Phi_{m+1}(\mu, s, r, t, x, y) \\
&= \sum_{k=0}^{\infty} \mathcal{H}_{m+1}^k \otimes [\partial_s \mathcal{H}_{m+1} + \partial_s \mathcal{H}_{m+1} \otimes \Phi_{m+1}] (\mu, s, r, t, x, y) \\
&= [\partial_s \mathcal{H}_{m+1} + \partial_s \mathcal{H}_{m+1} \otimes \Phi_{m+1}] (\mu, s, r, t, x, y) + \Phi_{m+1} \otimes [\partial_s \mathcal{H}_{m+1} + \partial_s \mathcal{H}_{m+1} \otimes \Phi_{m+1}] (\mu, s, r, t, x, y).
\end{aligned}$$

In order to deal with the differentiability of the map $s \in [0, t) \mapsto p_{m+1}(\mu, s, t, x, y)$, keeping in mind the parametrix expansion (8.89), we first study the differentiability of the map

$$s \in [0, r) \mapsto \int_{\mathbb{R}^d} \widehat{p}(s, r, x, z) \Phi_{m+1}(\mu, s, r, t, z, y)\, dz.$$

Since $\int_{\mathbb{R}^d} \partial_s \widehat{p}(s, r, x, z)\, dz = 0$, we deduce by the dominated convergence theorem that

$$\begin{aligned}
\partial_s \left(\int_{\mathbb{R}^d} \widehat{p}(s, r, x, z) \Phi_{m+1}(\mu, s, r, t, z, y)\, dz \right) &= \int_{\mathbb{R}^d} \partial_s \widehat{p}(s, r, x, z) \left(\Phi_{m+1}(\mu, s, r, t, z, y) - \Phi_{m+1}(\mu, s, r, t, x, y) \right) dz \\
&\quad + \int_{\mathbb{R}^d} \widehat{p}(s, r, x, z) \partial_s \Phi_{m+1}(\mu, s, r, t, z, y)\, dz \qquad (8.179) \\
&=: A_1 + A_2,
\end{aligned}$$

which is continuous with respect to $(\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0, r) \times \mathbb{R}^d$. We now control A_1 and A_2. For A_1, it follows from (8.229), (8.88) with $\gamma = \tilde{\eta}$ and the space-time inequality (8.231) that

$$\begin{aligned}
|A_1| &\leq K \int_{\mathbb{R}^d} (r - s)^{-1} \rho^0(r - s, z - x)(t - r)^{-\frac{\tilde{\eta}+1}{\alpha}} |z - x|^{\tilde{\eta}} \left[\rho^1(t - r, y - z) + \rho^1(t - r, y - x) \right] dz \\
&\leq K \int_{\mathbb{R}^d} (r - s)^{\frac{\tilde{\eta}}{\alpha} - 1} \rho^{-\tilde{\eta}}(r - s, z - x)(t - r)^{-\frac{\tilde{\eta}+1}{\alpha}} \left[\rho^{-\tilde{\eta}}(t - r, y - z) + \rho^{-\tilde{\eta}}(t - r, y - x) \right] dz.
\end{aligned}$$

Note that $\int_{\mathbb{R}^d} \rho^{-\tilde{\eta}}(r - s, z - x)\, dz$ is a constant independent of s and r . By the definition of $\rho^{-\tilde{\eta}}$ (8.225), the fact that $r > s$ and the convolution inequality (8.234), one has

$$|A_1| \leq K (r - s)^{\frac{\tilde{\eta}}{\alpha} - 1} (t - r)^{-\frac{\tilde{\eta}+1}{\alpha}} \left[\rho^{-\tilde{\eta}}(t - s, y - x) + (t - r)^{-\frac{d}{\alpha}} (1 + (t - s)^{-\frac{1}{\alpha}} |y - x|)^{-d-\alpha+\tilde{\eta}} \right]. \qquad (8.180)$$

Concerning A_2, it follows from (8.173) and (8.235) that

$$\begin{aligned}
|A_2| &\leq \int_{\mathbb{R}^d} K \rho^0(r - s, z - x) K_m (r - s)^{\frac{\eta}{\alpha} - 1} (t - r)^{-\frac{1}{\alpha}} \rho^1(t - r, y - z)\, dz \\
&\leq K_m (r - s)^{\frac{\eta}{\alpha} - 1} (t - r)^{-\frac{1}{\alpha}} \rho^0(t - s, y - x).
\end{aligned}$$

By the dominated convergence theorem justified by the controls previously obtained on A_1 and A_2, we obtain that the map $s \in [0, t) \mapsto \widehat{p} \otimes \Phi_{m+1}(\mu, s, t, x, y)$ is differentiable with

$$\begin{aligned}\partial_s\left(\widehat{p}\otimes\Phi_{m+1}\right)(\mu,s,t,x,y) &= -\Phi_{m+1}(\mu,s,t,x,y)+\partial_s\widehat{p}\otimes\Phi_{m+1}(\mu,s,t,x,y)+\widehat{p}\otimes\partial_s\Phi_{m+1}(\mu,s,t,x,y)\\ &= -\Phi_{m+1}(\mu,s,t,x,y)\\ &\quad+\int_s^t\int_{\mathbb{R}^d}\partial_s\widehat{p}(s,r,x,z)\left(\Phi_{m+1}(\mu,s,r,t,z,y)-\Phi_{m+1}(\mu,s,r,t,x,y)\right)dz\,dr\\ &\quad+\int_s^t\int_{\mathbb{R}^d}\widehat{p}(s,r,x,z)\partial_s\Phi_{m+1}(\mu,s,r,t,z,y)\,dz\,dr.\end{aligned}$$

Thus, the map $s\in[0,t)\mapsto p_{m+1}(\mu,s,t,x,y)$ is differentiable with

$$\begin{aligned}\partial_s p_{m+1}(\mu,s,t,x,y) &= \partial_s\widehat{p}(s,t,x,y)-\Phi_{m+1}(\mu,s,t,x,y)\\ &\quad+\partial_s\widehat{p}\otimes\Phi_{m+1}(\mu,s,t,x,y)+\widehat{p}\otimes\partial_s\Phi_{m+1}(\mu,s,t,x,y).\end{aligned}\tag{8.181}$$

Then, plugging (8.178) into (8.181), we obtain that

$$\begin{aligned}\partial_s p_{m+1}(\mu,s,t,x,y) &= \partial_s\widehat{p}(s,t,x,y)-\Phi_{m+1}(\mu,s,t,x,y)+\partial_s\widehat{p}\otimes\Phi_{m+1}(\mu,s,t,x,y)\\ &\quad+\widehat{p}\otimes\left(\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right)(\mu,s,t,x,y)\\ &\quad+\widehat{p}\otimes\Phi_{m+1}\otimes\left(\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right)(\mu,s,t,x,y).\end{aligned}$$

By the parametrix expansion (8.89) of p_{m+1}, we obtain the following representation formula

$$\begin{aligned}\partial_s p_{m+1}(\mu,s,t,x,y) &= \partial_s\widehat{p}(s,t,x,y)-\Phi_{m+1}(\mu,s,t,x,y)\\ &\quad+\partial_s\widehat{p}\otimes\Phi_{m+1}(\mu,s,t,x,y)\\ &\quad+p_{m+1}\otimes\left(\partial_s\mathcal{H}_{m+1}+\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right)(\mu,s,t,x,y)\\ &=: I_1+I_2+I_3.\end{aligned}$$

We can now prove that the estimate (8.94) is still true at step $m+1$ for some choice of the constant C which appears in (8.94).

For I_1, we note that (8.236) and (8.87) yield

$$|I_1|\le K(t-s)^{-1}\rho^0(t-s,y-x)\le K(t-s)^{-1}\rho^{-\tilde{\eta}}(t-s,y-x),$$

since $\alpha\in(1,2)$ and $\rho^0(t-s,y-x)\le\rho^{-\tilde{\eta}}(t-s,y-x)$.

Concerning I_2, it can be rewritten as

$$I_2=\int_s^t\int_{\mathbb{R}^d}\partial_s\widehat{p}(s,r,x,z)\left(\Phi_{m+1}(\mu,s,r,t,z,y)-\Phi_{m+1}(\mu,s,r,t,x,y)\right)dz\,dr.$$

Owing to the bound (8.180) obtained for A_1, which was defined in (8.179), we deduce that

$$
\begin{aligned}
|I_2| &\leq K \int_s^t (r-s)^{\frac{\tilde\eta}{\alpha}-1}(t-r)^{-\frac{\tilde\eta+1}{\alpha}} \left[\rho^{-\tilde\eta}(t-s,y-x) + (t-r)^{-\frac{d}{\alpha}}(1+(t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\right] dr \\
&\leq K(t-s)^{-\frac{1}{\alpha}}\rho^{-\tilde\eta}(t-s,y-x) + K(t-s)^{-\frac{1}{\alpha}-\frac{d}{\alpha}}(1+(t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta} \\
&\leq K(t-s)^{-\frac{1}{\alpha}}\rho^{-\tilde\eta}(t-s,y-x) \\
&\leq K(t-s)^{-1}\rho^{-\tilde\eta}(t-s,y-x).
\end{aligned}
$$

We now focus on I_3. Using the induction assumption (8.94), (8.171), the fact that $\eta+\tilde\eta<\alpha$ and the space-time inequality (8.231), we get that

$$
\begin{aligned}
&|\partial_s \mathcal{H}_{m+1}(\mu,s,r,t,x,y)| \\
&\leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\int_{\mathbb{R}^{2d}} |x'-y|^\eta (r-s)^{-1}\rho^{-\tilde\eta}(r-s,y-x') \\
&\qquad \sum_{k=1}^m C^k (r-s)^{(k-1)(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k-1}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right) dy\, d\mu(x') \\
&\leq K(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x) \\
&\qquad \sum_{k=1}^m C^k (r-s)^{(k-1)(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k-1}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

Since the kernel Φ_{m+1} yields a time-integrable singularity by (8.87), we deduce with the preceding inequality that

$$
\begin{aligned}
&|[\partial_s \mathcal{H}_{m+1} + \partial_s \mathcal{H}_{m+1}\otimes\Phi_{m+1}](\mu,s,r,t,x,y)| \\
&\leq K(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x) \\
&\qquad \sum_{k=1}^m C^k (r-s)^{(k-1)(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k-1}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

It follows from this inequality, (8.90) and the convolution inequality (8.234) that

$$
\begin{aligned}
|I_3| &\leq \int_s^t\int_{\mathbb{R}^d} \rho^0(r-s,z-x)K(r-s)^{\frac{\eta}{\alpha}-1}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z) \\
&\qquad \sum_{k=1}^m C^k (r-s)^{(k-1)(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k-1}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right) dz\, dr \\
&\leq K(t-s)^{-1}\rho^0(t-s,y-x)\sum_{k=1}^m C^k (r-s)^{k(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

Gathering the previous estimates on I_1, I_2 and I_3, we have

$$\begin{aligned}
&|\partial_s p_{m+1}(\mu,s,t,x,y)| \\
&\leq (t-s)^{-1}\rho^{-\tilde{\eta}}(t-s,y-x)K\left[1+\sum_{k=1}^{m}C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right] \\
&\leq (t-s)^{-1}\rho^{-\tilde{\eta}}(t-s,y-x)\sum_{k=1}^{m+1}C^k(t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(\frac{\eta}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)
\end{aligned}$$

provided that we choose $C \geq K$ in (8.94). This concludes the induction step.

8.8.3 Preparatory technical results

Lemma 8.30. • *For all $\gamma \in (0,1] \cap (0,(2\alpha-2)\wedge(\eta+\alpha-1))$, there exists a positive constant K independent of m and such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r \leq T$, $x, v_1, v_2 \in \mathbb{R}^d$, one has*

$$\begin{aligned}
&\left|\partial_v\frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](v_1)-\partial_v\frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](v_2)\right| \qquad (8.182)\\
&\leq K(r-s)^{\frac{\eta-1-\gamma}{\alpha}}|v_1-v_2|^\gamma\left[1+\sum_{k=1}^{m}C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}$$

• *For all $\gamma \in (0,1] \cap (0,(2\alpha-2)\wedge(\eta+\alpha-1))$, there exists a positive constant K independent of m and such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$, $x, y, v_1, v_2 \in \mathbb{R}^d$, one has*

$$\begin{aligned}
&\left|\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,r,t,x,y)(v_1)-\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,r,t,x,y)(v_2)\right| \qquad (8.183)\\
&\leq K(r-s)^{\frac{\eta-1-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^1(t-r,y-x)\\
&\qquad\left[1+\sum_{k=1}^{m}C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}$$

Proof of Lemma 8.30. First, note that we only need to show (8.182) since it implies (8.183) because of (8.164). We can write

$$\begin{aligned}
&\Delta_{v_1,v_2}\partial_v\frac{\delta}{\delta m}\left[b(r,x,[X_r^{s,\mu,(m)}])\right](\cdot)\\
&=\int_{\mathbb{R}^d}\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(z)\Delta_{v_1,v_2}\partial_x p_m(\mu,s,r,\cdot,z)\,dz\\
&\quad+\int_{\mathbb{R}^{2d}}\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(z)\Delta_{v_1,v_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,s,r,x',z)(\cdot)\,dz\,d\mu(x')\\
&=\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(z)-\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(v_2)\right)\Delta_{v_1,v_2}\partial_x p_m(\mu,s,r,\cdot,z)\,dz\\
&\quad+\int_{\mathbb{R}^{2d}}\left(\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(z)-\frac{\delta}{\delta m}b(r,x,[X_r^{s,\mu,(m)}])(x')\right)\\
&\qquad\qquad\Delta_{v_1,v_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,s,r,x',z)(\cdot)\,dz\,d\mu(x')\\
&=: I_1+I_2.
\end{aligned}$$

Let us first focus on I_1. We start by assuming that $|v_1 - v_2| \le (r-s)^{\frac{1}{\alpha}}$. Using the η-Hölder continuity of $\frac{\delta}{\delta m} b(r,x,\mu)(\cdot)$, (8.91), (8.232) since $|v_1 - v_2| \le (r-s)^{\frac{1}{\alpha}}$ and (8.231), we get that

$$\begin{aligned}
|I_1| &\le K \int_{\mathbb{R}^d} |z - v_2|^\eta (r-s)^{-\frac{\gamma+1}{\alpha}} |v_1 - v_2|^\gamma \rho^1(r-s, z-v_2)\, dz \\
&\le K(r-s)^{\frac{\eta-\gamma-1}{\alpha}} |v_1 - v_2|^\gamma.
\end{aligned}$$

Now assume that $|v_1 - v_2| > (r-s)^{\frac{1}{\alpha}}$. The gradient estimate (8.90) yields

$$\begin{aligned}
|I_1| &= \Bigg| \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} b(r,x,[X_r^{s,\mu,(m)}])(z) - \frac{\delta}{\delta m} b(r,x,[X_r^{s,\mu,(m)}])(v_1) \right) \partial_x p_m(\mu, s, r, v_1, z)\, dz \\
&\quad + \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} b(r,x,[X_r^{s,\mu,(m)}])(z) - \frac{\delta}{\delta m} b(r,x,[X_r^{s,\mu,(m)}])(v_2) \right) \partial_x p_m(\mu, s, r, v_2, z)\, dz \Bigg| \\
&\le K(r-s)^{\frac{\eta-1}{\alpha}} \\
&\le K(r-s)^{\frac{\eta-1-\gamma}{\alpha}} |v_1 - v_2|^\gamma.
\end{aligned}$$

We have thus shown that

$$|I_1| \le K(r-s)^{\frac{\eta-1-\gamma}{\alpha}} |v_1 - v_2|^\gamma.$$

Then, we have

$$\begin{aligned}
|I_2| \le K &\int_{\mathbb{R}^{2d}} |z - x'|^\eta (r-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}} |v_1 - v_2|^\gamma \rho^0(r-s, z-x') \\
&\quad \left(\sum_{k=1}^m C^k (r-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + \frac{\eta-1-\gamma}{\alpha} + j\left(1 + \frac{\eta-1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \right) dz\, d\mu(x') \\
\le K&(r-s)^{\frac{\eta-1-\gamma}{\alpha}+1+\frac{\eta-1}{\alpha}} |v_1 - v_2|^\gamma \\
&\quad \left(\sum_{k=1}^m C^k (r-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(1 + \frac{\eta-1-\gamma}{\alpha} + j\left(1 + \frac{\eta-1}{\alpha}\right), 1 - \frac{1}{\alpha}\right) \right).
\end{aligned}$$

□

8.8.4 Second part of the induction step

We prove here that the estimates (8.95), (8.96), (8.97), (8.98) and (8.99) hold true.

Proof of (8.95). It follows from (8.183), (8.90) and the convolution inequality (8.234) that for some positive constant K independent of m one has

$$\begin{aligned}
&\left|\Delta_{v_1,v_2}\left(\widehat{p}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,s,t,x,y)(\cdot)\right|\\
&\leq K\int_s^t\int_{\mathbb{R}^d}(r-s)^{\frac{\eta-1-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^1(t-r,y-x)\rho^0(r-s,z-x)\\
&\qquad\left[1+\sum_{k=1}^m C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]dz\,dr\\
&\leq K(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x)\left[1+\sum_{k=1}^m C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\right.\\
&\qquad\left.\prod_{j=1}^{k}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}$$

Starting from the representation formula (8.177), using this inequality, the bound (8.85) for $\mathcal{H}^k_{m+1}$, the convolution inequality (8.234) and finally summing over $k\geq 0$, we find that there exists a constant K independent of m such that

$$\begin{aligned}
&\left|\Delta_{v_1,v_2}\partial_v\frac{\delta}{\delta m}p_{m+1}(\mu,s,t,x,y)(\cdot)\right|\\
&\leq\sum_{k=0}^\infty\left|\Delta_{v_1,v_2}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\otimes\mathcal{H}^k_{m+1}\right)(\mu,s,t,x,y)(\cdot)\right|\\
&\leq K(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x)\left[1+\sum_{k=1}^m C^k(r-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\right.\\
&\qquad\left.\prod_{j=1}^{k}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\
&\leq(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x)\\
&\qquad\left(\sum_{k=1}^{m+1}C^k(t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-1-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right).
\end{aligned}$$

provided that we choose $C\geq K$ in (8.95). It ends the proof of the induction step for (8.95).

Proof of (8.96)**.** Let us first assume that $|v_1-v_2|\geq(t-s)^{\frac{1}{\alpha}}$. Using (8.92) and the fact the series appearing in this bound has a limit when m tends to infinity, one has in this case

$$\begin{aligned}
\left|\Delta_{v_1,v_2}\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(\cdot)\right|&\leq C(t-s)^{1-\frac{1}{\alpha}}\rho^0(t-s,y-x)\\
&\leq C(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x).
\end{aligned}$$

In the case where $|v_1 - v_2| < (t-s)^{\frac{1}{\alpha}}$, the mean value theorem and (8.93) yield

$$\begin{aligned}\left|\Delta_{v_1,v_2}\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(\cdot)\right| &\le C(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|\rho^0(t-s,y-x)\\ &\le C(t-s)^{\frac{\eta}{\alpha}-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x)\\ &\le C(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}|v_1-v_2|^\gamma\rho^0(t-s,y-x).\end{aligned}$$

Proof of (8.97)**.** We first note that since (8.92) has been proved before, (8.160) ensures that there exists a positive constant C such that for all $m \ge 1$

$$\left|\frac{\delta}{\delta m}\mathcal{H}_m(\mu,s,r,t,x,y)(v)\right| \le C(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x).$$

Using this inequality, (8.91) and the convolution inequality (8.234), we obtain that

$$\begin{aligned}&\left|\Delta_{x_1,x_2}\left[p_m\otimes\frac{\delta}{\delta m}\mathcal{H}_m\right](\mu,s,r,t,\cdot,y)(v)\right|\\ &=\left|\int_r^t\int_{\mathbb{R}^d}\Delta_{x_1,x_2}p_m(\mu,r,r',\cdot,z)\frac{\delta}{\delta m}\mathcal{H}_m(\mu,s,r',t,z,y)\,dz\,dr'\right|\\ &\le C\int_r^t\int_{\mathbb{R}^d}(r'-r)^{-\frac{\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(r'-r,z-x_1)+\rho^0(r'-r,z-x_2)\right](t-r')^{-\frac{1}{\alpha}}\rho^1(t-r',y-z)\,dz\,dr'\\ &\le C(t-r)^{1-\frac{1+\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(t-r,y-x_1)+\rho^0(t-r,y-x_2)\right].\end{aligned}$$

We conclude by the representation formula (8.176) and (8.85) (since the series appearing is convergent) that we have

$$\begin{aligned}\left|\Delta_{x_1,x_2}\frac{\delta}{\delta m}p_m(\mu,s,t,\cdot,y)(v)\right| &\le \sum_{k=0}^{\infty}\left|\Delta_{x_1,x_2}\left[p_m\otimes\frac{\delta}{\delta m}\mathcal{H}_m\otimes\mathcal{H}_m^k\right](\mu,s,t,\cdot,y)(v)\right|\\ &\le C(t-s)^{1-\frac{1+\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(t-s,y-x_1)+\rho^0(t-s,y-x_2)\right].\end{aligned}$$

Proof of (8.98)**.** We first note that since (8.93) has been proved before, (8.161) ensures that there exists a positive constant C such that for all $m \ge 1$

$$\left|\partial_v\frac{\delta}{\delta m}\mathcal{H}_m(\mu,s,r,t,x,y)(v)\right| \le C(r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x).$$

Using this inequality, (8.91), the convolution inequality (8.234) and the fact that $\gamma < \alpha-1-\eta$, we obtain that

$$\begin{aligned}&\left|\Delta_{x_1,x_2}\left[p_m\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_m\right](\mu,s,s,t,\cdot,y)(v)\right|\\ &=\left|\int_s^t\int_{\mathbb{R}^d}\Delta_{x_1,x_2}p_m(\mu,s,r,\cdot,z)\partial_v\frac{\delta}{\delta m}\mathcal{H}_m(\mu,s,r,t,z,y)\,dz\,dr\right|\\ &\le C\int_s^t\int_{\mathbb{R}^d}(r-s)^{-\frac{\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(r-s,z-x_1)+\rho^0(r-s,z-x_2)\right](r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\ &\le C(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(t-r,y-x_1)+\rho^0(t-r,y-x_2)\right].\end{aligned}$$

We conclude by the representation formula (8.177) and (8.85) that we have

$$\begin{aligned}\left|\Delta_{x_1,x_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,s,t,\cdot,y)(v)\right| &\leq \sum_{k=0}^{\infty}\left|\Delta_{x_1,x_2}\left[p_m\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_m\otimes\mathcal{H}_m^k\right](\mu,s,t,\cdot,y)(v)\right| \\ &\leq C(t-s)^{\frac{\eta-1-\gamma}{\alpha}+1-\frac{1}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(t-r,y-x_1)+\rho^0(t-r,y-x_2)\right].\end{aligned}$$

Proof of (8.99). Let us first assume that $W_1(\mu_1,\mu_2)\leq(t-s)^{\frac{1}{\alpha}}$. In this case, by definition of the linear derivative, the Kantorovich-Rubinstein theorem, and (8.93), we get

$$\begin{aligned}|\Delta_{\mu_1,\mu_2}p_m(\cdot,s,t,x,y)| &= \left|\int_0^1\int_{\mathbb{R}^d}\frac{\delta}{\delta m}p_m(\lambda\mu_1+(1-\lambda)\mu_2,s,t,x,y)(v)\,d(\mu_1-\mu_2)(v)\,d\lambda\right| \\ &\leq \sup_{\lambda\in[0,1],\,v\in\mathbb{R}^d}\left|\partial_v\frac{\delta}{\delta m}p_m(\lambda\mu_1+(1-\lambda)\mu_2,s,t,x,y)(v)\right|W_1(\mu_1,\mu_2) \\ &\leq C(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}W_1(\mu_1,\mu_2)\rho^0(t-s,y-x) \\ &\leq C(t-s)^{1-\frac{1+\gamma}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x).\end{aligned}$$

In the case where $W_1(\mu_1,\mu_2)>(t-s)^{\frac{1}{\alpha}}$, the parametrix expansion (8.89), (8.235) and (8.87) yield

$$\begin{aligned}|\Delta_{\mu_1,\mu_2}p_m(\cdot,s,t,x,y)| &\leq |\Delta_{\mu_1,\mu_2}[\widehat{p}\otimes\Phi_m](\cdot,s,t,x,y)| \\ &\leq |\widehat{p}\otimes\Phi_m(\mu_1,s,t,x,y)|+|\widehat{p}\otimes\Phi_m(\mu_2,s,t,x,y)| \\ &\leq C(t-s)^{1-\frac{1}{\alpha}}\rho^0(t-s,y-x) \\ &\leq C(t-s)^{1-\frac{1+\gamma}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x).\end{aligned}$$

8.8.5 Preparatory technical results

Lemma 8.31.
- *For any $\gamma\in[\eta,1]$, there exists a positive constant K such that for all $m\geq 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s<t\leq T$, $x\in\mathbb{R}^d$*

$$|b(t,x,[X_t^{s,\mu_1,(m)}])-b(t,x,[X_t^{s,\mu_2,(m)}])|\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}}W_1^\gamma(\mu_1,\mu_2). \tag{8.184}$$

- *For any $\gamma\in[\eta,1]$, there exists a positive constant K such that for all $m\geq 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s\leq r<t\leq T$, $x,y\in\mathbb{R}^d$*

$$|\mathcal{H}_m(\mu_1,s,r,t,x,y)-\mathcal{H}_m(\mu_2,s,r,t,x,y)|\leq K(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x). \tag{8.185}$$

- *For any $\gamma\in[\eta,1]$, there exists a positive constant K such that for all $m\geq 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s\leq r<t\leq T$, $x,y\in\mathbb{R}^d$*

$$|\Phi_m(\mu_1,s,r,t,x,y)-\Phi_m(\mu_2,s,r,t,x,y)|\leq K(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x). \tag{8.186}$$

Proof of Lemma 8.31. **Proof of (8.184) and (8.185).** By definition of the linear derivative and denoting by $m_l:=l[X_t^{s,\mu_1,(m)}]+(1-l)[X_t^{s,\mu_2,(m)}]$, we have

$$
\begin{aligned}
&b(t,x,[X_t^{s,\mu_1,(m)}]) - b(t,x,[X_t^{s,\mu_2,(m)}])\\
&= \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,y') - p_m(\mu_2,s,t,y')]\, dy'\, dl\\
&= \int_0^1 \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,m_l)(y') p_m(\mu_1,s,t,x',y')\, dy'\, d(\mu_1-\mu_2)(x') dl\\
&\quad + \int_0^1 \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_2,s,t,x',y')]\, dy'\, d\mu_2(x') dl\\
&=: I_1 + I_2.
\end{aligned}
$$

For I_1, we need to control, for $x', x'' \in \mathbb{R}^d$

$$
\int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'.
$$

In the case where $|x'-x''| \le (t-s)^{\frac{1}{\alpha}}$, we write

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'\\
&= \int_0^1 \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t,x,m_l)(y') - \frac{\delta}{\delta m} b(t,x,m_l)(l'x' + (1-l')x'')\right]\\
&\qquad\qquad \partial_x p_m(\mu_1,s,t,l'x'+(1-l')x'',y') \cdot (x'-x'')\, dy'\, dl'.
\end{aligned}
$$

Using the η-Hölder continuity of $\frac{\delta}{\delta m} b$ and (8.90), we obtain that

$$
\begin{aligned}
&\left|\int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'\right|\\
&\le K(t-s)^{\frac{\eta-1}{\alpha}} |x'-x''|\\
&\le K(t-s)^{\frac{\eta-\gamma}{\alpha}} |x'-x''|^{\gamma}.
\end{aligned}
$$

Assume now that $|x'-x''| > (t-s)^{\frac{1}{\alpha}}$. One can write

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'\\
&= \frac{\delta}{\delta m} b(t,x,m_l)(x') - \frac{\delta}{\delta m} b(t,x,m_l)(x'')\\
&\quad + \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t,x,m_l)(y') - \frac{\delta}{\delta m} b(t,x,m_l)(x')\right] p_m(\mu_1,s,t,x',y')\, dy'\\
&\quad + \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t,x,m_l)(y') - \frac{\delta}{\delta m} b(t,x,m_l)(x'')\right] p_m(\mu_1,s,t,x'',y')\, dy'
\end{aligned}
$$

It follows from the η-Hölder continuity of $\frac{\delta}{\delta m}b$, (8.90) and the space-time inequality (8.231) that

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,m_l)(y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\,dy' \\
&\leq K(|x'-x''|^\eta + (t-s)^{\frac{\eta}{\alpha}}) \\
&\leq K(|x'-x''|^\eta + (t-s)^{\frac{\eta}{\alpha}}) \\
&\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}}|x'-x''|^\gamma,
\end{aligned}
$$

since $\gamma \geq \eta$. Jensen's inequality yields

$$
|I_1| \leq K(t-s)^{\frac{\eta-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
$$

It remains to study I_2 which can be rewritten as

$$
I_2 = \int_0^1 \int_{\mathbb{R}^{2d}} \left[\frac{\delta}{\delta m} b(t,x,m_l)(y') - \frac{\delta}{\delta m} b(t,x,m_l)(x')\right] [p_m(\mu_1,s,t,x',y') - p_m(\mu_2,s,t,x',y')]\,dy'\,d\mu_2(x')dl. \tag{8.187}
$$

Thanks to (8.99), (8.90) and the space-time inequality (8.231), one has since $\alpha \in (1,2)$

$$
\begin{aligned}
|I_2| &\leq K(t-s)^{1-\frac{1+\gamma}{\alpha}+\frac{\eta}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \\
&\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned}
$$

This concludes the proof of (8.184). The proof of (8.185) immediately follows from the definition (8.82) of $\mathcal{H}_m$ and (8.235).

Proof of (8.186)**.** We prove by induction of $k \geq 2$ that there exists a constant K independent of k such that for any $k \geq 2$, $m \geq 1$, $\mu_1, \mu_2 \in \mathcal{P}_\beta(\mathbb{R}^d)$, $0 \leq s \leq r < t \leq T$, $x, y \in \mathbb{R}^d$

$$
\begin{aligned}
|\Delta_{\mu_1,\mu_2}\mathcal{H}_m^k(\cdot,s,r,t,x,y)| \leq (r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x) \\
\sum_{i=1}^{k-1} C^i (t-r)^{i(1-\frac{1}{\alpha})} \prod_{j=1}^{i} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right).
\end{aligned} \tag{8.188}
$$

We do not prove the base case $k = 2$ since it relies on the same computations as the induction step using (8.185). For the induction step, assume that (8.188) holds for $\mathcal{H}_m^k$ and let us prove it for $\mathcal{H}_m^{k+1}$. One has

$$
\begin{aligned}
|\Delta_{\mu_1,\mu_2}\mathcal{H}_m^{k+1}(\cdot,s,r,t,x,y)| &\leq \left|\int_r^t \int_{\mathbb{R}^d} \mathcal{H}_m(\mu_1,s,r,r',x,z)\Delta_{\mu_1,\mu_2}\mathcal{H}_m^k(\cdot,s,r',t,z,y)\,dz\,dr'\right| \\
&\quad + \left|\int_r^t \int_{\mathbb{R}^d} \Delta_{\mu_1,\mu_2}\mathcal{H}_m(\cdot,s,r,r',x,z)\mathcal{H}_m^k(\mu_2,s,r',t,z,y)\,dz\,dr'\right| \\
&=: I_1 + I_2.
\end{aligned}
$$

By (8.185) and the bound of $\mathcal{H}_m^k$ (8.85) (note that the series appearing in this bound in convergent and thus can be bounded independently of k and m), we have

$$\begin{aligned}
I_2 &\le K\int_r^t\int_{\mathbb{R}^d}(r-s)^{\frac{\eta-\gamma}{\alpha}}(r'-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(r'-r,z-x)(t-r')^{-\frac{1}{\alpha}}\rho^1(t-r',y-z)\,dz\,dr'\\
&\le K(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x)\mathcal{B}\left(1-\frac{1}{\alpha},1-\frac{1}{\alpha}\right).
\end{aligned}$$

By induction and using (8.85), one has

$$\begin{aligned}
I_1 \le{}& K\int_r^t\int_{\mathbb{R}^d}(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r,z-x)(r'-s)^{\frac{\eta-\gamma}{\alpha}}(t-r')^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r',y-z)\\
&\qquad\sum_{i=1}^{k-1}C^i(t-r')^{i(1-\frac{1}{\alpha})}\prod_{j=1}^{i}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\,dz\,dr'\\
\le{}& K(r-s)^{\frac{\eta-\gamma}{\alpha}}\int_r^t\int_{\mathbb{R}^d}(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r,z-x)(t-r')^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r',y-z)\\
&\qquad\sum_{i=1}^{k-1}C^i(t-r')^{i(1-\frac{1}{\alpha})}\prod_{j=1}^{i}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\,dz\,dr'\\
\le{}& K(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x)\sum_{i=1}^{k-1}C^i(t-r)^{(i+1)(1-\frac{1}{\alpha})}\prod_{j=1}^{i+1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)
\end{aligned}$$

Summing the bounds obtained for I_1 and I_2, we have proved that

$$\begin{aligned}
|\Delta_{\mu_1,\mu_2}\mathcal{H}_m^{k+1}(\cdot,s,r,t,x,y)| \le{}& (r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x)\\
&\sum_{i=1}^{k}C^i(t-r)^{i(1-\frac{1}{\alpha})}\prod_{j=1}^{i}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right),
\end{aligned}$$

if we choose $C\ge K$ since the beginning. Finally, summing (8.188) over $k\ge 1$ yields (8.186).

□

Lemma 8.32.
- *For all $\gamma\in[\eta,1]$, there exists $K>0$ such that for all $m\ge 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\le s<t\le T$, $x,v\in\mathbb{R}^d$*

$$\left|\frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu_1,(m)}])(v)-\frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu_2,(m)}])(v)\right|\le K(t-s)^{\frac{\eta-\gamma}{\alpha}}W_1^\gamma(\mu_1,\mu_2). \tag{8.189}$$

- *For all $\gamma\in(0,1]$, there exists $K>0$ such that for all $m\ge 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\le s<t\le T$, $x,v\in\mathbb{R}^d$*

$$\begin{aligned}
\left|\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v)-\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right|&\le K(t-s)^{-\frac{\gamma}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\\
&+K\int_{\mathbb{R}^{2d}}\left|\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}p_m(\cdot,s,t,x',y')\right|\,dy'\,d\mu_2(x').
\end{aligned} \tag{8.190}$$

- *For all $\gamma\in(0,1]\cap(0,\eta+\alpha-1)$, there exists $K>0$ such that for all $m\ge 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\le s<t\le T$, $x,v\in\mathbb{R}^d$*

$$\left|\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v) - \partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right| \leq K(t-s)^{\frac{\eta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)$$
$$+ K\int_{\mathbb{R}^{2d}} (1\wedge |x'-y'|^\eta)\left|\Delta_{\mu_1,\mu_2}\partial_v \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y')\right| dy'\, d\mu_2(x'). \quad (8.191)$$

- *For all $\gamma\in(0,1]$, there exists $K>0$ such that for all $m\geq 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s<t\leq T$, $x,v\in\mathbb{R}^d$*

$$\begin{aligned}
&\left|\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_1,s,r,t,x,y)(v) - \frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_2,s,r,t,x,y)(v)\right| \\
&\leq K(r-s)^{-\frac{\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x) \qquad (8.192)\\
&\quad + K(t-r)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^{2d}}\left|\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}p_m(\cdot,s,r,x',y')\right| dy'\, d\mu_2(x')\rho^1(t-r,y-x).
\end{aligned}$$

- *For all $\gamma\in(0,1]\cap(0,\eta+\alpha-1)$, there exists $K>0$ such that for all $m\geq 1$, $\mu_1,\mu_2\in\mathcal{P}_\beta(\mathbb{R}^d)$, $0\leq s<t\leq T$, $x,v\in\mathbb{R}^d$*

$$\begin{aligned}
&\left|\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_1,s,r,t,x,y)(v) - \partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_1,s,r,t,x,y)(v)\right| \\
&\leq K(r-s)^{\frac{\eta-\gamma-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-x) \qquad (8.193)\\
&\quad + K(t-r)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^{2d}}(1\wedge|x'-y'|^\eta)\left|\Delta_{\mu_1,\mu_2}\partial_v\frac{\delta}{\delta m}p_m(\cdot,s,r,x',y')\right| dy'\, d\mu_2(x')\rho^1(t-r,y-x).
\end{aligned}$$

Proof of Lemma 8.32. **Proof of (8.189).** By definition of the linear derivative and denoting by $m_l := l[X_t^{s,\mu_1,(m)}] + (1-l)[X_t^{s,\mu_2,(m)}]$, we have

$$\begin{aligned}
&\frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu_1,(m)}])(v) - \frac{\delta}{\delta m}b(t,x,[X_t^{s,\mu_2,(m)}])(v)\\
&= \int_0^1\int_{\mathbb{R}^d}\frac{\delta^2}{\delta m^2}b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,y') - p_m(\mu_2,s,t,y')]\, dy'\, dl\\
&= \int_0^1\int_{\mathbb{R}^{2d}}\frac{\delta^2}{\delta m^2}b(t,x,m_l)(v,y')p_m(\mu_1,s,t,x',y')\, dy'\, d(\mu_1-\mu_2)(x')dl\\
&\quad + \int_0^1\int_{\mathbb{R}^{2d}}\frac{\delta^2}{\delta m^2}b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_2,s,t,x',y')]\, dy'\, d\mu_2(x')dl\\
&=: I_1 + I_2.
\end{aligned}$$

For I_1, we need to control, for $x',x''\in\mathbb{R}^d$

$$\int_{\mathbb{R}^d}\frac{\delta^2}{\delta m^2}b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'.$$

In the case where $|x'-x''|\leq (t-s)^{\frac{1}{\alpha}}$, we write

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy' \\
&= \int_0^1 \int_{\mathbb{R}^d} \left[\frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y') - \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,l'x' + (1-l')x'')\right] \\
&\qquad\qquad \partial_x p_m(\mu_1,s,t,l'x' + (1-l')x'',y') \cdot (x'-x'')\, dy'\, dl'.
\end{aligned}$$

Using the η-Hölder continuity of $\frac{\delta^2}{\delta m^2} b(t,x,\mu)(v,\cdot)$ and (8.90), we obtain that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy'\right| \\
&\leq K(t-s)^{\frac{\eta-1}{\alpha}} |x'-x''| \\
&\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}} |x'-x''|^{\gamma}.
\end{aligned}$$

Assume now that $|x'-x''| > (t-s)^{\frac{1}{\alpha}}$. One can write

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy' \\
&= \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,x') - \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,x'') \\
&\quad + \int_{\mathbb{R}^d} \left[\frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y') - \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,x')\right] p_m(\mu_1,s,t,x',y')\, dy' \\
&\quad + \int_{\mathbb{R}^d} \left[\frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y') - \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,x'')\right] p_m(\mu_1,s,t,x'',y')\, dy'
\end{aligned}$$

It follows from the η-Hölder continuity of $\frac{\delta^2}{\delta m^2} b(t,x,\mu)(v,\cdot)$, (8.90) and the space-time inequality (8.231) that

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y')[p_m(\mu_1,s,t,x',y') - p_m(\mu_1,s,t,x'',y')]\, dy' \\
&\leq K(|x'-x''|^{\eta} + (t-s)^{\frac{\eta}{\alpha}}) \\
&\leq K(|x'-x''|^{\eta} + (t-s)^{\frac{\eta}{\alpha}}) \\
&\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}} |x'-x''|^{\gamma},
\end{aligned}$$

since $\gamma \geq \eta$. Jensen's inequality yields

$$|I_1| \leq K(t-s)^{\frac{\eta-\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2).$$

It remains to study I_2 which can be rewritten as

$$I_2 = \int_0^1 \int_{\mathbb{R}^{2d}} \left[\frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,y') - \frac{\delta^2}{\delta m^2} b(t,x,m_l)(v,x')\right] [p_m(\mu_1,s,t,x',y') - p_m(\mu_2,s,t,x',y')]\, dy'\, d\mu_2(x') dl. \tag{8.194}$$

Thanks to (8.99), (8.90) and the space-time inequality (8.231), one has since $\alpha \in (1,2)$

$$\begin{aligned}|I_2| &\le K(t-s)^{1-\frac{1+\gamma}{\alpha}+\frac{\eta}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2)\\ &\le K(t-s)^{\frac{\eta-\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2).\end{aligned}$$

This concludes the proof of (8.189).

Proof of (8.190)**.** We begin to treat the case where $W_1(\mu_1,\mu_2) \ge (t-s)^{\frac{1}{\alpha}}$. Using (8.158) and noting that the series appearing in the bound is convergent, one has

$$\begin{aligned}\left|\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\cdot,(m)}])\right](v)\right| &\le \left|\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v)\right| + \left|\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right|\\ &\le K\\ &\le K(t-s)^{-\frac{\gamma}{\alpha}} W_1(\mu_1,\mu_2)^{\gamma}.\end{aligned}$$

We now focus on the case where $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$. Using (8.162), one has

$$\begin{aligned}&\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\cdot,(m)}])\right](v)\\ &= \int_{\mathbb{R}^d} \Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m} b(t,x,[X_t^{s,\cdot,(m)}])(y) p_m(\mu_1,s,t,v,y)\,dy\\ &\quad + \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) \Delta_{\mu_1,\mu_2} p_m(\cdot,s,t,v,y)\,dy\\ &\quad + \int_{\mathbb{R}^{2d}} \Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m} b(t,x,[X_t^{s,\cdot,(m)}])(y) \frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v)\,dy\,d\mu_1(x')\\ &\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) \frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v)\,dy\,d(\mu_1-\mu_2)(x')\\ &\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) \Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m} p_m(\cdot,s,t,x',y)(v)\,dy\,d\mu_2(x')\\ &=: I_1+I_2+I_3+I_4+I_5.\end{aligned}$$

Thanks to (8.90) and (8.189) which is true for all $\gamma \in (0,1]$ since $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$, we obtain that

$$\begin{aligned}|I_1| &\le K(t-s)^{\frac{\eta-\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2)\\ &\le K(t-s)^{-\frac{\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2).\end{aligned}$$

It follows from the boundedness of $\frac{\delta}{\delta m} b$ and (8.99) that

$$\begin{aligned}|I_2| &\le K(t-s)^{1-\frac{1+\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2)\\ &\le K(t-s)^{-\frac{\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2).\end{aligned}$$

Owing to (8.189) which is true for all $\gamma \in (0,1]$ since $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$ and (8.92) (the series appearing in this estimates is convergent), one has

$$\begin{aligned}|I_3| &\leq K(t-s)^{\frac{\eta-\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\\ &\leq K(t-s)^{-\frac{\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).\end{aligned}$$

Let us now deal with I_4. We need to control

$$J := \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) \left[\frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v) - \frac{\delta}{\delta m} p_m(\mu_1,s,t,x'',y)(v)\right] dy.$$

Thanks to (8.97) and the boundedness $\frac{\delta}{\delta m} b$, we obtain that

$$\begin{aligned}|J| &\leq K \int_{\mathbb{R}^d} (t-s)^{1-\frac{1+\gamma}{\alpha}} |x'-x''|^\gamma \left[\rho^0(t-s,y-x') + \rho^0(t-s,y-x'')\right] dy\\ &\leq K(t-s)^{1-\frac{1+\gamma}{\alpha}} |x'-x''|^\gamma.\end{aligned}$$

Jensen's inequality implies that

$$\begin{aligned}|I_4| &\leq K(t-s)^{1-\frac{1+\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\\ &\leq K(t-s)^{-\frac{\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2).\end{aligned}$$

Finally, for I_5, one has

$$\begin{aligned}|I_5| &= \left|\int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) \Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y)(v)\, dy\, d\mu_2(x')\right|\\ &\leq K \int_{\mathbb{R}^{2d}} \left|\Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y)\right| dy\, d\mu_2(x').\end{aligned}$$

Gathering all the previous estimates, we have proved that

$$\begin{aligned}\left|\frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v) - \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right| &\leq K(t-s)^{-\frac{\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\\ &\quad + K\int_{\mathbb{R}^d} \left|\Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y')\right| dy\, d\mu_2(x').\end{aligned}$$

Proof of (8.191)**.** We begin to treat the case where $W_1(\mu_1,\mu_2) \geq (t-s)^{\frac{1}{\alpha}}$. Using (8.159) and noting that the series appearing in the bound is convergent, one has

$$\begin{aligned}\left|\Delta_{\mu_1,\mu_2} \partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\cdot,(m)}])\right](v)\right| &\leq \left|\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v)\right| + \left|\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right|\\ &\leq K(t-s)^{\frac{\eta-1}{\alpha}}\\ &\leq K(t-s)^{\frac{\eta-1-\gamma}{\alpha}} W_1(\mu_1,\mu_2)^\gamma.\end{aligned}$$

We now focus on the case where $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$. In this case, by (8.163), we have

$$
\begin{aligned}
&\Delta_{\mu_1,\mu_2}\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\cdot,(m)}])\right](v)\\
&= \int_{\mathbb{R}^d} \Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m} b(t,x,[X_t^{s,\cdot,(m)}])(y)\partial_x p_m(\mu_1,s,t,v,y)\,dy\\
&\quad + \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y)\Delta_{\mu_1,\mu_2}\partial_x p_m(\cdot,s,t,v,y)\,dy\\
&\quad + \int_{\mathbb{R}^{2d}} \Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m} b(t,x,[X_t^{s,\cdot,(m)}])(y)\partial_v \frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v)\,dy\,d\mu_1(x')\\
&\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y)\partial_v \frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v)\,dy\,d(\mu_1-\mu_2)(x')\\
&\quad + \int_{\mathbb{R}^{2d}} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y)\Delta_{\mu_1,\mu_2}\partial_v \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y)(v)\,dy\,d\mu_2(x')\\
&=: I_1+I_2+I_3+I_4+I_5.
\end{aligned}
$$

Thanks to (8.90) and (8.189) which is true for all $\gamma \in (0,1]$ since $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$, we obtain that

$$|I_1| \le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2).$$

Concerning I_2, it can be rewritten as

$$I_2 = \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y) - \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(v)\right)\Delta_{\mu_1,\mu_2}\partial_x p_m(\cdot,s,t,v,y)\,dy.$$

It follows from the η-Hölder continuity of $\frac{\delta}{\delta m}b(t,x,\mu)(\cdot)$, (8.100) since $\gamma \in (0,\eta+\alpha-1)$ and the space-time inequality (8.231) that

$$
\begin{aligned}
|I_2| &\le K(t-s)^{\frac{\eta}{\alpha}+1-\frac{1+\gamma+1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\\
&\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned}
$$

Owing to (8.189) which is true for all $\gamma \in (0,1]$ since $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$ and (8.93) (the series appearing in this estimates is convergent), one has

$$
\begin{aligned}
|I_3| &\le K(t-s)^{\frac{\eta-\gamma}{\alpha}+\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\\
&\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2).
\end{aligned}
$$

Let us now deal with I_4. We need to control

$$J := \int_{\mathbb{R}^d} \frac{\delta}{\delta m} b(t,x,[X_t^{s,\mu_2,(m)}])(y)\left[\partial_v \frac{\delta}{\delta m} p_m(\mu_1,s,t,x',y)(v) - \partial_v \frac{\delta}{\delta m} p_m(\mu_1,s,t,x'',y)(v)\right]dy.$$

Let us prove that

$$|J| \le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1+\frac{\eta-1}{\alpha}} |x'-x''|^\gamma.$$

Assume first that $|x'-x''| \le (t-s)^{\frac{1}{\alpha}}$. In this case using (8.177) and (8.164), we have

$\int_{\mathbb{R}^d} \partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x', y)(v)\, dy = 0$. We can thus rewrite

$$J = \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(y) - \frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(x'') \right] \\ \left[\partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x', y)(v) - \partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x'', y)(v) \right] dy$$

Thanks to (8.98), (8.232) since $|x' - x''| \le (t-s)^{\frac{1}{\alpha}}$ and the η-Hölder continuity of $\frac{\delta}{\delta m} b(t, x, \mu)(\cdot)$, we obtain that

$$|J| \le K \int_{\mathbb{R}^d} |y - x''|^\eta (t-s)^{\frac{\eta - 1 - \gamma}{\alpha} + 1 - \frac{1}{\alpha}} |x' - x''|^\gamma \rho^1(t-s, y - x'')\, dy.$$

The space-time inequality (8.231) yields

$$|J| \le K (t-s)^{\frac{\eta}{\alpha} + \frac{\eta - 1 - \gamma}{\alpha} + 1 - \frac{1}{\alpha}} |x' - x''|^\gamma.$$

Assume now that $|x' - x''| > (t-s)^{\frac{1}{\alpha}}$. In this case, we rewrite

$$J = \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(y) - \frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(x') \right] \partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x', y)(v)\, dy \\ + \int_{\mathbb{R}^d} \left[\frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(y) - \frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(x'') \right] \partial_v \frac{\delta}{\delta m} p_m(\mu_1, s, t, x'', y)(v)\, dy.$$

Then, we use (8.93) (the series appearing in the bound being convergent) and the η-Hölder continuity of $\frac{\delta}{\delta m} b(t, x, \mu)(\cdot)$ which yield

$$\begin{aligned} |J| &\le K (t-s)^{\frac{\eta}{\alpha} + \frac{\eta - 1}{\alpha} + 1 - \frac{1}{\alpha}} \\ &\le K (t-s)^{\frac{\eta}{\alpha} + \frac{\eta - \gamma - 1}{\alpha} + 1 - \frac{1}{\alpha}} |x' - x''|^\gamma. \end{aligned}$$

Jensen's inequality implies that

$$\begin{aligned} |I_4| &\le K (t-s)^{\frac{\eta - \gamma - 1}{\alpha} + 1 + \frac{\eta - 1}{\alpha}} W_1^\gamma(\mu_1, \mu_2) \\ &\le K (t-s)^{\frac{\eta - \gamma - 1}{\alpha}} W_1^\gamma(\mu_1, \mu_2). \end{aligned}$$

Finally, for I_5, one has

$$\begin{aligned} |I_5| &= \left| \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(y) - \frac{\delta}{\delta m} b(t, x, [X_t^{s,\mu_2,(m)}])(x') \right) \Delta_{\mu_1,\mu_2} \partial_v \frac{\delta}{\delta m} p_m(\cdot, s, t, x', y)(v)\, dy\, d\mu_2(x') \right| \\ &\le K \int_{\mathbb{R}^{2d}} (1 \wedge |x' - y|^\eta) \left| \Delta_{\mu_1,\mu_2} \partial_v \frac{\delta}{\delta m} p_m(\cdot, s, t, x', y) \right| dy\, d\mu_2(x'). \end{aligned}$$

Gathering all the previous estimates, we have proved that

$$\left|\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_1,(m)}])\right](v) - \partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{s,\mu_2,(m)}])\right](v)\right| \leq K(t-s)^{\frac{\eta-\gamma-1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)$$
$$+ K \int_{\mathbb{R}^d} (1 \wedge |x'-y|^\eta) \left|\Delta_{\mu_1,\mu_2} \partial_v \frac{\delta}{\delta m} p_m(\cdot,s,t,x',y')\right| dy\, d\mu_2(x').$$

Proof of (8.192) and (8.193). Both estimates are immediate consequences of (8.190) and (8.191) using the expression of $\frac{\delta}{\delta m}\mathcal{H}_{m+1}$ given by (8.164) and (8.90).

□

8.8.6 Third part of the induction step

We prove here that the estimates (8.100), (8.101), and (8.102) hold true.

Proof of (8.100). Assume first that $W_1(\mu_1,\mu_2) > (t-s)^{\frac{1}{\alpha}}$. In this case, the parametrix expansion (8.89) (which can be differentiated with respect to x), (8.235) and (8.87) yield

$$\begin{aligned}
\left|\Delta_{\mu_1,\mu_2} \partial_x^j p_m(\cdot,s,t,x,y)\right| &\leq \left|\Delta_{\mu_1,\mu_2}\left[\partial_x^j \widehat{p} \otimes \Phi_m\right](\cdot,s,t,x,y)\right| \\
&\leq \left|\partial_x^j \widehat{p} \otimes \Phi_m(\mu_1,s,t,x,y)\right| + \left|\partial_x^j \widehat{p} \otimes \Phi_m(\mu_2,s,t,x,y)\right| \\
&\leq C(t-s)^{1-\frac{1+j}{\alpha}} \rho^j(t-s,y-x) \\
&\leq C(t-s)^{1-\frac{1+j+\gamma}{\alpha}} W_1^\gamma(\mu_1,\mu_2) \rho^j(t-s,y-x).
\end{aligned}$$

We now focus on the case where $W_1(\mu_1,\mu_2) \leq (t-s)^{\frac{1}{\alpha}}$. Let us prove that the following representation formula holds true

$$\Delta_{\mu_1,\mu_2} \partial_x^j p_{m+1}(\cdot,s,t,x,y) = \sum_{k=0}^{\infty} \left(\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}\right) \otimes \mathcal{H}_{m+1}^k(\mu_2,s,t,x,y). \tag{8.195}$$

Indeed, using the representation formula (8.89) which can be differentiated with respect to $x \in \mathbb{R}^d$, we get that

$$\Delta_{\mu_1,\mu_2} \partial_x^j p_{m+1}(\cdot,s,t,x,y) = \partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}(\mu_2,s,t,x,y) + \Delta_{\mu_1,\mu_2}\partial_x^j p_{m+1} \otimes \mathcal{H}_{m+1}(\mu_1,s,t,x,y).$$

Hence, we easily prove by induction that for all $n \geq 1$, one has

$$\Delta_{\mu_1,\mu_2} \partial_x^j p_{m+1}(\cdot,s,t,x,y) = \sum_{k=0}^{n} \left(\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}\right) \otimes \mathcal{H}_{m+1}^k(\mu_2,s,t,x,y)$$
$$+ \Delta_{\mu_1,\mu_2}\partial_x^j p_{m+1} \otimes \mathcal{H}_{m+1}^{n+1}(\mu_1,s,t,x,y). \tag{8.196}$$

Thanks to (8.90) and (8.85), we deduce that $\Delta_{\mu_1,\mu_2}\partial_x^j p_{m+1} \otimes \mathcal{H}_{m+1} \otimes \mathcal{H}_{m+1}^{n+1}(\mu_1,s,t,x,y)$ converges to 0 as n tends to infinity. Then, it follows from (8.90), (8.85) and the convolution inequality (8.234) that

$$\begin{aligned}
&\left|\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}(\mu_2,s,t,x,y)\right| \\
&\le \int_s^t \int_{\mathbb{R}^d} \left|\partial_x^j p_{m+1}(\mu_2,s,r,x,z)\Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}(\cdot,s,r,t,z,y)\right| dz\,dr \\
&\le K\int_s^t \int_{\mathbb{R}^d} (r-s)^{-\frac{j}{\alpha}}\rho^j(r-s,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr \\
&\le K(t-s)^{-\frac{j}{\alpha}+1-\frac{1}{\alpha}}\rho^j(t-s,y-x).
\end{aligned}$$

From this estimate and (8.85), we deduce that for any $k \ge 1$

$$\begin{aligned}
\left|\left(\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}\right) \otimes \mathcal{H}_{m+1}^k(\mu_2,s,t,x,y)\right| &\le K(t-s)^{-\frac{j}{\alpha}+1-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\rho^j(t-s,y-x) \\
&\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right),1-\frac{j}{\alpha}+1-\frac{1}{\alpha}\right)\prod_{i=1}^{k-1}\mathcal{B}\left(i\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}$$

Thanks to the asymptotic behavior of the Beta function, we deduce that the series appearing in (8.196) is absolutely convergent, which yields the representation formula (8.195). Now, since $\gamma \in (0,1]$ if $j=0$ and $\gamma \in (0,\eta+\alpha-1)$ if $j=1$, we can chose $\tilde{\gamma} \ge \gamma$ such that $\tilde{\gamma} \in [\eta,1]$ if $j=0$ and $\tilde{\gamma} \in [\eta,\eta+\alpha-1)$ if $j=1$. It follows from (8.90), (8.185), the convolution inequality (8.234) and since $\tilde{\gamma} \in [\eta,\eta+\alpha-1)$ when $j=1$ that

$$\begin{aligned}
&\left|\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}(\mu_2,s,t,x,y)\right| \\
&\le \int_s^t \int_{\mathbb{R}^d} \left|\partial_x^j p_{m+1}(\mu_2,s,r,x,z)\Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}(\cdot,s,r,t,z,y)\right| dz\,dr \\
&\le K\int_s^t \int_{\mathbb{R}^d} (r-s)^{-\frac{j}{\alpha}}\rho^j(r-s,z-x)(r-s)^{\frac{\eta-\tilde{\gamma}}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^{\tilde{\gamma}}(\mu_1,\mu_2)\rho^1(t-r,y-z)\,dz\,dr \\
&\le K(t-s)^{\frac{\eta-\tilde{\gamma}-j}{\alpha}+1-\frac{1}{\alpha}}W_1^{\tilde{\gamma}}(\mu_1,\mu_2)\rho^j(t-s,y-x).
\end{aligned}$$

From this estimate and (8.85), we deduce that for any $k \ge 1$

$$\begin{aligned}
\left|\left(\partial_x^j p_{m+1} \otimes \Delta_{\mu_1,\mu_2}\mathcal{H}_{m+1}\right) \otimes \mathcal{H}_{m+1}^k(\mu_2,s,t,x,y)\right| &\le K(t-s)^{\frac{\eta-\tilde{\gamma}-j}{\alpha}+1-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}W_1^{\tilde{\gamma}}(\mu_1,\mu_2)\rho^j(t-s,y-x) \\
&\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right),1+\frac{\eta-\tilde{\gamma}-j}{\alpha}+1-\frac{1}{\alpha}\right)\prod_{i=1}^{k-1}\mathcal{B}\left(i\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}$$

Summing over $k \ge 0$, we find that, since $\gamma \le \tilde{\gamma}$ and $W_1(\mu_1,\mu_2) < (t-s)^{\frac{1}{\alpha}}$, we have

$$\begin{aligned}
\left|\Delta_{\mu_1,\mu_2}\partial_x^j p_{m+1}(\cdot,s,t,x,y)\right| &\le K(t-s)^{\frac{\eta-\tilde{\gamma}-j}{\alpha}+1-\frac{1}{\alpha}}W_1^{\tilde{\gamma}}(\mu_1,\mu_2)\rho^j(t-s,y-x) \\
&\le K(t-s)^{\frac{\eta-\gamma-j}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^j(t-s,y-x) \\
&\le K(t-s)^{1-\frac{1+\gamma+j}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^j(t-s,y-x).
\end{aligned}$$

This concludes the proof of (8.100).

Proof of (8.101)**.** We first assume that $W_1(\mu_1, \mu_2) > (t-s)^{\frac{1}{\alpha}}$. In this case, using (8.92), one has

$$
\begin{aligned}
&\left|\Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_{m+1}(\cdot, s, t, x, y)(v)\right| \\
&\leq K(t-s)^{1-\frac{1}{\alpha}} \rho^0(t-s, y-x) \\
&\leq K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x) \\
&\leq (t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x) \\
&\qquad\qquad \sum_{k=1}^{m+1} C^k (t-s)^{(k-1)(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(1-\frac{\gamma}{\alpha}+(j-1)\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right),
\end{aligned}
$$

provided that we choose $C \geq K$ in (8.101). We now treat the case $W_1(\mu_1,\mu_2) \leq (t-s)^{\frac{1}{\alpha}}$. By the representation formula (8.176), we have the following decomposition

$$
\begin{aligned}
\Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_{m+1}(\cdot, s, t, x, y)(v) &= p_{m+1} \otimes \Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} \mathcal{H}_{m+1}(\mu_2, s, t, x, y)(v) \\
&\quad + \Delta_{\mu_1,\mu_2} p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}(\mu_1, s, t, x, y)(v) \\
&\quad + \left(\Delta_{\mu_1,\mu_2}\left[p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}\right]\right) \otimes \Phi_{m+1}(\mu_1, s, t, x, y)(v) \\
&\quad + \left[p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}\right] \otimes \Delta_{\mu_1,\mu_2} \Phi_{m+1}(\mu_2, s, t, x, y)(v) \\
&=: I_1 + I_2 + I_3 + I_4.
\end{aligned}
$$

For I_1, using (8.90), (8.192), the induction assumption (8.101) and the convolution inequality (8.234), one has

$$
\begin{aligned}
|I_1| &\leq K \int_s^t \int_{\mathbb{R}^d} \rho^0(r-s, z-x) \Big[(r-s)^{-\frac{\gamma}{\alpha}} (t-r)^{-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^1(t-r, y-z) \\
&\qquad + (t-r)^{-\frac{1}{\alpha}} \int_{\mathbb{R}^{2d}} \left|\Delta_{\mu_1,\mu_2} \frac{\delta}{\delta m} p_m(\cdot, s, r, x', y')\right| dy'\, d\mu_2(x') \rho^1(t-r, y-z)\Big]\, dz\, dr \\
&\leq K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x) \\
&\quad + K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x) \sum_{k=1}^{m} C^k (t-s)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(1-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right).
\end{aligned}
$$

For I_2, it follows from (8.99), (8.160) (the series appearing in the bound being convergent) and the convolution inequality (8.234) that

$$
\begin{aligned}
|I_2| &\leq K \int_s^t \int_{\mathbb{R}^d} (r-s)^{1-\frac{1+\gamma}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(r-s, z-x) (t-r)^{-\frac{1}{\alpha}} \rho^1(t-r, y-z)\, dz\, dr \\
&\leq K(t-s)^{1-\frac{1+\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x) \\
&\leq K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}} W_1^{\gamma}(\mu_1,\mu_2) \rho^0(t-s, y-x).
\end{aligned}
$$

Concerning I_3, we note that it writes

$$
I_3 = (I_1 + I_2) \otimes \Phi_{m+1}(\mu_1, s, t, x, y)(v).
$$

Thus, using the preceding bounds obtained for I_1 and I_2, (8.87) and the convolution inequality (8.234), we get that

$$\begin{aligned}|I_3| &\le K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\
&\qquad\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^k \mathcal{B}\left(1-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\
&\le K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\
&\qquad\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^k \mathcal{B}\left(1-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].\end{aligned}$$

We finally deal with I_4. The convolution inequality (8.234), (8.90) and (8.160) (the series appearing in this bound being convergent) yield

$$\left|p_{m+1}\otimes\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_2,s,t,x,y)(v)\right|\le K(t-s)^{1-\frac{1}{\alpha}}\rho^0(t-s,y-x).$$

Then, it follows from (8.186), which is valid for all $\gamma\in(0,1]$ since $W_1(\mu_1,\mu_2)<(t-s)^{\frac{1}{\alpha}}$ that

$$\begin{aligned}|I_4| &\le K\int_s^t\int_{\mathbb{R}^d}K(r-s)^{1-\frac{1}{\alpha}}\rho^0(r-s,z-x)(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-z)\,dz\,dr\\
&\le K(t-s)^{\frac{\eta-\gamma}{\alpha}+1-\frac{1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\
&\le K(t-s)^{\frac{\eta-\gamma}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x).\end{aligned}$$

Gathering all the previous estimates, we have proved that

$$\begin{aligned}&\left|\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}p_{m+1}(\cdot,s,t,x,y)(v)\right|\\
&\le K(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\
&\qquad\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^k \mathcal{B}\left(1-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\
&\le (t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\sum_{k=1}^{m+1}C^k(t-s)^{(k-1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right),\end{aligned}$$

provided that we choose $C\ge K$ in (8.101). This ends the proof of the induction step for (8.101).

Proof of (8.102)**.** Assume first that $W_1(\mu_1,\mu_2) > (t-s)^{\frac{1}{\alpha}}$. In this case, (8.92) yields

$$\begin{aligned}
&\left|\Delta_{\mu_1,\mu_2}\partial_v \frac{\delta}{\delta m} p_{m+1}(\cdot,s,t,x,y)(v)\right| \\
&\leq K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s,y-x) \\
&\leq K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(t-s,y-x) \\
&\leq (t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(t-s,y-x) \\
&\qquad \sum_{k=1}^{m+1} C^k (t-s)^{(k-1)(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right),
\end{aligned}$$

provided that we choose $C \geq K$ in (8.101). We now treat the case $W_1(\mu_1,\mu_2) \leq (t-s)^{\frac{1}{\alpha}}$. By the representation formula (8.177), we have the following decomposition

$$\begin{aligned}
\Delta_{\mu_1,\mu_2}\partial_v\frac{\delta}{\delta m}p_{m+1}(\cdot,s,t,x,y)(v) &= p_{m+1}\otimes\Delta_{\mu_1,\mu_2}\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_2,s,t,x,y)(v) \\
&\quad + \Delta_{\mu_1,\mu_2}p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_1,s,t,x,y)(v) \\
&\quad + \left(\Delta_{\mu_1,\mu_2}\left[p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right]\right)\otimes\Phi_{m+1}(\mu_1,s,t,x,y)(v) \\
&\quad + \left[p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right]\otimes\Delta_{\mu_1,\mu_2}\Phi_{m+1}(\mu_2,s,t,x,y)(v) \\
&=: I_1+I_2+I_3+I_4.
\end{aligned}$$

For I_1, using (8.90), (8.193), the induction assumption (8.102), the space-time inequality (8.231) and the convolution inequality (8.234), one has since $\gamma \in (0,\eta+\alpha-1)$

$$\begin{aligned}
|I_1| &\leq K\int_s^t\int_{\mathbb{R}^d}\rho^0(r-s,z-x)\Big[(r-s)^{\frac{\eta-\gamma-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^1(t-r,y-z) \\
&\qquad +(t-r)^{-\frac{1}{\alpha}}\int_{\mathbb{R}^{2d}}(1\wedge|x'-y'|^{\eta})\left|\Delta_{\mu_1,\mu_2}\frac{\delta}{\delta m}p_m(\cdot,s,r,x',y')\right|\,dy'\,d\mu_2(x')\rho^1(t-r,y-z)\Big]\,dz\,dr \\
&\leq K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(t-s,y-x) \\
&\qquad \left[1+\sum_{k=1}^{m}C^k(t-s)^{k(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^{k}\mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}$$

For I_2, it follows from (8.99), (8.161) and the convolution inequality (8.234) that

$$\begin{aligned}
|I_2| &\leq K\int_s^t\int_{\mathbb{R}^d}(r-s)^{1-\frac{1+\gamma}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(r-s,z-x)(r-s)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr \\
&\leq K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1+\gamma}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(t-s,y-x) \\
&\leq K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^{\gamma}(\mu_1,\mu_2)\rho^0(t-s,y-x).
\end{aligned}$$

Concerning I_3, we note that it writes

$$I_3 = (I_1+I_2)\otimes\Phi_{m+1}(\mu_1,s,t,x,y)(v).$$

Thus, using the preceding bounds obtained for I_1 and I_2, (8.87) and the convolution inequality (8.234), we get that

$$\begin{aligned}|I_3| &\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\ &\qquad \left[1+\sum_{k=1}^m C^k (t-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^k \mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\ &\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}} W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\ &\qquad \left[1+\sum_{k=1}^m C^k (t-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)} \prod_{j=1}^k \mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].\end{aligned}$$

We finally deal with I_4. The convolution inequality (8.234), (8.90) and (8.161) (the series appearing in this bound being convergent) yield

$$\left|p_{m+1}\otimes \partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu_2,s,t,x,y)(v)\right| \le K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s,y-x).$$

Then, it follows from (8.186), which is valid for all $\gamma\in(0,1]$ since $W_1(\mu_1,\mu_2)<(t-s)^{\frac{1}{\alpha}}$ that

$$\begin{aligned}|I_4| &\le K\int_s^t\int_{\mathbb{R}^d} K(r-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(r-s,z-x)(r-s)^{\frac{\eta-\gamma}{\alpha}}(t-r)^{-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^1(t-r,y-z)\,dz\,dr\\ &\le K(t-s)^{\frac{\eta-1}{\alpha}+\frac{\eta-\gamma}{\alpha}+1-\frac{1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\ &\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x).\end{aligned}$$

Gathering all the previous estimates, we have proved that

$$\begin{aligned}&\left|\Delta_{\mu_1,\mu_2}\partial_v\frac{\delta}{\delta m}p_{m+1}(\cdot,s,t,x,y)(v)\right|\\ &\le K(t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\ &\qquad\left[1+\sum_{k=1}^m C^k(t-s)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^k\mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right]\\ &\le (t-s)^{\frac{\eta-\gamma-1}{\alpha}+1-\frac{1}{\alpha}}W_1^\gamma(\mu_1,\mu_2)\rho^0(t-s,y-x)\\ &\qquad\sum_{k=1}^{m+1}C^k(t-s)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(1+\frac{\eta-\gamma}{\alpha}+j\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right),\end{aligned}$$

provided that we choose $C\ge K$ in (8.102). This ends the proof of the induction step for (8.102).

8.8.7 Preparatory technical results

Lemma 8.33. • *For any $\gamma\in(0,1]$, $\tilde\eta\in(0,\eta\wedge(\alpha-1))$, there exists a positive constant K such that for all $m\ge1$, $t\in(0,T]$, $s_1,s_2\in[0,t)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $x,y,\in\mathbb{R}^d$*

$$|\Delta_{s_1,s_2}p_m(\mu,\cdot,t,x,y)|\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1)^\gamma}\rho^{-\tilde\eta}(t-s_1,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_2)^\gamma}\rho^{-\tilde\eta}(t-s_2,y-x)\right]. \quad (8.197)$$

- *For any $\gamma \in (0,1]$, there exists a positive constant K such that for all $m \geq 1$, $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x \in \mathbb{R}^d$*

$$\left|\Delta_{s_1,s_2} b(t,x,[X_t^{\cdot,\mu,(m)}])\right| \leq K|s_1-s_2|^\gamma \left[(t-s_1)^{\frac{\eta}{\alpha}-\gamma} + (t-s_2)^{\frac{\eta}{\alpha}-\gamma}\right]. \tag{8.198}$$

- *For any $\gamma \in (0,1]$, there exists a positive constant K such that for all $m \geq 1$, $0 \leq r < t \leq T$, $s_1, s_2 \in [0,r)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$*

$$|\Delta_{s_1,s_2}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)| \leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)|s_1-s_2|^\gamma \left[(r-s_1)^{\frac{\eta}{\alpha}-\gamma} + (r-s_2)^{\frac{\eta}{\alpha}-\gamma}\right], \tag{8.199}$$

and

$$|\Delta_{s_1,s_2}\Phi_{m+1}(\mu,\cdot,r,t,x,y)| \leq K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)|s_1-s_2|^\gamma \left[(r-s_1)^{\frac{\eta}{\alpha}-\gamma} + (r-s_2)^{\frac{\eta}{\alpha}-\gamma}\right]. \tag{8.200}$$

Proof of Lemma 8.33. **Proof of (8.197).** Assume first that $|s_1-s_2| > t - s_1 \vee s_2$. Then, using (8.90), we get that

$$\begin{aligned}
&|\Delta_{s_1,s_2} p_m(\mu,\cdot,t,x,y)| \\
&\leq K\left[\rho^0(t-s_1\vee s_2, y-x) + \rho^0(t-s_1\wedge s_2,y-x)\right] \\
&\leq K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}\rho^0(t-s_1\vee s_2,y-x) + \frac{(t-s_1\wedge s_2)^\gamma + |s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^\gamma}\rho^0(t-s_1\wedge s_2,y-x)\right] \\
&\leq K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}\rho^0(t-s_1\vee s_2,y-x) + \frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^\gamma}\rho^0(t-s_1\wedge s_2,y-x)\right].
\end{aligned}$$

We now focus on the case $|s_1-s_2| \leq t - s_1 \vee s_2$. For $\lambda \in [0,1]$, we set $s_\lambda := \lambda s_1 \vee s_2 + (1-\lambda)s_1\wedge s_2$. We have thus by (8.94) (the series appearing being convergent)

$$\begin{aligned}
&|\Delta_{s_1,s_2} p_m(\mu,\cdot,t,x,y)| \\
&\leq \int_0^1 |\partial_s p_m(\mu,s_\lambda,t,x,y)|\,|s_1-s_2|\,d\lambda \\
&\leq K|s_1-s_2|\int_0^1 (t-s_\lambda)^{-1}\rho^{-\tilde\eta}(t-s_\lambda,y-x)\,d\lambda \\
&\leq C|s_1-s_2|\int_0^1 (t-s_\lambda)^{-1-\frac{d}{\alpha}}(1+(t-s_\lambda)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\,d\lambda \\
&\leq K|s_1-s_2|^\gamma(t-s_1\vee s_2)^{1-\gamma}(t-s_1\vee s_2)^{-1-\frac{d}{\alpha}}\left[(1+(t-s_1\vee s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\right. \\
&\qquad\qquad\left.+(1+(t-s_1\wedge s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\right],
\end{aligned}$$

for some $\tilde\eta \in (0,\eta\wedge(\alpha-1))$. Since $|s_1-s_2| \leq t-s_1\vee s_2$, we easily check that $(t-s_1\vee s_2)^{-1} \leq 2(t-s_1\wedge s_2)^{-1}$. It follows that

$$
\begin{aligned}
&|\Delta_{s_1,s_2} p_m(\mu,\cdot,t,x,y)|\\
&\le K|s_1-s_2|^\gamma \Big[(t-s_1\vee s_2)^{-\gamma-\frac{d}{\alpha}}(1+(t-s_1\vee s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\\
&\qquad\qquad +(t-s_1\wedge s_2)^{-\gamma-\frac{d}{\alpha}}(1+(t-s_1\wedge s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha+\tilde\eta}\Big]\\
&\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\wedge s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\vee s_2,y-x)\right].
\end{aligned}
$$

This concludes the proof of (8.197).

Proof of (8.198)**.** By definition of the linear derivative and a centering argument, one can write setting $m_\lambda := \lambda[X_t^{s_1\vee s_2,\mu,(m)}]+(1-\lambda)[X_t^{s_1\wedge s_2,\mu,(m)}]$ for $\lambda\in[0,1]$

$$
\Delta_{s_1,s_2} b(t,x,[X_t^{\cdot,\mu,(m)}]) = \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}b(t,x,m_\lambda)(y)-\frac{\delta}{\delta m}b(t,x,m_\lambda)(x')\right)\Delta_{s_1,s_2}p_m(\mu,\cdot,t,x',y)\,d\mu(x')\,dy\,d\lambda.
$$

We deduce using (8.197) for some $\tilde\eta\in(0,\eta\wedge(\alpha-1))$, the boundedness of $\frac{\delta}{\delta m}b$ and the η-Hölder continuity of $\frac{\delta}{\delta m}b(t,x,\mu)(\cdot)$ that

$$
\begin{aligned}
&\left|\Delta_{s_1,s_2} b(t,x,[X_t^{\cdot,\mu,(m)}])\right|\\
&\le K\int_{\mathbb{R}^{2d}}|y-x'|^\eta\Big[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\wedge s_2,y-x')\\
&\qquad\qquad +\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\vee s_2,y-x')\Big]\,dy\,d\mu(x').
\end{aligned}
$$

Notice that because $\eta+\tilde\eta<\alpha$, $\rho^{-\tilde\eta-\eta}(t-s,\cdot)$ is integrable. By the space-time inequality (8.231) and since $\int_{\mathbb{R}^d}\rho^{-\tilde\eta-\eta}(t-s,y)\,dy$ is a finite constant independent of s and t, we obtain (8.198).

Proof of (8.199) **and** (8.200)**.** By the definition (8.82) of $\mathcal{H}_{m+1}$ and since $\widehat{p}(r,t,x,y)$ does not depend on μ and s, we immediately deduce (8.199) from (8.198) and (8.235).

Concerning the proof of (8.200), we start from the Volterra integral equation (8.83) which yields

$$
\begin{aligned}
\Delta_{s_1,s_2}\Phi_{m+1}(\mu,\cdot,r,t,x,y) &= \Delta_{s_1,s_2}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)+\Delta_{s_1,s_2}\mathcal{H}_{m+1}\otimes\Phi_{m+1}(\mu,s_1\vee s_2,r,t,x,y)\\
&\quad +\mathcal{H}_{m+1}\otimes\Delta_{s_1,s_2}\Phi_{m+1}(\mu,s_1\wedge s_2,r,t,x,y).
\end{aligned}
\tag{8.201}
$$

Using (8.199) and (8.87), we deduce that

$$
\begin{aligned}
&|\Delta_{s_1,s_2}\mathcal{H}_{m+1}\otimes\Phi_{m+1}(\mu,s_1\vee s_2,r,t,x,y)|\\
&\le K\int_r^t\int_{\mathbb{R}^d}(r'-r)^{-\frac{1}{\alpha}}\rho^1(r'-r,y-x)|s_1-s_2|^\gamma\left[(r'-s_1)^{\frac{\eta}{\alpha}-\gamma}+(r'-s_2)^{\frac{\eta}{\alpha}-\gamma}\right](t-r')^{-\frac{1}{\alpha}}\rho^1(t-r',y-z)\,dz\,dr'\\
&\le K(t-r)^{-\frac{1}{\alpha}+1-\frac{1}{\alpha}}\left[(r-s_1)^{\frac{\eta}{\alpha}-\gamma}+(r-s_2)^{\frac{\eta}{\alpha}-\gamma}\right]\rho^1(t-r,y-x).
\end{aligned}
$$

This inequality and (8.199) ensure that

$$|[\Delta_{s_1,s_2}\mathcal{H}_{m+1} + \Delta_{s_1,s_2}\mathcal{H}_{m+1}\otimes\Phi_{m+1}](\mu, s_1\vee s_2, r, t, x, y)| \\ \le K(t-r)^{-\frac{1}{\alpha}}\left[(r-s_1)^{\frac{\eta}{\alpha}-\gamma} + (r-s_2)^{\frac{\eta}{\alpha}-\gamma}\right]\rho^1(t-r, y-x). \quad (8.202)$$

As Φ_{m+1} yields a time-integrable singularity, we can iterate the implicit representation formula (8.201). We thus obtain

$$\Delta_{s_1,s_2}\Phi_{m+1}(\mu,\cdot,t,x,y) = [\Delta_{s_1,s_2}\mathcal{H}_{m+1} + \Delta_{s_1,s_2}\mathcal{H}_{m+1}\otimes\Phi_{m+1}](\mu, s_1\vee s_2, t, x, y) \\ + \sum_{k=1}^{\infty}\mathcal{H}_{m+1}^k\otimes[\Delta_{s_1,s_2}\mathcal{H}_{m+1} + \Delta_{s_1,s_2}\mathcal{H}_{m+1}\otimes\Phi_{m+1}](\mu, s_1\wedge s_2, t, x, y).$$

After standard computations using (8.202) and (8.85) that we omit, we conclude that (8.200) holds true.

□

8.8.8 Fourth part of the induction step

We prove here that the estimate (8.103) holds true.

Proof of (8.103)**.** We start by assuming that $|s_1-s_2| > t - s_1\vee s_2$. In this case, (8.90) directly yields

$$\begin{aligned}
&\left|\Delta_{s_1,s_2}\partial_x^j p_m(\mu,\cdot,t,x,y)\right| \\
&\le \left|\partial_x^j p_m(\mu,s_1,t,x,y)\right| + \left|\partial_x^j p_m(\mu,s_2,t,x,y)\right| \\
&\le K\left[(t-s_1\vee s_2)^{-\frac{j}{\alpha}}\rho^j(t-s_1\vee s_2, y-x) + (t-s_1\wedge s_2)^{-\frac{j}{\alpha}}\rho^j(t-s_1\wedge s_2, y-x)\right] \\
&\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\vee s_2, y-x) + \frac{|s_1-s_2|^\gamma + (t-s_1\vee s_2)^\gamma}{(t-s_1\wedge s_2)^\gamma(t-s_1\wedge s_2)^{\frac{j}{\alpha}}}\rho^j(t-s_1\wedge s_2, y-x)\right] \\
&\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\vee s_2, y-x) + \frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\wedge s_2, y-x)\right],
\end{aligned}$$

which proves (8.103).

We now turn to the case $|s_1-s_2| \le t - s_1\vee s_2$. Differentiating with respect to x the parametrix expansion (8.89), we get

$$\partial_x^j p_m(\mu,s,t,x,y) = \partial_x^j\widehat{p}(s,t,x,y) + \partial_x^j\widehat{p}\otimes\Phi_m(\mu,s,t,x,y).$$

We are going to use the following decomposition

$$\begin{aligned}\Delta_{s_1,s_2}\partial_x^j p_m(\mu,\cdot,t,x,y) &= \Delta_{s_1,s_2}\partial_x^j \widehat{p}(\cdot,t,x,y)\\ &\quad+\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\Delta_{s_1,s_2}\partial_x^j\widehat{p}(\cdot,r,x,z)\Phi_m(\mu,s_1\vee s_2,r,t,z,y)\,dz\,dr\\ &\quad+\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\partial_x^j\widehat{p}(s_1\wedge s_2,r,x,z)\Delta_{s_1,s_2}\Phi_m(\mu,\cdot,r,t,z,y)\,dz\,dr\\ &\quad-\int_{s_1\wedge s_2}^{s_1\vee s_2}\int_{\mathbb{R}^d}\partial_x^j\widehat{p}(s_1\wedge s_2,r,x,z)\Phi_m(\mu,s_1\wedge s_2,r,t,z,y)\,dz\,dr\\ &=: I_1+I_2+I_3+I_4.\end{aligned}$$

By (8.238), we obtain that

$$|I_1|\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\wedge s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\vee s_2,y-x)\right].$$

We now focus on I_2. Thanks to (8.238), (8.87) and the convolution inequality (8.234), one has

$$\begin{aligned}|I_2|&\le K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(r-s_1\wedge s_2,z-x)+\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(r-s_1\vee s_2,z-x)\right]\\ &\qquad\qquad(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\ &\le K\int_{s_1\wedge s_2}^t\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(r-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(r-s_1\wedge s_2,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\ &\quad+K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(r-s_1\vee s_2,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\ &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}-1+\frac{1}{\alpha}}}\rho^j(t-s_1\wedge s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}-1+\frac{1}{\alpha}}}\rho^j(t-s_1\vee s_2,y-x)\right]\\ &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\wedge s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}}\rho^j(t-s_1\vee s_2,y-x)\right].\end{aligned}$$

For I_3, it follows from (8.235), (8.200) and the convolution inequality (8.234) that

$$\begin{aligned}|I_3|&\le K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}(r-s_1\wedge s_2)^{-\frac{j}{\alpha}}\rho^j(r-s_1\wedge s_2,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\\ &\qquad\qquad|s_1-s_2|^\gamma\left[\frac{1}{(r-s_1\vee s_2)^\gamma}+\frac{1}{(r-s_1\wedge s_2)^\gamma}\right]dz\,dr\\ &\le K|s_1-s_2|^\gamma\int_{s_1\vee s_2}^t(r-s_1\vee s_2)^{-\frac{j}{\alpha}-\gamma}(t-r)^{-\frac{1}{\alpha}}\,dr\left[\rho^j(t-s_1\vee s_2,y-x)+\rho^j(t-s_1\wedge s_2,y-x)\right]\\ &\le K|s_1-s_2|^\gamma(t-s_1\vee s_2)^{-\frac{j}{\alpha}-\gamma+1-\frac{1}{\alpha}}\left[\rho^j(t-s_1\vee s_2,y-x)+\rho^j(t-s_1\wedge s_2,y-x)\right]\\ &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\frac{j}{\alpha}+\gamma}}\rho^j(t-s_1\vee s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\frac{j}{\alpha}+\gamma}}\rho^j(t-s_1\wedge s_2,y-x)\right],\end{aligned}$$

where the last inequality comes from the fact that $t-s_1\wedge s_1\le 2(t-s_1\vee s_2)$. We finally deal with I_4. Owing to (8.235) and (8.87), we have

$$
\begin{aligned}
|I_4| &\le K\int_{s_1\wedge s_2}^{s_1\vee s_2}\int_{\mathbb{R}^d}(r-s_1\wedge s_2)^{-\frac{j}{\alpha}}\rho^j(r-s_1\wedge s_2,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\
&\le K\int_{s_1\wedge s_2}^{s_1\vee s_2}(r-s_1\wedge s_2)^{-\frac{j}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\,dr\,\rho^j(t-s_1\wedge s_2,y-x)\\
&\le K(t-s_1\vee s_2)^{-\frac{1}{\alpha}}\int_{s_1\wedge s_2}^{s_1\vee s_2}(r-s_1\wedge s_2)^{-\frac{j}{\alpha}}\,dr\,\rho^j(t-s_1\wedge s_2,y-x)\\
&\le K(t-s_1\vee s_2)^{-\frac{1}{\alpha}}|s_1-s_2|^{1-\frac{j}{\alpha}}\left[\rho^j(t-s_1\vee s_2,y-x)+\rho^j(t-s_1\wedge s_2,y-x)\right].
\end{aligned}
$$

Since $\gamma\le 1-\frac{j}{\alpha}$ by assumption, we obtain

$$
\begin{aligned}
|I_4| &\le K(t-s_1\vee s_2)^{-\frac{1}{\alpha}-\gamma+1-\frac{j}{\alpha}}|s_1-s_2|^{\gamma}\left[\rho^j(t-s_1\vee s_2,y-x)+\rho^j(t-s_1\wedge s_2,y-x)\right]\\
&\le K\left[\frac{|s_1-s_2|^{\gamma}}{(t-s_1\vee s_2)^{\frac{j}{\alpha}+\gamma}}\rho^j(t-s_1\vee s_2,y-x)+\frac{|s_1-s_2|^{\gamma}}{(t-s_1\wedge s_2)^{\frac{j}{\alpha}+\gamma}}\rho^j(t-s_1\wedge s_2,y-x)\right],
\end{aligned}
$$

where the last inequality comes from the fact that $t-s_1\wedge s_1\le 2(t-s_1\vee s_2)$ and $1-\frac{1}{\alpha}>0$. This concludes the proof of (8.103).

8.8.9 Preparatory technical results

Lemma 8.34.
- *For any $\gamma\in(0,1]$, there exists a positive constant K such that for all $m\ge 1$, $t\in(0,T]$, $s_1,s_2\in[0,t)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $x,v\in\mathbb{R}^d$*

$$
\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(v)\right|\le K|s_1-s_2|^{\gamma}\left[(t-s_1)^{\frac{\eta}{\alpha}-\gamma}+(t-s_2)^{\frac{\eta}{\alpha}-\gamma}\right]. \tag{8.203}
$$

- *For any $\gamma\in(0,1]$, there exists a positive constant K such that for all $m\ge 1$, $t\in(0,T]$, $s_1,s_2\in[0,t)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $x,v\in\mathbb{R}^d$*

$$
\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}\left[b(t,x,[X_t^{\cdot,\mu,(m)}])\right](v)\right|\le K\frac{|s_1-s_2|^{\gamma}}{(t-s_1\vee s_2)^{\gamma}}+K\int_{\mathbb{R}^{2d}}\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x',y)\right|dy\,d\mu(x'). \tag{8.204}
$$

- *For any $\gamma\in(0,1]$, there exists a positive constant K such that for all $m\ge 1$, $t\in(0,T]$, $s_1,s_2\in[0,t)$, $r\in[s_1\vee s_2,t)$, $\mu\in\mathcal{P}_\beta(\mathbb{R}^d)$, $x,y,v\in\mathbb{R}^d$*

$$
\begin{aligned}
\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)(v)\right| &\le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\\
&\quad\left[\frac{|s_1-s_2|^{\gamma}}{(r-s_1\vee s_2)^{\gamma}}+\int_{\mathbb{R}^{2d}}\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_m(\mu,\cdot,r,x',y)\right|dy\,d\mu(x')\right].
\end{aligned} \tag{8.205}
$$

- *For any $\gamma\in(0,1]$, there exists a positive constant K such that for all $m\ge 1$, $t\in(0,T]$, $s_1,s_2\in$*

$[0,t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, v \in \mathbb{R}^d$

$$\left|\Delta_{s_1,s_2}\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{\cdot,\mu,(m)}])\right](v)\right| \le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}} + K\int_{\mathbb{R}^{2d}}(1\wedge|y-x'|^\eta)\left|\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x',y)\right| dy\, d\mu(x'). \quad (8.206)$$

- *For any $\gamma \in (0,1]$, there exists a positive constant K such that for all $m \ge 1$, $t \in (0,T]$, $s_1, s_2 \in [0,t)$, $r \in [s_1\vee s_2, t)$, $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $x, y, v \in \mathbb{R}^d$*

$$\left|\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)(v)\right| \le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x) \left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}} + \int_{\mathbb{R}^{2d}}(1\wedge|y-x'|^\eta)\left|\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,r,x',y)\right| dy\, d\mu(x')\right]. \quad (8.207)$$

Proof of Lemma 8.34. **Proof of (8.203).** By definition of the linear derivative and a centering argument, one can write, setting $m_\lambda := \lambda[X_t^{s_1\vee s_2,\mu,(m)}] + (1-\lambda)[X_t^{s_1\wedge s_2,\mu,(m)}]$, for $\lambda \in [0,1]$,

$$\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(v) = \int_{\mathbb{R}^{2d}}\left(\frac{\delta^2}{\delta m^2}b(t,x,m_\lambda)(v,y) - \frac{\delta^2}{\delta m^2}b(t,x,m_\lambda)(v,x')\right) \Delta_{s_1,s_2}p_m(\mu,\cdot,t,x',y)\, d\mu(x')\, dy\, d\lambda.$$

We deduce using (8.197) and uniform the η-Hölder continuity of $\frac{\delta^2}{\delta m^2}b(t,x,\mu)(v,\cdot)$ that

$$\begin{aligned}&\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(v)\right| \\ &\le K\int_{\mathbb{R}^{2d}}|y-x'|^\eta\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\wedge s_2,y-x') + \frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}\rho^{-\tilde\eta}(t-s_1\vee s_2,y-x')\right] dy\, d\mu(x').\end{aligned}$$

Notice that because $\eta+\tilde\eta<\alpha$, $\rho^{-\tilde\eta-\eta}(t-s,\cdot)$ is integrable. We conclude by the space-time inequality (8.231) and since $\int_{\mathbb{R}^d}\rho^{-\tilde\eta-\eta}(t-s,y)\, dy$ is a finite constant independent of s and t.

Proof of (8.204) and (8.205). Coming back to (8.162), we have the following decomposition

$$
\begin{aligned}
&\Delta_{s_1,s_2}\frac{\delta}{\delta m}\left[b(t,x,[X_t^{\cdot,\mu,(m)}])\right](v)\\
&=\int_{\mathbb{R}^d}\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(y)p_m(\mu,s_1\vee s_2,t,v,y)\,dy\\
&\quad+\int_{\mathbb{R}^d}\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)\Delta_{s_1,s_2}p_m(\mu,\cdot,t,v,y)\,dy\\
&\quad+\int_{\mathbb{R}^{2d}}\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(y)\frac{\delta}{\delta m}p_m(\mu,s_1\vee s_2,t,x',y)(v)\,d\mu(x')\,dy\\
&\quad+\int_{\mathbb{R}^{2d}}\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x',y)(v)\,d\mu(x')\,dy\\
&=: I_1+I_2+I_3+I_4.
\end{aligned}
$$

It immediately follows from (8.203) and (8.90) that

$$
\begin{aligned}
|I_1| &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1)^{\gamma-\frac{\eta}{\alpha}}}+\frac{|s_1-s_2|^\gamma}{(t-s_2)^{\gamma-\frac{\eta}{\alpha}}}\right]\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}.
\end{aligned}
$$

Then, the boundedness of $\frac{\delta}{\delta m}b$ and (8.103) yield

$$
\begin{aligned}
|I_2| &\le K\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(t-s_1)^\gamma}\rho^{-\tilde\eta}(t-s_1,y-v)+\frac{|s_1-s_2|^\gamma}{(t-s_2)^\gamma}\rho^{-\tilde\eta}(t-s_2,y-v)\,dy\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}.
\end{aligned}
$$

Similarly to I_1 and using (8.92), we deduce that

$$
|I_3|\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^\gamma}.
$$

Gathering the previous estimates on I_1, I_2, I_3 and noting that $\frac{\delta}{\delta m}b$ is bounded, we conclude that (8.204) holds true. Remark that (8.205) follows directly from the expression of $\frac{\delta}{\delta m}\mathcal{H}_{m+1}$ given in (8.164) and from (8.204).

Proof of (8.206)**.** We use (8.163), which yields the following decomposition

$$
\begin{aligned}
&\Delta_{s_1,s_2}\partial_v \frac{\delta}{\delta m}\left[b(t,x,[X_t^{\cdot,\mu,(m)}])\right](v)\\
&= \int_{\mathbb{R}^d} \Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(y)\partial_x p_m(\mu,s_1\vee s_2,t,v,y)\,dy\\
&\quad+\int_{\mathbb{R}^d}\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)\Delta_{s_1,s_2}\partial_x p_m(\mu,\cdot,t,v,y)\,dy\\
&\quad+\int_{\mathbb{R}^{2d}}\Delta_{s_1,s_2}\frac{\delta}{\delta m}b(t,x,[X_t^{\cdot,\mu,(m)}])(y)\partial_v\frac{\delta}{\delta m}p_m(\mu,s_1\vee s_2,t,x',y)(v)\,d\mu(x')\,dy\\
&\quad+\int_{\mathbb{R}^{2d}}\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x',y)(v)\,d\mu(x')\,dy\\
&=: I_1+I_2+I_3+I_4.
\end{aligned}
$$

It immediately follows from (8.203) and (8.90) that

$$
\begin{aligned}
|I_1| &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1)^{\gamma-\frac{\eta}{\alpha}}}+\frac{|s_1-s_2|^\gamma}{(t-s_2)^{\gamma-\frac{\eta}{\alpha}}}\right](t-s_1\vee s_2)^{-\frac{1}{\alpha}}\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}}.
\end{aligned}
$$

We rewrite I_2 in the following form

$$
I_2=\int_{\mathbb{R}^d}\left(\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)-\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(v)\right)\Delta_{s_1,s_2}\partial_x p_m(\mu,\cdot,t,v,y)\,dy.
$$

Then, the η-Hölder continuity of $\frac{\delta}{\delta m}b(t,x,\mu)(\cdot)$ and (8.103) yield

$$
|I_2|\le K\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(t-s_1)^{\gamma+\frac{1}{\alpha}}}|y-v|^\eta\rho^j(t-s_1,y-v)+\frac{|s_1-s_2|^\gamma}{(t-s_2)^{\gamma+\frac{1}{\alpha}}}|y-v|^\eta\rho^j(t-s_2,y-v)\,dy.
$$

The space-time inequality (8.231) ensures that

$$
|I_2|\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}}.
$$

By using (8.203) and (8.93), we deduce that

$$
\begin{aligned}
|I_3| &\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1)^{\gamma-\frac{\eta}{\alpha}}}+\frac{|s_1-s_2|^\gamma}{(t-s_2)^{\gamma-\frac{\eta}{\alpha}}}\right](t-s_1\vee s_2)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}}.
\end{aligned}
$$

Note that since $\int_{\mathbb{R}^d}\partial_v\frac{\delta}{\delta m}p_m(\mu,s,t,x,y)(v)\,dy=0$, one can rewrite I_4 as

$$
I_4=\int_{\mathbb{R}^{2d}}\left(\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(y)-\frac{\delta}{\delta m}b(t,x,[X_t^{s_1\wedge s_2,\mu,(m)}])(x')\right)\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x',y)(v)\,d\mu(x')\,dy.
$$

The boundedness of $\frac{\delta}{\delta m}b$ and the η-Hölder continuity of $\frac{\delta}{\delta m}b(t,x,\mu)(\cdot)$ yield

$$|I_4| \le \int_{\mathbb{R}^{2d}} (1 \wedge |y-x'|^\eta) \left| \Delta_{s_1,s_2} \partial_v \frac{\delta}{\delta m} p_m(\mu,\cdot,r,x',y) \right| dy\, d\mu(x').$$

Gathering the previous estimates on I_1, I_2, I_3 and I_4, we conclude that (8.206) holds true. Finally, notice that (8.207) follows directly from the expression of $\partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1}$ obtained by differentiating (8.164) with respect to v and from (8.206).

□

8.8.10 Fifth part of the induction step

We prove here that the estimates (8.104) and (8.105) hold true.

Proof of (8.104). We separate the proof of the induction step into two disjoint cases. First, we assume that $|s_1-s_2| > t - s_1 \vee s_2$. By (8.92) (the series appearing in the bound being convergent), we can write

$$\begin{aligned}
&\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x,y)(v)\right| \\
&\le \left|\frac{\delta}{\delta m}p_m(\mu,s_1,t,x,y)(v)\right| + \left|\frac{\delta}{\delta m}p_m(\mu,s_2,t,x,y)(v)\right| \\
&\le K\left[(t-s_1\vee s_2)^{1-\frac{1}{\alpha}}\rho^0(t-s_1\vee s_2, y-x) + (t-s_1\wedge s_2)^{1-\frac{1}{\alpha}}\rho^0(t-s_1\wedge s_2,y-x)\right] \\
&\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1\vee s_2,y-x) + \frac{|s_1-s_2|^\gamma+(t-s_1\vee s_2)^\gamma}{(t-s_1\wedge s_2)^\gamma (t-s_1\wedge s_2)^{\frac{1}{\alpha}-1}}\rho^0(t-s_1\wedge s_2,y-x)\right] \\
&\le K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1\vee s_2,y-x) + \frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}}\rho^0(t-s_1\wedge s_2,y-x)\right],
\end{aligned}$$

This shows that (8.104) holds true at step $m+1$ provided that we choose $C \ge K$ in (8.104). We now turn to the case $|s_1-s_2| \le t-s_1\vee s_2$. Using the representation formula (8.176), we get that

$$\begin{aligned}
\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_{m+1}(\mu,\cdot,t,x,y)(v) &= \Delta_{s_1,s_2}\left(p_{m+1}\otimes\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,\cdot,t,x,y)(v) \\
&\quad + \Delta_{s_1,s_2}\left(p_{m+1}\otimes\frac{\delta}{\delta m}\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right)(\mu,\cdot,t,x,y)(v) \\
&=: I_1 + I_2.
\end{aligned}$$

Then, we decompose I_1 in the following way

$$\begin{aligned}
I_1 &= \int_{s_1\vee s_2}^t \int_{\mathbb{R}^d} \Delta_{s_1,s_2}p_{m+1}(\mu,\cdot,r,x,z)\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s_1\vee s_2,r,t,z,y)\,dz\,dr \\
&\quad + \int_{s_1\vee s_2}^t \int_{\mathbb{R}^d} p_{m+1}(\mu,s_1\wedge s_2,r,x,z)\Delta_{s_1,s_2}\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,z,y)\,dz\,dr \\
&\quad - \int_{s_1\wedge s_2}^{s_1\vee s_2} \int_{\mathbb{R}^d} p_{m+1}(\mu,s_1\wedge s_2,r,x,z)\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s_1\wedge s_2,r,t,z,y)\,dz\,dr \\
&=: I_{1,1}+I_{1,2}+I_{1,3}.
\end{aligned}$$

It follows from (8.103), (8.160) (the series appearing in the bound being convergent) and the convolution inequality (8.234) that

$$\begin{aligned}|I_{1,1}| &\le K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^\gamma}\left[\rho^0(r-s_1\vee s_2,y-x)+\rho^0(r-s_1\wedge s_2,y-x)\right]\\ &\qquad\qquad (t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\ &\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}\left[\rho^0(t-s_1\vee s_2,y-x)\rho^0(t-s_1\wedge s_2,y-x)\right].\end{aligned}$$

We now deal with $I_{1,2}$. Using the induction assumption and (8.205), we deduce that

$$\begin{aligned}&\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)(v)\right|\\ &\le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^\gamma}+\int_{\mathbb{R}^{2d}}\left|\Delta_{s_1,s_2}\frac{\delta}{\delta m}p_m(\mu,\cdot,r,x',y)\right|dy\,d\mu(x')\right]\\ &\le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^\gamma}\\ &\quad+K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}+\frac{|s_1-s_2|^\gamma}{(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}}\right]\\ &\qquad\sum_{k=1}^m C^k(r-s_1\vee s_2)^{(k-1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(\left[2-\gamma-\frac{1}{\alpha}\right]\wedge 1+(j-1)\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).\end{aligned}\tag{8.208}$$

Notice that

$$\begin{cases}(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}\ge(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}, & \text{if } \gamma+\frac{1}{\alpha}-1\ge 0,\\ (r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}\ge(t-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}\ge 2^{\gamma+\frac{1}{\alpha}-1}(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}, & \text{if } \gamma+\frac{1}{\alpha}-1<0,\end{cases}$$

since $t-s_1\wedge s_2\le 2(t-s_1\vee s_2)$. Using these inequalities, one has

$$\begin{aligned}&\int_{s_1\vee s_2}^t(t-r)^{-\frac{1}{\alpha}}\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}+\frac{|s_1-s_2|^\gamma}{(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1}}\right](r-s_1\vee s_2)^{(k-1)\left(1-\frac{1}{\alpha}\right)}dr\\ &\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}(t-s_1\vee s_2)^{k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(1-\gamma+k\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\ &\quad+K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}(t-s_1\vee s_2)^{k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(1+(k-1)\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\ &\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1}}(t-s_1\vee s_2)^{k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(\left[2-\gamma-\frac{1}{\alpha}\right]\wedge 1+(k-1)\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right),\end{aligned}$$

since the Beta function is decreasing with respect to its first argument. From this inequality, (8.208) and (8.90), we obtain

$$|I_{1,2}| \leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} \rho^0(t - s_1 \wedge s_2, y - x) \tag{8.209}$$
$$\left[1 + \sum_{k=1}^{m} C^k (t - s_1 \vee s_2)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].$$

We now turn to estimate $I_{1,3}$. Thanks to (8.90) and (8.160), one has

$$\begin{aligned}
|I_{1,3}| &\leq \int_{s_1 \wedge s_2}^{s_1 \vee s_2} \rho^0(r - s_1 \wedge s_2, z - x)(t - r)^{-\frac{1}{\alpha}} \rho^1(t - r, y - z)\, dz\, dr \\
&\leq K \frac{|s_1 - s_2|}{(t - s_1 \vee s_2)^{\frac{1}{\alpha}}} \rho^0(t - s_1 \wedge s_2, y - x) \\
&\leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} \rho^0(t - s_1 \wedge s_2, y - x).
\end{aligned}$$

Gathering the preceding estimates on $I_{1,1}$ and $I_{1,2}$ and $I_{1,3}$, we have proved that

$$|I_1| \leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right] \tag{8.210}$$
$$\left[1 + \sum_{k=1}^{m} C^k (t - s_1 \vee s_2)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].$$

As done before for I_1, we decompose I_2 in the following way

$$\begin{aligned}
I_2 = &\int_{s_1 \vee s_2}^{t} \int_{\mathbb{R}^d} \Delta_{s_1, s_2} \left(p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}\right)(\mu, \cdot, r, x, z) \Phi_{m+1}(\mu, s_1 \vee s_2, r, t, z, y)\, dz\, dr \\
&+ \int_{s_1 \vee s_2}^{t} \int_{\mathbb{R}^d} \left(p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}\right)(\mu, s_1 \wedge s_2, r, x, z) \Delta_{s_1, s_2} \Phi_{m+1}(\mu, \cdot, r, t, z, y)\, dz\, dr \\
&- \int_{s_1 \wedge s_2}^{s_1 \vee s_2} \int_{\mathbb{R}^d} \left(p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}\right)(\mu, s_1 \wedge s_2, r, x, z) \Phi_{m+1}(\mu, s_1 \wedge s_2, r, t, z, y)\, dz\, dr \\
=: &\, I_{2,1} + I_{2,2} + I_{2,3}.
\end{aligned}$$

We follow the same lines of reasoning as for I_1. Using the bound (8.210) previously obtained for I_1 and (8.87), we have that

$$|I_{2,1}| \leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} (t - s_1 \vee s_2)^{1 - \frac{1}{\alpha}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right]$$
$$\left[1 + \sum_{k=1}^{m} C^k (t - s_1 \vee s_2)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^{k} \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].$$

Notice that (8.90), (8.160) and the convolution inequality (8.234) yield

$$\left|p_{m+1} \otimes \frac{\delta}{\delta m} \mathcal{H}_{m+1}(\mu, s, t, x, y)(v)\right| \leq K (t - s)^{1 - \frac{1}{\alpha}} \rho^0(t - s, y - x). \tag{8.211}$$

Using this inequality and (8.200), we deduce that

$$|I_{2,2}| \le K \int_{s_1 \vee s_2}^t \int_{\mathbb{R}^d} (t - s_1 \wedge s_2)^{1-\frac{1}{\alpha}} \rho^0(r - s_1 \wedge s_2, z - x) \frac{|s_1 - s_2|^\gamma}{(r - s_1 \vee s_2)^\gamma} (t-r)^{-\frac{1}{\alpha}} \rho^1(t-r, y-z)\, dz\, dr.$$

Since $t - s_1 \wedge s_2 \le 2(t - s_1 \vee s_2)$, we get

$$|I_{2,2}| \le K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} (t - s_1 \vee s_2)^{1-\frac{1}{\alpha}} \rho^0(t - s_1 \wedge s_2, y - x).$$

For $I_{2,3}$, (8.211) and (8.87) yield

$$\begin{aligned}
|I_{2,3}| &\le \int_{s_1 \wedge s_2}^{s_1 \vee s_2} (r - s_1 \wedge s_2)^{1-\frac{1}{\alpha}} \rho^0(r - s_1 \wedge s_2, z - x)(t-r)^{-\frac{1}{\alpha}} \rho^1(t-r, y-z)\, dz\, dr \\
&\le K \frac{|s_1 - s_2|^{1+1-\frac{1}{\alpha}}}{(t - s_1 \vee s_2)^{\frac{1}{\alpha}}} \rho^0(t - s_1 \wedge s_2, y - x) \\
&\le K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} (t - s_1 \vee s_2)^{1-\frac{1}{\alpha}} \rho^0(t - s_1 \wedge s_2, y - x),
\end{aligned}$$

since $|s_1 - s_2| \le t - s_1 \vee s_2$. Gathering the estimates obtained, we have proved that

$$\begin{aligned}
|I_2| \le K &\frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} (t - s_1 \vee s_2)^{1-\frac{1}{\alpha}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right] \\
&\left[1 + \sum_{k=1}^m C^k (t - s_1 \vee s_2)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^k \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right].
\end{aligned} \tag{8.212}$$

We conclude by (8.210), (8.212) and the fact that $t - s_1 \vee s_2 \le t - s_1 \wedge s_2 \le 2(t - s_1 \vee s_2)$ that

$$\begin{aligned}
&\left|\Delta_{s_1,s_2} \frac{\delta}{\delta m} p_{m+1}(\mu, \cdot, t, x, y)(v)\right| \\
&\le K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right] \\
&\qquad \left[1 + \sum_{k=1}^m C^k (t - s_1 \vee s_2)^{k(1-\frac{1}{\alpha})} \prod_{j=1}^k \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right] \\
&\le \left[\frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1}} \rho^0(t - s_1 \vee s_2, y - x) + \frac{|s_1 - s_2|^\gamma}{(t - s_1 \wedge s_2)^{\gamma + \frac{1}{\alpha} - 1}} \rho^0(t - s_1 \wedge s_2, y - x)\right] \\
&\qquad \sum_{k=1}^{m+1} C^k (t - s_1 \vee s_2)^{(k-1)(1-\frac{1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(\left[2 - \gamma - \frac{1}{\alpha}\right] \wedge 1 + (j-1)\left(1 - \frac{1}{\alpha}\right), 1 - \frac{1}{\alpha}\right),
\end{aligned}$$

provided that we choose $C \ge K$ in (8.104). This ends the proof of the induction step for (8.104).

Proof of (8.105)**.** We start by treating the case $|s_1 - s_2| > t - s_1 \vee s_2$. Reasoning as before by using

(8.93) (the series appearing in the bound being convergent), we get that

$$
\begin{aligned}
&\left|\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,t,x,y)(v)\right| \\
&\leq K\left[(t-s_1\vee s_2)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s_1\vee s_2,y-x)+(t-s_1\wedge s_2)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s_1\wedge s_2,y-x)\right] \\
&\leq K\left[\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\vee s_2,y-x)+\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\wedge s_2,y-x)\right].
\end{aligned}
$$

This shows that (8.105) holds true at step $m+1$ provided that we choose $C \geq K$ in (8.105). We now turn to the case $|s_1-s_2| \leq t-s_1\vee s_2$. By the representation formula (8.177), we write

$$
\begin{aligned}
\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_{m+1}(\mu,\cdot,t,x,y)(v) &= \Delta_{s_1,s_2}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,\cdot,t,x,y)(v) \\
&\quad+\Delta_{s_1,s_2}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\otimes\Phi_{m+1}\right)(\mu,\cdot,t,x,y)(v) \\
&=: I_1+I_2.
\end{aligned}
$$

Then, we decompose I_1 in the following way

$$
\begin{aligned}
I_1 &= \int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\Delta_{s_1,s_2}p_{m+1}(\mu,\cdot,r,x,z)\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s_1\vee s_2,r,t,z,y)\,dz\,dr \\
&\quad+\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}p_{m+1}(\mu,s_1\wedge s_2,r,x,z)\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,z,y)\,dz\,dr \\
&\quad-\int_{s_1\wedge s_2}^{s_1\vee s_2}\int_{\mathbb{R}^d}p_{m+1}(\mu,s_1\wedge s_2,r,x,z)\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s_1\wedge s_2,r,t,z,y)\,dz\,dr \\
&=: I_{1,1}+I_{1,2}+I_{1,3}.
\end{aligned}
$$

It follows from (8.103), (8.161) (the series appearing in the bound being convergent), the convolution inequality (8.234) and since $\gamma < 1+\frac{\eta-1}{\alpha}$ that

$$
\begin{aligned}
|I_{1,1}| &\leq K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^\gamma}\left[\rho^0(r-s_1\vee s_2,y-x)+\rho^0(r-s_1\wedge s_2,y-x)\right] \\
&\qquad\qquad(r-s_1\vee s_2)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr \\
&\leq K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\left[\rho^0(t-s_1\vee s_2,y-x)+\rho^0(t-s_1\wedge s_2,y-x)\right].
\end{aligned}
$$

We now deal with $I_{1,2}$. Using the induction assumption and (8.207), we deduce that

$$
\begin{aligned}
&\left|\Delta_{s_1,s_2}\partial_v \frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,\cdot,r,t,x,y)(v)\right|\\
&\le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}}+\int_{\mathbb{R}^{2d}}(1\wedge|y-x'|^\eta)\left|\Delta_{s_1,s_2}\partial_v\frac{\delta}{\delta m}p_m(\mu,\cdot,r,x',y)\right|\,dy\,d\mu(x')\right]\\
&\le K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1-\eta}{\alpha}}}\\
&\quad+K(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-x)\left[\frac{|s_1-s_2|^\gamma(r-s_1\vee s_2)^{\frac{\eta}{\alpha}}}{(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}+\frac{|s_1-s_2|^\gamma(r-s_1\wedge s_2)^{\frac{\eta}{\alpha}}}{(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\right]\\
&\qquad\sum_{k=1}^m C^k(r-s_1\vee s_2)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(\left[2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma\right]\wedge 1+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

Note that

$$
\begin{cases}
(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}\ge(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}, & \text{if } \gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}\ge 0,\\
(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}\ge 2^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}, & \text{if } \gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}<0,
\end{cases}
\tag{8.213}
$$

since $t-s_1\wedge s_2\le 2(t-s_1\vee s_2)$. It yields

$$
\begin{aligned}
&\int_{s_1\vee s_2}^t(t-r)^{-\frac{1}{\alpha}}\left[\frac{|s_1-s_2|^\gamma}{(r-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}}+\frac{|s_1-s_2|^\gamma}{(r-s_1\wedge s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-2\eta}{\alpha}}}\right](r-s_1\vee s_2)^{(k-1)\left(1+\frac{\eta-1}{\alpha}\right)}\,dr\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}(t-s_1\vee s_2)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\mathcal{B}\left(2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma+(k-1)\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\
&\quad+K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}(t-s_1\vee s_2)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\mathcal{B}\left(1+(k-1)\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\
&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}(t-s_1\vee s_2)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\\
&\qquad\qquad\mathcal{B}\left(\left[2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma\right]\wedge 1+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right),
\end{aligned}
$$

since the Beta function is decreasing with respect to its first argument. From this inequality, (8.208) and (8.90), we deduce that

$$
\begin{aligned}
|I_{1,2}|&\le K\frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\wedge s_2,y-x)\\
&\quad\left[1+\sum_{k=1}^m C^k(t-s_1\vee s_2)^{k\left(1+\frac{\eta-1}{\alpha}\right)}\prod_{j=1}^k\mathcal{B}\left(\left[2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma\right]\wedge 1+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}
$$

We now turn to estimate $I_{1,3}$. Thanks to (8.90) and (8.161), one has

$$
\begin{aligned}
|I_{1,3}| &\le \int_{s_1\wedge s_2}^{s_1\vee s_2} \rho^0(r-s_1\wedge s_2, z-x)(r-s_1\wedge s_2)^{\frac{\eta-1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr \\
&\le K\frac{|s_1-s_2|^{1+\frac{\eta-1}{\alpha}}}{(t-s_1\vee s_2)^{\frac{1}{\alpha}}}\rho^0(t-s_1\wedge s_2, y-x) \\
&\le K\frac{|s_1-s_2|^{\gamma}}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\rho^0(t-s_1\wedge s_2, y-x),
\end{aligned}
$$

since $\gamma < 1+\frac{\eta-1}{\alpha}$. Gathering the estimates obtained, we have proved that

$$
\begin{aligned}
|I_1| \le K&\frac{|s_1-s_2|^{\gamma}}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}\left[\rho^0(t-s_1\vee s_2,y-x)+\rho^0(t-s_1\wedge s_2,y-x)\right] \qquad (8.214)\\
&\left[1+\sum_{k=1}^m C^k(t-s_1\vee s_2)^{k(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^k \mathcal{B}\left(\left[2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma\right]\wedge 1+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}
$$

Then, we decompose I_2 in the following way

$$
\begin{aligned}
I_2 &= \int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\Delta_{s_1,s_2}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,\cdot,r,x,z)\Phi_{m+1}(\mu,s_1\vee s_2,r,t,z,y)\,dz\,dr \\
&+\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,s_1\wedge s_2,r,x,z)\Delta_{s_1,s_2}\Phi_{m+1}(\mu,\cdot,r,t,z,y)\,dz\,dr \\
&-\int_{s_1\wedge s_2}^{s_1\vee s_2}\int_{\mathbb{R}^d}\left(p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}\right)(\mu,s_1\wedge s_2,r,x,z)\Phi_{m+1}(\mu,s_1\wedge s_2,r,t,z,y)\,dz\,dr \\
&=: I_{2,1}+I_{2,2}+I_{2,3}.
\end{aligned}
$$

We follow the same lines of reasoning as for I_1. Using the bound (8.214) previously obtained for I_1 and (8.87), we show that

$$
\begin{aligned}
|I_{2,1}| \le K&\frac{|s_1-s_2|^{\gamma}}{(t-s_1\vee s_2)^{\gamma+\frac{1}{\alpha}-1+\frac{1-\eta}{\alpha}}}(t-s_1\vee s_2)^{1-\frac{1}{\alpha}}\left[\rho^0(t-s_1\vee s_2,y-x)+\rho^0(t-s_1\wedge s_2,y-x)\right]\\
&\left[1+\sum_{k=1}^m C^k(t-s_1\vee s_2)^{k(1+\frac{\eta-1}{\alpha})}\prod_{j=1}^k \mathcal{B}\left(\left[2\left(1+\frac{\eta-1}{\alpha}\right)-\gamma\right]\wedge 1+(j-1)\left(1+\frac{\eta-1}{\alpha}\right),1-\frac{1}{\alpha}\right)\right].
\end{aligned}
$$

Notice that it follows from (8.90), (8.161) and the convolution inequality (8.234) that

$$
\left|p_{m+1}\otimes\partial_v\frac{\delta}{\delta m}\mathcal{H}_{m+1}(\mu,s,t,x,y)(v)\right| \le K(t-s)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(t-s,y-x). \qquad (8.215)
$$

Using this inequality and (8.200), we deduce that

$$
|I_{2,2}| \le K\int_{s_1\vee s_2}^t\int_{\mathbb{R}^d}(r-s_1\wedge s_2)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}\rho^0(r-s_1\wedge s_2,z-x)\frac{|s_1-s_2|^{\gamma}}{(r-s_1\vee s_2)^{\gamma}}(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr.
$$

Reasoning as in (8.213) to control $(r-s_1\wedge s_2)^{\frac{\eta-1}{\alpha}+1-\frac{1}{\alpha}}$ since $t-s_1\vee s_2 \le t-s_1\wedge s_2 \le 2(t-s_1\vee s_2)$, we get that

$$|I_{2,2}| \leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} (t - s_1 \vee s_2)^{1 - \frac{1}{\alpha}} \rho^0(t - s_1 \wedge s_2, y - x).$$

For $I_{2,3}$, note that (8.215), (8.87) and the convolution inequality (8.234) yield

$$\begin{aligned} |I_{2,3}| &\leq \int_{s_1 \wedge s_2}^{s_1 \vee s_2} (r - s_1 \vee s_2)^{\frac{\eta - 1}{\alpha} + 1 - \frac{1}{\alpha}} \rho^0(r - s_1 \wedge s_2, z - x)(t - r)^{-\frac{1}{\alpha}} \rho^1(t - r, y - z)\, dz\, dr \\ &\leq K \frac{|s_1 - s_2|^{1 + \frac{\eta - 1}{\alpha} + 1 - \frac{1}{\alpha}}}{(t - s_1 \vee s_2)^{\frac{1}{\alpha}}} \rho^0(t - s_1 \wedge s_2, y - x) \\ &\leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} (t - s_1 \vee s_2)^{1 - \frac{1}{\alpha}} \rho^0(t - s_1 \wedge s_2, y - x), \end{aligned}$$

since $|s_1 - s_2| \leq t - s_1 \vee s_2$. Gathering the estimates obtained, we have proved that

$$\begin{aligned} |I_2| \leq K &\frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} (t - s_1 \vee s_2)^{1 - \frac{1}{\alpha}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right] \\ &\left[1 + \sum_{k=1}^m C^k (t - s_1 \vee s_2)^{k(1 + \frac{\eta - 1}{\alpha})} \prod_{j=1}^k \mathcal{B}\left(\left[2\left(1 + \frac{\eta - 1}{\alpha}\right) - \gamma\right] \wedge 1 + (j-1)\left(1 + \frac{\eta - 1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right]. \end{aligned} \tag{8.216}$$

We conclude from (8.214), (8.216) and the fact that $t - s_1 \wedge s_2 \leq 2(t - s_1 \vee s_2)$ that

$$\begin{aligned} &\left|\Delta_{s_1, s_2} \partial_v \frac{\delta}{\delta m} p_{m+1}(\mu, \cdot, t, x, y)(v)\right| \\ &\leq K \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} \left[\rho^0(t - s_1 \vee s_2, y - x) + \rho^0(t - s_1 \wedge s_2, y - x)\right] \\ &\qquad \left[1 + \sum_{k=1}^m C^k (t - s_1 \vee s_2)^{k(1 + \frac{\eta - 1}{\alpha})} \prod_{j=1}^k \mathcal{B}\left(\left[2\left(1 + \frac{\eta - 1}{\alpha}\right) - \gamma\right] \wedge 1 + (j-1)\left(1 + \frac{\eta - 1}{\alpha}\right), 1 - \frac{1}{\alpha}\right)\right] \\ &\leq \left[\frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{1}{\alpha} - 1 + \frac{1-\eta}{\alpha}}} \rho^0(t - s_1 \vee s_2, y - x) + \frac{|s_1 - s_2|^\gamma}{(t - s_1 \wedge s_2)^{\gamma + \frac{1}{\alpha} - 1}} \rho^0(t - s_1 \wedge s_2, y - x)\right] \\ &\qquad \sum_{k=1}^{m+1} C^k (t - s_1 \vee s_2)^{(k-1)(1 + \frac{\eta - 1}{\alpha})} \prod_{j=1}^{k-1} \mathcal{B}\left(\left[2\left(1 + \frac{\eta - 1}{\alpha}\right) - \gamma\right] \wedge 1 + (j-1)\left(1 + \frac{\eta - 1}{\alpha}\right), 1 - \frac{1}{\alpha}\right), \end{aligned}$$

provided that we choose $C \geq K$ in (8.105). This ends the proof of the induction step for (8.105).

8.9 Appendix: Differential calculus for functions of a measure variable

Let us fix $\beta \in [0, 2]$. We use the following convention $\mathcal{P}_0(\mathbb{R}^d) := \mathcal{P}(\mathbb{R}^d)$, endowed with the weak topology.

Definition 8.35 (Linear derivative)**.** A function $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ is said to have a linear derivative if there exists a function $\frac{\delta}{\delta m} u \in \mathcal{C}^0(\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d; \mathbb{R})$ satisfying the following properties.

1. For all compact subset $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists a constant $C_\mathcal{K} > 0$ such that

$$\forall \mu \in \mathcal{K},\, \forall v \in \mathbb{R}^d,\, \left|\frac{\delta u}{\delta m}(\mu)(v)\right| \leq C_\mathcal{K}(1 + |v|^\beta).$$

2. For all $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d)$, we have

$$u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta}{\delta m} u(t\mu + (1-t)\nu)(v)\, d(\mu - \nu)(v)\, dt.$$

The function u is said to have a linear derivative of order two if for all $v \in \mathbb{R}^d$, the map $\frac{\delta}{\delta m} u(\cdot)(v)$ admits a linear derivative $\frac{\delta^2}{\delta m^2}(\cdot)(v, \cdot)$ such that $\frac{\delta^2}{\delta m^2} u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d$ and for all compact subset $\mathcal{K} \subset \mathcal{P}_\beta(\mathbb{R}^d)$, there exists a constant $C_\mathcal{K} > 0$ such that

$$\forall \mu \in \mathcal{K},\, \forall v, v' \in \mathbb{R}^d,\, \left|\frac{\delta^2}{\delta m^2} u(\mu)(v, v')\right| \leq C_\mathcal{K}(1 + |v|^\beta + |v'|^\beta).$$

Definition 8.36. We define the space $\mathcal{C}^1([0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d))$ as the set of continuous functions $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ satisfying the following properties.

1. For any $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, the map $u(\cdot, \cdot, \mu)$ belongs to $\mathcal{C}^1([0,T] \times \mathbb{R}^d)$ with $\partial_t u$ and $\partial_x u$ continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$.
2. For any $(t, x) \in [0,T] \times \mathbb{R}^d$, the map $u(t, x, \cdot)$ admits a linear derivative $(\mu, v) \in \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \frac{\delta}{\delta m} u(t, x, \mu)(v)$ such that $\frac{\delta}{\delta m} u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$.
3. For any $(t, x, \mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, the map $\frac{\delta}{\delta m} u(t, x, \mu)$ is of class $\mathcal{C}^1$ on $\mathbb{R}^d$ and $\partial_v \frac{\delta}{\delta m} u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$.

We now introduce the notion of empirical projection.

Definition 8.37 (Empirical projection)**.** Fix $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$. For all $N \geq 1$, the empirical projection u^N of u is defined, for all $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$, by

$$u^N(\boldsymbol{x}) = u(\overline{\mu}^N_{\boldsymbol{x}}),$$

where $\overline{\mu}^N_{\boldsymbol{x}} = \frac{1}{N} \sum_{j=1}^N \delta_{x_j}$.

The following proposition is the analogue of [CD18a, Proposition 5.91] where $\beta = 2$, so we don't give the proof.

Proposition 8.38. *Let $u : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a function admitting a linear derivative $\frac{\delta}{\delta m} u$ such that all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d)$, $\frac{\delta}{\delta m} u(\mu)(\cdot) \in \mathcal{C}^1(\mathbb{R}^d)$ and $\partial_v \frac{\delta}{\delta m} u$ is continuous on $\mathcal{P}_\beta(\mathbb{R}^d) \times \mathbb{R}^d$. Then, for all $N \geq 1$, the empirical projection u^N of u is of class $\mathcal{C}^1$. Moreover for all $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$*

$$\partial_{x_i} u^N(x_1, \dots, x_N) = \frac{1}{N} \partial_v \frac{\delta}{\delta m} u(\overline{\mu}^N_{\boldsymbol{x}})(x_i).$$

The next proposition illustrates how a smooth flow of measures admitting a transition density can regularize a function defined on $\mathcal{P}_\beta(\mathbb{R}^d)$. It is clearly reminiscent of [CdRF21, Proposition 2.3]. We don't prove it since it can be done in a completely analogous manner.

Proposition 8.39 (Regularization by a smooth flow of density functions). *Let us fix $\phi : \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ a function admitting a linear derivative and consider a map $(\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0,T) \times \mathbb{R}^d \mapsto p(\mu, s, T, x, y)$, where $T > 0$ is fixed and such that $p(\mu, s, T, x, \cdot)$ is a density function. We define the measure-valued map $\Theta : (s, \mu) \in [0,T) \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \Theta(s,\mu)(dy) := (\int_{\mathbb{R}^d} p(\mu, s, T, x, y)\, d\mu(x))\, dy \in \mathcal{P}(\mathbb{R}^d)$. We assume that the following properties hold true.*

1. *For any compact subset $\mathcal{K}$ of $[0,T) \times \mathcal{P}_\beta(\mathbb{R}^d)$,*
$$\int_{\mathbb{R}^d} \sup_{(s,\mu)\in\mathcal{K}} |y|^\beta \Theta(s,\mu)(dy) < +\infty.$$
2. *For all $y \in \mathbb{R}^d$, the map $(\mu, s, x) \in \mathcal{P}_\beta(\mathbb{R}^d) \times [0,T) \times \mathbb{R}^d \mapsto p(\mu, s, T, x, y)$ belongs to $\mathcal{C}^1(\mathcal{P}_\beta(\mathbb{R}^d) \times [0,T) \times \mathbb{R}^d)$.*
3. *For any compact subset $\mathcal{K}$ of $\mathcal{P}_\beta(\mathbb{R}^d) \times [0,T) \times \mathbb{R}^d \times \mathbb{R}^d$ and for any $j \in \{0,1\}$*
$$\int_{\mathbb{R}^d} \sup_{(\mu,s,x,v)\in\mathcal{K}} \left\{ \left|\partial_v^j \frac{\delta}{\delta m} p(\mu, s, T, x, y)(v)\right| + \left|\partial_x^j p(\mu, s, T, x, y)\right| + |\partial_s p(\mu, s, T, x, y)| \right\} dy < +\infty.$$
4. *For any compact subset $\mathcal{K}$ of $\mathcal{P}_\beta(\mathbb{R}^d) \times [0,T)$, there exists a positive constant C such that*
$$\int_{\mathbb{R}^d} (1 + |y|^\beta) \sup_{(\mu,s)\in\mathcal{K}} |p(\mu, s, T, v, y)|\, dy \le C(1 + |v|^\beta),$$
$$\int_{\mathbb{R}^d} (1 + |y|^\beta) \sup_{(\mu,s)\in\mathcal{K}} \left|\frac{\delta}{\delta m} p(\mu, s, T, x, y)(v)\right| dy \le C(1 + |x|^\beta)(1 + |v|^\beta),$$
and
$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} \sup_{(\mu,s)\in\mathcal{K}} \left|\frac{\delta}{\delta m} p(\mu, s, T, x, y)(v)\right| dy \le C(1 + |v|^\beta).$$

Then, the function $(s, \mu) \in [0,T) \times \mathcal{P}_\beta(\mathbb{R}^d) \mapsto \phi(\Theta(s,\mu))$ belongs to $\mathcal{C}^1([0,T) \times \mathcal{P}_\beta(\mathbb{R}^d))$. Moreover, we have
$$\partial_s [\phi(\Theta(s,\mu))] = \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi(\Theta(s,\mu))(y) - \frac{\delta}{\delta m}\phi(\Theta(s,\mu))(x) \right) \partial_s p(\mu, s, T, x, y)\, dy\, d\mu(x),$$

$$\begin{aligned}\frac{\delta}{\delta m} [\phi(\Theta(s,\mu))](v) &= \int_{\mathbb{R}^d} \frac{\delta}{\delta m}\phi(\Theta(s,\mu))(y) p(\mu, s, T, v, y)\, dy \\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi(\Theta(s,\mu))(y) - \frac{\delta}{\delta m}\phi(\Theta(s,\mu))(x) \right) \frac{\delta}{\delta m} p(\mu, s, T, x, y)(v)\, dy\, d\mu(x),\end{aligned}$$

and
$$\begin{aligned}\partial_v \frac{\delta}{\delta m} [\phi(\Theta(s,\mu))](v) &= \int_{\mathbb{R}^d} \left(\frac{\delta}{\delta m}\phi(\Theta(s,\mu))(y) - \frac{\delta}{\delta m}\phi(\Theta(s,\mu))(v) \right) \partial_x p(\mu, s, T, v, y)\, dy \\ &\quad + \int_{\mathbb{R}^{2d}} \left(\frac{\delta}{\delta m}\phi(\Theta(s,\mu))(y) - \frac{\delta}{\delta m}\phi(\Theta(s,\mu))(x) \right) \partial_v \frac{\delta}{\delta m} p(\mu, s, T, x, y)(v)\, dy\, d\mu(x).\end{aligned}$$

We now focus on Itô's formula along the flow of probability measures associated with a jump process. Let us fix $Z^1 = (Z^1_t)_t$ and $Z^2 = (Z^2_t)_t$ two α-stable processes on $\mathbb{R}^d$ with $\alpha \in (1,2)$. Their associated Poisson random measures are respectively denoted by $\mathcal{N}^1$ and $\mathcal{N}^2$, their compensated Poisson random measures by $\widetilde{\mathcal{N}}^1$ and $\widetilde{\mathcal{N}}^2$ and their Lévy measures by ν^1 and ν^2. Since $\alpha \in (1,2)$, we can write for all $t \geq 0$

$$Z_t^1 = \int_0^t \int_{\mathbb{R}^d} z \widetilde{\mathcal{N}}^1(ds,dz) \quad \text{and} \quad Z_t^2 = \int_0^t \int_{\mathbb{R}^d} z \widetilde{\mathcal{N}}^2(ds,dz).$$

We fix $\beta \in (1,\alpha)$ and $\gamma \in (0,1]$ such that $\gamma > \alpha - 1$. We consider two jump processes $X = (X_t)_{t\in[0,T]}$ and $Y = (Y_t)_{t\in[0,T]}$ defined for all $t \in [0,T]$ by

$$X_t := X_0 + \int_0^t b_s \, ds + Z_t^1, \quad \text{and} \quad Y_t = Y_0 + \int_0^t \eta_s \, ds + Z_t^2, \tag{8.217}$$

where $X_0, Y_0 \in L^\beta(\Omega, \mathcal{F}_0)$, $b, \eta : [0,T] \times \Omega \to \mathbb{R}^d$ are bounded predictable processes. The distribution of X_t is denoted by μ_t.

We state in the next proposition Itô's formula, which is deduced from [Cav22a, Theorem 2] for the specific type of processes that are considered in the present work.

Proposition 8.40 (Itô's formula)**.** *Let $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d) \to \mathbb{R}$ be a continuous function satisfying the following properties.*

1. *The function u belongs to $\mathcal{C}^1([0,T]\times\mathbb{R}^d\times\mathcal{P}_\beta(\mathbb{R}^d))$ and $\partial_x u(t,\cdot,\mu)$ is γ-Hölder continuous uniformly with respect to t and μ.*

2. *For all compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*

$$\forall t \in [0,T], \, \forall (x,\mu) \in \mathcal{K}, \, \forall v \in \mathbb{R}^d, \, \left| \frac{\delta}{\delta m} u(t,x,\mu)(v) \right| \leq C_\mathcal{K}(1 + |v|^\beta).$$

3. *If $\gamma > 0$, for any compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, there exists $C_\mathcal{K} > 0$ such that*

$$\forall t \in [0,T], \, \forall (x,\mu) \in \mathcal{K}, \, \forall v, v' \in \mathbb{R}^d, \, \left| \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v) - \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v') \right| \leq C_\mathcal{K}|v - v'|^\gamma.$$

4. *For any compact $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_\beta(\mathbb{R}^d)$, we have*

$$\sup_{t\in[0,T]} \sup_{(x,\mu)\in\mathcal{K}} \int_{\mathbb{R}^d} \left| \partial_v \frac{\delta}{\delta m} u(t,x,\mu)(v) \right|^{\beta'} d\mu(v) < +\infty.$$

Then, the function $(t,x) \in [0,T]\times\mathbb{R}^d \mapsto u(t,x,\mu_t)$ is of class $\mathcal{C}^1$, with $\partial_x u(t,\cdot,\mu_t)$ γ-Hölder continuous *uniformly with respect to t. Moreover, we have almost surely for all $t \in [0,T]$*

$$\begin{aligned}
&u(t, Y_t, \mu_t) - u(0, Y_0, \mu_0)\\
&= \int_0^t \partial_t u(s, Y_s, \mu_s)\, ds + \int_0^t \overline{\mathbb{E}}\left(\partial_v \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_s) \cdot \bar{b}_s\right) ds\\
&\quad + \int_0^t \int_{\mathbb{R}^d} \overline{\mathbb{E}}\bigg[\frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-} + z) - \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-})\\
&\qquad\qquad - \partial_v \frac{\delta}{\delta m} u(s, Y_s, \mu_s)(\overline{X}_{s^-}) \cdot z\bigg]\, d\nu^1(z)\, ds \qquad (8.218)\\
&\quad + \int_0^t \partial_x u(s, Y_s, \mu_s) \cdot \eta_s\, ds + \int_0^t \int_{\mathbb{R}^d} u(s, Y_{s^-} + z, \mu_s) - u(s, Y_{s^-}, \mu_s)\, \widetilde{\mathcal{N}}^2(ds, dz)\\
&\quad + \int_0^t \int_{\mathbb{R}^d} [u(s, Y_{s^-} + z, \mu_s) - u(s, Y_{s^-}, \mu_s) - \partial_x u(s, Y_{s^-}, \mu_s) \cdot z]\, d\nu^2(z)\, ds
\end{aligned}$$

where $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ is an independent copy of $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\bar{b}, \overline{X})$ is a copy of (b, X).

8.10 Appendix: Parametrix expansion for stable-driven SDEs

Let us fix $Z = (Z_t)_t$ a rotationally invariant α-stable process on $\mathbb{R}^d$ with $\alpha \in (1,2)$. Its associated Poisson random measure is denoted by $\mathcal{N}$, the compensated Poisson random measure by $\widetilde{\mathcal{N}}$. Since $\alpha \in (1,2)$, we can write for all $t \geq 0$

$$Z_t = \int_0^t \int_{\mathbb{R}^d} z\, \widetilde{\mathcal{N}}(ds, dz).$$

The Lévy measure ν of Z is given by

$$d\nu(z) := \frac{dz}{|z|^{d+\alpha}}.$$

We consider a function $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying the following properties.

1. The function b is jointly continuous and globally bounded on $[0,T] \times \mathbb{R}^d$.
2. The function b is η-Hölder continuous on $\mathbb{R}^d$ uniformly in time, with $\eta \in (0,1]$, i.e. there exists $C > 0$ such that for all $t \in [0,T]$ and $x_1, x_2 \in \mathbb{R}^d$

$$|b(t, x_1) - b(t, x_2)| \leq C|x_1 - x_2|^\eta.$$

We fix $s \in [0,T)$ and we consider the following stable-driven SDE

$$\begin{cases} dX_t^{s,x} = b(t, X_t^{s,x})\, dt + dZ_t, & t \in [s,T],\\ X_s^{s,x} = x \in \mathbb{R}^d. \end{cases} \tag{8.219}$$

It is well-posed in the weak sense by [MP14]. The density of Z_t is denoted by $q(t,\cdot)$. We denote by $\Delta^{\frac{\alpha}{2}}$ the fractional Laplacian associated with Z defined for all $f \in \mathcal{C}_b^{1+\gamma}(\mathbb{R}^d;\mathbb{R})$, with $\gamma > \alpha - 1$ (i.e. f belongs to $\mathcal{C}_b^1(\mathbb{R}^d;\mathbb{R})$ and ∇f is γ-Hölder continuous) and for all $x \in \mathbb{R}^d$ by

$$\Delta^{\frac{\alpha}{2}} f(x) := \int_{\mathbb{R}^d} (f(x+z) - f(x) - \nabla f(x) \cdot z)\, d\nu(z). \tag{8.220}$$

We define for all $s \in [0,T)$, $t \in (s,T]$ and $x, y \in \mathbb{R}^d$

$$\begin{aligned} \widehat{p}(s,t,x,y) &:= q(t-s, y-x), \\ \mathcal{H}(s,t,x,y) &:= b(s,x) \cdot \partial_x \widehat{p}(s,t,x,y). \end{aligned} \tag{8.221}$$

Note that the proxy $\widehat{p}(s,t,x,\cdot)$ is the density at time $t > s$ of the solution to

$$\begin{cases} d\widehat{X}_t^{s,x} = dZ_t, \quad t \in [s,T], \\ \widehat{X}_s^{s,x} = x \in \mathbb{R}^d, \end{cases} \tag{8.222}$$

and $\mathcal{H}$ is the associated parametrix kernel. We also define the space-time convolution operator between to functions f and g by

$$f \otimes g(s,t,x,y) := \int_s^t \int_{\mathbb{R}^d} f(s,r,x,z) g(r,t,z,y)\, dz\, dr, \tag{8.223}$$

when it is well-defined. The space-time convolution iterates $\mathcal{H}^k$ of $\mathcal{H}$ are defined recursively by $\mathcal{H}^1 := \mathcal{H}$ and $\mathcal{H}^{k+1} := \mathcal{H} \otimes \mathcal{H}^k$. By convention $f \otimes \mathcal{H}^0$ is equal to f. Finally, we denote by Φ the solution to the following Volterra integral equation

$$\Phi(s,t,x,y) = \mathcal{H}(s,t,x,y) + \mathcal{H} \otimes \Phi(s,t,x,y),$$

which is given by the uniform convergent series

$$\Phi(s,t,x,y) = \sum_{k=1}^{\infty} \mathcal{H}^k(s,t,x,y). \tag{8.224}$$

Let us also define, for $k > -\alpha$ the function ρ^k by

$$\forall t > 0,\, x \in \mathbb{R}^d,\, \rho^k(t,x) := t^{-\frac{d}{\alpha}} (1 + t^{-\frac{1}{\alpha}} |x|)^{-d-\alpha-k}. \tag{8.225}$$

Theorem 8.41. *For any $s \in [0,T)$, $s < t \leq T$ and $x \in \mathbb{R}^d$, the distribution of $X_t^{s,x}$ has a density with respect to the Lebesgue measure denoted by $p(s,t,x,\cdot)$ and given by the absolutely convergent parametrix series*

$$\begin{aligned} p(s,t,x,y) &= \widehat{p}(s,t,x,y) + \sum_{k=1}^{\infty} \widehat{p} \otimes \mathcal{H}^k(s,t,x,y) \\ &= \widehat{p}(s,t,x,y) + \widehat{p} \otimes \Phi(s,t,x,y). \end{aligned} \tag{8.226}$$

For any $t \in (0,T]$ and $y \in \mathbb{R}^d$, $p(\cdot,t,\cdot,y)$ is of class $\mathcal{C}^1$ on $[0,t) \times \mathbb{R}^d$ and $p(\cdot,t,\cdot,y)$, $\partial_s p(\cdot,t,\cdot,y)$ and $\partial_x p(\cdot,t,\cdot,y)$ are continuous on $[0,t) \times \mathbb{R}^d$. The function $p(\cdot,t,\cdot,y)$ is solution to the following backward Kolmogorov PDE

$$\begin{cases} \partial_s p(s,t,x,y) + b(s,x) \cdot \partial_x p(s,t,x,y) + \Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x) = 0, \quad \forall (s,x) \in [0,t) \times \mathbb{R}^d, \\ p(s,t,x,\cdot) \underset{s \to t^-}{\longrightarrow} \delta_x, \end{cases} \tag{8.227}$$

where $p(s,t,x,\cdot) \underset{s\to t^-}{\longrightarrow} \delta_x$ *means that for all function* $f:\mathbb{R}^d\to\mathbb{R}$ *bounded and uniformly continuous, one has*

$$\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} f(y)p(s,t,x,y)\,dy - f(x)\right| \underset{s\to t^-}{\to} 0.$$

Moreover, p *satisfies the following estimates.*

- *There exists* $C>0$ *such that for all* $j\in\{0,1\}$, $0\le s<t\le T$ *and* $x,y\in\mathbb{R}^d$

$$|\partial_x^j p(s,t,x,y)| \le C(t-s)^{-\frac{j}{\alpha}}\rho^j(t-s,y-x). \tag{8.228}$$

$$|\Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x)| \le C(t-s)^{-1}\rho^0(t-s,y-x). \tag{8.229}$$

- *For all* $j\in\{0,1\}$ *and* $\gamma\in(0,1]$ *with* $\gamma\in(0,(2\alpha-2)\wedge(\eta+\alpha-1))$ *if* $n=1$, *there exists* $C>0$ *such that for all* $0\le s<t\le T$ *and* $x_1,x_2,y\in\mathbb{R}^d$

$$|\partial_x^j p(s,t,x_1,y)-\partial_x^j p(s,t,x_2,y)| \le C(t-s)^{-\frac{j+\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^j(t-s,y-x_1)+\rho^j(t-s,y-x_2)\right]. \tag{8.230}$$

Before proving Theorem 8.41, we recall some properties satisfied by the functions ρ^k.

Lemma 8.42. *The functions* ρ^k *satisfy the following properties.*

- *For all* $k>-\alpha$ *and* $\gamma\in[0,1]$ *with* $k-\gamma>-\alpha$, *we have for all* $t>0$ *and* $x\in\mathbb{R}^d$

$$|x|^\gamma t^{-\frac{\gamma}{\alpha}}\rho^k(t,x) \le \rho^{k-\gamma}(t,x). \tag{8.231}$$

- *Let us fix* $-\alpha<k_1\le k_2$. *Then, for all function* $y:(0,+\infty)\to\mathbb{R}^d$ *such that* $t\in(0,+\infty)\mapsto t^{-\frac{1}{\alpha}}y(t)$ *is bounded, there exists* $C>0$ *such that for all* $t>0$ *and* $x\in\mathbb{R}^d$

$$\rho^{k_2}(t,x+y(t)) \le C\rho^{k_1}(t,x). \tag{8.232}$$

- *For all* $k>-\alpha$ *and* $R>0$, *there exists* C *such that for all* $t>0$, $y\in\mathbb{R}^d$ *and* $x\in\mathbb{R}^d$ *with* $|x|\le R$

$$\rho^k(t,y+x) \le (1+ct^{-\frac{1}{\alpha}}R)^{d+\alpha+k}\rho^k(t,y). \tag{8.233}$$

- *For all* $k_1,k_2>-\alpha$, *there exists* $C>0$ *such that for all* $s\ge0$, $t>s$ *and* $y\in\mathbb{R}^d$

$$\int_{\mathbb{R}^d}\rho^{k_1}(t-s,y-z)\rho^{k_2}(s,z)\,dz \le C\rho^{k_1\wedge k_2}(t,y). \tag{8.234}$$

The following lemma gathers the properties that we need on the proxy $\widehat{p}$.

Lemma 8.43. *For all* $t\in(0,T]$, $y\in\mathbb{R}^d$, $\widehat{p}(\cdot,t,\cdot,y)$ *is of class* $\mathcal{C}^{1,\infty}$ *on* $[0,t)\times\mathbb{R}^d$. *Moreover, it satisfies the following gradient estimates.*

- *For all* $j\in\mathbb{N}$, *there exists* $C>0$ *such that for all* $t\in(0,T]$, $s\in[0,t)$, $x,y\in\mathbb{R}^d$, *we have*

$$|\partial_x^j \widehat{p}(s,t,x,y)| \le C(t-s)^{-\frac{j}{\alpha}}\rho^j(t-s,y-x). \tag{8.235}$$

- *There exists a constant $C > 0$ such that for all $j \in \{0, 1\}$, $t \in (0, T]$, $s \in [0, t)$, $x, y \in \mathbb{R}^d$*

$$|\partial_s \partial_x^j \widehat{p}(s, t, x, y)| \leq C(t - s)^{-1-\frac{j}{\alpha}} \rho^j(t - s, y - x). \tag{8.236}$$

- *For all $j \in \mathbb{N}$, there exists $C > 0$ such that for all $\gamma \in (0, 1]$, $t \in (0, T]$, $s \in [0, t)$, $x_1, x_2, y \in \mathbb{R}^d$, we have*

$$|\partial_x^j \widehat{p}(s, t, x_1, y) - \partial_x^j \widehat{p}(s, t, x_2, y)| \leq C(t - s)^{-\frac{j+\gamma}{\alpha}} |x_1 - x_2|^\gamma \left[\rho^j(t - s, y - x_1) + \rho^j(t - s, y - x_2)\right]. \tag{8.237}$$

- *For all $j \in \{0, 1\}$, there exists $C > 0$ such that for all $\gamma \in (0, 1]$, $t \in (0, T]$, $s_1, s_2 \in [0, t)$, $x, y \in \mathbb{R}^d$, we have*

$$\begin{aligned} &|\partial_x^j \widehat{p}(s_1, t, x, y) - \partial_x^j \widehat{p}(s_2, t, x, y)| \\ &\leq C \left[\frac{|s_1 - s_2|^\gamma}{(t - s_1 \wedge s_2)^{\gamma + \frac{j}{\alpha}}} \rho^j(t - s_1 \wedge s_2, y - x) + \frac{|s_1 - s_2|^\gamma}{(t - s_1 \vee s_2)^{\gamma + \frac{j}{\alpha}}} \rho^j(t - s_1 \vee s_2, y - x)\right]. \end{aligned} \tag{8.238}$$

Proof. Notice that (8.235) and (8.237) are quite standard since $\widehat{p}(s, t, x, y) = q(t - s, y - x)$, where $q(t-s, \cdot)$ is the density of the α-stable random variable Z_{t-s} (see [MZ22, Lemma 2.8]). We prove (8.236). To do this, we remark that the self-similarity of the α-stable process Z implies that

$$\widehat{p}(s, t, x, y) = q(t - s, y - x) = (t - s)^{-\frac{d}{\alpha}} q\left(1, \frac{y - x}{(t - s)^{\frac{1}{\alpha}}}\right) = (t - s)^{-\frac{d}{\alpha}} \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y).$$

This yields

$$\begin{aligned} \partial_s \widehat{p}(s, t, x, y) = &\frac{d}{\alpha}(t - s)^{-\frac{d}{\alpha}-1} \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y) \\ &+ (t - s)^{-\frac{d}{\alpha}} \partial_x \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y) \cdot \left(\frac{1}{\alpha}(t - s)^{-\frac{1}{\alpha}-1}(x - y)\right). \end{aligned} \tag{8.239}$$

Using (8.235), we obtain that

$$|\partial_s \widehat{p}(, s, t, xy)| \leq C(t - s)^{-1} \rho^0(t - s, y - x) + C(t - s)^{-1}((t - s)^{-\frac{1}{\alpha}} |x - y|) \rho^1(1, (t - s)^{-\frac{1}{\alpha}}(y - x)).$$

The space-time inequality (8.231) finally yields

$$|\partial_s \widehat{p}(, s, t, xy)| \leq C(t - s)^{-1} \rho^0(t - s, y - x).$$

By differentiating (8.239) with respect to x, one has

$$\begin{aligned} \partial_x \partial_s \widehat{p}(s, t, x, y) = &\frac{d}{\alpha}(t - s)^{-\frac{1}{\alpha}}(t - s)^{-\frac{d}{\alpha}-1} \partial_x \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y) \\ &+ (t - s)^{-\frac{1}{\alpha}}(t - s)^{-\frac{d}{\alpha}} \partial_x^2 \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y) \left(\frac{1}{\alpha}(t - s)^{-\frac{1}{\alpha}-1}(x - y)\right) \\ &+ \frac{1}{\alpha}(t - s)^{-\frac{1}{\alpha}-1}(t - s)^{-\frac{d}{\alpha}} \partial_x \widehat{p}(0, 1, (t - s)^{-\frac{1}{\alpha}} x, (t - s)^{-\frac{1}{\alpha}} y). \end{aligned} \tag{8.240}$$

As previously, it follows from (8.235) and the space-time inequality (8.231) that

$$|\partial_s \partial_x \widehat{p}(s,t,x,y)| \leq C(t-s)^{-1-\frac{j}{\alpha}} \rho^1(t-s, y-x).$$

We now use (8.236) to prove (8.238). We fix $j \in \{0,1\}$, $\gamma \in (0,1]$ and we start with the case $|s_1 - s_2| > t - s_1 \vee s_2$. In this case, using (8.235), we deduce that for some constant $C > 0$, one has

$$\begin{aligned}
&|\partial_x^j \widehat{p}(s_1,t,x,y) - \partial_x^j \widehat{p}(s_2,t,x,y)| \\
&\leq C \Big[(t-s_1 \vee s_2)^{-\frac{j}{\alpha}} \rho^j(t-s_1\vee s_2, y-x) + (t-s_1\wedge s_2)^{-\frac{j}{\alpha}} \rho^j(t-s_1\wedge s_2, y-x) \Big] \\
&\leq C \left[\frac{|s_1-s_2|^\gamma}{|t-s_1\vee s_2|^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\vee s_2, y-x) + \frac{|s_1-s_2|^\gamma + |t-s_1\vee s_2|^\gamma}{|t-s_1\wedge s_2|^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\wedge s_2, y-x) \right] \\
&\leq C \left[\frac{|s_1-s_2|^\gamma}{|t-s_1\vee s_2|^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\vee s_2, y-x) + \frac{|s_1-s_2|^\gamma}{|t-s_1\wedge s_2|^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\wedge s_2, y-x) \right].
\end{aligned}$$

We now focus on the case $|s_1 - s_2| \leq t - s_1 \vee s_2$. For $\lambda \in [0,1]$, we set $s_\lambda := \lambda s_1 + (1-\lambda) s_2$. We can thus write, thanks to (8.236),

$$\begin{aligned}
&|\partial_x^j \widehat{p}(s_1,t,x,y) - \partial_x^j \widehat{p}(s_2,t,x,y)| \\
&\leq \int_0^1 |\partial_s \partial_x^j \widehat{p}(s_\lambda,t,x,y)|\, |s_1-s_2| \, d\lambda \\
&\leq C|s_1-s_2| \int_0^1 (t-s_\lambda)^{-1-\frac{j}{\alpha}} \rho^j(t-s_\lambda, y-x)\, d\lambda \\
&\leq C|s_1-s_2| \int_0^1 (t-s_\lambda)^{-1-\frac{j+d}{\alpha}} (1+(t-s_\lambda)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha-j}\, d\lambda \\
&\leq C|s_1-s_2|^\gamma (t-s_1\vee s_2)^{1-\gamma} (t-s_1\vee s_2)^{-1-\frac{j+d}{\alpha}} \Big[(1+(t-s_1\vee s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha-j} \\
&\qquad\qquad + (1+(t-s_1\wedge s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha-j} \Big].
\end{aligned}$$

Since $|s_1 - s_2| \leq t - s_1 \vee s_2$, we easily check that $(t-s_1\vee s_2)^{-1} \leq 2(t-s_1\wedge s_2)^{-1}$. It follows that

$$\begin{aligned}
&|\partial_x^j \widehat{p}(s_1,t,x,y) - \partial_x^j \widehat{p}(s_2,t,x,y)| \\
&\leq C|s_1-s_2|^\gamma \Big[(t-s_1\vee s_2)^{-\gamma-\frac{j+d}{\alpha}} (1+(t-s_1\vee s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha-j} \\
&\qquad\qquad + (t-s_1\wedge s_2)^{-\gamma-\frac{j+d}{\alpha}} (1+(t-s_1\wedge s_2)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha-j} \Big] \\
&\leq C \left[\frac{|s_1-s_2|^\gamma}{(t-s_1\wedge s_2)^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\wedge s_2, y-x) + \frac{|s_1-s_2|^\gamma}{(t-s_1\vee s_2)^{\gamma+\frac{j}{\alpha}}} \rho^j(t-s_1\vee s_2, y-x) \right].
\end{aligned}$$

This concludes the proof.

□

Recall that the Beta function $\mathcal{B}$ is defined, for all $x, y > 0$, by

$$\mathcal{B}(x,y) := \int_0^1 (1-t)^{-1+x} t^{-1+y}\, dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

where Γ is the Gamma function.

The next proposition gathers the controls that we need on the parametrix kernel $\mathcal{H}$ and the solution de the Volterra integral equation Φ.

Proposition 8.44. *The following estimates hold true.*

- *There exists $C > 0$ such that for all $k \geq 1$, $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\mathcal{H}^k(s,t,x,y)| \leq C^k (t-s)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k-1} \mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^1(t-s, y-x). \tag{8.241}$$

- *For $\gamma \in (0, \eta]$ such that $\gamma < \alpha - 1$, there exists $C > 0$ depending on γ such that for all $k \geq 1$, $0 \leq s < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$|\mathcal{H}^k(s,t,x_1,y) - \mathcal{H}^k(s,t,x_2,y)| \leq C^k (t-s)^{-\frac{\gamma+1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} |x_1-x_2|^\gamma \prod_{j=1}^{k-1} \mathcal{B}\left(-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \left[\rho^1(t-s, y-x_1) + \rho^1(t-s, y-x_2)\right]. \tag{8.242}$$

- *There exists $C > 0$ such that for all $k \geq 1$, $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\widehat{p} \otimes \mathcal{H}^k(s,t,x,y)| \leq C^{k+1} (t-s)^{k\left(1-\frac{1}{\alpha}\right)} \prod_{j=1}^{k} \mathcal{B}\left(\frac{1}{\alpha}+j\left(1-\frac{1}{\alpha}\right), 1-\frac{1}{\alpha}\right) \rho^0(t-s, y-x). \tag{8.243}$$

- *The series (8.224) defining Φ is absolutely convergent and there exists $C > 0$ such that for all $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$*

$$|\Phi(s,t,x,y)| \leq C(t-s)^{-\frac{1}{\alpha}} \rho^1(t-s, y-x). \tag{8.244}$$

- *For $\gamma \in (0, \eta]$ such that $\gamma < \alpha - 1$, there exists $C > 0$ depending on γ such that for all $0 \leq s < t \leq T$ and $x_1, x_2, y \in \mathbb{R}^d$*

$$|\Phi(s,t,x_1,y) - \Phi(s,t,x_2,y)| \leq C(t-s)^{-\frac{\gamma+1}{\alpha}} |x_1-x_2|^\gamma \left[\rho^1(t-s, y-x_1) + \rho^1(t-s, y-x_2)\right]. \tag{8.245}$$

Proof. **Proof of (8.241).** We reason by induction on k. The base case $k = 1$ is clear since b is bounded and by Lemma 8.43. We assume now that (8.241) holds for $\mathcal{H}^k$ and we want to prove it for $\mathcal{H}^{k+1}$. We have thanks to Lemma 8.42.

$$
\begin{aligned}
|\mathcal{H}^{k+1}(s,t,x,y)| &= \left|\int_s^t \int_{\mathbb{R}^d} \mathcal{H}(s,r,x,z)\mathcal{H}^k(r,t,z,y)\,dz\,dr\right| \\
&\le \int_s^t \int_{\mathbb{R}^d} C(r-s)^{-\frac{1}{\alpha}}\rho^1(r-s,z-x)C^k(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)} \\
&\qquad\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-z)\,dz\,dr \\
&\le C^{k+1}\left(\int_s^t (r-s)^{-\frac{1}{\alpha}}(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\,dr\right)\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-s,y-x).
\end{aligned}
$$

Changing variables in $r = s+\lambda(t-s)$ yields

$$
\begin{aligned}
|\mathcal{H}^{k+1}(s,t,x,y)| &\le C^{k+1}(t-s)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\int_0^1 \lambda^{-\frac{1}{\alpha}}(1-\lambda)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\,d\lambda \\
&\qquad\qquad \prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-s,y-x) \\
&\le C^{k+1}(t-s)^{-\frac{1}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-s,y-x).
\end{aligned}
$$

Proof of (8.242). We start with the case $k=1$. We write

$$
\begin{aligned}
|\mathcal{H}(s,t,x_1,y)-\mathcal{H}(s,t,x_2,y)| &= |b(s,x_1)\cdot\partial_x\widehat{p}(s,t,x_1,y)-b(s,x_2)\cdot\partial_x\widehat{p}(s,t,x_2,y)| \\
&\le |b(s,x_1)|\,|\partial_x\widehat{p}(s,t,x_1,y)-\partial_x\widehat{p}(s,t,x_2,y)| \\
&\quad + |\partial_x\widehat{p}(s,t,x_2,y)|\,|b(s,x_1)-b(s,x_2)| \\
&=: I_1 + I_2.
\end{aligned}
$$

Using that b is bounded and Lemma 8.43, we deduce that

$$
I_1 \le C(t-s)^{-\frac{1+\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^1(t-s,y-x_1)+\rho^1(t-s,y-x_2)\right].
$$

Since b is uniformly η-Hölder continuous and bounded, and thus uniformly γ-Hölder continuous because $\gamma\le\eta$, one has by Lemma 8.43

$$
I_2 \le C(t-s)^{-\frac{1}{\alpha}}|x_1-x_2|^\gamma\rho^1(t-s,y-x_2).
$$

We now prove that (8.242) holds for $\mathcal{H}^{k+1}$, with $k\ge 1$. Using the case $k=1$, (8.241) and Lemma 8.42, we obtain that

$$
\begin{aligned}
&|\mathcal{H}^{k+1}(s,t,x_1,y)-\mathcal{H}^{k+1}(s,t,x_2,y)|\\
&=\left|\int_s^t\int_{\mathbb{R}^d}(\mathcal{H}(s,r,x_1,z)-\mathcal{H}(s,r,x_2,z))\mathcal{H}^k(r,t,z,y)\,dz\,dr\right|\\
&\leq\int_s^t\int_{\mathbb{R}^d}C(r-s)^{-\frac{1+\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^1(r-s,z-x_1)+\rho^1(r-s,z-x_2)\right]C^k(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\\
&\qquad\qquad\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-z)\,dz\,dr\\
&\leq C^{k+1}(t-s)^{-\frac{1+\gamma}{\alpha}+k\left(1-\frac{1}{\alpha}\right)}\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right),1-\frac{1+\gamma}{\alpha}\right)\prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\
&\qquad\qquad\left[\rho^1(t-s,y-x_1)+\rho^1(t-s,y-x_2)\right].
\end{aligned}
$$

We conclude noting that

$$
\begin{aligned}
&\mathcal{B}\left(k\left(1-\frac{1}{\alpha}\right),1-\frac{1+\gamma}{\alpha}\right)\prod_{j=1}^{k}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\\
&=\frac{\Gamma\left(k\left(1-\frac{1}{\alpha}\right)\right)\Gamma\left(1-\frac{1+\gamma}{\alpha}\right)}{\Gamma\left(1-\frac{1+\gamma}{\alpha}+k\left(1-\frac{1}{\alpha}\right)\right)}\prod_{j=1}^{k-1}\frac{\Gamma\left(j\left(1-\frac{1}{\alpha}\right)\right)\Gamma\left(1-\frac{1}{\alpha}\right)}{\Gamma\left((j+1)\left(1-\frac{1}{\alpha}\right)\right)}\\
&=\frac{\Gamma\left(k\left(1-\frac{1}{\alpha}\right)\right)\Gamma\left(1-\frac{1+\gamma}{\alpha}\right)}{\Gamma\left(1-\frac{1+\gamma}{\alpha}+k\left(1-\frac{1}{\alpha}\right)\right)}\frac{\left[\Gamma\left(1-\frac{1}{\alpha}\right)\right]^k}{\Gamma\left(k\left(1-\frac{1}{\alpha}\right)\right)}\\
&=\frac{\Gamma\left(1-\frac{1+\gamma}{\alpha}\right)\left[\Gamma\left(1-\frac{1}{\alpha}\right)\right]^k}{\Gamma\left(1-\frac{1+\gamma}{\alpha}+k\left(1-\frac{1}{\alpha}\right)\right)}\\
&=\prod_{j=1}^{k}\frac{\Gamma\left(-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right)\right)\Gamma\left(1-\frac{1}{\alpha}\right)}{\Gamma\left(-\frac{\gamma}{\alpha}+(j+1)\left(1-\frac{1}{\alpha}\right)\right)}\\
&=\prod_{j=1}^{k}\mathcal{B}\left(-\frac{\gamma}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right).
\end{aligned}
$$

Proof of (8.243)**.** It follows from the control of $\widehat{p}$ given by Lemma 8.43 and from (8.241). Indeed, one has thanks to Lemma 8.42

$$
\begin{aligned}
&|\widehat{p}\otimes\mathcal{H}^k(s,t,x,y)|\\
&\leq C\int_s^t\int_{\mathbb{R}^d}\rho^0(r-s,z-x)C^k(t-r)^{-\frac{1}{\alpha}+(k-1)\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^1(t-r,y-z)\,dz\,dr\\
&\leq C^{k+1}\frac{(t-s)^{k\left(1-\frac{1}{\alpha}\right)}}{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)\rho^0(t-s,y-x).
\end{aligned}
$$

We conclude noting that

$$
\begin{aligned}
\frac{1}{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\mathcal{B}\left(j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right) &= \frac{1}{k\left(1-\frac{1}{\alpha}\right)}\prod_{j=1}^{k-1}\frac{\Gamma\left(j\left(1-\frac{1}{\alpha}\right)\right)\Gamma\left(1-\frac{1}{\alpha}\right)}{\Gamma\left((j+1)\left(1-\frac{1}{\alpha}\right)\right)}\\
&= \frac{1}{k\left(1-\frac{1}{\alpha}\right)}\frac{\left[\Gamma\left(1-\frac{1}{\alpha}\right)\right]^k}{\Gamma\left(k\left(1-\frac{1}{\alpha}\right)\right)}\\
&= \frac{\left[\Gamma\left(1-\frac{1}{\alpha}\right)\right]^k}{\Gamma\left(1+k\left(1-\frac{1}{\alpha}\right)\right)}\\
&= \prod_{j=1}^{k}\mathcal{B}\left(\frac{1}{\alpha}+j\left(1-\frac{1}{\alpha}\right),1-\frac{1}{\alpha}\right)
\end{aligned}
$$

Finally, the proof of (8.244) and (8.245) follows directly from (8.241) and (8.242) using the asymptotic expansion of the Beta function.

□

Proof of Theorem 8.41. The existence of the density and its representation (8.226) is a consequence of (8.235), (8.243) and (8.244). Indeed, using the asymptotic expansion of the Beta function, we obtain that the series (8.226) is absolutely convergent, locally uniformly with respect to $(s,x)\in[0,t)\times\mathbb{R}^d$. We refer to [KK18] for a detailed presentation of the parametrix method for SDEs driven by stable processes. The permutation of the series and the convolution in the representation formula (8.226) is clearly justified by the dominated convergence theorem. The regularity of the density p with respect to x and the controls (8.228) and (8.230) follow from Lemma 8.43 for the proxy, i.e. the first term of the parametrix series (8.226). We now prove that they also hold for the other term $\widehat{p}\otimes\Phi$ of (8.226).

Proof of (8.228) **for** $j=1$**.** We use (8.244) and Lemma 8.43 to deduce that

$$
|\partial_x\widehat{p}\otimes\Phi(s,t,x,y)| \le C\int_s^t\int_{\mathbb{R}^d}(r-s)^{-\frac{1}{\alpha}}\rho^1(r-s,z-x)(t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr.
$$

Then, Lemma 8.42 yields

$$
\begin{aligned}
&|\partial_x\widehat{p}\otimes\Phi(s,t,x,y)|\\
&\le C\int_s^t(r-s)^{-\frac{1}{\alpha}}(t-r)^{-\frac{1}{\alpha}}\,dr\rho^1(t-s,y-x)\\
&\le C(t-s)^{-\frac{1}{\alpha}1-\frac{1}{\alpha}}\rho^1(t-s,y-x).
\end{aligned}
$$

This concludes the proof of (8.228) for $j=1$.

Proof of (8.230) **for** $j=0$**.** We use again the control of $\widehat{p}$ given by Lemma 8.43 and (8.244). Thanks to Lemma 8.42, we obtain

$$
\begin{aligned}
&|\widehat{p}\otimes\Phi(s,t,x_1,y)-\widehat{p}\otimes\Phi(s,t,x_2,y)|\\
&=\left|\int_s^t\int_{\mathbb{R}^d}(\widehat{p}(s,r,x_1,z)-\widehat{p}(s,r,x_2,z))\Phi(r,t,z,y)\,dz\,dr\right|\\
&\leq C\int_s^t\int_{\mathbb{R}^d}(r-s)^{-\frac{\gamma}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(r-s,z-x_1)+\rho^0(r-s,z-x_2)\right](t-r)^{-\frac{1}{\alpha}}\rho^1(t-r,y-z)\,dz\,dr\\
&\leq C(t-s)^{-\frac{\gamma}{\alpha}+1-\frac{1}{\alpha}}|x_1-x_2|^\gamma\left[\rho^0(t-s,y-x_1)+\rho^0(t-s,y-x_2)\right].
\end{aligned}
$$

Proof of (8.230) for $j=1$. We use the following decomposition

$$
\begin{aligned}
\partial_x\widehat{p}\otimes\Phi(s,t,x_1,y)-\partial_x\widehat{p}\otimes\Phi(s,t,x_2,y)&=\int_s^t\int_{\mathbb{R}^d}(\partial_x\widehat{p}(s,r,x_1,z)-\partial_x\widehat{p}(s,r,x_2,z))\Phi(r,t,z,y)\,dz\,dr\\
&=\int_{D_1}\int_{\mathbb{R}^d}(\partial_x\widehat{p}(s,r,x_1,z)-\partial_x\widehat{p}(s,r,x_2,z))\Phi(r,t,z,y)\,dz\,dr\\
&\quad+\int_{D_2}\int_{\mathbb{R}^d}(\partial_x\widehat{p}(s,r,x_1,z)-\partial_x\widehat{p}(s,r,x_2,z))\Phi(r,t,z,y)\,dz\,dr\\
&=: I_1+I_2,
\end{aligned}
$$

where $D_1:=\{r\in(s,t),\,|x_1-x_2|>(r-s)^{\frac{1}{\alpha}}\}$ and $D_2:=\{r\in(s,t),\,|x_1-x_2|\leq(r-s)^{\frac{1}{\alpha}}\}$. For I_1, one can write

$$
\begin{aligned}
I_1&=\int_{D_1}\int_{\mathbb{R}^d}\partial_x\widehat{p}(s,r,x_1,z)(\Phi(r,t,z,y)-\Phi(r,t,x_1,y))\,dz\,dr\\
&\quad-\int_{D_1}\int_{\mathbb{R}^d}\partial_x\widehat{p}(s,r,x_2,z)(\Phi(r,t,z,y)-\Phi(r,t,x_2,y))\,dz\,dr,\\
&=: I_{1,1}+I_{1,2},
\end{aligned}
$$

since $\int_{\mathbb{R}^d}\partial_x\widehat{p}(s,r,x,z)\,dz=0$ for all $x\in\mathbb{R}^d$. Since $\alpha\in(1,2)$ and $\gamma<(2\alpha-2)\wedge(\eta+\alpha-1)$, we can pick $\delta\in(0,(\alpha-1)\wedge\eta)$ such that $\gamma<\delta+\alpha-1$. Then, we can use Lemma 8.43 and (8.245) with δ, which yields

$$
|I_{1,1}|\leq C\int_{D_1}\int_{\mathbb{R}^d}(r-s)^{-\frac{1}{\alpha}}\rho^1(r-s,z-x_1)(t-r)^{-\frac{\delta+1}{\alpha}}|z-x_1|^\delta\left[\rho^1(t-r,y-z)+\rho^1(t-r,y-x_1)\right]\,dz\,dr.
$$

Using the space-time inequality (8.231) in Lemma 8.42, we deduce that

$$
|I_{1,1}|\leq C\int_{D_1}\int_{\mathbb{R}^d}(r-s)^{\frac{\delta-1}{\alpha}}\rho^{1-\delta}(r-s,z-x_1)(t-r)^{-\frac{\delta+1}{\alpha}}\left[\rho^1(t-r,y-z)+\rho^1(t-r,y-x_1)\right]\,dz\,dr.
$$

Let us note that $\int_{\mathbb{R}^d}\rho^{1-\delta}(r-s,z-x_1)\,dz$ is a constant independent of r,s,x_1, that if $r\in(s,t)$

$$
\rho^1(t-r,y-x_1)\leq(t-r)^{-\frac{d}{\alpha}}(1+(t-s)^{-\frac{1}{\alpha}}|y-x_1|)^{-d-\alpha-1}
$$

and that if $r\in D_1$, $(r-s)^{\frac{\gamma}{\alpha}}<|x_1-x_2|^\gamma$. Applying Lemma 8.42, we thus get that

$$
\begin{aligned}
|I_{1,1}| \leq C \Bigg[& \int_s^t (r-s)^{\frac{\delta-1-\gamma}{\alpha}} (t-r)^{-\frac{\delta+1}{\alpha}} \, dr |x_1 - x_2|^\gamma \rho^{1-\delta}(t-s, y-x_1) \\
& \int_s^t (r-s)^{\frac{\delta-1-\gamma}{\alpha}} (t-r)^{-\frac{\delta+1+d}{\alpha}} \, dr |x_1 - x_2|^\gamma (1 + (t-s)^{-\frac{1}{\alpha}} |y-x_1|)^{-d-\alpha-1} \Bigg] \\
\leq & \, C(t-s)^{-\frac{1+\gamma}{\alpha}+1-\frac{1}{\alpha}} |x_1 - x_2|^\gamma \rho^{1-\delta}(t-s, y-x_1).
\end{aligned}
$$

Similarly, we obtain

$$
|I_{1,2}| \leq C(t-s)^{-\frac{1+\gamma}{\alpha}+1-\frac{1}{\alpha}} |x_1 - x_2|^\gamma \rho^{1-\delta}(t-s, y-x_2).
$$

This proves (8.230) for I_1. We now focus on I_2. One can write

$$
\begin{aligned}
I_2 &= \int_{D_2} \int_{\mathbb{R}^d} (\partial_x \widehat{p}(s,r,x_1,z) - \partial_x \widehat{p}(s,r,x_2,z)) \Phi(r,t,z,y) \, dz \, dr \\
&= \int_{D_2} \int_{\mathbb{R}^d} (\partial_x \widehat{p}(s,r,x_1,z) - \partial_x \widehat{p}(s,r,x_2,z)) (\Phi(r,t,z,y) - \Phi(r,t,x_2,y)) \, dz \, dr.
\end{aligned}
$$

Note that if $r \in D_2$, we have

$$
\begin{aligned}
|\partial_x \widehat{p}(s,r,x_1,z) - \partial_x \widehat{p}(s,r,x_2,z)| &\leq C(r-s)^{-\frac{2}{\alpha}} |x_1 - x_2| \left[\rho^2(r-s, z-x_1) + \rho^2(r-s, z-x_1) \right] \\
&\leq C(r-s)^{-\frac{\gamma+1}{\alpha}} |x_1 - x_2|^\gamma \rho^2(r-s, z-x_2),
\end{aligned}
$$

since $|x_1 - x_2| \leq (r-s)^{\frac{1}{\alpha}}$ and by (8.232). Using (8.245) with $\delta \in (0, (\alpha-1) \wedge \eta)$ such that $\gamma < \delta + \alpha - 1$ and the space-time inequality (8.231), we get

$$
\begin{aligned}
|I_2| \leq C & \int_{D_2} \int_{\mathbb{R}^d} (r-s)^{-\frac{\gamma+1}{\alpha}} |x_1 - x_2|^\gamma \rho^2(r-s, z-x_2) (t-r)^{-\frac{\delta+1}{\alpha}} |z-x_2|^\delta \\
& \qquad \left[\rho^1(t-r, y-z) + \rho^1(t-r, y-x_2) \right] \, dz \, dr \\
\leq C & \int_{D_2} \int_{\mathbb{R}^d} (r-s)^{-\frac{\gamma+1}{\alpha}+\frac{\delta}{\alpha}} |x_1 - x_2|^\gamma \rho^{2-\delta}(r-s, z-x_2) (t-r)^{-\frac{\delta+1}{\alpha}} \\
& \qquad \left[\rho^1(t-r, y-z) + \rho^1(t-r, y-x_2) \right] \, dz \, dr.
\end{aligned}
$$

As done previously to deal with $I_{1,1}$, Lemma 8.42 yields

$$
|I_2| \leq C(t-s)^{-\frac{\gamma+1}{\alpha}+1-\frac{1}{\alpha}} \left[\rho^1(t-r, y-x_1) + \rho^1(t-r, y-x_2) \right].
$$

It concludes the proof of (8.230) for $j=1$.

Proof of (8.227). Let us now prove that $p(\cdot,t,\cdot,y)$ is a fundamental solution to (8.227). We fix $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$. From the Markov property satisfied by the SDE (8.219), stemming from the well-posedness of the related martingale problem, one has for all $h > 0$ such that $s - h \geq 0$

$$
p(s-h,t,x,y) = \mathbb{E}(p(s,t,X_s^{s-h,x},y)).
$$

Applying Itô's formula to the function $p(s,t,\cdot,y)$ which belongs to $\mathcal{C}_b^{1+\gamma}(\mathbb{R}^d;\mathbb{R})$ for $\gamma > \alpha - 1$, we obtain that

$$\begin{aligned} p(s,t,X_s^{s-h,x},y) = p(s,t,x,y) &+ \int_{s-h}^{s} b(r,X_r^{s-h,x}) \cdot \partial_x p(s,t,X_r^{s-h,x},y)\, dr \\ &+ \int_{s-h}^{s}\int_{\mathbb{R}^d} p(s,t,X_{r^-}^{s-h,x}+z,y) - p(s,t,X_{r^-}^{s-h,x},y)\, \widetilde{\mathcal{N}}(dr,dz) \\ &+ \int_{s-h}^{s}\int_{\mathbb{R}^d} p(s,t,X_{r^-}^{s-h,x}+z,y) - p(s,t,X_{r^-}^{s-h,x},y) - \partial_x p(s,t,X_{r^-}^{s-h,x},y)\cdot z\, d\nu(z)\, dr \end{aligned}$$

□

We can take the expectation in the preceding formula using (8.228), (8.230). It yields

$$\begin{aligned} p(s-h,t,x,y) = p(s,t,x,y) &+ \int_{s-h}^{s} \mathbb{E}(b(r,X_r^{s-h,x}) \cdot \partial_x p(s,t,X_r^{s-h,x},y))\, dr \\ &+ \int_{s-h}^{s}\int_{\mathbb{R}^d} \mathbb{E}(p(s,t,X_r^{s-h,x}+z,y) - p(s,t,X_r^{s-h,x},y) - \partial_x p(s,t,X_r^{s-h,x},y)\cdot z)\, d\nu(z)\, dr \end{aligned}$$

Let us prove that

$$\frac{1}{h}\int_{s-h}^{s} \mathbb{E}(b(r,X_r^{s-h,x}) \cdot \partial_x p(s,t,X_r^{s-h,x},y))\, dr \underset{h\to 0}{\longrightarrow} b(s,x)\cdot \partial_x p(s,t,x,y). \tag{8.246}$$

We can write

$$\begin{aligned} &\left|\frac{1}{h}\int_{s-h}^{s} \mathbb{E}(b(r,X_r^{s-h,x}) \cdot \partial_x p(s,t,X_r^{s-h,x},y))\, dr - b(s,x)\cdot \partial_x p(s,t,x,y)\right| \\ &\le \left|\frac{1}{h}\int_{s-h}^{s} \mathbb{E}(b(r,X_r^{s-h,x}) \cdot \partial_x p(s,t,X_r^{s-h,x},y)) - \mathbb{E}(b(r,X_s^{s-h,x}) \cdot \partial_x p(s,t,X_s^{s-h,x},y))\, dr\right| \\ &\quad + \left|\frac{1}{h}\int_{s-h}^{s} \mathbb{E}(b(r,X_s^{s-h,x}) \cdot \partial_x p(s,t,X_s^{s-h,x},y)) - (b(r,x)\cdot \partial_x p(s,t,x,y))\, dr\right| \\ &\quad + \left|\frac{1}{h}\int_{s-h}^{s} (b(r,x)\cdot \partial_x p(s,t,x,y)) - b(s,x)\cdot \partial_x p(s,t,x,y)\, dr\right| \\ &=: I_1 + I_2 + I_3. \end{aligned}$$

It is clear that I_3 converges to 0 as h tends to 0 since b is continuous. Concerning I_2, one has

$$\begin{aligned} I_2 &\le \frac{1}{h}\int_{s-h}^{s} \mathbb{E}|b(r,X_s^{s-h,x}) \cdot \partial_x p(s,t,X_s^{s-h,x},y)) - (b(r,x)\cdot \partial_x p(s,t,x,y)|\, dr \\ &\le \frac{1}{h}\int_{s-h}^{s} \mathbb{E}|b(r,X_s^{s-h,x})|\, |\partial_x p(s,t,X_s^{s-h,x},y) - \partial_x p(s,t,x,y)|\, dr \\ &\quad + \frac{1}{h}\int_{s-h}^{s} \mathbb{E}|b(r,X_s^{s-h,x}) - b(r,x)|\, |\partial_x p(s,t,x,y)|\, dr \\ &=: I_{2,1} + I_{2,2}. \end{aligned}$$

Since b is bounded and by the Hölder control (8.230), we obtain that for some constant $\gamma \in (0,1)$

$$\begin{aligned} I_{2,1} &\le C_{s,t,\gamma}\mathbb{E}|X_s^{s-h,x} - x|^{\gamma} \\ &\le C_{s,t,\gamma} h^{\frac{\gamma}{\alpha}}. \end{aligned}$$

The same reasoning based on the Hölder continuity of b proves that $I_2 \underset{h\to 0}{\longrightarrow} 0$.

Finally, we decompose I_1 in the following way

$$\begin{aligned}
I_1 &\leq \frac{1}{h}\int_{s-h}^{s} \mathbb{E}|b(r, X_r^{s-h,x})|\,|\partial_x p(s,t,X_r^{s-h,x},y) - \partial_x p(s,t,X_s^{s-h,x},y)|\,dr \\
&\quad + \frac{1}{h}\int_{s-h}^{s} \mathbb{E}|b(r, X_r^{s-h,x}) - b(r, X_s^{s-h,x})|\,|\partial_x p(s,t,X_s^{s-h,x},y)|\,dr \\
&=: I_{1,1} + I_{1,2}.
\end{aligned}$$

Note that for all $\gamma \in (0,1]$, there exists $C_{s,t,\gamma} > 0$ such that for all $r \in [s-h, s]$

$$\mathbb{E}|X_r^{s-h,x} - X_s^{s-h,x}|^\gamma \leq C_{s,t,\gamma}(s-r)^{\frac{\gamma}{\alpha}},$$

The same reasoning as done for I_2 can be applied since b and $\partial_x p(s,t,\cdot,y)$ are globally bounded and b is uniformly Hölder continuous. It yields $I_1 \underset{h\to 0}{\longrightarrow} 0$.

Let us now show that

$$\frac{1}{h}\int_{s-h}^{s}\int_{\mathbb{R}^d} \mathbb{E}(p(s,t,X_r^{s-h,x}+z,y) - p(s,t,X_r^{s-h,x},y) - \partial_x p(s,t,X_r^{s-h,x},y)\cdot z)\,d\nu(z)\,dr \underset{h\to 0}{\longrightarrow} \Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x). \quad (8.247)$$

One can write

$$\begin{aligned}
&\left|\frac{1}{h}\int_{s-h}^{s}\int_{\mathbb{R}^d} \mathbb{E}(p(s,t,X_r^{s-h,x}+z,y) - p(s,t,X_r^{s-h,x},y) - \partial_x p(s,t,X_r^{s-h,x},y)\cdot z)\,d\nu(z)\,dr - \Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x)\right| \\
&\leq \left|\int_{\mathbb{R}^d}\frac{1}{h}\int_{s-h}^{s} \mathbb{E}(p(s,t,X_r^{s-h,x}+z,y) - p(s,t,X_r^{s-h,x},y) - \partial_x p(s,t,X_r^{s-h,x},y)\cdot z)\,dr\,d\nu(z)\right. \\
&\qquad\qquad \left. - \int_{\mathbb{R}^d} \mathbb{E}(p(s,t,X_s^{s-h,x}+z,y) - p(s,t,X_s^{s-h,x},y) - \partial_x p(s,t,X_s^{s-h,x},y)\cdot z)\,d\nu(z)\right| \\
&\quad + \left|\int_{\mathbb{R}^d} \mathbb{E}(p(s,t,X_s^{s-h,x}+z,y) - p(s,t,X_s^{s-h,x},y) - \partial_x p(s,t,X_s^{s-h,x},y)\cdot z)\,d\nu(z) - \Delta^{\frac{\alpha}{2}} p(s,t,\cdot,y)(x)\right| \\
&=: J_1 + J_2.
\end{aligned}$$

For J_1, we obtain that

$$\begin{aligned}
J_1 \leq \int_{\mathbb{R}^d}\frac{1}{h}\int_{s-h}^{s}\int_0^1 \mathbb{E}\Big|&(\partial_x p(s,t,X_r^{s-h,x}+\lambda z,y) - \partial_x p(s,t,X_r^{s-h,x},y)) \\
&-(\partial_x p(s,t,X_s^{s-h,x}+\lambda z,y) - \partial_x p(s,t,X_s^{s-h,x},y))\Big|\;|z|\,d\lambda\,dr\,d\nu(z)
\end{aligned}$$

We are going to use the dominated convergence theorem in the integral with respect to ν. By the Hölder control (8.230) on $\partial_x p$, we deduce that for some $\gamma > \alpha - 1$, there exists a constant $C_{s,t,\gamma} > 0$ such that for all $r \in [s-h, s]$, $z \in \mathbb{R}^d$

$$\begin{aligned}
&\frac{1}{h}\int_{s-h}^{s}\int_0^1 \mathbb{E}\Big|(\partial_x p(s,t,X_r^{s-h,x}+\lambda z,y)-\partial_x p(s,t,X_r^{s-h,x},y)) \\
&\qquad\qquad\qquad -(\partial_x p(s,t,X_s^{s-h,x}+\lambda z,y)-\partial_x p(s,t,X_s^{s-h,x},y))\Big|\,|z|\,d\lambda\,dr \\
&\le C_{s,t,\gamma}\frac{1}{h}\int_{s-h}^{s}\mathbb{E}|X_r^{s-h,x}-X_s^{s-h,x}|^{\gamma}\,dr\,|z| \\
&\le C_{s,t,\gamma}\frac{1}{h}\int_{s-h}^{s}(r-s)^{\frac{\gamma}{\alpha}}\,dr\,|z| \\
&\le C_{s,t,\gamma}h^{\frac{\gamma}{\alpha}}|z|.
\end{aligned}$$

The right-hand side term tends to 0 when $h \to 0$. We start by justifying the domination on the ball B_1. In this case, we use again the Hölder continuity of $\partial_x p$ with respect to x, which yields

$$\begin{aligned}
&\frac{1}{h}\int_{s-h}^{s}\int_0^1 \mathbb{E}\Big|(\partial_x p(s,t,X_r^{s-h,x}+\lambda z,y)-\partial_x p(s,t,X_r^{s-h,x},y)) \\
&\qquad\qquad\qquad -(\partial_x p(s,t,X_s^{s-h,x}+\lambda z,y)-\partial_x p(s,t,X_s^{s-h,x},y))\Big|\,|z|\,d\lambda\,dr \\
&\le C_{s,t,\gamma}|z|^{1+\gamma}.
\end{aligned}$$

Since $\gamma > \alpha - 1$, the map $z \in B_1 \mapsto |z|^{1+\gamma}$ belongs to $L^1(B_1,\nu)$. The domination on B_1^c is clear since $\partial_x p(s,t,\cdot,y)$ is globally bounded by (8.228) and $z \in B_1^c \mapsto |z|$ belongs to $L^1(B_1^c,\nu)$.

We now deal with J_2. Thanks to the dominated convergence theorem, J_2 converges to 0 as h tends to 0. The Hölder control (8.230) of $\partial_x p$ ensures that for some constant $\gamma > \alpha - 1$, there exists $C_{s,t,\gamma} > 0$ such that

$$\mathbb{E}|p(s,t,X_s^{s-h,x}+z,y)-p(s,t,X_s^{s-h,x},y)-\partial_x p(s,t,X_s^{s-h,x},y)\cdot z| \le C_{s,t,\gamma}|z|^{1+\gamma}.$$

This proves the domination on the ball B_1. The domination on B_1^c is a consequence of (8.228), and the fact that $\alpha \in (1,2)$.

We have thus proved that the map $s \in [0,t) \mapsto p(s,t,x,y)$ is left-differentiable by (8.246) and (8.247). Moreover, since the map $(s,x) \in [0,t)\times\mathbb{R}^d \mapsto b(s,x)\cdot\partial_x p(s,t,x,y)+\Delta^{\frac{\alpha}{2}}p(s,t,\cdot,y)(x)$ is continuous, we deduce that the map $s \in [0,t) \mapsto p(s,t,x,y)$ is of class $\mathcal{C}^1$ and that it solves

$$\partial_s p(s,t,x,y)+b(s,x)\cdot\partial_x p(s,t,x,y)+\Delta^{\frac{\alpha}{2}}p(s,t,\cdot,y)(x)=0, \quad \forall (s,x)\in[0,t)\times\mathbb{R}^d.$$

Let us now fix $f:\mathbb{R}^d\to\mathbb{R}$ a bounded and uniformly continuous function. We fix $\varepsilon > 0$. There exists $\delta > 0$ such that for all $x,y \in \mathbb{R}^d$ with $|x-y| \le \delta$, we have $|f(x)-f(y)| \le \varepsilon$. Using (8.228), we obtain that

$$\begin{aligned}\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} f(y)p(s,t,x,y)\,dy - f(x)\right| &= \sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d}(f(y)-f(x))p(s,t,x,y)\,dy\right|\\ &\leq \varepsilon + C\|f\|_\infty \int_{|y|>\delta}(t-s)^{-\frac{d}{\alpha}}(1+(t-s)^{-\frac{1}{\alpha}}|y|)^{-d-\alpha}\,dy\\ &= \varepsilon + C\|f\|_\infty \int_{|z|>(t-s)^{-\frac{1}{\alpha}}\delta}(1+|z|)^{-d-\alpha}\,dz.\end{aligned}$$

We conclude, taking the lim sup when $s\to t$ in the preceding inequality, that $p(s,t,x,\cdot) \underset{s\to t^-}{\longrightarrow} \delta_x$.

Proof of (8.229)**.** We now prove that

$$|\Delta^{\frac{\alpha}{2}}p(s,t,\cdot,y)(x)| \leq C(t-s)^{-1}\rho^0(t-s,y-x),$$

using the PDE (8.227) and the fact that b is bounded and (8.228). Note that by symmetry, we have

$$\Delta^{\frac{\alpha}{2}}p(s,t,\cdot,y)(x) = \frac{1}{2}\int_{\mathbb{R}^d} p(s,t,x+z,y)+p(s,t,x-z,y)-2p(s,t,x,y)\,\frac{dz}{|z|^{d+\alpha}}.$$

We decompose it in the following way

$$\begin{aligned}|\Delta^{\frac{\alpha}{2}}p(s,t,\cdot,y)(x)| \leq &\int_{|z|\leq(t-s)^{\frac{1}{\alpha}}} |p(s,t,x+z,y)+p(s,t,x-z,y)-2p(s,t,x,y)|\,\frac{dz}{|z|^{d+\alpha}}\\ &+\int_{|z|>(t-s)^{\frac{1}{\alpha}}} |p(s,t,x+z,y)+p(s,t,x-z,y)-2p(s,t,x,y)|\,\frac{dz}{|z|^{d+\alpha}}\\ =&\int_{|z|\leq(t-s)^{\frac{1}{\alpha}}} \left|\int_0^1(\partial_x p(s,t,x+\lambda z,y)-\partial_x p(s,t,x-\lambda z,y))\cdot z\,d\lambda\right|\,\frac{dz}{|z|^{d+\alpha}}\\ &+\int_{|z|>(t-s)^{\frac{1}{\alpha}}} |p(s,t,x+z,y)+p(s,t,x-z,y)-2p(s,t,x,y)|\,\frac{dz}{|z|^{d+\alpha}}\\ =:&\, I_1+I_2.\end{aligned}$$

We start with I_1. Using the Hölder control (8.230), we obtain that for some $\gamma\in(0,(2\alpha-2)\wedge(\eta+\alpha-1))$ with $\gamma>\alpha-1$, there exists a constant $C>0$ such that

$$I_1 \leq C(t-s)^{-\frac{1+\gamma}{\alpha}}\int_{|z|\leq(t-s)^{\frac{1}{\alpha}}}\int_0^1(\rho^1(t-s,y-x-\lambda z)+\rho^1(t-s,y-x+\lambda z))\,d\lambda|z|^{1+\gamma}\,\frac{dz}{|z|^{d+\alpha}}.$$

Since $|z|\leq(t-s)^{\frac{1}{\alpha}}$, the space-time inequality (8.231) ensures that

$$\begin{aligned}I_1 &\leq C(t-s)^{-\frac{1+\gamma}{\alpha}}\int_{|z|\leq(t-s)^{\frac{1}{\alpha}}}\rho^1(t-s,y-x)|z|^{1+\gamma}\,\frac{dz}{|z|^{d+\alpha}}\\ &\leq C(t-s)^{-1}\rho^1(t-s,y-x).\end{aligned}\tag{8.248}$$

For I_2, we have with (8.228)

$$I_2 \leq C\int_{|z|>(t-s)^{\frac{1}{\alpha}}}\left[\rho^0(t-s,y-x-z)+\rho^0(t-s,y-x+z)+\rho^0(t-s,y-x)\right]\,\frac{dz}{|z|^{d+\alpha}}.$$

We distinguish two cases. Firstly, assume that $|y-x| \leq 2(t-s)^{\frac{1}{\alpha}}$. Then, using that for all $s<t$ and $y \in \mathbb{R}^d$ $\rho^0(t-s,x) \leq C(t-s)^{-\frac{d}{\alpha}}$, we get that

$$\begin{aligned} I_2 &\leq C(t-s)^{-\frac{d}{\alpha}} \int_{|z|>(t-s)^{\frac{1}{\alpha}}} \frac{dz}{|z|^{d+\alpha}} \\ &\leq C(t-s)^{-1-\frac{d}{\alpha}} \\ &\leq C(t-s)^{-1}\rho^1(t-s,y-x), \end{aligned} \tag{8.249}$$

since $|y-x| \leq 2(t-s)^{\frac{1}{\alpha}}$. Finally, if $|y-x| > 2(t-s)^{\frac{1}{\alpha}}$, we write

$$\begin{aligned} I_2 &\leq C\int_{(t-s)^{\frac{1}{\alpha}}<|z|\leq\frac{|x-y|}{2}} \left[\rho^0(t-s,y-x-z)+\rho^0(t-s,y-x+z)+\rho^0(t-s,y-x)\right]\frac{dz}{|z|^{d+\alpha}} \\ &\quad + C\int_{|z|>\frac{|x-y|}{2}} \left[\rho^0(t-s,y-x-z)+\rho^0(t-s,y-x+z)+\rho^0(t-s,y-x)\right]\frac{dz}{|z|^{d+\alpha}} \\ &=: I_{2,1}+I_{2,2}. \end{aligned}$$

Then, the reverse triangle inequality yields

$$\begin{aligned} I_{2,1} &\leq C\int_{(t-s)^{\frac{1}{\alpha}}<|z|\leq\frac{|x-y|}{2}} \left[\rho^0(t-s,y-x)+\rho^0(t-s,\frac{y-x}{2})\right]\frac{dz}{|z|^{d+\alpha}} \\ &\leq C\rho^0(t-s,y-x)\int_{(t-s)^{\frac{1}{\alpha}}<|z|} \frac{dz}{|z|^{d+\alpha}} \\ &\leq C(t-s)^{-1}\rho^0(t-s,y-x). \end{aligned} \tag{8.250}$$

For $I_{2,2}$, we have

$$\begin{aligned} I_{2,2} &\leq C|y-x|^{-d-\alpha}\int_{\mathbb{R}^d} \rho^0(t-s,y-x-z)+\rho^0(t-s,y-x+z)\,dz \\ &\quad + C\rho^0(t-s,y-x)\int_{|z|>\frac{|x-y|}{2}} \frac{dz}{|z|^{d+\alpha}} \\ &\leq C(t-s)^{\frac{-d-\alpha}{\alpha}}((t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha} + C\rho^0(t-s,y-x)|x-y|^{-\alpha} \\ &\leq C(t-s)^{\frac{-d-\alpha}{\alpha}}((t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha} + C(t-s)^{-1}\rho^0(t-s,y-x). \end{aligned}$$

Since $1 < \frac{1}{2}(t-s)^{-\frac{1}{\alpha}}|y-x|$, we deduce that for some constant $C>0$, one has

$$((t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha} \leq C(1+(t-s)^{-\frac{1}{\alpha}}|y-x|)^{-d-\alpha}.$$

It follows that

$$I_{2,2} \leq C(t-s)^{-1}\rho^0(t-s,y-x). \tag{8.251}$$

Combining (8.248), (8.249), (8.250) and (8.251), we have proved that (8.229) holds true.

Chapter 9

Itô-Krylov's formula for a flow of measures

This chapter corresponds to the article [Cav21]. It is in revision in ESAIM: Probability and Statistics.

Abstract. We prove Itô's formula for the flow of measures associated with an Itô process having a bounded drift and a uniformly elliptic and bounded diffusion matrix, and for functions in an appropriate Sobolev-type space. This formula is the almost analogue, in the measure-dependent case, of the Itô-Krylov formula for functions in a Sobolev space on $\mathbb{R}^+ \times \mathbb{R}^d$.

9.1 Introduction

We fix $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ a filtered probability space satisfying the usual conditions. Let $T > 0$ be a finite horizon of time, $d, d_1 \in \mathbb{N}^*$ with $d_1 \geq d$, and $(B_t)_{t\geq 0}$ a $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion of dimension d_1. We consider the Itô process on $\mathbb{R}^d$ defined, for $t \in [0, T]$, by

$$X_t := X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dB_s, \tag{9.1}$$

where $X_0 \in L^2(\Omega, \mathcal{F}_0; \mathbb{R}^d)$, $b : [0, T] \times \Omega \to \mathbb{R}^d$ and $\sigma : [0, T] \times \Omega \to \mathbb{R}^{d\times d_1}$ are progressively measurable processes. In the following, we will denote by μ_t the law of X_t and by a the matrix $\sigma\sigma^*$.

Let us fix a real-valued function u defined on the 2-Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$, i.e. the space of probability measures on $\mathbb{R}^d$ having a finite moment of order 2. In this paper, we are interested in Itô's formula for u and the flow of probability measures $(\mu_t)_{t\in[0,T]}$. This formula describes the dynamics of $t \mapsto u(\mu_t)$, essentially by computing its derivative (see (9.2) below). It has a wide range of applications for example in Mean-Field Games, McKean-Vlasov's control problems, McKean-Vlasov Stochastic Differential Equations (SDEs) but also in the study of interacting particle systems and the propagation of chaos. These applications will be detailed below.

Itô's formula for a flow of measures naturally requires differential calculus on the space of measures $\mathcal{P}_2(\mathbb{R}^d)$. We will use the linear (functional) derivative, which is a standard notion of differentiability for functions of measures relying on the convexity of $\mathcal{P}_2(\mathbb{R}^d)$. The function u admits a linear derivative if there exists a real-valued and continuous function $\frac{\delta u}{\delta m}$ defined on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, at most of quadratic growth with respect to the space variable uniformly on each compact set of $\mathcal{P}_2(\mathbb{R}^d)$, and such that for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$

$$u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta u}{\delta m}(t\mu + (1-t)\nu)(v)\, d(\mu - \nu)(v)\, dt.$$

The standard Itô formula for a flow of measures can be found in [BLPR17] (see Theorem 6.1) or in Section 3 of [CCD15] and Chapter 5 of [CD18a] (see Theorem 5.99) under less restrictive assumptions. It states that for all $t \in [0, T]$

$$u(\mu_t) = u(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot b_s\right) ds + \frac{1}{2}\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot a_s\right) ds, \tag{9.2}$$

where $x \cdot y$ denotes the usual scalar product of two vectors $x, y \in \mathbb{R}^d$ and $A \cdot B := \mathrm{Tr}(A^* B)$ the usual scalar product of two matrices $A, B \in \mathbb{R}^{d \times d}$. The common point between these results is that the function u has to be $\mathcal{C}^2$ in some sense. More precisely, it is always assumed that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the linear derivative $\frac{\delta u}{\delta m}(\mu)(\cdot)$ belongs to $\mathcal{C}^2(\mathbb{R}^d)$ or equivalently that the L-derivative $\partial_\mu u(\mu)(\cdot)$ belongs to $\mathcal{C}^1(\mathbb{R}^d)$ (see below for the definition of the L-derivative and its link with the linear derivative). This paper aims at proving Itô's formula (9.2) for functions u having a linear derivative $\frac{\delta u}{\delta m}$ that is not $\mathcal{C}^2$ with respect to the space variable.

We now fix the assumptions on the Itô process $(X_t)_{t \in [0,T]}$. In this paper, we always assume that the drift b and the diffusion matrix σ in (9.1) satisfy the following properties.

(A) There exists $K > 0$ such that almost surely

$$\forall t \in [0, T],\ |b_t| + |\sigma_t| \leq K.$$

(B) There exists $\delta > 0$ such that almost surely

$$\forall t \in [0, T],\ \forall \lambda \in \mathbb{R}^d,\ a_t \lambda \cdot \lambda \geq \delta |\lambda|^2.$$

Assumptions **(A)** and **(B)** stem from Section 2.10 of [Kry09]. Therein, Krylov deals with controlled diffusion processes and needs to apply the standard Itô formula for the so-called pay-off function which is not $\mathcal{C}^2$. That is why he proves an extension of the classical Itô formula for the Itô process $(X_t)_{t \in [0,T]}$ satisfying Assumptions **(A)** and **(B)**, and for a function $g : \mathbb{R}^d \to \mathbb{R}$ belonging to an appropriate Sobolev space. The crucial point is that $(X_t)_t$ satisfies the non-degeneracy Assumption **(B)**. It ensures that the noise does not degenerate and allows to produce a regularizing effect. Let us explain how. The non-degeneracy assumption leads to Krylov's inequality (see Theorem 9.20 taken from Section 2.3 of [Kry09]). This inequality, in turn, implies that for almost all $t \in [0, T]$, μ_t, the law of X_t, has a density $p(t, \cdot)$ with respect to the Lebesgue measure (see Proposition 9.22). Moreover, this density belongs to $L^{(d+1)'}([0, T] \times \mathbb{R}^d)$, where $(d+1)'$ denotes the conjugate exponent of $d+1$ defined in Section 9.2. The existence of densities together with the integrability property permit to assume Sobolev regularity for the function g. More precisely, Itô-Krylov's formula is established under the assumption that g is continuous on $\mathbb{R}^d$ and that ∇g belongs to the Sobolev space $W^{1,k}_{\mathrm{loc}}(\mathbb{R}^d)$, for $k \geq d+1$, i.e. that ∇g and $\nabla^2 g$ are in $L^k_{\mathrm{loc}}(\mathbb{R}^d)$ (see Section 2.10 of [Kry09]).

Our goal here is to take advantage of the regularizing effect of the noise, stemming from the existence of the densities $p(t, \cdot)$ and their integrability property, to establish an analogue of Itô-Krylov's formula in the measure-dependent case. Looking at Itô's formula for a flow of measures (9.2), the regularizing effect comes from the presence of expectations which average, with respect to the space variable, the derivatives of $\frac{\delta u}{\delta m}$ on all the trajectories of $(X_t)_t$. Indeed, the regularization by noise will only appear through the space variable of the linear derivative but not through its measure variable. This is not surprising since the space of measures $\mathcal{P}_2(\mathbb{R}^d)$ is somehow infinite dimensional while the noise is of finite dimension. Thus, we cannot expect a true regularization in the measure variable of $\frac{\delta u}{\delta m}$. The fact that a finite dimensional noise cannot have a complete regularizing effect in the space $\mathcal{P}_2(\mathbb{R}^d)$ is explained in [Mar20] in the context of McKean-Vlasov SDEs.

In order to prove Itô's formula (9.2) for u, it is clear that u needs to admit a linear derivative with at least distributional derivatives of order 1 and 2 with respect to the space variable in $L^k(\mathbb{R}^d)$ for some k, as for the standard Itô-Krylov formula. Let us describe more precisely our assumptions on u. As said before, for almost all $t \in [0,T]$, the law μ_t has a density $p(t,\cdot)$ such that p belongs to $L^{(d+1)'}([0,T]\times\mathbb{R}^d)$. Denoting by $\mathscr{P}(\mathbb{R}^d)$ the space of measures $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ having a density with respect to the Lebesgue measure in $L^{(d+1)'}(\mathbb{R}^d)$, our assumptions on the derivatives of $\frac{\delta u}{\delta m}(\mu)(\cdot)$ are only made for measures μ belonging to $\mathscr{P}(\mathbb{R}^d)$. This is natural since for almost all $t \in [0,T]$, μ_t belongs to $\mathscr{P}(\mathbb{R}^d)$, and the derivatives of $\frac{\delta u}{\delta m}$ are evaluated along the flow $(\mu_t)_{t\in[0,T]}$ and integrated in time. Moreover, because of the integrability property of the densities $p(t,\cdot)$, the derivatives of $\frac{\delta u}{\delta m}(\mu)(\cdot)$ do not need to be defined and continuous on the whole space $\mathbb{R}^d$ because they are somehow integrated against the densities $p(t,\cdot)$ (see (9.2)). We say "somehow" because it is not completely the case since b and a are random. But as they are bounded, we can omit them in some sense. More precisely, the integrability property of the densities leads us to assume that u admits a linear derivative such that for all $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\partial_v \frac{\delta u}{\delta m}(\mu)(\cdot)$ belongs to the Sobolev space $W^{1,k}(\mathbb{R}^d)$ defined in Section 9.2, with $k \geq d+1$. This is exactly the same condition as in the standard Itô-Krylov formula, except that we replace $W^{1,k}_{\text{loc}}(\mathbb{R}^d)$ by $W^{1,k}(\mathbb{R}^d)$. This is essentially explained by the expectations in Itô's formula (9.2). Indeed, the process $(X_t)_t$ cannot be localized by stopping times. Moreover, we assume that the map $\mu \in \mathscr{P}(\mathbb{R}^d) \mapsto \partial_v \frac{\delta u}{\delta m}(\mu)(\cdot) \in W^{1,k}(\mathbb{R}^d)$ is continuous for a distance on $\mathscr{P}(\mathbb{R}^d)$ satisfying the assumptions of Definition 9.3. This continuity assumption can be interpreted as the fact that the noise has no regularizing effect in the measure variable of the linear derivative, as explained above. The precise assumptions of our Itô-Krylov's formula are given in Definition 9.5 and Theorem 9.7. Eventually, we extend in Theorem 9.16 our formula to functions depending also on the time and space variables satisfying the assumptions of Definition 9.14.

We now focus on some applications of Itô's formula for a flow of measures. This one has been developed with the increasing interest for Mean-Field Games and McKean-Vlasov SDEs over the last decade. Mean-Field Games were initiated independently by Caines, Huang and Malhame in [CHM06] and by Lasry and Lions in [LL07]. The notion of Master equations has been introduced by Lions in his lectures at Collège de France [Lio] in order to describe Mean-Field Games. Master equations are Partial Differential Equations (PDEs) on the space of probability measures and can be derived with the help of Itô's formula. We refer to Lions' lectures [Lio], the notes written by Cardialaguet [Car10], and the books of Carmona and Delarue [CD18a, CD18b] for more details on Mean-Field Games and Master equations. We also mention Bensoussan, Frehse and Yam [BFY15] and Carmona, Delarue [CD14] where Master equations are derived, with the help of Itô's formula in [CD14]. The question of existence and uniqueness of classical solutions to Master equations was addressed by Cardaliaguet, Delarue, Lasry and Lions in [CDLL19] and by Chassagneux, Crisan and Delarue in [CCD15]. From a different point of view, Mou and Zhang deal with the well-posedness of Master equations in some weaker senses in [MZ20].

Moreover, Itô's formula appears to be the natural way to connect a McKean-Vlasov SDE (more precisely the associated semigroup $(P_t)_t$ acting on the space of functions of measures) to a PDE on the space of probability measures (the Master equation) in the same manner as for classical SDEs. It turns out to be a crucial tool to study the stochastic flow generated by a McKean-Vlasov SDE, as explained in Chapter 5 of [CD18a]. The link between McKean-Vlasov SDEs and PDEs on the space of measures is at the heart of the work of Buckdahn, Li, Peng and Rainer [BLPR17] where the authors prove that the PDE admits a unique classical solution expressed with the flow of measures associated with the McKean-Vlasov SDE. Moreover, in the parallel work [CCD15], Chassagneux, Crisan and Delarue adopt a similar approach and study the flow generated by a forward-backward stochastic system of McKean-Vlasov type under weaker assumptions on the coefficients of the equation. Both works are motivated by Mean-Field Games, and Itô's formula plays a key role. In [CM17], Crisan and McMurray prove that the Master equation admits a unique classical solution for some irregular terminal condition using Malliavin calculus. They point out a smoothing effect concerning the differentiability of the solution with respect to the measure even though there is no noise in the measure direction. Furthermore, the problem of propagation of chaos for the interacting particles system associated with the McKean-Vlasov SDE can also be addressed with the help of the associated PDE on the space of measures (see Chapter 5 of [CD18a]). It allows to obtain quantitative weak propagation of chaos estimates between the law of the solution to the McKean-Vlasov SDE and the empirical measure of the associated particle system. This approach was adopted for example by Chaudru de Raynal and Frikha in [CdRF22, CdRF21], by Delarue and Tse in [DT21] and by Chassagneux, Szpruch and Tse in [CST22]. Let us also mention that the Master equation satisfied by the semigroup has been recently used by Jourdain and Tse in [JT21] to study the mean-field fluctuation (CLT) of an interacting particle system. Finally, Itô's formula for a flow of measures is also important to deal with McKean-Vlasov control problems because it allows to derive a dynamic programming principle describing the value function of the problem as presented in Chapter 6 of [CD18a].

Recently, Itô's formula has been extended to flows of measures generated by càdlàg semi-martingales. It was achieved independently by Guo, Pham and Wei in [GPW20], who studied McKean-Vlasov control problems with jumps and by Talbi, Touzi and Zhang in [TTZ21] who worked on mean-field optimal stopping problems. In both works, dynamic programming principles are established thanks to Itô's formula for a flow of measures. Finally, we also mention that several Itô-Wentzell-Lions formulae for functional random fields of Itô type depending on measure flows have been established by dos Reis and Platonov in [dRP22].

Let us explain our choice to work with the linear derivative. Indeed, the L-derivative, which was introduced by Lions in his lectures at Collège de France [Lio], is also well-adapted to establish Itô's formula for a flow of measures. We say that u is L-differentiable if its lifting defined by

$$\tilde{u} : X \in L^2(\Omega; \mathbb{R}^d) \mapsto u(\mathcal{L}(X)) \in \mathbb{R},$$

where $\mathcal{L}(X)$ denotes the law of X, is Fréchet differentiable on $L^2(\Omega; \mathbb{R}^d)$. Moreover, there exists a $\mathbb{R}^d$-valued function $\partial_\mu u$ defined on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ such that the gradient of $\tilde{u}$ at $X \in L^2(\Omega; \mathbb{R}^d)$ is given by the random variable $\partial_\mu u(\mathcal{L}(X))(X)$. The function $\partial_\mu u$ is called the L-derivative of u. The advantage of the L-derivative is that it permits to use standard tools of differential calculus on Banach spaces. Of course, there is a link between the L-derivative and the linear derivative of u. Indeed, in general, the L-derivative $\partial_\mu u(\mu)(\cdot)$ is equal to the gradient of the linear derivative $\partial_v \frac{\delta u}{\delta m}(\mu)(\cdot)$ (see Propositions 5.48 and 5.51 in [CD18a] for the precise assumptions). Under our assumptions presented above, Sobolev embedding theorem ensures that for all $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\frac{\delta u}{\delta m}(\mu)(\cdot)$ belongs to $\mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$, and that $\partial_v \frac{\delta u}{\delta m}(\mu)(\cdot)$ is continuous and bounded on $\mathbb{R}^d$. We would be tempted to deduce that u admits a L-derivative given, as recalled above, by $\partial_v \frac{\delta u}{\delta m}(\mu)(\cdot)$. However, this term is assumed to exist only for measures $\mu \in \mathscr{P}(\mathbb{R}^d)$ and not for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. This is the case in Example 9.10, where this term is not well-defined for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ (see Remark 9.11). It seems therefore more restrictive to work with the L-derivative and thus justifies our choice to work with the linear derivative.

The paper is organized as follows. Section 9.2 gathers some notations and definitions used throughout the paper. In Section 9.3, more precisely in Definitions 9.5 and 9.14, we define the spaces of functions for which we will establish Itô-Krylov's formula. These formulas are given in Theorem 9.7 for functions defined on $\mathcal{P}_2(\mathbb{R}^d)$ and in Theorem 9.16 for functions depending also on the time and space variables. Moreover, we give examples of functions for which our formulas hold and we discuss our assumptions through them. The proofs of these examples are postponed to Section 9.7 for ease of reading. In Section 9.4, we give some preliminary results. We start with Krylov's inequality and its consequences on the existence of densities for the flow of measures $(\mu_t)_{t\in[0,T]}$ in Proposition 9.22. Then we recall some classical results on convolution and regularization. Finally, Sections 9.5 and 9.6 are respectively dedicated to the proofs Theorems 9.7 and 9.16.

9.2 Notations and definitions

9.2.1 General notations

Let us introduce some notations used several times in the article.

- B_R is the open ball centered at 0 and of radius R in $\mathbb{R}^d$ for the euclidean norm.
- p' is the conjugate exponent of $p \in [1,+\infty]$, defined by $\frac{1}{p}+\frac{1}{p'}=1$.
- $L^p_{\text{loc}}(\mathbb{R}^d)$ is the space of functions f such that for all $R>0$, $f \in L^p(B_R)$.
- $W^{m,k}(\mathcal{O})$ is the Sobolev space of functions $u \in L^k(\mathcal{O})$ admitting distributional derivatives of order between 1 and m in $L^k(\mathcal{O})$, where $\mathcal{O}$ is open in $\mathbb{R}^d$. It is equipped with the norm

$$\|u\|_{W^{m,k}(\mathcal{O})} = \sum_{\alpha\in\mathbb{N}^d,\,|\alpha|\le m} \|\partial^\alpha u\|_{L^k(\mathcal{O})}.$$

- $W^{m,k}_{\text{loc}}(\mathbb{R}^d)$ is the space of functions u such that for all $R>0$, u belongs to $W^{m,k}(B_R)$.
- $(\rho_n)_n$ is a mollifying sequence on $\mathbb{R}^d$, that is a sequence of non-negative $\mathcal{C}^\infty$ functions, such that for all n, $\int_{\mathbb{R}^d}\rho_n(x)\,dx=1$ and ρ_n is equal to 0 outside $B_{1/n}$. We assume that $\rho_n(x)=\rho_n(-x)$ for all x.
- $*$ denotes the convolution of two functions, when it is well-defined, or two probability measures.
- $\mathcal{B}(E)$ is the Borel σ-algebra where E is a metric space.

- A^* denotes the transpose of the matrix $A \in \mathbb{R}^{d\times d}$.
- $A \cdot B$ denotes the usual scalar product of two matrices $A, B \in \mathbb{R}^{d\times d}$ given by $A \cdot B := \mathrm{Tr}(A^* B)$.
- $\mathscr{P}(\mathbb{R}^d)$ is defined in Definition 9.3.
- $\mathcal{W}_1(\mathbb{R}^d)$ is defined in Definition 9.5.
- $\mathcal{W}_2(\mathbb{R}^d)$ is defined in Definition 9.14.

9.2.2 Spaces of measures and linear derivative

The set $\mathcal{P}(\mathbb{R}^d)$ is the space of probability measures on $\mathbb{R}^d$ equipped with the topology of weak convergence. The Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$ denotes the set of measures $\mu \in \mathcal{P}(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} |x|^2 \, d\mu(x) < +\infty$, equipped with the 2-Wasserstein distance W_2 defined for $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ by

$$W_2(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \, d\pi(x,y) \right)^{1/2},$$

where $\Pi(\mu,\nu)$ is the subset of $\mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ with marginal distributions μ and ν. We will work with the standard notion of linear derivative for functions of measures.

Definition 9.1 (Linear derivative)**.** A function $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to have a linear derivative if there exists a continuous function $(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \frac{\delta u}{\delta m}(\mu)(v) \in \mathbb{R}$, satisfying the following properties.

1. For all compact $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ $\displaystyle\sup_{v\in\mathbb{R}^d} \sup_{\mu\in\mathcal{K}} \left\{ (1+|v|^2)^{-1} \left| \frac{\delta u}{\delta m}(\mu)(v) \right| \right\} < +\infty$.
2. For all $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\displaystyle u(\mu) - u(\nu) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta u}{\delta m}(t\mu + (1-t)\nu)(v) \, d(\mu-\nu)(v) \, dt$.

Remark 9.2. Instead of the second point of the previous definition, it is equivalent to assume that for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$, $t \in [0,1] \mapsto u(t\mu + (1-t)\nu)$ is of class $\mathcal{C}^1$ with

$$\forall t \in [0,1], \; \frac{d}{dt} u(t\mu + (1-t)\nu) = \int_{\mathbb{R}^d} \frac{\delta u}{\delta m}(t\mu + (1-t)\nu)(v) \, d(\mu-\nu)(v).$$

One can find more details in Chapter 5 of [CD18a], in particular the connection with the L-derivative.

Let us fix $(\rho_n)_n$ a mollifying sequence on $\mathbb{R}^d$, that is a sequence of non-negative $\mathcal{C}^\infty$ functions, such that for all n, $\int_{\mathbb{R}^d} \rho_n(x) \, dx = 1$ and ρ_n is equal to 0 outside $B_{1/n}$. We assume that $\rho_n(x) = \rho_n(-x)$ for all x.

Definition 9.3. Let us define $\mathscr{P}(\mathbb{R}^d)$ as the space of measures $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ which admit a density $\frac{d\mu}{dx}$ with respect to the Lebesgue measure belonging to $L^{(d+1)'}(\mathbb{R}^d)$. We endow $\mathscr{P}(\mathbb{R}^d)$ with a general distance $d_{\mathscr{P}}$ satisfying the following properties.

(H1) For any $n \geq 1$, $\mu \in (\mathcal{P}_2(\mathbb{R}^d), W_2) \mapsto \mu * \rho_n \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$ is continuous.

(H2) For any $\mu \in \mathscr{P}(\mathbb{R}^d)$, $\mu * \rho_n \underset{n\to+\infty}{\longrightarrow} \mu$ for $d_{\mathscr{P}}$.

Note that for all $n \geq 1$ and for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu * \rho_n \in \mathscr{P}(\mathbb{R}^d)$. Indeed, its density is given by $x \mapsto \rho_n * \mu(x) = \int_{\mathbb{R}^d} \rho_n(x-y) \, d\mu(y)$. Jensen's inequality ensures that it belongs to $L^{(d+1)'}(\mathbb{R}^d)$. Considering the space $(\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$ comes in a natural way with Assumptions **(A)** and **(B)** on the Itô process X. As explained in the introduction, it implies the existence of a density $p \in L^1([0,T] \times \mathbb{R}^d; \mathbb{R}^+) \cap$

$L^{(d+1)'}([0,T]\times\mathbb{R}^d;\mathbb{R}^+)$ such that for almost all $t\in[0,T]$, the law of X_t is equal to $p(t,\cdot)\,dx$ and belongs to $\mathscr{P}(\mathbb{R}^d)$ (see Proposition 9.22). Let us give two examples for the distance $d_{\mathscr{P}}$.

Example 9.4. The Wasserstein distance W_2 clearly satisfies Assumptions **(H1)** and **(H2)** in Definition 9.3. Another family of examples is given by the distance d_k defined, for $k\in[d+1,+\infty[$, $\mu,\nu\in\mathscr{P}(\mathbb{R}^d)$, by

$$d_k(\mu,\nu)=\left\|\frac{d\mu}{dx}-\frac{d\nu}{dx}\right\|_{L^{k'}(\mathbb{R}^d)}.$$

Note that d_k is well-defined since for any $\mu\in\mathscr{P}(\mathbb{R}^d)$, $\frac{d\mu}{dx}\in L^1(\mathbb{R}^d)\cap L^{(d+1)'}(\mathbb{R}^d)$ which is included in $L^{k'}(\mathbb{R}^d)$ by interpolation. The proof is postponed Section 9.7.1.

9.3 Itô-Krylov's formula, ah-hoc spaces of functions and examples

Let us introduce now the Sobolev-type space of functions on $\mathcal{P}_2(\mathbb{R}^d)$ for which we will prove Itô's formula for a flow of measures.

Definition 9.5. Let $\mathcal{W}_1(\mathbb{R}^d)$ be the space of continuous functions $u:\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ having a linear derivative $\frac{\delta u}{\delta m}$ such that for all $\mu\in\mathscr{P}(\mathbb{R}^d)$, the function $\frac{\delta u}{\delta m}(\mu)(\cdot)$ admits distributional derivatives of order 1 and 2 in $L^k(\mathbb{R}^d)$, for a certain $k\geq d+1$, and satisfies the following properties.

1. The map $\mu\in(\mathscr{P}(\mathbb{R}^d),d_{\mathscr{P}})\mapsto\partial_v\frac{\delta u}{\delta m}(\mu)(\cdot)\in\left(W^{1,k}(\mathbb{R}^d)\right)^d$ is continuous for a certain distance $d_{\mathscr{P}}$ satisfying **(H1)** and **(H2)**.
2. There exists $\alpha\in\mathbb{N}$ such that $k\geq(1+\alpha)d$ and for all compact $\mathcal{K}\subset\mathcal{P}_2(\mathbb{R}^d)$ and for any $\mu\in\mathcal{K}\cap\mathscr{P}(\mathbb{R}^d)$

$$\left\|\partial_v\frac{\delta u}{\delta m}(\mu)(\cdot)\right\|_{L^k(\mathbb{R}^d)}+\left\|\partial_v^2\frac{\delta u}{\delta m}(\mu)(\cdot)\right\|_{L^k(\mathbb{R}^d)}\leq C_{\mathcal{K}}\left(1+\left\|\frac{d\mu}{dx}\right\|^{\alpha}_{L^{k'}(\mathbb{R}^d)}\right).$$

Remark 9.6. -The space $\mathcal{W}_1(\mathbb{R}^d)$ contains the functions which satisfy Assumption (1) in Definition 9.5 with $(\mathcal{P}_2(\mathbb{R}^d),W_2)$ instead of $(\mathscr{P}(\mathbb{R}^d),d_{\mathscr{P}})$. Indeed, the second point is clearly satisfied with $\alpha=0$ since $\mathcal{K}$ is compact.

-Assumption (2) in Definition 9.5 allows to control the growth of $\left\|\partial_v\frac{\delta u}{\delta m}(\mu)(\cdot)\right\|_{W^{1,k}(\mathbb{R}^d)}$ with respect to the measure μ. It allows us to take advantage of the continuity of the flow in $\mathcal{P}_2(\mathbb{R}^d)$ (because the control is assumed on compact subsets of $\mathcal{P}_2(\mathbb{R}^d)$), but also of its integrability properties proved in Lemmas 9.24 and 9.25. The form of the inequality suggests the integration of functions in $L^k(\mathbb{R}^d)$ with respect to μ, at least when the function u is linear in μ.

-Sobolev embedding theorem (see Corollary 9.14 in [Bre10]) ensures that for all $\mu\in\mathscr{P}(\mathbb{R}^d)$, $\frac{\delta u}{\delta m}(\mu)(\cdot)$ belongs to $\mathcal{C}^1(\mathbb{R}^d;\mathbb{R})$ and that $\partial_v\frac{\delta u}{\delta m}(\mu)(\cdot)$ is bounded and γ-Hölder, where $\gamma:=1-\frac{d}{k}$. Note that we do not need that $\frac{\delta u}{\delta m}(\mu)(\cdot)\in W^{2,k}(\mathbb{R}^d)$ since there is no integrability assumption made on the linear derivative.

Having this definition at hand, we can now state Itô-Krylov's formula for functions in $\mathcal{W}_1(\mathbb{R}^d)$.

Theorem 9.7 (Itô-Krylov's formula)**.** *Let u be a function in $\mathcal{W}_1(\mathbb{R}^d)$, which was defined in Definition 9.5. We have for all $t \in [0,T]$*

$$u(\mu_t) = u(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot b_s\right) ds + \frac{1}{2}\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot a_s\right) ds, \tag{9.3}$$

where $\partial_v^2 \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot a_s := Tr\Big(\partial_v^2 \frac{\delta u}{\delta m}(\mu_s)(X_s) a_s\Big)$ is the usual scalar product on $\mathbb{R}^{d\times d}$.

Remark 9.8. Notice that a function $u \in \mathcal{W}_1(\mathbb{R}^d)$ is assumed to have a linear derivative on the whole space $\mathcal{P}_2(\mathbb{R}^d)$. This seems a bit strong at first sight in comparison with the assumptions on its spatial derivatives that are only made for measures $\mu \in \mathscr{P}(\mathbb{R}^d)$. Indeed, we could consider working with a linear derivative defined only on the space of densities, as done for example in [BFY15]. However, in order to establish Itô-Krylov's formula by regularization, the function u needs to be continuous on the whole space $\mathcal{P}_2(\mathbb{R}^d)$ and not only on $\mathscr{P}(\mathbb{R}^d)$. Indeed, the flow $s \in [0,T] \mapsto \mu_s \in \mathcal{P}_2(\mathbb{R}^d)$ is continuous but μ_t does not necessarily belong to $\mathscr{P}(\mathbb{R}^d)$ for all $t \in [0,T]$. This is proved only for almost all t. Thus, as the function u has to be continuous on $\mathcal{P}_2(\mathbb{R}^d)$, we have chosen to assume the existence of a linear derivative on $\mathcal{P}_2(\mathbb{R}^d)$ even though we could have only required it on the space of densities.

Now, we focus on examples of functions belonging to $\mathcal{W}_1(\mathbb{R}^d)$. Let us start with the linear case.

Example 9.9 (Linear functional)*.* Fix $g \in \mathcal{C}^0(\mathbb{R}^d;\mathbb{R})$ admitting a distributional derivative such that $\nabla g \in (W^{1,k}(\mathbb{R}^d))^d$ for some $k \geq d+1$. Then, the function

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \qquad \mu & \mapsto \displaystyle\int_{\mathbb{R}^d} g(x)\, d\mu(x), \end{cases}$$

belongs to the space $\mathcal{W}_1(\mathbb{R}^d)$.

Indeed, Sobolev embedding theorem (see Corollary 9.14 in [Bre10]) implies that $\nabla g \in L^\infty(\mathbb{R}^d)$ since $k \geq d+1$. Thus g is at most of linear growth so that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\frac{\delta u}{\delta m}(\mu) = g$, which clearly satisfies Assumptions (1) and (2) (with $\alpha = 0$) in Definition 9.5.

Let us now focus on the multi-linear case.

Example 9.10 (Polynomials on the Wasserstein space)*.* Fix $N \geq 2$ and $g \in \mathcal{C}^0((\mathbb{R}^d)^N;\mathbb{R})$ such that

- there exists $C > 0$ such that for all $\boldsymbol{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$, $|g(\boldsymbol{x})| \leq C(1 + |x_1|^2 + \cdots + |x_N|^2)$,
- the distributional derivative ∇g belongs to $(W^{1,k}((\mathbb{R}^d)^N))^{Nd}$ for a certain $k \in [Nd, +\infty[$.

Then, the function

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \qquad \mu & \mapsto \displaystyle\int_{(\mathbb{R}^d)^N} g(x_1, \dots, x_N)\, d\mu(x_1) \dots d\mu(x_N), \end{cases}$$

belongs to the space $\mathcal{W}_1(\mathbb{R}^d)$ for $d_{\mathscr{P}} = d_k$.

The proof is postponed to Section 9.7.2.

Remark 9.11. - In Definition 9.5, the distributional derivatives of the linear derivative $\frac{\delta u}{\delta m}(\mu)$ are not necessarily integrable functions for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Of course, in Example 9.9, it is the case for all

$\mu \in \mathcal{P}_2(\mathbb{R}^d)$ as the linear derivative does not depend on the measure μ. However, in Example 9.10 for $N = 2$, the linear derivative is given by

$$\frac{\delta u}{\delta m}(\mu)(v) = \int_{\mathbb{R}^d} g(v, y)\, d\mu(y) + \int_{\mathbb{R}^d} g(y, v)\, d\mu(y). \tag{9.4}$$

Formally, the derivative with respect to v of the first integral in (9.4) is

$$\int_{\mathbb{R}^d} \partial_v g(v, y)\, d\mu(y).$$

This term is not well-defined for general measures $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ because we have only assumed that $\nabla g \in (W^{1,k}(\mathbb{R}^{2d}))^{2d}$ with $k \geq 2d$. Indeed, for $k = 2d$, we just know by Sobolev embedding theorem that ∇g belongs to $(L^r(\mathbb{R}^{2d})^{2d}$ with $r \in [2d, +\infty[$ (see Corollary 9.11 in [Bre10]). As we will see in the proof (Section 9.7.2), it is well-defined as an integrable function of v if we restrict to measures $\mu \in \mathscr{P}(\mathbb{R}^d)$. This also justifies why we have chosen to work with the linear derivative instead of the L-derivative. Indeed, the L-derivative of u would be equal to the gradient of the linear derivative $\partial_v \frac{\delta u}{\delta m}(\mu)(\cdot)$, which is not well-defined for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Thus, the function u does not need to be L-differentiable in the usual sense in our setting.

- Our assumptions on the derivatives of $\frac{\delta u}{\delta m}$ in Definition 9.5 deal with $\mathscr{P}(\mathbb{R}^d)$ instead of the whole space $\mathcal{P}_2(\mathbb{R}^d)$ essentially because in Itô's formula (9.3), these derivatives only appear under integrals along the flow $(\mu_s)_{s\in[0,T]}$, which belongs to $\mathscr{P}(\mathbb{R}^d)$ for almost all $s \in [0, T]$. However, we assume that u is continuous on $\mathcal{P}_2(\mathbb{R}^d)$ since the flow $s \in [0, T] \mapsto \mu_s \in \mathcal{P}_2(\mathbb{R}^d)$ is continuous but μ_t does not necessarily belong to $\mathscr{P}(\mathbb{R}^d)$ for all $t \in [0, T]$.

The next example focuses on the particular case of convolution which has to be treated differently than in Example 9.10 with $N = 2$ because of the structure of the convolution which mixes the two variables.

Example 9.12. Let $f \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{R})$ be a function such that the distributional derivative ∇f belongs to $(W^{1,k+1}(\mathbb{R}^d))^d$, for a certain $k \geq d$. Then, the function

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \mu & \mapsto \displaystyle\int_{\mathbb{R}^d} f * \mu\, d\mu, \end{cases}$$

belongs to $\mathcal{W}_1(\mathbb{R}^d)$ for $d_{\mathscr{P}} = W_2$.

Here, the particular structure of convolution enables us to work on the whole space $\mathcal{P}_2(\mathbb{R}^d)$ instead of $\mathscr{P}(\mathbb{R}^d)$, as explained in the first point of Remark 9.6. The proof is postponed to Section 9.7.3.

Finally, we give a non-linear example of functions belonging to $\mathcal{W}_1(\mathbb{R}^d)$.

Example 9.13. Let $F \in \mathcal{C}^1(\mathbb{R}; \mathbb{R})$ and $g \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{R})$ be such that the distributional derivative ∇g belongs to $(W^{1,k}(\mathbb{R}^d))^d$ for some $k \geq d + 1$. Then

$$u : \begin{cases} \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ \mu & \mapsto F\left(\int_{\mathbb{R}^d} g\, d\mu\right) \end{cases}$$

belongs to $\mathcal{W}_1(\mathbb{R}^d)$ for $d_{\mathscr{P}} = W_2$.

The proof is again postponed to Section 9.7.4.

We now deal with the extension of Itô's formula for functions depending also on the time and space variables. First, we define the space of functions generalizing the space $\mathcal{W}_1(\mathbb{R}^d)$.

Definition 9.14. Let $\mathcal{W}_2(\mathbb{R}^d)$ be the set of continuous functions $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ satisfying the following properties for a certain distance $d_{\mathscr{P}}$ satisfying **(H1)** and **(H2)**.

1. For all $(x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $u(\cdot,x,\mu) \in \mathcal{C}^1$ and $\partial_t u$ is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$.
2. There exists $k_1 \geq d+1$ such that for all $(t,\mu) \in [0,T] \times \mathscr{P}(\mathbb{R}^d)$, $u(t,\cdot,\mu) \in W^{2,k_1}_{\text{loc}}(\mathbb{R}^d)$ and for all $t \in [0,T]$ and $R > 0$
$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \partial_x u(t,\cdot,\mu) \in \left(W^{1,k_1}(B_R)\right)^d,$$
is continuous and $\partial_x u$ and $\partial_x^2 u$ are measurable with respect to $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d)$.
3. For all $(t,x) \in [0,T] \times \mathbb{R}^d$, $u(t,x,\cdot)$ admits a linear derivative $\frac{\delta u}{\delta m}(t,x,\cdot)(\cdot)$ which is continuous on $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, and such that for all $\mathcal{K} \subset \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ compact and $t \in [0,T]$, there exists $C > 0$ such that for all $v \in \mathbb{R}^d$
$$\sup_{(x,\mu)\in\mathcal{K}} \left| \frac{\delta u}{\delta m}(t,x,\mu)(v) \right| \, dx \leq C(1+|v|^2).$$
4. There exists $k_2 \geq 2d$ such that for all $(t,\mu) \in [0,T] \times \mathscr{P}(\mathbb{R}^d)$, $\frac{\delta u}{\delta m}(t,\cdot,\mu)(\cdot)$ admits distributional derivatives with respect to v of order 1 and 2 such that for all t and $R > 0$
$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}}) \mapsto \left(\partial_v \frac{\delta u}{\delta m}(t,\cdot,\mu)(\cdot)\,,\, \partial_v^2 \frac{\delta u}{\delta m}(t,\cdot,\mu)(\cdot)\right) \in (L^{k_2}(B_R \times \mathbb{R}^d))^d \times (L^{k_2}(B_R \times \mathbb{R}^d))^{d\times d},$$
is continuous and measurable with respect to $(t,x,\mu,v) \in [0,T] \times \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d$.
5. There exists $\alpha_1, \alpha_2 \in \mathbb{N}$ with $k_1 \geq (2\alpha_1+1)d$, $k_2 \geq (\alpha_2+2)d$ such that for all $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ compact and $R > 0$, there exists $C_{\mathcal{K},R} > 0$ such that for all $\mu \in \mathcal{K} \cap \mathscr{P}(\mathbb{R}^d)$
$$\begin{cases} \sup_{t\leq T} \left\{ \|\partial_x u(t,\cdot,\mu)\|_{L^{k_1}(B_R)} + \left\|\partial_x^2 u(t,\cdot,\mu)\right\|_{L^{k_1}(B_R)} \right\} \leq C_{\mathcal{K},R} \left(1 + \left\|\frac{d\mu}{dx}\right\|^{\alpha_1}_{L^{k_1'}(\mathbb{R}^d)}\right) \\ \sup_{t\leq T} \left\{ \left\|\partial_v \frac{\delta u}{\delta m}(t,\cdot,\mu)(\cdot)\right\|_{L^{k_2}(B_R\times\mathbb{R}^d)} + \left\|\partial_v^2 \frac{\delta u}{\delta m}(t,\cdot,\mu)(\cdot)\right\|_{L^{k_2}(B_R\times\mathbb{R}^d)} \right\} \leq C_{\mathcal{K},R} \left(1 + \left\|\frac{d\mu}{dx}\right\|^{\alpha_2}_{L^{k_2'}(\mathbb{R}^d)}\right). \end{cases}$$

Remark 9.15. - The space $\mathcal{W}_2(\mathbb{R}^d)$ contains the functions satisfying the four first assumptions of Definition 9.14 with $(\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$ replaced by $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ and also assuming that the functions in Assumptions (2) and (4) are continuous with respect to $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$. Indeed, Assumption (5) is automatically satisfied with $\alpha_1 = \alpha_2 = 0$ because $\mathcal{K}$ is compact.

- The bound in Assumption (3) is quite natural. If the supremum in this bound was taken only over a compact set of $\mathcal{P}_2(\mathbb{R}^d)$, it would be the definition of the linear derivative. But we also need to control $\frac{\delta u}{\delta m}$ locally uniformly in the space variable $x \in \mathbb{R}^d$ because of our regularization procedure through a convolution both in the space and measure variables. Assumptions (2), (4) and (5) are generalizations of those in Definition 9.5 adapted to the presence of the space and time variables. In Assumption (5), the condition on k_2 and α_2 changes a bit compared to the analogous assumption in Definition 9.5, essentially because it deals with functions on $\mathbb{R}^{2d}$ instead of $\mathbb{R}^d$. Let us mention that Assumption (5) in Definition 9.14 can be replaced by the integrability properties (9.12) established in Step 1 of the proof of the next theorem (see Section 9.6).

The next theorem is the natural extension of the formula for functions in $\mathcal{W}_2(\mathbb{R}^d)$. We still consider the flow of marginal distributions $(\mu_t)_{t\in[0,T]}$ of the process X defined by (9.1). Let $(\eta_s)_{s\in[0,T]}$ and $(\gamma_s)_{s\in[0,T]}$ be two progressively measurable processes, taking values respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d_1}$ and satisfying Assumptions **(A)** and **(B)**. We set, for all $t \le T$

$$\xi_t = \xi_0 + \int_0^t \eta_s \, ds + \int_0^t \gamma_s \, dB_s,$$

where ξ_0 is a $\mathcal{F}_0$-measurable random variable with values in $\mathbb{R}^d$.

Theorem 9.16 (Extension of Itô-Krylov's formula)**.** *Let u be a function in $\mathcal{W}_2(\mathbb{R}^d)$, which was defined in Definition 9.14. We have almost surely, for all $t \in [0,T]$*

$$\begin{aligned} u(t,\xi_t,\mu_t) = u(0,\xi_0,\mu_0) &+ \int_0^t (\partial_t u(s,\xi_s,\mu_s) + \partial_x u(s,\xi_s,\mu_s)\cdot\eta_s)\, ds + \frac{1}{2}\int_0^t \partial_x^2 u(s,\xi_s,\mu_s)\cdot \gamma_s\gamma_s^*\, ds \\ &+ \int_0^t \tilde{\mathbb{E}}\left(\partial_v \frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot \tilde{b_s}\right) ds + \frac{1}{2}\int_0^t \tilde{\mathbb{E}}\left(\partial_v^2 \frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot \tilde{a_s}\right) ds \\ &+ \int_0^t \partial_x u(s,\xi_s,\mu_s)\cdot(\gamma_s\, dB_s), \end{aligned} \tag{9.5}$$

where $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ is a copy of $(\Omega,\mathcal{F},\mathbb{P})$ and $(\tilde{X},\tilde{b},\tilde{\sigma})$ is an independent copy of (X,b,σ).

Let us now give examples of functions belonging to the space $\mathcal{W}_2(\mathbb{R}^d)$.

Example 9.17. Let $g \in \mathcal{C}^0(\mathbb{R}^{2d};\mathbb{R})$ be a function such that its distributional derivative ∇g belongs to $(W^{1,k}(\mathbb{R}^{2d}))^{2d}$ for some $k \ge 5d$. Then, the function

$$u : \begin{cases} \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ (x,\mu) & \mapsto \displaystyle\int_{\mathbb{R}^d} g(x,y)\, d\mu(y) \end{cases}$$

belongs to $\mathcal{W}_2(\mathbb{R}^d)$ for $d_{\mathscr{P}} = d_k$.

The proof is postponed to Section 9.7.5.

Example 9.18. Let $F \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R};\mathbb{R})$ be a function such that for all $R > 0$

$$y \in \mathbb{R} \mapsto \nabla F(\cdot,y) \in (W^{1,k_1}(B_R))^{d+1},$$

is well-defined and continuous for some $k_1 \ge d+1$. Let $g \in \mathcal{C}^0(\mathbb{R}^d;\mathbb{R})$ be such that the distributional derivative ∇g belongs to $(W^{1,k_2}(\mathbb{R}^d))^d$ for some $k_2 \ge 2d$. Then

$$u : \begin{cases} \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) & \to \mathbb{R} \\ (x,\mu) & \mapsto F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right) \end{cases}$$

belongs to $\mathcal{W}_2(\mathbb{R}^d)$ for $d_{\mathscr{P}} = W_2$.

The proof is again postponed to Section 9.7.6.

Remark 9.19. In the abstract, we said that our Itô-Krylov's formula for a flow of measure was the almost analogue of the standard Itô-Krylov formula. We used the word "almost" because Assumption (1) in Definition 9.14 is not completely satisfactory. Indeed, we do not assume Sobolev regularity with respect to time, as it is the case in Itô-Krylov's formula for functions defined on $[0,T]\times\mathbb{R}^d$. Of course if u is of the form $u(t,\mu)=\int_{\mathbb{R}^d} g(t,x)\,d\mu(x)$ with $g\in\mathcal{C}^0([0,T]\times\mathbb{R}^d;\mathbb{R})$ at most of quadratic growth in x uniformly in t, and such that the distributional derivatives $\partial_t g$, $\partial_x g$ and $\partial_x^2 g$ are in $L^k([0,T]\times\mathbb{R}^d)$ for some $k\geq d+1$, we will succeed in proving Itô-Krylov's formula for u.

Let us give the idea of the proof. We regularize u by setting $u^n(t,\mu):=\int_{\mathbb{R}^d} g*\rho_n(t,x)\,d\mu(x)$, where $(\rho_n)_n$ is a mollifying sequence on $\mathbb{R}\times\mathbb{R}^d$. The function u^n clearly satisfies the assumptions of the standard Itô formula for a flow of measures (see Proposition 5.102 in [CD18a]). It ensures that for all $t\in[0,T]$

$$\begin{aligned}u^n(t,\mu_t)=u^n(0,\mu_0)&+\int_0^t\mathbb{E}(\partial_t g*\rho_n(s,X_s))\,ds+\int_0^t\mathbb{E}\left(\partial_x g*\rho_n(s,X_s)\cdot b_s\right)\,ds\\&+\frac{1}{2}\int_0^t\mathbb{E}\left(\partial_x^2 g*\rho_n(s,X_s)\cdot a_s\right)\,ds.\end{aligned}\tag{9.6}$$

As g is continuous, $(g*\rho_n)_n$ converges to g uniformly on compact sets. It follows from the growth assumption on g that u^n converges point-wise to u. Using that $(\partial_t g*\rho_n)_n$ converges in $L^k([0,T]\times\mathbb{R}^d)$ to $\partial_t g$ as $n\to+\infty$, we deduce with Krylov's inequality in Corollary 9.21 that for all $t\in[0,T]$

$$\int_0^t\mathbb{E}(\partial_t g*\rho_n(s,X_s))\,ds\to\int_0^t\mathbb{E}(\partial_t g(s,X_s))\,ds.$$

The same holds with the two other integrals in (9.6). Taking the limit $n\to+\infty$ in (9.6) yields for all $t\in[0,T]$

$$\begin{aligned}u(t,\mu_t)=u(0,\mu_0)&+\int_0^t\mathbb{E}(\partial_t g(s,X_s))\,ds+\int_0^t\mathbb{E}\left(\partial_x g(s,X_s)\cdot b_s\right)\,ds\\&+\frac{1}{2}\int_0^t\mathbb{E}\left(\partial_x^2 g(s,X_s)\cdot a_s\right)\,ds.\end{aligned}$$

In the general case, when the dependence in μ of the function u is not explicit, we cannot apply Krylov's inequality. Indeed, consider a function $u:[0,T]\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ such that, for all $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, $u(\cdot,\mu)\in W^{1,k}([0,T])$. In Itô's formula for u, as in the classical formula, there should be the term $\int_0^t\partial_t u(s,\mu_s)\,ds$. The assumption does not imply that this term is well-defined. One possible hypothesis is to assume that for all compact $\mathcal{K}\subset\mathcal{P}_2(\mathbb{R}^d)$, $\sup_{\mu\in\mathcal{K}}|\partial_t u(\cdot,\mu)|\in L^1([0,T])$. Following our strategy to prove Itô-Krylov's formula, we would consider the mollified version of u defined by $u^n(t,\mu):=u(\cdot,\mu*\rho_n^1)*\rho_n^2(t)$, where $(\rho_n^1)_n$ and $(\rho_n^2)_n$ are mollifying sequences on $\mathbb{R}^d$ and on $\mathbb{R}$ respectively. Assume that we have proved Itô's formula for u^n. In order to take the limit and deduce Itô's formula for u, we would like to show that

$$\int_0^T|\partial_t u(\cdot,\mu_s*\rho_n^1)*\rho_n^2(s)-\partial_t u(s,\mu_s)|\,ds\to 0.$$

However, this convergence is not obvious in the general case since the presence of μ_s prevents us from using the classical results on convolution and we cannot apply Krylov's inequality if the dependence in the measure argument is not linear.

9.4 Preliminaries

9.4.1 Krylov's inequality and densities.

The key element to prove the theorem is Krylov's inequality. We recall it in the next theorem taken from [Kry09] (see Theorem 4 in Section 2.3).

Theorem 9.20 (Krylov's inequality)**.** *Let $b : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^+ \times \Omega \to \mathbb{R}^{d\times d_1}$ be two progressively measurable functions. We assume that $p, d_1 \geq d$. Moreover, assume that there exists $K > 0$ and $\delta > 0$ such that*

- ***(A1)*** $\forall (t,\omega) \in \mathbb{R}^+ \times \Omega$, $|b_t(\omega)| + |\sigma_t(\omega)| \leq K$
- ***(A2)*** $\forall (t,\omega) \in \mathbb{R}^+ \times \Omega$, $\forall \lambda \in \mathbb{R}^d$, $a_t(\omega)\lambda \cdot \lambda \geq \delta|\lambda|^2$, *where* $a = \sigma\sigma^*$.

For X_0 a $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random variable, we define the Itô process $X = (X_t)_t$, for all $t \in [0,T]$, by

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dB_s.$$

Let $\lambda > 0$ be a positive constant. Then, there exists a constant $N = N(d,p,\lambda,\delta,K)$ such that for all measurable function $f : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}$

$$\mathbb{E}\int_0^\infty e^{-\lambda t}|f(t,X_t)|\,dt \leq N\|f\|_{L^{p+1}(\mathbb{R}^+\times\mathbb{R}^d)}.$$

We will use the following corollary for a finite horizon of time.

Corollary 9.21. *If b and σ satisfy Assumptions **(A)** and **(B)**, there exists $N_1 = N_1(d,p,\delta,K,T)$ such that for all measurable function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, we have*

$$\mathbb{E}\int_0^T |f(s,X_s)|\,ds \leq N_1\|f\|_{L^{p+1}([0,T]\times\mathbb{R}^d)}.$$

Proof. We set $b_t = b_T$ and $\sigma_t = \sigma_T$ for $t > T$ to guarantee that Assumptions **(A1)** and **(A2)** are satisfied, without changing the process X on $[0,T]$. It remains to apply Krylov's inequality to $\tilde{f}(t,x) := f(t,x)\mathbf{1}_{t\in[0,T]}$, which gives the existence of $N_1 = N_1(d,p,\delta,K)$ such that

$$e^{-T}\mathbb{E}\int_0^T |f(s,X_s)|\,ds \leq N_1\|f\|_{L^{p+1}([0,T]\times\mathbb{R}^d)}.$$

□

Krylov's inequality also provides the existence of a density with respect to the Lebesgue measure for μ_s, the law of X_s, for almost all $s \in [0,T]$.

Proposition 9.22. *Under Assumptions **(A)** and **(B)** on the coefficients b and σ, there exists a function $p \in L^1([0,T]\times\mathbb{R}^d;\mathbb{R}^+) \cap L^{(d+1)'}([0,T]\times\mathbb{R}^d;\mathbb{R}^+)$ such that for all $f : [0,T]\times\mathbb{R}^d \to \mathbb{R}^+$ measurable*

$$\int_0^T \mathbb{E}f(s,X_s)\,ds = \int_{[0,T]\times\mathbb{R}^d} f(s,x)p(s,x)\,dx\,ds. \tag{9.7}$$

If τ is a stopping time such that $(X_t)_{t\in[0,T]}$ belongs to B_R almost surely on the set $\{\tau > 0\}$, then

$$\mathbb{E}\int_0^{\tau\wedge T} f(s,X_s)\,ds \le \int_{[0,T]\times B_R} f(s,x)p(s,x)\,dx\,ds. \tag{9.8}$$

Moreover, for almost all $s\in[0,T]$, $\mu_s = \mathcal{L}(X_s)$ is equal to $p(s,\cdot)\,dx$.

We give the proof for the sake of completeness.

Proof. We denote by μ the push-forward measure of $\lambda\otimes\mathbb{P}$, where λ is the Lebesgue measure on $[0,T]$, by the measurable map $(t,\omega)\in[0,T]\times\Omega\mapsto(t,X_t(\omega))\in[0,T]\times\mathbb{R}^d$ defined, for any $A\in\mathcal{B}([0,T])\otimes\mathcal{B}(\mathbb{R}^d)$, by

$$\mu(A) = \int_0^T \mathbb{E}\mathbf{1}_A(s,X_s)\,ds.$$

Note that μ is a finite measure on $[0,T]\times\mathbb{R}^d$. The monotone convergence theorem and Krylov's inequality ensure that for all $f:[0,T]\times\mathbb{R}^d\to\mathbb{R}^+$ measurable

$$\int_0^T \mathbb{E}f(s,X_s)\,ds = \int_{[0,T]\times\mathbb{R}^d} f(s,x)\,d\mu(s,x) \le C\|f\|_{L^{p+1}([0,T]\times\mathbb{R}^d)}.$$

Taking $f=\mathbf{1}_A$, for $A\in\mathcal{B}([0,T])\otimes\mathcal{B}(\mathbb{R}^d)$ with Lebesgue measure 0, we deduce that $\mu(A)=0$. Thus μ is absolutely continuous with respect to the Lebesgue measure on $[0,T]\times\mathbb{R}^d$. Radon-Nikodym's theorem provides the existence of $p\in L^1([0,T]\times\mathbb{R}^d;\mathbb{R}^+)$ such that for all measurable function $f:[0,T]\times\mathbb{R}^d\to\mathbb{R}^+$

$$\int_0^T \mathbb{E}f(s,X_s)\,ds = \int_{[0,T]\times\mathbb{R}^d} f(s,x)p(s,x)\,dx\,ds. \tag{9.9}$$

Krylov's inequality exactly proves that the map $f\in L^{d+1}([0,T]\times\mathbb{R}^d)\mapsto\int_{[0,T]\times\mathbb{R}^d} f(s,x)p(s,x)\,dx\,ds$ is a continuous linear form. Since the dual space of $L^{d+1}([0,T]\times\mathbb{R}^d)$ is $L^{(d+1)'}([0,T]\times\mathbb{R}^d)$, p belongs to $L^{(d+1)'}([0,T]\times\mathbb{R}^d)$.

To prove (9.8), it is enough to notice that

$$\mathbb{E}\int_0^{\tau\wedge T} f(s,X_s)\,ds \le \mathbb{E}\int_0^T f(s,X_s)\mathbf{1}_{B_R}(X_s)\,ds.$$

Next, we establish that for almost all $s\in[0,T]$, $\mu_s = p(s,\cdot)\,dx$. We fix $s\in[0,T]$, $n\ge 1$ large enough and $A\in\mathcal{B}(\mathbb{R}^d)$. Applying (9.9) with $f=\mathbf{1}_{[s-1/n,s+1/n]\times A}$, and using Fubini-Tonelli's theorem, we deduce that

$$\frac{n}{2}\int_{s-1/n}^{s+1/n}\mathbb{P}(X_t\in A)\,dt = \frac{n}{2}\int_{s-1/n}^{s+1/n}\int_A p(t,x)\,dx\,ds.$$

Since $t\mapsto\mathbb{P}(X_t\in A)$ is bounded and as Fubini's theorem implies that $t\mapsto\int_A p(t,x)\,dx$ belongs to $L^1([0,T])$, it follows from Lebesgue differentiation theorem (see Theorem 7.7 in [Rud87]) that for almost all $s\in[0,T]$

$$\mathbb{P}(X_s\in A) = \int_A p(s,x)\,dx.$$

We denote by $\mathcal{R}$ the set of all Borel sets in $\mathbb{R}^d$ of the form $\prod_{i=1}^d]a_i,b_i[$, with $a_i<b_i$ two rational numbers

for all i. The set $\mathcal{R}$ is at most countable, thus for almost $s \in [0,T]$

$$\forall A \in \mathcal{R}, \quad \mathbb{P}(X_s \in A) = \int_A p(s,x)\,dx.$$

The monotone class theorem enables us to conclude. $\square$

Note that for almost all $s \in [0,T]$, $p(s,\cdot) \in L^{(d+1)'}(\mathbb{R}^d)$ using Fubini-Tonelli's theorem. We deduce the following corollary.

Corollary 9.23. *For almost all $s \in [0,T]$, $\mu_s \in \mathscr{P}(\mathbb{R}^d)$.*

We now prove two lemmas dealing with the integrability of the density p.

Lemma 9.24. *Let p be the density given by Proposition 9.22. Then for all $k \geq d+1$*

$$s \in [0,T] \mapsto \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)} \in L^{k/d}([0,T]).$$

Proof. Using Jensen's inequality since $\frac{k}{k'} = k-1 \geq d$, we obtain that

$$\begin{aligned}\int_0^T \left(\int_{\mathbb{R}^d} p(s,x)^{k'}\,dx\right)^{\frac{k}{dk'}} ds &= \int_0^T \left(\int_{\mathbb{R}^d} p(s,x)^{k'-1} p(s,x)\,dx\right)^{\frac{k}{dk'}} ds \\ &\leq \int_0^T \int_{\mathbb{R}^d} p(s,x)^{\frac{k}{dk'}(k'-1)+1}\,dx\,ds.\end{aligned}$$

By definition of the conjugate exponent, we get

$$\int_0^T \int_{\mathbb{R}^d} p(s,x)^{\frac{k}{dk'}(k'-1)+1}\,dx\,ds = \int_0^T \int_{\mathbb{R}^d} p(s,x)^{\frac{1}{d}+1}\,dx\,ds,$$

which is finite since $(d+1)' = \frac{1}{d}+1$ and $p \in L^{(d+1)'}([0,T]\times\mathbb{R}^d)$. $\square$

Lemma 9.25. *Let p and q be two densities of two Itô processes of the form (9.1) and satisfying **(A)** and **(B)** given by Proposition 9.22. Then for $k, \alpha \in \mathbb{N}$ such that $k \geq \max\{d+1, d(\alpha+1)\}$, we have*

$$\int_0^T \|p(s,\cdot)\|^{\alpha}_{L^{k'}(\mathbb{R}^d)} \|q(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}\,ds < +\infty.$$

Proof. Owing to Lemma 9.24, the function $s \mapsto \|q(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}$ belongs to $L^1([0,T]) \cap L^{k/d}([0,T])$. Using Hölder's inequality, the proof is complete once we prove that $s \mapsto \|p(s,\cdot)\|^{\alpha}_{L^{k'}(\mathbb{R}^d)}$ belongs to $L^r([0,T])$ for some $r \geq \left(\frac{k}{d}\right)'$. Lemma 9.24 ensures that $s \mapsto \|p(s,\cdot)\|^{\alpha}_{L^{k'}(\mathbb{R}^d)} \in L^{\frac{k}{\alpha d}}([0,T])$ thus we have to prove that $\left(\frac{k}{d}\right)' \leq \frac{k}{\alpha d}$. This is equivalent to our assumption $k \geq d(\alpha+1)$. $\square$

9.4.2 Classical results on convolution and regularization.

Fix $p \in [1+\infty[$. We will need the two following basic lemmas, which we recall for the sake of clarity.

Lemma 9.26 (Convolution)**.** - *For all $f \in L^p(\mathbb{R}^d)$ and for all $g \in L^1(\mathbb{R}^d)$, the convolution $f * g$ is well-defined and belongs to $L^p(\mathbb{R}^d)$. Moreover, we have $\|f * g\|_{L^p} \leq \|f\|_{L^p}\|g\|_{L^1}$.*

- *For all $f \in L^p(\mathbb{R}^d)$ and for all $g \in L^{p'}(\mathbb{R}^d)$, the convolution $f * g$ is well-defined and belongs to $L^\infty(\mathbb{R}^d)$. Moreover, we have $\|f * g\|_{L^\infty} \leq \|f\|_{L^p}\|g\|_{L^{p'}}$.*

Lemma 9.27 (Regularization)**.** *Recall that $(\rho_n)_n$ is a mollifying sequence.*

- *Let $f \in L^1_{loc}(\mathbb{R}^d)$ and $\rho \in \mathcal{C}^\infty_c(\mathbb{R}^d)$. Then $f * \rho \in \mathcal{C}^\infty(\mathbb{R}^d)$ and $\forall \alpha \in \mathbb{N}^d$, $\partial^\alpha(f * \rho) = f * \partial^\alpha \rho$.*
- *If $f \in L^p(\mathbb{R}^d)$, then $f * \rho_n \xrightarrow{L^p} f$, and if $f \in \mathcal{C}^0(\mathbb{R}^d)$, $f * \rho_n \to f$ uniformly on compact sets.*
- *If $f \in L^p_{loc}(\mathbb{R}^d)$, then for all $R > 0$, $f * \rho_n \to f$ in $L^p(B_R)$.*

The following proposition will also be useful.

Proposition 9.28. *Let $f \in \mathcal{C}^0(\mathbb{R}^d)$ be a function admitting distributional derivatives of order* 1 *et* 2 *in $L^1_{loc}(\mathbb{R}^d)$. Then $f * \rho_n \in \mathcal{C}^\infty(\mathbb{R}^d)$ and for all $i, j \in \{1, \dots d\}$*

$$\begin{cases} \partial_{x_i}(f * \rho_n) &= \partial_{x_i} f * \rho_n \\ \partial_{x_i x_j}(f * \rho_n) &= \partial_{x_i x_j} f * \rho_n. \end{cases}$$

The next lemma deals with the convolution of a function $f \in L^p$ with $\mu \in \mathcal{P}(\mathbb{R}^d)$.

Lemma 9.29. *Let $f \in L^p(\mathbb{R}^d)$. Then $\mu \in \mathcal{P}(\mathbb{R}^d) \mapsto f * \mu \in L^p(\mathbb{R}^d)$ is continuous.*

Proof. Note that the convolution $f * \mu$ is well-defined as an element of $L^p(\mathbb{R}^d)$ thanks to Jensen's inequality which shows that

$$\forall f \in L^p(\mathbb{R}^d), \forall \mu \in \mathcal{P}(\mathbb{R}^d), \|f * \mu\|_{L^p} \leq \|f\|_{L^p}.$$

Let $(\mu_n)_n$ be a sequence of $\mathcal{P}(\mathbb{R}^d)$ weakly convergent to $\mu \in \mathcal{P}(\mathbb{R}^d)$. Using Skorokhod's representation theorem (see Theorem 6.7 in [Bil99]), there exists a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$, a sequence of random variables $(X_n)_n$ converging $\mathbb{P}'$-almost surely to a random variable X such that, the law of X_n is μ_n for all n and the law of X if μ. For any $a \in \mathbb{R}^d$, let us denote by $\tau_a f$ the translation of f defined, for all $x \in \mathbb{R}^d$, by $\tau_a f(x) := f(x - a)$. Jensen's inequality and Fubini-Tonelli's theorem yield

$$\begin{aligned} \|f * \mu_n - f * \mu\|^p_{L^p} &= \int_{\mathbb{R}^d} |\mathbb{E}'(f(x - X_n) - f(x - X))|^p \, dx \\ &\leq \int_{\mathbb{R}^d} \mathbb{E}'(|f(x - X_n) - f(x - X)|^p) \, dx \\ &= \mathbb{E}'(\|\tau_{X_n - X} f - f\|^p_{L^p}). \end{aligned}$$

It follows from the almost sure convergence of $(X_n)_n$ to X and the continuity of the translation operator in L^p that $\|\tau_{X_n - X} f - f\|^p_{L^p} \xrightarrow{a.s.} 0$. Moreover, the inequality

$$\begin{aligned} \|\tau_{X_n - X} f - f\|^p_{L^p} &\leq 2^{p-1}(\|\tau_{X_n - X} f\|^p_{L^p} + \|f\|^p_{L^p}) \\ &= 2^p \|f\|^p_{L^p}, \end{aligned}$$

enables us to conclude with the dominated convergence theorem. □

9.4.3 Convolution of probability measures

Lemma 9.30 (Contraction inequality)**.** *Fix $\mu, \nu, m \in \mathcal{P}_2(\mathbb{R}^d)$. Then, we have*

$$W_2(\mu * m, \nu * m) \leq W_2(\mu, \nu).$$

Proof. Let $\pi \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ be an optimal coupling between μ and ν. We consider a couple of random variables (X, Y) with law π, and a random variable Z independent of (X, Y) with law m. The law of $X + Z$ being $\mu * m$ and the law of $Y + Z$ being $\nu * m$, one has

$$W_2(\mu * m, \nu * m) \leq \|(X + Z) - (Y + Z)\|_{L^2} = W_2(\mu, \nu).$$

□

The next corollary follows from the fact that $\rho_n \xrightarrow{W_2} \delta_0$.

Corollary 9.31. *For all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu * \rho_n \xrightarrow{W_2} \mu$.*

9.4.4 Measurability

We will need the following lemma to guarantee that, for $u \in \mathcal{W}_1(\mathbb{R}^d)$, we can find versions of $\partial_v \frac{\delta u}{\delta m}$ and $\partial_v^2 \frac{\delta u}{\delta m}$ that are measurable with respect to $(\mu, v) \in \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d$.

Lemma 9.32. *Let $u : E \to L^k(\mathbb{R}^d)$ be a continuous function, where E is a metric space and $k > 1$. Then, for all $x \in E$, we can find a version of $u(x)$ such that $(x, v) \in E \times \mathbb{R}^d \mapsto u(x)(v)$ is measurable with respect to $\mathcal{B}(E) \otimes \mathcal{B}(\mathbb{R}^d)$.*

Proof. For $(x, v) \in E \times \mathbb{R}^d$, we define

$$\tilde{u}(x, v) = \varliminf_{n \to +\infty} \frac{1}{\lambda(B(v, 1/n))} \int_{B(v,1/n)} u(x)(y)\, dy = \varliminf_{n \to +\infty} u^n(x, v),$$

where λ denotes the Lebesgue measure on $\mathbb{R}^d$. From Lebesgue differentiation theorem (see Theorem 7.7 in [Rud87]), we deduce that for all $x \in E$, $\tilde{u}(x, \cdot) = u(x)$ λ-almost everywhere. We prove that for all $n \geq 1$, u^n is continuous. Note that $\frac{1}{\lambda(B(v,1/n))}$ does not depend on v. The continuity of u^n follows from the continuity of $x \in E \mapsto u(x) \in L^k(\mathbb{R}^d)$, $v \in \mathbb{R}^d \mapsto \mathbf{1}_{B(v,1/n)} \in L^{k'}(\mathbb{R}^d)$ (coming from the dominated convergence theorem), and of $(f, g) \in L^k(\mathbb{R}^d) \times L^{k'}(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} fg\, dx$. □

9.5 Proof of Theorem 9.7

The proof will be divided into three parts. Step 1 is dedicated to prove that all the terms in Itô-Krylov's formula (9.3) are well-defined. In Step 2, we regularize u by convolution of the measure argument with a mollifying sequence $(\rho_n)_n$. The effect of replacing $u(\mu)$ by $u(\mu * \rho_n)$ is that the linear derivative is regularized by convolution, in its space variable. Then, we apply the standard Itô's formula for a flow of measure. We finally take the limit $n \to +\infty$ in Step 3 with the help of Krylov's inequality.

Step 1: All the terms in (9.3) are well-defined.

Let us show that the two integrals in (9.3) are well-defined.

Measurability. Thanks to Lemma 9.32, we can find a version of $\partial_v \frac{\delta u}{\delta m}$ which is measurable with respect to $(\mu, v) \in \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d$. To conclude, we prove that $s \mapsto \mu_s \in \mathscr{P}(\mathbb{R}^d)$ is measurable. Indeed if it is the case, the function $(s,\omega) \in [0,T] \times \Omega \mapsto \partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s(\omega)).b_s(\omega)$ will be measurable by composition. First, note that $\mu_s \in \mathscr{P}(\mathbb{R}^d)$ for almost all $s \in [0,T]$ (see Corollary 9.23) so we can change μ_s on a negligible set of times s to ensure that $\mu_s \in \mathscr{P}(\mathbb{R}^d)$ for all $s \in [0,T]$. But $\mu_s = \lim_{n \to +\infty} \mu_s * \rho_n$ for $d_{\mathscr{P}}$ by Assumption **(H2)** in Definition 9.3. It remains to show that $s \mapsto \mu_s * \rho_n \in \mathscr{P}(\mathbb{R}^d)$ is continuous and thus mesurable for all n. This follows from the continuity of $s \mapsto \mu_s \in \mathcal{P}_2(\mathbb{R}^d)$ and also from Assumption **(H1)** in Definition 9.3.

Integrability. We can omit the coefficients b and a to prove the integrability properties because they are uniformly bounded. Taking advantage from the existence of a density coming from Proposition 9.22, we have by Hölder's inequality

$$\begin{aligned}\int_0^T \mathbb{E}\left|\partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s)\right| ds &= \int_0^T \int_{\mathbb{R}^d} \left|\partial_v \frac{\delta u}{\delta m}(\mu_s)(x)\right| p(s,x)\, dx\, ds \\ &\leq \int_0^T \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s)(\cdot)\right\|_{L^k(\mathbb{R}^d)} \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}\, ds \\ &\leq \int_0^T C\left(1 + \|p(s,\cdot)\|^{\alpha}_{L^{k'}(\mathbb{R}^d)}\right) \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}\, ds,\end{aligned}$$

for some constant C coming from Assumption (2) in Definition 9.5 because the flow $(\mu_s)_{s \leq T}$ is compact in $\mathcal{P}_2(\mathbb{R}^d)$ and belongs to $\mathscr{P}(\mathbb{R}^d)$ for almost all s. The last bound is finite thanks to Lemma 9.24 since $k \geq \max\{d(\alpha+1), d+1\}$. The same properties hold for the term involving $\partial^2_v \frac{\delta u}{\delta m}$.

Step 2: Itô's formula for the mollification of u.

For $n \geq 1$, we set $u^n : \mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto u(\mu * \rho_n)$. By standard arguments, for each $n \geq 1$, u^n has a linear derivative given by

$$\frac{\delta u^n}{\delta m}(\mu)(v) = \int_{\mathbb{R}^d} \frac{\delta u}{\delta m}(\mu * \rho_n)(x)\rho_n(v-x)\, dx = \frac{\delta u}{\delta m}(\mu * \rho_n) * \rho_n(v).$$

Now, we aim at applying the standard Itô formula for a flow of probability measures (see for example Theorem 5.99 in Chapter 5 of [CD18a] with the L-derivative) to u^n for a fixed $n \geq 1$.

(i) Regularity of $\frac{\delta u^n}{\delta m}(\mu)$ for a fixed $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Since for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu * \rho_n \in \mathscr{P}(\mathbb{R}^d)$, Proposition 9.28 implies that $\frac{\delta u^n}{\delta m}(\mu)(.) \in \mathcal{C}^\infty(\mathbb{R}^d)$ and for all $i,j \in \{1,\dots,d\}$

$$\partial_{v_i} \frac{\delta u^n}{\delta m}(\mu) = \partial_{v_i} \frac{\delta u}{\delta m}(\mu * \rho_n) * \rho_n \quad \text{and} \quad \partial_{v_i v_j} \frac{\delta u^n}{\delta m}(\mu) = \partial_{v_i v_j} \frac{\delta u}{\delta m}(\mu * \rho_n) * \rho_n.$$

(ii) Continuity of $\partial_v \frac{\delta u^n}{\delta m}$ and $\partial^2_v \frac{\delta u^n}{\delta m}$ with respect to (μ, v). Let $i \in \{1,\dots,d\}$, $(\mu_m)_m \in \mathcal{P}_2(\mathbb{R}^d)^{\mathbb{N}}$

and $(v_m)_m \in (\mathbb{R}^d)^{\mathbb{N}}$ be sequences converging respectively to μ and v. We have

$$\begin{aligned}
&\left|\partial_{v_i}\frac{\delta u^n}{\delta m}(\mu_m)(v_m) - \partial_{v_i}\frac{\delta u^n}{\delta m}(\mu)(v)\right| \\
&\leq \left|\partial_{v_i}\frac{\delta u^n}{\delta m}(\mu_m)(v_m) - \partial_{v_i}\frac{\delta u^n}{\delta m}(\mu)(v_m)\right| + \left|\partial_{v_i}\frac{\delta u^n}{\delta m}(\mu)(v_m) - \partial_{v_i}\frac{\delta u^n}{\delta m}(\mu)(v)\right| \\
&=: D_1 + D_2
\end{aligned}$$

D_2 converges to 0 when $m \to +\infty$ by (i). For D_1, the convolution inequality $L^k * L^{k'}$ gives that

$$\begin{aligned}
D_1 &= \left|\partial_{v_i}\frac{\delta u}{\delta_m}(\mu_m * \rho_n) * \rho_n(v_m) - \partial_{v_i}\frac{\delta u}{\delta_m}(\mu * \rho_n) * \rho_n(v_m)\right| \\
&\leq \left\|\partial_{v_i}\frac{\delta u}{\delta_m}(\mu_m * \rho_n) - \partial_{v_i}\frac{\delta u}{\delta_m}(\mu * \rho_n)\right\|_{L^k} \|\rho_n\|_{L^{k'}}.
\end{aligned}$$

Assumption **(H1)** in Definition 9.3 provides that $\mu_m * \rho_n \xrightarrow{d_{\mathscr{P}}} \mu * \rho_n$ when $m \to +\infty$. Finally, using the first assumption in Definition 9.5, we conclude that D_1 converges to 0 when $m \to +\infty$. This shows the continuity of $\partial_v \frac{\delta u^n}{\delta m}$ on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$. The same reasoning proves the joint continuity of $\partial_v^2 \frac{\delta u^n}{\delta m}$.

(iii) Boundedness of $\partial_v \frac{\delta u^n}{\delta m}$ and $\partial_v^2 \frac{\delta u^n}{\delta m}$. Let $\mathcal{K} \subset \mathcal{P}_2(\mathbb{R}^d)$ be a compact set. For $\mu \in \mathcal{K}$ and $v \in \mathbb{R}^d$, one has

$$\left|\partial_{v_i}\frac{\delta u^n}{\delta m}(\mu)(v)\right| \leq \left\|\partial_{v_i}\frac{\delta u}{\delta_m}(\mu * \rho_n)\right\|_{L^k} \|\rho_n\|_{L^{k'}}.$$

The set $\{\mu * \rho_n,\, \mu \in \mathcal{K}\}$ is compact in $(\mathscr{P}(\mathbb{R}^d), d_{\mathscr{P}})$ as the image of the compact $\mathcal{K}$ by the application $\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \mu * \rho_n \in \mathscr{P}(\mathbb{R}^d)$ which is continuous by Assumption **(H1)** in Definition 9.3. The first assumption in Definition 9.5 guarantees that $\sup_{\mu\in\mathcal{K}} \left\|\partial_{v_i}\frac{\delta u}{\delta_m}(\mu * \rho_n)\right\|_{L^k(\mathbb{R}^d)} < +\infty$ and thus

$$\sup_{v\in\mathbb{R}^d}\sup_{\mu\in\mathcal{K}} \left|\partial_v \frac{\delta u^n}{\delta m}(\mu)(v)\right| < \infty.$$

The same property holds for $\partial_v^2 \frac{\delta u^n}{\delta m}$.

We can thus apply Itô's formula of [CD18a] to obtain that for all $n \geq 1$ and for all $t \in [0, T]$

$$u^n(\mu_t) = u^n(\mu_0) + \int_0^t \mathbb{E}\left(\partial_v \frac{\delta u^n}{\delta m}(\mu_s)(X_s)\cdot b_s\right) ds + \frac{1}{2}\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta u^n}{\delta m}(\mu_s)(X_s)\cdot a_s\right) ds. \tag{9.10}$$

Step 3: Letting $n \to +\infty$.
Our aim is now to take the limit $n \to +\infty$ in (9.10). As for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mu * \rho_n \xrightarrow{W_2} \mu$ and u is continuous on $\mathcal{P}_2(\mathbb{R}^d)$, we deduce that $(u^n)_n$ converges pointwise to u. It remains to take the limit in the two integrals of (9.10). We show that

$$\int_0^t \mathbb{E}\left(\partial_v \frac{\delta u^n}{\delta m}(\mu_s)(X_s)\cdot b_s\right) ds \to \int_0^t \mathbb{E}\left(\partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s)\cdot b_s\right) ds. \tag{9.11}$$

Since b is uniformly bounded, it is enough to prove that

$$\mathbb{E}\int_0^T \left|\partial_v \frac{\delta u}{\delta m}(\mu * \rho_n) * \rho_n(X_s) - \partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s)\right| ds \to 0.$$

By Proposition 9.22, Hölder's inequality and then the $L^1 * L^k$ convolution inequality, one has

$$\begin{aligned}
&\mathbb{E}\int_0^T \left|\partial_v \frac{\delta u}{\delta m}(\mu_s * \rho_n) * \rho_n(X_s) - \partial_v \frac{\delta u}{\delta m}(\mu_s)(X_s)\right| ds \\
&\leq \int_0^T \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s * \rho_n) - \partial_v \frac{\delta u}{\delta m}(\mu_s)\right\|_{L^k(\mathbb{R}^d)} \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}\, ds \\
&\quad + \int_0^T \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s) * \rho_n - \partial_v \frac{\delta u}{\delta m}(\mu_s)\right\|_{L^k(\mathbb{R}^d)} \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}\, ds \\
&=: I_1 + I_2.
\end{aligned}$$

The integrand in I_1 converges to 0 for almost all s using Assumption (1) in Theorem 9.7 and the fact that $\mu_s * \rho_n \xrightarrow{d_{\mathscr{P}}} \mu_s$ for almost all s thanks to Assumption **(H2)** in Definition 9.3. Let us now prove that the dominated convergence theorem applies. The integrand is bounded by

$$\left[\sup_{n\geq 1}\left\|\partial_v \frac{\delta u}{\delta m}(\mu_s * \rho_n)\right\|_{L^k(\mathbb{R}^d)} + \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s)\right\|_{L^k(\mathbb{R}^d)}\right] \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)}.$$

Note that the set $\{\mu_s * \rho_n,\ s \in [0,T],\ n \geq 1\} \cup \{\mu_s,\ s \in [0,T]\}$ is compact in $\mathcal{P}_2(\mathbb{R}^d)$. Indeed, if $(s_k)_k \in [0,T]^{\mathbb{N}}$ and $(n_k)_k \in \mathbb{N}^{\mathbb{N}}$ are two sequences, we have to find a convergent subsequence from $(\mu_{s_k} * \rho_{n_k})_k$. Up to an extraction, we can assume that $(s_k)_k$ converges to some $s \in [0,T]$. There are two cases. If there exists l such that $n_k = l$ infinitely often, then $\mu_{s_k} * \rho_l \xrightarrow{W_2} \mu_s * \rho_l$ by the contraction inequality (see Lemma 9.30). Otherwise, we can assume that $(n_k)_k$ converges to $+\infty$. We use the triangle inequality to get

$$W_2(\mu_{s_k} * \rho_{n_k}, \mu_s) \leq W_2(\mu_{s_k} * \rho_{n_k}, \mu_s * \rho_{n_k}) + W_2(\mu_s * \rho_{n_k}, \mu_s).$$

The last term converges to 0 owing to Lemma 9.31, and the first is bounded by $W_2(\mu_{s_k}, \mu_s)$ by the contraction inequality (see Lemma 9.30), which converges to 0. Thus Assumption (2) in Definition 9.5 ensures that there exists $C > 0$ such that for almost all $s \in [0,T]$ and for all n

$$\left\|\partial_v \frac{\delta u}{\delta m}(\mu_s * \rho_n)\right\|_{L^k(\mathbb{R}^d)} + \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s)\right\|_{L^k(\mathbb{R}^d)} \leq C(1 + \|p(s,\cdot) * \rho_n\|^\alpha_{L^{k'}(\mathbb{R}^d)} + \|p(s,\cdot)\|^\alpha_{L^{k'}(\mathbb{R}^d)}).$$

It follows from the convolution inequality $L^{k'} * L^1$ that for almost all s

$$\begin{aligned}
&\left[\sup_{n\geq 1}\left\|\partial_v \frac{\delta u}{\delta m}(\mu_s * \rho_n)\right\|_{L^k(\mathbb{R}^d)} + \left\|\partial_v \frac{\delta u}{\delta m}(\mu_s)\right\|_{L^k(\mathbb{R}^d)}\right] \|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)} \\
&\leq 2C(1 + \|p(s,\cdot)\|^\alpha_{L^{k'}(\mathbb{R}^d)})\|p(s,\cdot)\|_{L^{k'}(\mathbb{R}^d)},
\end{aligned}$$

which is integrable on $[0,T]$ thanks to Lemma 9.24 since $k \geq \max\{d(\alpha+1), d+1\}$. We conclude by the dominated convergence theorem that I_1 converges to 0. The term I_2 also converges to 0 following the same method. Indeed, for almost all s, $\partial_v \frac{\delta u}{\delta m}(\mu_s)(\cdot) \in L^k(\mathbb{R}^d)$ thus the integrand converges to 0 by Lemma 9.27 and we conclude with the dominated convergence theorem. Therefore (9.11) is proved. Following the same lines, we take the limit $n \to +\infty$ in the last integral of (9.10) to obtain that for all $t \in [0, T]$

$$\int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta u^n}{\delta m}(\mu_s)(X_s) \cdot a_s\right) ds \to \int_0^t \mathbb{E}\left(\partial_v^2 \frac{\delta u}{\delta m}(\mu_s)(X_s) \cdot a_s\right) ds.$$

This concludes the proof of Theorem 9.7. □

9.6 Proof of Theorem 9.16

The strategy of the proof is the following. In Step 1, we prove some integrability results coming from Assumption (5) in Definition 9.14. Step 2 is devoted to prove that all the terms in Itô-Krylov's formula (9.5) are well-defined using a localization argument, Krylov's inequality, and Step 1. Moreover, we see that it is enough to prove the formula up to random times localizing the process ξ. Step 3 is dedicated to regularize u using convolutions both in space and measure variables. In Step 4 and 5, we follow the strategy of the proof of Theorem 5.102 in [CD18a] to prove Itô-Krylov's formula for u^n, the mollified version of u. Finally, Step 6 aims at taking the limit $n \to +\infty$ thanks to Krylov's inequality.

Note that there are three kind of integrals in Itô's formula (9.5): the terms involving standard time and space derivatives in the first line, those involving the linear derivative in the second line and the martingale term in the third line. We will treat them separately.

Step 1: Useful integrability results.

It follows from Assumption (5) in Definition 9.14 and Lemma 9.25 that for any $M > 0$ the following quantities are finite:

$$J_1(M) := \int_0^T \left[\sup_{n\geq 1} \|\partial_x u(s, \cdot, \mu_s * \rho_n)\|_{L^{k_1}(B_M)} + \sup_{n\geq 1} \left\|\partial_x^2 u(s, \cdot, \mu_s * \rho_n)\right\|_{L^{k_1}(B_M)}\right] \|q(s, \cdot)\|_{L^{k_1'}(B_M)} ds, \tag{9.12}$$

$$J_2(M) := \int_0^T \sup_{n\geq 1} \|\partial_x u(s, \cdot, \mu_s * \rho_n)\|^2_{L^{2k_1}(B_M)} \|q(s, \cdot)\|_{L^{k_1'}(B_M)} ds,$$

$$J_3(M) := \int_0^T \sup_{n\geq 1} \left\|\partial_v \frac{\delta u}{\delta m}(s, \cdot, \mu_s * \rho_n)(\cdot)\right\|_{L^{k_2}(B_M\times\mathbb{R}^d)} \|q(s, \cdot)\|_{L^{k_2'}(B_M)} \|p(s, \cdot)\|_{L^{k_2'}(\mathbb{R}^d)} ds,$$

$$J_4(M) := \int_0^T \sup_{n\geq 1} \left\|\partial_v^2 \frac{\delta u}{\delta m}(s, \cdot, \mu_s * \rho_n)(\cdot)\right\|_{L^{k_2}(B_M\times\mathbb{R}^d)} \|q(s, \cdot)\|_{L^{k_2'}(B_M)} \|p(s, \cdot)\|_{L^{k_2'}(\mathbb{R}^d)} ds.$$

To prove this, we follow the method employed in Step 3 of the preceding proof to justify the dominated convergence theorem. We just give details for $J_2(M)$ since it requires a bit more attention. Owing to Assumption (2) in Definition 9.14, we know that for all $(t, \mu) \in [0, T] \times \mathscr{P}(\mathbb{R}^d)$, $\partial_x u(t, \cdot, \mu) \in W^{1,k_1}(B)$. Sobolev embedding theorem (see Corollary 9.14 in [Bre10]) ensures that the embedding $W^{1,k_1}(B_M) \hookrightarrow L^{2k_1}(B_M)$ is continuous since $k_1 \geq d + 1$. Thus there exists $C > 0$ such that

$$\forall t \in [0, T], \forall \mu \in \mathscr{P}(\mathbb{R}^d), \|\partial_x u(t, \cdot, \mu)\|_{L^{2k_1}(B_M)} \leq C \left(\|\partial_x u(t, \cdot, \mu)\|_{L^{k_1}(B_M)} + \|\partial_x^2 u(t, \cdot, \mu)\|_{L^{k_1}(B_M)}\right).$$

Thanks to Assumption (5) in Definition 9.14, there exists a constant $C_M > 0$ such that for almost all s and for all $n \geq 1$

$$\sup_{n\geq 1} \|\partial_x u(s, \cdot, \mu_s * \rho_n)\|^2_{L^{2k_1}(B_M)} \leq C_M \left(1 + \|p(s,\cdot)\|^{2\alpha_1}_{L^{k_1'}(\mathbb{R}^d)}\right),$$

where we used the fact that $\{\mu_s * \rho_n,\ s \in [0,T],\ n \geq 1\}$ is relatively compact in $\mathcal{P}_2(\mathbb{R}^d)$ and the convolution inequality $L^{k_1'} * L^1$. We conclude with Lemma 9.25 since $k_1 \geq \max\{d(2\alpha_1+1), d+1\}$. Note that these integrability properties remain true if we replace $\mu_s * \rho_n$ by μ_s and remove the supremum. We justify it only for the second point. It follows from the continuity assumption (2) in Definition 9.14 that for almost all $s \in [0,T]$

$$\partial_x u(s, \cdot, \mu_s * \rho_n) \overset{W^{1,k_1}(B_M)}{\longrightarrow} \partial_x u(s,\cdot,\mu_s),$$

because $\mu_s * \rho_n \overset{d_{\mathscr{P}}}{\longrightarrow} \mu_s$ for almost all s. Sobolev embedding theorem guarantees that

$$\|\partial_x u(t,\cdot,\mu * \rho_n)\|_{L^{2k_1}(B_M)} \to \|\partial_x u(t,\cdot,\mu)\|_{L^{2k_1}(B_M)}.$$

Thus we obtain

$$\int_0^T \|\partial_x u(s,\cdot,\mu_s)\|^2_{L^{2k_1}(B_M)} \|q(s,\cdot)\|_{L^{k_1'}(B_M)}\, ds \leq J_2(M) < +\infty.$$

Step 2: Meaning of the terms in (9.5) and localization.

Let $(T_M)_M$ be the sequence of stopping times converging almost surely to T defined by

$$T_M = \inf\{t \in [0,T],\ |\xi_t| \geq M\} \wedge T.$$

Let $\xi_t^M = \xi_{t\wedge T_M}$, which is bounded by M on the set $\{T_M > 0\}$.

(i) Terms involving standard derivatives in (9.5). We prove that almost surely

$$\int_0^T |\partial_x u(s,\xi_s,\mu_s)\cdot\eta_s|\, ds < +\infty.$$

By Proposition 9.22 and Hölder's inequality, one has

$$\begin{aligned}\mathbb{E}\int_0^{T\wedge T_M} |\partial_x u(s,\xi_s,\mu_s)|\, ds &\leq \int_0^T \int_{B_M} |\partial_x u(s,x,\mu_s)| q(s,x)\, dx\, ds \\ &\leq \int_0^T \|\partial_x u(s,\cdot,\mu_s)\|_{L^{k_1}(B_M)} \|q(s,\cdot)\|_{L^{k_1'}(B_M)}\, ds \\ &\leq J_1(M),\end{aligned}$$

which is finite (see (9.12) in Step 1). We deduce that almost surely, for all $M \geq 1$

$$\int_0^{T\wedge T_M} |\partial_x u(s,\xi_s,\mu_s)|\, ds < \infty.$$

But it is clear that for almost all $\omega \in \Omega$ and for M bigger than some random constant $M(\omega) \geq 1$, $T_M(\omega) = T$. Thus, since η is uniformly bounded, $\int_0^T |\partial_x u(s,\xi_s,\mu_s).\eta_s|\, ds$ is finite almost surely. The other terms in the first line of Itô's formula (9.5) are treated with the same method.

(ii) Martingale term in (9.5). We need to prove that $\int_0^T |\partial_x u(s,\xi_s,\mu_s)|^2\,ds$ is almost surely finite. Reasoning as before, it is a consequence of the fact that J_2 is finite since we have

$$\int_0^T \|\partial_x u(s,\cdot,\mu_s)\|^2_{L^{2k_1}(B_M)}\|q(s,\cdot)\|_{L^{k_1'}(B_M)}\,ds \le J_2(M).$$

Therefore the martingale term in (9.5) is well-defined.

(iii) Terms involving the linear derivative in (9.5). We remark that $\tilde X$ and ξ can be seen as independent processes on the product space $\Omega\times\tilde\Omega$ with $\mathcal{L}(\tilde X_s)=p(s,\cdot)\,dx$ and $\mathcal{L}(\xi_s)=q(s,\cdot)\,dx$ for almost all s. Hölder's inequality gives that

$$\begin{aligned}
&\mathbb{E}\int_0^{T\wedge T_M}\tilde{\mathbb{E}}\left|\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde X_s)\right|\,ds\\
&\le \int_0^T\int_{B_M\times\mathbb{R}^d}\left|\partial_v\frac{\delta u}{\delta m}(s,x,\mu_s)(v)\right| q(s,x)p(s,v)\,dx\,dv\,ds\\
&\le \int_0^T\left\|\partial_v\frac{\delta u}{\delta m}(s,\cdot,\mu_s)(\cdot)\right\|_{L^{k_2}(B_M\times\mathbb{R}^d)}\|q(s,\cdot)\|_{L^{k_2'}(B_M)}\|p(s,\cdot)\|_{L^{k_2'}(\mathbb{R}^d)}\,ds\\
&= J_3(M),
\end{aligned}$$

which was defined in (9.12) and is finite. We deduce as previously that $\int_0^T\tilde{\mathbb{E}}\left|\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde X_s).\tilde{b_s}\right|\,ds$ is almost surely finite. The term involving $\partial_v^2\frac{\delta u}{\delta m}$ is dealt similarly.

Since all the terms in (9.5) are well-defined, it is enough to prove Itô-Krylov's formula for $u(t\wedge T_M,\xi_{t\wedge T_M},\mu_{t\wedge T_M})$ almost surely for all $t\in[0,T]$, and then take the limit $M\to+\infty$ using the continuity of the integrals in Itô-Krylov's formula with respect to t. So we fix $\tau:=T_M$ for $M\ge 1$ and we want to prove the formula up to time τ.

Step 3: Mollification of u.

Let u^n be the function defined by $u^n(t,x,\mu):=u(t,\cdot,\mu*\rho_n)*\rho_n(x)$. It is clearly continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, as u. Since $\partial_t u$ is jointly continuous, it follows from Leibniz's rule that u^n is $\mathcal{C}^1$ with respect to t and that we can differentiate under the integral i.e. for all $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$

$$\partial_t u^n(t,x,\mu)=\partial_t u(t,\cdot,\mu*\rho_n)*\rho_n(x),$$

which is also jointly continuous. As a result of Lemma 9.27 and Proposition 9.28, u^n is $\mathcal{C}^2$ with respect to x and we have

$$\partial_x u^n(t,x,\mu)=\partial_x u(t,\cdot,\mu*\rho_n)*\rho_n(x)\quad\text{and}\quad \partial_x^2 u^n(t,x,\mu)=\partial_x^2 u(t,\cdot,\mu*\rho_n)*\rho_n(x).$$

These two functions are continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ by the dominated convergence theorem and the fact that u is jointly continuous. We define $\tilde{\rho_n}$ by $\tilde{\rho_n}(x,v):=\rho_n(x)\rho_n(v)$ for all $x,v\in\mathbb{R}^d$. It is easy to see that $(\tilde{\rho_n})_n$ is a mollifying sequence on $\mathbb{R}^{2d}$. Next, we claim that for all $(t,x)\in[0,T]\times\mathbb{R}^d$, $u^n(t,x,\cdot)$

has a linear derivative given by

$$\frac{\delta u^n}{\delta m}(t,x,\mu)(v) := \frac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)*\tilde{\rho_n}(x,v). \tag{9.13}$$

This convolution is well-defined as $\frac{\delta u}{\delta m}$ is jointly continuous. To prove (9.13), note first that the bound of Assumption (3) in Definition 9.14 implies that for all $(t,x)\in[0,T]\times\mathbb{R}^d$, $\frac{\delta u^n}{\delta m}(t,x,\mu)(\cdot)$ is at most of quadratic growth, uniformly in μ on each compact set. Since for all $(t,x)\in[0,T]\times\mathbb{R}^d$, $\frac{\delta u}{\delta m}(t,x,\cdot)(\cdot)$ is continuous on $\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$, the dominated convergence theorem proves that $\frac{\delta u^n}{\delta m}(t,x,\cdot)(\cdot)$ is continuous. As explained in Remark 9.2, it is enough to compute, for $\mu,\nu\in\mathcal{P}_2(\mathbb{R}^d)$ and $\lambda\in[0,1]$, the derivative with respect to λ of $u^n(t,x,m_\lambda)$, where $m_\lambda=\lambda\mu+(1-\lambda)\nu$. As recalled in the proof of Theorem 9.7, when (t,x) are fixed

$$\frac{d}{d\lambda}u(t,x,m_\lambda*\rho_n)=\int_{\mathbb{R}^d}\frac{\delta u}{\delta m}(t,x,m_\lambda*\rho_n)*\rho_n(v)\,d(\mu-\nu)(v).$$

Thanks to the bound Assumption (3) in Definition 9.14 for all compact $K\subset\mathbb{R}^d$, one has

$$\sup_{x\in K}\sup_{\lambda\in[0,1]}\left|\frac{d}{d\lambda}u(t,x,m_\lambda*\rho_n)\right|\le C\left(1+\int_{\mathbb{R}^d}|v|^2\,d(\mu+\nu)(v)\right).$$

We can conclude with the help of Leibniz's rule and Fubini's theorem that

$$\frac{d}{d\lambda}u^n(t,x,m_\lambda)=\int_{\mathbb{R}^d}\frac{\delta u}{\delta m}(t,\cdot,m_\lambda*\rho_n)*\tilde{\rho_n}(x,v)\,d(\mu-\nu)(v).$$

It follows from the joint continuity of $\frac{\delta u}{\delta m}$ and Leibniz's rule that $\frac{\delta u^n}{\delta m}$ is $\mathcal{C}^2$ with respect to v and that

$$\begin{cases}\partial_v\dfrac{\delta u^n}{\delta m}(t,x,\mu)(v)= & \dfrac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)*\partial_v\tilde{\rho_n}(x,v)=\partial_v\dfrac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)*\tilde{\rho_n}(x,v)\\ \partial_v^2\dfrac{\delta u^n}{\delta m}(t,x,\mu)(v)= & \dfrac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)*\partial_v^2\tilde{\rho_n}(x,v)=\partial_v^2\dfrac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)*\tilde{\rho_n}(x,v).\end{cases}$$

Note that $\partial_v\frac{\delta u^n}{\delta m}$ and $\partial_v^2\frac{\delta u^n}{\delta m}$ are continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$ thanks to the dominated convergence theorem and the joint continuity of $\frac{\delta u}{\delta m}$. Moreover for all compact $\mathcal{K}\subset\mathcal{P}_2(\mathbb{R}^d)$ and for all $M>0$

$$\sup_{t\in[0,T]}\sup_{\mu\in\mathcal{K}}\sup_{|x|\le M}\sup_{v\in\mathbb{R}^d}\left|\partial_v\frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right|+\left|\partial_v^2\frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right|<+\infty. \tag{9.14}$$

Indeed, Hölder's inequality ensures that

$$\begin{aligned}&\sup_{|x|\le M}\sup_{v\in\mathbb{R}^d}\left|\partial_v\frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right|+\left|\partial_v^2\frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right|\\&\le\left[\left\|\partial_v\frac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)\right\|_{L^{k_2}(B_{M+1}\times\mathbb{R}^d)}+\left\|\partial_v^2\frac{\delta u}{\delta m}(t,\cdot,\mu*\rho_n)(\cdot)\right\|_{L^{k_2}(B_{M+1}\times\mathbb{R}^d)}\right]\|\tilde{\rho_n}\|_{L^{k_2'}(\mathbb{R}^{2d})},\end{aligned}$$

the ball B_{M+1} coming from the fact that the support of $\tilde{\rho_n}$ is included in B_1. Since $\mathcal{K}*\rho_n$ is compact in $\mathcal{P}_2(\mathbb{R}^d)$ and included in $\mathscr{P}(\mathbb{R}^d)$, Assumption (5) in Definition 9.14 ensures that there exists $C>0$ such that for all $\mu \in \mathcal{K}$

$$\sup_{t\in[0,T]}\sup_{|x|\le M}\sup_{v\in\mathbb{R}^d}\left|\partial_v \frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right| + \left|\partial_v^2 \frac{\delta u^n}{\delta m}(t,x,\mu)(v)\right| \le C\left(1+\left\|\frac{d\mu * \rho_n}{dx}\right\|^{\alpha_2}_{L^{k_2'}(\mathbb{R}^d)}\right)\|\tilde{\rho_n}\|_{L^{k_2'}(\mathbb{R}^{2d})}.$$

But we know that $\frac{d\mu*\rho_n}{dx}(x) = \int_{\mathbb{R}^d}\rho_n(x-y)\,d\mu(y)$. We conclude with Jensen's inequality that

$$\left\|\frac{d\mu * \rho_n}{dx}\right\|^{\alpha_2}_{L^{k_2'}(\mathbb{R}^d)} \le \|\rho_n\|^{\alpha_2}_{L^{k_2'}(\mathbb{R}^d)}.$$

This proves (9.14).

Step 4: Itô's formula (9.5) for u^n when the coefficients b and σ are continuous.

We claim that $(t,x)\mapsto U^n(t,x) := u^n(t,x,\mu_t) \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$. The regularity with respect to x is clear with the preceding properties on u^n. Let us thus focus on the regularity with respect to the time variable. For $(t,x)\in[0,T]\times\mathbb{R}^d$ fixed, the regularity assumption on u with respect to t and the standard Itô formula for a flow of measures applied to $u^n(t,x,\cdot)$ (see Theorem 5.99 in Chapter 5 of [CD18a]) ensure that we have for $h\in\mathbb{R}$ satisfying $t+h\ge 0$

$$\begin{aligned} u^n(t+h,x,\mu_{t+h}) - u^n(t,x,\mu_t) &= u^n(t+h,x,\mu_{t+h}) - u^n(t,x,\mu_{t+h}) + u^n(t,x,\mu_{t+h}) - u^n(t,x,\mu_t)\\ &= \int_t^{t+h}\partial_t u^n(s,x,\mu_{t+h})\,ds + \int_t^{t+h}\mathbb{E}\left(\partial_v\frac{\delta u^n}{\delta m}(t,x,\mu_s)(X_s)\cdot b_s\right)ds\\ &\quad+\frac{1}{2}\int_t^{t+h}\mathbb{E}\left(\partial_v^2\frac{\delta u^n}{\delta m}(t,x,\mu_s)(X_s)\cdot a_s\right)ds. \end{aligned} \tag{9.15}$$

The function $(s,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\mapsto \partial_t u^n(s,x,\mu)$ is continuous so

$$\frac{1}{h}\int_t^{t+h}\partial_t u^n(s,x,\mu_{t+h})\,ds \underset{h\to 0}{\longrightarrow} \partial_t u^n(t,x,\mu_t).$$

The two other terms in (9.15) can be dealt similarly. Indeed, the dominated convergence theorem justified by (9.14) ensures that the functions $(s,x)\in[0,T]\times\mathbb{R}^d\mapsto\mathbb{E}\left(\partial_v\frac{\delta u^n}{\delta m}(s,x,\mu_s)(X_s)\cdot b_s\right)$ and $(s,x)\in[0,T]\times\mathbb{R}^d\mapsto\mathbb{E}\left(\partial_v^2\frac{\delta u^n}{\delta m}(s,x,\mu_s)(X_s)\cdot a_s\right)$ are continuous. Then, it follows that $U^n\in\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^d)$ and that for all $(t,x)\in[0,T]\times\mathbb{R}^d$

$$\partial_t U^n(t,x) = \partial_t u^n(t,x,\mu_t) + \mathbb{E}\left(\partial_v\frac{\delta u^n}{\delta m}(t,x,\mu_t)(X_t)\cdot b_t\right) + \frac{1}{2}\mathbb{E}\left(\partial_v^2\frac{\delta u^n}{\delta m}(t,x,\mu_t)(X_t)\cdot a_t\right).$$

We can now apply the classical Itô formula for U^n and ξ, up to the random time τ defined at the end of

Step 2, to obtain that almost surely, for all $t \in [0,T]$

$$u^n(t\wedge\tau, \xi_{t\wedge\tau}, \mu_{t\wedge\tau}) = u^n(0,\xi_0,\mu_0) + \int_0^{t\wedge\tau} \partial_t u^n(s,\xi_s,\mu_s) + \partial_x u^n(s,\xi_s,\mu_s)\cdot\eta_s + \frac{1}{2}\partial_x^2 u^n(s,\xi_s,\mu_s)\cdot\gamma_s\gamma_s^*\,ds$$
$$+ \int_0^{t\wedge\tau} \tilde{\mathbb{E}}\left(\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}\right)ds + \frac{1}{2}\int_0^{t\wedge\tau}\tilde{\mathbb{E}}\left(\partial_v^2\frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{a_s}\right)ds \tag{9.16}$$
$$+ \int_0^{t\wedge\tau} \partial_x u^n(s,\xi_s,\mu_s)\cdot(\gamma_s\,dB_s).$$

Note that (9.16) does not require Assumptions **(A)** and **(B)** on the Itô process X. These assumptions will only be used in Step 6.

Step 5: Removing the continuity hypothesis on the coefficients b and σ.

We consider $(b^m)_m$ and $(\sigma^m)_m$ two sequences of continuous and progressively measurable processes such that

$$\mathbb{E}\int_0^T |b_s^n - b_s|^2 + |\sigma_s^n - \sigma_s|^4\,ds \to 0.$$

We set, for $t\le T$, $X_t^m := X_0 + \int_0^t b_s^m\,ds + \int_0^t \sigma_s^m\,dB_s$, and μ_t^m the law of X_t^m. Owing to Step 4, Itô's formula (9.16) holds true for X^m and ξ. Now, we aim at taking the limit $m\to+\infty$ in (9.16). Note that the set $\mathcal{K} := \{\mu_s^m,\, s\le T,\, m\ge 1\}\cup\{\mu_s,\, s\le T\}$ is compact in $\mathcal{P}_2(\mathbb{R}^d)$. Indeed, using Jensen's inequality and the Burkholder-Davis-Gundy (BDG) inequalities, it is clear that $\mathbb{E}\sup_{t\le T}|X_t^m - X_t|^2 \to 0$, thus $\sup_{t\le T} W_2(\mu_t^m,\mu_t)\to 0$. We deduce that almost surely, for all $t\in[0,T]$

$$u^n(t,\xi_t,\mu_t^m) \underset{m\to+\infty}{\longrightarrow} u^n(t,\xi_t,\mu_t).$$

Now, we take the limit $m\to+\infty$ in the integrals in Itô's formula (9.16).

(i) Martingale term in (9.16). Using BDG's inequality, there exists $C>0$ such that

$$\mathbb{E}\sup_{t\le T}\left|\int_0^{t\wedge\tau}(\partial_x u^n(s,\xi_s,\mu_s^m) - \partial_x u^n(s,\xi_s,\mu_s))\cdot(\gamma_s\,dB_s)\right|^2$$
$$\le C\mathbb{E}\int_0^{T\wedge\tau}|\partial_x u^n(s,\xi_s,\mu_s^m) - \partial_x u^n(s,\xi_s,\mu_s)|^2|\gamma_s|^2\,ds$$
$$\le C\mathbb{E}\int_0^{T}|\partial_x u^n(s,\xi_s,\mu_s^m) - \partial_x u^n(s,\xi_s,\mu_s)|^2\mathbf{1}_{B_M}(\xi_s)|\gamma_s|^2\,ds.$$

The dominated convergence theorem can be applied since γ is bounded and $\partial_x u^n$ is jointly continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$. It shows that, up to an extraction, almost surely

$$\forall t\le T,\ \int_0^{t\wedge\tau}\partial_x u^n(s,\xi_s,\mu_s^m)\cdot(\gamma_s\,dB_s) \underset{m\to+\infty}{\longrightarrow} \int_0^{t\wedge\tau}\partial_x u^n(s,\xi_s,\mu_s)\cdot(\gamma_s\,dB_s).$$

(ii) Terms involving the linear derivative in (9.16)**.** Let us write

$$\begin{aligned}
&\left|\int_0^{t\wedge\tau} \tilde{\mathbb{E}}\left(\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m})\cdot \tilde{b_s^m}\right) ds - \int_0^{t\wedge\tau} \tilde{\mathbb{E}}\left(\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot \tilde{b_s}\right) ds\right| \\
&\leq \int_0^{T\wedge\tau} \tilde{\mathbb{E}}\left|\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m})\right| |\tilde{b_s^m}-\tilde{b_s}|\, ds \\
&\quad + \int_0^{T\wedge\tau} \tilde{\mathbb{E}}\left|\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m}) - \partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\right| |\tilde{b_s}|\, ds \\
&=: I_1 + I_2
\end{aligned}$$

Cauchy-Schwarz's inequality ensures that

$$I_1 \leq \left(\int_0^{T\wedge\tau} \tilde{\mathbb{E}}\left|\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m})\right|^2 ds\right)^{1/2} \left(\int_0^{T} \tilde{\mathbb{E}}|\tilde{b_s^m}-\tilde{b_s}|^2\, ds\right)^{1/2}.$$

We conclude that I_1 converges to 0 thanks to the bound (9.14) proved in Step 3 and since ξ is bounded by M on the set $\{\tau > 0\}$. To show that $I_2 \to 0$, we use the fact that b is bounded by K to get

$$I_2 \leq K \int_0^{T\wedge\tau} \tilde{\mathbb{E}}\left|\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m}) - \partial_v \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\right| ds.$$

The continuity of $\partial_v \frac{\delta u^n}{\delta m}$ and the convergence in L^2 of $(\tilde{X_s^m})_m$ to $\tilde{X}_s$ ensure that for all $\omega \in \Omega$, $\left|\partial_v \frac{\delta u^n}{\delta m}(s,\xi_s(\omega),\mu_s^m)(\tilde{X_s^m}) - \partial_v \frac{\delta u^n}{\delta m}(s,\xi_s(\omega),\mu_s)(\tilde{X}_s)\right|$ converges in probability on $\tilde{\Omega}$ to 0 as m goes to infinity. Using a uniform integrability argument coming from (9.14), we deduce that I_2 converges to 0. Following the same strategy, one has for all $t \in [0,T]$

$$\int_0^{t\wedge\tau} \tilde{\mathbb{E}}\left(\partial_v^2 \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s^m)(\tilde{X_s^m})\cdot \tilde{a_s^m}\right) ds \underset{m\to+\infty}{\longrightarrow} \int_0^{t\wedge\tau} \tilde{\mathbb{E}}\left(\partial_v^2 \frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot \tilde{a_s}\right) ds.$$

(iii) Terms involving standard derivatives in (9.16)**.** It follows from the dominated convergence theorem that almost surely, for all $t \leq T$

$$\begin{aligned}
&\int_0^{t\wedge\tau} (\partial_t u^n(s,\xi_s,\mu_s^m) + \partial_x u^n(s,\xi_s,\mu_s^m)\cdot \eta_s)\, ds + \frac{1}{2}\int_0^{t\wedge\tau} \partial_x^2 u^n(s,\xi_s,\mu_s^m)\cdot \gamma_s\gamma_s^*\, ds \\
&\underset{m\to+\infty}{\longrightarrow} \int_0^{t\wedge\tau} (\partial_t u^n(s,\xi_s,\mu_s) + \partial_x u^n(s,\xi_s,\mu_s)\cdot \eta_s)\, ds + \frac{1}{2}\int_0^{t\wedge\tau} \partial_x^2 u^n(s,\xi_s,\mu_s)\cdot \gamma_s\gamma_s^*\, ds.
\end{aligned}$$

Indeed the functions $\partial_t u^n$, $\partial_x u^n$ and $\partial_x^2 u^n$ are jointly continuous on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and thus uniformly bounded on $[0,T]\times B_M\times\{\mu_s^m,\ s\in[0,T],\ m\geq 1\}$. Moreover, η and γ are also uniformly bounded.

This concludes Step 5.

Step 6: Letting $n \to +\infty$.

From Step 5, we deduce that Itô's formula (9.16) in Step 4 holds for u^n up to time τ. To conclude the proof, we need to take the limit $n \to +\infty$ in each term of (9.16). Then it remains to remove the stopping time τ as explained at the end of Step 2 (i.e. letting $\tau \to T$). The continuity of u ensures that almost surely, for all $t \leq T$, $u^n(t,\xi_t,\mu_t) \to u(t,\xi_t,\mu_t)$. We now focus on the integrals in Itô's formula (9.16).

(i) Martingale term in (9.16). Thanks to BDG's inequality, Hölder's inequality, and the boundedness of γ, we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\left|\int_0^{t\wedge\tau}(\partial_x u^n(s,\xi_s,\mu_s)-\partial_x u(s,\xi_s,\mu_s))\cdot(\gamma_s\,dB_s)\right|^2\\
&\le C\mathbb{E}\int_0^T|\partial_x u(s,\cdot,\mu_s*\rho_n)*\rho_n(\xi_s)-\partial_x u(s,\xi_s,\mu_s)|^2\mathbf{1}_{B_M}(\xi_s)\,ds\\
&=C\int_0^T\int_{B_M}|\partial_x u(s,\cdot,\mu_s*\rho_n)*\rho_n(x)-\partial_x u(s,x,\mu_s)|^2 q(s,x)\,dx\,ds\\
&\le C\int_0^T\|\partial_x u(s,\cdot,\mu_s*\rho_n)*\rho_n-\partial_x u(s,\cdot,\mu_s)\|^2_{L^{2k_1}(B_M)}\|q(s,\cdot)\|_{L^{k_1'}(B_M)}\,ds\\
&\le C\int_0^T\|\partial_x u(s,\cdot,\mu_s*\rho_n)*\rho_n-\partial_x u(s,\cdot,\mu_s)*\rho_n\|^2_{L^{2k_1}(B_M)}\|q(s,\cdot)\|_{L^{k_1'}(B_M)}\,ds\\
&\quad+C\int_0^T\|\partial_x u(s,\cdot,\mu_s)*\rho_n-\partial_x u(s,\cdot,\mu_s)\|^2_{L^{2k_1}(B_M)}\|q(s,\cdot)\|_{L^{k_1'}(B_M)}\,ds\\
&=: I_1+I_2.
\end{aligned}$$

We prove that I_1 and I_2 converge to 0. First note that, due to the convolution inequality $L^r * L^1$, we have for $f\in L^r_{\text{loc}}(\mathbb{R}^d)$ and for all $R>0$, $\|f*\rho_n\|_{L^r(B_R)}\le\|f\|_{L^r(B_{R+1})}$. The control on B_{R+1} follows from the fact that the support of each ρ_n is included in B_1. Hence

$$I_1\le C\int_0^T\|\partial_x u(s,\cdot,\mu_s*\rho_n)-\partial_x u(s,\cdot,\mu_s)\|^2_{L^{2k_1}(B_{M+1})}\|q(s,\cdot)\|_{L^{k_1'}(B_{M+1})}\,ds\ =:\tilde{I}_1.$$

As a consequence of Sobolev embedding theorem, for all t, the function

$$\mu\in(\mathscr{P}(\mathbb{R}^d),d_{\mathscr{P}})\mapsto\partial_x u(t,\cdot,\mu)\in L^\infty(B_{M+1})$$

is continuous. Since $\mu_s\in\mathscr{P}(\mathbb{R}^d)$ for almost all s and thanks to Assumption (2) in Definition 9.3, we deduce that the integrand in $\tilde{I}_1$ converges to 0 for almost all s. It follows from the dominated convergence theorem (see (9.12) in Step 1) that $\tilde{I}_1$ converges to 0, as well as I_1. We now focus on I_2. The integrand in I_2 converges to 0 for almost all s because $\partial_x u(s,\cdot,\mu_s)\in L^{2k_1}(B_M)$. We conclude with the dominated convergence theorem as previously. This shows that, up to an extraction, almost surely

$$\sup_{t\le T}\left|\int_0^{t\wedge\tau}\partial_x u^n(s,\xi_s,\mu_s)\cdot(\gamma_s\,dB_s)-\int_0^{t\wedge\tau}\partial_x u(s,\xi_s,\mu_s)\cdot(\gamma_s\,dB_s)\right|\to 0.$$

(ii) Terms involving the linear derivative in (9.16). Following the same strategy, we obtain using

Hölder's inequality

$$
\begin{aligned}
&\mathbb{E}\sup_{t\le T}\left|\int_0^{t\wedge\tau}\tilde{\mathbb{E}}\left(\partial_v\frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}\right)ds-\int_0^{t\wedge\tau}\tilde{\mathbb{E}}\left(\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}\right)ds\right|\\
&\le \mathbb{E}\tilde{\mathbb{E}}\int_0^{T\wedge\tau}\left|\partial_v\frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s).\tilde{b_s}-\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s).\tilde{b_s}\right|ds\\
&\le \mathbb{E}\tilde{\mathbb{E}}\int_0^{T}\left|\partial_v\frac{\delta u^n}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}-\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}\right|\mathbf{1}_{B_M}(\xi_s)\,ds\\
&\le C\int_0^T\int_{B_M\times\mathbb{R}^d}\left|\partial_v\frac{\delta u^n}{\delta m}(s,x,\mu_s)(v)-\partial_v\frac{\delta u}{\delta m}(s,x,\mu_s)(v)\right|q(s,x)p(s,v)\,dx\,dv\,ds\\
&\le C\int_0^T\left\|\partial_v\frac{\delta u}{\delta m}(s,\cdot,\mu_s*\rho_n)(\cdot)*\tilde{\rho_n}-\partial_v\frac{\delta u}{\delta m}(s,\cdot,\mu_s)(\cdot)\right\|_{L^{k_2}(B_M\times\mathbb{R}^d)}\|q(s,\cdot)\|_{L^{k_2'}(B_M)}\|p(s,\cdot)\|_{L^{k_2'}(\mathbb{R}^d)}ds.
\end{aligned}
$$

The dominated convergence theorem justified by Assumption (4) in Definition 9.14 and (9.12) in Step 1 ensures that this term converges to 0. The same argument holds true for the term involving $\partial_v^2\frac{\delta u}{\delta m}$.

(iii) Terms involving standard derivatives in (9.16)**.** The convergence of the term involving $\partial_t u^n$ in (9.16) follows from the continuity of $\partial_t u$ on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and the dominated convergence theorem since almost surely on the set $\{\tau>0\}$

$$
\sup_{s\in[0,T]}\sup_{n\ge1}|\partial_t u^n(s,\xi_s,\mu_s)|\le\sup_{s\in[0,T]}\sup_{n\ge1}\sup_{|x|\le M+1}|\partial_t u(s,x,\mu_s*\rho_n)|<+\infty.
$$

For the spatial derivatives, Hölder's inequality ensures that

$$
\begin{aligned}
&\mathbb{E}\sup_{t\le T}\left|\int_0^{t\wedge\tau}\partial_x u^n(s,\xi_s,\mu_s)\cdot\eta_s\,ds-\int_0^{t\wedge\tau}\partial_x u(s,\xi_s,\mu_s)\cdot\eta_s\,ds\right|\\
&\le C\int_0^T\|\partial_x u(s,\cdot,\mu_s*\rho_n)*\rho_n-\partial_x u(s,\cdot,\mu_s)\|_{L^{k_1}(B_M)}\|q(s,\cdot)\|_{L^{k_1'}(B_M)}ds.
\end{aligned}
$$

The right-hand side term converges to 0 with same reasoning as before. This shows that, up to an extraction, one has almost surely

$$
\sup_{t\le T}\left|\int_0^{t\wedge\tau}\partial_x u^n(s,\xi_s,\mu_s)\cdot\eta_s\,ds-\int_0^{t\wedge\tau}\partial_x u(s,\xi_s,\mu_s)\cdot\eta_s\,ds\right|\underset{n\to+\infty}{\longrightarrow}0.
$$

The term involving $\partial_x^2 u$ in (9.16) is dealt similarly.

Taking the limit $n\to+\infty$ in (9.16), up to an extraction, we conclude that almost surely, for all $t\in[0,T]$

$$
\begin{aligned}
u(t\wedge\tau,\xi_{t\wedge\tau},\mu_{t\wedge\tau})&=u(0,\xi_0,\mu_0)\\
&+\int_0^{t\wedge\tau}(\partial_t u(s,\xi_s,\mu_s)+\partial_x u(s,\xi_s,\mu_s)\cdot\eta_s)\,ds+\frac{1}{2}\int_0^{t\wedge\tau}\partial_x^2 u(s,\xi_s,\mu_s)\cdot\gamma_s\gamma_s^*\,ds\\
&+\int_0^{t\wedge\tau}\tilde{\mathbb{E}}\left(\partial_v\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{b_s}\right)ds+\frac{1}{2}\int_0^{t\wedge\tau}\tilde{\mathbb{E}}\left(\partial_v^2\frac{\delta u}{\delta m}(s,\xi_s,\mu_s)(\tilde{X}_s)\cdot\tilde{a_s}\right)ds\\
&+\int_0^{t\wedge\tau}\partial_x u(s,\xi_s,\mu_s)\cdot(\gamma_s\,dB_s).
\end{aligned}
$$

This ends the proof as explained in Step 2. □

9.7 Appendix: Proof of the examples

9.7.1 Proof of Example 9.4

(1) It follows from the contraction inequality in Lemma 9.30 and Corollary 9.31.

(2) To prove (**H1**), we fix $n \geq 1$ and $\mu_j \xrightarrow{W_2} \mu \in \mathcal{P}_2(\mathbb{R}^d)$. For $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, the density of $\nu * \rho_n$ is given by

$$x \in \mathbb{R}^d \mapsto \rho_n * \nu(x) = \int_{\mathbb{R}^d} \rho_n(x-y)\, d\nu(y).$$

Hence,

$$d_k(\mu_j * \rho_n, \mu * \rho_n) = \left\| \int_{\mathbb{R}^d} \rho_n(\cdot - y)\, d\mu_j(y) - \int_{\mathbb{R}^d} \rho_n(\cdot - y)\, d\mu(y) \right\|_{L^{k'}(\mathbb{R}^d)}.$$

Using Lemma 9.29, we conclude that $d_k(\mu_j * \rho_n, \mu * \rho_n) \underset{j \to +\infty}{\longrightarrow} 0$. For **(H2)**, let $\mu \in \mathscr{P}(\mathbb{R}^d)$ and denote by $f \in L^{k'}(\mathbb{R}^d)$ the density of μ. For $n \geq 1$, we have

$$\frac{d\mu * \rho_n}{dx} = f * \rho_n \xrightarrow{L^{k'}} f,$$

owing to Lemma 9.27.

□

9.7.2 Proof of Example 9.10

Let us give the detailed proof in the bilinear case $N = 2$. It is standard (see Example 4 page 389 in Chapter 5 of [CD18a]) that u has a linear derivative given by

$$\frac{\delta u}{\delta m}(\mu)(v) = \int_{\mathbb{R}^d} g(v,y)\, d\mu(y) + \int_{\mathbb{R}^d} g(y,v)\, d\mu(y).$$

We will only treat the first term since the other one can be dealt similarly.

Computation of the distributional derivatives and continuity: Let $\mu \in \mathscr{P}(\mathbb{R}^d)$ and $f \in L^{(d+1)'}(\mathbb{R}^d)$ be its density. By interpolation, we know that $f \in L^{r'}(\mathbb{R}^d)$ for all $r \geq d+1$. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ and $i \in \{1, \dots, d\}$. Using Fubini's theorem, justified by the quadratic growth of g and the fact that $f\, dx \in \mathcal{P}_2(\mathbb{R}^d)$, we have

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g(x,y) f(y)\, dy \right) \partial_{v_i}\varphi(v)\, dv = \int_{\mathbb{R}^d \times \mathbb{R}^d} g(v,y) f(y) \partial_{v_i}\varphi(v)\, dy\, dv.$$

Let us define $f_n(x) = \frac{1}{\mu(B_n)}(f \mathbf{1}_{B_n}) * \rho_n(x)$, for n large enough to have $\mu(B_n) > 0$. The function f_n is a probability density which is in $\mathcal{C}_c^\infty(\mathbb{R}^d)$. It easily follows from Lemma 9.26, Lemma 9.27 and the dominated convergence theorem that

$$f_n \xrightarrow{L^{k'}} f \quad \text{and} \quad f_n \xrightarrow{W_2} f. \tag{9.17}$$

For a fixed $n \geq 1$, we have by definition of the distributional derivative

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} g(v,y) f_n(y) \partial_{v_i} \varphi(v)\, dy\, dv = -\int_{\mathbb{R}^d \times \mathbb{R}^d} \partial_{v_i} g(v,y) f_n(y) \varphi(v)\, dy\, dv. \tag{9.18}$$

Our aim is to take the limit $n \to +\infty$ in both side of the previous equality. Using Fubini's theorem, the left-hand side term is equal to

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g(v,y) \partial_{v_i} \varphi(v)\, dv \right) f_n(y)\, dy.$$

Moreover, it converges to

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} g(v,y) \partial_{v_i} \varphi(v) f(y)\, dy\, dv.$$

Indeed, $f_n \xrightarrow{W_2} f$ and the function $y \mapsto \int_{\mathbb{R}^d} g(v,y) \partial_{v_i} \varphi(v)\, dv$ is continuous and at most of quadratic growth. For the right-hand side term, we prove that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \partial_{v_i} g(v,y) f_n(y) \varphi(v)\, dy\, dv \to \int_{\mathbb{R}^d \times \mathbb{R}^d} \partial_{v_i} g(v,y) f(y) \varphi(v)\, dy\, dv.$$

Note that the limit is well-defined using Hölder's inequality

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |\partial_{v_i} g(v,y) f(y) \varphi(v)|\, dy\, dv \leq \|f\|_{L^{k'}(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi(v) \|\partial_{v_i} g(v,\cdot)\|_{L^k(\mathbb{R}^d)}\, dv.$$

The right-hand side term is finite because $v \mapsto \|\partial_{v_i} g(v,\cdot)\|_{L^k(\mathbb{R}^d)} \in L^k(\mathbb{R}^d)$. The same inequality shows that

$$\left| \int_{\mathbb{R}^d \times \mathbb{R}^d} \partial_{v_i} g(v,y) (f_n(y) - f(y)) \varphi(v)\, dy\, dv \right| \leq \|f_n - f\|_{L^{k'}(\mathbb{R}^d)} \int_{\mathbb{R}^d} \varphi(v) \|\partial_{v_i} g(v,\cdot)\|_{L^k(\mathbb{R}^d)}\, dv \underset{n \to +\infty}{\longrightarrow} 0,$$

thanks to (9.17). Taking the limit $n \to +\infty$ in (9.18), we deduce that:

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} g(v,y) f(y) \partial_{v_i} \varphi(v)\, dy\, dv = -\int_{\mathbb{R}^d \times \mathbb{R}^d} \partial_{v_i} g(v,y) f(y) \varphi(v)\, dy\, dv.$$

Hence, the distributional derivative of $v \mapsto \int_{\mathbb{R}^d} g(v,y) f(y)\, dy$ is given by the function

$$v \mapsto \int_{\mathbb{R}^d} \partial_v g(v,y) f(y)\, dy.$$

Moreover, it belongs to $L^k(\mathbb{R}^d)$ because applying Hölder's inequality, one has

$$\begin{aligned} \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \partial_v g(v,y) f(y)\, dy \right|^k dv &\leq \int_{\mathbb{R}^d} \|\partial_v g(v,\cdot)\|^k_{L^k(\mathbb{R}^d)} \|f\|^k_{L^{k'}(\mathbb{R}^d)}\, dv \\ &= \|\partial_v g\|^k_{L^k(\mathbb{R}^d \times \mathbb{R}^d)} \|f\|^k_{L^{k'}(\mathbb{R}^d)}. \end{aligned}$$

Note that this inequality and the linearity in f justify that $\mu \in (\mathscr{P}(\mathbb{R}^d), d_k) \mapsto \int_{\mathbb{R}^d} \partial_v g(\cdot, y)\, d\mu(y) \in L^k(\mathbb{R}^d)$ is continuous with

$$\left\| \partial_v \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} \leq \left\| \frac{d\mu}{dx} \right\|_{L^{k'}(\mathbb{R}^d)} \|\nabla g\|_{L^k(\mathbb{R}^d \times \mathbb{R}^d)}. \tag{9.19}$$

Following the same lines, we show that the distributional derivative of order 2 of $\frac{\delta u}{\delta m}(\mu)$, for $\mu \in \mathscr{P}(\mathbb{R}^d)$, is given by the $\mathbb{R}^{d\times d}$-valued function

$$v \mapsto \int_{\mathbb{R}^d} \partial_v^2 g(v,y)\, d\mu(y) + \int_{\mathbb{R}^d} \partial_y^2 g(y,v)\, d\mu(y).$$

It is also a continuous function from $(\mathscr{P}(\mathbb{R}^d), d_k)$ into $L^k(\mathbb{R}^d)$. Indeed, as previously, we obtain:

$$\left\| \partial_v^2 \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} \leq \left\| \frac{d\mu}{dx} \right\|_{L^{k'}(\mathbb{R}^d)} \|\nabla^2 g\|_{L^k(\mathbb{R}^d\times\mathbb{R}^d)}. \tag{9.20}$$

Growth property: Using the inequalities (9.19) and (9.20) of the previous step, one has for all $\mu \in \mathscr{P}(\mathbb{R}^d)$

$$\left\| \partial_v \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} + \left\| \partial_v^2 \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} \leq \left\| \frac{d\mu}{dx} \right\|_{L^{k'}(\mathbb{R}^d)} \left[\|\nabla g\|_{L^k(\mathbb{R}^d\times\mathbb{R}^d)} + \|\nabla^2 g\|_{L^k(\mathbb{R}^d\times\mathbb{R}^d)} \right].$$

The second point in Definition 9.5 is thus satisfied with $\alpha = 1$ because we have supposed that $k \geq 2d$.

In the general case $N \geq 2$, one can show following the same lines that u admits a linear derivative and that for all $\mu \in \mathscr{P}(\mathbb{R}^d)$, its distributional derivative is given for all $v \in \mathbb{R}^d$ by

$$\partial_v \frac{\delta u}{\delta m}(\mu)(v) = \sum_{j=1}^N \int_{(\mathbb{R}^d)^{N-1}} \partial_{x_j} g(x_1, \dots, x_{j-1}, v, x_{j+1}, \dots, x_N)\, d\mu(x_1) \dots d\mu(x_{j-1})\, d\mu(x_{j+1}) \dots d\mu(x_N).$$

Denoting by f the density of μ and using Hölder's inequality, we obtain as previously that for all $j \in \{1, \dots, N\}$

$$\begin{aligned}&\int_{\mathbb{R}^d} \left| \int_{(\mathbb{R}^d)^{N-1}} \partial_{x_j} g(x_1, \dots, x_{j-1}, v, x_{j+1}, \dots, x_N)\, d\mu(x_1) \dots d\mu(x_{j-1})\, d\mu(x_{j+1}) \dots d\mu(x_N) \right|^k dv \\ &= \|\partial_{x_j} g\|^k_{L^k((\mathbb{R}^d)^N)} \|f\|^{(N-1)k}_{L^{k'}(\mathbb{R}^d)}.\end{aligned}$$

We easily show that $\mu \in (\mathscr{P}(\mathbb{R}^d), d_k) \mapsto \partial_v \frac{\delta u}{\delta m}(\mu) \in L^k(\mathbb{R}^d)$ is continuous and the same properties hold for the distributional derivative of order two. We deduce that $\mu \in \mathscr{P}(\mathbb{R}^d)$

$$\left\| \partial_v \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} + \left\| \partial_v^2 \frac{\delta u}{\delta m}(\mu) \right\|_{L^k(\mathbb{R}^d)} \leq \left\| \frac{d\mu}{dx} \right\|^{N-1}_{L^{k'}(\mathbb{R}^d)} \left[\|\nabla g\|_{L^k((\mathbb{R}^d)^N)} + \|\nabla^2 g\|_{L^k((\mathbb{R}^d)^N)} \right].$$

The second point in Definition 9.5 is thus satisfied with $\alpha = N-1$ because we have supposed that $k \geq Nd$.

□

9.7.3 Proof of Example 9.12

Note that $f * \mu$ and $u(\mu)$ are well-defined for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Indeed, it follows from Sobolev embedding theorem (see Corollary 9.14 in [Bre10]) that $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$ and $\partial_x f \in (L^\infty(\mathbb{R}^d))^d$. Thus f is at most of linear growth. Since f is continuous and at most of linear growth, it is easy to see that u has a linear derivative given by

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\, \forall v \in \mathbb{R}^d,\, \frac{\delta u}{\delta m}(\mu)(v) = f * \mu(v) + \tilde{f} * \mu(v),$$

where $\tilde{f}(x) = f(-x)$ (see Example 2 page 386 in Chapter 5 of [CD18a]). An easy computation based on Fubini's theorem shows that the distributional derivatives of order 1 and 2 of $\frac{\delta u}{\delta m}(\mu)$ are given by

$$\forall i, j \in \{1, \dots, d\}, \begin{cases} \partial_{v_i} \frac{\delta u}{\delta m}(\mu) & = \partial_{v_i} f * \mu + \partial_{v_i} \tilde{f} * \mu \\ \partial_{v_i v_j} \frac{\delta u}{\delta m}(\mu) & = \partial_{v_i v_j} f * \mu + \partial_{v_i v_j} \tilde{f} * \mu, \end{cases}$$

as elements of $L^{k+1}(\mathbb{R}^d)$. These functions are continuous with respect to $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ owing to Lemma 9.29. It remains to apply the first point in Remark 9.6 to conclude.

□

9.7.4 Proof of Example 9.13

The function u is well-defined and continuous because $\nabla g \in L^\infty(\mathbb{R}^d)$ and is continuous thanks to Sobolev embedding theorem. Thus g is at most of linear growth. It follows from the continuity of $\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} g\, d\mu$ that the function u admits a linear derivative given by

$$\forall (\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d,\, \frac{\delta u}{\delta m}(\mu)(v) = g(v) F'\left(\int_{\mathbb{R}^d} g\, d\mu\right).$$

We thus have in the sense of distributions

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\, \forall v \in \mathbb{R}^d,\, \partial_v \frac{\delta u}{\delta m}(\mu)(v) = \nabla g(v) F'\left(\int_{\mathbb{R}^d} g\, d\mu\right).$$

Moreover, the function

$$\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_v \frac{\delta u}{\delta m}(\mu)(\cdot) \in L^k(\mathbb{R}^d)$$

is continuous because $F \in \mathcal{C}^1(\mathbb{R}; \mathbb{R})$ and $\nabla g \in L^k(\mathbb{R}^d)$. The same reasoning proves that

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\, \forall v \in \mathbb{R}^d,\, \partial_v^2 \frac{\delta u}{\delta m}(\mu)(v) = \nabla^2 g(v) F'\left(\int_{\mathbb{R}^d} g\, d\mu\right),$$

and that the function

$$\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_v^2 \frac{\delta u}{\delta m}(\mu)(\cdot) \in L^k(\mathbb{R}^d)$$

is continuous. We conclude that $u \in \mathcal{W}_1(\mathbb{R}^d)$ with Remark 9.6.

□

9.7.5 Proof of Example 9.17

The function u is well-defined and continuous. Indeed, Sobolev embedding theorem implies that $\nabla g \in L^\infty(\mathbb{R}^{2d})$ and is continuous. Hence g is at most of linear growth. Following the same method as in the proof of Example 9.10, we obtain that

$$\forall \mu \in \mathscr{P}(\mathbb{R}^d),\, \partial_x u(\cdot, \mu) = \int_{\mathbb{R}^d} \partial_x g(\cdot, y)\, d\mu(y).$$

Moreover

$$\forall \mu \in \mathscr{P}(\mathbb{R}^d),\ \|\partial_x u(\cdot,\mu)\|_{L^k(\mathbb{R}^d)} \le \|\nabla g\|_{L^k(\mathbb{R}^{2d})} \left\| \frac{d\mu}{dx} \right\|_{L^{k'}(\mathbb{R}^d)}.$$

This yields the continuity of the function

$$\mu \in (\mathscr{P}(\mathbb{R}^d), d_k) \mapsto \partial_x u(\cdot,\mu) \in L^k(\mathbb{R}^d).$$

Moreover, keeping the notations of Definition 9.14, Assumption (2) is satisfied and setting $\alpha_1 = 1$, Assumption (5) is satisfied because we have supposed $k \ge 5d$. The same holds true for $\partial_x^2 u$. Since g is continuous and at most of linear growth, the linear derivative of u satisfies Assumption (3) in Definition 9.14 and is given, for all $x, v \in \mathbb{R}^d$ and for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, by

$$\frac{\delta u}{\delta m}(x,\mu)(v) = g(x,v).$$

As $\nabla g \in (W^{1,k}(\mathbb{R}^{2d}))^d$, Assumption (4) in Definition 9.14 is satisfied, as well as the growth property in Assumption (5) with $\alpha_2 = 0$.

□

9.7.6 Proof of Example 9.18

As in 9.7.5, the function u is well-defined and continuous because $\nabla g \in L^\infty(\mathbb{R}^d)$ and is continuous. Thus g is at most of linear growth. It is clear with the assumption on ∇F that for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $u(\cdot,\mu) \in W^{2,k_1}_{\text{loc}}(\mathbb{R}^d)$. It follows from the continuity of $\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} g\, d\mu$ that the function

$$\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_x u(\cdot,\mu) = \partial_x F\left(\cdot, \int_{\mathbb{R}^d} g\, d\mu\right) \in (W^{1,k_1}(B_R))^d,$$

is also continuous for all $R > 0$. Moreover, it is easy to show with Remark 9.2 that for all $x \in \mathbb{R}^d$, $u(x,\cdot)$ admits a linear derivative given by

$$\forall (\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d,\ \frac{\delta u}{\delta m}(x,\mu)(v) = g(v)\partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right).$$

Assumption (3) in Definition 9.14 is clearly satisfied because $\partial_y F$ is continuous. Next, we compute the derivatives of $\frac{\delta u}{\delta m}(\cdot,\mu)(\cdot)$ with respect to v in the sense of distributions. For $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^{2d})$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, Fubini's theorem ensures that

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} g(v)\partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right) \partial_v \phi(x,v)\, dx\, dv &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g(v) \partial_v \phi(x,v)\, dv \right) \partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right) dx \\
&= -\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \nabla g(v) \phi(x,v)\, dv \right) \partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right) dx \\
&= -\int_{\mathbb{R}^{2d}} \left(\nabla g(v) \partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right)\right) \phi(x,v)\, dx\, dv.
\end{aligned}$$

This proves exactly that

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\ \forall x, v \in \mathbb{R}^d,\ \partial_v \frac{\delta u}{\delta m}(x,\mu)(v) = \nabla g(v) \partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right).$$

Since $\nabla g \in L^{k_2}(\mathbb{R}^d)$ and $\partial_y F\left(\cdot, \int_{\mathbb{R}^d} g\, d\mu\right) \in L^\infty(B_R)$, for all $R > 0$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the function

$$(x, v) \in B_R \times \mathbb{R}^d \mapsto \partial_v \frac{\delta u}{\delta m}(x, \mu)(v)$$

belongs to $L^{k_2}(B_R \times \mathbb{R}^d)$. Moreover, the function

$$\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_v \frac{\delta u}{\delta m}(\cdot, \mu)(\cdot) \in L^{k_2}(B_R \times \mathbb{R}^d)$$

is continuous because $F \in \mathcal{C}^1(\mathbb{R}^d \times \mathbb{R}; \mathbb{R})$ and thus $y \mapsto \partial_y F(\cdot, y) \in L^\infty(B_R)$ is continuous. The same reasoning proves that

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d),\, \forall x, v \in \mathbb{R}^d,\, \partial_v^2 \frac{\delta u}{\delta m}(x, \mu)(v) = \nabla^2 g(v) \partial_y F\left(x, \int_{\mathbb{R}^d} g\, d\mu\right),$$

and that the function

$$\mu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \partial_v^2 \frac{\delta u}{\delta m}(\cdot, \mu)(\cdot) \in L^{k_2}(B_R \times \mathbb{R}^d)$$

is continuous for all $R > 0$. We conclude that $u \in \mathcal{W}_2(\mathbb{R}^d)$ with Remark 9.15.

$\square$

Chapter 10

Scaling limit of a kinetic inhomogeneous stochastic system in the quadratic potential

---

This chapter corresponds to the article [CL23], written in collaboration with Emeline Luirard. It has been accepted for publication in *Discrete and Continuous Dynamical Systems*.

Abstract. We consider a particle evolving in the quadratic potential and subject to a time-inhomogeneous frictional force and to a random force. The couple of its velocity and position is solution to a stochastic differential equation driven by a symmetric α-stable Lévy process with $\alpha \in (1, 2]$ and the frictional force is of the form $t^{-\beta} \operatorname{sgn}(v)|v|^{\gamma}$. We identify three regimes for the behavior in long-time of the couple velocity-position with a suitable rescaling, depending on the balance between the frictional force and the index of stability α of the noise.

10.1 Introduction and main results

10.1.1 Model and motivations

In this paper, we study the long-time behavior of a stochastic system modelling a particle, with velocity $V \in \mathbb{R}$ and position $X \in \mathbb{R}$. The particle evolves in the quadratic potential $\mathcal{U} : x \mapsto \frac{x^2}{2}$, and is subject to a time-inhomogeneous frictional force b and to a random force L. The dynamics of the particle is described by the following stochastic damping Hamiltonian system driven by an α-stable process L with $\alpha \in (1, 2]$

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}L_t - b(t, V_t)\,\mathrm{d}t - \nabla\mathcal{U}(X_t)\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0), \quad t_0 > 0 \text{ being fixed.} \end{cases} \tag{10.1}$$

The driving process L models a random force coming from the interaction of the particle with its environment represented by a surrounding heat bath. As in the classical Langevin model (see [Lan08]), L can be a Brownian motion denoted by B in the sequel. It corresponds to take $\alpha = 2$. It is natural to consider other types of noises such as Lévy processes, which are also largely used to model physical and biological systems (Lévy flights and anomalous diffusion), see e.g. [MJW01] for the physical point of view, [Dit99] in stochastic climate dynamics, and [JMW05] for the mathematical point of view. The case where L is an α-stable process is of particular interest. It is a generalization of the Brownian motion with jumps since it satisfies that for any $c > 0$, $(c^{\frac{1}{\alpha}} L_{t/c})_t$ has the same distribution as L (self-similarity property).

Degenerate systems like (10.1) have been intensively studied for several years. In particular, the existence and uniqueness of solutions to degenerate SDEs have been discussed in many works. These models are called degenerate because the noise is only present in one component of the system but can be transferred into others by drift terms. The well-posedness of these systems, when their deterministic version is ill-posed, can be proved by taking advantage of the regularizing effect of the noise and of its propagation through the whole system. The case of Brownian degenerate SDEs has been of course wildly explored, see e.g. [FFPV17], [WZ15], [Zha16], [CdR17], [HMC18] and references therein. The time-dependence is treated in the last four cited papers. The case of a Lévy driving process is more recent, see e.g. [Zha14] in a time-homogeneous setting, and [MM21] for drifts depending on time.

From another point of view, stochastic Hamiltonian systems, as (10.1) with $b = 0$, have been widely studied. An interesting problem is to understand their asymptotic behaviors. The Hamiltonian process associated with this system is defined, for $t \geq t_0$, by $H_t := \frac{1}{2}|V_t|^2 + \mathcal{U}(X_t)$. For example, the long-time dynamics of the Hamiltonian process under a suitable rescaling is studied in [AK94]. The case of time-homogeneous damping Hamiltonian systems is tackled in [Wu01] (see also references therein).

The long-time behavior of a particle evolving in a free potential, i.e. $\mathcal{U} = 0$, has already been studied, see e.g. [GO13], [FT21], [GL21a], [GL21b] and references therein. In this case, The velocity process can be studied independently on the position process. Even in the time-homogeneous case, various asymptotic behaviors can appear. Whenever the random force is supposed to be Brownian, a particular non-linear Langevin's type SDE was studied in [FT21]:

$$V_t = v_0 + B_t - \frac{\rho}{2}\int_0^t \frac{V_s}{1+V_s^2}\,\mathrm{d}s \quad \text{and} \quad X_t = x_0 + \int_0^t V_s\,\mathrm{d}s.$$

In that case, the frictional force asymptotically behaves as $-\dfrac{\rho}{v}$, which induces the velocity process to "behave", far away from zero, like a (signed) Bessel process of dimension $1-\rho$. Various asymptotic behaviors of the position process appear, depending on the moment order of Bessel excursion area (which depends itself on the value of ρ). More precisely, when $\rho \geq 5$, the moment is of order 2, hence, using a suitable rescaling, the authors show that the position process behaves asymptotically as a Brownian motion. An α-stable process appears as limiting dynamics when $\rho \in [1,5)$. The index of stability α is a function of ρ, which interpolates the power of the rescaling from $\frac{1}{2}$ (Brownian motion) to $\frac{3}{2}$ (integrated Bessel process). This last behavior occurs when $\rho \in (0,1)$. However, the tools used in [FT21], such as invariant measure, scale function and speed measure, are limited to time-homogeneous coefficients.

In [GO13], [GL21a] and [GL21b], the drift coefficient b is allowed to depend on time under an homogeneity condition. More precisely, the following system is considered

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}L_t - \rho\dfrac{\operatorname{sgn}(V_t)\,|V_t|^\gamma}{t^\beta}\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t. \end{cases}$$

The frictional force is time-inhomogeneous, depending on non-negative parameters β, γ and ρ. When the particle moves slowly, classical mechanics ensures that the frictional force is linear, i.e. $\gamma = 1$. Whereas in the turbulent regime, when the particle moves faster, thanks to fluid dynamics, the frictional force depends quadratically on the velocity, i.e. $\gamma = 2$. That is why in a broader framework, we assume that the frictional force has a space component of the form $v \mapsto -\rho \operatorname{sgn}(v) |v|^\gamma$. Moreover, the frictional force can depend on time through a friction coefficient $t \mapsto \rho_t$. For a particle evolving in a fluid, it can be the case for example when the viscosity of the fluid or the geometry of the particle change with time. For this reason, a time dependence is added to the function b in [GO13], [GL21a] and [GL21b]. In these works, it is assumed that $\rho_t = \frac{\rho}{t^\beta}$. The main goal behind the study of this model is to understand the competition between the frictional force, which tends to immobilize the system and the random force perturbing it. Notice that, by the self-similarity property satisfied by L, $\mathbb{E}[|L_t|]$ is proportional to $t^{\frac{1}{\alpha}}$. This shows that the noise L_t acts with a typical scale $t^{\frac{1}{\alpha}}$ and thus, when α decreases, it perturbs the velocity with higher typical values. The interest of the works mentioned above is to study the long-time behavior of the system through the prism of the competition between these two opposite actions.

Let us mention two relevant examples in the Brownian case before explaining the results obtained in [GL21a, GL21b]. When $\beta = 0$, the friction coefficient does not decrease with time. By ergodicity, the velocity converges towards its invariant distribution and thus, the rescaled position process $(\varepsilon^{\frac{1}{2}} X_{t/\varepsilon})_t$ behaves as a Brownian motion as ε tends to 0. When "$\beta = +\infty$", i.e. when there is no frictional force, the rescaled velocity-position process $(\varepsilon^{\frac{1}{2}} V_{t/\varepsilon}, \varepsilon^{\frac{3}{2}} X_{t/\varepsilon})_t$ converges in distribution towards $\left(B_t, \int_0^t B_s \,\mathrm{d}s\right)_t$. When $\beta > 0$, the frictional force is evanescent: it slows down the system but less and less efficiently as time increases and we expect a transition between the two extreme cases mentioned above, both on the limiting processes and on the rescaling.

In [GO13], the authors study the convergence in distribution, when t tends to $+\infty$, of $r_t V_t$, for a certain rate of convergence r_t in the case where L is a Brownian motion. In [GL21a], the authors extend the results obtained in [GO13] to the whole process given by the couple velocity-position. Namely, the authors study the limit in distribution of the rescaled process $(r_{\varepsilon,V} V_{t/\varepsilon}, r_{\varepsilon,X} X_{t/\varepsilon})_t$ for two appropriate rates of convergence $r_{\varepsilon,V}$ and $r_{\varepsilon,X}$. Results were further generalized in [GL21b] to an α-stable driving process. To be more precise, the authors highlight three regimes, depending on the balance between β, γ and α, the index of stability of L.

- Whenever the frictional force is sufficiently "small at infinity", i.e. if β is large enough, the rescaled process behaves as if there was no frictional force. It thus converges in distribution towards the Kolmogorov process $(L, \int_0^\cdot L)$, as in the particular case "$\beta = +\infty$" mentioned above in the Brownian case and with the same rescaling $(r_{\varepsilon,V}, r_{\varepsilon,X}) = (\varepsilon^{\frac{1}{\alpha}}, \varepsilon^{1+\frac{1}{\alpha}})$.
- When the two forces offset, the rescaling remains the same as in the preceding regime and the limiting process is still of kinetic form $(\mathcal{V}, \int_0^\cdot \mathcal{V})$, but the process $\mathcal{V}$ is henceforth ergodic.
- Whereas, when the drag force swings with the random process, i.e. when β is small enough, the limiting process is no longer kinetic and the rescaling is not the same as in the two preceding regimes. The rescaled velocity process converges in finite dimensional distributions towards a white noise. Here, the asymptotic behavior is somehow an interpolation between the two extreme cases $\beta = 0$ and "$\beta = +\infty$", which is explained by the slow decrease of the frictional force with time.

The proofs are essentially based on the self-similarity of the driving process and on moment estimates of the velocity process.

In this paper, we are interested in the long-time behavior of the solution to the following system of

SDEs, defined on the time interval $[t_0, +\infty)$, where $t_0 > 0$ and $x_0, v_0 \in \mathbb{R}$ are fixed

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}L_t - \operatorname{sgn}(V_t)\dfrac{|V_t|^\gamma}{t^\beta}\,\mathrm{d}t - X_t\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0), \end{cases} \tag{SKE}$$

Here $\gamma, \beta > 0$ and L is a symmetric α-stable process on $\mathbb{R}$ with $\alpha \in (1, 2]$. More precisely, our goal is to study the asymptotic behavior, as $\varepsilon \to 0$, of the rescaled velocity-position process

$$(Z_t^{(\varepsilon)})_t := \left(r_\varepsilon \begin{pmatrix} X_{t/\varepsilon} \\ V_{t/\varepsilon} \end{pmatrix} \right)_t,$$

for an appropriate rate of convergence r_ε. Our first motivation is to study how the presence of the quadratic potential influences the results obtained in [GL21a, GL21b] through a confining effect on the position X. Indeed, the confining effect is here related to the position of the particle and does not disappear asymptotically contrary to the frictional force. It has thus an effect both on the limiting processes and on the rescaling. Here, it is a competition of the quadratic potential and the frictional force, which confines and slows down the system, against the noise which perturbs it.

Notice that our system without noise and frictional force is nothing else than the classical harmonic oscillator

$$\begin{cases} v_t' = -x_t, \\ x_t' = v_t. \end{cases}$$

The intrinsic oscillatory behavior induced by the quadratic potential prevents the rescaled process $Z^{(\varepsilon)}$ from converging as a process. However, we prove that each of its one-dimensional marginal distributions converges. In order to obtain the convergence of the whole process, the key idea is to remove the oscillations present in the system. Namely, we set for $t \geq t_0$

$$\Theta_t := \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \quad \text{and} \quad Y_t := \Theta_t^{-1} \begin{pmatrix} X_t \\ V_t \end{pmatrix},$$

where Θ_t is the rotation on $\mathbb{R}^2$ of angle $-t$ and we study the behavior of $(r_\varepsilon Y_{t/\varepsilon})_t$ as ε tends to 0, for a certain rate of convergence r_ε.

10.1.2 Notations, main results and comments

Let us first introduce some notations used throughout the paper. For simplicity, we shall write $\mathcal{C}$ and $\mathcal{D}$ respectively for $\mathcal{C}((0,+\infty), \mathbb{R}^2)$, the space of continuous functions defined on $(0, +\infty)$ and $\mathcal{D}((0,+\infty), \mathbb{R}^2)$, the Skorokhod space of functions defined on $(0, +\infty)$ which are càdlàg on every compact subinterval of $(0, +\infty)$. For $x, y \in \mathbb{R}^2$, $\|x\|$ represents the Euclidean norm of x, and $x \cdot y$ the inner product of x and y. If $x \in \mathbb{R}^2$, for each $i \in \{1, 2\}$, $x^{(i)}$ denotes its i-th component. The minimum between two reals is denoted by $\wedge$. We call I_2 the identity matrix of dimension 2 and A^T is the transpose matrix of a matrix A. Finally, we denote by C some positive constant, which may change from line to line, and we use subscripts to indicate the parameters on which it depends when it is necessary. For the sake of simplicity, we denote by C_{t_0} a positive constant depending only on t_0, x_0 and v_0, which are fixed throughout the paper.

We can now state our results. The following theorem deals with convergences in distribution in the space $\mathcal{C}$ endowed with topology of uniform convergence on every compact set of $(0,+\infty)$.

Theorem 10.1 (Brownian case i.e. $\alpha = 2$)**.** *Define* $q := \frac{\beta}{\gamma+1}$, $r_\varepsilon := \varepsilon^{q\wedge\frac{1}{2}}$ *and set* $(Y_t^{(\varepsilon)})_{t\geq\varepsilon t_0} := \left(r_\varepsilon \Theta_{t/\varepsilon}^{-1}(X_{t/\varepsilon}, V_{t/\varepsilon})^T\right)_{t\geq\varepsilon t_0}$. *Let* $\mathcal{B}$ *be a standard two-dimensional Brownian motion on* $\mathbb{R}^2$.

(i) (Super-critical regime i.e. $2q > 1$*). The rescaled process* $Y^{(\varepsilon)}$ *converges in distribution towards* $\left(\mathcal{B}_{\frac{t}{2}}\right)_{t>0}$.

(ii) (Critical regime i.e. $2q = 1$*). Assume that* $\gamma = 1$*. The rescaled process* $Y^{(\varepsilon)}$ *converges in distribution towards* $\left(\frac{1}{\sqrt{2t}}\int_0^t \sqrt{s}\,\mathrm{d}\mathcal{B}_s\right)_{t>0}$*, which is the centered Gaussian process with covariance kernel* $K(s,t) = \frac{(s\wedge t)^2}{4\sqrt{st}} I_2$.

(iii) (Sub-critical regime i.e. $2q < 1$*). Assume that* $\gamma = 1$ *and* $\beta \in \left(\frac{1}{2}, 1\right)$*. The rescaled process* $Y^{(\varepsilon)}$ *converges in finite dimensional distributions towards the centered Gaussian process with covariance kernel* $K(s,t) = \frac{1}{2}s^\beta \mathbb{1}_{\{s=t\}} I_2$.

Let us denote by ψ the characteristic exponent of the symmetric stable process L. It follows from Theorem 14.15 p. 86 in [Sat99] that there exists $a > 0$ such that for all $\xi \in \mathbb{R}$,

$$\psi(\xi) = -a|\xi|^\alpha. \tag{10.2}$$

In the next theorem, the convergences occur in the space $\mathcal{D}$ endowed with the Skorokhod metric.

Theorem 10.2 (Stable case i.e. $\alpha \in (1,2)$)**.** *Assume that* $\gamma \in (0,\alpha)$*. Define* $q := \frac{\beta}{\gamma+\alpha-1}$, $r_\varepsilon := \varepsilon^{q\wedge\frac{1}{\alpha}}$ *and set* $(Y_t^{(\varepsilon)})_{t\geq\varepsilon t_0} := \left(r_\varepsilon \Theta_{t/\varepsilon}^{-1}(X_{t/\varepsilon}, V_{t/\varepsilon})^T\right)_{t\geq\varepsilon t_0}$*. Let* $\mathcal{L}$ *be the rotationally invariant stable process on* $\mathbb{R}^2$*, whose characteristic exponent is given by*

$$\xi \in \mathbb{R}^2 \mapsto -\widetilde{C}\,\|\xi\|^\alpha, \quad \textit{with } \widetilde{C} := \frac{a}{2\pi}\int_0^{2\pi} |\cos(x)|^\alpha\,\mathrm{d}x.$$

(i) (Super-critical regime i.e. $\alpha q > 1$*). The rescaled process* $Y^{(\varepsilon)}$ *converges in distribution towards* $(\mathcal{L}_t)_{t>0}$.

(ii) (Critical regime i.e. $\alpha q = 1$*). Assume that* $\gamma = 1$*. The rescaled process* $Y^{(\varepsilon)}$ *converges in distribution towards the Lévy-type process* $\left(\frac{1}{\sqrt{t}}\int_0^t \sqrt{s}\,\mathrm{d}\mathcal{L}_s\right)_{t>0}$.

(iii) (Sub-critical regime i.e. $\alpha q < 1$*). Assume that* $\gamma = 1$ *and* $\beta \in \left(\frac{1}{2}, 1\right)$*. Then, for all* $(t_1,\cdots,t_d) \in (0,+\infty)^d$, $\left(Y_{t_1}^{(\varepsilon)},\cdots,Y_{t_d}^{(\varepsilon)}\right)$ *converges in distribution towards the product measure* $\mu_{t_1}\otimes\cdots\otimes\mu_{t_d}$*, where* μ_t *is the distribution with characteristic function*

$$\xi \in \mathbb{R}^2 \mapsto \exp\left(-\frac{2}{\alpha}\widetilde{C}\,\|\xi\|^\alpha\, t^\beta\right).$$

Remark 10.3. The symmetry of L is only required to ensure the well-posedness of (SKE) when $\gamma < 1$ using [MM21].

Remark 10.4. At first sight, the parameter α does not seem to appear in the limiting process of Theorem 10.2 *(ii)* contrary to the Brownian case where there is a factor $\frac{1}{\sqrt{2}}$ in front of the stochastic integral.

The reason is that the constant α is hidden in the constant $\widetilde{C}$ which defines the α-stable process $\mathcal{L}$ in Theorem 10.2. Indeed, if we formally take $\alpha = 2$ in Theorem 10.2 (ii), we recover the limiting process of Theorem 10.1 (ii). Let us justify it. When $\alpha = 2$, the constant $\widetilde{C}$ can be computed explicitly and is equal to $\frac{1}{4}$. As the characteristic exponent of $\mathcal{L}$ is given by

$$\xi \in \mathbb{R}^2 \mapsto -\frac{\|\xi\|^2}{4},$$

we deduce that $(\mathcal{L}_t)_{t\geq 0} = (\mathcal{B}_{t/2})_{t\geq 0}$ in distribution. When $\alpha = 2$, the limiting process in 10.2 (ii) is thus given, for all $t > 0$, by

$$\frac{1}{\sqrt{2t}} \int_0^t \sqrt{s}\sqrt{2}\, \mathrm{d}\mathcal{B}_{s/2},$$

which is equal in distribution to the limiting process of Theorem 10.1 (ii) using the self-similarity of $\mathcal{B}$. We similarly notice that if we formally take $\alpha = 2$ in Theorem 10.2 (i) and (iii), we recover the same limiting processes as in Theorem 10.1 (i) and (iii) with the same rescaling.

Remark 10.5. The Hamiltonian process associated with the system is given by

$$H_t := \frac{1}{2}|V_t|^2 + \frac{1}{2}|X_t|^2 = \frac{1}{2}\|Z_t\|^2 = \frac{1}{2}\|Y_t\|^2.$$

Combining the preceding results with the continuous mapping theorem, we deduce the convergence of the rescaled energy process $(H_t^{(\varepsilon)})_{t>0} := (r_\varepsilon^2 H_{t/\varepsilon})_{t>0}$ as $\varepsilon \to 0$ either as a process in the critical and super-critical regimes, or for finite dimensional distributions in the sub-critical regime.
For example in the super-critical regime with $\alpha = 2$, the limiting energy process $(H_t^0)_{t\geq 0} := (\frac{1}{2}\|\mathcal{B}_{\frac{t}{2}}\|^2)_{t\geq 0}$ is the squared Bessel process, which is the solution to the following equation

$$\mathrm{d}H_t^0 = \sqrt{H_t^0}\,\mathrm{d}B_t + \frac{1}{2}\,\mathrm{d}t, \quad H_0^0 = 0,$$

where B is a standard one-dimensional Brownian motion. Note that we recover the limiting energy process obtained in Theorem 2.1 in [AK94] for the non-damped Hamiltonian system. The interpretation is that if the frictional force decreases sufficiently quickly as $t \to +\infty$, namely if β is large enough, then the rescaled Hamiltonian process converges as if there were no damping.

We obtain furthermore the convergence in distribution as $t \to +\infty$ of $t^{-q\wedge\frac{1}{\alpha}}(X_t, V_t)^T$ in the following corollary.

Corollary 10.6. *Let us define $(Z_t^{(\varepsilon)})_{t\geq\varepsilon t_0} := (r_\varepsilon(X_{t/\varepsilon}, V_{t/\varepsilon})^T)_{t\geq \varepsilon t_0}$, where $r_\varepsilon := \varepsilon^{q\wedge\frac{1}{\alpha}}$. The rescaled process $Z^{(\varepsilon)}$ does not converge in distribution. However, we deduce from Theorems 10.1 and 10.2 and under the same assumptions, the convergence in distribution of $r_{1/t}(X_t, V_t)^T$ towards explicit limits, as $t \to +\infty$.*

In the Brownian case, the limit is either $\mathcal{N}(0, \frac{1}{2}I_2)$ in the super-critical and sub-critical regimes, or $\mathcal{N}(0, \frac{1}{4}I_2)$ in the critical regime.

In the stable case, keeping the same notations as in Theorem 10.2, the characteristic function of the limit is given, for all $\xi \in \mathbb{R}^2$, by

(i) $\exp\left(-\widetilde{C}\,\|\xi\|^\alpha\right)$ *in the super-critical regime,*

(ii) $\exp\left(-\left(1+\frac{\alpha}{2}\right)^{-1}\widetilde{C}\,\|\xi\|^\alpha\right)$ *in the critical regime,*

(iii) $\exp\left(-\frac{2}{\alpha}\widetilde{C}\,\|\xi\|^{\alpha}\right)$ *in the sub-critical regime.*

The switch between the three regimes results in different scale parameters of the limiting distributions. Let us also notice that in the Brownian setting, the position X and the velocity V become independent in large time since the covariance matrix of the limiting Gaussian distribution is diagonal. However, this is false for the stable case. Indeed, the limit is a rotationally invariant stable distribution on $\mathbb{R}^2$, which cannot have independent coordinates.

As in [GL21a, GL21b], we highlight three regimes for the asymptotic behavior of the system. However, the rate of convergence of the position X is different from that found in [GL21a, GL21b], when $\mathcal{U} = 0$. Indeed, contrary to the free potential system, the position process is somehow more diffusive. This is due to the structure of our model. Namely, the presence of the quadratic potential allows the noise to propagate more efficiently from the velocity component to the position one (see [FFPV17] for more details). This explains why both the limiting processes and the rate of convergence are different between our work and [GL21a, GL21b]. Let us also note that the position process grows more slowly in our case than when $\mathcal{U} = 0$. For example, in the Brownian super-critical regime, X_t asymptotically behaves as $\mathcal{N}(0, \frac{t}{2})$ when t tends to infinity in our framework, but as $\mathcal{N}(0, \frac{t^3}{3})$ in the free potential one. This difference can also be seen in moment estimates established for the position process X (see Remarks 10.9 and 10.15). This is explained by the fact that the quadratic potential confines the position of the particle through a spring force.

10.1.3 Strategy and plan of the paper

In our model, the particle is no longer free, contrary to [GL21a, GL21b], and the equations on the position and the velocity are intrinsically linked to each other. Therefore, we can no longer separate by components the study of the velocity-position process. Writing the system (SKE) in a vector viewpoint, as done in [FFPV17], we use a variation of constants method to return to the study of a two-dimensional system in a free potential. We then adapt the methods used in [GL21a, GL21b]. In the super-critical regime, the proof is essentially based on the self-similarity of the driving process and on moment estimates of V and X. In the critical and sub-critical regimes, we need to restrict ourselves to a linear drag force, i.e. $\gamma = 1$, in order to rely on the study of the asymptotic behavior of the solution to the underlying non-autonomous ordinary differential equation (ODE). Whenever the driving process is Brownian, we take advantage of the theory of Gaussian processes. The convergence is thus characterized by the study of the mean and covariance functions. In the case of a stable driving process, we need to study the convergence, in distribution and as a process, of a Wiener-Lévy integral (i.e. the integral of a deterministic function integrated against a stable process). The key point here is to use the fact that a Wiener-Lévy integral is a process with independent increments.

Our paper is organized as follows. We consider the case of a Brownian driving process in Section 10.2, and we follow the same structure for an α-stable driving process in Section 10.3. For the sake of clarity, we opt for separating the two cases since the tools used are different. Finally, we state and prove some technical results in Section 10.4 and Section 10.5.

10.2 Study of the system driven by a Brownian motion

In this section, the driving process L is supposed to be a standard Brownian motion, i.e. $\alpha = 2$. It will be denoted by B to keep standard notations. To be precise, (SKE) becomes

$$\begin{cases} \mathrm{d}V_t = \mathrm{d}B_t - \operatorname{sgn}(V_t)\dfrac{|V_t|^\gamma}{t^\beta}\,\mathrm{d}t - X_t\,\mathrm{d}t, \\ \mathrm{d}X_t = V_t\,\mathrm{d}t, \\ (V_{t_0}, X_{t_0}) = (v_0, x_0). \end{cases} \tag{10.3}$$

The previous system can be written in a vector viewpoint. Indeed, we set, for all $t \geq t_0$ and $v \in \mathbb{R}$,

$$Z_t := \begin{pmatrix} X_t \\ V_t \end{pmatrix}, \; W_t := \begin{pmatrix} 0 \\ B_t \end{pmatrix}, \; A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \; \Gamma := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } F(t,v) := \begin{pmatrix} 0 \\ \operatorname{sgn}(v)\dfrac{|v|^\gamma}{t^\beta} \end{pmatrix}.$$

Thereby, the system (SKE) can be rewritten as

$$\begin{cases} \mathrm{d}Z_t = \Gamma\,\mathrm{d}W_t + AZ_t\,\mathrm{d}t - F(t, V_t)\,\mathrm{d}t, \\ Z_{t_0} = z_0 := (x_0, v_0)^T. \end{cases} \tag{10.4}$$

Notice that the matrix A is the rotation matrix of angle $\frac{\pi}{2}$ and that, for all $t \in \mathbb{R}$,

$$\Theta_t := e^{tA} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

We also define, for any $t \geq t_0$, $Y_t := e^{-tA} Z_t$. We easily check, with Itô's formula, that Y is given by

$$\mathrm{d}Y_t = e^{-tA}\Gamma\,\mathrm{d}W_t - e^{-tA}F(t, V_t)\,\mathrm{d}t. \tag{SDE_Y}$$

10.2.1 Existence up to explosion

Theorem 10.7. *The system of SDEs (10.3) admits a unique (global) strong solution if $\gamma \in (0,1]$. And if $\gamma > 1$, there exists a unique strong solution defined up to its explosion time τ_∞.*

Proof. In the case $\gamma > 1$, the coefficients of the SDE (10.4) are locally Lipschitz continuous with respect to the space variable, uniformly in time. By a standard localization argument as mentioned in [Kal02, p. 417, after (13)], there exists a unique solution up to explosion. We also refer to [Mao11, Theorem 2.8 p. 154] for a proof of the existence of a solution, up to explosion, in the more general case of path-dependent SDEs.

Assume now that $\gamma \leq 1$. We will use Theorem 1 in [HMC18]. Keeping the same notations, we have in our case, for any $(x, v) \in \mathbb{R}^2$ and $t \geq t_0$, $F_1(t, v, x) := -\operatorname{sgn}(v)|v|^\gamma t^{-\beta} - x$, $F_2(t, v, x) := v$ and $\sigma(t, v, x) = 1$. Assumptions **(ML)** and **(UE)** in [HMC18] are obviously satisfied. Let us now remark that F_1 is γ-Hölder with respect to $v \in \mathbb{R}$ uniformly with respect to $t \geq t_0$ and $x \in \mathbb{R}$, and is Lipschitz continuous with respect to x, uniformly with respect to t and v. With the notations used in [HMC18], we have $\beta_1 = \gamma$ and $\beta_2 = 1$. Thus, Assumption $(\boldsymbol{T_\beta})$ is satisfied. Finally, we check that Assumption $(\boldsymbol{H_\eta})$

is satisfied. Since $\partial_v F_2 = 1$, we can conclude, taking η small enough and $\mathcal{E}_1 = \{1\}$, that there exists a unique strong solution to (10.3).

□

10.2.2 Moment estimates and non-explosion

In this section, we state and prove moment estimates of Z. It will be useful to control some stochastic terms appearing later. For all $n \geq 0$, define the stopping time

$$\tau_n := \inf\{t \geq t_0,\ \|Z_t\| \geq n\}.$$

Set $\tau_\infty := \lim_{n\to+\infty} \tau_n$ the explosion time of Z.

Proposition 10.8. *The explosion time of Z is a.s. infinite and, for all $\kappa \geq 0$ and $t \geq t_0$*

$$\mathbb{E}\left[\|Z_t\|^\kappa\right] \leq C_{\kappa,t_0} t^{\frac{\kappa}{2}}. \tag{10.5}$$

Remark 10.9. Let us mention that the moment estimate obtained for the position process X is a priori smaller in our case than in the free potential case [GL21a]. It is explained by the confining effect of the quadratic potential.

Proof. The proof is adapted from [GL21a] to two-dimensional processes. For the sake of completeness, we sketch the proof in our context.
Using Itô's formula applied to the function $f : (x, v) \mapsto x^2 + v^2$ and the fact that for all $z \in \mathbb{R}^2$, $z \cdot Az = 0$, we deduce that, for all $t \geq t_0$,

$$\|Z_{t\wedge\tau_n}\|^2 \leq \|z_0\|^2 + \int_{t_0}^{t} 2\mathbb{1}_{\{s\leq\tau_n\}} Z_s \cdot (\Gamma\,\mathrm{d}W_s) - \int_{t_0}^{t\wedge\tau_n} 2Z_s \cdot F(s, V_s)\,\mathrm{d}s + (t - t_0).$$

Remark that for any $s \geq t_0$, $Z_s \cdot F(s, V_s) = V_s \operatorname{sgn}(V_s) |V_s|^\gamma s^{-\beta} \geq 0$. Taking expectation yields

$$\mathbb{E}\left[\|Z_{t\wedge\tau_n}\|^2\right] \leq \|z_0\|^2 + (t - t_0) \leq C_{t_0} t.$$

Thanks to Lemma 10.21, we can conclude that the explosion time of Z is a.s. infinite. Set $\kappa \in [0, 2]$, so, by Jensen's inequality and Fatou's lemma

$$\mathbb{E}\left[\|Z_{t\wedge\tau_\infty}\|^\kappa\right] \leq \left(\mathbb{E}\left[\|Z_{t\wedge\tau_\infty}\|^2\right]\right)^{\frac{\kappa}{2}} \leq \left(\liminf_{n\to\infty} \mathbb{E}\left[\|Z_{t\wedge\tau_n}\|^2\right]\right)^{\frac{\kappa}{2}} \leq C_{\kappa,t_0} t^{\frac{\kappa}{2}}. \tag{10.6}$$

This leads to (10.5).

When $\kappa > 2$, $v \mapsto \|v\|^\kappa$ is a $\mathcal{C}^2$-function, so by Itô's formula, for all $t \geq t_0$,

$$\|Z_{t\wedge\tau_n}\|^\kappa \leq \|z_0\|^\kappa + \int_{t_0}^{t\wedge\tau_n} \kappa \|Z_s\|^{\kappa-2} Z_s \cdot (\Gamma\,\mathrm{d}W_s) - \int_{t_0}^{t\wedge\tau_n} \kappa \|Z_s\|^{\kappa-2} Z_s \cdot F(s, V_s)\,\mathrm{d}s \\ + \int_{t_0}^{t\wedge\tau_n} C_\kappa \|Z_s\|^{\kappa-2}\,\mathrm{d}s.$$

In addition, it follows from the hypothesis on the sign of the drift function that

$$\|Z_{t\wedge\tau_n}\|^\kappa \leq \|z_0\|^\kappa + \int_{t_0}^{t} \kappa \mathbb{1}_{\{s\leq\tau_n\}} \|Z_s\|^{\kappa-2} Z_s \cdot (\Gamma \, \mathrm{d}W_s) + \int_{t_0}^{t\wedge\tau_n} C_\kappa \|Z_s\|^{\kappa-2} \, \mathrm{d}s. \tag{10.7}$$

Taking expectation in (10.7), we have

$$\mathbb{E}\left[\|Z_{t\wedge\tau_\infty}\|^\kappa\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[\|Z_{t\wedge\tau_n}\|^\kappa\right] \leq \|z_0\|^\kappa + \int_{t_0}^{t} C_\kappa \mathbb{E}\left[\|Z_s\|^{\kappa-2}\right] \mathrm{d}s.$$

When $0 \leq \kappa - 2 \leq 2$, we can upper bound $\mathbb{E}\left[\|Z_s\|^{\kappa-2}\right]$ by injecting (10.6) and get

$$\mathbb{E}\left[\|Z_{t\wedge\tau_\infty}\|^\kappa\right] \leq \|z_0\|^\kappa + \int_{t_0}^{t} C_{\kappa,t_0} s^{\frac{\kappa-2}{2}} \, \mathrm{d}s \leq C_{\kappa,t_0} s^{\frac{\kappa}{2}}.$$

The method is then applied inductively to prove the inequality for all $\kappa > 2$. □

10.2.3 Asymptotic behavior of the solution

We gather in this section the proof of Theorem 10.1. The strategy is to prove the convergence of the finite dimensional distributions (f.d.d.) of the process $Y^{(\varepsilon)}$, as $\varepsilon \to 0$, and its tightness in the critical and super-critical regimes. We first focus on the tightness.

Lemma 10.10. *If $2q \geq 1$, then the family $\left\{(\sqrt{\varepsilon} Y_{t/\varepsilon})_{t\geq\varepsilon t_0}, \ \varepsilon > 0\right\}$ is tight on every compact interval $[m, M]$, with $0 < m \leq M$.*

Proof. We use the Kolmogorov criterion stated in Problem 4.11 p. 64 in [KS98].
Take ε_0 small enough such that for all $\varepsilon \leq \varepsilon_0$, we have $\varepsilon t_0 \leq m$. Fix $m \leq s \leq t \leq M$ and $a > 4$. Define, for $t \geq \varepsilon t_0$, the local martingale term appearing in (SDE_Y)

$$M_t^{(\varepsilon)} := \sqrt{\varepsilon} \int_{t_0}^{t/\varepsilon} e^{-sA} \Gamma \, \mathrm{d}W_s = \sqrt{\varepsilon} \int_{t_0}^{t/\varepsilon} \begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix} \mathrm{d}B_s. \tag{10.8}$$

Using Jensen's inequality, moment estimates (see Proposition 10.8) and Burkholder-Davis-Gundy's inequality (see Theorem 4.4.22 p. 263 in [App09]), we have

$$\begin{aligned}
\mathbb{E}\left[\left\|Y_t^{(\varepsilon)} - Y_s^{(\varepsilon)}\right\|^a\right] &\leq C_a \mathbb{E}\left[\left\|M_t^{(\varepsilon)} - M_s^{(\varepsilon)}\right\|^a\right] + C_a \mathbb{E}\left[\left\|\sqrt{\varepsilon} \int_{s/\varepsilon}^{t/\varepsilon} e^{-uA} F(u, V_u) \, \mathrm{d}u\right\|^a\right] \\
&\leq C_a \mathbb{E}\left[\left\|M_t^{(\varepsilon)} - M_s^{(\varepsilon)}\right\|^a\right] + C_a \varepsilon^{1-\frac{a}{2}} (t-s)^{a-1} \mathbb{E}\left[\int_{s/\varepsilon}^{t/\varepsilon} \|F(u, V_u)\|^a \, \mathrm{d}u\right] \\
&\leq C_a \mathbb{E}\left[\left(\mathrm{Tr}\left(\left\langle M^{(\varepsilon)} - M_s^{(\varepsilon)}\right\rangle_t\right)\right)^{a/2}\right] + C_a \varepsilon^{1-\frac{a}{2}} (t-s)^{a-1} \int_{s/\varepsilon}^{t/\varepsilon} u^{\frac{\gamma a}{2} - \beta a} \, \mathrm{d}u \\
&\leq C_a (t-s)^{\frac{a}{2}} + C_{a,m,M} \varepsilon^{a(\beta - \frac{\gamma+1}{2})} (t-s)^{a-1} \\
&\leq C_{a,m,M} (t-s)^{\frac{a}{2}}.
\end{aligned}$$

Since $\frac{a}{2} > 2$ and $\beta \geq \frac{\gamma+1}{2}$ the upper bound is independent of $\varepsilon \leq 1$. Furthermore, by moment estimates (Proposition 10.8),

$$\sup_{\varepsilon\leq\varepsilon_0} \mathbb{E}\left[\left\|Y_m^{(\varepsilon)}\right\|\right] \leq \sqrt{m} < \infty.$$

Thus, Kolmogorov's criterion can be applied, proving the tightness result. $\square$

We will now prove the convergence of the finite-dimensional distributions of $Y^{(\varepsilon)}$. Thanks to the previous lemma, this will yield the weak convergence on every compact set (see Theorem 13.1 p. 139 in [Bil99]). The convergence in distribution on the whole space $\mathcal{C}$ will follow, for $2q \geq 1$, from Theorem 16.7 p. 174 in [Bil99], since all processes considered are continuous.

Convergence of the f.d.d. in the super-critical regime

Assume here that $2q > 1$. Recall that $(Y_t^{(\varepsilon)})_{t\geq\varepsilon t_0} := (\sqrt{\varepsilon}Y_{t/\varepsilon})_{t\geq\varepsilon t_0}$.

Proof of Theorem 10.1 (i).

Step 1. We first prove the convergence of the f.d.d. of the local martingale term $M^{(\varepsilon)} =: (M^{(\varepsilon,1)}, M^{(\varepsilon,2)})^T$ appearing in (SDE_Y).
Recall that the stochastic integral $M^{(\varepsilon)}$ was defined in (10.8). It is a centered Gaussian process with covariance kernel defined, for any $(s,t) \in [\varepsilon t_0, +\infty)^2$, by

$$K^{(\varepsilon)}(s,t) := \begin{pmatrix} \mathrm{Cov}(M_s^{(\varepsilon)}) & \mathrm{Cov}(M_s^{(\varepsilon)}, M_t^{(\varepsilon)}) \\ \mathrm{Cov}(M_t^{(\varepsilon)}, M_s^{(\varepsilon)}) & \mathrm{Cov}(M_t^{(\varepsilon)}) \end{pmatrix},$$

where

$$\mathrm{Cov}(M_s^{(\varepsilon)}, M_t^{(\varepsilon)}) = \begin{pmatrix} \mathrm{Cov}(M_s^{(\varepsilon,1)}, M_t^{(\varepsilon,1)}) & \mathrm{Cov}(M_s^{(\varepsilon,1)}, M_t^{(\varepsilon,2)}) \\ \mathrm{Cov}(M_s^{(\varepsilon,2)}, M_t^{(\varepsilon,1)}) & \mathrm{Cov}(M_s^{(\varepsilon,2)}, M_t^{(\varepsilon,2)}) \end{pmatrix},$$

and $\mathrm{Cov}(M_s^{(\varepsilon)}) = \mathrm{Cov}(M_s^{(\varepsilon)}, M_s^{(\varepsilon)})$. Thus, the convergence of the f.d.d. of $M^{(\varepsilon)}$ is reduced to the study of the limit of $K^{(\varepsilon)}$, when ε converges to 0. Let us fix $\varepsilon t_0 \leq s \leq t$. Using that $M^{(\varepsilon)}$ has independent increments and by Itô's isometry, we find that

$$\mathrm{Cov}(M_s^{(\varepsilon)}, M_t^{(\varepsilon)}) = \begin{pmatrix} \varepsilon \int_{t_0}^{s/\varepsilon} \sin(u)^2 \,\mathrm{d}u & -\varepsilon \int_{t_0}^{s/\varepsilon} \sin(u)\cos(u)\,\mathrm{d}u \\ -\varepsilon \int_{t_0}^{s/\varepsilon} \sin(u)\cos(u)\,\mathrm{d}u & \varepsilon \int_{t_0}^{s/\varepsilon} \cos(u)^2\,\mathrm{d}u \end{pmatrix}.$$

We get that, for all $0 < s \leq t$,

$$\mathrm{Cov}(M_s^{(\varepsilon)}, M_t^{(\varepsilon)}) \underset{\varepsilon\to 0}{\longrightarrow} \frac{1}{2}\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}.$$

We recognize the covariance kernel of the process $\left(\mathcal{B}_{\frac{t}{2}}\right)_{t>0}$, where $\mathcal{B}$ denotes a standard Brownian motion on $\mathbb{R}^2$. Since mean and covariance functions characterize Gaussian process (see Lemma 13.1 *(i)* p. 250 in [Kal02]), we have thus proved that $(M_t^{(\varepsilon)})_{t\geq\varepsilon t_0}$ converges in f.d.d. towards $\left(\mathcal{B}_{\frac{t}{2}}\right)_{t>0}$.
Step 2. Pick $T > 0$. We prove that

$$\mathbb{E}\left[\sup_{\varepsilon t_0\leq t\leq T} \left\| Y_t^{(\varepsilon)} - M_t^{(\varepsilon)} \right\|\right] \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

Let us fix $\varepsilon > 0$ small enough such that $\varepsilon t_0 \leq T$. We have

$$\sup_{\varepsilon t_0\leq t\leq T} \left\| Y_t^{(\varepsilon)} - M_t^{(\varepsilon)} \right\| \leq \sqrt{\varepsilon}\,\|z_0\| + \sqrt{\varepsilon}\int_{t_0}^{T/\varepsilon} \left\| e^{-sA}F(s,V_s)\right\| \mathrm{d}s.$$

We use moment estimates (Proposition 10.8) to get

$$\begin{aligned}\mathbb{E}\left[\sqrt{\varepsilon}\int_{t_0}^{T/\varepsilon}\left\|e^{-sA}F(s,V_s)\right\|\mathrm{d}s\right] &= \mathbb{E}\left[\sqrt{\varepsilon}\int_{t_0}^{T/\varepsilon}\|F(s,V_s)\|\,\mathrm{d}s\right] \\ &\le \mathbb{E}\left[\sqrt{\varepsilon}\int_{t_0}^{T/\varepsilon}|V_s|^{\gamma}\,s^{-\beta}\,\mathrm{d}s\right] \\ &\le \sqrt{\varepsilon}C_{\gamma,t_0}\int_{t_0}^{T/\varepsilon}s^{\frac{\gamma}{2}-\beta}\,\mathrm{d}s \\ &\le C_{\gamma,t_0}(\varepsilon^{\beta-\frac{\gamma+1}{2}}T^{\frac{\gamma}{2}-\beta+1}+\sqrt{\varepsilon}t_0^{\frac{\gamma}{2}-\beta+1}). \end{aligned} \tag{10.9}$$

Hence, setting $r := \min(\beta - \frac{\gamma+1}{2}, \frac{1}{2})$, which is positive by assumption, we get

$$\mathbb{E}\left[\sup_{\varepsilon t_0\le t\le T}\left\|Y_t^{(\varepsilon)}-M_t^{(\varepsilon)}\right\|\right] = \underset{\varepsilon\to 0}{O}(\varepsilon^r).$$

We conclude the proof using Theorem 3.1 p. 27 in [Bil99].

□

Convergence of the f.d.d. in the critical and sub-critical regimes

In this section, we consider the linear case, i.e. $\gamma = 1$. Pick $\beta \in \left(\frac{1}{2}, 1\right]$. Recall that

$$(Y_t^{(\varepsilon)})_{t\ge\varepsilon t_0} := (\varepsilon^q Y_{t/\varepsilon})_{t\ge\varepsilon t_0},$$

where $q = \frac{\beta}{\gamma+1}$.

Proof of Theorem 10.1 (ii) and (iii). Leaving out the Brownian term, the underlying ODE of our system is the following

$$x''(t) + \frac{x'(t)}{t^{\beta}} + x(t) = 0, \quad t \ge t_0. \tag{10.10}$$

Pick the basis of solutions given in Lemma 10.20 and denote by R its resolvent matrix which is the solution to

$$R_t' = \begin{pmatrix} 0 & 1 \\ -1 & -\frac{1}{t^{\beta}}\end{pmatrix} R_t, \quad t \ge t_0.$$

It follows by differentiating the inverse function that

$$(R^{-1})_t' = -R_t^{-1}\begin{pmatrix} 0 & 1 \\ -1 & -\frac{1}{t^{\beta}}\end{pmatrix}, \quad t \ge t_0.$$

Using Itô's product rule for $R_t^{-1}Z_t$, the fact that the quadratic covariation of $(R_t^{-1})_{t\ge t_0}$ and $(Z_t)_{t\ge t_0}$ is equal to zero and (10.4) with $\gamma = 1$, we get that for all $t \ge t_0$,

$$\begin{aligned} R_t^{-1}Z_t &= R_{t_0}^{-1}Z_{t_0} + \int_{t_0}^{t} R_s^{-1}\,\mathrm{d}Z_s + \int_{t_0}^{t}(R^{-1})_s' Z_s\,\mathrm{d}s \\ &= R_{t_0}^{-1}Z_{t_0} + \int_{t_0}^{t} R_s^{-1}\Gamma\,\mathrm{d}W_s. \end{aligned}$$

Let us define f the rate of decrease of R (see Lemma 10.20) by

$$\forall t > 0, \ f(t) := \begin{cases} \frac{1}{\sqrt{t}} & \text{if } \beta = 1, \\ \exp\left(-\frac{t^{1-\beta}}{2(1-\beta)}\right) & \text{else.} \end{cases} \tag{10.11}$$

Set, for $t \geq \varepsilon t_0$,

$$\Phi_t := \frac{e^{-tA} R_t}{f(t)} \quad \text{and} \quad \widetilde{M}_t^{(\varepsilon)} := \varepsilon^q f\left(\frac{t}{\varepsilon}\right) \int_{t_0}^{t/\varepsilon} R_s^{-1} \Gamma \, \mathrm{d}W_s. \tag{10.12}$$

Pick $t \geq \varepsilon t_0$. To study the convergence of $Y^{(\varepsilon)}$ we decompose it into

$$Y_t^{(\varepsilon)} = \varepsilon^q f\left(\frac{t}{\varepsilon}\right) \Phi_{t/\varepsilon} R_{t_0}^{-1} Z_0 + \Phi_{t/\varepsilon} \widetilde{M}_t^{(\varepsilon)}. \tag{10.13}$$

Step 1. Convergence of Φ and simplification the problem.

Using the asymptotic expansion of the resolvent matrix (Lemma 10.20), we can write, for $t \geq \varepsilon t_0$,

$$\Phi_t = I_2 + \underset{t \to \infty}{O}\left(t^{1-2\beta}\right).$$

As a consequence, since $1 - 2\beta < 0$, $\Phi_{t/\varepsilon}$ converges to the identity matrix I_2, as $\varepsilon \to 0$ and for any $t > 0$. Let us notice that, for any $t > 0$, $\varepsilon^q f\left(\frac{t}{\varepsilon}\right)$ converges to 0, as $\varepsilon \to 0$. We thus obtain that

$$\varepsilon^q f\left(\frac{t}{\varepsilon}\right) \Phi_{t/\varepsilon} R_{t_0}^{-1} Z_0 \underset{\varepsilon \to 0}{\longrightarrow} 0.$$

Therefore, we can forget the first term appearing in the decomposition (10.13) of $Y^{(\varepsilon)}$ (see Theorem 3.1 p. 27 in [Bil99]). It is thus enough to prove the convergence of the f.d.d. of the centered Gaussian process $(\Phi_{t/\varepsilon} \widetilde{M}_t^{(\varepsilon)})_t$ to deduce the convergence of the f.d.d. of $Y^{(\varepsilon)}$ towards the same limit.

Step 2. Computation of the covariance kernel of $(\Phi_{t/\varepsilon} \widetilde{M}_t^{(\varepsilon)})_t$ and convergence in f.d.d.

We have, for all $\varepsilon t_0 \leq s \leq t < +\infty$,

$$\operatorname{Cov}\left(\Phi_{s/\varepsilon} \widetilde{M}_s^{(\varepsilon)}, \Phi_{t/\varepsilon} \widetilde{M}_t^{(\varepsilon)}\right) = \Phi_{s/\varepsilon} \operatorname{Cov}(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)}) \Phi_{t/\varepsilon}^T. \tag{10.14}$$

Using the expression of the Wronskian obtained in Lemma 10.20, we obtain, for all $t \geq \varepsilon t_0$,

$$\widetilde{M}_t^{(\varepsilon)} = \varepsilon^q f(t/\varepsilon) \int_{t_0}^{t/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}B_u.$$

It is a centered Gaussian process and for any $\varepsilon t_0 \leq s \leq t$, we have

$$\operatorname{Cov}(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)}) = \varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-4} \begin{pmatrix} y_2^2(u) & -y_2(u) y_1(u) \\ -y_2(u) y_1(u) & y_1^2(u) \end{pmatrix} \mathrm{d}u.$$

Using the asymptotic expansion of the solutions and Lemma 10.23, we get, for all $\varepsilon t_0 < s \leq t$,

$$\operatorname{Cov}(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)}) = \varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-2} \begin{pmatrix} \sin^2(u) & -\sin(u)\cos(u) \\ -\sin(u)\cos(u) & \cos^2(u) \end{pmatrix} \mathrm{d}u \\ + \underset{\varepsilon\to 0}{O}\left(\varepsilon^{2\beta-1} f(t/\varepsilon) f(s/\varepsilon)^{-1}\right).$$

Moreover, using asymptotic expansions of these integrals (see Lemmas 10.22 and 10.23),

$$\varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-2} \cos^2(u)\,\mathrm{d}u = \varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \frac{1}{2} \int_{t_0}^{s/\varepsilon} f(u)^{-2}\,\mathrm{d}u \\ + \underset{\varepsilon\to 0}{o}\left(f(t/\varepsilon) f(s/\varepsilon)^{-1}\right).$$

The same equality holds for

$$\varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-2} \sin^2(u)\,\mathrm{d}u,$$

and we have

$$\varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-2} \cos(u)\sin(u)\,\mathrm{d}u = \underset{\varepsilon\to 0}{o}\left(f(t/\varepsilon) f(s/\varepsilon)^{-1}\right).$$

Thanks to Lemma 10.23, this leads to

$$\begin{aligned}\operatorname{Cov}(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)}) &= \left[\frac{1}{2}\varepsilon^\beta f(t/\varepsilon) f(s/\varepsilon) \int_{t_0}^{s/\varepsilon} f(u)^{-2}\,\mathrm{d}u\right] I_2 + \underset{\varepsilon\to 0}{o}\left(f(t/\varepsilon) f(s/\varepsilon)^{-1}\right) \\ &= k_\beta \frac{f(t/\varepsilon)}{f(s/\varepsilon)} s^\beta I_2 + \underset{\varepsilon\to 0}{o}\left(f(t/\varepsilon) f(s/\varepsilon)^{-1}\right),\end{aligned}$$

where

$$k_\beta := \begin{cases} \frac{1}{4} & \text{if } \beta = 1, \\ \frac{1}{2} & \text{else.} \end{cases}$$

It follows from the definition of f given in (10.11) and the fact that $s \le t$ that

$$\operatorname{Cov}(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)}) \underset{\varepsilon\to 0}{\longrightarrow} \begin{cases} \frac{1}{4}\frac{(s\wedge t)^2}{\sqrt{st}} I_2 & \text{if } \beta = 1, \\ \frac{1}{2} s^\beta \mathbb{1}_{\{s=t\}} I_2 & \text{else.} \end{cases}$$

Using the preceding convergence, (10.14) and Step 1, we have proved the convergence of the f.d.d. of $Y^{(\varepsilon)}$. Note that whenever $2q = 1$, i.e. $\beta = 1$ since $\gamma = 1$, we recognize the covariance kernel of the process $\left(\frac{1}{\sqrt{2t}} \int_0^t \sqrt{s}\,\mathrm{d}\mathcal{B}_s\right)_{t>0}$, where $\mathcal{B}$ denotes a standard Brownian motion on $\mathbb{R}^2$.

$\square$

Remark 10.11.

- The proof relies on the asymptotic expansion of the resolvent matrix of (10.10). We were able to prove it only for $\beta \in \left(\frac{1}{2}, 1\right]$. However, if $\beta = 0$, the resolvent matrix is explicit and following the same lines, we can prove that $\left(Z_{t/\varepsilon}\right)_{t\ge\varepsilon t_0}$ converges in f.d.d. towards a centered Gaussian process with covariance kernel $(s,t) \mapsto \frac{1}{2} I_2 \mathbb{1}_{\{s=t\}}$. This behavior can be explained by the fact that the frictional force does not decrease along time. This cancels somehow the rotation bearing, which prevents $Z^{(\varepsilon)}$ from converging as a process when $\beta > 0$.

- The asymptotic expansion of the resolvant matrix is also known in the super-critical regime, i.e. $\beta > 1$. Therefore, one can prove the result in the linear case, i.e. $\gamma = 1$, following the same lines.

10.2.4 Proof of Corollary 10.6

Proof of Corollary 10.6. We start by proving the convergence in distribution of $r_{1/T} Z_T$, as $T \to +\infty$. We claim that it follows from Theorem 10.1. Indeed, it is enough to remark that the convergence results stated in Theorem 10.1 imply the convergence in distribution of the marginal distribution at time $t = 1$ of $Y^{(\varepsilon)}$. Let us also recall that $Z_T = e^{TA} Y_T$. Setting $T = \frac{1}{\varepsilon}$, the convergence of $r_{1/T} Z_T$ is therefore a direct consequence of Lemma 10.24.
We now show that the rescaled process $Z^{(\varepsilon)}$ does not converge in distribution. We do the proof only in the super-critical regime. Assume by contradiction that it is the case. Hence, each of its coordinates shall converge too. We thus have the convergence of the rescaled process $X^{(\varepsilon)}$. Using (SDE_Y), we can write

$$\sqrt{\varepsilon} X_{t/\varepsilon} = \sqrt{\varepsilon} x_0 + \sqrt{\varepsilon} \int_{t_0}^{t/\varepsilon} \sin\left(\frac{t}{\varepsilon} - s\right) \mathrm{d}B_s - \sqrt{\varepsilon} \int_{t_0}^{t/\varepsilon} \sin\left(\frac{t}{\varepsilon} - s\right) F(s, V_s)\, \mathrm{d}s.$$

As in the proof of Theorem 10.1 (i), the last term converges in probability uniformly on compact intervals towards zero. Hence, the following term shall converge in distribution

$$I_t^{(\varepsilon)} := \sqrt{\varepsilon} \int_{t_0}^{t/\varepsilon} \sin\left(\frac{t}{\varepsilon} - s\right) \mathrm{d}B_s.$$

The process $(I_t^{(\varepsilon)})_{t \geq \varepsilon t_0}$ is Gaussian, thereby its limit shall be Gaussian too and its covariance function shall converge (see Lemma 13.1 *(i)* p. 250 in [Kal02]). However, using Itô's isometry, one can compute, for $\varepsilon t_0 \leq s \leq t$,

$$\begin{aligned}
\mathbb{E}\left[I_t^{(\varepsilon)} I_s^{(\varepsilon)}\right] &= \varepsilon \int_{t_0}^{s/\varepsilon} \sin\left(\frac{t}{\varepsilon} - u\right) \sin\left(\frac{s}{\varepsilon} - u\right) \mathrm{d}u \\
&= \varepsilon \frac{1}{2} \left[\cos\left(\frac{t-s}{\varepsilon}\right)\left(\frac{s}{\varepsilon} - t_0\right) + \frac{1}{2}\left(\sin\left(\frac{t-s}{\varepsilon}\right) - \sin\left(\frac{t+s}{\varepsilon} - 2t_0\right)\right)\right] \\
&= \frac{1}{2} s \cos\left(\frac{t-s}{\varepsilon}\right) + \underset{\varepsilon \to 0}{o}(1).
\end{aligned}$$

This term does not converge if $s \neq t$, and that concludes the proof. □

10.3 Study of the system driven by an α-stable process

In this section, L is a symmetric α-stable Lévy process. We call ν its Lévy measure, which can be written as $\nu(\mathrm{d}z) = a\, |z|^{-1-\alpha} \mathbb{1}_{\{z \neq 0\}}\, \mathrm{d}z$ with $a > 0$. As a Lévy measure, it satisfies $\int_{\mathbb{R}^*} (1 \wedge z^2) \nu(\mathrm{d}z) < +\infty$. We denote by N the Poisson random measure associated with L and by $\widetilde{N}$ its compensated Poisson measure. Using Lévy-Itô's decomposition (see [Sat99, Remark 14.6 and Theorem 14.7 iii)]), we have, for all $t \geq 0$,

$$L_t = \int_0^t \int_{\mathbb{R}^*} z \widetilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{10.15}$$

As in the previous section, we set, for all $t \geq t_0$ and $v \in \mathbb{R}$,

$$Z_t := \begin{pmatrix} X_t \\ V_t \end{pmatrix}, \ S_t := \begin{pmatrix} 0 \\ L_t \end{pmatrix}, \ A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ \Gamma := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } F(t,v) := \begin{pmatrix} 0 \\ \operatorname{sgn}(v)\dfrac{|v|^\gamma}{t^\beta} \end{pmatrix}.$$

Thereby, the system (SKE) can be rewritten as

$$\begin{cases} \mathrm{d}Z_t = \Gamma\,\mathrm{d}S_t + AZ_t\,\mathrm{d}t - F(t,V_t)\,\mathrm{d}t, \\ Z_{t_0} = z_0 := (x_0,v_0)^T. \end{cases} \tag{10.16}$$

We define, for any $t \geq t_0$, $Y_t := e^{-tA}Z_t$. We easily check, with Itô's formula, that Y is given by

$$\mathrm{d}Y_t = e^{-tA}\Gamma\,\mathrm{d}S_t - e^{-tA}F(t,V_t)\,\mathrm{d}t. \tag{SDE_Y}$$

10.3.1 Existence up to explosion

Theorem 10.12. *The system* (SKE) *admits a unique weak solution if $\gamma \in (0,1]$. If $\gamma > 1$, there exists a unique strong solution defined up to its explosion time τ_∞.*

Proof. In the case $\gamma > 1$, the coefficients of the SDE (10.16) satisfied by $Z = (X,V)$ are locally Lipschitz continuous with respect to the space variable, uniformly in time. So we can apply Theorem 6.2.11 in [App09] (see Theorem 119 in [Sit05] for a detailed proof), which ensures the existence of a unique solution to (10.16) defined up to explosion.

Assume now that $\gamma \leq 1$. We check that we can use Theorem 1 in [MM21]. Using the same notations, we have

$$\widetilde{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

and for any $(t,x_1,x_2) \in [t_0,+\infty) \times \mathbb{R}^2$, $F_1(t,x_1,x_2) = -\operatorname{sgn}(x_1)|x_1|^\gamma t^{-\beta}$, $F_2(t,x_1,x_2) = 0$ and $\sigma(t,x_1,x_2) = 1$. Assumptions **(UE)** and **(ND)** are clearly satisfied. Since F_2 does not depend on x_1 and since $[\widetilde{A}]_{2,1} = 1$ is different from 0, we deduce that Assumption **(H)** is satisfied. We easily check that Theorem 1 in [MM21] can be applied with $\beta_1 = \gamma$, and $\beta_2 = 1$.

□

Remark 10.13. For $\alpha \in (0,2)$, employing the technique of Picard iteration and the interlacing procedure, one can deduce that (10.16) has a unique solution in the linear setting $\gamma = 1$ (see [App09, p. 375]).

10.3.2 Moment estimates and non-explosion

Let Z be the unique solution up to explosion time to (10.16). As in the continuous setting, define, for all $r \geq 0$, the stopping time

$$\tau_r := \inf\{t \geq t_0,\ \|Z_t\| \geq r\}.$$

Set $\tau_\infty := \lim_{r\to+\infty}\tau_r$ the explosion time of Z. For the sake of simplicity, since there is no jump on the position component, for $z \in \mathbb{R}$, we shall write $Z_{s-} + z$ for $(X_s, V_{s-} + z)$ in the following.
We adapt the proof of [GL21b] to two-dimensional processes.

Proposition 10.14. *For any $\gamma \geq 0$ and $\beta \geq 0$, the explosion time τ_∞ is a.s. infinite and for $\kappa \in (0, \alpha)$, there exists C_{κ,t_0} such that*

$$\forall t \geq t_0, \ \mathbb{E}\left[\|Z_t\|^\kappa\right] \leq C_{\kappa,t_0} t^{\frac{\kappa}{\alpha}}. \tag{10.17}$$

Remark 10.15. Note that, as in the Brownian case, the moment estimates obtained for the position process X is a priori smaller in our case than in the free potential case [GL21b]. It is explained by the confining effect of the quadratic potential.

Proof. The key idea is to slice the small and big jumps in a non-homogeneous way with respect to the characteristic scale of an α-stable process $\xi \mapsto \xi^{\frac{1}{\alpha}}$.

Pick $\xi \geq t_0$. Using (10.15), the α-stable symmetric Lévy driving process can be written as

$$L_t - L_{t_0} = \int_{t_0}^t \int_{|z| \leq \xi^{\frac{1}{\alpha}}} z \widetilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_{t_0}^t \int_{|z| > \xi^{\frac{1}{\alpha}}} z N(\mathrm{d}s, \mathrm{d}z).$$

Indeed, the term

$$\int_{t_0}^t \int_{|z| > \xi^{\frac{1}{\alpha}}} z \nu(\mathrm{d}z) \, \mathrm{d}s$$

is equal to 0 since ν is a symmetric measure.

Step 1. We first apply Itô's formula (see Theorem 4.4.7 p. 251 in [App09]) and estimate the expectation of each term for $\kappa \leq 1$, in order to get (10.17).

Fix $\eta > 0$ to be chosen latter and define the $\mathcal{C}^2$-function $f : (x,v) \mapsto (\eta + x^2 + v^2)^{\kappa/2}$. We use the fact that for all $y \in \mathbb{R}^2$, $y \cdot Ay = 0$, and observe that for any $s \geq t_0$ and $(x,v) \in \mathbb{R}^2$, $(x,v)^T \cdot F(s,v) = |v|^{\gamma+1} s^{-\beta} \geq 0$. In the sequel, we write, for simplicity, $f(Z_{s^-} + z)$ for $f(X_{s^-}, V_{s^-} + z)$. For all $t \geq t_0$, by Itô's formula, we have

$$f(Z_{t \wedge \tau_r}) \leq f(z_0) + M_t + R_t + S_t,$$

where

$$M_t := \int_{t_0}^t \int_{0 < |z| < \xi^{\frac{1}{\alpha}}} \mathbb{1}_{\{s \leq \tau_r\}} (f(Z_{s-} + z) - f(Z_{s-})) \widetilde{N}(\mathrm{d}s, \mathrm{d}z),$$

$$R_t := \int_{t_0}^t \int_{|z| \geq \xi^{\frac{1}{\alpha}}} \mathbb{1}_{\{s \leq \tau_r\}} (f(Z_{s-} + z) - f(Z_{s-})) N(\mathrm{d}s, \mathrm{d}z), \tag{10.18}$$

$$S_t := \int_{t_0}^t \int_{0 < |z| < \xi^{\frac{1}{\alpha}}} \mathbb{1}_{\{s \leq \tau_r\}} \left[f(Z_{s-} + z) - f(Z_{s-}) - \nabla f(Z_{s-}).z\right] \nu(\mathrm{d}z) \, \mathrm{d}s. \tag{10.19}$$

Moreover, remark that for all $k > \alpha$,

$$\int_{0 < |z| < \xi^{\frac{1}{\alpha}}} |z|^k \, \nu(\mathrm{d}z) = \frac{2a}{k - \alpha} \xi^{\frac{k}{\alpha} - 1}, \tag{10.20}$$

and for all $k < \alpha$,

$$\int_{|z| \geq \xi^{\frac{1}{\alpha}}} |z|^k \, \nu(\mathrm{d}z) = \frac{2a}{\alpha - k} \xi^{\frac{k}{\alpha} - 1}. \tag{10.21}$$

We estimate expectations of M, R and S.
To that end, we first show that the local martingale $(M_t)_{t \geq t_0}$ is a martingale. Fix $q \geq 2$ and $r \geq 0$.

Moreover, we set

$$I_t(q) := \int_{t_0}^{t} \int_{0<|z|<\xi^{\frac{1}{\alpha}}} \mathbb{1}_{\{s\le\tau_r\}} \left|f(Z_{s-}+z) - f(Z_{s-})\right|^q \nu(\mathrm{d}z)\,\mathrm{d}s.$$

Thanks to Taylor-Lagrange inequality, for all $\|(x,v)\| \le r$ and $|z| \le \xi^{\frac{1}{\alpha}}$,

$$|f(x,v+z) - f(x,v)| \le \sup\{\|\nabla f(y)\|\,,\ \|y\| \in [0, r+\xi^{1/\alpha}]\}\,|z| \le C_{r,\xi,\kappa}\,|z|\,,$$

so we have

$$I_t(q) \le C_{r,\xi,\kappa} \int_{t_0}^{t} \int_{0<|z|<\xi^{\frac{1}{\alpha}}} \mathbb{1}_{\{s\le\tau_r\}}\,|z|^q\,\nu(\mathrm{d}z)\,\mathrm{d}s.$$

Hence, it is a finite quantity, since $q \ge 2$ and (10.20) holds. Therefore, for $q \ge 2$, by Kunita's inequality (see Theorem 4.4.23 p. 265 in [App09]), there exists $D_q > 0$ such that

$$\mathbb{E}\left[\sup_{t_0\le s\le t} |M_s|^q\right] \le D_q\left(\mathbb{E}\left[I_t(2)^{\frac{q}{2}}\right] + \mathbb{E}\left[I_t(q)\right]\right) < +\infty.$$

Hence, by Theorem 51 p. 38 in [Pro05], M is a martingale.
We estimate now the finite variation part S defined in (10.19). Note that for all $(x,v) \in \mathbb{R}^2$, the Hessian matrix of f is given by

$$\mathrm{Hess}(f)(x,v) = \kappa(x^2+v^2+\eta)^{\frac{\kappa}{2}-1}\begin{pmatrix} 1+(\kappa-2)\dfrac{x^2}{x^2+v^2+\eta} & (\kappa-2)\dfrac{xv}{x^2+v^2+\eta} \\ (\kappa-2)\dfrac{xv}{x^2+v^2+\eta} & 1+(\kappa-2)\dfrac{v^2}{x^2+v^2+\eta}\end{pmatrix}.$$

Its matrix norm is bounded by $C_\kappa \eta^{\frac{\kappa}{2}-1}$.
Assume that $|z| < \xi^{\frac{1}{\alpha}}$. Using Taylor- Lagrange's inequality and injecting (10.20) we get the almost sure following bound, for all $s \ge t_0$,

$$\left|\int_{0<|z|<\xi^{\frac{1}{\alpha}}} \left(f(Z_{s-}+z) - f(Z_{s-}) - \nabla f(Z_{s-})\cdot z\right)\nu(\mathrm{d}z)\right| \le C_\kappa \eta^{\frac{\kappa}{2}-1}\frac{2a}{2-\alpha}\xi^{\frac{2}{\alpha}-1}. \tag{10.22}$$

It remains to study the Poisson integral R defined in (10.18). Recall that $\kappa \le 1$. Let us note that for all $x, v, z \in \mathbb{R}$, by Hölder property of power functions, one has

$$\begin{aligned}|f(x,v+z) - f(x,v)| &\le \left|\left(\eta+x^2+(v+z)^2\right)^{\frac{\kappa}{2}} - \left(x^2+(v+z)^2\right)^{\frac{\kappa}{2}}\right| \\ &\quad + \left|\left(x^2+(v+z)^2\right)^{\frac{\kappa}{2}} - \left(x^2+v^2\right)^{\frac{\kappa}{2}}\right| + \left|\left(x^2+v^2\right)^{\frac{\kappa}{2}} - \left(\eta+x^2+v^2\right)^{\frac{\kappa}{2}}\right| \\ &\le 2\eta^{\frac{\kappa}{2}} + \left|\|(x,v+z)\|^\kappa - \|(x,v)\|^\kappa\right| \\ &\le 2\eta^{\frac{\kappa}{2}} + |z|^\kappa\,.\end{aligned}$$

Injecting (10.21), we deduce that

$$\int_{|z|\ge\xi^{\frac{1}{\alpha}}} |f(Z_{s-}+z) - f(Z_{s-})|\,\nu(\mathrm{d}z) \le \eta^{\frac{\kappa}{2}}\frac{2a}{\alpha}\xi^{-1} + \frac{2a}{\alpha-\kappa}\xi^{\frac{\kappa}{\alpha}-1}. \tag{10.23}$$

Moment estimate of the Poisson integral follows from Theorem 2.3.7 p. 106 in [App09].
Gathering (10.23) and (10.22), we obtain

$$\mathbb{E}\left[\|Z_{t\wedge\tau_r}\|^{\kappa}\right] \quad\leq\quad \mathbb{E}\left[f(Z_{t\wedge\tau_r})\right] \quad\leq\quad \mathbb{E}\left[f(Z_{t_0})\right] \;+\; t\xi^{-1}\left(\eta^{\kappa/2}\frac{2a}{\alpha}+\frac{2a}{\alpha-\kappa}\xi^{\frac{\kappa}{\alpha}}+C_\kappa\eta^{\frac{\kappa}{2}-1}\frac{2a}{2-\alpha}\xi^{\frac{2}{\alpha}}\right).$$

Choosing $\eta=t^{\frac{2}{\alpha}}$ and $\xi=t$, we get

$$\mathbb{E}\left[\|Z_{t\wedge\tau_r}\|^{\kappa}\right]\leq \mathbb{E}\left[f(Z_{t_0})\right]+t^{\frac{\kappa}{\alpha}}\left(\frac{2a}{\alpha}+\frac{2a}{\alpha-\kappa}+C_\kappa\frac{2a}{2-\alpha}\right)\leq C_{\kappa,t_0}t^{\frac{\kappa}{\alpha}}.$$

Thanks to Lemma 10.21, we can conclude that the explosion time of Z is a.s. infinite, and letting $r\to+\infty$ with Fatou's lemma, for all $\kappa\in[0,1]$,

$$\mathbb{E}\left[\|Z_t\|^{\kappa}\right]\leq C_{\kappa,t_0}t^{\frac{\kappa}{\alpha}}. \tag{10.24}$$

Step 2. Pick $\kappa\in(1,\alpha)$. We estimate R in another way, using again Theorem 2.3.7 p. 106 in [App09].

By the Hölder property of power function and (10.21), we get

$$\begin{aligned}\int_{|z|\geq\xi^{\frac{1}{\alpha}}}|f(Z_{s-}+z)-f(Z_{s-})|\,\nu(\mathrm{d}z)&\leq\int_{|z|\geq\xi^{\frac{1}{\alpha}}}\left|2zV_{s-}+z^2\right|^{\frac{\kappa}{2}}\nu(\mathrm{d}z)\\&\leq C_\kappa\left(\frac{2a}{\alpha-\kappa}\xi^{\frac{\kappa}{\alpha}-1}+|V_{s-}|^{\frac{\kappa}{2}}\frac{2a}{\alpha-\frac{\kappa}{2}}\xi^{\frac{\kappa}{2\alpha}-1}\right).\end{aligned}\tag{10.25}$$

Gathering (10.22) and (10.25), one has

$$\begin{aligned}\mathbb{E}\left[\|Z_{t\wedge\tau_r}\|^{\kappa}\right]\leq\mathbb{E}\left[f(Z_{t_0})\right]+t&\left(C_\kappa\frac{2a}{\alpha-\kappa}\xi^{\frac{\kappa}{\alpha}-1}+C_\kappa\eta^{\frac{\kappa}{2}-1}\frac{2a}{2-\alpha}\xi^{\frac{2}{\alpha}-1}\right)\\&+C_\kappa\frac{2a}{\alpha-\frac{\kappa}{2}}\xi^{\frac{\kappa}{2\alpha}-1}\int_{t_0}^{t}\mathbb{E}\left[|V_s|^{\frac{\kappa}{2}}\right]\mathrm{d}s.\end{aligned}$$

Injecting (10.24) applied with $\frac{\kappa}{2}$, choosing $\eta=t^{\frac{2}{\alpha}}$ and $\xi=t$, we get

$$\mathbb{E}\left[\|Z_{t\wedge\tau_r}\|^{\kappa}\right]\leq C_{\kappa,t_0,\alpha}t^{\frac{\kappa}{\alpha}}.$$

The conclusion of the proof follows, letting $r\to+\infty$.

□

10.3.3 Asymptotic behavior of the solution

We gather in this section the proof of Theorem 10.2. The strategy is to prove the convergence of the f.d.d. of the process $Y^{(\varepsilon)}$, and then its tightness both in the super-critical and critical regimes. We first prove the tightness when $\alpha q\geq 1$. Recall that $q=\frac{\beta}{\gamma+\alpha-1}$.

Lemma 10.16. *Assume that $\alpha q\geq 1$, then the family $\left\{(\varepsilon^{\frac{1}{\alpha}}Y_{t/\varepsilon})_{t\geq\varepsilon t_0},\ \varepsilon>0\right\}$ is tight on every compact interval $[m,M]$, for $0<m\leq M$.*

Proof. We check the Aldous's tightness criterion stated in Theorem 16.10 p. 178 in [Bil99]. Let a, η, T be positive reals. Let τ be a discrete stopping time with finite range $\mathcal{T}$, bounded by T and fix $\delta > 0$ and $\varepsilon > 0$ small enough to be chosen later. Let us define the Wiener-Lévy integral appearing in (SDE_Y) $M^{(\varepsilon)}$ by, for all $t \geq \varepsilon t_0$,

$$\begin{aligned} M_t^{(\varepsilon)} &:= \varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{t/\varepsilon} e^{-sA}\Gamma \, \mathrm{d}S_s \\ &= \varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{t/\varepsilon} \begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix} \mathrm{d}L_s \\ &= \varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{t/\varepsilon} \int_{\mathbb{R}} \begin{pmatrix} -\sin(s) \\ \cos(s) \end{pmatrix} z\widetilde{N}(\mathrm{d}s, \mathrm{d}z), \end{aligned} \tag{10.26}$$

using the representation (10.15) of L. Notice that it is a martingale. By the triangle inequality, one has

$$\mathbb{E}\left[\left\|Y_{\tau+\delta}^{(\varepsilon)} - Y_\tau^{(\varepsilon)}\right\|\right] \leq \mathbb{E}\left[\left\|M_{\tau+\delta}^{(\varepsilon)} - M_\tau^{(\varepsilon)}\right\|\right] + \mathbb{E}\left[\varepsilon^{\frac{1}{\alpha}} \int_{\tau/\varepsilon}^{(\tau+\delta)/\varepsilon} |V_u|^\gamma \, u^{-\beta} \, \mathrm{d}u\right]. \tag{10.27}$$

Writing $M^{(\varepsilon)} =: \left(M^{(\varepsilon),1}, M^{(\varepsilon),2}\right)^T$, the quadratic variations of $M^{(\varepsilon),1}$ and $M^{(\varepsilon),2}$ satisfy, by [App09] (see (4.15) p. 257), for all $t \geq \varepsilon t_0$

$$\begin{aligned} \mathrm{Tr}\left(\left[M_t^{(\varepsilon)}, M_t^{(\varepsilon)}\right]\right) &:= \left[M_t^{(\varepsilon),1}, M_t^{(\varepsilon),1}\right] + \left[M_t^{(\varepsilon),2}, M_t^{(\varepsilon),2}\right] \\ &= \int_{t_0}^{t/\varepsilon} \int_{\mathbb{R}} \left\| \varepsilon^{\frac{1}{\alpha}} \begin{pmatrix} -\sin(u) \\ \cos(u) \end{pmatrix} z \right\|^2 N(\mathrm{d}u, \mathrm{d}z). \end{aligned}$$

Using Burkholder-Davis-Gundi's inequality (see Theorem 48 p. 193 in [Pro05]), we deduce that for some constant $C > 0$ independent of τ, δ and ε which may change from line to line

$$\begin{aligned} \mathbb{E}\left[\left\|M_{\tau+\delta}^{(\varepsilon)} - M_\tau^{(\varepsilon)}\right\|\right] &\leq C\mathbb{E}\left[\left(\mathrm{Tr}\left(\left[M_{\tau+\delta}^{(\varepsilon)}, M_{\tau+\delta}^{(\varepsilon)}\right] - \left[M_\tau^{(\varepsilon)}, M_\tau^{(\varepsilon)}\right]\right)\right)^{\frac{1}{2}}\right] \\ &\leq C\mathbb{E}\left[\left(\int_{\tau/\varepsilon}^{(\tau+\delta)/\varepsilon} \int_{\mathbb{R}} |\varepsilon^{\frac{1}{\alpha}} z|^2 \, N(\mathrm{d}u, \mathrm{d}z)\right)^{\frac{1}{2}}\right] \\ &= C\mathbb{E}\left[\left(\left[L_{\tau+\delta}^{(\varepsilon)}, L_{\tau+\delta}^{(\varepsilon)}\right] - \left[L_\tau^{(\varepsilon)}, L_\tau^{(\varepsilon)}\right]\right)^{\frac{1}{2}}\right], \end{aligned}$$

where $(L_t^{(\varepsilon)})_{t\geq 0} := (\varepsilon^{\frac{1}{\alpha}} L_{t/\varepsilon})_{t\geq 0}$ has the same distribution as L by its self-similarity property. Using this and the lower-bound in Burkholder-Davis-Gundi's inequality, we deduce that

$$\begin{aligned}
\mathbb{E}\left[\left\|M^{(\varepsilon)}_{\tau+\delta} - M^{(\varepsilon)}_{\tau}\right\|\right] &\leq C\mathbb{E}\left[\sup_{0\leq s\leq\delta} |L_{\tau+s} - L_\tau|\right] \\
&= C\mathbb{E}\left[\sup_{0\leq s\leq\delta} |L_s|\right] \\
&= C\mathbb{E}\left[\sup_{0\leq s\leq\delta} \delta^{\frac{1}{\alpha}}|L_{s/\delta}|\right] \\
&= C\delta^{\frac{1}{\alpha}}\mathbb{E}\left[\sup_{0\leq s\leq 1} |L_s|\right],
\end{aligned}$$

where the first equality stems from the strong Markov property satisfied by L and the second one from the self-similarity property of L again. Note that $\mathbb{E}\left[\sup_{0\leq s\leq 1}|L_s|\right]$ is finite thanks to [LP08] (see Section 3) since $\alpha > 1$.

Since $\tau \in [m, M]$ a.s., the last term in (10.27) can be handled as in (10.9) using moment estimates of V (see Proposition 10.14). It yields

$$\mathbb{E}\left[\varepsilon^{\frac{1}{\alpha}} \int_{\tau/\varepsilon}^{(\tau+\delta)/\varepsilon} |V_u|^\gamma u^{-\beta}\,\mathrm{d}u\right] \leq C_{m,M}\varepsilon^{\beta-\frac{\gamma+\alpha-1}{\alpha}}.$$

Since $\eta > 0$ and by Markov's inequality, we obtain for δ and ε small enough

$$\mathbb{P}\left(\left\|Y^{(\varepsilon)}_{\tau+\delta} - Y^{(\varepsilon)}_{\tau}\right\| \geq a\right) \leq \frac{C\delta^{\frac{1}{\alpha}} + C_{m,M}\varepsilon^{\beta-\frac{\gamma+\alpha-1}{\alpha}}}{a} \leq \eta.$$

Moreover, by Markov's inequality and the moment estimates again, we deduce that for all $t \in [m, M]$,

$$\lim_{a\to+\infty} \limsup_{\varepsilon\to 0} \mathbb{P}\left(\left\|Y^{(\varepsilon)}_t\right\| \geq a\right) \leq \lim_{a\to+\infty} \limsup_{\varepsilon\to 0} \frac{\mathbb{E}\left[\left\|Y^{(\varepsilon)}_t\right\|\right]}{a} \leq \lim_{a\to+\infty} \frac{Ct^{\frac{1}{\alpha}}}{a} = 0.$$

By Corollary and Theorem 16.8 p. 175 in [Bil99], this concludes the proof of the tightness on every compact interval of $(0, +\infty)$. □

We will now prove the convergence of the f.d.d. of $Y^{(\varepsilon)}$. Thanks to the previous lemma, this will yield the weak convergence on every compact set (see Theorem 13.1 p. 139 in [Bil99]) in the super-critical and critical regimes. The convergence in distribution on the whole space $\mathcal{D}$ will follow from Theorem 16.7 p. 174 in [Bil99], since all processes considered are càdlàg.

Convergence of the f.d.d. in the super-critical regime

Assume that $\alpha q > 1$. Recall that $(Y^{(\varepsilon)}_t)_{t\geq\varepsilon t_0} := (\varepsilon^{\frac{1}{\alpha}} Y_{t/\varepsilon})_{t\geq\varepsilon t_0}$.

Proof of Theorem 10.2 (i).

Step 1. We first prove the convergence of the f.d.d. of the Wiener-Lévy integral appearing in (SDE_Y). Recall that the local martingale $M^{(\varepsilon)}$ was defined in (10.26).

STEP 1A. We begin with the convergence in distribution of $M^{(\varepsilon)}_{s,t} := M^{(\varepsilon)}_t - M^{(\varepsilon)}_s$, for $\varepsilon t_0 \le s \le t$. To this end, we study the characteristic function $\phi^{(\varepsilon)}_{s,t}$ of $M^{(\varepsilon)}_{s,t}$. Let us recall that ψ denotes the characteristic exponent of L, and is given, for all $\xi \in \mathbb{R}$, by

$$\psi(\xi) = -a|\xi|^\alpha.$$

The characteristic function of the Wiener-Lévy integral can be computed as p. 105 in [Sat99], hence one has, for all $\xi := (u, v) \in \mathbb{R}^2$,

$$\begin{aligned}
\phi^{(\varepsilon)}_{s,t}(\xi) &= \mathbb{E}\left(\exp\left[-iu\varepsilon^{\frac{1}{\alpha}} \int_{s/\varepsilon}^{t/\varepsilon} \sin(y)\,\mathrm{d}L_y + iv\varepsilon^{\frac{1}{\alpha}} \int_{s/\varepsilon}^{t/\varepsilon} \cos(y)\,\mathrm{d}L_y\right]\right) \\
&= \mathbb{E}\left(\exp\left[i\varepsilon^{\frac{1}{\alpha}} \int_{s/\varepsilon}^{t/\varepsilon} (-u\sin(y) + v\cos(y))\,\mathrm{d}L_y\right]\right) \\
&= \exp\left(\int_{s/\varepsilon}^{t/\varepsilon} \psi\left(\varepsilon^{\frac{1}{\alpha}}[-u\sin(y) + v\cos(y)]\right)\mathrm{d}y\right) \\
&= \exp\left(-a\varepsilon \int_{s/\varepsilon}^{t/\varepsilon} |-u\sin(y) + v\cos(y)|^\alpha\,\mathrm{d}y\right).
\end{aligned}$$

Using Lemma 10.22, we deduce that $\phi^{(\varepsilon)}_{s,t}(\xi)$ converges, as $\varepsilon \to 0$, to

$$\exp\left(-a(t-s)\frac{1}{2\pi}\int_0^{2\pi} |-u\sin(y) + v\cos(y)|^\alpha\,\mathrm{d}y\right).$$

STEP 1B. We now compute explicitly the scale parameter of the stable limiting process. We denote by λ the uniform probability distribution on the circle $\mathbb{S}^1$. Thanks to a change of variable and the symmetry of λ, setting $\omega := \frac{\xi}{\|\xi\|}$ for $\xi = (u, v) \in \mathbb{R}^2 \setminus \{0\}$, we have

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} |-u\sin(y) + v\cos(y)|^\alpha\,\mathrm{d}y &= \int_{\mathbb{S}^1} |\xi\cdot\lambda|^\alpha\,\mathrm{d}\lambda \\
&= \|\xi\|^\alpha \int_{\mathbb{S}^1} |\omega\cdot\lambda|^\alpha\,\mathrm{d}\lambda.
\end{aligned}$$

Since λ is rotationally invariant, we deduce that $\int_{\mathbb{S}^1} |\omega\cdot\lambda|^\alpha\,\mathrm{d}\lambda$ does not depend on $\omega \in \mathbb{S}^1$. Taking $\omega = (1, 0)^T$, we set

$$\widetilde{C} := \frac{a}{2\pi}\int_0^{2\pi} |\cos(x)|^\alpha\,\mathrm{d}x. \tag{10.28}$$

We have thus proved that, for any $\xi \in \mathbb{R}^2$,

$$\phi^{(\varepsilon)}_{s,t}(\xi) \underset{\varepsilon\to 0}{\longrightarrow} \exp\left(-(t-s)\widetilde{C}\|\xi\|^\alpha\right).$$

Thus, the following convergence in distribution holds

$$M^{(\varepsilon)}_{s,t} = M^{(\varepsilon)}_t - M^{(\varepsilon)}_s \underset{\varepsilon\to 0}{\Longrightarrow} \mathcal{L}_{t-s}. \tag{10.29}$$

Following the same lines, we show that, for any $t > 0$,

$$M^{(\varepsilon)}_t \underset{\varepsilon\to 0}{\Longrightarrow} \mathcal{L}_t. \tag{10.30}$$

Step 1c. We now prove the convergence in f.d.d. of $M^{(\varepsilon)}$ to $\mathcal{L}$, as ε tends to 0.
Let us fix $0 < t_1 \le t_2 \le \cdots \le t_d$. Note that $(M_t^{(\varepsilon)})_{t \ge \varepsilon t_0}$ is a càdlàg process with independent increments, since the integrands in its definition are deterministic and because L is a Lévy process. Thus, the random variables $(M_{t_1}^{(\varepsilon)}, M_{t_1,t_2}^{(\varepsilon)}, \ldots, M_{t_{d-1},t_d}^{(\varepsilon)})$ are mutually independent. We deduce from the convergence results established in (10.29) and (10.30), and the fact that $\mathcal{L}$ has stationary and independent increments that

$$(M_{t_1}^{(\varepsilon)}, M_{t_1,t_2}^{(\varepsilon)}, \ldots, M_{t_{d-1},t_d}^{(\varepsilon)}) \underset{\varepsilon \to 0}{\Longrightarrow} (\mathcal{L}_{t_1}, \mathcal{L}_{t_2} - \mathcal{L}_{t_1}, \ldots, \mathcal{L}_{t_d} - \mathcal{L}_{t_{d-1}}).$$

The continuous mapping theorem yields the convergence in f.d.d. of $M^{(\varepsilon)}$ to $\mathcal{L}$.

Step 2. Pick $T > 0$. We prove that

$$\mathbb{E}\left[\sup_{\varepsilon t_0 \le t \le T} \left\| Y_t^{(\varepsilon)} - M_t^{(\varepsilon)} \right\|\right] \underset{\varepsilon \to 0}{\longrightarrow} 0.$$

Let us fix $\varepsilon > 0$ small enough such that $\varepsilon t_0 \le T$. We have

$$\sup_{\varepsilon t_0 \le t \le T} \left\| Y_t^{(\varepsilon)} - M_t^{(\varepsilon)} \right\| \le \varepsilon^{\frac{1}{\alpha}} \|z_0\| + \varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{T/\varepsilon} \left\| e^{-sA} F(s, V_s) \right\| \mathrm{d}s.$$

We use moment estimates (Proposition 10.14) to get

$$\begin{aligned}
\mathbb{E}\left[\varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{T/\varepsilon} \left\| e^{-sA} F(s, V_s) \right\| \mathrm{d}s\right] &= \mathbb{E}\left[\varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{T/\varepsilon} \| F(s, V_s) \| \,\mathrm{d}s\right] \\
&\le \mathbb{E}\left[\varepsilon^{\frac{1}{\alpha}} \int_{t_0}^{T/\varepsilon} |V_s|^{\gamma} s^{-\beta} \,\mathrm{d}s\right] \\
&\le \varepsilon^{\frac{1}{\alpha}} C_{\gamma,t_0} \int_{t_0}^{T/\varepsilon} s^{\frac{\gamma}{\alpha} - \beta} \,\mathrm{d}s \\
&\le C_{\gamma,t_0} (\varepsilon^{\beta - \frac{\gamma+\alpha-1}{\alpha}} T^{\frac{\gamma}{\alpha} - \beta + 1} + \varepsilon^{\frac{1}{\alpha}} t_0^{\frac{\gamma}{\alpha} - \beta + 1}).
\end{aligned}$$

Hence, setting $r := \min(\beta - \frac{\gamma+\alpha-1}{\alpha}, \frac{1}{\alpha})$, which is positive by assumption, we get

$$\mathbb{E}\left[\sup_{\varepsilon t_0 \le t \le T} \left\| Y_t^{(\varepsilon)} - M_t^{(\varepsilon)} \right\|\right] = \underset{\varepsilon \to 0}{O}(\varepsilon^r). \tag{10.31}$$

The conclusion follows from Theorem 3.1 p. 27 in [Bil99].

□

Convergence of the f.d.d. in the critical and sub-critical regime

In this section, we consider the linear case, i.e. $\gamma = 1$ and we assume that $\beta \in \left(\frac{1}{2}, 1\right]$. Recall that

$$(Y_t^{(\varepsilon)})_{t \ge \varepsilon t_0} = \left(\varepsilon^q Y_{t/\varepsilon}\right)_{t \ge \varepsilon t_0}.$$

Proof of Theorem 10.2 (ii) and (iii). The proof follows the same lines as in the Brownian setting. Leaving out the noise, recall that the underlying ODE is the following

$$x''(t) + \frac{x'(t)}{t^\beta} + x(t) = 0, \quad t \geq t_0. \tag{10.32}$$

We pick again the basis of solutions given by Lemma 10.20, and we still denote by R its resolvent matrix and by f its rate of decrease. Recall that it is given by

$$\forall t > 0,\ f(t) := \begin{cases} \frac{1}{\sqrt{t}} & \text{if } \beta = 1, \\ \exp\left(-\frac{t^{1-\beta}}{2(1-\beta)}\right) & \text{else.} \end{cases} \tag{10.33}$$

We set, for all $t \geq \varepsilon t_0$,

$$\widetilde{M}_t^{(\varepsilon)} := \varepsilon^q f(t/\varepsilon) \int_{t_0}^{t/\varepsilon} R_s^{-1} \Gamma \, \mathrm{d}S_s. \tag{10.34}$$

Keeping the same notations as in the Brownian case, we decompose $(Y_t)_{t \geq t_0} = (e^{-tA} Z_t)_{t \geq t_0}$ into

$$\varepsilon^q Y_{t/\varepsilon} = \varepsilon^q f(t/\varepsilon) \Phi_{t/\varepsilon} R_{t_0}^{-1} Z_0 + \Phi_{t/\varepsilon} \widetilde{M}_t^{(\varepsilon)}.$$

Reasoning as in the Brownian case, it remains to study the convergence of $\widetilde{M}^{(\varepsilon)}$ since the first term converges towards 0. Using the expression of the Wronskian obtained in Lemma 10.20, we obtain, for all $t \geq \varepsilon t_0$,

$$\widetilde{M}_t^{(\varepsilon)} = \varepsilon^q f(t/\varepsilon) \int_{t_0}^{t/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}L_u.$$

Let us fix $0 < s < t$. We study the convergence in distribution of the couple $(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)})$ when ε tends to 0. The convergence in distribution of a general d-dimensional distribution $(\widetilde{M}_{t_1}^{(\varepsilon)}, \dots, \widetilde{M}_{t_d}^{(\varepsilon)})$ relies on the same computations.

Let us fix $(\xi_1, \xi_2) \in \mathbb{R}^2 \times \mathbb{R}^2$. Using that L has independent increments, the characteristic function $\widetilde{\phi}_{s,t}^{(\varepsilon)}$ of $(\widetilde{M}_s^{(\varepsilon)}, \widetilde{M}_t^{(\varepsilon)})$ is given by

$$\begin{aligned} \widetilde{\phi}_{s,t}^{(\varepsilon)}(\xi_1, \xi_2) &= \mathbb{E}\left[\exp\left(i\varepsilon^q \left[f(s/\varepsilon)\, \xi_1 \cdot \int_{t_0}^{s/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}L_u \right.\right.\right. \\ &\qquad \left.\left.\left. + f(t/\varepsilon)\, \xi_2 \cdot \int_{t_0}^{t/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}L_u \right]\right)\right] \\ &= \mathbb{E}\left[\exp\left(i\varepsilon^q \left(f(s/\varepsilon)\, \xi_1 + f(t/\varepsilon)\, \xi_2\right) \cdot \int_{t_0}^{s/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}L_u \right)\right] \\ &\qquad \times \mathbb{E}\left[\exp\left(i\varepsilon^q f(t/\varepsilon)\, \xi_2 \cdot \int_{s/\varepsilon}^{t/\varepsilon} f(u)^{-2} \begin{pmatrix} -y_2(u) \\ y_1(u) \end{pmatrix} \mathrm{d}L_u \right)\right]. \end{aligned} \tag{10.35}$$

Let us recall that the characteristic exponent of L is given, for all $\xi \in \mathbb{R}$, by

$$\psi(\xi) = -a|\xi|^\alpha.$$

The characteristic function of the Wiener-Lévy integral can be computed as p. 105 in [Sat99]. Indeed, if

$G:\mathbb{R}\to\mathbb{R}$ is a continuous function, then we have for all $z\in\mathbb{R}$

$$\mathbb{E}\left[\exp\left(iz\int_s^t G(u)\,\mathrm{d}L_u\right)\right]=\exp\left(-a\int_s^t |G(u)z|^\alpha\,\mathrm{d}u\right).$$

Using this with $z=1$ and

$$G:u\in[t_0,+\infty)\mapsto \varepsilon^q\left(f\left(s/\varepsilon\right)\xi_1+f\left(t/\varepsilon\right)\xi_2\right)\cdot\left[f(u)^{-2}\begin{pmatrix}-y_2(u)\\ y_1(u)\end{pmatrix}\right]$$

to compute the first expectation in (10.35) and the corresponding function G for the second expectation, one has

$$\begin{aligned}\widetilde{\phi}_{s,t}^{(\varepsilon)}(\xi_1,\xi_2)=\exp&\left(-a\varepsilon^\beta\int_{t_0}^{s/\varepsilon}f(u)^{-2\alpha}\left|\left(f\left(s/\varepsilon\right)\xi_1+f\left(t/\varepsilon\right)\xi_2\right)\cdot\begin{pmatrix}-y_2(u)\\ y_1(u)\end{pmatrix}\right|^\alpha\mathrm{d}u\right)\\ &\times\exp\left(-a\varepsilon^\beta\int_{s/\varepsilon}^{t/\varepsilon}f(u)^{-2\alpha}\left|f\left(t/\varepsilon\right)\xi_2\cdot\begin{pmatrix}-y_2(u)\\ y_1(u)\end{pmatrix}\right|^\alpha\mathrm{d}u\right).\end{aligned}$$

Using the asymptotic expansion of the resolvent matrix (Lemma 10.20), we can write, for any $u\geq t_0$,

$$\begin{pmatrix}-y_2(u)\\ y_1(u)\end{pmatrix}=f(u)\left[\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}+g(u)\right],$$

where $g:[t_0,+\infty)\to\mathbb{R}^2$ is a function satisfying for all $u\geq t_0$,

$$|g(u)|\leq Cu^{1-2\beta}.$$

We set

$$K_1^{(\varepsilon)}:=\exp\left(-a\varepsilon^\beta\int_{t_0}^{s/\varepsilon}f(u)^{-\alpha}\left|\left(f\left(s/\varepsilon\right)\xi_1+f\left(t/\varepsilon\right)\xi_2\right)\cdot\left[\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}+g(u)\right]\right|^\alpha\mathrm{d}u\right) \tag{10.36}$$

and

$$K_2^{(\varepsilon)}:=\exp\left(-a\varepsilon^\beta\int_{s/\varepsilon}^{t/\varepsilon}f(u)^{-\alpha}\left|f\left(\frac{t}{\varepsilon}\right)\xi_2\cdot\left[\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}+g(u)\right]\right|^\alpha\mathrm{d}u\right).$$

We thus obtain

$$\widetilde{\phi}_{s,t}^{(\varepsilon)}(\xi_1,\xi_2)=K_1^{(\varepsilon)}\times K_2^{(\varepsilon)}, \tag{10.37}$$

Step 1. We start by justifying that we can omit g to study the limit when $\varepsilon\to0$. More precisely, we prove that, for all function $\zeta:\mathbb{R}\to\mathbb{R}^2$ such that $\|\zeta(\varepsilon)\|\,f(s/\varepsilon)^{-1}=\underset{\varepsilon\to0}{O}(1)$,

$$R^{(\varepsilon)}:=\varepsilon^\beta\int_{t_0}^{s/\varepsilon}f(u)^{-\alpha}\left|\left|\zeta(\varepsilon)\cdot\left[\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}+g(u)\right]\right|^\alpha-\left|\zeta(\varepsilon)\cdot\left[\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right]\right|^\alpha\right|\mathrm{d}u\underset{\varepsilon\to0}{\longrightarrow}0. \tag{10.38}$$

Thanks to the mean value theorem applied to $|\cdot|^\alpha$ (since $\alpha\geq1$), and the domination of g, we obtain that, for some constant $C>0$,

$$R^{(\varepsilon)}\leq C\varepsilon^\beta\,\|\zeta(\varepsilon)\|^\alpha\int_{t_0}^{s/\varepsilon}f(u)^{-\alpha}u^{1-2\beta}\,\mathrm{d}u=\underset{\varepsilon\to0}{O}(\|\zeta(\varepsilon)\|^\alpha\,f(s/\varepsilon)^{-\alpha}\varepsilon^{2\beta-1}),$$

where the last equality follows from Lemma 10.23. This proves (10.38) since $\beta > \frac{1}{2}$.

STEP 2. We focus on the first term $K_1^{(\varepsilon)}$ defined in (10.36). Since f is decreasing, notice that

$$\zeta(\varepsilon) := f\left(s/\varepsilon\right)\xi_1 + f\left(t/\varepsilon\right)\xi_2 = \underset{\varepsilon\to 0}{O}(f(s/\varepsilon)).$$

Then we have to study the convergence of $I^{(\varepsilon)}$ defined by

$$I^{(\varepsilon)} := a\varepsilon^{\beta}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\left|\left(f\left(s/\varepsilon\right)\xi_1 + f\left(t/\varepsilon\right)\xi_2\right)\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha}\mathrm{d}u.$$

Its limit differs according to the value of β.

STEP 2A. Assume first that $\beta = 1$. Then, using the expression of f (see (10.33)),

$$I^{(\varepsilon)} = a\varepsilon^{1+\frac{\alpha}{2}}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\left|\left(\frac{\xi_1}{\sqrt{s}} + \frac{\xi_2}{\sqrt{t}}\right)\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha}\mathrm{d}u.$$

We proved in Step 1B of the super-critical regime that there exists a constant $\widetilde{C} > 0$ given in (10.28) such that, for all $\zeta \in \mathbb{R}^2$,

$$\frac{a}{2\pi}\int_0^{2\pi}\left|\zeta\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha}\mathrm{d}u = \widetilde{C}\|\zeta\|^{\alpha}. \tag{10.39}$$

Using Lemma 10.22, we can compute the following asymptotic expansion

$$I^{(\varepsilon)} = \varepsilon^{1+\frac{\alpha}{2}}\widetilde{C}\left\|\frac{\xi_1}{\sqrt{s}} + \frac{\xi_2}{\sqrt{t}}\right\|^{\alpha}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\,\mathrm{d}u + \underset{\varepsilon\to 0}{o}\left(\varepsilon^{1+\frac{\alpha}{2}}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\,\mathrm{d}u\right).$$

Therefore, it follows from Lemma 10.23 that

$$K_1^{(\varepsilon)} \underset{\varepsilon\to 0}{\longrightarrow} \exp\left(-\widetilde{C}\left(1+\frac{\alpha}{2}\right)^{-1}\left\|\frac{\xi_1}{\sqrt{s}} + \frac{\xi_2}{\sqrt{t}}\right\|^{\alpha} s^{1+\frac{\alpha}{2}}\right).$$

STEP 2B. Let us consider now $\beta \in \left(\frac{1}{2}, 1\right)$. Let us notice that $I^{(\varepsilon)}$ can be decomposed into the sum

$$I^{(\varepsilon)} = I_1^{(\varepsilon)} + I_2^{(\varepsilon)} \tag{10.40}$$

of the two following terms

$$I_1^{(\varepsilon)} := a\varepsilon^{\beta}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\left|f\left(\frac{s}{\varepsilon}\right)\xi_1\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha}\mathrm{d}u$$

and

$$I_2^{(\varepsilon)} := a\varepsilon^{\beta}\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\left[\left|\left(f\left(s/\varepsilon\right)\xi_1 + f\left(t/\varepsilon\right)\xi_2\right)\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha} - \left|f\left(s/\varepsilon\right)\xi_1\cdot\begin{pmatrix}-\sin(u)\\ \cos(u)\end{pmatrix}\right|^{\alpha}\right]\mathrm{d}u.$$

Using again the mean value theorem and Lemma 10.23, we get that for some positive constant $C_{\|\xi_1\|,\|\xi_2\|}$,

$$|I_2^{(\varepsilon)}| \le C\varepsilon^{\beta} f\left(s/\varepsilon\right)^{\alpha-1} f\left(t/\varepsilon\right)\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha}\,\mathrm{d}u = \underset{\varepsilon\to 0}{O}\left(f(t/\varepsilon)f(s/\varepsilon)^{-1}\right) = \underset{\varepsilon\to 0}{o}(1), \tag{10.41}$$

since $\beta < 1$. Using Lemma 10.22, we can compute the following asymptotic expansion of $I_1^{(\varepsilon)}$

$$\begin{aligned}I_1^{(\varepsilon)} &= a\varepsilon^\beta f\left(\frac{s}{\varepsilon}\right)^\alpha \int_{t_0}^{s/\varepsilon} f(u)^{-\alpha} \left| \xi_1 \cdot \begin{pmatrix} -\sin(u) \\ \cos(u) \end{pmatrix} \right|^\alpha \mathrm{d}u \\ &= a\varepsilon^\beta f\left(\frac{s}{\varepsilon}\right)^\alpha \left[\left(\frac{1}{2\pi} \int_0^{2\pi} \left| \xi_1 \cdot \begin{pmatrix} -\sin(u) \\ \cos(u) \end{pmatrix} \right|^\alpha \mathrm{d}u \right) \int_{t_0}^{s/\varepsilon} f(u)^{-\alpha} \,\mathrm{d}u + \underset{\varepsilon\to 0}{o} \left(\int_{t_0}^{s/\varepsilon} f(u)^{-\alpha} \,\mathrm{d}u \right) \right].\end{aligned}$$

Thanks to (10.39) and the asymptotic expansion's results given in Lemma 10.23, there exists an explicit constant $k_{\beta,\alpha}$ given in Lemma 10.23, such that

$$I_1^{(\varepsilon)} \underset{\varepsilon\to 0}{\longrightarrow} k_{\beta,\alpha} \widetilde{C} s^\beta \|\xi_1\|^\alpha. \tag{10.42}$$

Combining (10.38), (10.40), (10.41) and (10.42), we have proved that $K_1^{(\varepsilon)}$, defined in (10.36), converges as $\varepsilon \to 0$ towards

$$\exp\left(-k_{\beta,\alpha} \widetilde{C} s^\beta \|\xi_1\|^\alpha \right).$$

Step 3. It remains to deal with the limit of $K_2^{(\varepsilon)}$. Notice that

$$\zeta(\varepsilon) := f\left(t/\varepsilon\right) \xi_2 = \underset{\varepsilon\to 0}{O} \left(f\left(t/\varepsilon\right) \right) = \underset{\varepsilon\to 0}{O} \left(f\left(s/\varepsilon\right) \right).$$

Hence, thanks to Step 1, we are reduced to study, for $r \in \{s, t\}$,

$$J_r^{(\varepsilon)} := a\varepsilon^\beta f\left(\frac{t}{\varepsilon}\right)^\alpha \int_{t_0}^{r/\varepsilon} f(u)^{-\alpha} \left| \xi_2 \cdot \begin{pmatrix} -\sin(u) \\ \cos(u) \end{pmatrix} \right|^\alpha \mathrm{d}u. \tag{10.43}$$

Asymptotic expansion's results (Lemmas 10.22 and 10.23) and (10.39) yield

$$\begin{aligned}J_r^{(\varepsilon)} &= \widetilde{C} \left\|\xi_2\right\|^\alpha \varepsilon^\beta f\left(t/\varepsilon\right)^\alpha \int_{t_0}^{r/\varepsilon} f(u)^{-\alpha} \,\mathrm{d}u + \underset{\varepsilon\to 0}{o} \left(f\left(t/\varepsilon\right)^\alpha f\left(r/\varepsilon\right)^{-\alpha} \right) \\ &= \widetilde{C} \left\|\xi_2\right\|^\alpha k_{\beta,\alpha} r^\beta f\left(t/\varepsilon\right)^\alpha f\left(r/\varepsilon\right)^{-\alpha} + \underset{\varepsilon\to 0}{o} \left(f\left(t/\varepsilon\right)^\alpha f\left(r/\varepsilon\right)^{-\alpha} \right).\end{aligned}$$

Hence,

$$J_t^{(\varepsilon)} \underset{\varepsilon\to 0}{\longrightarrow} \widetilde{C} \left\|\xi_2\right\|^\alpha k_{\beta,\alpha} t^\beta \tag{10.44}$$

and

$$J_s^{(\varepsilon)} \underset{\varepsilon\to 0}{\longrightarrow} \widetilde{C} \left\|\xi_2\right\|^\alpha k_{\beta,\alpha} s^\beta \left(\frac{s}{t}\right)^{\frac{\alpha}{2}} \mathbb{1}_{\{\beta=1\}}. \tag{10.45}$$

Since

$$K_2^{(\varepsilon)} = \exp\left(-J_t^{(\varepsilon)} + J_s^{(\varepsilon)} \right),$$

we thus obtain that, for all $0 < s \leq t$,

$$\widetilde{\phi}_{s,t}^{(\varepsilon)}(\xi_1, \xi_2) \underset{\varepsilon\to 0}{\longrightarrow} \begin{cases} \exp\left(-k_{\beta,\alpha} \widetilde{C} s^\beta \|\xi_1\|^\alpha \right) \exp\left(-k_{\beta,\alpha} \widetilde{C} t^\beta \|\xi_2\|^\alpha \right) & \text{if } \beta < 1, \\ \exp\left(-k_{\beta,\alpha} \widetilde{C} \left[\left\| \frac{\xi_1}{\sqrt{s}} + \frac{\xi_2}{\sqrt{t}} \right\|^\alpha s^{1+\frac{\alpha}{2}} + \|\xi_2\|^\alpha t - \|\xi_2\|^\alpha \left(\frac{s}{t}\right)^{\frac{\alpha}{2}} s \right] \right) & \text{if } \beta = 1. \end{cases}$$

Step 4. We can compute the characteristic function of the process $\left(\frac{1}{\sqrt{t}} \int_0^t \sqrt{s} \,\mathrm{d}\mathcal{L}_s \right)_{t>0}$ in the same manner, and thus recognize the limiting process in the critical regime.

□

Remark 10.17.

- As in the Brownian setting, if $\beta = 0$, the resolvent matrix is explicit and following the same lines, we can prove that $\left(Z_{t/\varepsilon}\right)_{t\geq\varepsilon t_0}$ converges in f.d.d. towards the product of the measure μ, whose characteristic function is given by

$$\xi \mapsto \exp\left(-a\int_0^{+\infty} e^{-\alpha u}\left\|\xi\cdot h(u)\right\|^\alpha \mathrm{d}u\right),$$

  h being an explicit periodic function depending on the resolvent matrix.

- As in the Brownian setting, since the asymptotic expansion of the resolvant matrix is also known in the super-critical regime, i.e. $\beta > 1$, one can prove the result in the linear case, i.e. $\gamma = 1$, following the same lines.

10.3.4 Proof of Corollary 10.6

Proof of Corollary 10.6. We start by proving the convergence in distribution of $r_{1/T}Z_T$. Reasoning as in the Brownian setting, it follows from Theorem 10.2 that $r_{1/T}Y_T$ converges. The conclusion is a consequence of Lemma 10.24, noting that the limiting distribution is invariant under rotations thanks to the expression of its characteristic function.
Let us now prove that the rescaled process $Z^{(\varepsilon)}$ does not converge in distribution. We state the proof in the super-critical regime. Assume by contradiction that it is the case. Reasoning as in the Brownian case, we prove that this implies the convergence in distribution of the process $I^{(\varepsilon)}$ defined, for $t \geq \varepsilon t_0$, by

$$I_t^{(\varepsilon)} := \varepsilon^{\frac{1}{\alpha}}\int_{t_0}^{t/\varepsilon}\sin\left(\frac{t}{\varepsilon}-u\right)\mathrm{d}L_u.$$

In particular, for $s < t$, the random variable $I_t^{(\varepsilon)} - I_s^{(\varepsilon)}$ shall converge in distribution.

Let us denote by $\phi^{(\varepsilon)}$ the characteristic function of $I_t^{(\varepsilon)} - I_s^{(\varepsilon)}$, which is supposed to converge on $\mathbb{R}$. Using that L has independent increments, we have

$$\begin{aligned}\phi^{(\varepsilon)}(1) &= \mathbb{E}\left[\exp\left(\varepsilon^{\frac{1}{\alpha}}\int_{s/\varepsilon}^{t/\varepsilon}\sin\left(\frac{t}{\varepsilon}-u\right)\mathrm{d}L_u\right)\right]\mathbb{E}\left[\exp\left(\varepsilon^{\frac{1}{\alpha}}\int_{t_0}^{s/\varepsilon}\sin\left(\frac{t}{\varepsilon}-u\right)-\sin\left(\frac{s}{\varepsilon}-u\right)\mathrm{d}L_u\right)\right]\\ &=: \phi^{(\varepsilon),1}\phi^{(\varepsilon),2}.\end{aligned}$$

Recall that ψ defined in (10.2), denotes the characteristic exponent of L. Using a change of variables, we have in particular

$$\begin{aligned}\phi^{(\varepsilon),1} &= \exp\left(\int_{s/\varepsilon}^{t/\varepsilon}\psi\left(\varepsilon^{\frac{1}{\alpha}}\sin\left(\frac{t}{\varepsilon}-u\right)\right)\mathrm{d}u\right) = \exp\left(-a\varepsilon\int_{s/\varepsilon}^{t/\varepsilon}\left|\sin\left(\frac{t}{\varepsilon}-u\right)\right|^\alpha \mathrm{d}u\right)\\ &= \exp\left(-a\varepsilon\int_0^{(t-s)/\varepsilon}\left|\sin\left(u\right)\right|^\alpha \mathrm{d}u\right).\end{aligned}$$

Lemma 10.22 ensures that $\phi^{(\varepsilon),1}$ has a limit when ε converges to 0. Similarly, we obtain

$$\begin{aligned}
\phi^{(\varepsilon),2} &= \exp\left(\int_{t_0}^{s/\varepsilon} \psi\left(\varepsilon^{\frac{1}{\alpha}} \sin\left(\frac{t}{\varepsilon}-u\right) - \sin\left(\frac{s}{\varepsilon}-u\right)\right) \mathrm{d}u\right) \\
&= \exp\left(-a\varepsilon \int_{t_0}^{s/\varepsilon} \left|\sin\left(\frac{t}{\varepsilon}-u\right) - \sin\left(\frac{s}{\varepsilon}-u\right)\right|^{\alpha} \mathrm{d}u\right) \\
&= \exp\left(-a2^{\alpha}\left|\sin\left(\frac{t-s}{2\varepsilon}\right)\right|^{\alpha} \varepsilon \int_{t_0}^{s/\varepsilon} \left|\cos\left(\frac{t+s}{2\varepsilon}-u\right)\right|^{\alpha} \mathrm{d}u\right) \\
&= \exp\left(-a2^{\alpha}\left|\sin\left(\frac{t-s}{2\varepsilon}\right)\right|^{\alpha} \varepsilon \int_{\frac{t-s}{2\varepsilon}}^{\frac{t+s}{2\varepsilon}-t_0} \left|\cos\left(u\right)\right|^{\alpha} \mathrm{d}u\right).
\end{aligned} \tag{10.46}$$

The change of variables $u = v + \pi$ yields, for all $\varepsilon > 0$,

$$\int_0^{2\pi} |\cos(u)|^{\alpha} \operatorname{sgn}\left(\sin\left(\frac{t-s}{2\varepsilon}\right)\cos(u)\right) \mathrm{d}u = 0.$$

Thus, Lemma 10.22 ensures that

$$\varepsilon \int_{\frac{t-s}{2\varepsilon}}^{\frac{t+s}{2\varepsilon}-t_0} |\cos(u)|^{\alpha} \mathrm{d}u \quad \underset{\varepsilon\to 0}{\longrightarrow} \quad \frac{s}{2\pi}\int_0^{2\pi} |\cos(u)|^{\alpha} \mathrm{d}u.$$

Coming back to (10.46), we see that $\phi^{(\varepsilon),2}$ does not converge when ε tends to 0. This is a contradiction. $\square$

10.4 Appendix: Study of the deterministic underlying ODE

The deterministic ODE behind the system, i.e. without frictional force and without noise, is the following

$$x''(t) + \frac{x'(t)}{t^{\beta}} + x(t) = 0, \quad t \geq t_0. \tag{10.47}$$

The solutions form a vector space of dimension 2. Let us take two solutions y_1 and y_2 which are linearly independent. Then, we introduce the fundamental system of solutions (resolvent matrix) R to (10.47) defined, for $t \geq t_0$, by

$$R_t = \begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}.$$

It satisfies, for all $t \geq t_0$,

$$R_t' = \begin{pmatrix} 0 & 1 \\ -1 & -\frac{1}{t^{\beta}} \end{pmatrix} R_t.$$

We recall that the Wronskian w is defined, for all $t \geq t_0$, by

$$w(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t).$$

Let us finally set, for $t > 0$,

$$f(t) := \begin{cases} \frac{1}{\sqrt{t}} & \text{if } \beta = 1, \\ \exp\left(-\frac{t^{1-\beta}}{2(1-\beta)}\right) & \text{else.} \end{cases} \tag{10.48}$$

Lemma 10.18. *Pick $\beta \in \left(\frac{1}{2}, +\infty\right)$ and consider a solution y to (10.47). Then, there exist $a \in \mathbb{R}$ and $\phi \in [0, 2\pi)$ such that*

$$y(t) = af(t)\cos(t+\phi) + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right)$$

and

$$y'(t) = -af(t)\sin(t+\phi) + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right).$$

Proof. Let us set, for $t \geq t_0$, $u(t) = f(t)^{-1}y(t)$. We easily check that u satisfies

$$u''(t) + u(t)\left[1 + h(t)\right] = 0,$$

where $h(t) := \frac{f''(t)}{f(t)} + \frac{f'(t)}{f(t)t^\beta}$. Since for all $t \geq t_0$

$$\frac{f'(t)}{f(t)} = -\frac{1}{2t^\beta},$$

we obtain that $h(t) = \underset{t\to+\infty}{O}(t^{-(2\beta)\wedge(\beta+1)})$. Following the proof of the method of variation of parameters, there exists $a_0, b_0 \in \mathbb{R}$ such that, for any $t \geq t_0$,

$$u(t) = a_0\cos(t) + b_0\sin(t) - \int_{t_0}^t u(s)h(s)\sin(t-s)\ \mathrm{d}s.$$

Using that $h \in L^1((t_0, +\infty))$ since $\beta > \frac{1}{2}$, we obtain by Grönwall's lemma that u is bounded on $[t_0, +\infty)$. We deduce that the functions $s \mapsto u(s)h(s)\cos(s)$ and $s \mapsto u(s)h(s)\sin(s)$ belong to $L^1((t_0, +\infty))$. Thus, up to changing the constants a_0 and b_0, one has, for all $t \geq t_0$,

$$u(t) = a_0\cos(t) + b_0\sin(t) - \sin(t)\int_t^\infty u(s)h(s)\cos(s)\ \mathrm{d}s + \cos(t)\int_t^\infty u(s)h(s)\sin(s)\ \mathrm{d}s. \tag{10.49}$$

It follows from the fact that u is bounded that

$$u(t) = a_0\cos(t) + b_0\sin(t) + \underset{t\to+\infty}{O}\left(\int_t^\infty \frac{\mathrm{d}s}{s^{(2\beta)\wedge(\beta+1)}}\right).$$

Thus, there exist $a \in \mathbb{R}$ and $\phi \in [0, 2\pi)$ such that

$$u(t) = a\cos(t+\phi) + \underset{t\to+\infty}{O}(t^{-(2\beta-1)\wedge\beta}).$$

This proves the asymptotic expansion of y. Differentiating (10.49) and using that $h(t) = \underset{t\to+\infty}{O}(t^{-(2\beta)\wedge(\beta+1)})$, we prove that

$$u'(t) = -a\sin(t+\phi) + \underset{t\to+\infty}{O}(t^{-(2\beta-1)\wedge\beta}).$$

Since u is bounded and $f'(t) = \underset{t\to+\infty}{O}(f(t)t^{-\beta})$, we finally obtain that

$$y'(t) = f'(t)u(t) + f(t)u'(t) = f(t)u'(t) + \underset{t\to+\infty}{O}(f(t)t^{-\beta}).$$

This concludes the proof of the asymptotic expansion of y'. □

Remark 10.19. Note that if $\beta = 1$, the Bessel functions of the first kind J_0 and of the second kind Y_0

form a basis of solutions. Their asymptotic expansions can be found in [Wat44, Chap VII].

Lemma 10.20. *Pick $\beta \in \left(\frac{1}{2}, +\infty\right)$. There exists a basis of solutions y_1 and y_2 to (10.47) such that the resolvent matrix R satisfies*

$$R_t = f(t)\begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right) = f(t)e^{tA} + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right).$$

Moreover, its Wronskian w is given for any $t \geq t_0$ by

$$w(t) = f(t)^2.$$

Proof. It is well-known that the Wronskian satisfies, for all $t \geq t_0$,

$$w'(t) = -\frac{1}{t^\beta}w(t).$$

Thus, there exists $w_0 \in \mathbb{R}\backslash\{0\}$ such that, for all $t \geq t_0$, $w(t) = w_0 f(t)^2$. Moreover, thanks to Lemma 10.18, for $i \in \{1, 2\}$, there exist $a_i \in \mathbb{R}$ and $\phi_i \in [0, 2\pi)$ such that

$$y_i(t) = a_i f(t)\cos(t + \phi_i) + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right)$$

and

$$y_i'(t) = -a_i f(t)\sin(t + \phi_i) + \underset{t\to\infty}{O}\left(f(t)t^{-(2\beta-1)\wedge\beta}\right).$$

As a consequence,

$$w(t) = -a_1 a_2 f(t)^2 \sin(\phi_2 - \phi_1) + \underset{t\to\infty}{O}\left(f(t)^2 t^{-(4\beta-2)\wedge 2\beta}\right).$$

But since $w(t) = w_0 f(t)^2$, it implies that $a_i \neq 0$ and $\phi_2 \not\equiv \phi_1[\pi]$.
Up to dividing by a_i, we can assume that $a_i = 1$, and up to considering a linear combination of y_1 and y_2, we can assume that $\phi_1 = 0$ and $\phi_2 = -\frac{\pi}{2}$. Thus, we have $w_0 = 1$. This concludes the proof. ☐

10.5 Appendix: Some technical results

We collect here some technical results used in our proofs. Recall first a sufficient condition for the non-explosion of the solution to a SDE. The proof can be found in [GL21a].

Lemma 10.21. *Let $(Y_t)_{t\geq t_0}$ be a càdlàg process, solution to a SDE. For all $n \geq 0$, define the stopping time*

$$\tau_n := \inf\{t \geq t_0,\ \|Y_t\| \geq n\}. \tag{10.50}$$

Set $\tau_\infty := \lim_{n\to+\infty}\tau_n$ the explosion time of Y. Assume that there exist two measurable and non-negative functions ϕ and b such that

(i) ϕ is non-decreasing and $\lim_{n\to\infty}\phi(n) = +\infty$,

(ii) b is finite-valued,

(iii) and for all $t \geq t_0$,

$$\sup_{n\geq 0}\mathbb{E}\left[\phi(|Y_{t\wedge\tau_n}|)\right] \leq b(t).$$

Then $\tau_\infty = +\infty$ a.s.

We now state and prove a result on the periodic-averaging phenomenon.

Lemma 10.22. *Let us fix $t_0 > 0$ and $h : [t_0, +\infty) \to \mathbb{R}$ a continuous m-periodic function, with $m > 0$. Let $g : [t_0, +\infty) \to \mathbb{R}^+$ be a continuously differentiable function which is not integrable on $[t_0, +\infty)$. We assume moreover that*

(i) $g'(t) = \underset{t\to+\infty}{o}(g(t))$,

(ii) $g(t) = \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right)$.

Then,

$$\int_{t_0}^t g(u)h(u)\,\mathrm{d}u = \left[\frac{1}{m}\int_{t_0}^{t_0+m} h(u)\,\mathrm{d}u\right]\int_{t_0}^t g(u)\,\mathrm{d}u + \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right).$$

Let us remark that the functions g_1 and g_2 defined for $t \in \mathbb{R}$, $r \geq 0$ and $\beta \in [0, 1)$ by $g_1(t) := t^r$ and $g_2(t) := \exp(rt^{1-\beta})$, satisfy the preceding assumptions made on g.

Proof. Let us define $\widetilde{h} := h - \frac{1}{m}\int_{t_0}^{t_0+m} h(u)\,\mathrm{d}u$, and $\widetilde{H}$ a primitive of $\widetilde{h}$. The function $\widetilde{H}$ is bounded on $[t_0, +\infty)$ since the average of $\widetilde{h}$ on its period is equal to 0. To prove the lemma, we only need to justify that

$$\int_{t_0}^t g(u)\widetilde{h}(u)\,\mathrm{d}u = \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right).$$

By integration by parts, we obtain that, for all $t \geq t_0$,

$$\int_{t_0}^t g(u)\widetilde{h}(u)\,\mathrm{d}u = g(t)\widetilde{H}(t) - g(t_0)\widetilde{H}(t_0) - \int_{t_0}^t g'(u)\widetilde{H}(u)\,\mathrm{d}u.$$

Using the fact that $\widetilde{H}$ is bounded, that $g'(t) = \underset{t\to+\infty}{o}(g(t))$ and that $\int_{t_0}^\infty g(u)\,\mathrm{d}u = +\infty$, we deduce that

$$\int_{t_0}^t g'(u)\widetilde{H}(u)\,\mathrm{d}u = \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right).$$

The conclusion follows from the fact that $g(t)\widetilde{H}(t) - g(t_0)\widetilde{H}(t_0) = \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right)$, since we have assumed that $g(t) = \underset{t\to+\infty}{o}\left(\int_{t_0}^t g(u)\,\mathrm{d}u\right)$ and that $\int_{t_0}^\infty g(u)\,\mathrm{d}u = +\infty$.

□

Lemma 10.23. *Let f be given by (10.48) for $\beta \in \left(\frac{1}{2}, 1\right]$, and pick $\alpha \in (1, 2]$. Define*

$$k_{\beta,\alpha} := \begin{cases} (1+\alpha/2)^{-1} & \textit{if } \beta = 1, \\ \frac{2}{\alpha} & \textit{else.} \end{cases}$$

Then for any $t > 0$, we have

$$\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha} u^{1-2\beta}\,\mathrm{d}u = \underset{\varepsilon\to 0}{O}(f(t/\varepsilon)^{-\alpha}\varepsilon^{\beta-1}),$$

and

$$\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}\,\mathrm{d}u = k_{\beta,\alpha} f\left(\frac{t}{\varepsilon}\right)^{-\alpha}\left(\frac{t}{\varepsilon}\right)^{\beta} + \underset{\varepsilon\to 0}{o}(f(t/\varepsilon)^{-\alpha}\varepsilon^{-\beta}).$$

Proof. When $\beta = 1$, the results follow from direct computations because of the expression of f. Assume now that $\beta \in \left(\frac{1}{2}, 1\right)$. For the first point, the integration by parts formula ensures that

$$\begin{aligned}\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha} u^{1-2\beta}\, \mathrm{d}u &= \frac{2}{\alpha}\left[f(u)^{-\alpha}u^{1-\beta}\right]_{t_0}^{t/\varepsilon} - \frac{2}{\alpha}(1-\beta)\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}u^{-\beta}\, \mathrm{d}u \\ &= \underset{\varepsilon\to 0}{O}(f(t/\varepsilon)^{-\alpha}\varepsilon^{\beta-1}) + \underset{\varepsilon\to 0}{O}(f(t/\varepsilon)^{-\alpha}) \\ &= \underset{\varepsilon\to 0}{O}(f(t/\varepsilon)^{-\alpha}\varepsilon^{\beta-1}).\end{aligned}$$

For the second asymptotic expansion, it follows again from an integration by parts that

$$\begin{aligned}\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}\, \mathrm{d}u &= \int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}u^{-\beta}u^{\beta}\, \mathrm{d}u \\ &= \frac{2}{\alpha}\left[f(u)^{-\alpha}u^{\beta}\right]_{t_0}^{t/\varepsilon} - \frac{2}{\alpha}\beta\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}u^{\beta-1}\, \mathrm{d}u.\end{aligned}$$

Remarking that $f(u)^{-\alpha}u^{\beta-1} = \underset{u\to+\infty}{o}(f(u)^{-\alpha})$, since $\beta < 1$, we deduce that

$$\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}u^{\beta-1}\, \mathrm{d}u = \underset{\varepsilon\to 0}{o}\left(\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}\, \mathrm{d}u\right).$$

We obtain that

$$\int_{t_0}^{t/\varepsilon} f(u)^{-\alpha}\, \mathrm{d}u \underset{\varepsilon\to 0}{\sim} \frac{2}{\alpha} f\left(\frac{t}{\varepsilon}\right)^{-\alpha}\left(\frac{t}{\varepsilon}\right)^{\beta}$$

This ends the proof. □

Lemma 10.24. *Let $(X_n)_n$ be a sequence of random variables with values in $\mathbb{R}^2$, and which converges in distribution to a random variable X. We assume that the distribution of X is invariant under rotations, i.e. for any orthogonal matrix $R \in \mathcal{M}_2(\mathbb{R})$, the random variables X and RX have the same distribution. Then for all sequence $(R_n)_n$ of orthogonal matrices in $\mathcal{M}_2(\mathbb{R})$, we have*

$$R_n X_n \underset{n\to+\infty}{\Longrightarrow} X.$$

Proof. Let us denote by ϕ_Z the characteristic function of a random variable Z. Using Theorem 5.3 p. 86 in [Kal02], we know that $(\phi_{X_n})_n$ converges to ϕ_X uniformly on every compact subset of $\mathbb{R}^2$. The characteristic function of the random variable $Y_n := R_n X_n$ is given by

$$\xi \mapsto \phi_{Y_n}(\xi) = \phi_{X_n}(R_n\xi).$$

Thus, by assumption, we have, for all $\xi \in \mathbb{R}^2$,

$$\phi_X(R_n\xi) = \phi_X(\xi).$$

It follows that, for any $\xi \in \mathbb{R}^2$ and $n \geq 0$,

$$|\phi_{Y_n}(\xi) - \phi(\xi)| = |\phi_{X_n}(R_n\xi) - \phi_X(R_n\xi)| \leq \sup_{z\in\mathbb{R}^2, \|z\|=\|\xi\|} |\phi_{X_n}(z) - \phi_X(z)|,$$

which converges to 0, as $n \to +\infty$. This ends the proof of the lemma. □

Bibliography

[AK94] S. Albeverio and A. Klar. Long time behavior of nonlinear stochastic oscillators: The one-dimensional Hamiltonian case. *J. Math. Phys.*, 35(8):4005–4027, August 1994.

[App09] David Applebaum. *Levy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2009.

[Bas11] Richard F. Bass. *Stochastic processes*. Cambridge University Press, 2011.

[BFY15] Alain Bensoussan, Jens Frehse, and Sheung Chi Phillip Yam. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées*, 103(6):1441 – 1474, 2015.

[Bil99] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Probability and Statistics Section. Wiley, 2nd ed edition, 1999.

[BLPR17] Rainer Buckdahn, Juan Li, Shige Peng, and Catherine Rainer. Mean-field stochastic differential equations and associated pdes. *Ann. Probab.*, 45(2):824–878, 03 2017.

[Bre10] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer New York, 2010.

[Car10] Pierre Cardaliaguet. Notes on Mean Field Games, from P.L. Lions lectures at Collège de France, 2010.

[Cav21] Thomas Cavallazzi. Itô-Krylov's formula for a flow of measures. arXiv:2110.05251, 2021.

[Cav22a] Thomas Cavallazzi. Ito's formula for the flow of measures of Poisson stochastic integrals and applications. arXiv:2211.16131, 2022.

[Cav22b] Thomas Cavallazzi. Quantitative weak propagation of chaos for stable-driven McKean-Vlasov SDEs. arXiv:2212.01079, 2022.

[Cav23] Thomas Cavallazzi. Well-posedness and propagation of chaos for Lévy-driven McKean-Vlasov SDEs under Lipschitz assumptions. arXiv:2301.08594, 2023.

[CCD15] Jean-François Chassagneux, Dan Crisan, and François Delarue. A probabilistic approach to classical solutions of the master equation for large population equilibria. hal-01144845, 2015.

[CD14] René Carmona and Francois Delarue. *The Master Equation for Large Population Equilibriums*. Stochastic Analysis and Applications 2014. Springer, 2014.

[CD15] René Carmona and François Delarue. Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics. *The Annals of Probability*, 43(5):2647 – 2700, 2015.

[CD18a] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games with Applications I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, 2018.

[CD18b] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games with Applications II*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, 2018.

[CD22a] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. arXiv:2203.00446, May 2022.

[CD22b] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. arXiv:2106.14812, May 2022.

[CDLL19] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre Louis Lions. *The master equation and the convergence problem in mean field games*, volume 381 of *AMS-201*. Princeton University Press, 2019.

[CdR17] P. E. Chaudru de Raynal. Strong existence and uniqueness for degenerate SDE with Hölder drift. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 53(1):259–286, February 2017.

[CdRF21] Paul-Eric Chaudru de Raynal and Noufel Frikha. From the backward Kolmogorov PDE on the Wasserstein space to propagation of chaos for Mckean-Vlasov SDEs. *Journal de Mathématiques Pures et Appliquées*, 156:1–124, 2021.

[CdRF22] Paul-Eric Chaudru de Raynal and Noufel Frikha. Well-Posedness for Some Non-Linear Diffusion Processes and Related PDE on the Wasserstein Space. *Journal de Mathématiques Pures et Appliquées*, 153:1–167, 2022.

[CdRJM23] Paul-Eric Chaudru de Raynal, Jean-François Jabir, and Stéphane Menozzi. Multidimensional Stable Driven McKean-Vlasov SDEs with Distributional Interaction Kernel: Critical Thresholds and Related Models. arXiv:2205.11866, May 2023.

[CHM06] Peter E. Caines, Minyi Huang, and Roland P. Malhamé. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information and Systems*, 6(3):221 – 252, 2006.

[CHZ20a] Zhen-Qing Chen, Zimo Hao, and Xicheng Zhang. Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes. *Electronic Journal of Probability*, 25:1 – 23, 2020.

[CHZ20b] Zhen-Qing Chen, Zimo Hao, and Xicheng Zhang. Hölder regularity and gradient estimates for SDEs driven by cylindrical α-stable processes. *Electronic Journal of Probability*, 25:1–23, 2020.

[CL23] Thomas Cavallazzi and Emeline Luirard. Scaling limit of a kinetic inhomogeneous stochastic system in the quadratic potential. arXiv:2206.13276, accepted for publication in *Discrete and Continuous Dynamical Systems*, 2023.

[CM17] Dan Crisan and Eamon McMurray. Smoothing properties of McKean-Vlasov SDEs. *Probab. Theory Related Fields*, 171:97–148, 2017. arXiv:1702.01397.

[CST22] Jean-François Chassagneux, Lukasz Szpruch, and Alvin Tse. Weak quantitative propagation of chaos via differential calculus on the space of measures. *The Annals of Applied Probability*, 32(3):1929 – 1969, 2022.

[CSZ18] Zhen-Qing Chen, Renming Song, and Xicheng Zhang. Stochastic flows for lévy processes with hölder drifts. *Revista Matematica Iberoamericana*, 34(4), 2018.

[Dit99] Peter D. Ditlevsen. Observation of α-stable noise induced millennial climate changes from an ice-core record. *Geophysical Research Letters*, 26:1441–1444, 1999.

[dRP22] Gonçalo dos Reis and Vadim Platonov. Itô-Wentzell-Lions Formula for Measure Dependent Random Fields under Full and Conditional Measure Flows. *Potential Analysis*, 2022.

[DT21] François Delarue and Alvin Tse. Uniform in time weak propagation of chaos on the torus. arXiv:2104.14973, 2021.

[Ebe16] Andreas Eberle. Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases. *Probability Theory and Related Fields*, 166:851– 886, 2016.

[FFPV17] Ennio Fedrizzi, Franco Flandoli, Enrico Priola, and Julien Vovelle. Regularity of stochastic kinetic equations. *Electron. J. Probab.*, 22, 2017.

[FG15] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707, August 2015.

[FIR17] Franco Flandoli, Elena Issoglio, and Francesco Russo. Multidimensional stochastic differential equations with distributional drift. *Transactions of the American Mathematical Society*, 369:1665–1688, 2017.

[FKM21] Noufel Frikha, Valentin Konakov, and Stéphane Menozzi. Well-posedness of some non-linear stable driven SDEs. *Discrete and Continuous Dynamical Systems - Series A*, 41(2):849–898, 2021.

[FL21] Noufel Frikha and Libo Li. Well-posedness and approximation of some one-dimensional Lévy-driven non-linear SDEs. *Stochastic Processes and their Applications*, 132:76–107, February 2021.

[Fla11] Franco Flandoli. *Random perturbation of PDEs and fluid dynamic models*. Volume 2015 of Lecture Notes in Mathematics. Springer, Heidelberg, 2011.

[Fou13] Nicolas Fournier. On pathwise uniqueness for stochastic differential equations driven by stable Lévy processes. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, 49(1):138–159, 2013.

[FT21] Nicolas Fournier and Camille Tardif. One dimensional critical Kinetic Fokker-Planck equations, Bessel and stable processes. *Communications in Mathematical Physics*, 381, January 2021.

[GBM22] Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases. *Electronic Journal of Probability*, 27:1 – 44, 2022.

[GL21a] Mihai Gradinaru and Emeline Luirard. Asymptotic behavior for a time-inhomogeneous Kolmogorov type diffusion. to appear in ESAIM : PS; arXiv:2004.11576, July 2021.

[GL21b] Mihai Gradinaru and Emeline Luirard. Kinetic time-inhomogeneous lévy-driven model. arXiv:2112.07287, December 2021.

[GO13] Mihai Gradinaru and Yoann Offret. Existence and asymptotic behaviour of some time-inhomogeneous diffusions. *Ann. Inst. H. Poincaré Probab. Statist.*, 49(1):182–207, February 2013.

[GPW20] Xin Guo, Huyên Pham, and Xiaoli Wei. Itô's formula for flow of measures on semimartingales. arXiv:2010.05288, 2020.

[Gra92a] Carl Graham. McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications*, 40(1):69–82, 1992.

[Gra92b] Carl Graham. Nonlinear diffusion with jumps. *Annales de l'I.H.P.*, 28(3):393–402, 1992.

[Gä88] Jürgen Gärtner. On the McKean-Vlasov Limit for Interacting Diffusions. *Mathematische Nachrichten*, 137(1):197–248, 1988.

[Han22] Yi Han. Entropic propagation of chaos for mean field diffusion with L^p interactions via hierarchy, linear growth and fractional noise. arXiv:2205.02772, 2022.

[HHMT20] Jasper Hoeksema, Thomas Holding, Mario Maurelli, and Oliver Tse. Large deviations for singularly interacting diffusions. arXiv:2002.01295, 2020.

[HMC18] Igor Honore, Stéphane Menozzi, and Paul-Eric Chaudru de Raynal. Strong regularization by Brownian noise propagating through a weak Hörmander structure, October 2018.

[Jab19] Jean-Francois Jabir. Rate of propagation of chaos for diffusive stochastic particle systems via girsanov transformation. arXiv:1907.09096, 2019.

[JMW05] Benjamin Jourdain, Sylvie Méléard, and Wojbor A. Woyczynski. A Probabilistic Approach for Nonlinear Equations Involving the Fractional Laplacian and a Singular Operator. *Potential Analysis*, 23:55–81, 2005.

[JMW07] Benjamin Jourdain, Sylvie Méléard, and Wojbor Woyczynski. Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:1–29, 2007. arXiv:0707.2723.

[Jou97] Benjamin Jourdain. Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized burgers' equations. *ESAIM: PS*, 1:339–355, 1997.

[JT21] Benjamin Jourdain and Alvin Tse. Central limit theorem over non-linear functionals of empirical measures with applications to the mean-field fluctuation of interacting particle systems. *Electronic Journal of Probability*, 26, 2021.

[JW18] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Inventiones mathematicae*, 214(1):523–591, 2018.

[Kac56] Mark Kac. Foundations of kinetic theory. *in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 3:171–197, 1956.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability.* Probability and Its Applications. Springer New York, New York, NY, 2002.

[KK18] Victoria Knopova and Alexei Kulik. Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(1):100 – 140, 2018.

[Kry09] Nikolaj V. Krylov. *Controlled Diffusion Processes.* Number 14 in Stochastic Modelling and Applied Probability. Springer, 2009.

[KS98] Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1998.

[Lac18] Daniel Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electronic Communications in Probability*, 23(none):1 – 11, 2018.

[Lac21] Daniel Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. arXiv:2105.02983, 2021.

[Lan08] Paul Langevin. Sur la théorie du mouvement brownien. *Comptes-Rendus de l'Académie des Sciences*, 146:530–532, January 1908.

[Lio] Pierre-Louis Lions. Cours au Collège de France. `https://www.college-de-france.fr/site/pierre-louis-lions/_course.htm`.

[LL07] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[LP08] Harald Luschgy and Gilles Pagès. Moment estimates for Lévy Processes. *Electronic Journal of Probability*, 13:422–434, 2008.

[Mal03] Florent Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Annals of Applied Probability*, 13(2):540–560, 2003.

[Mao11] Xuerong Mao. *Stochastic Differential Equations and Applications.* Woodhead Publishing, second edition edition, 2011.

[Mar20] Victor Marx. Infinite-dimensional regularization of mckean-vlasov equation with a wasserstein diffusion. arXiv:2002.10157, 2020.

[McK66] Henry P McKean. A class of markov processes associated with nonlinear parabolic equations. *Proc Natl Acad Sci U S A*, 56:1907–1911, 1966.

[McK67] Henry P McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967.

[MJCB14] Ralf Metzler, Jae-Hyung Jeon, Andrey G. Cherstvy, and Eli Barkai. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.*, 16:24128–24164, 2014.

[MJW01] J. Adin Mann Jr. and Wojbor A. Woyczynski. Growing fractal interfaces in the presence of self-similar hopping surface diffusion. *Physica A: Statistical Mechanics and its Applications*, 291:159–183, 2001.

[MM21] L. Marino and S. Menozzi. Weak well-posedness for degenerate SDEs driven by Lévy processes, December 2021.

[MMW15] Stéphane Mischler, Clément Mouhot, and Bernt Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probability Theory and Related Fields*, 161(1-2):1–59, 2015.

[MP14] Remigijus Mikulevicius and Henrikas Pragarauskas. On the Cauchy problem for integro-differential operators in Hölder classes and the uniqueness of the martingale problem. *Potential Analysis*, 40:539–563, 2014.

[MV21] Yuliya Mishura and Alexander Veretennikov. Existence and uniqueness theorems for solutions of mckean–vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*, 103:59–101, 2021.

[MZ20] Chenchen Mou and Jianfeng Zhang. Wellposedness of second order master equations for mean field games with nonsmooth data. arXiv:1903.09907, 2020.

[MZ22] Stéphane Menozzi and Xicheng Zhang. Heat kernel of supercritical SDEs with unbounded drifts. arXiv:2012.14775, 2022.

[Mé96] Sylvie Méléard. *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, pages 42–95. Springer Berlin Heidelberg, 1996.

[NBK+20] Neelima, Sani Biswas, Chaman Kumar, Gonçalo dos Reis, and Christoph Reisinger. Well-posedness and tamed Euler schemes for McKean-Vlasov equations driven by Lévy noise. arXiv:2010.08585, 2020.

[Pro05] Philip E. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin Heidelberg, second edition, 2005.

[Rud87] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, 3rd ed edition, 1987.

[RZ21] Michael Röckner and Xicheng Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27(2):1131 – 1158, 2021.

[Sat99] Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. Number 68 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[Sch87] Michael Scheutzow. Uniqueness and non-uniqueness of solutions of vlasov-mckean equations. *Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics*, 43(2):246–256, 1987.

[Sit05] Rong Situ. *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering*. Mathematical and Analytical Techniques with Applications to Engineering. Springer US, 2005.

[ST85] T. Shiga and H. Tanaka. Central limit theorem for a system of markovian particles with mean field interactions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 69(3):439–459, 1985.

[SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer-Verlag Berlin Heidelberg New-York, 1979.

[Szn91] Alain-Sol Sznitman. *Topics in Propagation of Chaos*, volume 1464 of *Lecture Notes in Mathematics*, pages 165–251. Springer Berlin Heidelberg, 1991.

[Tom20] Milica Tomašević. Propagation of chaos for stochastic particle systems with singular mean-field interaction of $L^q - L^p$ type. hal-03086253f, December 2020.

[TTZ21] Mehdi Talbi, Nizar Touzi, and Jianfeng Zhang. Dynamic programming equation for the mean field optimal stopping problem. arXiv:2103.05736, 2021.

[Ver81] A J Veretennikov. On strong solutions and explicit formulas for solutions of stochastic intregral equations. *Mathematics of the USSR-Sbornik*, 39(3):387, apr 1981.

[Vil09] Cédric Villani. *Optimal transportation : old and new*. Springer, 2009.

[Wat44] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambrige University Press, second edition, 1944.

[Wu01] Liming Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Processes and their Applications*, 91(2):205–238, February 2001.

[WZ15] Feng-Yu Wang and Xicheng Zhang. Degenerate SDEs in Hilbert Spaces with Rough Drifts, January 2015.

[Zha14] Xicheng Zhang. Densities for sdes driven by degenerate α-stable processes. *The Annals of Probability*, 42(5):1885–1910, September 2014.

[Zha16] Xicheng Zhang. Stochastic Hamiltonian flows with singular coefficients, June 2016.

[Zvo74] A. K. Zvonkin. A Transformation of the Phase Space of a Diffusion Process that removes the Drift. *Sbornik: Mathematics*, 22(1):129–149, February 1974.

COLLEGE DOCTORAL BRETAGNE MATHS, TELECOMS INFORMATIQUE, SIGNAL SYSTEMES, ELECTRONIQUE

Université de Rennes

Titre : Équations différentielles stochastiques dirigées par des bruits de Lévy : systèmes de particules en interaction de type champ moyen et processus de McKean-Vlasov

Mot clés : Équations différentielles stochastiques, processus de Lévy, processus de McKean-Vlasov, champ moyen, propagation du chaos, formule d'Itô le long d'un flot de mesures

Résumé : Cette thèse porte en grande partie sur l'étude des Équations Différentielles Stochastiques (EDS) non-linéaires au sens de McKean-Vlasov. Les bruits directeurs que nous considérons sont des processus de Lévy, en grande majorité des processus stables. On s'intéresse à quantifier la propagation du chaos au sens faible pour le système de particules en interaction de type champ moyen associé, sous des hypothèses höldériennes sur les coefficients. Cela se fait à travers l'étude du semi-groupe, agissant sur les fonctions définies sur l'espace des mesures de probabilité, associé à l'EDS de McKean-Vlasov. En particulier, on exhibe des propriétés régularisantes du semi-groupe, et on décrit sa dynamique grâce à la formule d'Itô le long d'un flot de mesures de probabilités. Cette formule est l'un des outils importants de cette thèse. Premièrement, on la prouve pour une grande classe de processus à sauts. Deuxièmement, on prouve, grâce à l'inégalité de Krylov, la formule d'Itô le long d'un flot de mesures pour des fonctions appartenant à un espace de type Sobolev. Dans la dernière partie de cette thèse, on s'intéresse à un système cinétique inhomogène en temps spécifique, qui est dirigé par un processus stable. On étudie son comportement asymptotique après changement d'échelle.

Title: Lévy-driven Stochastic differential equations: mean-field interacting particle systems and McKean-Vlasov processes

Keywords: Stochastic differential equations, Lévy processes, McKean-Vlasov processes, mean-field, propagation of chaos, Ito's formula along a flow of measures

Abstract: This thesis is mainly devoted to the study of nonlinear Stochastic Differential Equations (SDE), in the sense of McKean-Vlasov. The driving noises that we consider are Lévy processes, mostly stable processes. We are interested in quantifying the propagation of chaos, in the weak sense, for the associated mean-field interacting particle system, under Hölder assumptions on the coefficients. This is achieved through the study of the semigroup, acting on functions defined on the space of probability measures, associated with the McKean-Vlasov SDE. In particular, we exhibit regularizing properties of the semigroup, and we describe its dynamics thanks to Itô's formula along a flow of probability measures. This formula is one of the crucial tools of this thesis. Firstly, we prove it for a large class of jump processes. Secondly, we prove, relying on Krylov's inequality, Itô's formula along a flow of probability measures for functions belonging to a Sobolev-type space. In the last part of this thesis, we are interested in a specific time-inhomogenous kinetic system, which is driven by a stable process. We study its asymptotic behavior after rescaling.