Well-posedness and propagation of chaos for Lévy-driven McKean-Vlasov SDEs under Lipschitz assumptions

Thomas Cavallazzi

WELL-POSEDNESS AND PROPAGATION OF CHAOS FOR LÉVY-DRIVEN MCKEAN-VLASOV SDES UNDER LIPSCHITZ ASSUMPTIONS

THOMAS CAVALLAZZI

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.

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 [15, 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

$$\begin{cases} dX_t = b_t(X_t, \mu_t)\, dt + \sigma_t(X_{t^-}, \mu_t)\, dZ_t, \quad t \in [0,T], \\ \mu_t := [X_t], \\ X_0 = \xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d), \end{cases} \tag{1.1}$$

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, $\mathcal{M}_d(\mathbb{R})$

Date: January 19, 2023.

2000 *Mathematics Subject Classification.* 60H10, 60G51, 60H30.

Key words and phrases. McKean-Vlasov stochastic differential equations, Lévy processes, propagation of chaos.

being the space of matrices of size $d \times d$ on $\mathbb{R}$. The first motivation to study (1.1) lies into its connexion with the following mean-field interacting particle system

$$\begin{cases} dX_t^{i,N} = b_t(X_t^{i,N}, \overline{\mu}_t^N)\, dt + \sigma_t(X_{t^-}^{i,N}, \overline{\mu}_t^N)\, dZ_t^i, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ \overline{\mu}_t^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}, \\ X_0^{i,N} = \xi^i, \end{cases} \tag{1.2}$$

where $(Z^i, \xi^i)_{i\geq 1}$ are i.i.d. with same distribution as (Z, ξ). The link between (1.1) and (1.2) is that for any $k \geq 1$, the dynamics of k particles is expected to be described by k independent copies of(1.1) 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 [13] and then investigated by Sznitman [16] when Z is a Brownian motion. For a detailed review on the topic of propagation of chaos, we refer the reader to [4, 5].

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

$$|b_t(x,\mu) - b_t(y,\nu)| + |\sigma_t(x,\mu) - \sigma_t(y,\nu)| \leq C(|x-y| + W_\beta(\mu,\nu)), \tag{1.3}$$

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

1.1. Well-posedness of the McKean-Vlasov SDE (1.1).

Theorem 1. Under Assumption (H1), there exists a unique strong solution $(X_t)_{t\in[0,T]}$ to (1.1) 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

$$\mathbb{E}\sup_{t\leq T} |X_t|^\beta < +\infty. \tag{1.4}$$

Remark 1. 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 [1, Theorem 2.4.16], we can thus consider a general Lévy process Z having a finite moment of order $\beta \in [1,2]$.

Let us compare our result with the existing literature. When $\beta = 2$, the well-posedness of (1.1) was proved by Jourdain, Méléard and Woyczynski [12]. 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 1 is proved by Graham in [11, Theorem 2.2]. The main differences are the following. Firstly, in [11], 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 [11] that X_0 has a finite moment of order 2, which is not the case in Theorem 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

$$\left|\int_{B_1^c} \sigma_t(x,\mu) z \, d\nu(z)\right|^2 + \int_{B_1^c} |\sigma_t(x,\mu)z|^2 \, d\nu(z) \leq C(1+|x|^2). \tag{1.5}$$

It suggests that σ is bounded with respect to its measure variable, which is not the case in Theorem 1. Moreover, (1.5) 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 1 since $\beta \in [1,2]$. In the non-degenerate case, i.e. when σ is uniformly elliptic, we refer to [8]. In this work, Frikha, Menozzi and Konakov prove the well-posedness of (1.1) 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 1, we do not assume that the diffusion coefficient σ is uniformly elliptic. Moreover, another assumption made in [8] 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 2. Notice that when $\beta \in (0,1)$, the uniqueness result of Theorem 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

$$\begin{cases} y'(t) = |y(t)|^\beta, \quad t \in [0,T], \\ y(0) = 0. \end{cases}$$

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 (1.1). We refer to Appendix A for a proof of this result.

1.2. **Propagation of chaos for the interacting particle system (1.2).** We now focus on the propagation of chaos for the interacting particle system (1.2). Under Assumption (H1), the SDE (1.2) admits a unique strong solution by [1, 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 (1.1), 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 $(Z^i)_{i\geq 1}$.

Theorem 2. Assume that Assumption (H1) holds true with W_1 instead of W_β in the Lipschitz control (1.3) and with $\sigma = \mathrm{Id}$. Then, there exists a constant $C > 0$, independent of d and N, such that for all $N \geq 1$

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_t^{i,N} - X_t^{i,\infty}| \leq C \begin{cases} N^{\frac{1}{\beta}-1}, & \text{if} \quad d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \beta < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \beta > \frac{d}{d-1}, \end{cases} \tag{1.6}$$

and

$$\sup_{t\in[0,T]} \mathbb{E} W_1(\overline{\mu}_t^N, \mu_t) \leq C \begin{cases} N^{\frac{1}{\beta}-1}, & \text{if} \quad d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \beta < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \beta > \frac{d}{d-1}. \end{cases} \tag{1.7}$$

Remark 3. The method used in the proof of Theorem 2 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 [10], 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 [12] 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. [14], where quantitative strong propagation of chaos is proved in L^2, relaxing the assumptions of [12]. In the one-dimensional case, Frikha and Li [9] 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 2. 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 (1.2) and the limiting McKean-Vlasov SDE (1.1) 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 2, we have the following quantitative propagation of chaos result.

Theorem 3. There exists a positive constant C independent of d and N such that for all $N \geq 1$

$$\sup_{i\leq N} \mathbb{E} \sup_{t\leq T} |X_t^{i,N} - X_t^{i,\infty}| \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}, \end{cases} \tag{1.8}$$

and

$$\sup_{t\in[0,T]} \mathbb{E} W_1(\overline{\mu}_t^N, \mu_t) \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1,2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}. \end{cases} \tag{1.9}$$

2. Proof of Theorem 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 [1, Theorem 6.2.9], we deduce that the SDE

$$\begin{cases} dX_t^\mu = b_t(X_t^\mu, \mu_t)\, dt + \sigma_t(X_{t^-}^\mu, \mu_t)\, dZ_t, \quad t \in [0,T], \\ X_0^\mu = \xi \in L^\beta(\Omega, \mathcal{F}_0; \mathbb{R}^d), \end{cases} \tag{2.1}$$

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 [6], we get that

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

The map

$$\phi : \begin{cases} \mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d)) & \to \mathcal{C}^0([0,T];\mathcal{P}_\beta(\mathbb{R}^d)) \\ \mu & \mapsto ([X_t^\mu])_{t\in[0,T]} \end{cases} \tag{2.2}$$

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 (1.1). 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 $\mathbb{E}\sup_{s\le t}|X_s^\mu - X_s^\nu|^\beta$, for $t\in[0,T]$. We employ the method used by Fournier in the proof of [6, Proposition 2], which was used in the context of McKean-Vlasov SDEs by Frikha and Li in [9] to prove the moment estimation (1.4). The first step is to consider the SDE (2.1) 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 $t\in[0,T]$

$$X_t^\mu = \xi_1 + \int_0^t b_s(X_s^\mu,\mu_s)\,ds + \int_0^t\int_{B_1} \sigma_s(X_{s^-}^\mu,\mu_s) z\,\widetilde{\mathcal{N}}(ds,dz), \tag{2.3}$$

and

$$X_t^\nu = \xi_2 + \int_0^t b_s(X_s^\nu,\nu_s)\,ds + \int_0^t\int_{B_1} \sigma_s(X_{s^-}^\nu,\nu_s) z\,\widetilde{\mathcal{N}}(ds,dz). \tag{2.4}$$

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 Burkholder-Davis-Gundy (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 $t\in[0,T]$

$$\mathbb{E}\left(\sup_{s\le t}|X_s^\mu - X_s^\nu|^2 \mid \xi_1,\xi_2\right) \\ \le C\left[|\xi_1-\xi_2|^2 + \int_0^t \mathbb{E}\left(|X_s^\mu - X_s^\nu|^2 \mid \xi_1,\xi_2\right)\,ds + \int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right].$$

Gronwall's lemma ensures that

$$\mathbb{E}\left(\sup_{s\le t}|X_s^\mu - X_s^\nu|^2 \mid \xi_1,\xi_2\right) \le C|\xi_1-\xi_2|^2 + C\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds.$$

It follows from Jensen's inequality that

$$\mathbb{E}\left(\sup_{s\le t}|X_s^\mu - X_s^\nu|^\beta \mid \xi_1,\xi_2\right) \le \left(\mathbb{E}\left(\sup_{s\le t}|X_s^\mu - X_s^\nu|^2 \mid \xi_1,\xi_2\right)\right)^{\frac{\beta}{2}} \\ \le C|\xi_1-\xi_2|^\beta + C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}.$$

Taking the expectation yields for all $t\in[0,T]$

$$\mathbb{E}\left(\sup_{s\le t}|X_s^\mu - X_s^\nu|^\beta\right) \le C\mathbb{E}|\xi_1-\xi_2|^\beta + C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}. \tag{2.5}$$

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 (2.1) are

respectively solutions to (2.3) and (2.4) with $\xi_1 = \xi_2$. Thus, using (2.5) with the conditional expectation with respect to $\mathcal{G}$ instead of the expectation, we deduce that

$$\mathbb{E}\left(\sup_{s<t\wedge T_1}|X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right) \leq C\left(\int_0^{t\wedge T_1} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}. \tag{2.6}$$

Let us now deal with the first big jump of Z, which occurs at time T_1. We have

$$X_{T_1}^\mu - X_{T_1}^\nu = X_{T_1^-}^\mu - X_{T_1^-}^\nu + \left(\sigma_{T_1}(X_{T_1^-}^\mu,\mu_{T_1}) - \sigma_{T_1}(X_{T_1^-}^\nu,\nu_{T_1})\right)\Delta Z_1.$$

It follows from the Lipschitz assumption on σ that

$$|X_{T_1}^\mu - X_{T_1}^\nu|^\beta \leq C|X_{T_1^-}^\mu - X_{T_1^-}^\nu|^\beta(1+|\Delta Z_1|^\beta) + W_\beta^\beta(\mu_{T_1},\nu_{T_1})|\Delta Z_1|^\beta.$$

Since ΔZ_1 is independent of $\mathcal{G}$ and $\mathbb{E}|\Delta Z_1|^\beta < +\infty$, we deduce by (2.6) that almost surely on the set $\{T_1 \leq T\}$

$$\mathbb{E}\left(|X_{T_1}^\mu - X_{T_1}^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1\leq t} \leq C\left[\left(\int_0^{t\wedge T_1} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + W_\beta^\beta(\mu_{T_1},\nu_{T_1})\right].$$

We thus have by the preceding inequality and (2.6)

$$\begin{aligned}
&\mathbb{E}\left(|X_{t\wedge T_1}^\mu - X_{t\wedge T_1}^\nu|^\beta \mid \mathcal{G}\right)\\
&= \mathbb{E}\left(|X_{T_1}^\mu - X_{T_1}^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1\leq t} + \mathbb{E}\left(|X_t^\mu - X_t^\nu|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_1>t}\\
&\leq C\left[\left(\int_0^{t\wedge T_1} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + W_\beta^\beta(\mu_{T_1},\nu_{T_1})\right].
\end{aligned}$$

Following the same lines and using (2.5), we prove that for any $n \geq 1$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\leq s<t\wedge T_{n+1}}|X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\\
&\qquad\qquad \leq C\left[\mathbb{E}\left(|X_{t\wedge T_n}^\mu - X_{t\wedge T_n}^\nu|^\beta \mid \mathcal{G}\right) + \left(\int_{t\wedge T_n}^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}\right],
\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\leq s\leq t\wedge T_{n+1}}|X_s^\mu - X_s^\nu|^\beta \mid \mathcal{G}\right)\\
&\quad \leq C\left[\mathbb{E}\left(|X_{t\wedge T_n}^\mu - X_{t\wedge T_n}^\nu|^\beta \mid \mathcal{G}\right) + \left(\int_{t\wedge T_n}^{t\wedge T_{n+1}} W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + W_\beta^2(\mu_{T_{n+1}},\nu_{T_{n+1}})\right],
\end{aligned} \tag{2.8}$$

where C is independent of n. Reasoning by induction, we deduce that for a constant $C > 1$ depending only on T, we have

$$\begin{aligned}
&\mathbb{E}\left(\sup_{t\wedge T_n\le s<t\wedge T_{n+1}}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\\
&\le C^{n+1}\left[\sum_{k=0}^{n}\left(\int_{t\wedge T_k}^{t\wedge T_{k+1}} W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W^\beta_\beta(\mu_{T_k},\nu_{T_k})\right]\\
&\le C^{n+1}\left[(n+1)^{1-\frac{\beta}{2}}\left(\int_0^{t\wedge T_{n+1}} W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W^\beta_\beta(\mu_{T_k},\nu_{T_k})\right],
\end{aligned}$$

by Jensen's inequality and with the convention that $T_0 = 0$. Thus, for a certain constant $K > C$, one has for all $j \le n$

$$\mathbb{E}\left(\sup_{t\wedge T_j\le s<t\wedge T_{j+1}}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right) \le K^{j+2}\left[\left(\int_0^{t\wedge T_{n+1}} W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W^\beta_\beta(\mu_{T_k},\nu_{T_k})\right].$$

Summing the preceding inequality over $j \in \{0,\dots,n\}$, we deduce that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s<t\wedge T_{n+1}}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\\
&\qquad\le \frac{K^{n+3}}{1-K}\left[\left(\int_0^{t\wedge T_{n+1}} W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^{n} W^\beta_\beta(\mu_{T_k},\nu_{T_k})\right]. \quad (2.9)
\end{aligned}$$

Let us denote by $(N_t)_{t\ge0}$ the Poisson process associated with the jumping times $(T_n)_{n\ge1}$ which has an intensity $\lambda = \nu(B_1^c)$. One has

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\mathbf{1}_{N_t=n}\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\mathbf{1}_{t<T_1}\right) + \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_n\le t<T_{n+1}}\right)\\
&\le \mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s<t\wedge T_1}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\mathbf{1}_{t<T_1}\right)\\
&\quad + \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\sup_{0\le s<t\wedge T_{n+1}}|X^\mu_s - X^\nu_s|^\beta \mid \mathcal{G}\right)\mathbf{1}_{T_n\le t<T_{n+1}}\right)
\end{aligned}$$

Using (2.6) and (2.9), we obtain that for any $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta\right)\\
&\le C\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\right)^{\frac{\beta}{2}}\\
&\quad+\sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\mathbb{E}\left(\left[\left(\int_0^{t\wedge T_{n+1}} W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \sum_{k=1}^n W^\beta_\beta(\mu_{T_k},\nu_{T_k})\right]\mathbf{1}_{T_n\le t<T_{n+1}}\right) \qquad (2.10)\\
&\le C\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\right)^{\frac{\beta}{2}}\\
&\quad+\sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\mathbb{P}(N_t=n)\left(\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \mathbb{E}\left(\sum_{k=1}^n W^\beta_\beta(\mu_{T_k},\nu_{T_k})\;\middle|\; N_t=n\right)\right).
\end{aligned}$$

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

$$(t_1,\dots,t_n)\in[0,t]^n \mapsto \frac{n!}{t^n}\mathbf{1}_{t_1<\dots<t_n}.$$

This yields

$$\begin{aligned}
\mathbb{E}\left(\sum_{k=1}^n W^\beta_\beta(\mu_{T_k},\nu_{T_k})\;\middle|\; N_t=n\right) &= \int_{[0,t]^n}\frac{n!}{t^n}\mathbf{1}_{t_1<\dots<t_n}\sum_{k=1}^n W^\beta_\beta(\mu_{t_k},\nu_{t_k})\,dt_1\dots\,dt_n\\
&= \int_{[0,t]^n}\frac{1}{t^n}\sum_{k=1}^n W^\beta_\beta(\mu_{t_k},\nu_{t_k})\,dt_1\dots\,dt_n\\
&= \sum_{k=1}^n \frac{1}{t^n}t^{n-1}\int_0^t W^\beta_\beta(\mu_s,\nu_s)\,ds\\
&= \frac{n}{t}\int_0^t W^\beta_\beta(\mu_s,\nu_s)\,ds.
\end{aligned}$$

Injecting this equality in (2.10), we get

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta\right)\\
&\le\sum_{n=1}^\infty \frac{K^{n+3}}{1-K}\frac{(\lambda t)^n}{n!}e^{-\lambda t}\left(\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \frac{n}{t}\int_0^t W^\beta_\beta(\mu_s,\nu_s)\,ds\right)\\
&\quad + C\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\right)^{\frac{\beta}{2}}.
\end{aligned}$$

This proves the existence of a constant $C>0$ depending only on T such that for all $t\in[0,T]$

$$\mathbb{E}\left(\sup_{0\le s\le t}|X^\mu_s - X^\nu_s|^\beta\right) \le C\left[\left(\int_0^t W^2_\beta(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}} + \int_0^t W^\beta_\beta(\mu_s,\nu_s)\,ds\right]. \qquad (2.11)$$

Note that (2.11) is true if $\beta \in (0,1)$ since we have only used that $0 < \beta \le 2$. Changing again the constant C, Hölder's inequality yields for all $t \in [0,T]$

$$\sup_{0\le s\le t} W_\beta^\beta(\phi(\mu)_s, \phi(\nu)_s) \le \mathbb{E}\left(\sup_{0\le s\le t} |X_s^\mu - X_s^\nu|^\beta\right) \le C\left(\int_0^t W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}. \tag{2.12}$$

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.

3. Proof of Theorem 2 and Theorem 3

Proof of Theorem 2. To prove (1.6), 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 $N \ge 1$

$$\sup_{i\le N} \mathbb{E}\sup_{t\le T} |X_t^{i,N} - X_t^{i,\infty}| \le C\int_0^T \mathbb{E}W_1(\tilde{\mu}_s^N, \mu_s)\,ds. \tag{3.1}$$

We conclude using [7, 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 (1.7) follows from (1.6) and [7] 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 3. Let us prove (1.8). 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 $i \geq 1$ and $t \in [0,T]$

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

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 $i \in \{1,\dots,N\}$, $X^{i,\infty}_N$ as the unique solution to

$$\begin{cases} dX^{i,\infty}_{N,t} &= AX^{i,\infty}_{N,t}\,dt + A'\mathbb{E}X^{i,\infty}_{N,t}\,dt + B\,dZ^i_{N,t}, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ \mu_{N,t} &:= [X^{i,\infty}_{N,t}], \\ X^{i,\infty}_{N,0} &= \xi^i. \end{cases} \tag{3.2}$$

For any $N \geq 1$ fixed, the random variables $(X^{i,\infty}_N)_{i\leq N}$ are i.i.d. We proceed similarly for the particle system by defining $(X^{i,N}_N)_{i\leq N}$ as the unique solution to

$$\begin{cases} dX^{i,N}_{N,t} = AX^{i,N}_{N,t}\,dt + A'\frac{1}{N}\sum_{k=1}^N X^{k,N}_{N,t}\,dt + B\,dZ^i_{N,t}, \quad t \in [0,T], \quad i \in \{1,\dots,N\}, \\ \overline{\mu}^N_{N,t} := \frac{1}{N}\sum_{j=1}^N \delta_{X^{j,N}_{N,t}}, \\ X^{i,N}_{N,0} = \xi^i, \end{cases} \tag{3.3}$$

The first objective is to control the L^1-error respectively between $X^{i,N}_N$ and $X^{i,N}$ and between $X^{i,\infty}_N$ and $X^{i,\infty}$ for all $i \in \{1,\dots,N\}$. We write for all $t \in [0,T]$

$$X^{i,N}_{N,t} - X^{i,N}_t = \int_0^t (b(X^{i,N}_{N,s},\overline{\mu}^N_{N,s}) - b(X^{i,N}_s,\overline{\mu}^N_s))\,ds + B\int_0^t\int_{B_N^c} z\widetilde{\mathcal{N}}^i(ds,dz).$$

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 $N \geq 1$ and $t \in [0,T]$, which can change from line to line, such that for all $t \in [0,T]$

$$\sup_{i\leq N}\mathbb{E}\sup_{r\leq t}|X^{i,N}_{N,r} - X^{i,N}_r| \leq C\left[\int_0^t \sup_{i\leq N}\mathbb{E}|X^{i,N}_{N,s} - X^{i,N}_s|\,ds + \int_0^t \frac{1}{N}\sum_{j=1}^N \mathbb{E}|X^{j,N}_{N,s} - X^{j,N}_s|\,ds \right.$$
$$\left. + \sup_{i\leq N}\mathbb{E}\left(\int_0^t\int_{B_N^c}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac{1}{2}}\right].$$

Using the subadditivity of the square root, one has for all $t \in [0,T]$

$$\begin{aligned} &\sup_{i\leq N}\mathbb{E}\sup_{r\leq t}|X^{i,N}_{N,r} - X^{i,N}_r| \\ &\leq C\left[\int_0^t \sup_{i\leq N}\mathbb{E}|X^{i,N}_{N,s} - X^{i,N}_s|\,ds + \int_0^t\int_{B_N^c}|z|\,d\nu(z)\,ds\right] \\ &\leq C\left[\int_0^t \sup_{i\leq N}\mathbb{E}|X^{i,N}_{N,s} - X^{i,N}_s|\,ds + \int_N^\infty r\,\frac{dr}{r^{1+\alpha}}\right]. \end{aligned}$$

Gronwall's inequality ensures that

$$\sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X^{i,N}_{N,t} - X^{i,N}_t| \leq CN^{1-\alpha}. \tag{3.4}$$

We similarly get that for some constant $C > 0$ independent of N

$$\sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X^{i,\infty}_{N,t} - X^{i,\infty}_t| \leq CN^{1-\alpha}. \tag{3.5}$$

The triangle inequality, (3.4) and (3.5) yield

$$\sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X^{i,N}_t - X^{i,\infty}_t| \leq CN^{1-\alpha} + \sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X^{i,N}_{N,t} - X^{i,\infty}_{N,t}|. \tag{3.6}$$

The second term in the right hand-side of (3.6) is controlled as in the proof of Theorem 2. We get that there exists $C > 0$ such that for all $N \geq 1$

$$\sup_{i\leq N}\mathbb{E}\sup_{t\leq T}|X^{i,N}_{N,t} - X^{i,\infty}_{N,t}| \leq C\int_0^T \mathbb{E}W_1(\tilde{\mu}^N_{N,s}, \mu_{N,s})\, ds. \tag{3.7}$$

As the random variables $(X^{i,\infty}_N)_{i\leq N}$ are i.i.d. when N is fixed, we are going to use [7, Theorem 1] to control $\mathbb{E}W_1(\tilde{\mu}^N_{N,s}, \mu_{N,s})$ uniformly with respect to $s \in [0, T]$. We start by controlling the moments of $X^{i,\infty}_N$. Let us fix $\beta \in [1, \alpha]$. Gronwall's inequality ensures that there exists $C > 0$ such that for any $t \in [0, T]$

$$\mathbb{E}|X^{i,\infty}_{N,t}|^\beta \leq C\sup_{s\leq t}\mathbb{E}|Z^i_{N,s}|^\beta.$$

If $\beta < \alpha$ and since Z admits a finite moment of order β, it is clear that

$$\sup_{N\geq 1}\sup_{s\leq T}\mathbb{E}|X^{i,\infty}_{N,s}|^\beta \leq \sup_{N\geq 1}\sup_{s\leq T}\mathbb{E}|Z^i_{N,s}|^\beta < +\infty. \tag{3.8}$$

If $\beta = \alpha$, BDG's and Jensen's inequalities yield

$$\begin{aligned}
\mathbb{E}|Z^i_{N,t}|^\alpha &\leq C\mathbb{E}\left(\int_0^t\int_{B_N}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\\
&\leq C\left[\left(\mathbb{E}\int_0^t\int_{B_1}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}} + \mathbb{E}\left(\int_0^t\int_{B_N\setminus B_1}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\right]\\
&\leq C\left[\left(t\int_{B_1}|z|^2\,d\nu(z)\right)^{\frac{\alpha}{2}} + \mathbb{E}\left(\int_0^t\int_{B_N\setminus B_1}|z|^2\mathcal{N}^i(ds,dz)\right)^{\frac{\alpha}{2}}\right] \qquad (3.9)\\
&\leq C\left[1 + \int_1^N r^\alpha\frac{dr}{r^{1+\alpha}}\right]\\
&\leq C\ln(N).
\end{aligned}$$

If $d = 1$, $d = 2$ or $d \geq 3$ and $\alpha < \frac{d}{d-1}$, it follows from (3.9) and [7] that for all $N \geq 1$

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_{N,t}, \mu_{N,t}) \leq C(\ln(N))^{\frac{1}{\alpha}}\begin{cases}(N^{-\frac{1}{2}} + N^{\frac{1}{\alpha}-1}), & \text{if} \quad d = 1,\\ (N^{-\frac{1}{2}}\ln(N+1) + N^{\frac{1}{\alpha}-1}), & \text{if} \quad d = 2,\\ (N^{-\frac{1}{d}} + N^{\frac{1}{\alpha}-1}), & \text{if} \quad d \geq 3 \text{ and } \alpha \neq \frac{d}{d-1}.\end{cases}$$

Since $\alpha < 2$ and if $d \geq 3$, we have assumed that $\alpha < \frac{d}{d-1}$ and thus $\frac{1}{d} > 1 - \frac{1}{\alpha}$, we deduce that for all $N \geq 1$

$$\sup_{t\in[0,T]}\mathbb{E}W_1(\tilde{\mu}^N_{N,s}, \mu_{N,s}) \leq C(\ln(N))^{\frac{1}{\alpha}}N^{\frac{1}{\alpha}-1}. \tag{3.10}$$

In the case where $d \geq 3$ and $\alpha > \frac{d}{d-1}$. Let us introduce $\beta \in \left(\frac{d}{d-1}, \alpha\right)$. By (3.8) with this choice of β and [7], for all $N \geq 1$

$$\sup_{t\in[0,T]}W_1(\tilde{\mu}^N_{N,t}, \mu_{N,t}) \leq CN^{-\frac{1}{d}}. \tag{3.11}$$

This ends the proof of (1.8) thanks to (3.7) and (3.6) since $\frac{\alpha-1}{\alpha} < \alpha - 1$ because in the case where $d \geq 3$ and $\alpha > \frac{d}{d-1}$, then $\alpha - 1 \geq \frac{1}{d}$.

For the proof of (1.9), keeping the same notations as in the proof of Theorem 2, we have

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

The first term in the right hand-side of the preceding inequality is controlled by (1.8). For the second one, we use the following decomposition

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N, \mu_t) &\leq \sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_t^N, \tilde{\mu}_{N,t}^N) + \sup_{t\in[0,T]} \mathbb{E}W_1(\tilde{\mu}_{N,t}^N, \mu_{N,t}) + \sup_{t\in[0,T]} \mathbb{E}W_1(\mu_{N,t}, \mu_t) \\ &=: I_1 + I_2 + I_3.\end{aligned}$$

Using (3.4) and (3.5), we get that

$$I_1 + I_3 \leq CN^{1-\alpha}. \tag{3.12}$$

For I_2, the inequalities (3.11) and (3.10) prove that for all $N \geq 1$

$$I_2 \leq C \begin{cases} (\ln(N))^{\frac{1}{\alpha}} N^{\frac{1}{\alpha}-1}, & \text{if} \quad d = 1, 2 \quad \text{or} \quad d \geq 3 \text{ and } \alpha < \frac{d}{d-1}, \\ N^{-\frac{1}{d}}, & \text{if} \quad d \geq 3 \text{ and } \alpha > \frac{d}{d-1}. \end{cases} \tag{3.13}$$

Gathering (3.12) and (3.13) ends the proof of (1.9) as previously.

Appendix A. Existence of a solution to (1.1) when $\beta \in (0, 1)$

Let us fix $\beta \in (0, 1)$. We have seen in Remark 2 that in this case, uniqueness for (1.1) fails to be true under Assumption (H1). However, the existence of solutions to (1.1) is given in the following proposition.

Proposition 1. 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 (1.1) for all $\xi \in L^{\beta+\delta}(\Omega, \mathcal{F}_0; \mathbb{R}^d)$. Moreover, we have

$$\mathbb{E} \sup_{t\leq T} |X_t|^{\beta+\delta} < +\infty. \tag{A.1}$$

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 [2, 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 1, the standard SDE

$$\begin{cases} dX_t^\mu = b_t(X_t^\mu,\mu_t)\,dt + \sigma_t(X_{t^-}^\mu,\mu_t)\,dZ_t, \quad t \in [0,T], \\ X_0^\mu = \xi \in L^\beta(\Omega,\mathcal{F}_0) \end{cases} \tag{A.2}$$

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

$$\phi : \begin{cases} \mathcal{P}_\beta(D_T) & \to \mathcal{P}_\beta(D_T) \\ \mu & \mapsto [(X_t^\mu)_{t\in[0,T]}]. \end{cases} \tag{A.3}$$

The goal is to prove that ϕ has at least one fixed point using Schauder's fixed point theorem. By the estimation (2.11) obtained in the proof of Theorem 1, we have

$$\mathcal{W}_\beta(\phi(\mu),\phi(\nu)) \le \mathbb{E}\left(\sup_{0\le s\le T}|X_s^\mu - X_s^\nu|^\beta\right) \le C\left(\int_0^T W_\beta^2(\mu_s,\nu_s)\,ds\right)^{\frac{\beta}{2}}. \tag{A.4}$$

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 [3, 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 [17, 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 (A.4), and using the dominated convergence theorem justified since

$$\sup_{n\ge1}\int_{D_T}\|f\|_\infty^\beta\,d\mu^n(f) < +\infty,$$

we conclude that ϕ is continuous. Following the same lines as to prove (2.11), we show that for some constant $C = C_T > 0$

$$\mathbb{E}\sup_{t\le T}|X_t^\mu|^\beta \le C\left[1 + \left(\int_0^T M_\beta^2(\mu_s)\,ds\right)^{\frac{\beta}{2}}\right]. \tag{A.5}$$

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 (A.5) 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 [2, 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 $t\in[0,T]$

$$\lim_{A\to+\infty}\sup_n \mathbb{P}(|X_t^{\mu^n}|\geq A)=0. \tag{A.6}$$

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 $\epsilon>0$

$$\mathbb{P}(|X^{\mu^n}_{\tau_n+\delta_n}-X^{\mu^n}_{\tau_n}|\geq\epsilon)\underset{n\to+\infty}{\longrightarrow}0.$$

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

$$\mathbb{P}(|X^{\mu^n}_{\tau_n+\delta_n}-X^{\mu^n}_{\tau_n}|\geq\epsilon)\leq\mathbb{P}(T_A^n\leq T)+\frac{1}{\epsilon^\beta}\mathbb{E}(|X^{\mu^n}_{\tau_n+\delta_n}-X^{\mu^n}_{\tau_n}|^\beta\mathbf{1}_{T_A^n\geq T}). \tag{A.7}$$

Notice first that

$$\begin{aligned}\mathbb{P}(T_A^n\leq T)&\leq\mathbb{P}(\sup_{t\leq T}|X_t^{\mu^n}|^\beta\geq A^\beta)\\ &\leq\frac{R}{A^\beta}.\end{aligned} \tag{A.8}$$

Then, by the triangle inequality, we have

$$\begin{aligned}\mathbb{E}(|X^{\mu^n}_{\tau_n+\delta_n}-X^{\mu^n}_{\tau_n}|^\beta\mathbf{1}_{T_A^n\geq T}) &\leq \mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n} b_s(X_s^{\mu^n},\mu_s^n)\,ds\right|^\beta\mathbf{1}_{T_A^n\geq T}\right)\\ &\quad+\mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n}\int_{B_1}\sigma_s(X_{s^-}^{\mu^n},\mu_s^n)z\,\widetilde{\mathcal{N}}(ds,dz)\right|^\beta\mathbf{1}_{T_A^n\geq T}\right)\\ &\quad+\mathbb{E}\left(\left|\int_{\tau_n}^{\tau_n+\delta_n}\int_{B_1^c}\sigma_s(X_{s^-}^{\mu^n},\mu_s^n)z\,\mathcal{N}(ds,dz)\right|^\beta\mathbf{1}_{T_A^n\geq T}\right)\\ &=: I_1+I_2+I_3\end{aligned}$$

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 (A.7), (A.8), 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 [6, 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 (1.1). The moment estimation (A.1) directly follows from [6, Proposition 2]. □

References

[1] David Applebaum. *Levy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2009.

[2] Richard F. Bass. *Stochastic processes*. Cambridge University Press, 2011.

[3] Patrick Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Statistics. Probability and Statistics Section. Wiley, 2nd ed edition, 1999.

[4] 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.
[5] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: a review of models, methods and applications. II. Applications. arXiv:2106.14812, May 2022.
[6] 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.
[7] 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.
[8] 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.
[9] Noufel Frikha and Libo Li. Well-posedness and approximation of some one-dimensional Lévy-driven nonlinear SDEs. *Stochastic Processes and their Applications*, 132:76–107, February 2021.
[10] Carl Graham. McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications*, 40(1):69–82, 1992.
[11] Carl Graham. Nonlinear diffusion with jumps. *Annales de l'I.H.P.*, 28(3):393–402, 1992.
[12] 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.
[13] 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.
[14] 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.
[15] Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. Number 68 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.
[16] Alain-Sol Sznitman. *Topics in Propagation of Chaos*, volume 1464 of *Lecture Notes in Mathematics*, pages 165–251. Springer Berlin Heidelberg, 1991.
[17] Cédric Villani. *Optimal transportation : old and new*. Springer, 2009.

Université de Rennes 1, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France
Email address: thomas.cavallazzi@univ-rennes1.fr