Well-Posedness for McKean-Vlasov SDEs Driven by Multiplicative Stable Noises

Chang-Song Deng [a)], **Xing Huang** [b)],

a)School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

dengcs@whu.edu.cn

b)Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

xinghuang@tju.edu.cn

December 11, 2025

Abstract

We establish the well-posedness for a class of McKean-Vlasov SDEs driven by symmetric α-stable Lévy process ($1/2 < \alpha \leq 1$), where the drift coefficient is Hölder continuous in space variable, while the noise coefficient is Lipscitz continuous in space variable, and both of them satisfy the Lipschitz condition in distribution variable with respect to Wasserstein distance. If the drift coefficient does not depend on distribution variable, our methodology developed in this paper applies to the case $\alpha \in (0, 1]$. The main tool relies on heat kernel estimates for (distribution independent) stable SDEs and Banach's fixed point theorem.

AMS subject Classification: 60G52, 60H10.
Keywords: McKean-Vlasov SDEs, α-stable process, distribution dependent noise, Wasserstein distance.

1 Introduction

It is well-known that many complex physical, biological, and other scientific phenomena can be modeled by interacting particle systems, which attract much attention in recent years due to their importance both in theory and in applications. When the number of particles goes to infinity, the equation for one single particle in the mean field interacting particle system tends to the so-called McKean-Vlasov SDE, which was first introduced by McKean in [9]. This is related to the propagation of chaos, see for instance [15]. As a fundamental issue in the study of McKean-Vlasov SDEs, the well-posedness has been intensively investigated for Gaussian noise case, see [1, 6, 11, 12, 16, 18] and references therein for more details.

As far as we know, however, the results concerning the well-posedness for McKean-Vlasov SDEs with jump noises are still quite limited. The well-posedness is established in [7] for Lévy-driven McKean-Vlasov SDEs without drift. In [17], the authors consider strong well-posedness for density dependent SDEs with additive α-stable noise ($1 < \alpha < 2$), where the drift is assumed to be C_b^β with $\beta \in (1-\alpha/2, 1)$ in space variable, and Lipschitz continuous in distribution variable with respect to the L^θ-Wasserstein distance ($1 < \theta < \alpha$). In [10], the authors prove the well-posedness for stable McKean-Vlasov SDEs under the assumption that the coefficients have bounded and Hölder continuous flat derivatives (also called linear functional derivatives); in the supercritical case, i.e. the stability index $\alpha < 1$, it is necessary to require $\alpha > 2/3$ (see [10, Theorem 2.2]). In the very recent work [5], we establish the well-posedness for McKean-Vlasov SDEs driven by α-stable noise ($1 < \alpha < 2$), where the noise coefficient depends only on time and distribution variables.

As a continuation of [5], in this paper, we consider the following stable McKean-Vlasov SDE with stable index $\alpha \in (0, 1]$:

$$\mathrm{d}X_t = b_t(X_t, \mathscr{L}_{X_t})\,\mathrm{d}t + \sigma_t(X_{t-}, \mathscr{L}_{X_t})\,\mathrm{d}L_t, \quad t \in [0, T], \tag{1.1}$$

where $T > 0$ is a fixed constant, $(L_t)_{t\geq 0}$ is a d-dimensional rotationally invariant α-stable Lévy process with infinitesimal generator $-\frac{1}{2}(-\triangle)^{\alpha/2}$, $\mathscr{L}_{X_t}$ is the law of X_t, and for the space $\mathscr{P}$ of all probability measures on $\mathbb{R}^d$ equipped with the weak topology,

$$b : [0, T] \times \mathbb{R}^d \times \mathscr{P} \to \mathbb{R}^d, \quad \sigma : [0, T] \times \mathbb{R}^d \times \mathscr{P} \to \mathbb{R}^d \otimes \mathbb{R}^d$$

are measurable.

For $\kappa \in (0, 1]$, let

$$\mathscr{P}_\kappa := \big\{\gamma \in \mathscr{P} : \ \gamma(|\cdot|^\kappa) < \infty\big\},$$

which is a Polish space under the L^κ-Wasserstein distance

$$\mathbb{W}_\kappa(\gamma, \tilde\gamma) := \inf_{\pi \in \mathscr{C}(\gamma, \tilde\gamma)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^\kappa\, \pi(\mathrm{d}x, \mathrm{d}y), \quad \gamma, \tilde\gamma \in \mathscr{P}_\kappa,$$

where $\mathscr{C}(\gamma, \tilde\gamma)$ is the set of all couplings of γ and $\tilde\gamma$. By [3, Theorem 5.10], the following dual formula

$$\mathbb{W}_\kappa(\gamma, \tilde\gamma) = \sup_{[f]_\kappa \leq 1} |\gamma(f) - \tilde\gamma(f)|, \quad \gamma, \tilde\gamma \in \mathscr{P}_\kappa$$

holds, where $[f]_\kappa$ denotes the Hölder seminorm (of exponent κ) of $f : \mathbb{R}^d \to \mathbb{R}$ defined by $[f]_\kappa := \sup_{x \neq y} \frac{|f(x)-f(y)|}{|x-y|^\kappa}$.

To derive the well-posedness for (1.1), we make the following assumptions.

(A) There exist $\beta \in (0, 1)$ satisfying $2\beta + \alpha > 2$, $K > 1$ and $\eta \in (0, \alpha)$ with $\alpha + \eta > 1$ such that for all $t \in [0, T]$, $x, y \in \mathbb{R}^d$ and $\gamma, \tilde\gamma \in \mathscr{P}_\eta$,

$$|b_t(x, \gamma) - b_t(y, \tilde\gamma)| \leq K\left(\mathbb{W}_\eta(\gamma, \tilde\gamma) + \big\{|x-y|^\beta \vee |x-y|\big\}\right), \tag{1.2}$$

and

$$\tag{1.3} \begin{cases} |b_t(0,\delta_0)| \le K, \\ \|\sigma_t(x,\gamma) - \sigma_t(y,\tilde{\gamma})\| \le K\big(|x-y| + \mathbb{W}_\eta(\gamma,\tilde{\gamma})\big), \\ K^{-1}I \le (\sigma_t\sigma_t^*)(x,\gamma) \le KI, \end{cases}$$

where I is the $d \times d$ identity matrix.

(A') $\alpha \in (0,1]$ and $b_t(x,\gamma) = b_t(x)$ does not depend on γ. There exist $\beta \in (0,1)$ satisfying $2\beta + \alpha > 2$, $K > 1$ and $\eta \in (0,\alpha)$ such that for all $t \in [0,T]$, $x, y \in \mathbb{R}^d$ and $\gamma, \tilde{\gamma} \in \mathscr{P}_\eta$,

$$|b_t(x) - b_t(y)| \le K\big\{|x-y|^\beta \vee |x-y|\big\}$$

and (1.3) hold.

Denote by $C([0,T];\mathscr{P}_k)$ the set of all continuous maps from $[0,T]$ to $\mathscr{P}_k$ under the metric $\mathbb{W}_k$. Throughout the paper the constant C denotes a positive constant which may depend on $T, d, \alpha, \beta, \eta, K$; its value may change, without further notice, from line to line.

Our main result is the following theorem:

Theorem 1.1. *Assume (A) or (A'). Then (1.1) is strongly/weakly well-posed in $\mathscr{P}_\eta$, and the solution satisfies $\mathscr{L}_{X_\cdot} \in C([0,T];\mathscr{P}_\eta)$ and*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^\eta\right] < C\left(1 + \mathbb{E}\big[|X_0|^\eta\big]\right).$$

Remark 1.2. *If $\alpha < 1$, to ensure the well-posedness, it is required in [10, Theorem 2.2] that $\alpha > 2/3$. Noting that $\eta \in (0,\alpha)$ and $\alpha + \eta > 1$ in (A) imply that $\alpha \in (\frac{1}{2}, 1]$, we can handle the case $\alpha \in (1/2, 1)$. Moreover, when $b_t(x,\gamma)$ does not depend on γ, α can take all values in $(0,1]$ if β is close enough to 1.*

Remark 1.3. *With respect to assumption $\alpha + \eta > 1$ in (A), in the proof of Lemma 2.3, it is essential to derive a nice estimate for J_1. Moreover, by Example 4.2 in the Appendix, the condition $\alpha + \eta > 1$ in (A) is necessary in the sense that if $\alpha + \eta = 1$, then we cannot expect the uniqueness for the solution to (1.1).*

The remainder of the paper is organized as follows: In Section 2, we make some preparations and the proof of Theorem 1.1 is presented in Section 3. A counterexample is provided in the Appendix for non-uniqueness of solutions to stable McKean-Vlasov SDEs

2 Some preparations

Let $\gamma \in \mathscr{P}_\eta$ and $\mu \in C([0,T];\mathscr{P}_\eta)$, where $\eta \in (0,\alpha)$. Consider the following (distribution independent) SDE with initial distribution $\mathscr{L}_{X_{s,s}^{\gamma,\mu}} = \gamma$:

$$\tag{2.1} \mathrm{d}X_{s,t}^{\gamma,\mu} = b_t(X_{s,t}^{\gamma,\mu}, \mu_t)\,\mathrm{d}t + \sigma_t(X_{s,t-}^{\gamma,\mu}, \mu_t)\,\mathrm{d}L_t, \quad 0 \le s \le t \le T.$$

By [4, Theorem 1.1] and a standard localization argument, (2.1) has a unique strong solution under (A) or (A'). For simplicity, we denote $X_t^{\gamma,\mu} = X_{0,t}^{\gamma,\mu}$. Moreover, if $\gamma = \delta_x$ is the Dirac measure concentrated at $x \in \mathbb{R}^d$, we write $X_{s,t}^{x,\mu} = X_{s,t}^{\delta_x,\mu}$.

By [8], $\mathscr{L}_{X_{s,t}^{x,\mu}}$ is absolutely continuous with respect to the Lebesgue measure, and we denote by $p_{s,t}^{\mu}(x,\cdot)$ the corresponding density function. Denote by $P_{s,t}^{\mu}$ the inhomogeneous Markov semigroup associated with $X_{s,t}^{x,\mu}$, i.e. for $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$P_{s,t}^{\mu} f(x) = \mathbb{E} f(X_{s,t}^{x,\mu}) = \int_{\mathbb{R}^d} p_{s,t}^{\mu}(x,y) f(y)\,\mathrm{d}y.$$

Here and in the sequel, $\mathscr{B}_b(\mathbb{R}^d)$ denotes the set of all bounded measurable functions on $\mathbb{R}^d$. As before, write $p_t^{\mu}(x,\cdot) = p_{0,t}^{\mu}(x,\cdot)$ and $P_t^{\mu} = P_{0,t}^{\mu}$ for $t \in [0,T]$ and $\mu \in C([0,T];\mathscr{P}_\eta)$. We will use the following notation for $\mu \in C([0,T];\mathscr{P}_\eta)$

$$\mathscr{A}_t^{\mu} f(\cdot) := \int_{\mathbb{R}^d\setminus\{0\}} \left[f(\cdot + \sigma_t(\cdot,\mu_t)y) - f(\cdot) - \langle \sigma_t(\cdot,\mu_t)y, \nabla f(\cdot)\rangle \mathbb{1}_{\{|y|\le 1\}} \right] \Pi(\mathrm{d}y), \tag{2.2}$$

where

$$\Pi(\mathrm{d}y) := \frac{\alpha\Gamma(\frac{d+\alpha}{2})}{2^{2-\alpha}\pi^{d/2}\Gamma(1-\frac{\alpha}{2})}\frac{\mathrm{d}y}{|y|^{d+\alpha}}$$

is the Lévy measure of L_t.

Lemma 2.1. *Assume (A) or (A'). For any $0 \le r < t \le T$, $\mu^1, \mu^2 \in C([0,T];\mathscr{P}_\eta)$, and $f \in \mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta \le 1$,*

$$|\nabla P_{r,t}^{\mu^2} f| \le C(t-r)^{-\frac{1}{\alpha}+\frac{\eta}{\alpha}}, \tag{2.3}$$

$$|(\mathscr{A}_r^{\mu^1} - \mathscr{A}_r^{\mu^2}) P_{r,t}^{\mu^2} f| \le C(t-r)^{-1+\frac{\eta}{\alpha}} \mathbb{W}_\eta(\mu_r^1, \mu_r^2). \tag{2.4}$$

Proof. (i) Denote by p_t^{α} the density function of L_t (with respect to the Lebesgue measure). Note that (1.2) implies that

$$|b_t(x,\gamma) - b_t(y,\gamma)| \le K(|x-y|^\beta \vee |x-y|), \quad t \in [0,T], x,y \in \mathbb{R}^d, \gamma \in \mathscr{P}_\eta.$$

Hence, it follows from [8, Theorem 1.1 (iii) and (i)] that for any $0 \le r < t \le T$,

$$|\nabla p_{r,t}^{\mu^2}(\cdot,y)(x)| \le C(t-r)^{-1/\alpha} p_{t-r}^{\alpha}\left(\theta_{r,t}(x) - y\right),$$

where $\{\theta_{r,t}\}_{T\ge t\ge r\ge 0}$ is a flow with

$$\theta_{s,t}(x) = x + \int_s^t b_r\left(\theta_{r,t}(x), \mu_r^2\right)\mathrm{d}r, \quad 0 \le s \le t \le T.$$

One can refer to [2, 14] for more applications of $\theta_{r,t}$ on investigating the properties of the solution associated to drifted fractional diffusion. Noting that

$$\nabla_x \underbrace{\int_{\mathbb{R}^d} p_{r,t}^{\mu^2}(x,y)\,\mathrm{d}y}_{=1} = 0,$$

we obtain for all $f \in \mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta \le 1$,

$$\begin{aligned}
|\nabla P_{r,t}^{\mu^2} f(x)| &= \left| \int_{\mathbb{R}^d} \nabla p_{r,t}^{\mu^2}(\cdot, y)(x) f(y)\,\mathrm{d}y \right| \\
&= \left| \int_{\mathbb{R}^d} \nabla p_{r,t}^{\mu^2}(\cdot, y)(x) [f(y) - f(\theta_{r,t}(x))]\,\mathrm{d}y \right| \\
&\le \int_{\mathbb{R}^d} |\nabla p_{r,t}^{\mu^2}(\cdot, y)(x)| \times |f(y) - f(\theta_{r,t}(x))|\,\mathrm{d}y \\
&\le C(t-r)^{-1/\alpha} \int_{\mathbb{R}^d} p_{t-r}^{\alpha}\left(\theta_{r,t}(x) - y\right) |\theta_{r,t}(x) - y|^\eta\,\mathrm{d}y \\
&= C(t-r)^{-1/\alpha} \int_{\mathbb{R}^d} p_{t-r}^{\alpha}(z) |z|^\eta\,\mathrm{d}y \\
&= C(t-r)^{-1/\alpha} \mathbb{E}\big[|L_{t-r}|^\eta\big] \\
&= C(t-r)^{(\eta-1)/\alpha} \mathbb{E}\big[|L_1|^\eta\big].
\end{aligned}$$

This implies (2.3) since $\mathbb{E}\big[|L_1|^\eta\big] < \infty$.

(ii) By (1.3), it is not hard to get

$$\begin{aligned}
&\left| \frac{|\det(\sigma_t^{-1}(\cdot, \mu_t^1))|}{|\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha}} - \frac{|\det(\sigma_t^{-1}(\cdot, \mu_t^2))|}{|\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha}} \right| \\
&= \left| \frac{|\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha} |\det(\sigma_t^{-1}(\cdot, \mu_t^1))| - |\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha} |\det(\sigma_t^{-1}(\cdot, \mu_t^2))|}{|\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha} |\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha}} \right| \\
&\le \left| \frac{|\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha} [|\det(\sigma_t^{-1}(\cdot, \mu_t^1))| - |\det(\sigma_t^{-1}(\cdot, \mu_t^2))|]}{|\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha} |\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha}} \right| \\
&\quad + \left| \frac{|\det(\sigma_t^{-1}(\cdot, \mu_t^2))| [|\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha} - |\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha}]}{|\sigma_t^{-1}(\cdot, \mu_t^1) y|^{d+\alpha} |\sigma_t^{-1}(\cdot, \mu_t^2) y|^{d+\alpha}} \right| \\
&\le C \frac{\mathbb{W}_\eta(\mu_t^1, \mu_t^2)}{|y|^{d+\alpha}}.
\end{aligned}$$

Since we can rewrite (2.2) as a principal value (p.v.) integral:

$$\begin{aligned}
\mathscr{A}_t^{\nu} f(\cdot) &= \frac{1}{2}\,\text{p.v.} \int_{\mathbb{R}^d} \big[f(\cdot + \sigma_t(\cdot, \nu_t) y) + f(\cdot - \sigma_t(\cdot, \nu_t) y) - 2f(\cdot)\big]\,\Pi(\mathrm{d}y) \\
&= \frac{\alpha \Gamma(\frac{d+\alpha}{2})}{2^{3-\alpha} \pi^{d/2} \Gamma(1 - \frac{\alpha}{2})}\,\text{p.v.} \int_{\mathbb{R}^d} \big[f(\cdot + y) + f(\cdot - y) - 2f(\cdot)\big] \frac{|\det(\sigma_t^{-1}(\cdot, \nu_t))|}{|\sigma_t^{-1}(\cdot, \nu_t) y|^{d+\alpha}}\,\mathrm{d}y,
\end{aligned}$$

it holds that

$$
\begin{aligned}
&|(\mathscr{A}_r^{\mu^1}-\mathscr{A}_r^{\mu^2})P_{r,t}^{\mu^2}f|\\
&=C\left|\mathrm{p.v.}\int_{\mathbb{R}^d}\left[P_{r,t}^{\mu^2}f(\cdot+y)+P_{r,t}^{\mu^2}f(\cdot-y)-2P_{r,t}^{\mu^2}f(\cdot)\right]\left[\frac{|\det(\sigma_t^{-1}(\cdot,\mu_t^1))|}{|\sigma_t^{-1}(\cdot,\mu_t^1)y|^{d+\alpha}}-\frac{|\det(\sigma_t^{-1}(\cdot,\mu_t^2))|}{|\sigma_t^{-1}(\cdot,\mu_t^2)y|^{d+\alpha}}\right]\mathrm{d}y\right|\\
&\le C\,\mathrm{p.v.}\int_{\mathbb{R}^d}\left|P_{r,t}^{\mu^2}f(\cdot+y)+P_{r,t}^{\mu^2}f(\cdot-y)-2P_{r,t}^{\mu^2}f(\cdot)\right|\left|\frac{|\det(\sigma_t^{-1}(\cdot,\mu_t^1))|}{|\sigma_t^{-1}(\cdot,\mu_t^1)y|^{d+\alpha}}-\frac{|\det(\sigma_t^{-1}(\cdot,\mu_t^2))|}{|\sigma_t^{-1}(\cdot,\mu_t^2)y|^{d+\alpha}}\right|\mathrm{d}y\\
&\le C\mathbb{W}_\eta(\mu_t^1,\mu_t^2)\,\mathrm{p.v.}\int_{\mathbb{R}^d}\left|P_{r,t}^{\mu^2}f(\cdot+y)+P_{r,t}^{\mu^2}f(\cdot-y)-2P_{r,t}^{\mu^2}f(\cdot)\right|\frac{\mathrm{d}y}{|y|^{d+\alpha}}\\
&=C\mathbb{W}_\eta(\mu_t^1,\mu_t^2)|\mathscr{D}^\alpha P_{r,t}^{\mu^2}f|,
\end{aligned}
$$

where $\mathscr{D}^\alpha$ is a fractional derivative operator of order α (cf. [8, (1.23)]) with

$$
|\mathscr{D}^\alpha f|(x):=\int_{\mathbb{R}^d}|f(x+y)+f(x-y)-2f(x)|\,\frac{\mathrm{d}y}{|y|^{d+\alpha}}.
$$

By [8, Theorem 1.1 (ii)], for any $0\le r<t\le T$,

$$
|\mathscr{D}^\alpha p_{r,t}^{\mu^2}(\cdot,z)|(x)\le C(t-r)^{-1}p_{t-r}^\alpha\left(\theta_{r,t}(x)-z\right).
$$

Then we get for all $f\in\mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta\le 1$,

$$
\begin{aligned}
|\mathscr{D}^\alpha P_{r,t}^{\mu^2}f|(x)&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\left[p_{r,t}^{\mu^2}(x+y,z)+p_{r,t}^{\mu^2}(x-y,z)-2p_{r,t}^{\mu^2}(x,z)\right]f(z)\,\mathrm{d}z\right|\frac{\mathrm{d}y}{|y|^{d+\alpha}}\\
&=\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}\left[p_{r,t}^{\mu^2}(x+y,z)+p_{r,t}^{\mu^2}(x-y,z)-2p_{r,t}^{\mu^2}(x,z)\right][f(z)-f(\theta_{r,t}(x))]\,\mathrm{d}z\right|\frac{\mathrm{d}y}{|y|^{d+\alpha}}\\
&\le\int_{\mathbb{R}^d}|\mathscr{D}^\alpha p_{r,t}^{\mu^2}(\cdot,z)|(x)|f(\theta_{r,t}(x))-f(z)|\,\mathrm{d}z\\
&\le C(t-r)^{-1}\int_{\mathbb{R}^d}p_{t-r}^\alpha(\theta_{r,t}(x)-z)|\theta_{r,t}(x)-z|^\eta\,\mathrm{d}z\\
&=C(t-r)^{-1}\mathbb{E}\big[|L_{t-r}|^\eta\big]\\
&=C(t-r)^{-1+\frac{\eta}{\alpha}}\mathbb{E}\big[|L_1|^\eta\big].
\end{aligned}
$$

This together with the above estimate implies (2.4). □

Lemma 2.2. *Let $0\le s<t\le T$, $\mu^1,\mu^2\in C([0,T];\mathscr{P}_\eta)$ for $\eta\in(0,\alpha)$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$. Then*

$$
P_{s,t}^{\mu^1}f=P_{s,t}^{\mu^2}f+\int_s^t P_{s,r}^{\mu^1}\big\langle b_r(\cdot,\mu_r^1)-b_r(\cdot,\mu_r^2),\nabla P_{r,t}^{\mu^2}f\big\rangle\,\mathrm{d}r+\int_s^t P_{s,r}^{\mu^1}(\mathscr{A}_r^{\mu_r^1}-\mathscr{A}_r^{\mu_r^2})P_{r,t}^{\mu^2}f\,\mathrm{d}r.
$$

If furthermore b does not depend on distribution variable, then

$$
P_{s,t}^{\mu^1}f=P_{s,t}^{\mu^2}f+\int_s^t P_{s,r}^{\mu^1}(\mathscr{A}_r^{\mu_r^1}-\mathscr{A}_r^{\mu_r^2})P_{r,t}^{\mu^2}f\,\mathrm{d}r.
$$

Proof. By a standard approximation argument, it suffices to prove the desired assertion for $f \in C_b^2(\mathbb{R}^d)$. By the backward Kolmogorov equation, see [8, Theorem 1.1], it holds that

$$\frac{\partial P_{r,t}^{\mu^2} f}{\partial r} = -\langle b_r(\cdot, \mu_r^2), \nabla P_{r,t}^{\mu^2} f\rangle - \mathscr{A}_r^{\mu^2}(P_{r,t}^{\mu^2} f), \quad 0 \le r < t \le T,$$

where $\mathscr{A}_r^{\nu} f$ is given by (2.2). By Itô's formula, we have the forward Kolmogorov equation

$$\frac{\partial P_{s,r}^{\mu^1} f}{\partial r} = P_{s,r}^{\mu^1}[\langle b_r(\cdot, \mu_r^1), \nabla f\rangle + \mathscr{A}_r^{\mu^1} f], \quad 0 \le s < r \le T.$$

Hence, we have

$$\begin{aligned} P_{s,t}^{\mu^1} f - P_{s,t}^{\mu^2} f &= \int_s^t \frac{\partial}{\partial r}[P_{s,r}^{\mu^1} P_{r,t}^{\mu^2} f]\,\mathrm{d}r \\ &= \int_s^t P_{s,r}^{\mu^1}\big\langle b_r(\cdot, \mu_r^1) - b_r(\cdot, \mu_r^2), \nabla P_{r,t}^{\mu^2} f\big\rangle \mathrm{d}r + \int_s^t P_{s,r}^{\mu^1}(\mathscr{A}_r^{\mu_r^1} - \mathscr{A}_r^{\mu_r^2}) P_{r,t}^{\mu^2} f\,\mathrm{d}r, \end{aligned}$$

which implies the first assertion. Clearly, the second assertion follows immediately from the first one. □

Lemma 2.3. *If (A) holds, then for all $\gamma \in \mathscr{P}_\eta$, $\mu^i \in C([0,T]; \mathscr{P}_\eta)$, $i = 1, 2$, and $\delta > 0$,*

$$\sup_{t\in[0,T]} \mathrm{e}^{-\delta t}\mathbb{W}_\eta\big(\mathscr{L}_{X_t^{\gamma,\mu^1}}, \mathscr{L}_{X_t^{\gamma,\mu^2}}\big) \le C\left(\delta^{\frac{1}{\alpha}-\frac{\eta}{\alpha}-1} + \delta^{-\frac{\eta}{\alpha}}\right) \sup_{t\in[0,T]} \mathrm{e}^{-\delta t}\mathbb{W}_\eta(\mu_t^1, \mu_t^2).$$

If (A') holds, then for all $\gamma \in \mathscr{P}_\eta$, $\mu^i \in C([0,T]; \mathscr{P}_\eta)$, $i = 1, 2$, and $\delta > 0$,

$$\sup_{t\in[0,T]} \mathrm{e}^{-\delta t}\mathbb{W}_\eta\big(\mathscr{L}_{X_t^{\gamma,\mu^1}}, \mathscr{L}_{X_t^{\gamma,\mu^2}}\big) \le C\delta^{-\frac{\eta}{\alpha}} \sup_{t\in[0,T]} \mathrm{e}^{-\delta t}\mathbb{W}_\eta(\mu_t^1, \mu_t^2).$$

Proof. Assume (A). It follows form the definition of $\mathbb{W}_\eta$ and Lemma 2.2 that

$$\begin{aligned} \mathbb{W}_\eta\big(\mathscr{L}_{X_t^{\gamma,\mu^1}}, \mathscr{L}_{X_t^{\gamma,\mu^2}}\big) &= \sup_{f\in\mathscr{B}_b(\mathbb{R}^d),[f]_\eta\le 1} \left|\int_{\mathbb{R}^d} \big[P_t^{\mu^1} f(x) - P_t^{\mu^2} f(x)\big]\,\gamma(\mathrm{d}x)\right| \\ &\le \sup_{f\in\mathscr{B}_b(\mathbb{R}^d),[f]_\eta\le 1} \left|\int_{\mathbb{R}^d} \gamma(\mathrm{d}x) \int_0^t P_{0,r}^{\mu^1}\big\langle b_r(\cdot, \mu_r^1) - b_r(\cdot, \mu_r^2), \nabla P_{r,t}^{\mu^2} f\big\rangle(x)\,\mathrm{d}r\right| \\ &\quad + \sup_{f\in\mathscr{B}_b(\mathbb{R}^d),[f]_\eta\le 1} \left|\int_{\mathbb{R}^d} \gamma(\mathrm{d}x) \int_0^t P_{0,r}^{\mu^1}\{(\mathscr{A}_r^{\mu_r^1} - \mathscr{A}_r^{\mu_r^2}) P_{r,t}^{\mu^2} f\}(x)\,\mathrm{d}r\right| \\ &=: \sum_{i=1}^2 \sup_{f\in\mathscr{B}_b(\mathbb{R}^d),[f]_\eta\le 1} \mathsf{J}_i. \end{aligned}$$

By (1.2) and (2.3), we derive that for all $f \in \mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta \le 1$,

$$\big\|\big\langle b_r(\cdot, \mu_r^1) - b_r(\cdot, \mu_r^2), \nabla P_{r,t}^{\mu^2} f(\cdot)\big\rangle\big\|_\infty \le \big\|b_r(\cdot, \mu_r^1) - b_r(\cdot, \mu_r^2)\big\|_\infty \big\|\nabla P_{r,t}^{\mu^2} f(\cdot)\big\|_\infty$$

$$\leq C(t-r)^{-\frac{1}{\alpha}+\frac{\eta}{\alpha}}\mathbb{W}_\eta(\mu_r^1,\mu_r^2).$$

Observing $\alpha+\eta>1$, we get for all $t\in[0,T]$, $\delta>0$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta\leq 1$,

$$\begin{aligned}
\mathsf{J}_1 &\leq C\int_0^t (t-r)^{-\frac{1}{\alpha}+\frac{\eta}{\alpha}}\mathbb{W}_\eta(\mu_r^1,\mu_r^2)\,\mathrm{d}r\\
&= C\mathrm{e}^{\delta t}\int_0^t \mathrm{e}^{-\delta r}\mathbb{W}_\eta(\mu_r^1,\mu_r^2)\cdot(t-r)^{-\frac{1}{\alpha}+\frac{\eta}{\alpha}}\mathrm{e}^{-\delta(t-r)}\,\mathrm{d}r\\
&\leq C\mathrm{e}^{\delta t}\sup_{s\in[0,T]}\mathrm{e}^{-\delta s}\mathbb{W}_\eta(\mu_s^1,\mu_s^2)\times\int_0^t (t-r)^{-\frac{1}{\alpha}+\frac{\eta}{\alpha}}\mathrm{e}^{-\delta(t-r)}\,\mathrm{d}r\\
&\leq C\delta^{\frac{1}{\alpha}-\frac{\eta}{\alpha}-1}\mathrm{e}^{\delta t}\sup_{s\in[0,T]}\mathrm{e}^{-\delta s}\mathbb{W}_\eta(\mu_s^1,\mu_s^2),
\end{aligned}$$

where in the last inequality we have used the fact that for any $\epsilon\in(0,1)$,

$$\sup_{t\in[0,T]}\int_0^t (t-r)^{-\epsilon}\mathrm{e}^{-\delta(t-r)}\,\mathrm{d}r\leq\int_0^\infty r^{-\epsilon}\mathrm{e}^{-\delta r}\,\mathrm{d}r=\Gamma(1-\epsilon)\,\delta^{\epsilon-1}. \tag{2.5}$$

By (2.4) and (2.5), for all $t\in[0,T]$, $\delta>0$ and $f\in\mathscr{B}_b(\mathbb{R}^d)$ with $[f]_\eta\leq 1$,

$$\begin{aligned}
\mathsf{J}_2 &\leq C\int_0^t (t-r)^{-1+\frac{\eta}{\alpha}}\mathbb{W}_\eta(\mu_r^1,\mu_r^2)\,\mathrm{d}r\\
&= C\mathrm{e}^{\delta t}\int_0^t \mathrm{e}^{-\delta r}\mathbb{W}_\eta(\mu_r^1,\mu_r^2)\cdot(t-r)^{-1+\frac{\eta}{\alpha}}\mathrm{e}^{-\delta(t-r)}\,\mathrm{d}r\\
&\leq C\mathrm{e}^{\delta t}\sup_{s\in[0,T]}\mathrm{e}^{-\delta s}\mathbb{W}_\eta(\mu_s^1,\mu_s^2)\times\int_0^t (t-r)^{-1+\frac{\eta}{\alpha}}\mathrm{e}^{-\delta(t-r)}\,\mathrm{d}r\\
&\leq C\delta^{-\frac{\eta}{\alpha}}\mathrm{e}^{\delta t}\sup_{s\in[0,T]}\mathrm{e}^{-\delta s}\mathbb{W}_\eta(\mu_s^1,\mu_s^2).
\end{aligned}$$

Combining the bounds for J_i, $i=1,2$, we obtain that for all $\delta>0$

$$\begin{aligned}
\sup_{t\in[0,T]}\mathrm{e}^{-\delta t}\mathbb{W}_\eta\big(\mathscr{L}_{X_t^{\gamma,\mu^1}},\mathscr{L}_{X_t^{\gamma,\mu^2}}\big) &\leq\sum_{i=1}^2\sup_{t\in[0,T]}\sup_{f\in\mathscr{B}_b(\mathbb{R}^d),[f]_\eta\leq 1}\mathrm{e}^{-\delta t}\mathsf{J}_i\\
&\leq C\left(\delta^{\frac{1}{\alpha}-\frac{\eta}{\alpha}-1}+\delta^{-\frac{\eta}{\alpha}}\right)\sup_{s\in[0,T]}\mathrm{e}^{-\delta s}\mathbb{W}_\eta(\mu_s^1,\mu_s^2).
\end{aligned}$$

This yields the first assertion. One can prove the second assertion by repeating the argument above (with $\mathsf{J}_1=0$). □

3 Proof of Theorem 1.1

Proof of Theorem 1.1. It follows from Lemma 2.3 that for $\delta>0$ large enough, the map

$$\mu\mapsto\mathscr{L}_{X_\cdot^{\gamma,\mu}}$$

is strictly contractive in $C([0,T];\mathscr{P}_\eta)$ under the complete metric

$$\sup_{t\in[0,T]} \mathrm{e}^{-\delta t}\mathbb{W}_\eta(\mu_t^1,\mu_t^2)$$

for $\mu^1,\mu^2\in C([0,T];\mathscr{P}_\eta)$. Then it has a unique fixed point $\mu^*=\mu^*(\gamma)\in C([0,T];\mathscr{P}_\eta)$ such that $\mu^*=\mathscr{L}_{X^{\gamma,\mu^*}_\cdot}$, and $X_t=X_t^{\gamma,\mu^*}$ is the unique solution to (1.1) with $\mathscr{L}_{X_0}=\gamma\in\mathscr{P}_\eta$.

To prove the moment estimate, we will use a (random) time-change argument. Let S_t be an $\frac{\alpha}{2}$-stable subordinator with the following Laplace transform:

$$\mathbb{E}\left[\mathrm{e}^{-rS_t}\right]=\mathrm{e}^{-2^{-1}t(2r)^{\alpha/2}},\quad r>0,\, t\geq 0,$$

and let W_t be a d-dimensional standard Brownian motion, which is independent of S_t. The time-changed process $L_t:=W_{S_t}$ is a d-dimensional rotationally symmetric α-stable Lévy process such that $\mathbb{E}\,\mathrm{e}^{\mathrm{i}\langle\xi,L_t\rangle}=\mathrm{e}^{-t|\xi|^\alpha/2}$ for $\xi\in\mathbb{R}^d$, see e.g. [13]. Using the subordination representation, (1.1) can be written in the following form

$$X_t=X_0+\int_0^t b_r(X_r,\mathscr{L}_{X_r})\,\mathrm{d}r+\int_0^t\sigma_r(X_{r-},\mathscr{L}_{X_r})\,\mathrm{d}W_{S_r},$$

where $\mathscr{L}_{X_0}\in\mathscr{P}_\eta$. It is easy to see that (A) or (A') implies for $x\in\mathbb{R}^d$ and $\gamma\in\mathscr{P}_\eta$,

$$\sup_{t\in[0,T]}|b_t(x,\gamma)|\leq C\big(1+|x|+\gamma(|\cdot|^\eta)\big).$$

Since σ is bounded due to (1.3), we obtain for all $s\in[0,T]$,

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,s]}|X_t|^\eta\right]&\leq C\mathbb{E}\left[|X_0|^\eta\right]+C\mathbb{E}\left[\int_0^s|b_r(X_r,\mathscr{L}_{X_r})|^\eta\,\mathrm{d}r\right]\\&\quad+C\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t\sigma_r(X_r,\mathscr{L}_{X_r})\,\mathrm{d}W_{S_r}\right|^\eta\right]\\&\leq C\mathbb{E}\left[|X_0|^\eta\right]+C\int_0^s\left(1+\mathbb{E}\left[|X_r|^\eta\right]\right)\mathrm{d}r+C\mathbb{E}\left[S_T^{\eta/2}\right]\\&\leq C\left(1+\mathbb{E}\left[|X_0|^\eta\right]\right)+C\int_0^s\mathbb{E}\left[\sup_{t\in[0,r]}|X_t|^\eta\right]\mathrm{d}r,\end{aligned}$$

which together with the Gronwall inequality yields that

$$\mathbb{E}\left[\sup_{t\in[0,s]}|X_t|^\eta\right]\leq C\left(1+\mathbb{E}\left[|X_0|^\eta\right]\right)\mathrm{e}^{Cs}\leq C\left(1+\mathbb{E}\left[|X_0|^\eta\right]\right),\quad s\in[0,T].$$

This completes the proof. □

4 Appendix

Lemma 4.1. *Let $\alpha \in (1/2, 1)$ and L_1 be a (stable) random variable with $\mathbb{E}\,\mathrm{e}^{\mathrm{i}\xi L_1} = \mathrm{e}^{-|\xi|^\alpha}$. Then there exist $c > 0$ and $\varrho > 0$ such that*

$$c = \alpha \mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1)|c + \varrho L_1|^{1-\alpha}\right]. \tag{4.1}$$

Proof. Let

$$g(c, \varrho) := c - \alpha \mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1)|c + \varrho L_1|^{1-\alpha}\right], \quad c > 0, \varrho \geq 0.$$

It follows from the dominated convergence theorem that for $c > 0$ and $\varrho \geq 0$,

$$\begin{aligned}
&\lim_{\epsilon_1 \to 0, \epsilon_2 \to 0} g(c + \epsilon_1, \varrho + \epsilon_2) \\
&\quad = c - \alpha \lim_{\epsilon_1 \to 0, \epsilon_2 \to 0} \mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1 + \epsilon_1 + \epsilon_2 L_1) \mathbb{1}_{\{c + \varrho L_1 \neq 0\}} |c + \varrho L_1 + \epsilon_1 + \epsilon_2 L_1|^{1-\alpha}\right] \\
&\quad = c - \alpha \mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1) \mathbb{1}_{\{c + \varrho L_1 \neq 0\}} |c + \varrho L_1|^{1-\alpha}\right] \\
&\quad = g(c, \varrho),
\end{aligned}$$

which means that $g(c, \varrho)$ is continuous for $c > 0$ and $\varrho \geq 0$. Using the dominated convergence theorem again, we know that for $c > 0$,

$$\lim_{\varrho \downarrow 0} g(c, \varrho) = c - \alpha c^{1-\alpha} = c^{1-\alpha}(c^\alpha - \alpha).$$

Pick $0 < c_1 < c_2$ such that $c_1^\alpha < \alpha$ and $c_2^\alpha > \alpha$. Then we get

$$\begin{aligned}
\lim_{\varrho \downarrow 0} g(c_1, \varrho) &= c_1^{1-\alpha}(c_1^\alpha - \alpha) < 0, \\
\lim_{\varrho \downarrow 0} g(c_2, \varrho) &= c_2^{1-\alpha}(c_2^\alpha - \alpha) > 0.
\end{aligned}$$

Now we conclude that there exist $c \in (c_1, c_2)$ and $\varrho > 0$ such that $g(c, \varrho) = 0$, and this completes the proof. $\square$

Example 4.2. *Let $d = 1$ and L_t be a symetric α-stable process on $\mathbb{R}$ with $\mathbb{E}\,\mathrm{e}^{\mathrm{i}\xi L_t} = \mathrm{e}^{-t|\xi|^\alpha}$ $(1/2 < \alpha < 1)$. Let*

$$b(\gamma) := \int_{\mathbb{R}} \operatorname{sgn}(x)|x|^{1-\alpha}\, \gamma(\mathrm{d}x), \quad \gamma \in \mathscr{P}.$$

It is easy to see that

$$|b(\gamma) - b(\tilde{\gamma})| \leq 2^\alpha \mathbb{W}_{1-\alpha}(\gamma, \tilde{\gamma}) \leq 2\mathbb{W}_{1-\alpha}(\gamma, \tilde{\gamma}), \quad \gamma, \tilde{\gamma} \in \mathscr{P}_{1-\alpha}.$$

By Lemma 4.1, we can pick two constants $c > 0$ and $\varrho > 0$ such that (4.1) holds. Consider the McKean-Vlasov SDE on $\mathbb{R}$:

$$\mathrm{d}X_t = b(\mathscr{L}_{X_t})\,\mathrm{d}t + \varrho\,\mathrm{d}L_t. \tag{4.2}$$

Since

$$b(\mathscr{L}_{L_t}) = \mathbb{E}\left[\operatorname{sgn}(L_t)|L_t|^{1-\alpha}\right] = 0,$$

we know that $X_t = L_t$ *is a solution to* (4.2) *with* $X_0 = 0$. *Next, we will show that* $X_t = ct^{1/\alpha} + \varrho L_t$ *also solves* (4.2). *To this aim, we use the scaling property of* L_t *to get that for all* $s \in (0, T]$,

$$\begin{aligned} b(\mathscr{L}_{cs^{1/\alpha}+\varrho L_s}) &= \mathbb{E}\left[\operatorname{sgn}(cs^{1/\alpha} + \varrho s^{1/\alpha} L_1)|cs^{1/\alpha} + \varrho s^{1/\alpha} L_1|^{1-\alpha}\right] \\ &= s^{\frac{1}{\alpha}-1}\mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1)|c + \varrho L_1|^{1-\alpha}\right], \end{aligned}$$

which, together with (4.1), *implies that*

$$\begin{aligned} \int_0^t b(\mathscr{L}_{cs^{1/\alpha}+\varrho L_s})\,\mathrm{d}s &= \int_0^t s^{\frac{1}{\alpha}-1}\,\mathrm{d}s \times \mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1)|c + \varrho L_1|^{1-\alpha}\right] \\ &= \alpha t^{1/\alpha}\mathbb{E}\left[\operatorname{sgn}(c + \varrho L_1)|c + \varrho L_1|^{1-\alpha}\right] \\ &= ct^{1/\alpha}. \end{aligned}$$

This means that $X_t = ct^{1/\alpha} + \varrho L_t$ *is a solution to* (4.2) *with* $X_0 = 0$. *Thus, the SDE* (4.2) *with initial value* $X_0 = 0$ *has at least two strong solutions:* L_t *and* $ct^{1/\alpha} + \varrho L_t$, *where* $c > 0$ *and* $\varrho > 0$ *are two constants satisfying* (4.1).

Acknowledgement. C.-S. Deng is supported by National Natural Science Foundation of China (12371149) and Natural Science Foundation of Hubei Province of China (2022CF-B129). X. Huang is supported by National Key R&D Program of China (No. 2022Y-FA1006000) and National Natural Science Foundation of China (12271398).

References

[1] P.-E. Chaudru de Raynal, N. Frikha, *Well-posedness for some non-linear SDEs and related PDE on the Wasserstein space*, J. Math. Pures Appl. 159 (2022), 1–167.

[2] P.-E. Chaudru de Raynal, S. Menozzi, E. Priola, *Schauder estimates for drifted fractional operators in the supercritical case*, J. Funct. Anal. 278(2020), no.8, 108425, 57 pp.

[3] M.-F. Chen, *From Markov Chains to Non-Equilibrium Particle Systems.* Word Scientific, Singapore 2004 (2nd edn).

[4] Z.-Q. Chen, X. Zhang, G. Zhao, *Supercritical SDEs driven by multiplicative stable-like Lévy processes*, Trans. Amer. Math. Soc. 374 (2021), 7621–7655.

[5] C.-S. Deng, X. Huang, *Well-posedness for McKean-Vlasov SDEs with distribution dependent stable noises*, arXiv:2306.10970.

[6] X. Huang, F.-Y. Wang, *Singular McKean-Vlasov (reflecting) SDEs with distribution dependent noise*, J. Math. Anal. Appl. 514 (2022), no. 1, Paper No. 126301, 21 pp.

[7] B. Jourdain, S. Méléard, W. A. Woyczynski, *Nonlinear SDEs driven by Lévy processes and related PDEs*, ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008), 1–29.

[8] S. Menozzi, X. Zhang, *Heat kernel of supercritical nonlocal operators with unbounded drifts*, J. Éc. polytech. Math. 9 (2022), 537–579.

[9] H.P. McKean Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Nat. Acad. Sci. U.S.A. 56 (1966), 1907–1911.

[10] N. Frikha, V. Konakov, S. Menozzi, *Well-posedness of some non-linear stable driven SDEs*, Discrete Contin. Dyn. Syst. 41 (2021), 849–898.

[11] P. Ren, *Singular McKean-Vlasov SDEs: Well-posedness, regularities and Wang's Harnack inequality*, Stochastic Process. Appl. 156 (2023), 291–311.

[12] M. Röckner, X. Zhang, *Well-posedness of distribution dependent SDEs with singular drifts*, Bernoulli 27 (2021), 1131–1158.

[13] K. Sato, *Lévy processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge 1999.

[14] L. Silvestre, *On the differentiability of the solution to an equation with drift and fractional diffusion*, Indiana Univ. Math. J. 61(2012), 557-584.

[15] A.-S. Sznitman, *Topics in propagation of chaos*, in: Lect. Notes in Math., vol. 1464, Springer-Verlag, Berlin 1991.

[16] F.-Y. Wang, *Distribution dependent reflecting stochastic differential equations*, Sci. China Math. 66 (2023), 2411–2456.

[17] M. Wu, Z. Hao, *Well-posedness of density dependent SDE driven by α-stable process with Hölder drifts*. Stochastic Process. Appl. 164 (2023) 416–442.

[18] G. Zhao, *On distribution dependent SDEs with singular drifts*, arXiv:2003.04829v3.