Journal of Theoretical Probability (2026) 39:10
https://doi.org/10.1007/s10959-025-01465-2

A Note on Lévy-Driven McKean–Vlasov Stochastic Differential Equations Under Monotonicity

Jianhai Bao[1] · Yao Liu[2] · Jian Wang[3]

Received: 7 December 2024 / Revised: 3 July 2025 / Accepted: 1 November 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract

In this note, under weak monotonicity and weak coercivity, we address strong well-posedness of McKean–Vlasov stochastic differential equations (SDEs) driven by Lévy jump processes, where the coefficients are Lipschitz continuous (with respect to the measure variable) under the L^β-Wasserstein distance for $\beta \in [1, 2]$. Moreover, the issues of on the weak propagation of chaos (i.e., convergence in distribution via convergence of the empirical measure) and strong propagation of chaos (i.e., at the level paths by coupling) are explored simultaneously. To treat the strong well-posedness of McKean–Vlasov SDEs we are interested in, we investigate strong well-posedness of classical time-inhomogeneous SDEs with jumps under a local weak monotonicity and a global weak coercivity. Such a result is of independent interest, and, most importantly, can provide an available reference on strong well-posedness of Lévy-driven SDEs under the monotonicity condition, which has been is missing for a long time. Based on the theory derived, along with the interlacing technique and the Banach fixed point theorem, the strong well-posedness of McKean–Vlasov SDEs driven by Lévy jump processes can be established. Additionally, as a potential extension, strong well-posedness and conditional propagation of chaos are treated for Lévy-driven McKean–Vlasov SDEs with common noise under a weak monotonicity.

✉ Jian Wang
jianwang@fjnu.edu.cn

Jianhai Bao
jianhaibao@tju.edu.cn

Yao Liu
liuyaomath@163.com

[1] Center for Applied Mathematics, Tianjin University, Tianjin 300072, People's Republic of China

[2] School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350007, People's Republic of China

[3] School of Mathematics and Statistics & Key Laboratory of Analytical Mathematics and Applications (Ministry of Education) & Fujian Provincial Key Laboratory of Statistics and Artificial Intelligence, Fujian Normal University, 350007 Fuzhou, People's Republic of China

Published online: 22 November 2025

Keywords McKean-Vlasov SDE-Lévy process · weak monotonicity-weak coercivity · well-posedness-propagation of chaos

Mathematics Subject Classification (2020) 60G51 · 60J25 · 60J76

1 Introduction and Main Results

1.1 Background

The treatment of strong/weak well-posedness is a starter to explore qualitative/quantitative studies of SDEs under consideration. In the past few decades, strong well-posedness of SDEs with Brownian motion noises has been investigated under various scenarios; see, for instance, [23, 29] for the Lipschitz continuity and the linear growth; [29, Theorem 3.4] concerning the local Lipschitz condition plus the linear growth; [29, Theorem 3.5] regarding the local Lipschitz continuity along with the Lyapunov condition, and [33, Theorem 3.1.1] with regard to the local weak monotonicity besides the global weak coercivity. Moreover, under the local $L^q(L^p)$-condition, strong well-posedness of singular SDEs has also advanced greatly via Zvonkin's transformation; see, e.g., [37, 38] and references therein.

At the same time, there is a considerable amount of literature concerned with strong well-posedness of SDEs driven by Lévy jump process. As we know, under the standard assumption that coefficients are globally Lipschitz and of linear growth, SDEs with pure jumps are strongly well-posed; see, for example, [3, 23]. In case drifts and Brownian diffusions are locally Lipschitz and the jump coefficient is globally Lipschitz, in addition to a weak coercivity, strong well-posedness of SDEs with jumps was explored in [2]. With contrast to SDEs with Brownian motion noises, strong well-posedness of SDEs driven by pure jump processes is rare under the weak monotonicity (which is also termed as the one-sided Lipschitz condition) and the weak coercivity. As stated in [28], many authors quote strong well-posedness of SDEs with jumps under the one-sided Lipschitz condition by *claiming that it is totally well known nevertheless without providing any reference or referring to references which do not contain it at all*. This phenomenon is further stressed in [31] as follows: "However, we could not find a reference in the literature that covers our setting completely." Based on the point of view above, via a truncation approach, [28] addressed strong well-posedness of SDEs driven by Brownian motions and compensated Poisson random measures, where a local weak monotonicity and a global weak coercivity were imposed. Unsatisfactorily, due to the limitation of the method adopted, the local weak monotonicity and the global weak coercivity put in [28] cannot go back to the classical one (see, e.g., [33, (3.1.3) and (3.1.4)] for more details) when the pure jump term involved vanishes. Additionally, we would like to mention [17] for a much more general setup, where the driven noise is a square-integrable semimartingale.

In recent years, there is great progresses as well on strong well-posedness of McKean–Vlasov SDEs driven by Brownian motions; see, e.g., monographs [5, 32, 36]. We also would like to mention that [10] explored strong well-posedness of McKean–Vlasov SDEs, which allow drifts and diffusions to be of super-linear growth in measure and state variables. Meanwhile, strong well-posedness of regular McKean–Vlasov SDEs with jumps has also attracted a lot of interest; see, e.g., [11, 14, 27, 31]. In detail, [11, 27] is concerned with the additive noise and the drift involved in [14] is of linear growth with respect to the state variable. Furthermore, strong well-posedness of McKean–Vlasov SDEs with singular interaction kernels and symmetric α-stable noises has been tackled in [12, 13] and [16, 18] by the aid of the (two-step) fixed point theorem and the nonlinear martingale problem, respectively. Additionally, via a Fourier-based Picard iteration approach, [1] considered strong well-posedness of a class of McKean–Vlasov SDEs with Lévy jumps, where the underlying drift coefficient is affine in the state variable. When the coefficients are Lipschitz continuous with respect to the measure variable under the L^2-Wasserstein distance and non-globally Lipschitz continuous with respect to the state variable, strong well-posedness of Lévy-driven McKean–Vlasov SDEs has been treated in [31, 35]. In case that the associated coefficients are L^β-Wasserstein continuous (for $1 \le \beta \le 2$) as far as the measure variable is concerned, and Lipschitz continuous in the state variable, the paper [7] probed into well-posedness of Lévy-driven McKean–Vlasov SDEs.

Inspired by the aforementioned literature, in this note we aim to investigate the strong well-posedness for a class of Lévy-driven McKean–Vlasov SDEs under a weak monotonicity and a weak coercivity, which will weaken the associated conditions and improve the corresponding results in, e.g., [1, 7, 11, 31, 35] in various aspects.

1.2 Well-Posedness of McKean–Vlasov SDEs

More precisely, in this note we focus on the following McKean–Vlasov SDE on $\mathbb{R}^d$:

$$\begin{aligned} \mathrm{d}X_t &= b(X_t, \mathscr{L}_{X_t})\,\mathrm{d}t + \int_U f(X_{t-}, \mathscr{L}_{X_t}, z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z) \\ &\quad + \int_V g(X_{t-}, \mathscr{L}_{X_t}, z)\,N(\mathrm{d}t, \mathrm{d}z). \end{aligned} \tag{1.1}$$

Herein, $\mathscr{L}_{X_t}$ stands for the law of X_t; $X_{t-} := \lim_{s\uparrow t} X_s$; $b : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d$, and $f, g : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ are measurable maps, where $\mathscr{P}(\mathbb{R}^d)$ means the family of probability measures on $\mathbb{R}^d$; $U, V \subset \mathbb{R}^d_0 := \mathbb{R}^d \backslash \{\mathbf{0}\}$ so that $U \cap V = \emptyset$; N is a Poisson random measure, carried on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with the intensity measure $\mathrm{d}t \times \nu(\mathrm{d}z)$ for a σ-finite measure $\nu(\mathrm{d}z)$, and $\widetilde{N}(\mathrm{d}t, \mathrm{d}z) := N(\mathrm{d}t, \mathrm{d}z) - \mathrm{d}t\,\nu(\mathrm{d}z)$ represents the associated compensated Poisson measure. Furthermore, we shall assume that for some $\beta \in [1, 2]$ and any fixed $x \in \mathbb{R}^d$ and $\mu \in \mathscr{P}(\mathbb{R}^d)$,

$$\nu(|f(x, \mu, \cdot)|^2 \mathbb{1}_U(\cdot)) + \nu((1 \vee |\cdot|^\beta \vee |g(x, \mu, \cdot)|^\beta)\mathbb{1}_V(\cdot)) < \infty, \tag{1.2}$$

where $\nu(f) := \int_{\mathbb{R}^d} f(x)\nu(\mathrm{d}x)$ for a ν-integrable function $f : \mathbb{R}^d \to \mathbb{R}$. For $p > 0$, denote by $\mathbb{W}_p$ the L^p-Wasserstein distance:

$$\mathbb{W}_p(\mu_1, \mu_2) := \inf_{\pi \in \mathscr{C}(\mu_1,\mu_2)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p \, \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{1\vee p}}, \quad \mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d),$$

where $\mathscr{P}_p(\mathbb{R}^d) := \{\mu \in \mathscr{P}(\mathbb{R}^d) : \mu(|\cdot|^p) < \infty\}$ and $\mathscr{C}(\mu_1, \mu_2)$ is the set of couplings for μ_1 and μ_2.

To guarantee the well-posedness of the SDE (1.1), the following assumptions are in force.

($\mathbf{A}_1$) for fixed $\mu \in \mathscr{P}_\beta(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$, $\mathbb{R}^d \ni x \mapsto b(x, \mu)$ and $\mathbb{R}^d \ni x \mapsto f(x, \mu, z)$ are continuous and locally bounded, and there exists a constant $L_1 > 0$ such that for any $x, y, z \in \mathbb{R}^d$, and $\mu_1, \mu_2 \in \mathscr{P}_\beta(\mathbb{R}^d)$,

$$\begin{aligned} &2\langle b(x, \mu_1) - b(y, \mu_2), x - y\rangle + \nu(|f(x, \mu_1, \cdot) - f(y, \mu_2, \cdot)|^2 \mathbb{1}_U(\cdot)) \\ &\le L_1(|x-y|^2 + \mathbb{W}_\beta(\mu_1, \mu_2)^2), \end{aligned} \tag{1.3}$$

and

$$|g(x, \mu_1, z) - g(y, \mu_2, z)| \le L_1(1 + |z|)(|x-y| + \mathbb{W}_\beta(\mu_1, \mu_2)); \tag{1.4}$$

($\mathbf{A}_2$) there exists a constant $L_2 > 0$ such that for any $x \in \mathbb{R}^d$ and $\mu \in \mathscr{P}_\beta(\mathbb{R}^d)$,

$$2\langle x, b(x, \mu)\rangle + \nu(|f(x, \mu, \cdot)|^2 \mathbb{1}_U(\cdot)) \le L_2(1 + |x|^2 + \mu(|\cdot|^\beta)^{\frac{2}{\beta}});$$

($\mathbf{A}_3$) for any $T, R > 0$ and $\mu \in C([0, T]; \mathscr{P}_\beta(\mathbb{R}^d))$,

$$\int_0^T \Big(\sup_{\{|x|\le R\}} |b(x, \mu_t)| + \int_U \sup_{\{|x|\le R\}} |f(x, \mu_t, z)|^2 \nu(\mathrm{d}z) \Big) \, \mathrm{d}t < \infty.$$

The first main result in this paper is stated as follows.

Theorem 1.1 *Assume that Assumptions* ($\mathbf{A}_1$)–($\mathbf{A}_3$) *hold, and suppose further* $X_0 \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$. *Then, the McKean–Vlasov SDE* (1.1) *admits a unique strong solution* $(X_t)_{t\ge 0}$ *satisfying that, for any fixed* $T > 0$, *there exists a constant* $C_T > 0$ *such that*

$$\mathbb{E}|X_t|^\beta \le C_T(1 + \mathbb{E}|X_0|^\beta), \quad 0 \le t \le T. \tag{1.5}$$

In addition, if Assumption ($\mathbf{A}_2$) *is replaced by the following stronger one: for some* $L_3 > 0$,

$$\langle x, b(x, \mu)\rangle \vee \nu(|f(x, \mu, \cdot)|^2 \mathbb{1}_U(\cdot)) \le L_3\big(1 + |x|^2 + \mu(|\cdot|^\beta)^{\frac{2}{\beta}}\big), \tag{1.6}$$

then, for any $T > 0$, there exists a constant $C'_T > 0$ such that

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t|^\beta\Big) \le C'_T(1+\mathbb{E}|X_0|^\beta). \tag{1.7}$$

Below, we make some comments on Theorem 1.1 and assumptions mentioned above.

Remark 1.2 (i) (1.3) shows that b, f satisfy the so-called one-sided Lipschitz condition so they are allowed to be non-globally Lipschitz with respect to the state variable as the following example reveals. For $x, z \in \mathbb{R}^d$ and $\mu \in \mathscr{P}_\beta(\mathbb{R}^d)$, let

$$\begin{aligned} b(x,\mu) &= C_1x - C_2x|x|^2 + \mu(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\mathbf{1},\\ f(x,\mu,z) &= C_3z(1+C_4|x|^2+\mu(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}),\\ g(x,\mu,z) &= (\mathbf{1}+z)(1+|x|+\mu(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}), \end{aligned}$$

where $C_1, C_2, C_3, C_4 > 0$, $\mathbf{1} := (1,\cdots,1)^\top \in \mathbb{R}^d$, and $h : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz. Then, Assumptions ($\mathbf{A}_1$), ($\mathbf{A}_2$) and ($\mathbf{A}_3$) are valid, respectively, provided $\nu(|\cdot|^2\mathbb{1}_U(\cdot)) < \infty$ and $C_2 > 12C_3^2C_4^2\nu(|\cdot|^2\mathbb{1}_U(\cdot))$.

(ii) For the case $\beta \in (0, 1)$, there is no uniqueness of the McKean–Vlasov SDE (1.1) as [7, Remark 2] shown. See also Remark 3.1(i) for additional comments. So, in this work we focus only on the setting $\beta \in [1, 2]$. As for the case $U = \{z \in \mathbb{R}^d : 0 < |z| \le 1\}$, $V = \{z \in \mathbb{R}^d : |z| > 1\}$, $f(x,\mu,z) = \sigma(x,\mu)z$, and $g(x,\mu,z) = \sigma(x,\mu)z$ with $\sigma : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^d$, the strong well-posedness of (1.1) was addressed in [7, Theorem 1] when b, σ are Lipschitz and L^β-Wasserstein Lipschitz with respect to the spatial variable and the measure variable, respectively. (In particular, in this special case, by (1.2) the jump process under consideration is allowed to be an α-stable process with $\alpha \in (1, \beta)$.) However, in Theorem 1.1, b and f might be non-globally Lipschitz with respect to the state variables as the previous example demonstrates. In addition, [11] addressed the strong well-posedness of the McKean–Vlasov SDE (1.1) with additive noise under the following condition: there is some $L_0 > 0$ such that for all $x, y \in \mathbb{R}^d$ and $\mu_1, \mu_2 \in \mathscr{P}_\beta(\mathbb{R}^d)$

$$\langle x-y, b(x,\mu_1)-b(y,\mu_2)\rangle \le L_0(|x-y| + \mathbb{W}_\beta(\mu_1,\mu_2))|x-y|.$$

Apparently, the preceding condition is rigorous than the one imposed in (1.3).

(iii) In addition to the fixed point theorem used in the proof of Theorem 1.1, Yamada–Watanabe's principle is another approach that is applied widely to prove strong well-posedness, e.g., see [19–21] concerning McKean–Vlasov diffusions and [11, 22] for McKean–Vlasov SDEs with additive Lévy noise. In particular, under the following local Lipschitz continuity: for some $L_1 > 0$, and any $x, y \in \mathbb{R}^d$, $\mu_1, \mu_2 \in \mathscr{P}_1(\mathbb{R}^d)$,

$$|b(x,\mu_1)-b(y,\mu_2)| + \nu(|f(x,\mu_1,z)-f(y,\mu_2,z)|\mathbb{1}_U(|\cdot|))$$

$$\leq L_1(1+|x|+|y|+\mu_1(|\cdot|)+\mu_2(|\cdot|))(|x-y|+\mathbb{W}_1(\mu_1,\mu_2)),$$

strong well-posedness of a kind of McKean–Vlasov equations with jumps was investigated in [14] by Yamada–Watanabe's principle. On the other hand, the strong existence is implied by weak existence of McKean–Vlasov SDEs and strong well-posedness of the corresponding decoupled SDEs. In literature, the statement above is called the modified Yamada–Watanabe principle [20, Lemma 3.4]. In general, some kind of growth condition needs to be imposed to verify the tightness of the sequence of Euler-type approximation equations in order to prove the weak existence. For example, the growth condition: for some $L_2 > 0$, and any $x \in \mathbb{R}^d$, $\mu \in \mathscr{P}_1(\mathbb{R}^d)$,

$$|b(x,\mu)|+|f(x,\mu,z)| \leq L_2(1+|z|)(1+|x|+\mu(|\cdot|))$$

was set in [14]. Clearly, such condition is stronger than Assumption ($\mathbf{A}_2$) in our paper. See Remark 2.3 for comments on the setting of classical SDEs with jumps.

1.3 Propagation of Chaos

The McKean–Vlasov SDE (1.1) arises naturally in the framework of the limit for the mean-field interacting particle system in the form:

$$\begin{cases} \mathrm{d}X_t^{i,n} = b(X_t^{i,n},\bar{\mu}_t^n)\,\mathrm{d}t + \displaystyle\int_U f(X_{t-}^{i,n},\bar{\mu}_{t-}^n,z)\,\widetilde{N}^i(\mathrm{d}t,\mathrm{d}z) \\ \qquad\qquad + \displaystyle\int_V g(X_{t-}^{i,n},\bar{\mu}_{t-}^n,z)\,N^i(\mathrm{d}t,\mathrm{d}z), \\ X_0^{i,n} = X_0^i, \qquad i=1,2,\cdots,n, \end{cases} \tag{1.8}$$

where $\bar{\mu}_t^n := \frac{1}{n}\sum_{i=1}^n \delta_{X_t^{i,n}}$, and $\{N^i(\mathrm{d}t,\mathrm{d}z)\}_{1\leq i\leq n}$ are independent Poisson measures with the intensity measure $\mathrm{d}t \times \nu(\mathrm{d}z)$. The link between (1.1) and (1.8) lies in that the dynamics of the particle system (1.8) are expected to be described by (1.1) when the number of particles n goes to infinity. This property is the so-called propagation of chaos, which was originally studied by Kac [25] for the Boltzmann equation and was further developed by Sznitman [34]. The propagation of chaos can be interpreted in the weak sense (i.e., in the distribution through the convergence of the empirical measure $\bar{\mu}_t^n$) and in the strong sense (i.e., from the point of view of paths via coupling); see [7–9] and references within.

In this subsection, our purpose is to prove quantitative propagation of chaos both in the weak sense and the strong sense, respectively, concerning the mean-field interacting particle system (1.8) with $f(x,z) = f(x,\mu,z)$ (i.e., f is unrelated to the measure variable). For this, we need to replace (1.3) in Assumption ($\mathbf{A}_1$) by the following stronger version:

($\mathbf{A}_1'$) assume that $\beta \in (1,2]$; for fixed $\mu \in \mathscr{P}_\beta(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$, $\mathbb{R}^d \ni x \mapsto b(x,\mu)$ and $\mathbb{R}^d \ni x \mapsto f(x,z)$ are continuous and locally bounded, and for some

$p \in [1, \beta)$, there exists a constant $L_4 > 0$ such that for any $x, y, z \in \mathbb{R}^d$, and $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)$,

$$\begin{aligned} &2\langle b(x, \mu_1) - b(y, \mu_2), x - y\rangle + \nu(|f(x, \cdot) - f(y, \cdot)|^2 \mathbb{1}_U(\cdot)) \\ &\leq L_4(|x - y| + \mathbb{W}_p(\mu_1, \mu_2))|x - y|, \end{aligned} \tag{1.9}$$

and

$$|g(x, \mu_1, z) - g(y, \mu_2, z)| \leq L_4(1 + |z|)(|x - y| + \mathbb{W}_p(\mu_1, \mu_2)). \tag{1.10}$$

Under ($\mathbf{A}_1'$), ($\mathbf{A}_2$) and ($\mathbf{A}_3$), besides (1.2), SDEs (1.1) and (1.8) have unique strong solutions for $X_0 \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$ by taking Theorem 1.1 into consideration.

Let $\{(X_t^i)_{t\geq 0}\}_{1\leq i\leq n}$ be n-independent versions of the unique solution to the SDE (1.1) with $f(x, z) = f(x, \mu, z)$. In particular, each $(X_t^i)_{t\geq 0}$, $1 \leq i \leq n$, shares the same distribution. The following theorem provides quantitative characterizations of strong/weak propagation of chaos in finite time.

Theorem 1.3 *Assume that Assumptions* ($\mathbf{A}_1'$), ($\mathbf{A}_2$) *and* ($\mathbf{A}_3$) *hold, and suppose further* $X_0^i \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$ *for any* $1 \leq i \leq n$. *Let* $(\mu_t)_{t\geq 0}$ *be the common distribution of* $(X_t^i)_{t\geq 0}$ *for all* $1 \leq i \leq n$, *and* $\bar{\mu}_t^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_t^{i,n}}$, *where* $\{(X_t^{i,n})_{t\geq 0}\}_{1\leq i\leq n}$ *is the solution to the mean-filed interacting particle system* (1.8) *with* $f(x, z) = f(x, \mu, z)$. *Then, for any fixed* $T > 0$, *there is a constant* $C_T > 0$ *such that*

$$\mathbb{E}\mathbb{W}_p^p(\bar{\mu}_t^n, \mu_t) \leq C_T \phi_{p,\beta,d}(n), \quad t \in [0, T], \tag{1.11}$$

where the quantity $\phi_{p,\beta,d}(n)$ *is defined as below:*

$$\phi_{p,\beta,d}(n) := \begin{cases} & d = 1, 2; \\ n^{-(1-\frac{p}{\beta})}, & d = 3 \text{ and } 1 \leq p < \frac{3}{2};\ d = 3 \text{ and } \frac{3}{2} \leq p < \beta \text{ and } \beta < \frac{3p}{3-p}; \\ & d \geq 4 \text{ and } \beta < \frac{dp}{d-p}; \\ n^{-\frac{p}{d}}, & d = 3,\ 1 \leq p < \frac{3}{2} \text{ and } \beta \geq \frac{3p}{3-p};\ d \geq 4 \text{ and } \beta \geq \frac{dp}{d-p}. \end{cases} \tag{1.12}$$

Furthermore, for any $0 \leq q_1 < q_2 < 1$, *there exists a constant* $\hat{C}_T > 0$ *such that*

$$\mathbb{E}\Big(\sup_{0\leq t\leq T} |X_t^{i,n} - X_t^i|^{pq_1}\Big) \leq \frac{q_2}{q_2 - q_1}\big(\hat{C}_T \phi_{p,\beta,n}(n)\big)^{q_1}. \tag{1.13}$$

To proceed, concerning Theorem 1.3, we make a remark on the structure of f and the prerequisite (1.9).

Remark 1.4 Recall that the well-posedness of (1.1) is explored under ($\mathbf{A}_1$) via the interlacing trick. So, it is quite natural to investigate the issue on propagation of chaos under ($\mathbf{A}_1$). While the empirical measure involved in (1.8) is random, so the above technique does not work any more. In turn, we reinforce Assumption (1.3) as (1.9).

Once f is dependent on the measure variable (as given in (1.8)), the assumption (1.9) can be formulated as below: for any $x, y \in \mathbb{R}^d$, and $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)$,

$$2\langle b(x,\mu_1) - b(y,\mu_2), x - y\rangle + \nu(|f(x,\mu_1,\cdot) - f(y,\mu_2,\cdot)|^2 \mathbb{1}_U(\cdot)) \\ \le L_4(|x-y| + \mathbb{W}_p(\mu_1,\mu_2))|x-y|.$$

In the preceding inequality, in case of $b \equiv \mathbf{0}$, one has $\nu(|f(x,\mu_1,\cdot) - f(x,\mu_2,\cdot)|^2 \mathbb{1}_U(\cdot)) = 0$ for arbitrary $\mu_1, \mu_2 \in \mathscr{P}_p(\mathbb{R}^d)$. Accordingly, we can conclude that f is irrelevant to the measure variable (at least when $b \equiv \mathbf{0}$). Let b be defined as in Remark 1.2(i) and $f(x,z) = C_3 z(1 + C_4|x|^2)$ for some $C_3, C_4 > 0$. For this case, (1.9) is valid when the assumptions in Remark 1.2(i) are satisfied by examining the proof of Remark 1.2(i); see the end of Sect. 3 for more details.

The content of this paper is organized as follows. In Sect. 2, via a Picard iteration approach, we investigate strong well-posedness of classical time-inhomogeneous SDEs with Lévy noises under a local weak monotonicity and a weak coercivity, which is quite interesting in its own right. Also, by invoking the interlacing technique, a uniform moment estimate in a finite horizon is established in Sect. 2. Based on the theory derived in Sect. 2, along with the Banach fixed point theorem and the interlacing technique, the proof of Theorem 1.1 is complete in Sect. 3. In addition, the remaining part of Sect. 3 is devoted to the proof of Theorem 1.3, which is concerned with the weak propagation of chaos and the associated strong version. In the last section, we extend accordingly Theorems 1.1 and 1.3 to Lévy-driven McKean–Vlasov SDEs with common noise.

2 Well-Posedness of Classical SDEs with Lévy Noises

The fixed point theorem is one of the powerful tools to investigate well-posedness of McKean–Vlasov SDEs under variant settings. For this purpose, the corresponding distribution-frozen SDE (which, in literature, is also named as a decoupled SDE) needs to be considered. In other words, by the aid of the decoupled SDE (which definitely is a time-inhomogeneous SDE), along with the fixed point theorem, the well-posedness of McKean–Vlasov SDEs can be treated. Inspired by the aforementioned routine, in this section, we focus on the following time-inhomogeneous SDE: for any $t \ge 0$,

$$\mathrm{d}X_t = b_t(X_t)\,\mathrm{d}t + \int_U f_t(X_{t-}, z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z) + \int_V g_t(X_{t-}, z)\,N(\mathrm{d}t, \mathrm{d}z), \quad (2.1)$$

where $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$, and $f, g : [0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are jointly measurable; the subsets U, V, and the random measures $N(\mathrm{d}t, \mathrm{d}z)$, $\widetilde{N}(\mathrm{d}t, \mathrm{d}z)$ are untouched as those in (1.1). In this section, we assume that for all $t \ge 0$ and $x \in \mathbb{R}^d$,

$$\nu(|f_t(x,\cdot)|^2 \mathbb{1}_U(\cdot)) + \nu((|g_t(x,\cdot)|^\beta \vee 1)\mathbb{1}_V(\cdot)) < \infty,$$

where $\beta \in (0, 2]$. Herein, we emphasize that the results to be derived in this section hold true for all $\beta \in (0, 2]$ instead of $\beta \in [1, 2]$. In general, the second stochastic integral and the third one on the right hand side of (2.1) are concerned with small jumps and big jumps, respectively.

2.1 Main Results

Inspired by the diffusive setting (e.g., [33, Theorem 3.1.1]), to address the well-posedness of (2.1), we impose the following local weak monotonicity, weak coercivity, and local integrability on the coefficients. In detail, we shall assume that

($\mathbf{H}_1$) for each fixed $t \geq 0$ and $z \in \mathbb{R}^d$, $\mathbb{R}^d \ni x \mapsto b_t(x)$, $\mathbb{R}^d \ni x \mapsto f_t(x, z)$ and $\mathbb{R}^d \ni x \mapsto g_t(x, z)$ are continuous and locally bounded; for any fixed $R > 0$, there exists an increasing and locally integrable function $[0, \infty) \ni t \mapsto K_t(R)$ such that for any $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq R$ and $t \geq 0$,

$$2\langle x - y, b_t(x) - b_t(y)\rangle + \nu(|f_t(x, \cdot) - f_t(y, \cdot)|^2 \mathbb{1}_U(\cdot)) \leq K_t(R)|x - y|^2; \tag{2.2}$$

($\mathbf{H}_2$) there exists an increasing and locally integrable function $\varphi : [0, \infty) \to (0, \infty)$ such that for $x \in \mathbb{R}^d$ any $t \geq 0$,

$$\begin{aligned}&2\langle x, b_t(x)\rangle + \nu(|f_t(x, \cdot)|^2 \mathbb{1}_U(\cdot)) \\ &\qquad + \beta^{-1} 2^{\frac{\beta}{2}+1}(1 + |x|^2)^{1-\frac{\beta}{2}} \nu(|g_t(x, \cdot)|^\beta \mathbb{1}_V(\cdot)) \leq \varphi(t)(1 + |x|^2);\end{aligned}$$

($\mathbf{H}_3$) for any $R, T > 0$,

$$\int_0^T \Big(\sup_{\{|x| \leq R\}} |b_t(x)| + \int_U \sup_{\{|x| \leq R\}} |f_t(x, z)|^2 \nu(\mathrm{d}z) \Big) \mathrm{d}t < \infty.$$

Under assumptions above, (2.1) is strongly well-posed as the theorem below states.

Theorem 2.1 *Assume that* ($\mathbf{H}_1$)–($\mathbf{H}_3$) *hold true, and suppose further* $X_0 \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$. *Then,* (2.1) *has a unique strong solution* $(X_t)_{t \geq 0}$ *satisfying that, for any* $T > 0$, *there exists a constant* $C_T > 0$ *such that*

$$\sup_{0 \leq t \leq T} \mathbb{E}|X_t|^\beta \leq C_T(1 + \mathbb{E}|X_0|^\beta). \tag{2.3}$$

To achieve a uniform moment estimate in a finite horizon, we strengthen ($\mathbf{H}_2$) as follows:

($\mathbf{H}_2'$) there exists an increasing and locally integrable function $\phi : [0, \infty) \to (0, \infty)$ such that for any $x \in \mathbb{R}^d$ and $t \geq 0$,

$$\langle x, b_t(x)\rangle \vee \nu(|f_t(x, \cdot)|^2 \mathbb{1}_U(\cdot)) \vee \big(\nu(|g_t(x, \cdot)|^\beta \mathbb{1}_V(\cdot))\big)^{\frac{2}{\beta}} \leq \phi(t)(1 + |x|^2).$$

With Assumption ($\mathbf{H}_2'$) at hand, a stronger version of (2.3) can be obtained.

Theorem 2.2 *Assume that* ($\mathbf{H}_1$), ($\mathbf{H}_2'$) *and* ($\mathbf{H}_3$) *hold true, and suppose further* $X_0 \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$. *Then, for any fixed* $T > 0$, *there is a constant* $C_T' > 0$ *such that*

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t|^\beta\Big) \le C_T'(1+\mathbb{E}|X_0|^\beta). \tag{2.4}$$

Before we move to the next subsection, we make the following remark.

Remark 2.3 In literature, there are several ways to show strong well-posedness of SDEs under consideration. In particular, as long as the weak existence and the pathwise uniqueness are available, the strong well-posedness can be derived by leveraging on the Yamada–Watanabe theorem; see [24] for classical SDEs driven by semimartingales. Concerning the aforementioned method, one needs to examine tightness of the solution processes associated with the approximated SDEs. To this end, in general, some growth conditions (with respect to the state variable) related to coefficients are imposed to show the equicontinuity in probability. Nevertheless, by adopting the procedure in the present work, the growth condition on each coefficient can be neglected.

2.2 Well-Posedness of SDEs with Small Jumps

In this subsection, we adopt a two-step strategy to explore the well-posedness of (2.1). Firstly, we establish the well-posedness of the SDE (without big jumps): for any $t > 0$,

$$\mathrm{d}Y_t = b_t(Y_t)\,\mathrm{d}t + \int_U f_t(Y_{t-}, z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z), \quad Y_0 = X_0. \tag{2.5}$$

Afterward, the well-posedness of (2.1) can be tackled by splicing together big jumps via the so-called interlacing technique (see, e.g., [3, p. 236] and [23, p. 244–246]).

Proposition 2.4 *Under Assumptions of Theorem 2.1 with* $g \equiv \mathbf{0}$, (2.5) *has a unique strong solution* $(Y_t)_{t\ge0}$ *satisfying that, for any* $T > 0$, *there exists a constant* $C_T > 0$ *such that*

$$\sup_{0\le t\le T}\mathbb{E}|Y_t|^\beta \le C_T(1+\mathbb{E}|Y_0|^\beta). \tag{2.6}$$

To address the well-posedness of (2.5), we appeal to the Picard iteration approach. So, in the sequel, we work with the following iterated SDE: for any $t > 0$ and integer $n \ge 1$,

$$\mathrm{d}Y_t^{(n)} = b_t(Y_{t_n}^{(n)})\,\mathrm{d}t + \int_U f_t(Y_{t_n}^{(n)}, z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z), \quad Y_0^{(n)} = Y_0, \tag{2.7}$$

where $t_n := \lfloor tn \rfloor / n$ with $\lfloor \cdot \rfloor$ being the floor function. Below, for the notation brevity, we set

$$p_t^{(n)} := Y_{t_n}^{(n)} - Y_t^{(n)} \quad \text{and} \quad p_{t-}^{(n)} := Y_{t_n}^{(n)} - Y_{t-}^{(n)}.$$

Additionally, for any $R > 0$, we define the stopping time $\tau_R^{(n)}$ by

$$\tau_R^{(n)} = \inf\{t \geq 0 : |Y_t^{(n)}| > R/2\}.$$

To accomplish the proof of Proposition 2.4, we prepare several preliminary lemmas, where the following one shows that, for fixed $R > 0$, $\mathbb{1}_{(0,\tau_R^{(n)}]}(t)|p_t^{(n)}| \to 0$ in probability as $n \to \infty$.

Lemma 2.5 *Under* $(\mathbf{H}_3)$, *for any* $R, \varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\big(|p_t^{(n)}| \geq \varepsilon, 0 < t \leq \tau_R^{(n)}\big) = 0. \tag{2.8}$$

Proof From (2.7), we obviously have for any $t \geq 0$,

$$p_t^{(n)} = -\int_{t_n}^t b_s(Y_{s_n}^{(n)})\,\mathrm{d}s - \int_{t_n}^t \int_U f_s(Y_{s_n}^{(n)}, z)\,\widetilde{N}(\mathrm{d}s, \mathrm{d}z).$$

This thus implies that for any $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}\big(|p_t^{(n)}| \geq \varepsilon, 0 < t \leq \tau_R^{(n)}\big) &\leq \mathbb{P}\bigg(\int_0^{t\wedge\tau_R^{(n)}} |b_s(Y_{s_n}^{(n)})|\mathbb{1}_{[t_n,t]}(s)\,\mathrm{d}s \geq \frac{\varepsilon}{2}\bigg) \\
&\quad + \mathbb{P}\bigg(\bigg|\int_0^{t\wedge\tau_R^{(n)}} \int_U f_s(Y_{s_n}^{(n)}, z)\mathbb{1}_{[t_n,t]}(s)\,\widetilde{N}(\mathrm{d}s, \mathrm{d}z)\bigg| \geq \frac{\varepsilon}{2}\bigg) \\
&=: \Gamma_1(t, n, R, \varepsilon) + \Gamma_2(t, n, R, \varepsilon).
\end{aligned}$$

On the one hand, via Chebyshev's inequality and the definition of $\tau_R^{(n)}$, we deduce that

$$\Gamma_1(t, n, R, \varepsilon) \leq \frac{2}{\varepsilon}\mathbb{E}\bigg(\int_0^{t\wedge\tau_R^{(n)}} |b_s(Y_{s_n}^{(n)})|\mathbb{1}_{[t_n,t]}(s)\,\mathrm{d}s\bigg) \leq \frac{2}{\varepsilon}\int_{t_n}^t \sup_{\{|x|\leq R/2\}} |b_s(x)|\,\mathrm{d}s. \tag{2.9}$$

On the other hand, using Chebyshev's inequality once more followed by Itô's isometry and the notion of $\tau_R^{(n)}$ yields that

$$\begin{aligned}\Gamma_2(t,n,R,\varepsilon) &\le \frac{4}{\varepsilon^2}\mathbb{E}\bigg|\int_0^{t\wedge\tau_R^{(n)}}\int_U f_s(Y_{s_n}^{(n)},z)\mathbb{1}_{[t_n,t]}(s)\,\widetilde{N}(\mathrm{d}s,\mathrm{d}z)\bigg|^2\\ &= \frac{4}{\varepsilon^2}\mathbb{E}\bigg(\int_0^{t\wedge\tau_R^{(n)}}\int_U |f_s(Y_{s_n}^{(n)},z)|^2\mathbb{1}_{[t_n,t]}(s)\,\nu(\mathrm{d}z)\mathrm{d}s\bigg)\\ &\le \frac{4}{\varepsilon^2}\int_{t_n}^t\int_U \sup_{\{|x|\le R/2\}}|f_s(x,z)|^2\,\nu(\mathrm{d}z)\mathrm{d}s.\end{aligned}\tag{2.10}$$

Therefore, (2.8) is valid by combining (2.9) with (2.10), and taking ($\mathbf{H}_3$) into account. □

Roughly speaking, the next lemma indicates that the life time of $(Y_t^{(n)})_{t\ge0}$ goes to infinity.

Lemma 2.6 *Assume that* ($\mathbf{H}_2$) *with* $g\equiv\mathbf{0}$ *and* ($\mathbf{H}_3$) *holds true, and suppose* $Y_0\in L^\beta(\Omega\to\mathbb{R}^d,\mathscr{F}_0,\mathbb{P})$. *Then, for any fixed* $T>0$,

$$\lim_{R\to\infty}\lim_{n\to\infty}\mathbb{P}(\tau_R^{(n)}\le T)=0.\tag{2.11}$$

Proof Apparently, we have

$$\big\{\tau_R^{(n)}\le T\big\}=\Big\{\tau_R^{(n)}\le T,\sup_{t\in[0,\tau_R^{(n)}]}|Y_t^{(n)}|\ge R/4\Big\}+\Big\{\tau_R^{(n)}\le T,\sup_{t\in[0,\tau_R^{(n)}]}|Y_t^{(n)}|<R/4\Big\}.\tag{2.12}$$

In terms of the definition of $\tau_R^{(n)}$, it is obvious that the second event on the right hand side of (2.12) is empty. Hence, the following implication and equivalence

$$\big\{\tau_R^{(n)}\le T\big\}\subseteq\Big\{\sup_{t\in[0,T\wedge\tau_R^{(n)}]}|Y_t^{(n)}|\ge R/4\Big\}=\Big\{\big|Y^{(n)}_{\tau_R^{(n),*}}\big|\ge R/4\Big\}$$

are available, where

$$\tau_R^{(n),*}:=T\wedge\tau_R^{(n)}\wedge\inf\big\{t\ge0:|Y_t^{(n)}|\ge R/4\big\}.$$

By means of Chebyshev's inequality, one has

$$\mathbb{P}\big(\tau_R^{(n)}\le T\big)\le\frac{1}{(1+(R/4)^2)^{\beta/2}}\mathbb{E}V_\beta\big(Y^{(n)}_{\tau_R^{(n),*}}\big),$$

where $V_\beta(x) := (1+|x|^2)^{\beta/2}$ for all $x \in \mathbb{R}^d$. Next, we define a function below:

$$\Psi_t(z_1,z_2) = \big(\varphi(t)|z_1| + |b_t(z_2)|\big)|z_1|, \quad t \ge 0, z_1, z_2 \in \mathbb{R}^d. \tag{2.13}$$

As long as there exists a constant $\bar{C}_T > 0$ such that

$$\mathbb{E}V_\beta\big(Y^{(n)}_{\tau_R^{(n),*}}\big) \le \bar{C}_T\bigg(\mathbb{E}V_\beta(Y_0) + \mathbb{E}\Big(\int_0^{\tau_R^{(n),*}} \Psi_s\big(p^{(n)}_{s-}, Y^{(n)}_{s_n}\big)\,\mathrm{d}s\Big)\bigg) \tag{2.14}$$

and

$$\lim_{n\to\infty} \mathbb{E}\Big(\int_0^{\tau_R^{(n),*}} \Psi_s\big(p^{(n)}_{s-}, Y^{(n)}_{s_n}\big)\,\mathrm{d}s\Big) = 0, \tag{2.15}$$

the statement (2.11) is verifiable by approaching firstly $n \to \infty$, followed by sending $R \to \infty$. Based on the preceding analysis, it amounts to verifying, respectively, (2.14) and (2.15) for the establishment of (2.11).

Applying the Itô formula, we deduce from (2.7) that

$$\begin{aligned}\mathrm{d}\Big(\mathrm{e}^{-\beta\int_0^t \varphi(s)\,\mathrm{d}s}\, V_\beta\big(Y_t^{(n)}\big)\Big) = \mathrm{e}^{-\beta\int_0^t \varphi(s)\,\mathrm{d}s}\Big(-\beta\varphi(t)V_\beta\big(Y_t^{(n)}\big) + (\mathscr{L}_t^0 V_\beta)\big(Y_t^{(n)}, Y_{t_n}^{(n)}\big)\Big)\,\mathrm{d}t \\ + \mathrm{d}M_t^{(n),\beta},\end{aligned}$$

where $(M_t^{(n),\beta})_{t\ge 0}$ is a local martingale, and for $h \in C^2(\mathbb{R}^d)$, $x, y \in \mathbb{R}^d$ and $t \ge 0$,

$$\begin{aligned}(\mathscr{L}_t^0 h)(x,y) :=& \langle \nabla h(x), b_t(y)\rangle \\ &+ \nu\big((h(x+f_t(y,\cdot)) - h(x) - \langle \nabla h(x), f_t(y,\cdot)\rangle)\mathbb{1}_U(\cdot)\big).\end{aligned} \tag{2.16}$$

For $F_\beta(r) := (1+r)^{\frac{\beta}{2}}$ on $[0,\infty)$, note readily that $F'_\beta(r) = \frac{\beta}{2}(1+r)^{\frac{\beta}{2}-1}$ and $F''_\beta(r) < 0$ for $\beta \in (0,2]$. Thus, an application of Taylor's expansion yields that for any $x, y \in \mathbb{R}^d$ and $t \ge 0$,

$$\begin{aligned}(\mathscr{L}_t^0 V_\beta)(x,y) &\le \frac{1}{2}\beta(1+|x|^2)^{\frac{\beta}{2}-1}\big(2\langle x, b_t(y)\rangle + \nu(|f_t(y,\cdot)|^2\mathbb{1}_U)\big) \\ &= \frac{1}{2}\beta(1+|x|^2)^{\frac{\beta}{2}-1}\big(2\langle y, b_t(y)\rangle + \nu(|f_t(y,\cdot)|^2\mathbb{1}_U)\big) \\ &\quad + \frac{1}{2}\beta(1+|x|^2)^{\frac{\beta}{2}-1}\langle x-y, b_t(y)\rangle \\ &\le \frac{1}{2}\beta\varphi(t)(1+|x|^2)^{\frac{\beta}{2}-1}(1+|y|^2) + |x-y|\cdot|b_t(y)| \\ &\le \beta\varphi(t)V_\beta(x) + \big(2\varphi(t)|x-y| + |b_t(y)|\big)|x-y|,\end{aligned} \tag{2.17}$$

where in the second inequality we used $(\mathbf{H}_2)$ with $g \equiv \mathbf{0}$ and $\beta \in (0, 2]$, and in the last display we utilized $\beta \in (0, 2]$ and $|y|^2 \le 2|x|^2 + 2|x-y|^2$. Subsequently, (2.17), besides $\varphi > 0$, implies that

$$\begin{aligned} \mathrm{e}^{-\beta \int_0^T \varphi(s)\,\mathrm{d}s}\, \mathbb{E} V_\beta\big(Y^{(n)}_{\tau_R^{(n),*}}\big) &\le \mathbb{E}\Big(\mathrm{e}^{-\beta \int_0^{\tau_R^{(n),*}} \varphi(s)\,\mathrm{d}s}\, V_\beta\big(Y^{(n)}_{\tau_R^{(n),*}}\big)\Big) \\ &\le \mathbb{E} V_\beta(Y_0) + 2\mathbb{E}\Big(\int_0^{\tau_R^{(n),*}} \Psi_s\big(p^{(n)}_{s-}, Y^{(n)}_{s_n}\big)\,\mathrm{d}s\Big), \end{aligned} \tag{2.18}$$

where Ψ was defined in (2.13). As a result, (2.14) is attainable by making use of the locally integrable property of φ.

We proceed to verify (2.15). Due to the definition of $\tau_R^{(n)}$, we infer that for any $t \in (0, \tau_R^{(n),*}]$,

$$\Psi_t\big(p^{(n)}_{t-}, Y^{(n)}_{t_n}\big) \le \Lambda(t, R)|p^{(n)}_{t-}| \le R\Lambda(t, R),$$

where $\Lambda(t, R) := R\varphi(t) + \sup_{|x|\le R/2} |b_t(x)|$. Hence, by virtue of $\tau_R^{(n),*} \le T \wedge \tau_R^{(n)}$, it is easy to see that for any integer $m \ge 1$,

$$\begin{aligned} \int_0^{\tau_R^{(n),*}} \Psi_s\big(p^{(n)}_{s-}, Y^{(n)}_{s_n}\big)\,\mathrm{d}s &\le \int_0^{\tau_R^{(n),*}} \Lambda(s, R)|p^{(n)}_{s-}|\,\mathrm{d}s \\ &\le m \int_0^{\tau_R^{(n)}} |p^{(n)}_{s-}|\,\mathrm{d}s + R\int_0^T \Lambda(s, R)\mathbb{1}_{\{\Lambda(s,R)>m\}}\,\mathrm{d}s. \end{aligned} \tag{2.19}$$

Next, Lemma 2.5, Fubini's theorem as well as the dominated convergence theorem yield that

$$\lim_{n\to\infty} \mathbb{E}\Big(\int_0^{\tau_R^{(n)}} |p^{(n)}_{s-}|\,\mathrm{d}s\Big) = 0.$$

Consequently, (2.15) is achievable by exploiting the locally integrable property of $t \mapsto \Lambda(t, R)$, thanks to $(\mathbf{H}_3)$, and sending $n \to \infty$, followed by $m \to \infty$. □

With Lemmas 2.5 and 2.6 at hand, we move forward to show that $(Y_t^{(n)})_{t\ge 0}$ is a Cauchy sequence in the sense of uniform convergence in probability, which is stated precisely as follows.

Lemma 2.7 *Assume that* $(\mathbf{H}_1)$, $(\mathbf{H}_2)$ *with* $g \equiv \mathbf{0}$ *and* $(\mathbf{H}_3)$ *are satisfied, and suppose that* $Y_0 \in L^\beta(\Omega \to \mathbb{R}^d, \mathscr{F}_0, \mathbb{P})$. *Then, for any* $T, \varepsilon > 0$,

$$\lim_{n,m\to\infty} \mathbb{P}\Big(\sup_{t\in[0,T]} |Y_t^{(n)} - Y_t^{(m)}| \ge \varepsilon\Big) = 0. \tag{2.20}$$

Proof Below, we set $Q_t^{n,m} := Y_t^{(n)} - Y_t^{(m)}$, $t \geq 0$, for notation brevity. It is obvious that

$$\mathbb{P}\Big(\sup_{t\in[0,T]} \big|Q_t^{n,m}\big| \geq \varepsilon\Big) \leq \mathbb{P}\Big(\sup_{t\in[0,\tau_R^{n,m}]} \big|Q_t^{n,m}\big| \geq \varepsilon\Big) + \mathbb{P}\big(\tau_R^{(n)} < T\big) + \mathbb{P}\big(\tau_R^{(m)} < T\big),$$

where $\tau_R^{n,m} := T \wedge \tau_R^{(n)} \wedge \tau_R^{(m)}$. Note that the following equivalence

$$\Big\{\sup_{t\in[0,\tau_R^{n,m}]} \big|Q_t^{n,m}\big| \geq \varepsilon\Big\} = \Big\{\big|Q_{\tau_R^{n,m,\varepsilon}}^{n,m}\big| \geq \varepsilon\Big\}$$

holds true, where

$$\tau_R^{n,m,\varepsilon} := \tau_R^{n,m} \wedge \inf\big\{t \geq 0 : \big|Q_t^{n,m}\big| \geq \varepsilon\big\}.$$

Thus, (2.20) can be obtained from the fact that

$$\lim_{R\to\infty}\lim_{n\to\infty} \mathbb{P}\big(\tau_R^{(n)} < T\big) + \lim_{R\to\infty}\lim_{m\to\infty} \mathbb{P}\big(\tau_R^{(m)} < T\big) = 0,$$

which is attainable by invoking Lemma 2.6, and provided that

$$\lim_{R\to\infty}\lim_{n,m\to\infty} \mathbb{P}\big(\big|Q_{\tau_R^{n,m,\varepsilon}}^{n,m}\big| \geq \varepsilon\big) = 0. \tag{2.21}$$

From (2.7), one obviously has for any $t \geq 0$,

$$\mathrm{d}Q_t^{n,m} = b_t^{n,m}\,\mathrm{d}t + \int_U f_t^{n,m}(z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z),$$

in which

$$b_t^{n,m} := b_t(Y_{t_n}^{(n)}) - b_t(Y_{t_m}^{(m)}) \quad \text{and} \quad f_t^{n,m}(z) := f_t(Y_{t_n}^{(n)}, z) - f_t(Y_{t_m}^{(m)}, z).$$

Applying Itô's formula followed by using (2.2) and $K_\cdot(R) > 0$ yields that for $0 < t < \tau_R^{n,m,\varepsilon}$,

$$\begin{aligned}
&\mathrm{d}\Big(\mathrm{e}^{-2\int_0^t K_s(R)\,\mathrm{d}s}\,\big|Q_t^{n,m}\big|^2\Big)\\
&\leq \mathrm{e}^{-2\int_0^t K_s(R)\,\mathrm{d}s}\Big(-2K_t(R)\big|Q_t^{n,m}\big|^2 + 2\big\langle Y_{t_n}^{(n)} - Y_{t_m}^{(m)}, b_t^{n,m}\big\rangle + \int_U |f_t^{n,m}(z)|^2\,\nu(dz)\Big)\,\mathrm{d}t\\
&\quad + 2\,\mathrm{e}^{-2\int_0^t K_s(R)\,\mathrm{d}s}\big\langle p_t^{(n)} - p_t^{(m)}, b_t(Y_{t_n}^{(n)}) - b_t(Y_{t_m}^{(m)})\big\rangle\,\mathrm{d}t + \mathrm{d}M_t^{n,m}\\
&\leq K_t(R)\,\mathrm{e}^{-2\int_0^t K_s(R)\,\mathrm{d}s}\Big(-2\big|Q_t^{n,m}\big|^2 + \big|Y_{t_n}^{(n)} - Y_{t_m}^{(m)}\big|^2\Big)\,\mathrm{d}t\\
&\quad + 2\big|p_t^{(n)} - p_t^{(m)}\big|\cdot\big|b_t(Y_{t_n}^{(n)}) - b_t(Y_{t_m}^{(m)})\big|\,\mathrm{d}t + \mathrm{d}M_t^{n,m}\\
&\leq 2K_t(R)\,\mathrm{e}^{-2\int_0^t K_s(R)\,\mathrm{d}s}\big|p_t^{(n)} - p_t^{(m)}\big|^2\,\mathrm{d}t + 2\big|p_t^{(n)} - p_t^{(m)}\big|\cdot\big|b_t(Y_{t_n}^{(n)})
\end{aligned}$$

$$
\begin{aligned}
&- b_t(Y^{(m)}_{t_m})\big|\,\mathrm{d}t + \mathrm{d}M^{n,m}_t \\
&\le \Gamma(t,R)\big(|p^{(n)}_t| + |p^{(m)}_t|\big)\,\mathrm{d}t + \mathrm{d}M^{n,m}_t,
\end{aligned}
$$

where $(M^{n,m}_t)_{t\ge 0}$ is a local martingale and $\Gamma(t,R) := 4RK_t(R)+4\sup_{|x|\le R/2}|b_t(x)|$. Subsequently, with the aid of $Y^{(n)}_0 = Y^{(m)}_0 = Y_0$ and $\tau^{n,m,\varepsilon}_R \le \tau^{(n)}_R \wedge \tau^{(m)}_R \wedge T$, we arrive at

$$
\begin{aligned}
\mathbb{E}\big|Q^{n,m}_{\tau^{n,m,\varepsilon}_R}\big|^2 \le \mathrm{e}^{2\int_0^T K_s(R)\,\mathrm{d}s}\Bigg(&\mathbb{E}\Big(\int_0^{\tau^{(n)}_R\wedge T}\Gamma(s,R)|p^{(n)}_{s-}|\,\mathrm{d}s\Big) \\
&+ \mathbb{E}\Big(\int_0^{\tau^{(m)}_R\wedge T}\Gamma(s,R)|p^{(m)}_{s-}|\,\mathrm{d}s\Big)\Bigg).
\end{aligned}
\tag{2.22}
$$

Thus, (2.21) follows from Chebyshev's inequality and by noting

$$
\lim_{n\to\infty}\mathbb{E}\Big(\int_0^{\tau^{(n)}_R\wedge T}\Gamma(s,R)|p^{(n)}_{s-}|\,\mathrm{d}s\Big) + \lim_{n\to\infty}\mathbb{E}\Big(\int_0^{\tau^{(m)}_R\wedge T}\Gamma(s,R)|p^{(m)}_{s-}|\,\mathrm{d}s\Big) = 0,
$$

which can indeed be established by repeating exactly the procedure to derive (2.15) (in particular, see the argument for (2.19)). □

Now we turn to complete the proof of Proposition 2.4.

Proof of Proposition 2.4 *Strong existence*. Since Lemma 2.7 is available, the proof on strong well-posedness of (2.5) is more or less standard; see, e.g., [33, Theorem 3.1.1] and [31, Theorem 1.5] for more details. Nevertheless, we herein provide a sketch to make the content self-contained.

Note that the space $L^\beta(\Omega, \mathscr{D}([0,T];\mathbb{R}^d))$ is complete w.r.t. locally uniform convergence in probability, where $\mathscr{D}([0,T];\mathbb{R}^d)$ stands for the set of cadlag functions $f : [0,T] \to \mathbb{R}^d$. Then, Lemma 2.7 implies that there is a cadlag, $(\mathscr{F}_t)_{t\ge 0}$-adapted process $(Y_t)_{t\ge 0}$ such that

$$
\sup_{t\in[0,T]} |Y^{(n)}_t - Y_t| \to 0 \quad \text{in probability as } n\to\infty.
$$

Therefore, a subsequence, still written as $(Y^{(n)}_t)_{t\in[0,T]}$, can be extracted so $\mathbb{P}$-a.s.,

$$
\sup_{t\in[0,T]} |Y^{(n)}_t - Y_t| \to 0 \quad \text{as } n\to\infty.
$$

This thus implies that

$$
\sup_{t\in[0,T]} S_t < \infty \quad \mathbb{P}\text{-a.s.} \quad \text{with } S_t := \sup_{n\ge 1}|Y^{(n)}_t|. \tag{2.23}
$$

Define the following stopping time: for any $R > 0$,

$$\tau_R = T \wedge \inf\big\{t \in [0, T] : S_t > R\big\}.$$

The continuity of $x \mapsto b_t(x)$, Assumption ($\mathbf{H}_3$) and the dominated convergence theorem yield that for $t \le \tau_R$,

$$\lim_{n\to\infty} \int_0^t b_s(Y_{s_n}^{(n)})\,\mathrm{d}s = \int_0^t b_s(Y_s)\,\mathrm{d}s, \quad \mathbb{P}\text{-a.s.}$$

Furthermore, applying Itô's isometry and combining the dominated convergence theorem with ($\mathbf{H}_3$) and the continuity of $x \mapsto f_t(x, z)$ enable us to derive that

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\Big|\int_0^{\tau_R}\int_U f_s(Y_{s_n}^{(n)}, z) - f_s(Y_s, z)\widetilde{N}(\mathrm{d}s, \mathrm{d}z)\Big|^2 \\
&= \lim_{n\to\infty} \mathbb{E}\Big(\int_0^{\tau_R}\int_U \big|f_s(Y_{s_n}^{(n)}, z) - f_s(Y_s, z)\big|^2 \nu(\mathrm{d}z)\mathrm{d}s\Big) = 0.
\end{aligned}$$

Then, for a subsequence, still written as $(Y_t^{(n)})_{0\le t\le T}$, we have for any $t \le \tau_R$,

$$\lim_{n\to\infty} \int_0^t \int_U f_s(Y_{s_n}^{(n)}, z)\,\widetilde{N}(\mathrm{d}s, \mathrm{d}z) = \int_0^t \int_U f_s(Y_{s-}, z)\,\widetilde{N}(\mathrm{d}s, \mathrm{d}z), \quad \mathbb{P}\text{-a.s.}$$

Based on the preceding analysis, we conclude that $(Y_t)_{t\in[0,T]}$ is a strong solution to (2.5) once we note $\lim_{R\to\infty} \tau_R = T$ by taking advantage of (2.23).

Uniqueness. For any $R, T > 0$, define $\tau_R^{Y,T} = T \wedge \tau_R^Y$ with $\tau_R^Y := \inf\{t > 0 : |Y_t| > R\}$, where $(Y_t)_{t\ge0}$ is a strong solution to (2.5). By taking $x = y$ in (2.17), we obviously obtain from ($\mathbf{H}_2$) that

$$\begin{aligned}
(\mathscr{L}_t^0 V_\beta)(x, x) &\le \frac{1}{2}\beta(1 + |x|^2)^{\frac{\beta}{2}-1}\big(2\langle x, b_t(x)\rangle + \nu(|f_t(x, \cdot)|^2 \mathbb{1}_U(\cdot))\big) \\
&\le \frac{1}{2}\beta\varphi(t)V_\beta(x),
\end{aligned} \tag{2.24}$$

where $\mathscr{L}_t^0$ was defined as in (2.16). Subsequently, by the aid of the locally integrable property of φ, we can deduce that for some constant $C(T) > 0$,

$$\mathbb{E}\big|Y_{\tau_R^{Y,T}}\big|^\beta \le C(T)(1 + \mathbb{E}|Y_0|^\beta). \tag{2.25}$$

This further implies that

$$R^\beta \mathbb{P}(\tau_R^Y < T) \le \mathbb{E}\big(\big|Y_{\tau_R^Y}\big|^\beta \mathbb{1}_{\{\tau_R^Y < T\}}\big) \le C(T)(1 + \mathbb{E}|Y_0|^\beta).$$

As a consequence, $\tau_R^{X,T} \to T$, $\mathbb{P}$-a.s., as $R \to \infty$.

Below, let $(Y_t)_{t\in[0,T]}$ and $(\tilde{Y}_t)_{t\in[0,T]}$ be two solutions to (2.5) with the same initial value, and set $\tau_R^* := \tau_R^{Y,T} \wedge \tau_R^{\tilde{Y},T}$. By following the line to derive (2.22), we deduce that for any $\varepsilon > 0$ and $t \in [0, T]$,

$$\mathbb{E}\Big(\mathrm{e}^{-\frac{1}{2}\beta \int_0^{t\wedge\tau_R^*} K_s(R)\,\mathrm{d}s} \big(\varepsilon + |Y_{t\wedge\tau_R^*} - \tilde{Y}_{t\wedge\tau_R^*}|^2\big)^{\frac{\beta}{2}} \Big) = 0$$

and so

$$\mathbb{E}\Big(\mathrm{e}^{-\frac{1}{2}\beta \int_0^{t\wedge\tau_R^*} K_s(R)\,\mathrm{d}s} \big|Y_{t\wedge\tau_R^*} - \tilde{Y}_{t\wedge\tau_R^*}\big|^\beta \Big) = 0.$$

This, combining $K_\cdot(R) > 0$ with $\tau_R^{Y,T} \to T$ and $\tau_R^{\tilde{Y},T} \to T$, $\mathbb{P}$-a.s., as $R \to \infty$, and utilizing Fatou's lemma, give that for any $t \in (0, T]$,

$$\mathbb{E}|Y_t - \tilde{Y}_t|^\beta \le \liminf_{R\to\infty} \mathbb{E}\big|Y_{t\wedge\tau_R^*} - \tilde{Y}_{t\wedge\tau_R^*}\big|^\beta = 0.$$

Whence, the uniqueness of strong solution follows.

Moment estimate. The establishment of (2.6) can be done by examining the procedure to derive (2.25), so we herein omit the corresponding details. □

With the help of Proposition 2.4, we move to apply the so-called interlacing technique (see, e.g., [3, p. 112–113]) to construct the unique solution to the SDE (2.1) and thus verify Theorem 2.1.

Proof of Theorem 2.1 Let for any $t \ge 0$,

$$Z_t^V = \int_0^t \int_V z N(\mathrm{d}s, \mathrm{d}z) \quad \text{and} \quad D_p^V = \{t \ge 0 : Z_t^V \ne Z_{t-}^V, \triangle Z_t \in V\}, \tag{2.26}$$

where $\triangle Z_t := Z_t - Z_{t-}$, the increment of $Z_\cdot$ at time t. Note that D_p^V is a countable set so that it can be rewritten as $D_p^V = \{\sigma_1, \cdots, \sigma_n, \cdots\}$, where, $n \mapsto \sigma_n$ is increasing, and, for each fixed $n \ge 1$, σ_n is a finite stopping time satisfying $\lim_{n\to\infty} \sigma_n = \infty$ a.s. by taking $\nu(\mathbb{1}_V) < \infty$ into consideration. Let $p_n = \triangle Z_{\sigma_n}, n \ge 1$, i.e., the jump amplitude at the jumping time σ_n. Then, $(p_n)_{n\ge1}$ is an i.i.d sequence of random variables with the common distribution $\nu|_V/\nu(\mathbb{1}_V)$ and independent of $(\sigma_n)_{n\ge1}$. Now, we set

$$X_t^{(1)} = \begin{cases} Y_t, & 0 \le t < \sigma_1, \\ Y_{\sigma_1} + g_{\sigma_1-}(Y_{\sigma_1-}, p_1), & t = \sigma_1, \end{cases}$$

where $(Y_t)_{t\ge0}$ solves (2.5). Obviously, the process $(X_t^{(1)})_{t\in[0,\sigma_1]}$ is the unique solution to (2.1) on $[0, \sigma_1]$. Next, we set

$$X_t^{(2)} = \begin{cases} X_t^{(1)}, & 0 \le t \le \sigma_1, \\ Y_t + g_{\sigma_1-}(Y_{\sigma_1-}, p_1), & \sigma_1 < t < \sigma_2, \\ Y_{\sigma_2-} + \sum_{i=1}^2 g_{\sigma_i-}(Y_{\sigma_i-}, p_i), & t = \sigma_2, \end{cases}$$

which is the unique solution to (2.1) on $[0, \sigma_2]$. Continuing successively the previous procedure, the global solution $(X_t)_{t\geq 0}$ to (2.1) can be determined uniquely. In particular, $(X_t)_{t\geq 0}$ can be written as follows:

$$X_t = Y_t + \sum_{i=1}^{\infty} \big(g_{\sigma_i -}(Y_{\sigma_i -}, p_i)\mathbb{1}_{\{\sigma_i \leq t\}}\big) = Y_t + \int_0^t \int_V g_s(Y_{s-}, z\, N(\mathrm{d}s, \mathrm{d}z).$$

Below, we proceed to prove the statement (2.3). In retrospect, $V_\beta(x) = (1+|x|^2)^{\frac{\beta}{2}}$ for all $x \in \mathbb{R}^d$. Applying Itô's formula yields that

$$\begin{aligned}
&\mathrm{d}\Big(\mathrm{e}^{-\int_0^t (\frac{\beta}{2}\varphi(s)+\nu(\mathbb{1}_V))\,\mathrm{d}s}\, V_\beta(X_t)\Big)\\
&= \mathrm{e}^{-\int_0^t (\frac{\beta}{2}\varphi(s)+\nu(\mathbb{1}_V))\,\mathrm{d}s}\Big(-\big(\beta\varphi(t)/2+\nu(\mathbb{1}_V)\big)V_\beta(X_t) + (\mathscr{L}_t V_\beta)(X_t)\Big)\,\mathrm{d}t + \mathrm{d}M_t,
\end{aligned}$$

where $(M_t)_{t\geq 0}$ is a martingale, and for $h \in C^2(\mathbb{R}^d)$

$$(\mathscr{L}_t h)(x) := (\mathscr{L}_t^0 h)(x, x) + \nu\big((h(x + g_t(x, \cdot)) - h(x))\mathbb{1}_V(\cdot)\big)$$

with $\mathscr{L}_t^0$ being given by (2.16). Furthermore, by invoking the inequality: $(a+b)^\theta \leq a^\theta + b^\theta$, $a, b > 0$ and $\theta \in (0, 1]$, it follows that for any $x, y \in \mathbb{R}^d$,

$$V_\beta(x+y) - V_\beta(x) \leq \big||y|^2 + 2\langle x, y\rangle\big|^{\frac{\beta}{2}} \leq 2^{\beta/2}|y|^\beta + |x|^\beta. \tag{2.27}$$

This, together with (2.24) and $(\mathbf{H}_2)$, leads to

$$\begin{aligned}
(\mathscr{L}_t V_\beta)(x) &\leq \frac{\beta}{2}(1+|x|^2)^{\beta/2-1}\Big(2\langle x, b_t(x)\rangle + \nu(|f_t(x,\cdot)|^2\mathbb{1}_U(\cdot))\\
&\qquad + \frac{2^{\beta/2+1}}{\beta}(1+|x|^2)^{1-\beta/2}\nu(|g_t(x,\cdot)|^\beta\mathbb{1}_V(\cdot))\Big) + |x|^\beta\,\nu(\mathbb{1}_V)\\
&\leq (\beta\varphi(t)/2 + \nu(\mathbb{1}_V))V_\beta(x).
\end{aligned}$$

Consequently, the desired assertion (2.3) follows. □

Before the end of this section, we finish the proof of Theorem 2.2, which is concerned with a stronger moment estimate.

Proof of Theorem 2.2 The proof is inspired essentially by that of [7, Theorem 1] but with some essential modifications. According to Theorem 2.1, under Assumptions $(\mathbf{H}_1)$, $(\mathbf{H}_2')$ and $(\mathbf{H}_3)$, the SDE (2.1) has a unique strong solution.

Recall that $D_p^V = \{\sigma_1, \cdots, \sigma_n, \cdots\}$, and $(p_n)_{n\geq 1}$ is the Poisson point process associated with the Lévy process $(Z_t^V)_{t\geq 0}$ given in (2.26). More details on $(\sigma_n)_{n\geq 1}$

and $(p_n)_{n\geq 1}$ are presented in the beginning part of the proof of Theorem 2.1. Obviously, for any $T>0$, we have

$$\begin{aligned}\mathbb{E}\Big(\sup_{0\le t\le T}|X_t|^\beta\Big)&\le\mathbb{E}\Big(\mathbb{1}_{\{0\le T<\sigma_1\}}\sup_{0\le t\le T}|X_t|^\beta\Big)+\sum_{n=1}^\infty\mathbb{E}\Big(\mathbb{1}_{\{\sigma_n\le T<\sigma_{n+1}\}}\sum_{k=0}^{n-1}\sup_{\sigma_k\le t<\sigma_{k+1}}|X_t|^\beta\Big)\\&\quad+\sum_{n=1}^\infty\mathbb{E}\Big(\mathbb{1}_{\{\sigma_n\le T<\sigma_{n+1}\}}\sup_{\sigma_n\le t\le T}|X_t|^\beta\Big)\\&=:\Gamma_1+\Gamma_2+\Gamma_3.\end{aligned}$$

Therefore, to achieve (2.4), it is sufficient to show that for some constant $C_T^*>0$,

$$\Gamma_i\le C_T^*(1+\mathbb{E}|X_0|^\beta),\quad i=1,2,3. \tag{2.28}$$

For the validity of (2.28), we firstly verify that there exists a constant $C_T^{**}>0$ such that for any $0\le k\le(n-1)^+$ and $\sigma_n\le T$,

$$\mathbb{E}\Big(\sup_{\sigma_k\le s<\sigma_{k+1}}|X_s|^\beta\big|\mathscr{G}_{n,k}\Big)\le C_T^{**}\big(1+|X_{\sigma_k}|^\beta\big), \tag{2.29}$$

where $\sigma_0=0$, and

$$\mathscr{G}_{n,k}:=\sigma(\sigma_i,1\le i\le n+1)\vee\mathscr{F}_{T\wedge\sigma_k},\quad n,k\ge 0.$$

By Itô's formula, it follows from $(\mathbf{H}_2')$ that for any $t\in[\sigma_k,\sigma_{k+1})$ with $0\le k\le(n-1)^+$,

$$\begin{aligned}\mathbb{E}\Big(\sup_{\sigma_k\le s\le t}|X_s|^2\big|\mathscr{G}_{n,k}\Big)&\le|X_{\sigma_k}|^2+3\int_{\sigma_k}^t\phi(s)\big(1+\mathbb{E}(|X_s|^2|\mathscr{G}_{n,k})\big)\,\mathrm{d}s\\&\quad+\mathbb{E}\Big(\sup_{\sigma_k\le s\le t}|M_{k,s}|\big|\mathscr{G}_{n,k}\Big),\end{aligned} \tag{2.30}$$

where

$$M_{k,t}:=\int_{\sigma_k^V}^t\int_U\big(2\langle X_{s-},f_s(X_{s-},z)\rangle+|f_s(X_{s-},z)|^2\big)\widetilde{N}(\mathrm{d}s,\mathrm{d}z).$$

Next, by applying BDG's inequality (see, e.g., [30, Theorem 1]), in addition to $(\mathbf{H}_2')$, there exist constants $c_1,c_2>0$ such that for any for any $t\in[\sigma_k,\sigma_{k+1})$ with $0\le k\le(n-1)^+$,

$$\begin{aligned}\mathbb{E}\Big(\sup_{\sigma_k\le s\le t}|M_{k,s}|\big|\mathscr{G}_{n,k}\Big)&\le c_1\mathbb{E}\bigg(\Big(\int_{\sigma_k}^t\int_U|X_s|^2|f_s(X_{s-},z)|^2\,\nu(\mathrm{d}z)\,\mathrm{d}s\Big)^{1/2}\big|\mathscr{G}_{n,k}\bigg)\\&\quad+c_1\int_{\sigma_k}^t\int_U\mathbb{E}\big(|f_s(X_{s-},z)|^2\big|\mathscr{G}_{n,k}\big)\,\nu(\mathrm{d}z)\,\mathrm{d}s\end{aligned}$$

$$\leq \frac{1}{2}\mathbb{E}\Big(\sup_{\sigma_k \leq s \leq t} |X_s|^2 \big| \mathscr{G}_{n,k}\Big) + c_2 \int_{\sigma_k}^t \phi(s)\big(1 + \mathbb{E}(|X_s|^2 | \mathscr{G}_{n,k})\big)\,\mathrm{d}s.$$

Subsequently, plugging the estimate above back into (2.30) followed by Gronwall's inequality yields that there exists a constant $c_3 > 0$ such that for any $0 \leq k \leq (n-1)^+$ and $\sigma_n \leq T$,

$$\mathbb{E}\Big(\sup_{\sigma_k \leq s < \sigma_{k+1}} |X_s|^2 \big| \mathscr{G}_{n,k}\Big) \leq c_3\, \mathrm{e}^{c_3 \int_0^T \phi(u)\mathrm{d}u}\, \big(1 + |X_{\sigma_k}|^2\big).$$

This, combining with Jensen's inequality, implies (2.29).

Let $N_t = \sharp\{s \in [0,t] : Z_s^V \neq Z_{s-}^V\}$, which is a Poisson process with the intensity $\nu(\mathbb{1}_V)t$ (so $\mathbb{E}N_T = \nu(\mathbb{1}_V)T$). Making use of the definition of $\mathscr{G}_{n,k}$ and (2.29) gives us that

$$\begin{aligned}
\Gamma_2 &= \sum_{n=1}^{\infty} \mathbb{E}\Big(\mathbb{1}_{\{\sigma_n \leq T < \sigma_{n+1}\}} \sum_{k=0}^{n-1} \mathbb{E}\Big(\sup_{\sigma_k \leq t < \sigma_{k+1}} |X_t|^\beta \big| \mathscr{G}_{n,k}\Big)\Big) \\
&\leq C_T^{**} \sum_{n=1}^{\infty} \mathbb{E}\Big(\mathbb{1}_{\{\sigma_n \leq T < \sigma_{n+1}\}} \sum_{k=0}^{n-1} \big(1 + |X_{\sigma_k}|^\beta\big)\Big) \\
&\leq C_T^{**}\Big(1 + \sup_{0 \leq t \leq T} \mathbb{E}|X_t|^\beta\Big) \sum_{n=1}^{\infty} \mathbb{P}(N_T = n)n \\
&\leq \nu(\mathbb{1}_V) T C_T^{**}\Big(1 + \sup_{0 \leq t \leq T} \mathbb{E}|X_t|^\beta\Big),
\end{aligned}$$

where in the second identity we exploited the fact that $\mathbb{1}_{\{\sigma_n \leq T < \sigma_{n+1}\}}$ is independent of $\sum_{k=0}^{n-1}(1 + |X_{\sigma_k}|^\beta)$. Hence, (2.28) with $i = 2$ is available by the aid of (2.3).

On the other hand, by following exactly the strategy to derive (2.29), we can obtain that for $T \in [\sigma_n, \sigma_{n+1})$,

$$\mathbb{E}\Big(\sup_{\sigma_n \leq t \leq T} |X_t|^\beta \big| \mathscr{G}_{n,n}\Big) \leq C_T^{**}\big(1 + |X_{\sigma_n}|^\beta\big).$$

In particular, it is ready to see that

$$\Gamma_1 \leq \mathbb{E}\Big(\mathbb{E}\Big(\sup_{0 \leq t \leq T} |X_t|^\beta \big| \mathscr{G}_{0,0}\Big)\Big) \leq C_T^{**}(1 + \mathbb{E}|X_0|^\beta)$$

so that (2.28) holds true for $i = 1$. Furthermore, we can obtain that

$$\begin{aligned}
\Gamma_3 &= \sum_{n=1}^{\infty} \mathbb{E}\Big(\mathbb{1}_{\{\sigma_n \leq T < \sigma_{n+1}\}} \mathbb{E}\Big(\sup_{\sigma_n \leq t \leq T} |X_t|^\beta \big| \mathscr{G}_{n,n}\Big)\Big) \\
&\leq C_T^{**} \sum_{n=1}^{\infty} \mathbb{E}\Big(\mathbb{1}_{\{\sigma_n \leq T < \sigma_{n+1}\}} \big(1 + |X_{\sigma_n}|^\beta\big)\Big)
\end{aligned}$$

$$
\begin{aligned}
&\leq 2C_T^{**}\sum_{n=1}^{\infty}\big(\mathbb{P}(N_T=n)\big(1+\mathbb{E}|X_{T\wedge\sigma_n-}|^{\beta}+\mathbb{E}|g_{T\wedge\sigma_n-}(X_{T\wedge\sigma_n-},p_n)|^{\beta}\big)\big)\\
&\leq 2C_T^{**}\big(1+\phi(T)^{\frac{\beta}{2}}\big)\sum_{n=1}^{\infty}\big(\mathbb{P}(N_T=n)\big(1+\mathbb{E}|X_{T\wedge\sigma_n-}|^{\beta}\big)\big)\\
&\leq 2C_T^{**}\big(1+\phi(T)^{\frac{\beta}{2}}\big)\Big(1+\sup_{0\leq t\leq T}\mathbb{E}|X_t|^{\beta}\Big),
\end{aligned}
$$

where in the second inequality we employed that $X_{\sigma_n}=X_{\sigma_n-}+g_{\sigma_n-}(X_{\sigma_n-},p_n)$ and the strong Markov property, and in the third inequality we took advantage of $(\mathbf{H}_2')$ and the nondecreasing property of ϕ as well as the fact that $(a+b)^{\beta/2}\leq a^{\beta/2}+b^{\beta/2}$ for all $a,b\geq 0$. Finally, (2.28) with $i=3$ is verifiable on the basis of (2.3). Therefore, the proof is finished. □

3 Proofs of Theorems 1.1, 1.3 and Remark 1.2 (ii)

3.1 Proof of Theorem 1.1

Based on the warm-up Theorem 2.1, along with the fixed point theorem, in the following part we aim to complete the proof of Theorem 1.1.

Proof of Theorem 1.1 For a fixed horizon $T>0$, define the following path space:

$$
\mathscr{C}_T^{X_0}=\Big\{\mu\in C([0,T];\mathscr{P}_\beta(\mathbb{R}^d)):\sup_{t\in[0,T]}\mu_t(|\cdot|^{\beta})<\infty,\mu_0=\mathscr{L}_{X_0}\Big\},
$$

where $X_0\in L^{\beta}(\Omega\to\mathbb{R}^d,\mathscr{F}_0,\mathbb{P})$ is the initial value of $(X_t)_{t\geq0}$, and

$$
C([0,T];\mathscr{P}_\beta(\mathbb{R}^d)):=\{\mu:[0,T]\to\mathscr{P}_\beta(\mathbb{R}^d)\text{ is weakly continuous}\}.
$$

For $\gamma>0$, $(\mathscr{C}_T^{X_0},\mathbb{W}_{\beta,\gamma})$ is a complete metric space, where

$$
\mathbb{W}_{\beta,\gamma}(\mu,\widetilde{\mu}):=\sup_{0\leq t\leq T}\big(\mathrm{e}^{-\gamma t}\,\mathbb{W}_\beta(\mu_t,\widetilde{\mu}_t)\big),\quad \mu,\widetilde{\mu}\in\mathscr{C}_T^{X_0}.
$$

In the sequel, we work with the decoupled SDE associated with (1.1): for any $\mu\in\mathscr{C}_T^{X_0}$,

$$
\mathrm{d}X_t^{\mu}=b(X_t^{\mu},\mu_t)\,\mathrm{d}t+\int_U f(X_{t-}^{\mu},\mu_t,z)\,\widetilde{N}(\mathrm{d}t,\mathrm{d}z)+\int_V g(X_{t-}^{\mu},\mu_t,z)\,N(\mathrm{d}t,\mathrm{d}z) \tag{3.1}
$$

with the initial value $X_0^{\mu}=X_0$. By setting for any $t\geq0$, $x,z\in\mathbb{R}^d$ and $\mu\in\mathscr{C}_T^{X_0}$,

$$
b_t^{\mu}(x):=b(x,\mu_t),\quad f_t^{\mu}(x,z):=f(x,\mu_t,z)\quad\text{and}\quad g_t^{\mu}(x,z):=g(x,\mu_t,z),
$$

the SDE (3.1) can be reformulated so it fits into the framework (2.1). From (1.3), it is easy to see that for any $t \geq 0, x, y \in \mathbb{R}^d$ and $\mu \in \mathscr{C}_T^{X_0}$,

$$2\langle b_t^\mu(x) - b_t^\mu(y), x - y\rangle + \nu(|f_t^\mu(x,\cdot) - f_t^\mu(y,\cdot)|^2 \mathbb{1}_U(\cdot)) \leq L_1|x-y|^2$$

so that (2.2) holds true with b_t and f_t being replaced by b_t^μ and f_t^μ, respectively. In addition, by virtue of (1.4) and ($\mathbf{A}_2$), it follows that for some constants $c_1, c_2, c_3 > 0$,

$$\begin{aligned} &2\langle x, b_t^\mu(x)\rangle + \nu(|f_t^\mu(x,\cdot)|^2\mathbb{1}_U(\cdot)) + 2^{\frac{\beta}{2}+1}\beta^{-1}(1+|x|^2)^{1-\frac{\beta}{2}}\nu(|g_t^\mu(x,\cdot)|^\beta \mathbb{1}_V(\cdot)) \\ &\leq L_2(1+|x|^2+\mu_t(|\cdot|^\beta)^{\frac{2}{\beta}}) \\ &\quad + c_1(1+|x|^2)^{1-\frac{\beta}{2}}\big(\nu((1+|\cdot|^\beta)\mathbb{1}_V(\cdot))(|x|^\beta+\mu_t(|\cdot|^\beta)) + \nu(|g(\mathbf{0},\delta_{\mathbf{0}},\cdot)|^\beta\mathbb{1}_V(\cdot))\big) \\ &\leq c_2(1+|x|^2+\mu_t(|\cdot|^\beta)^{\frac{2}{\beta}}) \\ &\leq c_3\Big(1+\sup_{0\leq s\leq t}\mu_s(|\cdot|^\beta)^{\frac{2}{\beta}}\Big)(1+|x|^2), \quad 0\leq t\leq T, \end{aligned} \tag{3.2}$$

where we used (1.2) in the first inequality, and the second inequality is valid due to Young's inequality. Therefore, ($\mathbf{H}_2$) holds true thanks to $\mu \in \mathscr{C}_T^{X_0}$. Furthermore, ($\mathbf{A}_3$) implies ($\mathbf{H}_3$) directly. Consequently, according to Theorem 2.1, (3.1) has the unique strong solution $(X_t^\mu)_{t\geq 0}$ under ($\mathbf{A}_1$)–($\mathbf{A}_3$), along with (1.2).

Now, we define a mapping $\mathscr{C}_T^{X_0} \ni \mu \mapsto \Phi(\mu)$ by

$$(\Phi(\mu))_t = \mathscr{L}_{X_t^\mu}, \quad t\in[0,T]. \tag{3.3}$$

In the sequel, we shall claim, respectively, that (i) $\Phi : \mathscr{C}_T^{X_0} \to \mathscr{C}_T^{X_0}$, and (ii) Φ is contractive under $\mathbb{W}_{\beta,\gamma}$ for some $\gamma > 0$ large enough. Once (i) and (ii) are available, the classical Banach fixed point theorem yields that the map Φ defined in (3.3) has a unique fixed point, still written as μ, so $(\Phi(\mu))_t = \mu_t = \mathscr{L}_{X_t^\mu}$ for any $t\in[0,T]$. Thus, the measure variable $(\mu_t)_{t\in[0,T]}$ in (3.1) can be replaced by $(\mathscr{L}_{X_t^\mu})_{t\in[0,T]}$. Accordingly, (1.1) has a unique strong solution.

For the validity of (i), we need to show that there is a constant $C_{T,\mu} > 0$ such that

$$\sup_{0\leq t\leq T}\mathbb{E}|X_t^\mu|^\beta \leq C_{T,\mu}(1+\mathbb{E}|X_0|^\beta), \quad \mu\in\mathscr{C}_T^{X_0}, T>0. \tag{3.4}$$

and that

$$\mathscr{L}_{X_t^\mu} \xrightarrow{w} \mathscr{L}_{X_0^\mu} \quad \text{as } t\to 0. \tag{3.5}$$

Indeed, (3.4) follows from (2.3) and (3.2). For any $R > 0$ and $h \in \mathrm{Lip}_b(\mathbb{R}^d)$ (i.e., the set of bounded Lipschitz functions on $\mathbb{R}^d$), one obviously has for any $t\in[0,T]$

$$\mathbb{E}|h(X_t^\mu) - h(X_0^\mu)| \leq \|h\|_{\mathrm{Lip}}\mathbb{E}|X_{t\wedge\tau_R}^\mu$$

$$- X_0^\mu| + 2\|h\|_\infty \mathbb{P}(\tau_R \le T) =: I_1(t,R) + I_2(T,R),$$

where $\tau_R := \{t > 0 : |X_t^\mu| > R\}$ and $\|h\|_{\rm Lip}$ is the Lipschitz constant of h. Based on this, to verify (3.5), it suffices to show that $\lim_{t\to 0} I_1(t,R) = 0$ for fixed R and $\lim_{R\to\infty} I_2(T,R) = 0$, respectively. By Itô's formula, in addition to the local boundedness (see ($\mathbf{A}_1$) and ($\mathbf{A}_3$) for more details), it follows that for any $\mu \in \mathscr{C}_T^{X_0}$, $\varepsilon > 0$, and $t \in [0,T]$,

$$\begin{aligned}
&\mathbb{E}(\varepsilon + |X^\mu_{t\wedge\tau_R} - X_0^\mu|^2)^{\frac{\beta}{2}}\\
&\le \varepsilon^{\frac{\beta}{2}} + \mathbb{E}\bigg(\int_0^{t\wedge\tau_R}\Big[\frac12\beta(\varepsilon + |X_s^\mu - X_0^\mu|^2)^{\frac{\beta}{2}-1}\\
&\qquad \times\Big(2\langle X_s^\mu - X_0^\mu, b(X_s^\mu,\mu_s)\rangle + \nu(|f(X_s^\mu,\mu_s,\cdot)|^2\mathbb{1}_U(\cdot))\\
&\qquad\quad + 2^{\frac{\beta}{2}+1}\beta^{-1}(\varepsilon + |X_s^\mu - X_0^\mu|^2)^{1-\frac{\beta}{2}}\nu(|g(X_s^\mu,\mu_s,\cdot)|^\beta\mathbb{1}_V(\cdot))\Big)\Big]\,\mathrm{d}s\bigg)\\
&\quad + \nu(V)\mathbb{E}\bigg(\int_0^{t\wedge\tau_R}|X^\mu_{s\wedge\tau_R} - X_0^\mu|^\beta\,\mathrm{d}s\bigg)\\
&\le \varepsilon^{\frac{\beta}{2}} + C_{\varepsilon,R}t,
\end{aligned}$$

where $C_{\varepsilon,R}$ is a positive constant depending on the parameters ε, R. See the arguments in the end of the proof for Theorem 2.1. The preceding estimate, besides Jensen's inequality for $\beta \in [1,2]$, implies that for any $\varepsilon, R > 0$,

$$\mathbb{E}|X^\mu_{t\wedge\tau_R} - X_0^\mu| \le \mathbb{E}(\varepsilon + |X^\mu_{t\wedge\tau_R} - X_0^\mu|^2)^{\frac12} \le (\varepsilon^{\frac{\beta}{2}} + C_{\varepsilon,R}t)^{\frac{1}{\beta}}.$$

This apparently leads to $\lim_{t\to 0} I_1(t,R) = 0$. Furthermore, $\lim_{R\to\infty} I_2(T,R) = 0$ can be handled by following the argument to derive (2.25) and using (3.2). Indeed, we have

$$\mathbb{P}(\tau_R \le T) \le \frac{\mathbb{E}\big(1 + \big|X^\mu_{T\wedge\tau_R}\big|^2\big)^{\beta/2}}{(1+R^2)^{\beta/2}}.$$

Then, $\lim_{R\to\infty} I_2(T,R) = 0$ is available by taking (3.4) into consideration.

We turn to show (ii). Recall that $(Z_t^V)_{t\ge 0}$ is defined as in (2.26). Furthermore, $(p_n)_{n\ge 1}$ (i.e., the sequence concerning jumping amplitude of $(Z_t^V)_{t\ge0}$) are i.i.d random variables with the common distribution $\nu|_V/\nu(V)$ and independent of $(\sigma_n)_{n\ge1}$ (i.e., the sequence of jumping time of $(Z_t^V)_{t\ge0}$); see the first paragraph of the proof of Theorem 2.1 for further details. Note obviously that

$$\begin{aligned}
\mathbb{W}_{\beta,\gamma}(\Phi(\mu),\Phi(\tilde\mu))^\beta \le &\sup_{0\le t\le T}\Big(\mathrm{e}^{-\beta\gamma t}\,\mathbb{E}\big(\mathbb{1}_{\{0\le t<\sigma_1\}}\mathbb{E}\big(|\Upsilon_t^{\mu,\tilde\mu}|^\beta\big|\mathscr{G}_0\big)\big)\Big)\\
&+\sup_{0\le t\le T}\Big(\mathrm{e}^{-\beta\gamma t}\sum_{n=1}^\infty\mathbb{E}\big(\mathbb{1}_{\{\sigma_n\le t<\sigma_{n+1}\}}\mathbb{E}\big(|\Upsilon_t^{\mu,\tilde\mu}|^\beta\big|\mathscr{G}_n\big)\big)\Big)
\end{aligned}$$

$$= \Gamma_1(\mu, \tilde{\mu}) + \Gamma_2(\mu, \tilde{\mu}),$$

where

$$\Upsilon_t^{\mu,\tilde{\mu}} := X_t^{\mu} - X_t^{\tilde{\mu}} \quad \text{and} \quad \mathscr{G}_n := \sigma\{\sigma_i, 1 \le i \le n+1\} \vee \mathscr{F}_{\sigma_n}.$$

Thus, to obtain (ii), it remains to show that there are constants $C_1(T), C_2(T) > 0$ such that

$$\Gamma_1(\mu, \tilde{\mu}) \le C_1(T)/\gamma^{\frac{\beta}{2}} \mathbb{W}_{\beta,\gamma}(\mu, \tilde{\mu})^{\beta}, \tag{3.6}$$

and

$$\Gamma_2(\mu, \tilde{\mu}) \le C_2(T)\big(1/\gamma + 1/\gamma^{\frac{\beta}{2}}\big)\mathbb{W}_{\beta,\gamma}(\mu, \tilde{\mu})^{\beta}. \tag{3.7}$$

Let $\sigma_0 = \sigma_{0-} = 0$. For the establishments of (3.6) and (3.7), we start to verify that for any $n \ge 0$ and $t \in [\sigma_n, \sigma_{n+1})$,

$$\mathbb{E}(|\Upsilon_t^{\mu,\tilde{\mu}}|^{\beta}|\mathscr{G}_n) \le \bigg(|\Upsilon_{\sigma_n}^{\mu,\tilde{\mu}}|^{\beta} + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_n}^t \mathbb{W}_{\beta}(\mu_s, \tilde{\mu}_s)^2 \mathrm{d}s\Big)^{\frac{\beta}{2}}\bigg)\,\mathrm{e}^{L_1(t-\sigma_n)}. \tag{3.8}$$

Indeed, applying Itô's formula, in addition to (1.3), we obtain that for any $t \in [\sigma_n, \sigma_{n+1})$,

$$\mathrm{d}|\Upsilon_t^{\mu,\tilde{\mu}}|^2 \le L_1\big(|\Upsilon_t^{\mu,\tilde{\mu}}|^2 + \mathbb{W}_{\beta}(\mu_t, \tilde{\mu}_t)^2\big)\,\mathrm{d}t + \mathrm{d}\bar{M}_t,$$

where $\bar{M}_\cdot$ is a martingale. Thus, Gronwall's inequality yields that for any $t \in [\sigma_n, \sigma_{n+1})$,

$$\mathbb{E}(|\Upsilon_t^{\mu,\tilde{\mu}}|^2|\mathscr{G}_n) \le \Big(|\Upsilon_{\sigma_n}^{\mu,\tilde{\mu}}|^2 + L_1 \int_{\sigma_n}^t \mathbb{W}_{\beta}(\mu_s, \tilde{\mu}_s)^2 \mathrm{d}s\Big)\,\mathrm{e}^{L_1(t-\sigma_n)}.$$

Whence, (3.8) follows from Jensen's inequality.

By virtue of (3.8), in addition to $X_0^{\mu} = X_0^{\tilde{\mu}}$, we have

$$\begin{aligned}\Gamma_1(\mu, \tilde{\mu}) &\le L_1^{\frac{\beta}{2}}\,\mathrm{e}^{L_1 T} \sup_{0\le t\le T}\bigg(\mathrm{e}^{-\beta\gamma t}\Big(\int_0^t \mathbb{W}_{\beta}(\mu_s, \tilde{\mu}_s)^2 \mathrm{d}s\Big)^{\frac{\beta}{2}}\bigg)\\ &\le (L_1/(2\gamma))^{\frac{\beta}{2}}\,\mathrm{e}^{L_1 T}\,\mathbb{W}_{\beta,\gamma}(\mu, \tilde{\mu})^{\beta}.\end{aligned}$$

Hence, (3.6) follows right now. Again, by virtue of (3.8), we find that

$$\begin{aligned}\Gamma_2(\mu, \tilde{\mu}) &\le \mathrm{e}^{L_1 T} \sup_{0\le t\le T}\bigg(\mathrm{e}^{-\beta\gamma t}\sum_{n=1}^{\infty}\mathbb{E}\Big(\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\Big(|\Upsilon_{\sigma_n}^{\mu,\tilde{\mu}}|^{\beta} + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_n}^t \mathbb{W}_{\beta}(\mu_s, \tilde{\mu}_s)^2 \mathrm{d}s\Big)^{\frac{\beta}{2}}\Big)\\ &\le \mathrm{e}^{L_1 T} \sup_{0\le t\le T}\bigg(\mathrm{e}^{-\beta\gamma t}\sum_{n=1}^{\infty}\mathbb{E}\big(\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)|\Upsilon_{\sigma_n}^{\mu,\tilde{\mu}}|^{\beta}\big)\bigg) + (L_1/(2\gamma))^{\frac{\beta}{2}}\,\mathrm{e}^{L_1 T}\,\mathbb{W}_{\beta,\gamma}(\mu, \tilde{\mu})^{\beta}\end{aligned}$$

$$=: \mathrm{e}^{L_1 T}\,\hat{\Gamma}_2(\mu,\tilde{\mu}) + (L_1/(2\gamma))^{\frac{\beta}{2}}\,\mathrm{e}^{L_1 T}\,\mathbb{W}_{\beta,\gamma}(\mu,\tilde{\mu})^{\beta}.$$

Next, because of

$$|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n}| \le |\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}| + |g(X^{\mu}_{\sigma_n-},\mu_{\sigma_n},p_n) - g(X^{\tilde{\mu}}_{\sigma_n-},\tilde{\mu}_{\sigma_n},p_n)|, \tag{3.9}$$

it follows from (1.4) that

$$\begin{aligned}|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n}|^{\beta} &\le 2|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}|^{\beta} + 8L_1^{\beta}\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}|^{\beta} + \mathbb{W}_{\beta}(\mu_{\sigma_n},\tilde{\mu}_{\sigma_n})^{\beta}\big)(1+|p_n|^{\beta})\\ &\le \xi_n\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}|^{\beta} + \mathbb{W}_{\beta}(\mu_{\sigma_n},\tilde{\mu}_{\sigma_n})^{\beta}\big),\end{aligned} \tag{3.10}$$

where $\xi_n := 2 + 8L_1^{\beta}(1+|p_n|^{\beta}))$, and $\sigma_0 = \sigma_{0-} = 0$. Due to the stationarity of $(p_n)_{t\ge 0}$, one has $c_* := 2+8L_1^{\beta}(1+\mathbb{E}|p_n|^{\beta})$ (which is independent of n). Furthermore, applying (3.8) and (3.10) repeatedly yields that

$$\begin{aligned}\mathbb{E}\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}|^{\beta}\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\big) &= \mathbb{E}|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}|^{\beta}\mathbb{E}\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\\ &\le \mathrm{e}^{L_1 T}\,\mathbb{E}\bigg(\Big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_{n-1}}|^{\beta} + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_{n-1}}^{\sigma_n}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\Big)\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\bigg)\\ &\le \mathrm{e}^{L_1 T}\,\mathbb{E}\bigg(\Big(\xi_{n-1}\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_{n-1}-}|^{\beta} + \mathbb{W}_{\beta}(\mu_{\sigma_{n-1}},\tilde{\mu}_{\sigma_{n-1}})^{\beta}\big)\\ &\qquad + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_{n-1}}^{\sigma_n}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\Big)\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\bigg)\\ &= \mathrm{e}^{L_1 T}\,\mathbb{E}\bigg(\Big(c_*\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_{n-1}-}|^{\beta} + \mathbb{W}_{\beta}(\mu_{\sigma_{n-1}},\tilde{\mu}_{\sigma_{n-1}})^{\beta}\big)\\ &\qquad + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_{n-1}}^{\sigma_n}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\Big)\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\bigg)\\ &\le \mathrm{e}^{L_1 T}\,\mathbb{E}\bigg[\bigg(c_*\,\mathrm{e}^{L_1 T}\,\Big(c_*\big(|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_{n-2}-}|^{\beta} + \mathbb{W}_{\beta}(\mu_{\sigma_{n-2}},\tilde{\mu}_{\sigma_{n-2}})^{\beta}\big)\\ &\qquad + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_{n-2}}^{\sigma_{n-1}}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\Big)\\ &\qquad + c_*\mathbb{W}_{\beta}(\mu_{\sigma_{n-1}},\tilde{\mu}_{\sigma_{n-1}})^{\beta}\\ &\qquad + L_1^{\frac{\beta}{2}}\Big(\int_{\sigma_{n-1}}^{\sigma_n}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\bigg)\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\bigg]\\ &\le \cdots\\ &\le \mathrm{e}^{L_1 T}(c_*\,\mathrm{e}^{L_1 T})^{n-1}\sum_{i=0}^{n-1}\Big(c_*\mathbb{E}\big(\mathbb{W}_{\beta}(\mu_{\sigma_i},\tilde{\mu}_{\sigma_i})^{\beta}\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)\big)\\ &\qquad + L_1^{\frac{\beta}{2}}\,\mathbb{E}\Big(\int_0^t\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s\Big)^{\frac{\beta}{2}}\,\mathbb{P}(N_t = n)\Big),\end{aligned}$$

where in the first identity we used the independence of $\big|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}\big|^{\beta}$ and $\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)$; the first inequality is valid by invoking (3.8) and the independence between $\big|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_n-}\big|^{\beta}$ and $\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)$ as well as the independence between $\int_{\sigma_{n-1}}^{\sigma_n}\mathbb{W}_{\beta}(\mu_s,\tilde{\mu}_s)^2\mathrm{d}s$ and

$\mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)$; the second inequality holds true from (3.10); in the second identity we employed that ξ_{n-1} is independent of $|\Upsilon^{\mu,\tilde{\mu}}_{\sigma_{n-1}-}|^\beta \mathbb{1}_{[\sigma_n,\sigma_{n+1})}(t)$ and that $(\sigma_n)_{n\ge 1}$ is independent of $(p_n)_{n\ge 1}$; in the last inequality we used $c_* > 1$ and $\Upsilon^{\mu,\tilde{\mu}}_0 = \mathbf{0}$ due to $X^\mu_0 = X^{\tilde{\mu}}_0 = X_0$. Note that $\mathbb{P}(N_t = n) = \mathrm{e}^{-\lambda t}(\lambda t)^n/n!$ for $\lambda := \nu(\mathbb{1}_V)$ and recall from [7, p. 8] that

$$\sum_{k=1}^n \mathbb{E}(\mathbb{W}_\beta(\mu_{\sigma_k}, \tilde{\mu}_{\sigma_k})^\beta | N_t = n) = \frac{n}{t}\int_0^t \mathbb{W}_\beta(\mu_s, \tilde{\mu}_s)^\beta \,\mathrm{d}s.$$

Subsequently, we find that there exists a constant $C_3(T) > 0$ such that

$$\hat{\Gamma}_2(\mu, \tilde{\mu}) \le C_3(T)\big(1/\gamma + 1/(\gamma)^{\frac{\beta}{2}}\big)\mathbb{W}_{\beta,\gamma}(\mu, \tilde{\mu})^\beta.$$

This thus implies (3.7). Based on the preceding analysis, we conclude that Φ is contractive by choosing $\gamma > 0$ large enough so the statement (ii) follows.

Furthermore, according to (1.4) and (1.2),

$$\big(\nu(|g(x,\mu,\cdot)|^\beta \mathbb{1}_V(\cdot))\big)^{\frac{2}{\beta}} \le c_0(1 + |x|^2 + \mu(|\cdot|^\beta)^{\frac{2}{\beta}}).$$

This along with (1.6) yields that Assumption ($\mathbf{H}_2'$) is satisfied for the decoupled SDE (3.1). Therefore, in terms of Theorem 2.2, there exists a constant $c_T > 0$ such that

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |X^\mu_t|^\beta\Big) \le c_T\Big(1 + \mathbb{E}|X_0|^\beta + \sup_{0\le t\le T} \mu_t(|\cdot|^\beta)\Big).$$

Thus, the assertion (1.4) follows by noting that $(X_t)_{t\ge 0}$ and $(X^\mu_t)_{t\ge 0}$ with the alternative $\mu_t = \mathscr{L}_{X_t}$ share the same law on the path space $C([0,T];\mathbb{R}^d)$, and by making use of (1.5). □

Before the end of this section, we make some comments.

Remark 3.1 (i) It is quite natural to directly derive, via an approximate argument and Itô's formula, that Φ constructed in the proof of Theorem 1.1 is contractive. Nevertheless, some issues might be encountered when such a direct approach is adopted. To demonstrate the underlying difficulty, we set $f(x,\mu,z) = f(z)$ and $g(x,\mu,z) \equiv \mathbf{0}$ for simplicity. Thus, the chain rule, together with (1.3), shows formally that for $\beta \in [1,2]$,

$$\begin{aligned}\mathrm{d}|\Upsilon^{\mu,\tilde{\mu}}_t|^\beta &= \beta|\Upsilon^{\mu,\tilde{\mu}}_t|^{\beta-2}\langle \Upsilon^{\mu,\tilde{\mu}}_t, b(X^\mu_t,\mu_t) - b(X^{\tilde{\mu}}_t,\tilde{\mu}_t)\rangle\,\mathrm{d}t\\ &\le (L_1\beta|\Upsilon^{\mu,\tilde{\mu}}_t|^\beta + L_1\beta|\Upsilon^{\mu,\tilde{\mu}}_t|^{\beta-2}\mathbb{W}_\beta(\mu_t,\tilde{\mu}_t)^2).\rangle\,\mathrm{d}t\\ &(L_1\beta_1\cdots\tilde{\mu}_t)^2),\end{aligned}$$

Obviously, the second term in the inequality above cannot be dominated by the linear combination of $|\Upsilon_t^{\mu,\tilde{\mu}}|^\beta$ and $\mathbb{W}_\beta(\mu_t,\tilde{\mu}_t)^\beta$ when, in particular, the quantity $|\Upsilon_t^{\mu,\tilde{\mu}}|$ approaches to zero.
Additionally, if (1.3) is replaced by the following stronger one:

$$\begin{aligned}&2\langle b(x,\mu_1)-b(y,\mu_2),x-y\rangle+\nu(|f(x,\mu_1,\cdot)-f(y,\mu_2,\cdot)|^2\mathbb{1}_U(\cdot))\\&\quad\le L_1(|x-y|+\mathbb{W}_\beta(\mu_1,\mu_2))|x-y|,\quad x,y\in\mathbb{R}^d,\mu_1,\mu_2\in\mathscr{P}_\beta(\mathbb{R}^d),\end{aligned}$$

then the proof concerning the contraction of Φ will become much more straightforward by the aid of an approximate argument and Itô's formula. For the case mentioned previously, (1.4) can be weakened in the form: for any $x,y\in\mathbb{R}^d$ and $\mu_1,\mu_2\in\mathscr{P}_\beta(\mathbb{R}^d)$,

$$\nu(|g(x,\mu_1,\cdot)-g(y,\mu_2,\cdot)|^\beta\mathbb{1}_V(\cdot))\le L_1(|x-y|^\beta+\mathbb{W}_\beta^\beta(\mu_1,\mu_2)).$$

One can see the proof of Theorem 4.1 in the next subsection for details.

(ii) For the case $\beta\in(0,1)$, the proof above no longer works due to the definition of the L^β-Wasserstein distance $\mathbb{W}_\beta$. In particular, the contractivity of Φ under $\mathbb{W}_{\beta,\gamma}$ is unavailable even for $\gamma>0$ large enough. Nonetheless, concerning the case $\beta\in(0,1)$, it is possible to demonstrate existence of the strong solution via the Schauder fixed point theorem; see [7, Proposition 1] for related details.

(iii) Under Assumptions $(\mathbf{A}_1)$–$(\mathbf{A}_3)$, we can also derive that, for fixed $T>0$ and any $p\in[1,\beta)$ with $\beta\in(1,2]$, there exists a constant $C_T'>0$ such that

$$\mathbb{E}\Big(\sup_{0\le t\le T}|X_t|^p\Big)\le C_T'(1+\mathbb{E}|X_0|^\beta).$$

This can be achieved via the stochastic Gronwall inequality; see the derivation of (4.5) in Theorem 4.1 for more details.

3.2 Proof of Theorem 1.3

In this part, we move to finish the proof of Theorem 1.3.

Proof of Theorem 1.3 Below, to shorten the notation, for all $t\in[0,T]$ and $1\le i\le n$, we set $Q_t^i:=X_t^i-X_t^{i,n}$ and $\tilde{\mu}_t^n:=\frac{1}{n}\sum_{j=1}^n\delta_{X_t^j}$ for $t\in[0,T]$. By invoking the triangle inequality and the basic inequality: $(a+b)^p\le 2^{p-1}(a^p+b^p)$ for $a,b\ge0$, it follows that for any $t\in[0,T]$,

$$\begin{aligned}\mathbb{W}_p^p(\mu_t,\bar{\mu}_t^n)&\le 2^{p-1}\big(\mathbb{W}_p^p(\bar{\mu}_t^n,\tilde{\mu}_t^n)+\mathbb{W}_p^p(\mu_t,\tilde{\mu}_t^n)\big)\\&\le 2^{p-1}\frac{1}{n}\sum_{j=1}^n|Q_t^j|^p+2^{p-1}\mathbb{W}_p^p(\mu_t,\tilde{\mu}_t^n),\end{aligned}\tag{3.11}$$

where in the second inequality we exploited the fact that $\mathbb{W}_p^p(\bar{\mu}_t^n, \tilde{\mu}_t^n) \le \frac{1}{n}\sum_{j=1}^n |Q_t^j|^p$ since $\frac{1}{n}\sum_{j=1}^n \delta_{(X_t^j, X_t^{j,n})}$ is a coupling of $\bar{\mu}_t^n$ and $\tilde{\mu}_t^n$. Consequently, the assertion (1.11) follows from Gronwall's inequality, provided that there exist constants $C_1(T), C_2(T) > 0$ such that for any $t \in [0, T]$ and $1 \le j \le n$,

$$\mathbb{E}|Q_t^j|^p \le C_1(T)\int_0^t \mathbb{E}\mathbb{W}_p^p(\mu_s, \bar{\mu}_s^n)\,\mathrm{d}s \tag{3.12}$$

and

$$\mathbb{E}\mathbb{W}_p^p(\mu_t, \tilde{\mu}_t^n) \le C_2(T)\phi_{p,\beta,d}(n), \tag{3.13}$$

where $\phi_{p,\beta}(n, d)$ was defined as in (1.12), and the number $C_2(T)$ depends on the initial moment $\mathbb{E}|X_0|^\beta$.

Set for $\varepsilon > 0$, $r \in \mathbb{R}$ and $x \in \mathbb{R}^d$,

$$U_{\varepsilon,r}(x) := (\varepsilon + |x|^2)^{\frac{r}{2}}. \tag{3.14}$$

It is easy to see that for $\varepsilon > 0$, $r \in \mathbb{R}$ and $x \in \mathbb{R}^d$,

$$\nabla U_{\varepsilon,r}(x) = rxU_{\varepsilon,r-2}(x). \tag{3.15}$$

Thus, we obtain from Itô's formula and $(\mathbf{A}_1')$ that there exist constants $c_1, c_2 > 0$ such that for any $t \in [0, T]$ and $\varepsilon > 0$,

$$\begin{aligned} U_{\varepsilon,p}(Q_t^i) &\le \varepsilon^{\frac{p}{2}} + \hat{M}_t^i \\ &\quad + c_1\int_0^t \Big(U_{\varepsilon,p-2}(Q_s^i)|Q_s^i|\big(|Q_s^i| + \mathbb{W}_p(\mu_s, \bar{\mu}_s^n)\big) + |Q_s^i|^p + \mathbb{W}_p^p(\mu_s, \bar{\mu}_s^n)\Big)\,\mathrm{d}s \\ &\le \varepsilon^{\frac{p}{2}} + c_2\int_0^t \Big(U_{\varepsilon,p}(Q_s^i) + \frac{1}{n}\sum_{j=1}^n |Q_s^j|^p + \mathbb{W}_p^p(\mu_s, \bar{\mu}_s^n)\Big)\,\mathrm{d}s + \hat{M}_t^i, \end{aligned} \tag{3.16}$$

where $(\hat{M}_t^i)_{t\ge 0}$ is a local martingale. In particular, the first inequality in (3.16) follows exactly the line to derive (4.10) and we also took advantage of $X_0^i = X_0^{i,n}$ herein, and we made use of Young's inequality and (3.11) in the second inequality. Then, taking expectations on both sides of (3.16) followed by sending $\varepsilon \to 0$, the estimate (3.16) enables us to deduce that for any $t \in [0, T]$,

$$\max_{1\le i\le N}\mathbb{E}|Q_t^i|^p \le 2c_2\int_0^t \max_{1\le i\le N}\mathbb{E}|Q_s^i|^p\,\mathrm{d}t + c_2\int_0^t \mathbb{E}\mathbb{W}_p^p(\mu_s, \bar{\mu}_s^n)\,\mathrm{d}s.$$

Whence, (3.12) follows from Gronwall's inequality.

In terms of [15, Theorem 1], for all $1 \le p < \beta$, there exists a constant $c_3 > 0$ such that

$$\mathbb{E}\mathbb{W}_p^p(\mu_t, \tilde{\mu}_t^n) \le c_3\big(\mathbb{E}|X_t^i|^\beta\big)^{\frac{p}{\beta}}\phi_{p,\beta,d}(n), \quad t \in [0, T].$$

As a result, (3.13) is available by taking (1.5) into account.

Next, by applying the stochastic Gronwall inequality (see, e.g., [38, Lemma 3.7]) and then approaching $\varepsilon \to 0$, we derive from (3.16) that for any $0 < q_1 < q_2 < 1$,

$$\Big(\mathbb{E}\Big(\sup_{0\le s\le t}|Q_t^i|^{pq_1}\Big)\Big)^{\frac{1}{q_1}} \le c_2\Big(\frac{q_2}{q_2-q_1}\Big)^{\frac{1}{q_1}} \mathrm{e}^t \int_0^t \Big(\frac{1}{n}\sum_{j=1}^n \mathbb{E}|Q_s^j|^p + \mathbb{E}\mathbb{W}_p^p(\mu_s, \tilde{\mu}_s^n)\Big)\,\mathrm{d}s.$$

This, together with (3.12) and (3.13), implies that there exists a constant $C_3(T) > 0$ such that for any $0 < q_1 < q_2 < 1$ and $t \in [0, T]$,

$$\mathbb{E}\Big(\sup_{0\le s\le t}|Q_t^i|^{pq_1}\Big) \le \frac{q_2}{q_2-q_1}\big(C_3(T)\phi_{p,\beta,d}(n)\big)^{q_1}.$$

Thus, (1.13) follows immediately. □

To proceed, we make a comment on the method proving Theorem 1.3.

Remark 3.2 By applying Itô's formula and BDG's inequality, along with [15, Theorem 1], we can also prove (1.13) with pq_1 therein being replaced by β as soon as the order of the initial moment is greater than β. In this regard, the methods based, respectively, on the stochastic Gronwall inequality and BDG's inequality share the same feature. Regarding the latter approach, one further needs to handle the term $\mathbb{E}(\sup_{0\le t\le T} M_t^i)$, where

$$M_t^i := \int_0^t \int_U \big(|X_{s-}^{i,n} - X_{s-}^i + f(X_{s-}^{i,n}, z) - f(X_{s-}^i, z)|^p - |X_{s-}^{i,n} - X_{s-}^i|^p\big)\,\widetilde{N}^i(\mathrm{d}s, \mathrm{d}z).$$

To this end, some additional assumptions associated with f (e.g., f is Lipschitz in the state variable) have to be exerted, provided that mere (1.9) is imposed.

At the end of this part, we present the proof of the statement in Remark 1.2(ii).

Proof of Remark 1.2 (ii) Below, we stipulate $x, y \in \mathbb{R}^d$ and $\mu_1, \mu_2 \in \mathscr{P}_\beta(\mathbb{R}^d)$. It is easy to see that

$$\begin{aligned}
&2\langle b(x,\mu_1) - b(y,\mu_2), x-y\rangle + \nu(|f(x,\mu_1,\cdot) - f(y,\mu_2,\cdot)|^2 \mathbb{1}_U(\cdot))\\
&\le 2\langle x-y, C_1(x-y) - C_2(x|x|^2 - y|y|^2)\rangle s\\
&\quad + 2d^{\frac{1}{2}}|x-y|\big|\mu_1(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}} - \mu_2(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\big|\\
&\quad + C_3^2\nu(|\cdot|^2\mathbb{1}_U(\cdot))\big(C_4|x|^2 - C_4|y|^2 + \mu_1(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}} - \mu_2(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\big)^2
\end{aligned}$$

and that for $z \in \mathbb{R}^d$,

$$\begin{aligned}
|g(x,\mu_1,z) - g(y,\mu_2,z)| = |\mathbf{1}+z|\big||x| - |y| + \mu_1(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\\
- \mu_2(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\big|.
\end{aligned}$$

Via the Minkowski inequality, one obviously has

$$\begin{aligned}
&\big|\mu_1(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}} - \mu_2(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\big| \\
&= \Big|\Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} |h(x-z_1)|^\beta \pi(\mathrm{d}z_1,\mathrm{d}z_2)\Big)^{\frac{1}{\beta}} - \Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} |h(x-z_2)|^\beta \pi(\mathrm{d}z_1,\mathrm{d}z_2)\Big)^{\frac{1}{\beta}}\Big| \\
&\le \Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} |h(x-z_1)-h(x-z_2)|^\beta \pi(\mathrm{d}z_1,\mathrm{d}z_2)\Big)^{\frac{1}{\beta}} \\
&\le \|h\|_{\mathrm{Lip}}\Big(\int_{\mathbb{R}^d\times\mathbb{R}^d} |z_1-z_2|^\beta \pi(\mathrm{d}z_1,\mathrm{d}z_2)\Big)^{\frac{1}{\beta}},
\end{aligned} \tag{3.17}$$

where $\|h\|_{\mathrm{Lip}}$ means the Lipschitz constant of h, and $\pi \in \mathscr{C}(\mu_1,\mu_2)$. Thus, taking infimum with respect to π on both sides of (3.17) yields that

$$\big|\mu_1(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}} - \mu_2(|h(x-\cdot)|^\beta)^{\frac{1}{\beta}}\big| \le \|h\|_{\mathrm{Lip}}\mathbb{W}_\beta(\mu_1,\mu_2). \tag{3.18}$$

Next, a direct calculation (see [4, (2.3)] for more details) shows that

$$-\langle x-y, x|x|^2 - y|y|^2\rangle \le (1-(|x|^2+|y|^2)/6)|x-y|^2,$$

and it is easy to see that

$$(|x|^2-|y|^2)^2 \le 2(|x|^2+|y|^2)|x-y|^2.$$

Consequently, (1.3) follows from the basic inequality: $2ab \le a^2+b^2$, $a, b \in \mathbb{R}$, and making use of $\nu(|\cdot|^2\mathbb{1}_U(\cdot)) < \infty$ and $C_2 > 12C_3^2C_4^2\nu(|\cdot|^2\mathbb{1}_U(\cdot))$. Apparently, (1.4) is verifiable based on (3.18). Therefore, Assumption ($\mathbf{A}_1$) is examinable. In terms of definitions of b and f, it is easy to see that ($\mathbf{A}_2$) is satisfied in case of $\nu(|\cdot|^2\mathbb{1}_U(\cdot)) < \infty$ and $C_2 > 12C_3^2C_4^2\nu(|\cdot|^2\mathbb{1}_U(\cdot))$, and that ($\mathbf{A}_3$) holds true readily. The proof is thus complete. □

4 Extension to McKean–Vlasov SDEs with Common Noise

In this section, we consider the following McKean–Vlasov SDE with common noise:

$$\begin{aligned}
\mathrm{d}X_t =& b(X_t, \mathscr{L}_{X_t|\mathscr{F}_t^{N^0}})\,\mathrm{d}t + \int_U f(X_{t-},z)\,\widetilde{N}(\mathrm{d}t,\mathrm{d}z) \\
&+ \int_V g(X_{t-},\mathscr{L}_{X_t|\mathscr{F}_t^{N^0}},z)\,N(\mathrm{d}t,\mathrm{d}z) + \int_U f^0(X_{t-},z)\,\widetilde{N}^0(\mathrm{d}t,\mathrm{d}z) \\
&+ \int_V g^0(X_{t-},\mathscr{L}_{X_t|\mathscr{F}_t^{N^0}},z)\,N^0(\mathrm{d}t,\mathrm{d}z),
\end{aligned} \tag{4.1}$$

where $b : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}^d$, $f, f^0 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, and $g, g^0 : \mathbb{R}^d \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ are measurable maps; $U, V \subset \mathbb{R}^d\backslash\{\mathbf{0}\}$ so that $U \cap V = \emptyset$; Poisson measures

$N(\mathrm{d}t,\mathrm{d}z)$ and $N^0(\mathrm{d}t,\mathrm{d}z)$ correspond to the idiosyncratic noise and the common noise with Lévy measure ν and ν^0, respectively, while $\widetilde{N}(\mathrm{d}t,\mathrm{d}z)$ and $\widetilde{N}^0(\mathrm{d}t,\mathrm{d}z)$ are their associated compensated Poisson measures.

To distinguish between the underlying sources of randomness, we introduce complete probability spaces $(\Omega^1,\mathscr{F}^1,\mathbb{P}^1)$ and $(\Omega^0,\mathscr{F}^0,\mathbb{P}^0)$, whose respective filtrations $(\mathscr{F}_t^1)_{t\geq 0}$ and $(\mathscr{F}_t^0)_{t\geq 0}$ satisfy the usual conditions. In (4.1), $N(\mathrm{d}t,\mathrm{d}z)$ and $N^0(\mathrm{d}t,\mathrm{d}z)$ shall be supported, respectively, on $\Omega^1\times\mathbb{R}_+\times\mathbb{R}^d$ and $\Omega^0\times\mathbb{R}_+\times\mathbb{R}^d$. Throughout this section, we shall work on the product probability space $(\Omega,\mathscr{F},\mathbb{F},\mathbb{P})$, where $\Omega:=\Omega^0\times\Omega^1$, $(\mathscr{F},\mathbb{P})$ is the completion of $(\mathscr{F}^0\otimes\mathscr{F}^1,\mathbb{P}^0\otimes\mathbb{P}^1)$ and $\mathbb{F}$ is the complete and right-continuous argumentation of $(\mathscr{F}_t^0\otimes\mathscr{F}_t^1)_{t\geq 0}$. Moreover, we write $\mathbb{E}^0$, $\mathbb{E}^1$ and $\mathbb{E}$ as the expectations on $(\Omega^0,\mathscr{F}^0,\mathbb{P}^0)$, $(\Omega^1,\mathscr{F}^1,\mathbb{P}^1)$ and $(\Omega,\mathscr{F},\mathbb{F},\mathbb{P})$, respectively. Note that $\mathscr{L}_{X_t|\mathscr{F}_t^{N^0}}$ represents the conditional distribution with respect to the σ-algebra $\mathscr{F}_t^{N^0}:=\sigma\{Z_s^0:s\leq t\}$, in which

$$Z_t^0:=\int_0^t\int_U z\,\widetilde{N}^0(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_V z\,N^0(\mathrm{d}s,\mathrm{d}z).$$

Furthermore, in the subsequent analysis, we shall assume that the initial value X_0 is an $\mathscr{F}_0^1$-measurable random variable, so $(Z_t^0)_{t\geq 0}$ is the solely common source of noise.

4.1 Well-Posedness of McKean–Vlasov SDEs with Common Noise

To carry out the study on the well-posedness of the SDE (4.1), we impose the following assumptions.

($\mathbf{B}_0$) there is $\beta\in[1,2]$ so that

$$\begin{aligned}\nu(|f(\mathbf{0},\cdot)|^2\mathbb{1}_V(\cdot))+\nu^0(|f^0(\mathbf{0},\cdot)|^2\mathbb{1}_V(\cdot))+\nu((1\vee|\cdot|^\beta\vee|g(\mathbf{0},\delta_{\mathbf{0}},\cdot)|^\beta\mathbb{1}_V(\cdot))\\+\nu^0((1\vee|\cdot|^\beta\vee|g^0(\mathbf{0},\delta_{\mathbf{0}},\cdot)|^\beta\mathbb{1}_V(\cdot))<\infty;\end{aligned}$$

($\mathbf{B}_1$) for fixed $\mu\in\mathscr{P}_\beta(\mathbb{R}^d)$ and $z\in\mathbb{R}^d$, $\mathbb{R}^d\ni x\mapsto b(x,\mu)$, $\mathbb{R}^d\ni x\mapsto f(x,\mu,z)$ and $\mathbb{R}^d\ni x\mapsto f^0(x,\mu,z)$ are continuous and locally bounded, and there exists a constant $K_1>0$ such that for any $x,y,z\in\mathbb{R}^d$, and $\mu_1,\mu_2\in\mathscr{P}_\beta(\mathbb{R}^d)$,

$$\begin{aligned}&2\langle b(x,\mu_1)-b(y,\mu_2),x-y\rangle+\nu(|f(x,\cdot)-f(y,\cdot)|^2\mathbb{1}_U(\cdot))\\&+\nu^0(|f^0(x,\cdot)-f^0(y,\cdot)|^2\mathbb{1}_U(\cdot))\leq K_1|x-y|(|x-y|+\mathbb{W}_\beta(\mu_1,\mu_2))\end{aligned}\tag{4.2}$$

and

$$\begin{aligned}&\nu(|g(x,\mu_1,\cdot)-g(y,\mu_2,\cdot)|^\beta\mathbb{1}_V(\cdot))+\nu^0(|g^0(x,\mu_1,\cdot)-g^0(y,\mu_2,\cdot)|^\beta\mathbb{1}_V(\cdot))\\&\leq K_1(|x-y|^\beta+\mathbb{W}_\beta(\mu_1,\mu_2)^\beta);\end{aligned}\tag{4.3}$$

($\mathbf{B}_2$) there exists a constant $K_2 > 0$ such that for any $x \in \mathbb{R}^d$ and $\mu \in \mathscr{P}_\beta(\mathbb{R}^d)$,

$$\begin{aligned} &2\langle x, b(x,\mu)\rangle + \nu(|f(x,\cdot)|^2 \mathbb{1}_U(\cdot)) + \nu^0(|f^0(x,\cdot)|^2 \mathbb{1}_U(\cdot)) \\ &\leq K_2\big(1 + |x|^2 + |x|\mu(|\cdot|^\beta)^{\frac{1}{\beta}}\big); \end{aligned}$$

($\mathbf{B}_3$) for any $T, R > 0$ and $\mu \in C([0,T]; \mathscr{P}_\beta(\mathbb{R}^d))$,

$$\int_0^T \bigg(\sup_{\{|x|\leq R\}} |b(x,\mu_t)| + \int_U \sup_{\{|x|\leq R\}} |f(x,z)|^2 \nu(\mathrm{d}z) + \int_U \sup_{\{|x|\leq R\}} |f^0(x,z)|^2 \nu^0(\mathrm{d}z) \bigg) \mathrm{d}t < \infty.$$

The main result in this part is stated as follows.

Theorem 4.1 *Assume that Assumptions ($\mathbf{B}_0$)–($\mathbf{B}_3$) hold, and suppose further $X_0 \in L^\beta(\Omega^1 \to \mathbb{R}^d, \mathscr{F}_0^1, \mathbb{P}^1)$. Then, the McKean–Vlasov SDE with common noise (4.1) admits a unique strong solution $(X_t)_{t\geq 0}$ satisfying that, for any fixed $T > 0$, there exists a constant $C_T > 0$ such that*

$$\mathbb{E}|X_t|^\beta \leq C_T(1 + \mathbb{E}|X_0|^\beta), \quad 0 \leq t \leq T. \tag{4.4}$$

Furthermore, if $\beta \in (1,2]$, then for all $p \in [1,\beta)$ and $T > 0$, there exists a constant $C'_T > 0$ such that

$$\mathbb{E}\Big(\sup_{0\leq t\leq T} |X_t|^p\Big) \leq C'_T(1 + \mathbb{E}|X_0|^\beta). \tag{4.5}$$

Proof To begin with, we introduce some notations. Let

$$L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d)) := \big\{\mu : \Omega^0 \to \mathscr{P}_\beta(\mathbb{R}^d) \big| \, \mathbb{E}^0(\mu(|\cdot|^\beta)) < \infty\big\}.$$

Then $\big(L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d)), \mathcal{W}_\beta\big)$ is a complete metric space (see, e.g., [26, Lemma 1.2]) endowed with the metric:

$$\mathcal{W}_\beta(\mu_1,\mu_2) := \big(\mathbb{E}^0 \mathbb{W}_\beta^\beta(\mu_1,\mu_2)\big)^{\frac{1}{\beta}}, \quad \mu_1,\mu_2 \in L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d)),$$

so $C\big([0,T]; L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d))\big)$ is also a complete metric space for any fixed $T > 0$. In addition, we set for a fixed horizon $T > 0$,

$$\mathscr{D}_T^{X_0} = \Big\{\mu \in C([0,T]; L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d))) : \mu_0 = \mathscr{L}_{X_0}, \sup_{t\in[0,T]} \mathbb{E}^0(\mu_t(|\cdot|^\beta)) < \infty\Big\},$$

in which $X_0 \in L^\beta(\Omega^1 \to \mathbb{R}^d, \mathscr{F}_0^1, \mathbb{P}^1)$ is the initial value of $(X_t)_{t\geq 0}$, and

$$C([0,T]; L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d))) := \big\{\mu : [0,T] \times \Omega^0 \to \mathscr{P}_\beta(\mathbb{R}^d) \text{ is weakly continuous}\big\}.$$

For $\eta > 0$, $(\mathscr{D}_T^{X_0}, \mathcal{W}_{\beta,\eta})$ is a complete metric space equipped with the metric

$$\mathcal{W}_{\beta,\eta}(\mu, \widetilde{\mu}) := \sup_{0\le t\le T} \left(\mathrm{e}^{-\eta t}\, \mathcal{W}_\beta(\mu_t, \widetilde{\mu}_t)\right), \quad \mu, \widetilde{\mu} \in \mathscr{D}_T^{X_0}.$$

For $\mu \in \mathscr{D}_T^{X_0}$, we focus on the following SDE with random coefficients:

$$\begin{aligned} X_t^\mu =& X_0 + \int_0^t b(X_s^\mu, \mu_s)\,\mathrm{d}t + \int_0^t \int_U f(X_{s-}^\mu, z)\,\widetilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_V g(X_{s-}^\mu, \mu_s, z)\,N(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_U f^0(X_{s-}^\mu, z)\,\widetilde{N}^0(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_V g^0(X_{s-}^\mu, \mu_s, z)\,N^0(\mathrm{d}s, \mathrm{d}z). \end{aligned} \tag{4.6}$$

Under $(\mathbf{B}_0)$–$(\mathbf{B}_3)$, for each $\mu \ni \mathscr{D}_T^{X_0}$, (4.6) has a unique solution $(X_t^\mu)_{t\ge 0}$ with the aid of Theorem 2.1 (which is still available to the SDE (2.1) with random coefficients). Accordingly, we can define a map $\mathscr{D}_T^{X_0} \ni \mu \mapsto \Gamma(\mu)$ by

$$(\Gamma(\mu))_t = \mathscr{L}_{X_t^\mu|\mathscr{F}_t^{N^0}}, \quad t \in [0, T]. \tag{4.7}$$

By Itô's formula, it follows from $(\mathbf{B}_0)$, (4.3) and $(\mathbf{B}_2)$ that for some $C_1, C_2 > 0$,

$$\begin{aligned} \mathrm{d}(1 + |X_t^\mu|^2)^{\frac{\beta}{2}} \le& \frac{K_2\beta}{2}(1 + |X_t^\mu|^2)^{\frac{\beta}{2}-1}\left(1 + |X_t^\mu|^2 + |X_t^\mu|\mu_t(|\cdot|^\beta)^{\frac{1}{\beta}}\right)\mathrm{d}t \\ &+ C_1(1 + |X_t^\mu|^\beta + \mu_t(|\cdot|^\beta))\,\mathrm{d}t + \mathrm{d}\hat{M}_t^\mu \\ \le& C_2\left((1 + |X_t^\mu|^2)^{\frac{\beta}{2}} + \mu_t(|\cdot|^\beta)\right)\mathrm{d}t + \mathrm{d}\hat{M}_t^\mu, \end{aligned} \tag{4.8}$$

where $(\hat{M}_t^\mu)_{t\ge 0}$ is a martingale. Then, via Gronwall's inequality, we have

$$\mathbb{E}|X_t^\mu|^\beta \le \left(\mathbb{E}^1(1 + |X_0^\mu|^2)^{\frac{\beta}{2}} + C_2 \int_0^T \mathbb{E}^0\mu_t(|\cdot|^\beta)\,\mathrm{d}t\right)\mathrm{e}^{C_2 T}, \quad t \in [0, T].$$

This, together with the fact that

$$\mathbb{E}^0((\Gamma(\mu))_t(|\cdot|^\beta)) = \mathbb{E}^0\left(\mathbb{E}^1(|X_t^\mu|^\beta|\mathscr{F}_t^{N^0})\right) = \mathbb{E}|X_t^\mu|^\beta,$$

implies that for $\mu \in \mathscr{D}_T^{X_0}$,

$$\sup_{0\le t\le T} \mathbb{E}^0\left((\Gamma(\mu))_t(|\cdot|^\beta)\right) < \infty.$$

Next, note that for any $h \in \mathrm{Lip}_b(\mathbb{R}^d)$ and $t \in [0, T]$,

$$\begin{aligned}\big|\mathbb{E}^0\big((\Gamma(\mu))_t(h(\cdot))\big) - \mathbb{E}^0\big((\Gamma(\mu))_0(h(\cdot))\big)\big| =&\mathbb{E}^0\big|\mathbb{E}^1\big((h(X_t^\mu) - h(X_0^\mu))\big|\mathscr{F}_t^{N^0}\big)\big| \\ \leq&\mathbb{E}|h(X_t^\mu) - h(X_0^\mu)|.\end{aligned}$$

Thus, we can conclude $\Gamma(\mu) \in C([0, T]; L^{0,\beta}(\mathscr{P}_\beta(\mathbb{R}^d)))$ by following the line to derive (3.5) so that we arrive at $\Gamma(\mu) \in \mathscr{D}_T^{X_0}$.

In the sequel, we shall claim that Γ is contractive under $\mathcal{W}_{\beta,\eta}$ for some appropriate $\eta > 0$. According to (4.6), for $R_t^{\mu,\tilde{\mu}} := X_t^\mu - X_t^{\tilde{\mu}}$ with $\mu, \tilde{\mu} \in \mathscr{D}_T^{X_0}$, we have

$$\begin{aligned}\mathrm{d}R_t^{\mu,\tilde{\mu}} =&B_t\,\mathrm{d}t + \int_U F_t(z)\,\widetilde{N}(\mathrm{d}t, \mathrm{d}z) + \int_V G_t(z)\,N(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_U F_t^0(z)\,\widetilde{N}^0(\mathrm{d}t, \mathrm{d}z) + \int_V G_t^0(z)\,N^0(\mathrm{d}s, \mathrm{d}z),\end{aligned}$$

where $B_t := b(X_t^\mu, \mu_t) - b(X_t^{\tilde{\mu}}, \tilde{\mu}_t)$ and

$$F_t(z) := f(X_{t-}^\mu, z) - f(X_{t-}^{\tilde{\mu}}, z), \quad F_t^0(z) := f^0(X_{t-}^\mu, z) - f^0(X_{t-}^{\tilde{\mu}}, z),$$

$$G_t(z) := g(X_{t-}^\mu, \mu_t, z) - g(X_{t-}^{\tilde{\mu}}, \tilde{\mu}_t, z), \quad G_t^0(z) := g^0(X_{t-}^\mu, \mu_t, z) - g^0(X_{t-}^{\tilde{\mu}}, \tilde{\mu}_t, z).$$

Recall that $U_{\varepsilon,\beta}$ is defined as in (3.14). Then, applying Itô's formula and making use of $(\mathbf{B}_1)$ and (3.15) yield that

$$\begin{aligned}\mathrm{d}U_{\varepsilon,\beta}(R_t^{\mu,\tilde{\mu}}) \leq& \frac{\beta}{2}U_{\varepsilon,\beta-2}(R_t^{\mu,\tilde{\mu}})\big(2\langle R_t^{\mu,\tilde{\mu}}, B_t\rangle + \nu(|F_t(\cdot)|^2\mathbb{1}_U(\cdot)) + \nu^0(|F_t^0(\cdot)|^2\mathbb{1}_U(\cdot))\big)\,\mathrm{d}t \\ &+ 2^{\frac{\beta}{2}}\big(\nu(|G_t(\cdot)|^\beta\mathbb{1}_V(\cdot)) + \nu^0(|G_t^0(\cdot)|^\beta\mathbb{1}_V(\cdot))\big)\,\mathrm{d}t \\ &+ |R_t^{\mu,\tilde{\mu}}|^\beta(\nu(\mathbb{1}_V) + \nu^0(\mathbb{1}_V))\,\mathrm{d}t + \mathrm{d}\hat{M}_t \\ \leq& \frac{\beta K_1}{2}U_{\varepsilon,\beta-1}(R_t^{\mu,\tilde{\mu}})(|R_t^{\mu,\tilde{\mu}}| + \mathbb{W}_\beta(\mu_t, \tilde{\mu}_t))\,\mathrm{d}t \\ &+ c_1\,(|R_t^{\mu,\tilde{\mu}}|^\beta + \mathbb{W}_\beta^\beta(\mu_t, \tilde{\mu}_t))\,\mathrm{d}t + |R_t^{\mu,\tilde{\mu}}|^\beta(\nu(\mathbb{1}_V) + \nu^0(\mathbb{1}_V))\,\mathrm{d}t + \mathrm{d}\hat{M}_t,\end{aligned} \tag{4.9}$$

where $c_1 := 2^{\frac{\beta}{2}}K_1^\beta(\nu((1 + |\cdot|)^\beta\mathbb{1}_V(\cdot)) + \nu^0((1 + |\cdot|)^\beta\mathbb{1}_V(\cdot))) < \infty$ thanks to $(\mathbf{B}_0)$. Whereafter, integrating from 0 to t followed by taking expectations on both sides of (4.9), and applying Young's inequality and the fact that $X_t^\mu = X_t^{\tilde{\mu}} = X_0$, we obtain there exists a constant $C_T^* > 0$ that for any $t \in [0, T]$,

$$\mathbb{E}U_{\varepsilon,\beta}(R_t^{\mu,\tilde{\mu}}) \leq C_T^*\int_0^t \big(\mathbb{E}U_{\varepsilon,\beta}(R_s^{\mu,\tilde{\mu}}) + \mathbb{E}^0\mathbb{W}_\beta^\beta(\mu_s, \tilde{\mu}_s)\big)\,\mathrm{d}s. \tag{4.10}$$

This, combining with Gronwall's inequality and approaching $\varepsilon \to 0$, leads to

$$\mathbb{E}|R_t^{\mu,\tilde{\mu}}|^\beta \leq C_T^* e^{C_T^* T}\int_0^t \mathbb{E}^0\mathbb{W}_\beta^\beta(\mu_s, \tilde{\mu}_s)\,\mathrm{d}s, \quad t \in [0, T].$$

Correspondingly, we derive that

$$\begin{aligned}\mathcal{W}_{\beta,\eta}^{\beta}(\Gamma(\mu),\Gamma(\tilde{\mu})) &\leq \sup_{0\leq t\leq T}\big(\mathrm{e}^{-\eta\beta t}\,\mathbb{E}^0\big(\mathbb{E}^1\big(|R_t^{\mu,\tilde{\mu}}|^\beta|\mathscr{F}_t^{N^0}\big)\big)\big)\\ &\leq C_T^* e^{C_T^* T}\sup_{0\leq t\leq T}\Big(\int_0^t \mathrm{e}^{-\eta\beta(t-s)}\,\mathrm{e}^{-\eta\beta s}\,\mathbb{E}^0\mathbb{W}_\beta^\beta(\mu_s,\tilde{\mu}_s)\,\mathrm{d}s\Big)\\ &\leq C_T^* e^{C_T^* T}/(\eta\beta)\mathcal{W}_{\beta,\eta}^{\beta}(\mu,\tilde{\mu}).\end{aligned}$$

As a consequence, we conclude that Γ is contractive under $\mathcal{W}_{\beta,\eta}$ for $\eta > 0$ large enough so the strong well-posedness of (4.1) is available via the Banach fixed point theorem.

The assertion (4.4) follows by following the procedure to derive (4.8) and applying Gronwall's inequality. Next, by virtue of the stochastic Gronwall inequality (see, e.g., [38, Lemma 3.7]), we obtain from (4.8) that for any $0 < q_1 < q_2 < 1$ and $\mu \in \mathscr{D}_T^{X_0}$,

$$\Big(\mathbb{E}\Big(\sup_{0\leq t\leq T}(1+|X_t^\mu|^2)^{\frac{q_1\beta}{2}}\Big)\Big)^{\frac{1}{q_1}} \leq \Big(\frac{q_2}{q_2-q_1}\Big)^{\frac{1}{q_1}}\mathrm{e}^T\Big(\mathbb{E}^1(1+|X_0^\mu|^2)^{\frac{\beta}{2}} + C_2\int_0^T \mathbb{E}^0\mu_t(|\cdot|^\beta)\,\mathrm{d}t\Big).$$

In particular, we take $\mu \in \mathscr{D}_T^{X_0}$ as the fixed point of $\Gamma(\mu)$, defined in (4.7), such that $X_t^\mu = X_t$ for any $t \in [0,T]$ and

$$\Big(\mathbb{E}\Big(\sup_{0\leq t\leq T}(1+|X_t|^2)^{\frac{q_1\beta}{2}}\Big)\Big)^{\frac{1}{q_1}} \leq \Big(\frac{q_2}{q_2-q_1}\Big)^{\frac{1}{q_1}}\mathrm{e}^T\Big(\mathbb{E}^1(1+|X_0|^2)^{\frac{\beta}{2}} + C_2\int_0^T \mathbb{E}|X_t|^\beta\,\mathrm{d}t\Big).$$

As a result, (4.5) holds true from (4.4). □

At the end of this subsection, we make a remark concerning Assumptions ($\mathbf{B}_1$) and ($\mathbf{B}_2$).

Remark 4.2 As far as the decoupled SDE associated with the McKean–Vlasov SDE (1.1) is concerned, the frozen measure variable is deterministic so the interlacing technique is applicable and moreover the corresponding technical condition is weaker; see ($\mathbf{A}_1$) and ($\mathbf{A}_2$) for more details. However, with regard to the SDE with random coefficients corresponding to the conditional McKean–Vlasov SDE (4.1), the underlying measure-valued process is no longer deterministic but random. Thus, the interlacing trick adopt in the proof of Theorem 1.1 is unusable. Furthermore, once we replace the term $|X_t^\mu|\mu_t(|\cdot|^\beta)^{\frac{1}{\beta}}$ in (4.8) by $\mu_t(|\cdot|^\beta)^{\frac{2}{\beta}}$, we need to estimate correspondingly the quantity $\mathbb{E}((1+|X_t^\mu|^2)^{\frac{\beta}{2}-1}\mu_t(|\cdot|^\beta)^{\frac{2}{\beta}})$. In case that μ_t is deterministic, it is easy to bound the term mentioned. Nevertheless, $\mathbb{E}^0(\mu_t(|\cdot|^\beta)^{\frac{2}{\beta}})$ might explode for $\mu \in \mathscr{D}_T^{X_0}$ (in this case, $(\mu_t)_{t\geq 0}$ is a measure-valued stochastic process). On the basis of the aforementioned analysis, we impose Assumptions ($\mathbf{B}_1$) and ($\mathbf{B}_2$), which is a little bit stronger than Assumptions ($\mathbf{A}_1$) and ($\mathbf{A}_2$), to offset the singularity arising from the spatial variables.

4.2 Conditional Propagation of Chaos for McKean–Vlasov SDEs with Common Noise

In this subsection, we are still concerned with the Lévy-driven McKean–Vlasov SDE with common noise (4.1), which describes the asymptotic behavior of the mean-field interacting particle system below:

$$\begin{cases} \mathrm{d}\bar{X}_t^{i,n} = b(\bar{X}_t^{i,n}, \bar{\mu}_t^n)\,\mathrm{d}t + \int_U f(\bar{X}_{t-}^{i,n}, z)\,\widetilde{N}^i(\mathrm{d}t, \mathrm{d}z) + \int_V g(\bar{X}_{t-}^{i,n}, \bar{\mu}_{t-}^n, z)\,N^i(\mathrm{d}t, \mathrm{d}z) \\ \qquad + \int_U f^0(\bar{X}_{t-}^{i,n}, z)\,\widetilde{N}^{0,i}(\mathrm{d}t, \mathrm{d}z) + \int_V g^0(\bar{X}_{t-}^{i,n}, \bar{\mu}_{t-}^n, z)\,N^{0,i}(\mathrm{d}t, \mathrm{d}z), \\ \bar{X}_0^{i,n} = X_0^i, \qquad i = 1, 2, \cdots, n, \end{cases} \tag{4.11}$$

where $\bar{\mu}_t^n := \frac{1}{n}\sum_{i=1}^n \delta_{\bar{X}_t^{i,n}}$, $\bar{\mu}_{t-}^n := \frac{1}{n}\sum_{i=1}^n \delta_{\bar{X}_{t-}^{i,n}}$, and $\{N^i(\mathrm{d}t, \mathrm{d}z)\}_{1\le i\le n}$ (resp. $\{N^{0,i}(\mathrm{d}t, \mathrm{d}z)\}_{1\le i\le n}$) are independent Poisson measures with intensity measure $\mathrm{d}t \times \nu(\mathrm{d}z)$ (resp. $\mathrm{d}t \times \nu^0(\mathrm{d}z)$).

Throughout this subsection, we will assume that $\beta \in (1, 2]$ and work under Assumptions $(\mathbf{B}_0)$–$(\mathbf{B}_3)$ with β involved in Assumption $(\mathbf{B}_1)$ replaced by $p \in [1, \beta)$. It is easy to see that (4.11) has a unique strong solution $(\bar{X}_t^{i,n})_{t\ge 0}$. Denote by $\{(X_t^i)_{t\ge 0}\}_{1\le i\le n}$ n-independent versions of the unique solution to (4.1). In particular, $(\mu_t)_{t\ge 0}$ is their common distribution,

It is worth noting that in the presence of common noise, all particles in the stochastic system (4.11) are not asymptotically independent any more and the classical propagation of chaos no longer holds. However, [6, Theorem 2.12] puts forward the conditional propagation of chaos, which reveals that, conditioned on the σ-algebra associated with common noise, all particles are asymptotically independent and the empirical measure converges to the common conditional distribution of each particle. The specific result upon conditional propagation of chaos in our setting is as follows.

Theorem 4.3 *Assume that $\beta \in (1, 2]$, that Assumptions $(\mathbf{B}_0)$–$(\mathbf{B}_3)$ hold with β involved in Assumption $(\mathbf{B}_1)$ replaced by some $p \in [1, \beta)$, and suppose further $X_0^i \in L^\beta(\Omega^1 \to \mathbb{R}^d, \mathscr{F}_0^1, \mathbb{P}^1)$ for any $1 \le i \le n$. Then, for any fixed $T > 0$, there exists a constant $\bar{C}_T > 0$ such that*

$$\mathbb{E}\mathbb{W}_p^p(\bar{\mu}_t^n, \mu_t) \le \bar{C}_T \phi_{p,\beta,d}(n), \quad t \in [0, T],$$

where $\phi_{p,\beta}(n, d)$ was defined as in (1.12). Furthermore, for fixed $T > 0$ and any $0 \le q_1 < q_2 < 1$, there exists a constant $\hat{C}_T > 0$ such that

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |\bar{X}_t^{i,n} - X_t^i|^p\Big) \le \frac{q_2}{q_2 - q_1}\big(\hat{C}_T \phi_{p,\beta,d}(n)\big)^{q_1}. \tag{4.12}$$

Proof The structure of proof is largely analogous to Theorem 1.3, so we omit it here. □

Acknowledgements We would like to thank two referees for their constructive comments and suggestions. The research of Jianhai Bao is supported by the National Key R&D Program of China (2022YFA1006004) and the National Natural Science Foundation of China (12071340). The research of Jian Wang is supported by the National Key R&D Program of China (2022YFA1006003) and the National Natural Science Foundation of China (12225104 and 12531007).

Author Contributions All the authors contribute equally to the manuscript.

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Agarwal, A., Pagliarani, S.: A Fourier-based Picard-iteration approach for a class of McKean-Vlasov SDEs with Lévy jumps. Stochastics **93**, 592–624 (2021)
2. Albeverio, S., Brzeźniak, Z., Wu, J.-L.: Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients. J. Math. Anal. Appl. **371**, 309–322 (2010)
3. Applebaum, D.: Lévy Processes and Stochastic Calculus, 2nd edn. Cambridge University Press, Cambridge (2011)
4. Bao, J., Majka, Mateusz B., Wang, J.: Geometric ergodicity of modified Euler schemes for SDEs with super-linearity, arXiv:2412.19377
5. Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games. Springer, Cham (2018)
6. Carmona, R., Delarue, F.: Probabilistic theory of mean field games with applications II: mean field games with common noise and master equations. Springer, Cham (2018)
7. Cavallazzi, T.: Well-posedness and propagation of chaos for Lévy-driven McKean-Vlasov SDEs under Lipschitz assumptions, Electron. Commun. Probab. 30, 1-14 (2025)
8. Chaintron, L.P., Diez, A.: Propagation of chaos: a review of models, methods and applications I: Models and methods. Kinet. Relat. Models **15**, 895–1015 (2022)
9. Chaintron, L.P., Diez, A.: Propagation of chaos: a review of models, methods and applications II: Applications. Kinet. Relat. Models **15**, 1017–1173 (2022)
10. Chen, X., dos Reis, G., Stockinger, W.: Wellposedness, exponential ergodicity and numerical approximation of fully super-linear McKean-Vlasov SDEs and associated particle systems, Electron. J. Probab. *30* (2025), Paper No. 23, 50 pp.
11. Deng, C.S., Huang, X.: Harnack inequalities for McKean-Vlasov SDEs driven by subordinate Brownian motions. J. Math. Anal. Appl. **519**, 126763 (2023)
12. Deng, C.S., Huang, X.: Well-posedness for McKean-Vlasov SDEs with distribution dependent stable noises, Acta Math. Sin. (Engl. Ser.) *41* 1269–1278 (2025)
13. Deng, C.S., Huang, X.: Well-Posedness for McKean-Vlasov SDEs driven by multiplicative stable noises, arXiv:2401.11384
14. Erny, X.: Well-posedness and propagation of chaos for McKean-Vlasov equations with jumps and locally Lipschitz coefficients. Stoch. Proc. Appl. **150**, 192–214 (2022)
15. Fournier, N., Guillin, A.: On the rate of convergence in Wasserstein distance of the empirical measure. Probab. Theory Related Fields **162**, 707–738 (2015)
16. Frikha, N., Konakov, V., Menozzi, S.: Well-posedness of some non-linear stable driven SDEs. Discrete Contin. Dyn. Syst. **41**, 849–898 (2021)
17. Gyöngy, I., Krylov, N.V.: On stochastic equations with respect to semimartingales I, Stochastics, **4** , 1–21 (1980/81)

18. Hao, Z., Ren, C., Wu, M.: Supercritical McKean-Vlasov SDE driven by cylindrical α-stable process, arXiv:2410.18611
19. Hong, W., Hu, S., Liu, W.: McKean-Vlasov SDE and SPDE with locally monotone coefficients. Ann. Appl. Probab. **34**, 2136–2189 (2024)
20. Huang, X.: Path-distribution dependent SDEs with singular coefficients. Electron. J. Probab. **26**, 1–21 (2021)
21. Huang, X., Wang, F.-Y.: Distribution dependent SDEs with singular coefficients. Stochastic Process. Appl. **129**, 4747–4770 (2019)
22. Huang, X., Yang, F.-F.: Distribution-dependent SDEs with Hölder continuous drift and α-stable noise. Numer. Algorithms **86**, 813–831 (2021)
23. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, 2nd edn. North-Holland, Kodansha (1989)
24. Jean, J.: Weak and strong solutions of stochastic differential equations. Stochastics **3**, 171–191 (1980)
25. Kac, M.: Foundations of Kinetic Theory, in: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, **3**, University of California Press, Berkeley and Los Angeles, 171–197 (1956)
26. Kumar, C., Neelima, Reisinger, C., Stockinger, W., : Well-posedness and tamed schemes for McKean-Vlasov equations with common noise. Ann. Appl. Probab. **32**, 3283–3330 (2022)
27. Liang, M., Majka, M.B., Wang, J.: Exponential ergodicity for SDEs and McKean-Vlasov processes with Lévy noise. Ann. Inst. Henri Poincaré Probab. Stat. **57**, 1665–1701 (2021)
28. Majka, M.B.: A note on existence of global solutions and invariant measures for jump SDEs with locally one-sided Lipschitz drift. Probab. Math. Statist. **40**, 37–55 (2020)
29. Mao, X.: Stochastic differential equations and applications. Woodhead Publishing, Sawston (2007)
30. Marinelli, C., Röckner, M.: On maximal inequalities for purely discontinuous martingales in infinite dimensions, in: Séminaire de Probabilités XLVI, Lecture Notes in Math., vol. **2123**, Cham, Springer International Publishing, 293–315 (2014)
31. Mehri, S., Scheutzow, M., Stannat, W., Zangeneh, Bian Z.: Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. Ann. Appl. Probab. **30**, 175–207 (2020)
32. Méléard, S.: Asymptotic behaviour of some interacting particle systems, McKean-Vlasov and Boltzmann models, in: Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995), Lecture Notes in Math., vol. **1642**, Springer, Berlin, 42–95 (1996)
33. Prévôt, C., Röckner, M.: A concise course on stochastic partial differential equations. Springer, Berlin (2007)
34. Sznitman, A.-S.: Topics in Propagation of Chaos, in: École d'Été de Probabilités de Saint-Flour XIX-1989, Lecture Notes in Math., vol. **1464**, Springer, Berlin, 165–251 (1991)
35. Tran, N.-K., Kieu, T.-T., Luong, D.-T., Ngo, H.-L.: On the infinite time horizon approximation for Lévy-driven McKean-Vlasov SDEs with non-globally Lipschitz continuous and super-linearly growth drift and diffusion coefficients, J. Math. Anal. Appl., **543** , no. 2, part 2, Paper No. 128982 (2025)
36. Wang, F.-Y., Ren, P.: Distribution dependent stochastic differential equations. World Scientific Publishing Co Pte. Ltd., Hackensack, NJ (2024)
37. Xia, P., Xie, L., Zhang, X., Zhao, G.: $L^q(L^p)$-theory of stochastic differential equations. Stochastic Process. Appl. **130**, 5188–5211 (2020)
38. Xie, L., Zhang, X.: Ergodicity of stochastic differential equations with jumps and singular coefficients. Ann. Inst. Henri Poincaré Probab. Stat. **56**, 175–229 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.