Bernoulli **32**(2), 2026, 1432–1455
https://doi.org/10.3150/25-BEJ1913

Long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos for mean field interacting particle system

XING HUANG[1,a], FEN-FEN YANG[2,3,b] and CHENGGUI YUAN[4,c]

[1] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China,* [a]*xinghuang@tju.edu.cn*
[2] *Department of Mathematics, Shanghai University, Shanghai 200444, China,* [b]*yangfenfen@shu.edu.cn*
[3] *Newtouch Center for Mathematics of Shanghai University, Shanghai, 200444 China*
[4] *Department of Mathematics, Swansea University, Bay campus, SA1 8EN, UK,* [c]*c.yuan@swansea.ac.uk*

In this paper, a general result on the long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos (PoC), one type of the PoC with regularization effect, is derived for mean field interacting particle system driven by Lévy noise, where TV is the total variation distance and $\mathbb{W}_{\ell_1}$ is the L^1-Wasserstein distance. By using the method of coupling, the general result is applied to mean field interacting particle system driven by Brownian motion and $\alpha(\alpha > 1)$-stable noise respectively, where the non-interacting drift is assumed to be dissipative in long distance.

Keywords: α-stable noise; McKean-Vlasov SDEs; mean field interacting particle system; quantitative PoC; reflection coupling; total variation distance

1. Introduction

Let (E,ρ) be a Polish space and $\mathscr{P}(E)$ be the set of all probability measures on E equipped with the weak topology. Let Z_t be an n-dimensional Lévy process on some complete filtration probability space $(\Omega,\mathscr{F},(\mathscr{F}_t)_{t\geq 0},\mathbb{P})$, the characteristic function of which has the form

$$\mathbb{E}\mathrm{e}^{i\langle \xi, Z_t\rangle} = \exp\left\{ i\langle \eta,\xi\rangle t - \frac{1}{2}\langle a\xi,\xi\rangle t + t\int_{\mathbb{R}^n-\{0\}} \mathrm{e}^{i\langle z,\xi\rangle} - 1 - i\langle z,\xi\rangle 1_{\{|z|\leq 1\}}\nu(\mathrm{d}z)\right\},\quad \xi\in\mathbb{R}^n,$$

where $\eta\in\mathbb{R}^n$, a is an $n\times n$ non-negative definite symmetric matrix and ν is the Lévy measure with $\int_{\mathbb{R}^n}(1\wedge|z|^2)\nu(\mathrm{d}z)<\infty$.

Let $N\geq 1$ be an integer and $(Z_t^i)_{1\leq i\leq N}$ be i.i.d. copies of Z_t. Consider the mean field interacting particle system

$$\mathrm{d}X_t^{i,N} = b_t(X_t^{i,N},\hat{\mu}_t^N)\mathrm{d}t + \sigma_t(X_{t-}^{i,N})\mathrm{d}Z_t^i,\quad 1\leq i\leq N, \tag{1.1}$$

where $b:[0,\infty)\times\mathbb{R}^d\times\mathscr{P}(\mathbb{R}^d)\to\mathbb{R}^d$, $\sigma:[0,\infty)\times\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^n$ are measurable and bounded on bounded sets, $\hat{\mu}_t^N$ stands for the empirical distribution of $(X_t^{i,N})_{1\leq i\leq N}$, i.e. $\hat{\mu}_t^N=\frac{1}{N}\sum_{j=1}^N\delta_{X_t^{j,N}}$, and the initial distribution of (1.1) is exchangeable. We also consider the non-interacting particle system

$$\mathrm{d}X_t^i = b_t(X_t^i,\mathscr{L}_{X_t^i})\mathrm{d}t + \sigma_t(X_{t-}^i)\mathrm{d}Z_t^i,\quad 1\leq i\leq N, \tag{1.2}$$

where $\mathscr{L}_{X_t^i}$ is the distribution of X_t^i and the initial values $(X_0^i)_{1\leq i\leq N}$ are i.i.d.. Note that (1.2) consists of N independent SDEs behaving as

$$\mathrm{d}X_t = b_t(X_t,\mathscr{L}_{X_t})\mathrm{d}t + \sigma_t(X_{t-})\mathrm{d}Z_t, \tag{1.3}$$

which is called McKean-Vlasov SDE and was first introduced in McKean (1966).

The PoC means that the joint distribution of any k particles of (1.1) at time $t > 0$ converges weakly to that of (1.2) (i.e. $(X_t^{i,N})_{1\le i\le N<\infty}$ exhibits Kac's chaotic property at $t > 0$), provided it does so at time 0. The PoC can be used to describe the dynamical evolution of Kac's chaotic property (also known as Boltzmann's property), which was introduced by Kac (1954–1955) in his derivation of the spatially homogeneous Boltzmann equation by taking limit on the master equation for Poisson-like processes. One can also refer to Sznitman (1991) for more details on PoC.

Recently, numerous results have been established on quantitative PoC in various settings, particularly for systems driven by Brownian motion. We begin by reviewing several key results relevant to our current study. The first two types of results pertain to systems with Brownian motion noise, while the third type extends to systems driven by Lévy noise.

1.1. Entropy-entropy and TV-entropy type PoC

When $n = d$, $\sigma = I_{d\times d}$, the entropy method was developed in Bresch, Jabin and Wang (2023), Jabin and Wang (2018, 2016) to derive the entropy-entropy type PoC:

$$\mathrm{Ent}((P_t^{[k],N})^*\mu_0^N|(P_t^*\mu_0)^{\otimes k}) \le \frac{k}{N}\mathrm{Ent}(\mu_0^N|\mu_0^{\otimes N}) + \frac{ck}{N}, \quad t\in[0,T], 1\le k\le N \tag{1.4}$$

for some constant $c > 0$ depending on $T > 0$, here $P_t^*\mu_0$ is the distribution of the solution to (1.3) with initial distribution $\mu_0 \in \mathscr{P}(\mathbb{R}^d)$, $(P_t^{[k],N})^*\mu_0^N$ is the distribution of $(X_t^{i,N})_{1\le i\le k}$ with initial exchangeable distribution $\mu_0^N \in \mathscr{P}((\mathbb{R}^d)^N)$, $\mu_0^{\otimes k}$ denotes the k independent product of μ_0, and the relative entropy of two probability measures is defined as

$$\mathrm{Ent}(\nu|\mu) = \begin{cases} \nu(\log(\frac{\mathrm{d}\nu}{\mathrm{d}\mu})), & \nu \ll \mu; \\ \infty, & \text{otherwise.} \end{cases}$$

Recently, Lacker (2023) applies the BBGKY argument to estimate $\mathrm{Ent}((P_t^{[k],N})^*\mu_0^N\ |(P_t^*\mu_0)^{\otimes k})$ directly and then derives the sharp rate $\frac{k^2}{N^2}$ for entropy-entropy type PoC in the case of Lipschitzian or bounded interaction. Combining the BBGKY argument in Lacker (2023) and the uniform in time log-Sobolev inequality for $\mathscr{L}_{X_t^i}$, Lacker and Flem (2023) shows that the sharp long time entropy-entropy type PoC, which together with the Pinsker inequality $\|\mu-\nu\|_{TV}^2 \le 2\mathrm{Ent}(\nu|\mu)$ for the total variation distance

$$\|\gamma-\tilde{\gamma}\|_{TV} = \sup_{\|f\|_\infty\le 1}|\gamma(f)-\tilde{\gamma}(f)|,$$

implies the sharp uniform in time TV-entropy type PoC, i.e.

$$\|(P_t^{[k],N})^*\mu_0^N - (P_t^*\mu_0)^{\otimes k}\|_{TV} \le \sqrt{2\mathrm{Ent}(\mu_0^N\circ(\pi_k)^{-1}|\mu_0^{\otimes k}) + \frac{ck}{N}}, \quad t\ge 0, 1\le k\le N, \tag{1.5}$$

where for any $1\le k\le N$, π_k is the projection from $(\mathbb{R}^d)^N$ to $(\mathbb{R}^d)^k$ defined by

$$\pi_k(x) = (x^1, x^2, \cdots, x^k), \quad x = (x^1, x^2, \cdots, x^N) \in (\mathbb{R}^d)^N.$$

Quite recently, combining Wang's Harnack inequality with power, Monmarché, Ren and Wang (2024) presents explicit conditions for the uniform in time log-Sobolev inequality for $\mathscr{L}_{X_t^i}$ in the non-convex settings, which, together with Lacker and Flem (2023), implies the long time entropy-entropy type PoC with the sharp rate $\frac{k^2}{N^2}$.

1.2. Entropy-$\mathbb{W}_2^2$ and TV-$\mathbb{W}_2$ type PoC

For kinetic mean field interacting particle system, the authors in (Chen et al., 2024, Theorem 2.3) adopt the synchronous coupling technique to derive $\mathbb{W}_2$-$\mathbb{W}_2$ type PoC. They then apply the log-Harnack inequality due to the coupling by change of measure to obtain the entropy-$\mathbb{W}_2^2$ type PoC, i.e. for $s+1 \geq t > s \geq 0$,

$$\mathrm{Ent}((P_t^{[k],N})^*\mu_0^N | (P_t^*\mu_0)^{\otimes k}) \leq \frac{C_1 k}{N(t-s)^3}\mathbb{W}_2((P_s^{[N],N})^*\mu_0^N, (P_s^*\mu_0)^{\otimes N})^2 + \frac{k}{N}C_2, \quad 1 \leq k \leq N$$

for some constant $C_1 > 0$ and a constant C_2 depending on t, s and the variance of $P_s^*\mu_0$, where the L^2-Wasserstein distance is defined by

$$\mathbb{W}_2(\mu, \nu) = \inf_{\pi \in \mathbf{C}(\mu,\nu)} \left(\int_{(\mathbb{R}^d)^k \times (\mathbb{R}^d)^k} |x-y|^2 \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{2}}$$

for $\mathbf{C}(\mu, \nu)$ being the set of all couplings of μ and ν. The Pinsker inequality implies the TV-$\mathbb{W}_2$ type PoC, i.e. for $s+1 \geq t > s \geq 0$,

$$\|(P_t^{[k],N})^*\mu_0^N - (P_t^*\mu_0)^{\otimes k}\|_{TV} \leq \frac{\sqrt{C_1 k}\,\mathbb{W}_2((P_s^{[N],N})^*\mu_0^N, (P_s^*\mu_0)^{\otimes N})}{\sqrt{N}(t-s)^{\frac{3}{2}}} + \sqrt{\frac{k}{N}C_2}, \ 1 \leq k \leq N.$$

Both the entropy-$\mathbb{W}_2^2$ type PoC and TV-$\mathbb{W}_2$ type PoC reflect the regularization effect of the stochastic noise which allows the initial distribution $\mu_0^{\otimes N}$ to be singular with respect to μ_0^N.

1.3. $\mathbb{W}_{\ell^1}$-$\mathbb{W}_{\ell^1}$ type PoC

In Brownian motion noise case, when there exists a partially dissipative non-interacting drift, the authors in Durmus et al. (2020) adopt the asymptotic reflection coupling to derive the long time $\mathbb{W}_{\ell_1}$-$\mathbb{W}_{\ell_1}$ type PoC:

$$\mathbb{W}_{\ell_1}((P_t^{[k],N})^*\mu_0^N, (P_t^*\mu_0)^{\otimes k}) \leq \varepsilon(t)\frac{k}{N}\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N}) + c\frac{k}{\sqrt{N}}, \quad t \geq 0, 1 \leq k \leq N \tag{1.6}$$

for some constant $c > 0$ and $\lim_{t\to\infty} \varepsilon(t) = 0$, here the L^1-Wasserstein distance $\mathbb{W}_{\ell_1}$ is given by

$$\mathbb{W}_{\ell_1}(\mu, \nu) = \inf_{\pi \in \mathbf{C}(\mu,\nu)} \left(\int_{(\mathbb{R}^d)^k \times (\mathbb{R}^d)^k} \rho_{\ell_1}(x, y) \pi(\mathrm{d}x, \mathrm{d}y) \right), \quad \mu, \nu \in \mathscr{P}_1((\mathbb{R}^d)^k),$$

where

$$\rho_{\ell_1}(x, y) = \sum_{i=1}^k |x^i - y^i|, \quad x = (x^1, x^2, \cdots, x^k), y = (y^1, y^2, \cdots, y^k) \in (\mathbb{R}^d)^k,$$

and

$$\mathscr{P}_1((\mathbb{R}^d)^k) := \left\{ \mu \in \mathscr{P}((\mathbb{R}^d)^k) : \int_{(\mathbb{R}^d)^k} \rho_{\ell_1}(x, 0)\mu(\mathrm{d}x) < \infty \right\}.$$

The asymptotic reflection coupling was introduced in Wang (2015) to study the ergodicity of nonlinear monotone SPDEs. One can also refer to (Liu, Wu and Zhang, 2021, Theorem 2.11(b)) for PoC in $\mathbb{W}_{\ell_1}$

if $\mu_0^N = \mu_0^{\otimes N}$. The asymptotic reflection coupling can be also applied to study the long-time behavior of one-dimensional McKean-Vlasov SDEs with common noise in Bao and Wang (2025). In general Lévy noise case, Liang, Majka and Wang (2021) derives the long time $\mathbb{W}_{\ell_1}$-$\mathbb{W}_{\ell_1}$ type PoC (1.6) for interacting particle system by using the asymptotic refined basic coupling.

However, the initial conditions required for the entropy-entropy type PoC and entropy-$\mathbb{W}_2^2$ type PoC in the case of Brownian motion noise appear somewhat restrictive. To illustrate this, consider the initial condition $X_0^{1,N} = X_0^{2,N} = \cdots = X_0^{N,N}$ with $\mathbb{P}(X_0^{i,N} = x + \sqrt{N}) = 1 - \mathbb{P}(X_0^{i,N} = x) = \frac{1}{N}$, and $X_0^i = x, i \geq 1$. In this case, we observe that $\mathrm{Ent}(\mu_0^N | \mu_0^{\otimes N}) = \infty$ and $\frac{1}{N}\mathbb{W}_2(\mu_0^N, \mu_0^{\otimes N})^2 = 1$. This prevents Kac's chaotic property at $t > 0$. However, it holds $\frac{1}{N}\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N}) = \frac{1}{\sqrt{N}}$. This encourages us to weaken the initial conditions and present a new type PoC: TV-$\mathbb{W}_{\ell_1}$ type PoC, see for instance (1.14) below.

Furthermore, fewer results are available for the PoC in the total variation distance in Lévy noise case. We will extend the TV-$\mathbb{W}_{\ell_1}$ type PoC to the case of Lévy noise. The main tool we employ in the proof is the Duhamel formula in the literature (McKean and Singer, 1967, (3a)):

$$P_t^1 f - P_t^2 f = \int_0^t [P_s^1\{\mathscr{L}^1 - \mathscr{L}^2\}P_{t-s}^2 f]\mathrm{d}s, \tag{1.7}$$

where P_t^1 and P_t^2 are two Markov semigroups with generators $\mathscr{L}^1$ and $\mathscr{L}^2$ respectively.

1.4. The main contribution in Brownian motion noise case

Since Brownian motion noise case has been more extensively studied in the literature, we present our main contributions in this setting for clarity and accessibility. The general Lévy noise case will be discussed in Section 2.

Let $\{W_t^i\}_{i\geq 1}$ and $\{B_t^i\}_{i\geq 1}$ be independent d-dimensional Brownian motions and n-dimensional Brownian motions, where $\{W_t^i\}_{i\geq 1}$ is independent of $\{B_t^i\}_{i\geq 1}$. Let $\beta > 0$, $b^{(0)} : \mathbb{R}^d \to \mathbb{R}^d$, $b^{(1)} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^n$ be measurable and bounded on bounded sets. Consider

$$\mathrm{d}X_t^{i,N} = b^{(0)}(X_t^{i,N})\mathrm{d}t + \frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^{i,N}, X_t^{m,N})\mathrm{d}t + \sqrt{\beta}\mathrm{d}W_t^i + \sigma(X_t^{i,N})\mathrm{d}B_t^i \tag{1.8}$$

for $1 \leq i \leq N$ and the independent McKean-Vlasov SDEs:

$$\mathrm{d}X_t^i = b^{(0)}(X_t^i)\mathrm{d}t + \int_{\mathbb{R}^d} b^{(1)}(X_t^i, y)\mathscr{L}_{X_t^i}(\mathrm{d}y)\mathrm{d}t + \sqrt{\beta}\mathrm{d}W_t^i + \sigma(X_t^i)\mathrm{d}B_t^i, \;\; 1 \leq i \leq N. \tag{1.9}$$

To derive the long time TV-$\mathbb{W}_{\ell_1}$ type PoC, we make the following assumptions.

(A) There exists a constant $K_\sigma > 0$ such that

$$\frac{1}{2}\|\sigma(x_1) - \sigma(x_2)\|_{HS}^2 \leq K_\sigma |x_1 - x_2|^2, \;\; x_1, x_2 \in \mathbb{R}^d. \tag{1.10}$$

$b^{(0)}$ is continuous and there exist $R > 0$, $K_1 \geq 0$ and $K_2 > K_\sigma$ such that

$$\langle x_1 - x_2, b^{(0)}(x_1) - b^{(0)}(x_2)\rangle \leq \gamma(|x_1 - x_2|)|x_1 - x_2| \tag{1.11}$$

with

$$\gamma(r) = \begin{cases} K_1 r, & r \le R; \\ \left\{-\frac{K_1+K_2}{R}(r-R)+K_1\right\} r, & R \le r \le 2R; \\ -K_2 r, & r > 2R. \end{cases}$$

Moreover, there exists $K_b \ge 0$ such that

$$|b^{(1)}(x,y) - b^{(1)}(\tilde{x},\tilde{y})| \le K_b(|x-\tilde{x}| + |y-\tilde{y}|), \quad x,\tilde{x},y,\tilde{y} \in \mathbb{R}^d. \tag{1.12}$$

For any $p > 1$, let $\mathscr{P}_p(\mathbb{R}^d) = \{\mu \in \mathscr{P}(\mathbb{R}^d) : \|\mu\|_p := (\mu(|\cdot|^p))^{\frac{1}{p}} < \infty\}$.

Theorem 1.1. *Assume (A). Let $\mu_0 \in \mathscr{P}_{1+\delta}(\mathbb{R}^d)$ for some $\delta \in (0,1)$ and $\mu_0^N \in \mathscr{P}_1((\mathbb{R}^d)^N)$ be exchangeable. Let*

$$D := \int_0^\infty s \mathrm{e}^{\frac{1}{2\beta}\int_0^s \{\gamma(v)+K_\sigma v\}\mathrm{d}v} \mathrm{d}s.$$

If

$$K_b < \frac{2\beta^2}{(K_2 - K_\sigma)D^2}, \tag{1.13}$$

then there exists a positive constant c independent of k and N such that

$$\begin{aligned} &\|(P_t^{[k],N})^* \mu_0^N - (P_t^* \mu_0)^{\otimes k}\|_{TV} \\ &\le k c \mathrm{e}^{-\lambda t} \frac{\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N})}{N} + c\{1 + \|\mu_0\|_{1+\delta}\} k N^{-\frac{\delta}{1+\delta}}, \quad 1 \le k \le N, t \ge 1, \end{aligned} \tag{1.14}$$

with

$$\lambda = \frac{(K_2 - K_\sigma)D}{\beta}\left(\frac{2\beta^2}{(K_2 - K_\sigma)D^2} - K_b\right).$$

Typically in the quantitative PoC, the distance at time t is controlled by the same one at time 0 plus an additional term that vanishes as the number of particles N tends to infinity. However, in (1.14), we use a weaker distance to control a stronger one. We now analyze our findings and compare them with the aforementioned results.

(i) PoC: Compared to existing results on TV-entropy and TV-$\mathbb{W}_2$ type PoC, we establish quantitative Kac's chaos in total variation distance under a weaker initial condition. Specifically, our results show that even when

$$\lim_{N\to\infty} \frac{\mathrm{Ent}(\mu_0^N | \mu_0^{\otimes N})}{N} > 0 \quad \text{and} \quad \lim_{N\to\infty} \frac{\mathbb{W}_2(\mu_0^N, \mu_0^{\otimes N})^2}{N} > 0,$$

the condition $\lim_{N\to\infty} \frac{\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N})}{N} = 0$ is sufficient to guarantee quantitative Kac's chaos in total variation distance at $t \ge 1$.

(ii) "Generation of chaos" introduced in Lukkarinen (2023) is a relatively new concept that is gaining popularity, see Rosenzweig and Serfaty (2025) for the entropic generation of chaos of overdamped mean field Langevin dynamics with singular interactions and references therein for various generation of chaos. The term refers to the phenomenon where a mean field interacting system evolves over time toward a Kac's chaotic state, even if the initial distribution does not exhibit the Kac's chaotic property. Our work contributes to the generation of chaos by:

a) Exponential decay estimate: Demonstrating convergence to chaos with an explicit rate $e^{-\lambda t}$.

b) Weaker initial conditions: By employing the $\mathbb{W}_{\ell_1}$ distance, our results accommodate initial distributions that deviate significantly from chaoticity, broadening the scope of applicability.

(iii) Advantages of $\mathbb{W}_{\ell_1}$ over $\mathbb{W}_2$: Compared to the $\mathbb{W}_2$ distance, the weaker $\mathbb{W}_{\ell_1}$ distance offers a key advantage: it relaxes assumptions on the coefficients. Notably, it eliminates the need for strong convexity in the drift, thereby broadening the applicability of our results to a wider class of interacting particle systems.

The paper is organized in the following: In Section 2, we give a general result on the long time TV-$\mathbb{W}_{\ell_1}$ type PoC for mean field interacting particle system driven by general Lévy noise. We prove Theorem 1.1 in Section 3. The α-stable noise case is investigated in Section 4. Finally, some auxiliary lemmas which are used in the proof of the main results are postponed and carried out in a supplementary paper Huang, Yang and Yuan (2026).

2. A general result on long time TV-$\mathbb{W}_{\ell_1}$ type PoC

Let $b^{(0)}:\mathbb{R}^d\to\mathbb{R}^d$, $b^{(1)}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^n$ be measurable and bounded on bounded set. Recall that $(Z_t^i)_{i\geq 1}$ are i.i.d. n-dimensional Lévy processes. Consider the mean field interacting particle system

$$\mathrm{d}X_t^{i,N}=b^{(0)}(X_t^{i,N})\mathrm{d}t+\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^{i,N},X_t^{m,N})\mathrm{d}t+\sigma(X_{t-}^{i,N})\mathrm{d}Z_t^i,\quad 1\leq i\leq N, \tag{2.1}$$

and the non-interacting particle system

$$\mathrm{d}X_t^{i}=b^{(0)}(X_t^{i})\mathrm{d}t+\int_{\mathbb{R}^d} b^{(1)}(X_t^{i},y)\mathscr{L}_{X_t^i}(\mathrm{d}y)\mathrm{d}t+\sigma(X_{t-}^{i})\mathrm{d}Z_t^i,\quad 1\leq i\leq N. \tag{2.2}$$

We assume that SDEs (2.1) and (2.2) are well-posed. As in Section 1, let $P_t^*\mu_0=\mathscr{L}_{X_t^i}$ with $\mathscr{L}_{X_0^i}=\mu_0\in\mathscr{P}(\mathbb{R}^d)$, which is independent of i. For simplicity, we denote $\mu_t=P_t^*\mu_0$. For any exchangeable $\mu^N\in\mathscr{P}((\mathbb{R}^d)^N)$, $1\leq k\leq N$, $(P_t^{[k],N})^*\mu^N$ denotes the distribution of $(X_t^{i,N})_{1\leq i\leq k}$ with initial distribution μ^N.

To derive the long time TV-$\mathbb{W}_{\ell_1}$ type PoC, for any $s\geq 0$, consider the decoupled SDE

$$\mathrm{d}X_{s,t}^{i,\mu,z}=b^{(0)}(X_{s,t}^{i,\mu,z})\mathrm{d}t+\int_{\mathbb{R}^d} b^{(1)}(X_{s,t}^{i,\mu,z},y)\mu_t(\mathrm{d}y)\mathrm{d}t+\sigma(X_{s,t-}^{i,\mu,z})\mathrm{d}Z_t^i,\quad t\geq s \tag{2.3}$$

with $X_{s,s}^{i,\mu,z}=z\in\mathbb{R}^d$. Let

$$P_{s,t}^{i,\mu}f(z):=\mathbb{E}f(X_{s,t}^{i,\mu,z}),\quad f\in\mathscr{B}_b(\mathbb{R}^d),z\in\mathbb{R}^d,i\geq 1,0\leq s\leq t.$$

We also assume that (2.3) is well-posed so that $P_{s,t}^{i,\mu}$ does not depend on i and we denote

$$P_{s,t}^{\mu}=P_{s,t}^{i,\mu},\quad i\geq 1. \tag{2.4}$$

Moreover, for any $x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k$, $F \in \mathscr{B}_b((\mathbb{R}^d)^k)$, and $(s_1, s_2, \cdots, s_k) \in [0,t]^k$ define

$$(P^\mu_{s_1,t} \otimes P^\mu_{s_2,t} \otimes \cdots \otimes P^\mu_{s_k,t})F(x) := \mathbb{E}F(X^{1,\mu,x^1}_{s_1,t}, X^{2,\mu,x^2}_{s_2,t}, \cdots, X^{k,\mu,x^k}_{s_k,t}). \tag{2.5}$$

In particular, we denote

$$(P^\mu_{s,t})^{\otimes k}F(x) := (P^\mu_{s,t} \otimes P^\mu_{s,t} \otimes \cdots \otimes P^\mu_{s,t})F(x), \quad 0 \le s \le t. \tag{2.6}$$

For simplicity, we write $P^\mu_t = P^\mu_{0,t}$. For any $F \in C^1((\mathbb{R}^d)^k)$, $1 \le i \le k$, $x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k$, let $\nabla_i F(x)$ denote the gradient with respect to x^i. We now state the general result.

Theorem 2.1. *Let $\mu^N_0 \in \mathscr{P}_1((\mathbb{R}^d)^N)$ be exchangeable and $\mu_0 \in \mathscr{P}_p(\mathbb{R}^d)$ for some $p \ge 1$. Assume that the following conditions hold.*

(i) *For any $1 \le k \le N$, $F \in C^2_b((\mathbb{R}^d)^k)$ with $\|F\|_\infty \le 1$, $t \ge 0$, it holds*

$$\begin{aligned}&\int_{(\mathbb{R}^d)^k} F(x)\{(P^{[k],N}_t)^*\mu^N_0\}(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k} \{(P^\mu_t)^{\otimes k}F\}(x)(\mu^N_0 \circ \pi_k^{-1})(\mathrm{d}x)\\ &= \int_0^t \sum_{i=1}^k \int_{(\mathbb{R}^d)^N} \left\langle B^i_s(x), [\nabla_i (P^\mu_{s,t})^{\otimes k}F](\pi_k(x))\right\rangle \{(P^{[N],N}_s)^*\mu^N_0\}(\mathrm{d}x)\mathrm{d}s\end{aligned} \tag{2.7}$$

with

$$B^i_s(x) = \frac{1}{N}\sum_{m=1}^N b^{(1)}(x^i, x^m) - \int_{\mathbb{R}^d} b^{(1)}(x^i, y)\mu_s(\mathrm{d}y), \quad x = (x^1, x^2, \cdots, x^N) \in (\mathbb{R}^d)^N.$$

(ii) *There exists a measurable function $\varphi : (0,\infty) \to (0,\infty)$ with $\int_0^T \varphi(s)\mathrm{d}s < \infty, T > 0$ such that*

$$|\nabla P^\mu_{r,t} f| \le \varphi((t-r) \wedge 1)\|f\|_\infty, \quad f \in \mathscr{B}_b(\mathbb{R}^d), 0 \le r < t < T. \tag{2.8}$$

(iii) *There exist an increasing function $g : (0,\infty) \to (0,\infty)$ and a decreasing $\Delta : (0,\infty) \to (0,\infty)$ with $\lim_{N\to\infty} \Delta(N) = 0$ such that*

$$\begin{aligned}&\int_{(\mathbb{R}^d)^N} \left|B^1_s(x)\right| \{(P^{[N],N}_s)^*\mu^N_0\}(\mathrm{d}x)\\ &\le g(s)\left\{\frac{1}{N}\mathbb{W}_{\ell_1}(\mu^N_0, \mu_0^{\otimes N}) + \Delta(N)\{1 + \|\mu_0\|_p\}\right\}.\end{aligned} \tag{2.9}$$

(iv) *There exist functions $\varepsilon : (0,\infty) \to (0,\infty)$ with $\lim_{t\to\infty} \varepsilon(t) = 0$ and $\tilde{\Delta} : (0,\infty) \to (0,\infty)$ with $\lim_{N\to\infty} \tilde{\Delta}(N) = 0$ such that for any $t \ge 0$,*

$$\mathbb{W}_{\ell_1}((P^{[N],N}_t)^*\mu^N_0, (P^*_t\mu_0)^{\otimes N}) \le \varepsilon(t)\mathbb{W}_{\ell_1}(\mu^N_0, \mu_0^{\otimes N}) + \{1 + \|\mu_0\|_p\}N\tilde{\Delta}(N). \tag{2.10}$$

Moreover, there exists a constant $c_0 > 0$ such that

$$\sup_{t\ge 0}(P^*_t\mu_0)(|\cdot|^p) < c_0(1 + \mu_0(|\cdot|^p)). \tag{2.11}$$

Then there exists a constant $c > 0$ independent of t and N such that

$$\|(P_t^{[k],N})^*\mu_0^N - (P_t^*\mu_0)^{\otimes k}\|_{TV} \le ck\varepsilon(t-1)\frac{\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N})}{N} + c\{1+\|\mu_0\|_p\}k(\Delta(N)+\tilde{\Delta}(N)), \quad 1\le k\le N, t\ge 1. \tag{2.12}$$

Remark 2.2. (1) Different from the classical Duhamel formula (1.7), the left hand side of (2.7) only contains $(P_t^{[k],N})^*$ while the right hand side involves in $(P_t^{[N],N})^*$ due to the interaction. In Section 3 and Section 4 below, explicit conditions on the coefficients will be presented to ensure that (2.7) holds.

(2) In the study of entropy-entropy type PoC in Bresch, Jabin and Wang (2023), Jabin and Wang (2018, 2016) or $\mathbb{W}_{\ell_1}$-$\mathbb{W}_{\ell_1}$ type PoC (1.6), due to the tensor property of relative entropy or the property

$$\mathbb{W}_{\ell_1}((P_t^{[k],N})^*\mu_0^N, (P_t^*\mu_0)^{\otimes k}) \le \frac{k}{N}\mathbb{W}_{\ell_1}((P_t^{[N],N})^*\mu_0^N, (P_t^*\mu_0)^{\otimes N}), \tag{2.13}$$

one can derive the estimate of $\mathrm{Ent}((P_t^{[N],N})^*\mu_0^N|(P_t^*\mu_0)^{\otimes N})$ or $\mathbb{W}_{\ell_1}((P_t^{[N],N})^*\mu_0^N, (P_t^*\mu_0)^{\otimes N})$ for N particles first and then obtain the local PoC for k particles with $k\le N$. However, (2.13) does not hold if $\mathbb{W}_{\ell_1}$ is replaced by TV. This is the reason why we directly consider k particles with $k\le N$ instead of N particles in (2.7).

(3) As a hot topic related to the long time PoC, the ergodicity for McKean-Vlasov SDEs attracts much attention, see for instance Liang, Majka and Wang (2021), Song (2020), Wang (2023) and references therein for more details.

Proof. Firstly, for any $\gamma, \tilde{\gamma} \in \mathscr{P}(\mathbb{R}^n)$, since $C_b^2(\mathbb{R}^n)$, the functions from $\mathbb{R}^n$ to $\mathbb{R}$ having bounded and continuous up to second order derivatives, is dense in $\mathscr{B}_b(\mathbb{R}^n)$ under $L^1(\gamma+\tilde{\gamma})$, we have

$$\|\gamma-\tilde{\gamma}\|_{TV} = \sup_{\|f\|_\infty\le 1, f\in C_b^2(\mathbb{R}^n)} |\gamma(f)-\tilde{\gamma}(f)|, \quad \gamma,\tilde{\gamma}\in\mathscr{P}(\mathbb{R}^n). \tag{2.14}$$

Let $F\in C_b^2((\mathbb{R}^d)^k)$ with $\|F\|_\infty\le 1$. For any $(x^1,x^2,\cdots,x^{i-1},x^{i+1},\cdots,x^k)\in(\mathbb{R}^d)^{k-1}$, define

$$\begin{aligned}&[\mathscr{T}_{s,t}^{x^1,x^2,\cdots,x^{i-1},x^{i+1},\cdots,x^k}]F(z)\\ &= \mathbb{E}F(X_{s,t}^{1,\mu,x^1}, X_{s,t}^{2,\mu,x^2},\cdots,X_{s,t}^{i-1,\mu,x^{i-1}}, z, X_{s,t}^{i+1,\mu,x^{i+1}}\cdots, X_{s,t}^{k,\mu,x^k}), \quad z\in\mathbb{R}^d, \quad 0\le s\le t.\end{aligned}$$

This together with (2.6) and (2.8) implies that

$$\begin{aligned}|\nabla_i(P_{s,t}^\mu)^{\otimes k}F|(x^1,x^2,\cdots,x^k) &= |\nabla\{P_{s,t}^\mu[\mathscr{T}_{s,t}^{x^1,x^2,\cdots,x^{i-1},x^{i+1},\cdots,x^k}]F\}|(x^i)\\ &\le \varphi((t-s)\wedge 1), \quad 1\le i\le k, 0\le s<t.\end{aligned} \tag{2.15}$$

Then, using (2.7), (2.15), and the fact that $\{(P_s^{[N],N})^*\mu_0^N\}$ is exchangeable, we have

$$\begin{aligned}&\left|\int_{(\mathbb{R}^d)^k} F(x)\{(P_t^{[k],N})^*\mu_0^N\}(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)(\mu_0^N\circ\pi_k^{-1})(\mathrm{d}x)\right|\\ &\le \int_0^t k\int_{(\mathbb{R}^d)^N}\left|B_s^1(x)\right|\{(P_s^{[N],N})^*\mu_0^N\}(\mathrm{d}x)\varphi((t-s)\wedge 1)\mathrm{d}s.\end{aligned}$$

This combined with (2.9) implies that for any $t \geq 0$ and $1 \leq k \leq N$,

$$\left|\int_{(\mathbb{R}^d)^k} F(x)\{(P_t^{[k],N})^*\mu_0^N\}(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)(\mu_0^N\circ\pi_k^{-1})(\mathrm{d}x)\right|$$
$$\leq \int_0^t \varphi(s\wedge 1)\mathrm{d}sg(t)\left\{\frac{k}{N}\mathbb{W}_{\ell_1}(\mu_0^N,\mu_0^{\otimes N}) + k\Delta(N)\{1+\|\mu_0\|_p\}\right\}. \tag{2.16}$$

Next, for any $\tilde{\pi} \in \mathbf{C}(\mu_0^N \circ \pi_k^{-1}, \mu_0^{\otimes k})$, we conclude

$$\left|\int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)(\mu_0^N\circ\pi_k^{-1})(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)\mu_0^{\otimes k}(\mathrm{d}x)\right|$$
$$\leq \int_{(\mathbb{R}^d)^k\times(\mathbb{R}^d)^k}|\{(P_t^\mu)^{\otimes k}F\}(x) - \{(P_t^\mu)^{\otimes k}F\}(y)|\tilde{\pi}(\mathrm{d}x,\mathrm{d}y).$$

This together with (2.15) and $\mathbb{W}_{\ell_1}(\mu_0^N\circ\pi_k^{-1},\mu_0^{\otimes k}) \leq \frac{k}{N}\mathbb{W}_{\ell_1}(\mu_0^N,\mu_0^{\otimes N})$ implies that

$$\left|\int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)(\mu_0^N\circ\pi_k^{-1})(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k}\{(P_t^\mu)^{\otimes k}F\}(x)\mu_0^{\otimes k}(\mathrm{d}x)\right|$$
$$\leq \varphi(t\wedge 1)\mathbb{W}_{\ell_1}(\mu_0^N\circ\pi_k^{-1},\mu_0^{\otimes k}) \leq \varphi(t\wedge 1)\frac{k}{N}\mathbb{W}_{\ell_1}(\mu_0^N,\mu_0^{\otimes N}).$$

Finally, it follows from (2.16) as well as the triangle inequality that

$$\|(P_t^{[k],N})^*\mu_0^N - (P_t^*\mu_0)^{\otimes k}\|_{TV} \leq \left(\int_0^t \varphi(s\wedge 1)\mathrm{d}sg(t) + \varphi(t\wedge 1)\right)\frac{k}{N}\mathbb{W}_{\ell_1}(\mu_0^N,\mu_0^{\otimes N})$$
$$+ \int_0^t \varphi(s\wedge 1)\mathrm{d}sg(t)k\Delta(N)\{1+\|\mu_0\|_p\}. \tag{2.17}$$

By the definition of $(P_t^{[k],N})^*$, it is not difficult to see that

$$(P_t^{[k],N})^*\mu^N = \{(P_t^{[N],N})^*\mu^N\}\circ(\pi_k)^{-1}, \quad 1 \leq k \leq N. \tag{2.18}$$

Combining (2.18) with the definition of P_t^* yields

$$(P_{t+s}^{[k],N})^*\mu_0^N = \{(P_{t+s}^{[N],N})^*\mu_0^N\}\circ\pi_k^{-1} = \left\{(P_s^{[N],N})^*\{(P_t^{[N],N})^*\mu_0^N\}\right\}\circ\pi_k^{-1}$$
$$= (P_s^{[k],N})^*\{(P_t^{[N],N})^*\mu_0^N\},$$
$$P_{t+s}^*\mu_0 = P_s^*P_t^*\mu_0, \quad 1 \leq k \leq N, s \geq 0, t \geq 0.$$

Then for any $t \geq 1$, we derive from (2.17) for $t = 1$ that

$$\|(P_t^{[k],N})^*\mu_0^N - (P_t^*\mu_0)^{\otimes k}\|_{TV} = \|(P_1^{[k],N})^*\{(P_{t-1}^{[N],N})^*\mu_0^N\} - (P_1^*P_{t-1}^*\mu_0)^{\otimes k}\|_{TV}$$
$$\leq \left(\int_0^1 \varphi(s)\mathrm{d}sg(1) + \varphi(1)\right)\frac{k}{N}\mathbb{W}_{\ell_1}((P_{t-1}^{[N],N})^*\mu_0^N,(P_{t-1}^*\mu_0)^{\otimes N})$$
$$+ \int_0^1 \varphi(s)\mathrm{d}sg(1)k\Delta(N)\{1+\|P_{t-1}^*\mu_0\|_p\}.$$

This, combined with (2.10) and (2.11), gives (2.12). □

3. Proof of Theorem 1.1

Before giving the proof of Theorem 1.1, we first provide two remarks related to Theorem 1.1.

Remark 3.1. Under **(A)**, (1.8) and (1.9) are well-posed. By (1.11), $b^{(0)}$ is dissipative at long distance and (1.11) is equivalent to that there exist $C_1 \geq 0, C_2 > 0$ and $r_0 > 0$ such that for any $x, y \in \mathbb{R}^d$,

$$\langle x - y, b^{(0)}(x) - b^{(0)}(y)\rangle \leq C_1|x - y|^2 1_{\{|x-y|\leq r_0\}} - C_2|x - y|^2 1_{\{|x-y|>r_0\}}. \tag{3.1}$$

Remark 3.2. (1) The condition (1.13) requires that the Lipschitz constant of $b^{(1)}$ is sufficiently small, which is crucial to ensure $\lambda > 0$. A similar condition is imposed in (Durmus et al., 2020, (**H2**(ii)). In fact, such a smallness condition of the interaction is essential when one investigates the long-time behavior of McKean-Vlasov SDEs. A classical example is

$$\mathrm{d}X_t = -X_t\mathrm{d}t + \epsilon\mathbb{E}(X_t)\mathrm{d}t + \mathrm{d}W_t,$$

where $\epsilon = 1$, it admits infinitely many invariant probability measures, i.e. a phase transition occurs. When $\beta = 1$, $K_\sigma = 0$, by (Durmus et al., 2020, Theorem 2, Remark 4), the $\mathbb{W}_{\ell_1}$-$\mathbb{W}_{\ell_1}$ type PoC is derived under the assumption that the Lipschitz constant K_b of $b^{(1)}$ is sufficiently small and there exists a continuous function $\kappa : [0,\infty) \to \mathbb{R}$ with $\liminf_{r\to\infty} \kappa(r) > 0$ such that

$$\langle x_1 - x_2, b^{(0)}(x_1) - b^{(0)}(x_2)\rangle \leq -\kappa(|x_1 - x_2|)|x_1 - x_2|^2. \tag{3.2}$$

Note that $\liminf_{r\to\infty} \kappa(r) > 0$ implies that there exist constants $r_0 > 0$ and $C_2 > 0$ to ensure $\kappa(r) \geq C_2$ for any $r > r_0$, while the continuity of κ yields $\sup_{r\in[0,r_0]} |\kappa(r)| \leq C_1$ for some constant $C_1 > 0$. These observations show that (3.2) is equivalent to (3.1) and hence equivalent to (1.11) as explained in Remark 3.1. Hence, the conditions in Theorem 1.1 are consistent with those in (Durmus et al., 2020, Theorem 2, Remark 4).

(2) When $\sigma = 0$, $b^{(0)}, b^{(1)} \in C^1$, $b^{(0)} = \nabla V_1 + \nabla V_2$, $b^{(1)}(x, y) = \nabla_x W(x, y)$, V_1 is ρ-strongly convex and the coefficients satisfy

$$\|V_2\|_\infty + \|W\|_\infty + \|\nabla_x W\|_\infty < \infty, \ \ \beta > \frac{8}{\rho}\|\nabla_x W\|_\infty \exp\left\{2\frac{\|V_2\|_\infty + 2\|W\|_\infty}{\beta}\right\},$$

the density u_0 of μ_0 with respect to the invariant probability measure of $\mathscr{L}_{X_t^i}$ satisfies $\log u_0 = \bar{u} + \tilde{u}$ for bounded $\bar{u}$ and Lipschitz continuous $\tilde{u}$, and the initial distribution μ_0 and μ_0^N have finite moments of all orders, (Monmarché, Ren and Wang, 2024, Corollary 3.8) derives the sharp long time entropy-entropy type PoC with rate $\frac{k^2}{N^2}$, which implies the sharp long time TV-entropy type PoC (1.5) with rate $\frac{k}{N}$. In Theorem 1.1 above, $b^{(1)}$ is allowed to be unbounded, and $\mu_0 \in \mathscr{P}_{1+\delta}(\mathbb{R}^d)$ for some $\delta \in (0, 1)$, $\mu_0^N \in \mathscr{P}_1((\mathbb{R}^d)^N)$ and μ_0 can be singular with the invariant probability measure of $\mathscr{L}_{X_t^i}$. So, Theorem 1.1 is not covered in (Monmarché, Ren and Wang, 2024, Corollary 3.8).

(3) From the comparison above, one of the advantages of coupling method is to allow weaker conditions on initial distribution, i.e. μ_0^N can be singular with $\mu_0^{\otimes N}$ and the coefficients to be more general, whereas the cost is to reduce the rate of PoC since we need to estimate

$$\mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_s^i, X_s^m) - \int_{\mathbb{R}^d} b^{(1)}(X_s^i, y)\mathscr{L}_{X_s^i}(\mathrm{d}y)\right|,$$

which is provided in Lemma S2 of supplementary document Huang, Yang and Yuan (2026), and the central limit theorem indicates that the sharp rate is $N^{\frac{1}{2}}$ when $\mathscr{L}_{X_s^i}$ has finite second moment.

We are now in a position to prove Theorem 1.1.

Proof of Theorem 1.1. It follows from (2.14) that for any $\gamma^n \to \gamma$ and $\zeta^n \to \zeta$ weakly in $\mathscr{P}((\mathbb{R}^d)^k)$ as $n \to \infty$,

$$\|\gamma - \zeta\|_{var} = \sup_{|f|\le 1, f\in C_b((\mathbb{R}^d)^k)} |\gamma(f) - \zeta(f)| = \sup_{|f|\le 1, f\in C_b((\mathbb{R}^d)^k)} \lim_{n\to\infty} |\gamma^n(f) - \zeta^n(f)|$$
$$\le \liminf_{n\to\infty} \|\gamma^n - \zeta^n\|_{var}. \tag{3.3}$$

Combining (3.3) with Lemma S1 in Huang, Yang and Yuan (2026), we may and do assume that there exists a constant $K_0 > 0$ such that

$$|b^{(0)}(x_1) - b^{(0)}(x_2)| \le K_0|x_1 - x_2|, \ \ x_1, x_2 \in \mathbb{R}^d. \tag{3.4}$$

Following that, we intend to verify the conditions (i)-(iv) in Theorem 2.1 one by one.

(1) Take $\mathscr{F}_0$-measurable random variables $(X_0^{i,N})_{1\le i\le N}$ and $(X_0^i)_{1\le i\le N}$ such that

$$\mathscr{L}_{(X_0^{i,N})_{1\le i\le N}} = \mu_0^N, \ \ \mathscr{L}_{(X_0^i)_{1\le i\le N}} = \mu_0^{\otimes N}.$$

Fix $t > 0$. Recall that $P_{s,t}^\mu$ and $(P_{s,t}^\mu)^{\otimes k}$ are defined in (2.4) and (2.6) respectively. By (3.4), (1.10), (1.12) and the fact that $\beta \neq 0$, the backward Kolmogorov equation

$$\frac{\mathrm{d}P_{s,t}^\mu f}{\mathrm{d}s} = -\mathscr{L}_s^\mu P_{s,t}^\mu f, \ \ f \in C_b^2(\mathbb{R}^d), \|f\|_\infty \le 1, s \in [0,t] \tag{3.5}$$

holds with

$$\mathscr{L}_s^\mu = \langle b^{(0)}, \nabla\rangle + \left\langle \int_{\mathbb{R}^d} b^{(1)}(\cdot, y)\mu_s(\mathrm{d}y), \nabla \right\rangle + \frac{1}{2}\mathrm{Tr}[(\beta I_{d\times d} + \sigma\sigma^*)\nabla^2].$$

Recall that for any $F \in C^1((\mathbb{R}^d)^k)$, $1 \le i \le k$ and $x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k$, $\nabla_i F(x)$ represents the gradient with respect to x^i. Simply denote $\nabla_i^2 = \nabla_i \nabla_i$.

Next, we fix $F \in C_b^2((\mathbb{R}^d)^k)$ with $\|F\|_\infty \le 1$. Define

$$(\mathscr{L}_s^\mu)^i F(x) = \langle b^{(0)}(x^i), \nabla_i F(x)\rangle + \left\langle \int_{\mathbb{R}^d} b^{(1)}(x^i, y)\mu_s(\mathrm{d}y), \nabla_i F(x) \right\rangle$$
$$+ \frac{1}{2}\mathrm{Tr}[(\beta I_{d\times d} + (\sigma\sigma^*)(x^i))\nabla_i^2 F(x)], \ \ x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k, 1 \le i \le k,$$

and

$$(\mathscr{L}_s^\mu)^{\otimes k} F(x) = \sum_{i=1}^k (\mathscr{L}_s^\mu)^i F(x), \ \ x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k.$$

We now claim that

$$\frac{\mathrm{d}(P_{s,t}^\mu)^{\otimes k} F}{\mathrm{d}s} = -(\mathscr{L}_s^\mu)^{\otimes k} (P_{s,t}^\mu)^{\otimes k} F, \ \ s \in [0,t]. \tag{3.6}$$

In fact, for any $(s_1, s_2, \cdots, s_k) \in [0,t]^k$ and $x = (x^1, x^2, \cdots, x^k) \in (\mathbb{R}^d)^k$, define

$$\Psi_F(s_1, s_2, \cdots, s_k, x) = (P^\mu_{s_1,t} \otimes P^\mu_{s_2,t} \otimes \cdots \otimes P^\mu_{s_k,t})F(x),$$

and

$$\begin{aligned}&\{[\mathscr{T}^{x^1,x^2,\cdots,x^{i-1},x^{i+1},\cdots,x^k}_{s_1,s_2,\cdots,s_{i-1},s_{i+1},\cdots,s_k,t}]F\}(z)\\&= \mathbb{E}F(X^{1,\mu,x^1}_{s_1,t}, X^{2,\mu,x^2}_{s_2,t}, \cdots, X^{i-1,\mu,x^{i-1}}_{s_{i-1},t}, z, X^{i+1,\mu,x^{i+1}}_{s_{i+1},t} \cdots, X^{k,\mu,x^k}_{s_k,t}), \quad z \in \mathbb{R}^d, 1 \le i \le k,\end{aligned}$$

where $(P^\mu_{s_1,t} \otimes P^\mu_{s_2,t} \otimes \cdots \otimes P^\mu_{s_k,t})$ is given in (2.5). Then it follows from Fubini's theorem that

$$\Psi_F(s_1, s_2, \cdots, s_k, x) = P^\mu_{s_i,t}\{[\mathscr{T}^{x^1,x^2,\cdots,x^{i-1},x^{i+1},\cdots,x^k}_{s_1,s_2,\cdots,s_{i-1},s_{i+1},\cdots,s_k,t}]F\}(x^i), \quad 1 \le i \le k.$$

Combining this with (3.5), the definition of $(\mathscr{L}^\mu_{s_i})^i$ and Fubini's theorem, we conclude that

$$\frac{\partial}{\partial s_i}\Psi_F(s_1, s_2, \cdots, s_k, x) = -(\mathscr{L}^\mu_{s_i})^i(P^\mu_{s_1,t} \otimes P^\mu_{s_2,t} \otimes \cdots \otimes P^\mu_{s_k,t})F, \quad 1 \le i \le k,$$

which together with $(P^\mu_{s,t})^{\otimes k}F(x) = \Psi_F(s, s, \cdots, s, x)$ and the definition of $(\mathscr{L}^\mu_s)^{\otimes k}$ yields (3.6).

Now, we now prove condition (i) in Theorem 2.1. Recall that

$$B^i_s(x) = \frac{1}{N}\sum_{m=1}^N b^{(1)}(x^i, x^m) - \int_{\mathbb{R}^d} b^{(1)}(x^i, y)\mu_s(\mathrm{d}y), \quad x = (x^1, x^2, \cdots, x^N) \in (\mathbb{R}^d)^N.$$

Combining (3.6) with Itô's formula, for any $s \in [0,t]$, we have

$$\begin{aligned}&\mathrm{d}[(P^\mu_{s,t})^{\otimes k}F](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)\\&= \left[-(\mathscr{L}^\mu_s)^{\otimes k}(P^\mu_{s,t})^{\otimes k}F\right](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)\mathrm{d}s\\&+ \sum_{i=1}^k \langle b^{(0)}(X^{i,N}_s), \nabla_i[(P^\mu_{s,t})^{\otimes k}F](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)\rangle \mathrm{d}s\\&+ \sum_{i=1}^k \left\langle \frac{1}{N}\sum_{m=1}^N b^{(1)}(X^{i,N}_s, X^{m,N}_s), \nabla_i[(P^\mu_{s,t})^{\otimes k}F](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)\right\rangle \mathrm{d}s\\&+ \frac{1}{2}\sum_{i=1}^k \mathrm{Tr}[(\beta I_{d\times d} + (\sigma\sigma^*)(X^{i,N}_s))\nabla^2_i[(P^\mu_{s,t})^{\otimes k}F](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)]\mathrm{d}s + \mathrm{d}M_s\\&= \sum_{i=1}^k \left\langle B^i_s(X^{1,N}_s, X^{2,N}_s, \cdots, X^{N,N}_s), [\nabla_i(P^\mu_{s,t})^{\otimes k}F](X^{1,N}_s, X^{2,N}_s, \cdots, X^{k,N}_s)\right\rangle \mathrm{d}s + \mathrm{d}M_s\end{aligned}$$

for some martingale M_s. Integrating with respect to s from 0 to t and taking expectation, we arrive at

$$\begin{aligned}&\int_{(\mathbb{R}^d)^k} F(x)\{(P^{[k],N}_t)^*\mu^N_0\}(\mathrm{d}x) - \int_{(\mathbb{R}^d)^k}\{(P^\mu_t)^{\otimes k}F\}(x)(\mu^N_0 \circ \pi_k^{-1})(\mathrm{d}x)\\&= \int_0^t \sum_{i=1}^k \int_{(\mathbb{R}^d)^N}\left\langle B^i_s(x), [\nabla_i(P^\mu_{s,t})^{\otimes k}F](\pi_k(x))\right\rangle\{(P^{[N],N}_s)^*\mu^N_0\}(\mathrm{d}x)\mathrm{d}s. \qquad (3.7)\end{aligned}$$

So, the condition (i) in Theorem 2.1 follows.

(2) By (1.10)-(1.12) and (Wang, 2018, Theorem 4.1 (1)) for $\kappa_2(t)=0$, there exists a constant $c_0>0$ independent of K_0 such that

$$|\nabla P^{\mu}_{r,t} f| \le \frac{c_0}{(t-r)^{1/2}\wedge 1}\|f\|_\infty, \quad 0\le r<t, f\in \mathscr{B}_b(\mathbb{R}^d). \tag{3.8}$$

One can refer to (Priola and Wang, 2006, Corollary 3.5) for (3.8) in the time homogeneous case and (Wang, 2011, Theorem 1.1 (1)) for log-Harnack inequality. Hence, condition (ii) in Theorem 2.1 holds.

(3) Firstly, (1.11) implies that

$$\langle x_1-x_2, b^{(0)}(x_1)-b^{(0)}(x_2)\rangle \le K_1|x_1-x_2|^2, \quad x_1,x_2\in\mathbb{R}^d. \tag{3.9}$$

Then, it is standard to derive from (1.10), (1.12) and (3.9) that

$$\mathbb{E}((1+|X_t^1|^2)^{\frac{1+\delta}{2}}) \le c_0(t)\mu_0(1+|\cdot|^{1+\delta}), \quad t\ge 0 \tag{3.10}$$

for some increasing function $c_0:[0,\infty)\to[0,\infty)$. Let $Z_t^{i,N}=X_t^i-X_t^{i,N}$. By the Itô-Tanaka formula, or equivalently using $\psi_\varepsilon(x)=\sqrt{x^2+\varepsilon}$ to approximate $|x|$ as $\varepsilon\to 0$, (1.10), (1.12) and (3.9), we derive

$$\begin{aligned} \mathrm{d}|Z_t^{i,N}| &\le K_1|Z_t^{i,N}|\mathrm{d}t + K_b|Z_t^{i,N}|\mathrm{d}t + K_\sigma|Z_t^{i,N}|\mathrm{d}t + \frac{1}{N}\sum_{m=1}^N K_b|Z_t^{m,N}|\mathrm{d}t \\ &\quad + \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^i,X_t^m) - \int_{\mathbb{R}^d} b^{(1)}(X_t^i,y)\mu_t(\mathrm{d}y)\right|\mathrm{d}t \\ &\quad + \left\langle [\sigma(X_t^i)-\sigma(X_t^{i,N})]\mathrm{d}B_t^i, \frac{Z_t^{i,N}}{|Z_t^{i,N}|}1_{\{|Z_t^{i,N}|\neq 0\}}\right\rangle, \end{aligned}$$

where we used the fact

$$\begin{aligned} &\left\langle \int_{\mathbb{R}^d} b^{(1)}(X_t^i,y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^{i,N},X_t^{m,N}), \frac{Z_t^{i,N}}{|Z_t^{i,N}|}1_{\{|Z_t^{i,N}|\neq 0\}}\right\rangle \\ &\le \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^{i,N},X_t^{m,N}) - \frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^i,X_t^m)\right| \\ &+ \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^i,X_t^m) - \int_{\mathbb{R}^d} b^{(1)}(X_t^i,y)\mu_t(\mathrm{d}y)\right| \\ &\le K_b|Z_t^{i,N}| + \frac{1}{N}\sum_{m=1}^N K_b|Z_t^{m,N}| + \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^i,X_t^m) - \int_{\mathbb{R}^d} b^{(1)}(X_t^i,y)\mu_t(\mathrm{d}y)\right|. \end{aligned}$$

Moreover, Lemma S2 in Huang, Yang and Yuan (2026) and (3.10) imply that we can find an increasing function $c:[0,\infty)\to[0,\infty)$ such that

$$\mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^i,X_t^m) - \int_{\mathbb{R}^d} b^{(1)}(X_t^i,y)\mu_t(\mathrm{d}y)\right| \le c(t)\{1+\|\mu_0\|_{1+\delta}\}N^{-\frac{\delta}{1+\delta}}. \tag{3.11}$$

Applying Gronwall's inequality and (3.11), we get

$$\sum_{i=1}^N \mathbb{E}|Z_s^{i,N}| \le \mathrm{e}^{(K_1+2K_b+K_\sigma)s} \sum_{i=1}^N \mathbb{E}|Z_0^{i,N}| \\ + \mathrm{e}^{(K_1+2K_b+K_\sigma)s} sc(s)\{1+\|\mu_0\|_{1+\delta}\} N N^{-\frac{\delta}{1+\delta}}. \tag{3.12}$$

Since $\{X_s^{i,N}\}_{i=1}^N$ are exchangeable, we derive from (3.12) and (3.11) that

$$\begin{aligned}
&\int_{(\mathbb{R}^d)^N} \left|B_s^1(x)\right| \{(P_s^{[N],N})^*\mu_0^N\}(\mathrm{d}x) = \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left|B_s^i(X_s^{1,N}, X_s^{2,N}, \cdots, X_s^{N,N})\right| \\
&= \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_s^{i,N}, X_s^{m,N}) - \int_{\mathbb{R}^d} b^{(1)}(X_s^{i,N}, y)\mu_s(\mathrm{d}y)\right| \\
&\le \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_s^{i,N}, X_s^{m,N}) - \int_{\mathbb{R}^d} b^{(1)}(X_s^{i,N}, y)\mu_s(\mathrm{d}y)\right. \\
&\qquad \left. - \left(\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_s^i, X_s^m) - \int_{\mathbb{R}^d} b^{(1)}(X_s^i, y)\mu_s(\mathrm{d}y)\right)\right| \\
&+ \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_s^i, X_s^m) - \int_{\mathbb{R}^d} b^{(1)}(X_s^i, y)\mu_s(\mathrm{d}y)\right| \\
&\le 3K_b \frac{1}{N}\sum_{i=1}^N \mathbb{E}|X_s^{i,N} - X_s^i| + c(s)\{1+\|\mu_0\|_{1+\delta}\} N^{-\frac{\delta}{1+\delta}} \\
&\le 3K_b \mathrm{e}^{(K_1+2K_b+K_\sigma)s} \frac{1}{N}\sum_{i=1}^N \mathbb{E}|Z_0^{i,N}| \\
&+ \{3K_b \mathrm{e}^{(K_1+2K_b+K_\sigma)s} sc(s) + c(s)\}\{1+\|\mu_0\|_{1+\delta}\} N^{-\frac{\delta}{1+\delta}}.
\end{aligned}$$

Letting $g(s) = \max\{3K_b\mathrm{e}^{(K_1+2K_b+K_\sigma)s}, 3K_b\mathrm{e}^{(K_1+2K_b+K_\sigma)s} sc(s) + c(s)\}$, and taking infimum with respect to $(X_0^{i,N}, X_0^i)_{1\le i\le N}$ with $\mathscr{L}_{(X_0^{i,N})_{1\le i\le N}} = \mu_0^N, \mathscr{L}_{(X_0^i)_{1\le i\le N}} = \mu_0^{\otimes N}$, we get

$$\int_{(\mathbb{R}^d)^N} \left|B_s^1(x)\right| \{(P_s^{[N],N})^*\mu_0^N\}(\mathrm{d}x) \\ \le g(s)\left\{\frac{1}{N}\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N}) + \{1+\|\mu_0\|_{1+\delta}\} N^{-\frac{\delta}{1+\delta}}\right\}. \tag{3.13}$$

Therefore, the condition (iii) in Theorem 2.1 is verified.

(4) To verify condition (iv) in Theorem 2.1 under **(A)**, we adopt the technique of asymptotic reflection coupling. For any $\varepsilon \in (0,1]$, let $\pi_R^\varepsilon \in [0,1]$ and π_S^ε be two Lipschitz continuous function on $[0,\infty)$ satisfying

$$\pi_R^\varepsilon(x) = \begin{cases} 1, & x \ge \varepsilon; \\ 0, & x \le \frac{\varepsilon}{2} \end{cases}, \quad (\pi_R^\varepsilon)^2 + (\pi_S^\varepsilon)^2 = 1. \tag{3.14}$$

Let $\{\tilde{W}_t^i\}_{i\geq 1}$ be independent Brownian motions and independent of $\{W_t^i, B_t^i\}_{i\geq 1}$. Construct

$$\begin{aligned}\mathrm{d}\tilde{X}_t^i &= b^{(0)}(\tilde{X}_t^i)\mathrm{d}t + \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\mathrm{d}t \\ &\quad + \sqrt{\beta}\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)\mathrm{d}W_t^i + \sqrt{\beta}\pi_S^\varepsilon(|\tilde{Z}_t^{i,N}|)\mathrm{d}\tilde{W}_t^i + \sigma(\tilde{X}_t^i)\mathrm{d}B_t^i,\end{aligned}$$

and

$$\begin{aligned}\mathrm{d}\tilde{X}_t^{i,N} &= b^{(0)}(\tilde{X}_t^{i,N})\mathrm{d}t + \frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^{i,N}, \tilde{X}_t^{m,N})\mathrm{d}t \\ &\quad + \sqrt{\beta}\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)(I_{d\times d} - 2\tilde{U}_t^{i,N}\otimes\tilde{U}_t^{i,N})\mathrm{d}W_t^i \\ &\quad + \sqrt{\beta}\pi_S^\varepsilon(|\tilde{Z}_t^{i,N}|)\mathrm{d}\tilde{W}_t^i + \sigma(\tilde{X}_t^{i,N})\mathrm{d}B_t^i,\end{aligned}$$

where $\tilde{Z}_t^{i,N} = \tilde{X}_t^i - \tilde{X}_t^{i,N}$, $\tilde{U}_t^{i,N} = \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}$ and $\mathscr{L}_{(\tilde{X}_0^{i,N})_{1\leq i\leq N}} = \mu_0^N$ and $\mathscr{L}_{(\tilde{X}_0^i)_{1\leq i\leq N}} = \mu_0^{\otimes N}$. By the Itô-Tanaka formula, (1.10), (1.11) and (1.12), we have

$$\begin{aligned}\mathrm{d}|\tilde{Z}_t^{i,N}| &\leq \left\langle b^0(\tilde{X}_t^i) - b^0(\tilde{X}_t^{i,N}), \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}\right\rangle \mathrm{d}t + K_b|\tilde{Z}_t^{i,N}|\mathrm{d}t \\ &\quad + \frac{1}{2}\|\sigma(\tilde{X}_t^{i,N}) - \sigma(\tilde{X}_t^i)\|_{HS}^2 \frac{1}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}\mathrm{d}t + \frac{1}{N}\sum_{m=1}^N K_b|\tilde{Z}_t^{m,N}|\mathrm{d}t \\ &\quad + \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^i, \tilde{X}_t^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right|\mathrm{d}t \\ &\quad + \left\langle [\sigma(\tilde{X}_t^i) - \sigma(\tilde{X}_t^{i,N})]\mathrm{d}B_t^i, \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}\right\rangle \\ &\quad + 2\sqrt{\beta}\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)\left\langle \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}, \mathrm{d}W_t^i\right\rangle \\ &\leq \gamma(|\tilde{Z}_t^{i,N}|)\mathrm{d}t + K_b|\tilde{Z}_t^{i,N}|\mathrm{d}t + K_\sigma|\tilde{Z}_t^{i,N}|\mathrm{d}t + \frac{1}{N}\sum_{m=1}^N K_b|\tilde{Z}_t^{m,N}|\mathrm{d}t \\ &\quad + \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^i, \tilde{X}_t^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right|\mathrm{d}t \\ &\quad + \left\langle [\sigma(\tilde{X}_t^i) - \sigma(\tilde{X}_t^{i,N})]\mathrm{d}B_t^i, \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}\right\rangle \\ &\quad + 2\sqrt{\beta}\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)\left\langle \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}, \mathrm{d}W_t^i\right\rangle,\end{aligned}$$

where we used

$$\left\langle \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y) - \frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^{i,N}, \tilde{X}_t^{m,N}), \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|} 1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}} \right\rangle$$
$$\leq K_b|\tilde{Z}_t^{i,N}| + \frac{1}{N}\sum_{m=1}^N K_b|\tilde{Z}_t^{m,N}| + \left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^i, \tilde{X}_t^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right|.$$

Let $\tilde{\gamma}(v) = \gamma(v) + K_\sigma v, v \geq 0$, and define

$$f(r) = \int_0^r \mathrm{e}^{-\frac{1}{2\beta}\int_0^u \tilde{\gamma}(v)\mathrm{d}v} \int_u^\infty s\mathrm{e}^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\mathrm{d}s\mathrm{d}u, \quad r \geq 0.$$

Then one can see that

$$\begin{cases} f'(r) = \mathrm{e}^{-\frac{1}{2\beta}\int_0^r \tilde{\gamma}(v)\mathrm{d}v} \int_r^\infty s\mathrm{e}^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\mathrm{d}s > 0, \\ f''(r) = -\frac{1}{2\beta}\tilde{\gamma}(r)f'(r) - r. \end{cases} \tag{3.15}$$

Moreover, noting that $\gamma(v) \geq -K_2 v$, we derive

$$\int_0^\infty s\mathrm{e}^{\frac{1}{2\beta}\int_0^s \{\gamma(v)+K_\sigma v\}\mathrm{d}v}\mathrm{d}s \geq \int_0^\infty s\mathrm{e}^{\frac{-(K_2-K_\sigma)s^2}{4\beta}}\mathrm{d}s = \frac{2\beta}{K_2 - K_\sigma}.$$

Combining this with (1.13), we have

$$K_b < \frac{K_2 - K_\sigma}{2}. \tag{3.16}$$

Recall that

$$\gamma(r) = \begin{cases} K_1 r, & r \leq R; \\ \{-\frac{K_1+K_2}{R}(r-R) + K_1\}r, & R \leq r \leq 2R; \\ -K_2 r, & r > 2R \end{cases}$$

for $K_2 > K_b + K_\sigma$ due to (3.16). Letting $\ell_0 = \{1 + \frac{K_1+K_\sigma}{K_1+K_2}\}R$, it is not difficult to see that

$$\tilde{\gamma}(r) \geq 0, r \in [0, \ell_0]; \quad \tilde{\gamma}(r) < 0, r \in (\ell_0, \infty).$$

By (3.15), we derive

$$f''(r) \leq 0, \quad r \in [0, \ell_0]. \tag{3.17}$$

In view of the definition of γ and $\tilde{\gamma}$, we conclude that $\frac{r}{-\tilde{\gamma}(r)}$ is decreasing in (ℓ_0, ∞). This combined with the integration by parts formula gives

$$\int_r^\infty s\mathrm{e}^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\mathrm{d}s = \int_r^\infty \frac{2\beta s}{\tilde{\gamma}(s)}\left(\frac{\mathrm{d}}{\mathrm{d}s}\mathrm{e}^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\right)\mathrm{d}s \leq -\frac{2\beta r}{\tilde{\gamma}(r)}\mathrm{e}^{\frac{1}{2\beta}\int_0^r \tilde{\gamma}(v)\mathrm{d}v}, \quad r > \ell_0,$$

which together with (3.15) yields $f''(r) \leq 0, \ r \in (\ell_0, \infty)$. This as well as (3.17) means that $f'' \leq 0$ so that $f(r) \leq f'(0)r$ and $\frac{f(r)}{r}$ is decreasing on $(0, \infty)$. As a result, we derive from (3.15) that

$$\inf_{r>0} \frac{f(r)}{r} = \lim_{r\to\infty} \frac{f(r)}{r} = \lim_{r\to\infty} f'(r) = \lim_{r\to\infty} \frac{\int_r^\infty s\mathrm{e}^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\mathrm{d}s}{\mathrm{e}^{\frac{1}{2\beta}\int_0^r \tilde{\gamma}(v)\mathrm{d}v}} = \frac{2\beta}{K_2 - K_\sigma}.$$

So, we conclude that

$$\frac{2\beta}{K_2 - K_\sigma} r \le f(r) \le f'(0) r, \quad r \ge 0. \tag{3.18}$$

By Itô's formula and $f'' \le 0$, we have

$$\begin{aligned}
\mathrm{d} f(|\tilde{Z}_t^{i,N}|) &\le f'(|\tilde{Z}_t^{i,N}|)\tilde{\gamma}(|\tilde{Z}_t^{i,N}|)\mathrm{d}t + 2\beta f''(|\tilde{Z}_t^{i,N}|)\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)^2 \mathrm{d}t \\
&+ f'(|\tilde{Z}_t^{i,N}|)\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^i, \tilde{X}_t^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right| \mathrm{d}t \\
&+ f'(|\tilde{Z}_t^{i,N}|)\left\langle [\sigma(\tilde{X}_t^i) - \sigma(\tilde{X}_t^{i,N})]\mathrm{d}B_t^i, \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|} 1_{\{|\tilde{Z}_t^{i,N}| \ne 0\}}\right\rangle \\
&+ f'(|\tilde{Z}_t^{i,N}|) 2\sqrt{\beta}\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)\left\langle \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|} 1_{\{|\tilde{Z}_t^{i,N}| \ne 0\}}, \mathrm{d}W_t^i\right\rangle \\
&+ f'(|\tilde{Z}_t^{i,N}|) K_b |\tilde{Z}_t^{i,N}| \mathrm{d}t + f'(|\tilde{Z}_t^{i,N}|)\frac{1}{N}\sum_{m=1}^N K_b |\tilde{Z}_t^{m,N}| \mathrm{d}t.
\end{aligned} \tag{3.19}$$

It follows from the second equation of (3.15) and $\|f'\|_\infty = f'(0)$ that

$$\begin{aligned}
& f'(|\tilde{Z}_t^{i,N}|)\tilde{\gamma}(|\tilde{Z}_t^{i,N}|) + 2\beta f''(|\tilde{Z}_t^{i,N}|)\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)^2 \\
&\le \left(f'(|\tilde{Z}_t^{i,N}|)\tilde{\gamma}(|\tilde{Z}_t^{i,N}|) + 2\beta f''(|\tilde{Z}_t^{i,N}|)\right)\pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)^2 + \|f'\|_\infty \left\{\sup_{s\in[0,\varepsilon]} \gamma^+(s) + K_\sigma \varepsilon\right\} \\
&\le -2\beta|\tilde{Z}_t^{i,N}| + 2\beta|\tilde{Z}_t^{i,N}|\left(1 - \pi_R^\varepsilon(|\tilde{Z}_t^{i,N}|)^2\right) + \|f'\|_\infty \left\{\sup_{s\in[0,\varepsilon]} \gamma^+(s) + K_\sigma \varepsilon\right\} \\
&\le -2\beta|\tilde{Z}_t^{i,N}| + 2\beta\varepsilon + f'(0)\left\{\sup_{s\in[0,\varepsilon]} \gamma^+(s) + K_\sigma \varepsilon\right\}.
\end{aligned}$$

This combined with (3.18) and (3.19) gives

$$\begin{aligned}
\mathrm{d}\sum_{i=1}^N f(|\tilde{Z}_t^{i,N}|) &\le -\left\{\frac{2\beta}{f'(0)} - f'(0)\frac{(K_2 - K_\sigma)}{\beta} K_b\right\}\sum_{i=1}^N f(|\tilde{Z}_t^{i,N}|)\mathrm{d}t \\
&+ 2\sum_{i=1}^N \beta\varepsilon \mathrm{d}t + f'(0)\sum_{i=1}^N \left\{\sup_{s\in[0,\varepsilon]} \gamma^+(s) + K_\sigma\varepsilon\right\}\mathrm{d}t \\
&+ \sum_{i=1}^N f'(|\tilde{Z}_t^{i,N}|)\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_t^i, \tilde{X}_t^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right|\mathrm{d}t \\
&+ \sum_{i=1}^N f'(|\tilde{Z}_t^{i,N}|)\left\langle [\sigma(\tilde{X}_t^i) - \sigma(\tilde{X}_t^{i,N})]\mathrm{d}B_t^i, \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|} 1_{\{|\tilde{Z}_t^{i,N}| \ne 0\}}\right\rangle
\end{aligned}$$

$$+\sum_{i=1}^{N} f'(|\tilde{Z}_t^{i,N}|)2\sqrt{\beta}\pi_R^{\varepsilon}(|\tilde{Z}_t^{i,N}|)\left\langle \frac{\tilde{Z}_t^{i,N}}{|\tilde{Z}_t^{i,N}|}1_{\{|\tilde{Z}_t^{i,N}|\neq 0\}}, \mathrm{d}W_t^i \right\rangle.$$

Let $\lambda = \frac{2\beta}{f'(0)} - f'(0)\frac{(K_2-K_\sigma)}{\beta}K_b$. Then (1.13) and the fact that $f'(0) = \int_0^\infty s e^{\frac{1}{2\beta}\int_0^s \tilde{\gamma}(v)\mathrm{d}v}\mathrm{d}s$ imply $\lambda > 0$. Hence, it follows that

$$\begin{aligned}
\sum_{i=1}^{N}\mathbb{E}f(|\tilde{Z}_t^{i,N}|) &\leq \exp\{-\lambda t\}\sum_{i=1}^{N}\mathbb{E}f(|\tilde{Z}_0^{i,N}|) \\
&\quad + N\int_0^t \exp\{-\lambda(t-s)\}\left\{2\beta\varepsilon + f'(0)\left\{\sup_{s\in[0,\varepsilon]}\gamma^+(s) + K_\sigma\varepsilon\right\}\right\}\mathrm{d}s \\
&\quad + \int_0^t \exp\{-\lambda(t-s)\}f'(0)N\mathbb{E}\left|\frac{1}{N}\sum_{m=1}^{N} b^{(1)}(\tilde{X}_s^i, \tilde{X}_s^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_s^i, y)\mu_s(\mathrm{d}y)\right|\mathrm{d}s.
\end{aligned} \tag{3.20}$$

Next, we prove that there exists a constant $c_0 > 0$ such that

$$\mathbb{E}|\tilde{X}_t^i|^{1+\delta} \leq c_0(1+\mathbb{E}|\tilde{X}_0^i|^{1+\delta}), \quad t \geq 0. \tag{3.21}$$

It follows from Itô's formula that

$$\begin{aligned}
\mathrm{d}(1+|\tilde{X}_t^i|^2) &= 2\langle \tilde{X}_t^i, b^{(0)}(\tilde{X}_t^i)\rangle\mathrm{d}t + 2\left\langle \tilde{X}_t^i, \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_t^i, y)\mu_t(\mathrm{d}y)\right\rangle\mathrm{d}t \\
&\quad + \beta d\mathrm{d}t + \|\sigma(\tilde{X}_t^i)\|_{HS}^2\mathrm{d}t + \mathrm{d}\tilde{M}_t, \ t \geq 0
\end{aligned}$$

for some martingale $\tilde{M}_t$. By **(A)**, we can find a constant $C_0 > 0$ such that

$$\begin{aligned}
&2\langle x, b^{(0)}(x)\rangle + 2\left\langle x, \int_{\mathbb{R}^d} b^{(1)}(x,y)\mu_t(\mathrm{d}y)\right\rangle + \beta d + \|\sigma(x)\|_{HS}^2 \\
&\leq (2K_1+2K_2)|x|^2 1_{\{|x|\leq 2R\}} - (2K_2-2K_\sigma)|x|^2 + 2\langle x, b^{(0)}(0)\rangle + \beta d \\
&+ 2\sqrt{2K_\sigma}\|\sigma(0)\|_{HS}|x| + \|\sigma(0)\|_{HS}^2 + 2|x|K_b(|x| + \mu_t(|\cdot|)) + 2|x||b^{(1)}(0,0)| \\
&\leq (2K_1+2K_2)4R^2 + \beta d + \|\sigma(0)\|_{HS}^2 - (2K_2-2K_\sigma-4K_b)|x|^2 \\
&+ 2|x|(|b^{(0)}(0)| + \sqrt{2K_\sigma}\|\sigma(0)\|_{HS} + |b^{(1)}(0,0)|) - 2K_b|x|^2 + 2(1+|x|^2)^{\frac{1}{2}}K_b\mu_t(|\cdot|) \\
&\leq C_0 - (K_2-K_\sigma-2K_b)(1+|x|^2) - 2K_b(1+|x|^2) + 2(1+|x|^2)^{\frac{1}{2}}K_b\mu_t(|\cdot|) \\
&= C_0 - (K_2-K_\sigma-2K_b)(1+|x|^2) \\
&+ (1+|x|^2)^{\frac{1-\delta}{2}}\{-2K_b(1+|x|^2)^{\frac{1+\delta}{2}} + 2(1+|x|^2)^{\frac{\delta}{2}}K_b\mu_t(|\cdot|)\} \\
&\leq C_0 - (K_2-K_\sigma-2K_b)(1+|x|^2) \\
&+ (1+|x|^2)^{\frac{1-\delta}{2}}2K_b\frac{1}{1+\delta}\left\{-(1+|x|^2)^{\frac{1+\delta}{2}} + \mu_t((1+|\cdot|^2)^{\frac{1+\delta}{2}})\right\}, \quad x \in \mathbb{R}^d.
\end{aligned}$$

This together with the Itô formula gives

$$\begin{aligned}\mathrm{d}(1+|\tilde{X}_t^i|^2)^{\frac{1+\delta}{2}} &\le C_1\mathrm{d}t - \frac{1+\delta}{2}(K_2-K_\sigma-2K_b)(1+|\tilde{X}_t^i|^2)^{\frac{1+\delta}{2}}\mathrm{d}t \\ &\quad + K_b\left\{-(1+|\tilde{X}_t^i|^2)^{\frac{1+\delta}{2}} + \mu_t((1+|\cdot|^2)^{\frac{1+\delta}{2}})\right\} + \mathrm{d}\bar{M}_t,\ t\ge 0.\end{aligned}$$

Combining this with (3.16) yields (3.21). Then, applying Lemma S2 from Huang, Yang and Yuan (2026) implies

$$\mathbb{E}\left|\frac{1}{N}\sum_{m=1}^N b^{(1)}(\tilde{X}_s^i,\tilde{X}_s^m) - \int_{\mathbb{R}^d} b^{(1)}(\tilde{X}_s^i,y)\mu_s(\mathrm{d}y)\right| \le \tilde{c}_0\{1+\|\mu_0\|_{1+\delta}\}N^{-\frac{\delta}{1+\delta}},\quad s\ge 0 \tag{3.22}$$

for some constant $\tilde{c}_0>0$. Note that different from (3.11), $\tilde{c}_0$ in (3.22) is independent of s. Substituting (3.22) into (3.20) and applying (3.18), we can find some constant $c_1>0$ such that

$$\begin{aligned}\sum_{i=1}^N \mathbb{E}|\tilde{Z}_t^{i,N}| &\le c_1\mathrm{e}^{-\lambda t}\sum_{i=1}^N \mathbb{E}|\tilde{Z}_0^{i,N}| + c_1 N\{1+\|\mu_0\|_{1+\delta}\}N^{-\frac{\delta}{1+\delta}} \\ &\quad + c_1\left\{2\beta\varepsilon + f'(0)\left\{\sup_{s\in[0,\varepsilon]}\gamma^+(s) + K_\sigma\varepsilon\right\}\right\}.\end{aligned}$$

Letting $\varepsilon\to 0$, we derive

$$\mathbb{W}_{\ell_1}((P_t^{[N],N})^*\mu_0^N,(P_t^*\mu_0)^{\otimes N}) \le c_1\mathrm{e}^{-\lambda t}\sum_{i=1}^N \mathbb{E}|\tilde{Z}_0^{i,N}| + c_1 N\{1+\|\mu_0\|_{1+\delta}\}N^{-\frac{\delta}{1+\delta}}.$$

Taking infimum with respect to $(\tilde{X}_0^{i,N},\tilde{X}_0^i)_{1\le i\le N}$ with $\mathscr{L}_{(\tilde{X}_0^{i,N})_{1\le i\le N}}=\mu_0^N$, $\mathscr{L}_{(\tilde{X}_0^i)_{1\le i\le N}}=\mu_0^{\otimes N}$, we get

$$\mathbb{W}_{\ell_1}((P_t^{[N],N})^*\mu_0^N,(P_t^*\mu_0)^{\otimes N}) \le c_1\mathrm{e}^{-\lambda t}\mathbb{W}_{\ell_1}(\mu_0^N,\mu_0^{\otimes N}) + c_1 N\{1+\|\mu_0\|_{1+\delta}\}N^{-\frac{\delta}{1+\delta}}.$$

Therefore, condition (iv) in Theorem 2.1 holds. Finally, applying Theorem 2.1, we complete the proof. □

4. α-stable noise case

Recall that a d-dimensional rotationally invariant α-stable process has Lévy measure $\nu^\alpha(\mathrm{d}z)=\frac{c_{d,\alpha}}{|z|^{d+\alpha}}\mathrm{d}z$ for some constant $c_{d,\alpha}>0$ and the generator $-(-\Delta)^{\frac{\alpha}{2}}$ is defined by

$$-(-\Delta)^{\frac{\alpha}{2}}f(x)=\int_{\mathbb{R}^d-\{0\}}\{f(x+z)-f(x)-\langle\nabla f(x),z\rangle 1_{\{|z|\le 1\}}\}\nu^\alpha(\mathrm{d}z),\quad f\in C_b^2(\mathbb{R}^d),\|f\|_\infty\le 1.$$

Let $b^{(0)}$, $b^{(1)}$ and $\{Z_t^i\}_{i\ge 1}$ be introduced in Section 2 with $n=d$ and $\sigma=I_{d\times d}$. The equations (2.1) and (2.2) reduce to

$$\mathrm{d}X_t^{i,N}=b^{(0)}(X_t^{i,N})\mathrm{d}t+\frac{1}{N}\sum_{m=1}^N b^{(1)}(X_t^{i,N},X_t^{m,N})\mathrm{d}t+\mathrm{d}Z_t^i,\quad 1\le i\le N,$$

$$\mathrm{d}X_t^i = b^{(0)}(X_t^i)\mathrm{d}t + \int_{\mathbb{R}^d} b^{(1)}(X_t^i, y)\mathscr{L}_{X_t^i}(\mathrm{d}y)\mathrm{d}t + \mathrm{d}Z_t^i, \quad 1 \le i \le N,$$

respectively. We make the following assumptions.

(B1) The generator of Z_t^i is $-(-\Delta)^{\frac{\alpha}{2}}$ for some $\alpha \in (1,2)$.

(B2) $b^{(0)}$ is continuous. There exist $\ell_0 > 0$, $K_1 \ge 0, K_2 > 0$ and $K_b \ge 0$ such that

$$\langle x-y, b^{(0)}(x) - b^{(0)}(y)\rangle \le K_1|x-y|^2 1_{\{|x-y|\le \ell_0\}} - K_2|x-y|^2 1_{\{|x-y|>\ell_0\}}, \tag{4.1}$$

and

$$|b^{(1)}(x,y) - b^{(1)}(\tilde{x},\tilde{y})| \le K_b(|x-\tilde{x}| + |y-\tilde{y}|), \quad x,\tilde{x},y,\tilde{y} \in \mathbb{R}^d.$$

Recall that for two measures $\zeta, \tilde{\zeta}$ on $\mathbb{R}^d$, $\zeta \wedge \tilde{\zeta} = \zeta - (\zeta - \tilde{\zeta})^+$. Let

$$J^\alpha(s) = \inf_{x\in\mathbb{R}^d, |x|\le s} (\nu^\alpha \wedge (\delta_x * \nu^\alpha))(\mathbb{R}^d), \quad s \ge 0, \tag{4.2}$$

here

$$\delta_x * \nu^\alpha(\mathrm{d}z) = \nu^\alpha(\mathrm{d}z - x) = \frac{c_{d,\alpha}}{|z-x|^{d+\alpha}}\mathrm{d}z, \quad x \in \mathbb{R}^d,$$

and hence

$$(\nu^\alpha \wedge (\delta_x * \nu^\alpha))(\mathrm{d}z) = \frac{c_{d,\alpha}}{(|z| \vee |z-x|)^{d+\alpha}}\mathrm{d}z, \quad x \in \mathbb{R}^d.$$

By (Luo and Wang, 2019, Example 1.2), there exist constants $\kappa > 0$ and $\tilde{c}_{d,\alpha} > 0$ such that

$$J^\alpha(s) \ge \tilde{c}_{d,\alpha} s^{-\alpha}, \quad s \in (0, \kappa]. \tag{4.3}$$

In fact, for any $r > 0$ and $x \in \mathbb{R}^d$ with $|x| = r$, it holds

$$\begin{aligned}(\nu^\alpha \wedge (\delta_x * \nu^\alpha))(\mathbb{R}^d) &= \int_{\mathbb{R}^d} \frac{c_{d,\alpha}}{(|z| \vee |z-x|)^{d+\alpha}}\mathrm{d}z \ge \int_{|z|\le\frac{r}{2}} \frac{c_{d,\alpha}}{(|z| \vee |z-x|)^{d+\alpha}}\mathrm{d}z \\ &\ge \frac{2^{d+\alpha}}{3^{d+\alpha}} c_{d,\alpha} r^{-d-\alpha} \int_{|z|\le\frac{r}{2}} \mathrm{d}z = \tilde{c}_{d,\alpha} r^{-\alpha}.\end{aligned}$$

So, we have

$$J^\alpha(s) = \inf_{0\le r\le s} \inf_{x\in\mathbb{R}^d, |x|=r} (\nu^\alpha \wedge (\delta_x * \nu^\alpha))(\mathbb{R}^d) \ge \inf_{0\le r\le s} \tilde{c}_{d,\alpha} r^{-\alpha} = \tilde{c}_{d,\alpha} s^{-\alpha}, \quad s > 0.$$

Moreover, let $\eta \in (0,1)$ and take

$$\sigma_\eta(r) = \frac{\tilde{c}_{d,\alpha}(\kappa \wedge (2\ell_0))^{2-\alpha}}{2(2\ell_0)^{1+\eta}} r^\eta, \quad r \in [0, 2\ell_0]. \tag{4.4}$$

Then $\sigma_\eta \in C([0,2\ell_0]) \cap C^2((0,2\ell_0])$ and it is a nondecreasing and concave function. Note that

$$\frac{\tilde{c}_{d,\alpha}(\kappa \wedge (2\ell_0))^{2-\alpha}}{2(2\ell_0)^{1+\eta}} = \inf_{r\in[0,2\ell_0]} \frac{\tilde{c}_{d,\alpha}(\kappa \wedge r)^{2-\alpha}}{2r^{1+\eta}}.$$

This, together with (4.3) and (4.4) implies that

$$\sigma_\eta(r) \le \frac{1}{2r} J^\alpha(\kappa \wedge r)(\kappa \wedge r)^2, \quad r \in [0, 2\ell_0]. \tag{4.5}$$

Let

$$g_\eta(r) = \left(1 + \frac{K_1}{K_2}\right) \int_0^r \frac{1}{\sigma_\eta(s)} \mathrm{d}s, \quad r \in [0, 2\ell_0], \ c_1 = \mathrm{e}^{-2K_2 g_\eta(2\ell_0)}. \tag{4.6}$$

Theorem 4.1. *Assume **(B1)**-**(B2)**. Let $\mu_0 \in \mathscr{P}_{1+\delta}(\mathbb{R}^d)$ for some $\delta \in (0, \alpha - 1)$ and $\mu_0^N \in \mathscr{P}_1((\mathbb{R}^d)^N)$ be exchangeable. If $K_b < \frac{2c_1^2 K_2}{(1+c_1)^2}$, then there exist positive constants c and λ such that*

$$\begin{aligned} &\|(P_t^{[k],N})^* \mu_0^N - (P_t^* \mu_0)^{\otimes k}\|_{TV} \\ &\le k c \mathrm{e}^{-\lambda t} \frac{\mathbb{W}_{\ell_1}(\mu_0^N, \mu_0^{\otimes N})}{N} + c\{1 + \|\mu_0\|_{1+\delta}\} k N^{-\frac{\delta}{1+\delta}}, \quad 1 \le k \le N, t \ge 1. \end{aligned} \tag{4.7}$$

Remark 4.2. (1) The condition $\mathbb{E}|Z_t^i|^2 < \infty$ in (Liang, Majka and Wang, 2021, Theorem 1.2) is no longer required in Theorem 4.1, which is attributed to Lemma S2 in Huang, Yang and Yuan (2026).

(2) One may try to adopt the entropy method in Bresch, Jabin and Wang (2023), Jabin and Wang (2018, 2016) to derive the PoC in relative entropy in α-stable noise case. However, there exist essential difficulties. Let us illustrate it in detail. Let $n = d$, $b^i : \mathbb{R}^d \to \mathbb{R}^d, i = 1, 2$ be measurable. Denote by P_t^i the associated semigroup to the generator

$$\mathscr{L}^i = \langle b^i, \nabla \rangle - (-\Delta)^{\frac{\alpha}{2}}, \quad i = 1, 2.$$

By formal calculation and the Kolmogorov forward equation for P_t^1 and the Kolmogorov backward equation for P_t^2, for a smooth and positive f, it holds that

$$\begin{aligned} P_t^1 \log(f) - \log(P_t^2 f) &= \int_0^t \left(\frac{\mathrm{d} P_s^1 \log(P_{t-s}^2 f)}{\mathrm{d}s} \right) \mathrm{d}s \\ &= \int_0^t [P_s^1 \{(P_{t-s}^2 f)^{-1} \langle b^1 - b^2, \nabla P_{t-s}^2 f \rangle\}] \mathrm{d}s \\ &+ \int_0^t \left[-P_s^1 (-\Delta)^{\frac{\alpha}{2}} \{\log(P_{t-s}^2 f)\} + P_s^1 \left\{ \frac{(-\Delta)^{\frac{\alpha}{2}} P_{t-s}^2 f}{P_{t-s}^2 f} \right\} \right] \mathrm{d}s. \end{aligned} \tag{4.8}$$

When $\alpha = 2$, (4.8), together with the chain rule

$$\Delta\{\log(f)\} = -f^{-2} |\nabla f|^2 + f^{-1} \Delta f \tag{4.9}$$

and the Cauchy-Schwartz inequality, yields that

$$P_t^1 \log f - \log(P_t^2 f) \le \frac{1}{2} \int_0^t P_s^1 (|b^1 - b^2|^2) \mathrm{d}s.$$

This inequality coincides with the estimate derived by the Girsanov transform. One can also refer to Ren and Wang (2025) for the entropy estimates of two diffusion processes with different drifts and diffusion coefficients.

Unfortunately, when $\alpha \in (0,2)$, $(-\Delta)^{\frac{\alpha}{2}}\{\log(f)\}$ is not so explicit as (4.9). Hence, it seems challenging to derive an entropy estimate for two semigroups with different drifts in the α-stable noise case. To our best knowledge, the PoC in relative entropy in the Lévy noise case, even in the case of α-stable noise, remains an open problem.

Proof of Theorem 4.1. Similar to the proof of Theorem 1.1, by the Yosida approximation in (Wang and Wang, 2014, part (c) of proof of Theorem 2.1) and (3.3), it is sufficient to prove (4.7) for Lipschitz continuous $b^{(0)}$. So, in the following, we assume that $b^{(0)}$ is Lipschitz continuous and we will verify conditions (i)-(iv) in Theorem 2.1 one by one. We should remark that the proof of conditions (i)-(iii) is similar to that of Theorem 1.1.

(1) Take $(X_0^{i,N})_{1\le i\le N}$ and $(X_0^i)_{1\le i\le N}$ such that $\mathscr{L}_{(X_0^{i,N})_{1\le i\le N}} = \mu_0^N$ and $\mathscr{L}_{(X_0^i)_{1\le i\le N}} = \mu_0^{\otimes N}$. Fix $t > 0$. Recall that $P_{s,t}^\mu$ and $(P_{s,t}^\mu)^{\otimes k}$ are defined in (2.4) and (2.6) respectively. Since $b^{(0)}$ and $b^{(1)}$ are Lipschitz continuous, the backward Kolmogorov equation

$$\frac{\mathrm{d}P_{s,t}^\mu f}{\mathrm{d}s} = -\mathscr{L}_s^\mu P_{s,t}^\mu f, \quad f \in C_b^2(\mathbb{R}^d), \|f\|_\infty \le 1 \tag{4.10}$$

holds with

$$\mathscr{L}_s^\mu = \left\langle b^{(0)} + \int_{\mathbb{R}^d} b^{(1)}(\cdot, y)\mu_s(\mathrm{d}y), \nabla \right\rangle - (-\Delta)^{\frac{\alpha}{2}}.$$

For any $F \in C_b^2((\mathbb{R}^d)^k)$ with $\|F\|_\infty \le 1$, $1 \le i \le k$, $x = (x^1, \cdots, x^k) \in (\mathbb{R}^d)^k$, denote

$$-(-\Delta_i)^{\frac{\alpha}{2}} F(x) = \int_{\mathbb{R}^d} \{F(x^1, \cdots, x^i + z, \cdots, x^k) - F(x) - \langle \nabla_i F(x), z\rangle 1_{\{|z|\le 1\}}\}\nu^\alpha(\mathrm{d}z),$$

and define

$$(\mathscr{L}_s^\mu)^{\otimes k} F(x) = \sum_{i=1}^k \left\{ \left\langle b^{(0)}(x^i) + \int_{\mathbb{R}^d} b^{(1)}(x^i, y)\mu_s(\mathrm{d}y), \nabla_i F(x) \right\rangle - (-\Delta_i)^{\frac{\alpha}{2}} F(x) \right\}.$$

Repeating the argument to derive (3.6) from (3.5), it follows from (4.10) that

$$\frac{\mathrm{d}(P_{s,t}^\mu)^{\otimes k} F}{\mathrm{d}s} = -(\mathscr{L}_s^\mu)^{\otimes k} (P_{s,t}^\mu)^{\otimes k} F, \quad F \in C_b^2((\mathbb{R}^d)^k), \|F\|_\infty \le 1, s \in [0,t].$$

This together with the procedure to derive (3.7) from (3.6) implies (i) in Theorem 2.1.

(2) By (Wang and Wang, 2014, Corollary 2.2(2)), there exists a constant $c_0 > 0$ independent of the Lipschitz constant of $b^{(0)}$ such that

$$|\nabla P_{r,t}^\mu f| \le c_0 \frac{1}{(t-r)^{1/\alpha} \wedge 1} \|f\|_\infty, \quad 0 \le r < t, f \in \mathscr{B}_b(\mathbb{R}^d).$$

This means that (ii) in Theorem 2.1 holds.

(3) Note that (4.1) implies that

$$\langle x_1 - x_2, b^{(0)}(x_1) - b^{(0)}(x_2)\rangle \le K_1 |x_1 - x_2|^2. \tag{4.11}$$

It is standard to derive from (4.11) and **(B1)-(B2)** that

$$\mathbb{E}((1 + |X_t^1|^2)^{\frac{1+\delta}{2}}) \le c_0(t)\mu_0(1 + |\cdot|^{1+\delta}), \quad t \ge 0 \tag{4.12}$$

for some increasing function $c_0 : [0, \infty) \to [0, \infty)$. Let $Z_t^{i,N} = X_t^i - X_t^{i,N}$. It follows from Itô's formula and **(B2)** that

$$\mathrm{d}|Z_t^{i,N}| \le K_1|Z_t^{i,N}|\mathrm{d}t + K_b|Z_t^{i,N}|\mathrm{d}t + \frac{1}{N}\sum_{m=1}^{N} K_b|Z_t^{m,N}|\mathrm{d}t + \left|\frac{1}{N}\sum_{m=1}^{N} b^{(1)}(X_t^i, X_t^m) - \int_{\mathbb{R}^d} b^{(1)}(X_t^i, y)\mu_t(\mathrm{d}y)\right|\mathrm{d}t.$$

Applying Gronwall's inequality, Lemma S2 in Huang, Yang and Yuan (2026) and (4.12), we get

$$\sum_{i=1}^{N} \mathbb{E}|Z_s^{i,N}| \le \mathrm{e}^{(K_1+2K_b)s}\sum_{i=1}^{N} \mathbb{E}|Z_0^{i,N}| + c(s)\{1 + \|\mu_0\|_{1+\delta}\}NN^{-\frac{\delta}{1+\delta}}$$

for some increasing function $c : [0, \infty) \to [0, \infty)$. By the same argument to derive (3.13) from (3.12), (iii) in Theorem 2.1 holds.

(4) By Lemma S4 in Huang, Yang and Yuan (2026), we derive the condition (iv) in Theorem 2.1 and the proof is completed. □

Funding

The first author was supported by National Key R&D Program of China (No. 2022YFA1006000) and NNSFC (12271398). The second author was supported by NNSFC (12101390, 12426656).

Supplementary Material

Supplement to "Long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos for mean field interacting particle system" (DOI: 10.3150/25-BEJ1913SUPP; .pdf). Some lemmas for the proofs of main results are postponed and carried out in a supplementary paper.

References

Bao, J. and Wang, J. (2025). Long time behavior of one-dimensional McKean-Vlasov SDEs with common noise. *J. Math. Anal. Appl.* **552** 129819. MR4925916 https://doi.org/10.1016/j.jmaa.2025.129819

Bresch, D., Jabin, P.-E. and Wang, Z. (2023). Mean-field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.* **172** 2591–2641. MR4658923 https://doi.org/10.1215/00127094-2022-0088

Chen, F., Lin, Y., Ren, Z. and Wang, S. (2024). Uniform-in-time propagation of chaos for kinetic mean field Langevin dynamics. *Electron. J. Probab.* **29** 1–43. MR4701825 https://doi.org/10.1214/24-EJP1079

Durmus, A., Eberle, A., Guillin, A. and Zimmer, R. (2020). An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** 5387–5398. MR4163850 https://doi.org/10.1090/proc/14612

Huang, X., Yang, F.-F. and Yuan, C. (2026). Supplement to "Long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos for mean field interacting particle system." https://doi.org/10.3150/25-BEJ1913SUPP

Jabin, P.-E. and Wang, Z. (2016). Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.* **271** 3588–3627. MR3558251 https://doi.org/10.1016/j.jfa.2016.09.014

Jabin, P.-E. and Wang, Z. (2018). Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.* **214** 523–591. MR3858403 https://doi.org/10.1007/s00222-018-0808-y

Kac, M. (1954–1955). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. III* 171–197. University of California Press. MR0084985

Lacker, D. (2023). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.* **4** 377–432. MR4595391 https://doi.org/10.2140/pmp.2023.4.377

Lacker, D. and Flem, L.L. (2023). Sharp uniform-in-time propagation of chaos. *Probab. Theory Related Fields* **187** 443–480. MR4634344 https://doi.org/10.1007/s00440-023-01192-x

Liang, M., Majka, M.B. and Wang, J. (2021). Exponential ergodicity for SDEs and McKean-Vlasov processes with Lévy noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1665–1701. MR4291453 https://doi.org/10.1214/20-aihp1123

Liu, W., Wu, L. and Zhang, C. (2021). Long-time behaviors of mean-field interacting particle systems related to McKean-Vlasov equations. *Comm. Math. Phys.* **387** 179–214. MR4312362 https://doi.org/10.1007/s00220-021-04198-5

Lukkarinen, J. (2023). Generation and propagation of chaos in the stochastic Kac model. Talk at Rutgers.

Luo, D. and Wang, J. (2019). Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stoch. Process. Appl.* **129** 3129–3173. MR3985558 https://doi.org/10.1016/j.spa.2018.09.003

McKean, H.P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907

McKean, H.P. and Singer, I.M. (1967). Curvature and the eigenvalues of the Laplacian. *J. Differential Geom.* **1** 43–69. MR0217739

Monmarché, P., Ren, Z. and Wang, S. (2024). Time-uniform log-Sobolev inequalities and applications to propagation of chaos. *Electron. J. Probab.* **29** 1–38. MR4816013

Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236** 244–264. MR2227134 https://doi.org/10.1016/j.jfa.2005.12.010

Ren, P. and Wang, F.-Y. (2025). Bi-coupling method and applications. *Probab. Theory Related Fields*. https://doi.org/10.1007/s00440-025-01394-5

Rosenzweig, M. and Serfaty, S. (2025). Modulated logarithmic Sobolev inequalities and generation of chaos. *Ann. Fac. Sci. Toulouse Math. (6)* **34** 107–134. https://doi.org/10.5802/afst.1807

Song, Y. (2020). Gradient estimates and exponential ergodicity for mean-field SDEs with jumps. *J. Theoret. Probab.* **33** 201–238. MR4064299 https://doi.org/10.1007/s10959-018-0845-x

Sznitman, A.-S. (1991). Topics in propagation of chaos. In *"École d'Été de Probabilités de Sain-Flour XIX-1989". Lecture Notes in Mathematics* **1464** 165–251. Berlin: Springer. MR1108185 https://doi.org/10.1007/BFb0085169

Wang, F.-Y. (2011). Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds. *Ann. Probab.* **39** 1449–1467. MR2857246 https://doi.org/10.1214/10-AOP600

Wang, F.-Y. (2015). Asymptotic couplings by reflection and applications for nonlinear monotone SPDES. *Nonlinear Anal.* **117** 169–188. MR3316613 https://doi.org/10.1016/j.na.2015.01.012

Wang, F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stoch. Process. Appl.* **128** 595–621. MR3739509 https://doi.org/10.1016/j.spa.2017.05.006

Wang, F.-Y. (2023). Exponential ergodicity for non-dissipative McKean-Vlasov SDEs. *Bernoulli* **29** 1035–1062. MR4550214 https://doi.org/10.3150/22-bej1489

Wang, F.-Y. and Wang, J. (2014). Harnack inequalities for stochastic equations driven by Lévy noise. *J. Math. Anal. Appl.* **410** 513–523. MR3109860 https://doi.org/10.1016/j.jmaa.2013.08.013

Received July 2024 and revised July 2025