Mean field analysis of interacting network model with jumps
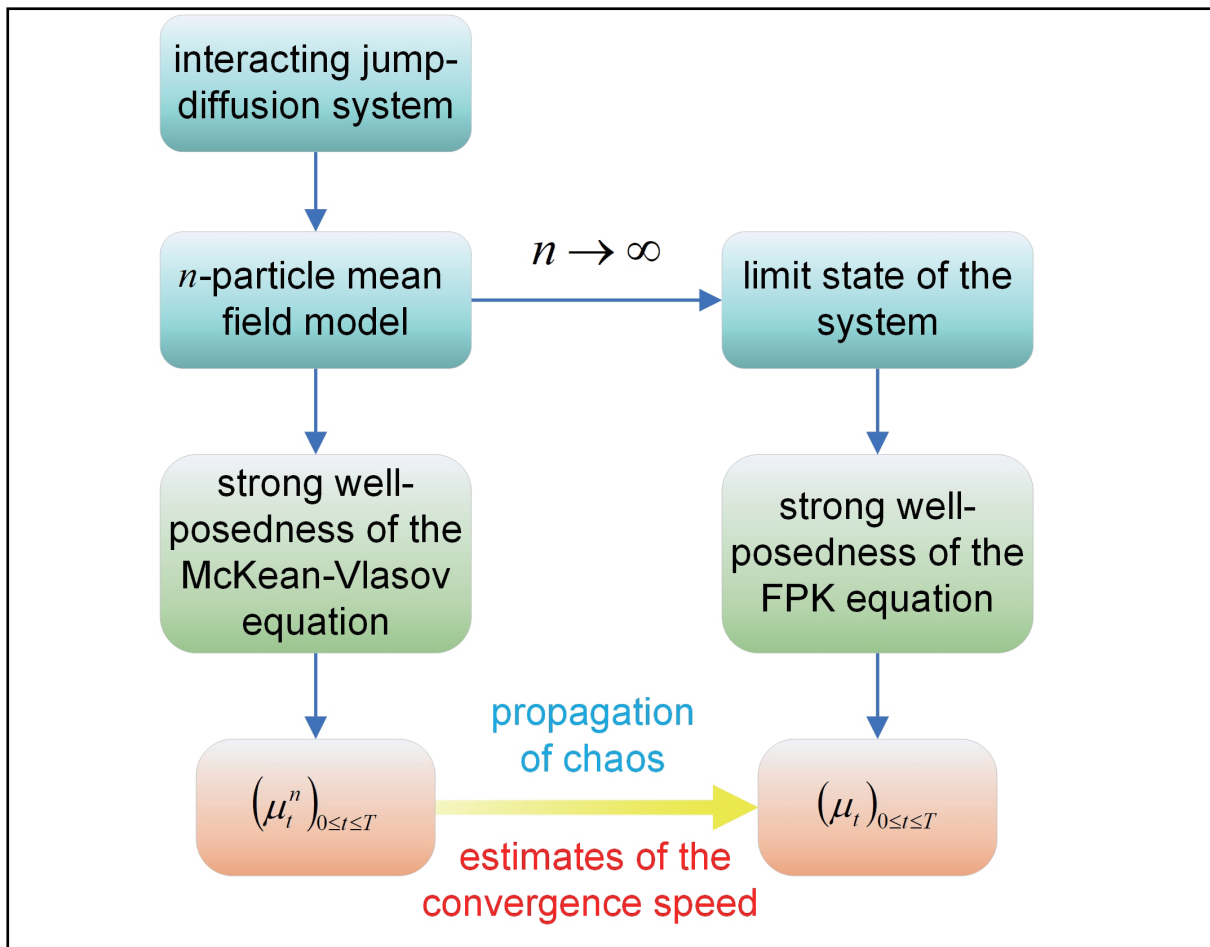
Zeqian Li ✉

School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Zeqian Li, E-mail: lzq7890@mail.ustc.edu.cn

Graphical abstract

An estimate on the speed of the convergence of the empirical measure to the limit measure is given with respect to a Wasserstein-like distance.

Public summary

■ Constructing the n-particle system with mean filed interaction, whose dynamic follows a couple of SDEs with jumps.

■ Studying the existence of a solution to the corresponding McKean-Vlasov limit equation.

■ Establishing the propagation of chaos under a proper metric, giving an estimate of the speed of the convergence and proving the uniqueness of the solution to the McKean-Vlasov limit equation.

Mean field analysis of interacting network model with jumps

Zeqian Li ✉

School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Zeqian Li, E-mail: lzq7890@mail.ustc.edu.cn

Abstract: This paper considers an n-particle jump-diffusion system with mean filed interaction, where the coefficients are locally Lipschitz continuous. We address the convergence as $n \to \infty$ of the empirical measure of the jump-diffusions to the solution of a deterministic McKean–Vlasov equation. The strong well-posedness of the associated McKean–Vlasov equation and a corresponding propagation of chaos result are proven. In particular, we also provide precise estimates of the convergence speed with respect to a Wasserstein-like metric.

Keywords: Mean-field theory; McKean–Vlasov equations; FPK equations; propagation of chaos; particle system

CLC number: O211.63   **Document code:** A

2020 Mathematics Subject Classification: 60K35; 92B20; 60J60; 60F99

1 Introduction

The interaction of particle systems is commonly observed in both physical and biological science. The behavior of such systems as the number of particles strength approaches infinity has sparked widespread interest, and the related study has been a long-standing issue. For instance, the interacting particles can stand for the stars in a galaxy[1], the units of a neural network[2] or the animals of a cluster[3].

Mean field analysis is utilized to investigate systems that comprise a large number of interacting elements. When the scale of the particle system approaches infinity, the behavior of each element is assumed to be independent of others except for an average effect caused by the collective behavior of the system as a whole. Mean field analysis is widely employed in various fields, for example, describing nuclear structure and low-energy dynamics[4], simulating bio-inspired neuronal networks[5, 6], and modeling interbank lending and borrowing activities[7]. And its mathematical properties have been studied from various perspectives in recent years (see Refs. [8, 9] for example). To describe such a mean field phenomenon, the following mathematical framework is considered: the dynamics of the system with n particles are governed by n SDEs, and the mean field interactions can be expressed as a dependence on the empirical measure of the coefficients related to the equation. As n tends to infinity, the dynamic of the limit system follows the McKean–Vlasov equation where the coefficients exhibit a natural dependence on the probability distribution associated with its solution.

The convergence result for the empirical measure with respect to the n-particle system is called the propagation of chaos property. The concept of propagation of chaos was first proposed by Ref. [10] to describe the asymptotic independence between particles that arises as the spatially extension tends to infinity. This convergence implies that the statistical properties of the system can be effectively described by the limiting measure, allowing for simplifications in the analysis and modeling of the system. For a more comprehensive overview, one can refer to Ref. [11].

Since the seminal work Ref. [12], there have been various studies on well-posedness of the McKean–Vlasov equations and the propagation of chaos for the mean field model without jumps, even under a less strict hypothesis than the Lipschitz continuity of the coefficients. For example, under relaxed regularity conditions, Ref. [13] is a recent work that establishes weak and strong well-posedness results and Ref. [14] considers also the propagation of chaos. Ref. [15] investigates the uniform-in-time propagation of chaos for the mean-field weakly interacting particle system.

In this paper, we focus on the mean field coupled jump-diffusion particle system and its limit McKean–Vlasov equation. The jump term is modeled as a double stochastic Poisson process with the state-dependent intensity. The coefficients of corresponding McKean–Vlasov equation are locally Lipschitz continuous and satisfy a growth condition. We establish the strong well-posedness of the relevant McKean–Vlasov equations and prove a convergence result for the particle system as it extends spatially.

The first novelty of this paper is that we prove the strong well-posedness of the solution to the McKean–Vlasov equation under our framework. This result of the model with jumps has already been studied under globally Lipschitz continuous assumptions on the coefficients[16]. For the locally Lipschitz case, it becomes more challenging. Ref. [17] uses Osgood's lemma instead of Grönwall's lemma to deal with the locally Lipschitz case, and Ref. [18] applies the Euler approximation to construct a solution of the McKean–Vlasov equation. In this paper, we first establish the existence of a local (strong) solution by a classical truncation argument, then we provide a criterion using Lyapunov functions and thereby prove that under the given growth conditions, such solution is a global one. As an example of our locally

Lipschitz system, we show that a biological science model called FitzHugh-Nagumo neuron networks (which was first proposed by Ref. [19], with clarification note by Ref. [20]) falls into our model.

The second novelty of our paper is that we establish the propagation of chaos property and estimate the order of the convergence rate of our model. Under continuous frameworks, Refs.[21, 22] study the convergence rate. Refs.[17, 18] prove the propagation of chaos with respect to the jump model under locally Lipschitz assumption, but without a conclusion on the convergence rate. In this paper, to prove the smoothness of the propagator, we study the well-posedness of an integro-differential parabolic equation. With the help of this smoothness property, we ultimately provide an estimate for the order of the convergence rate with respect to a Wasserstein-like distance.

The paper is organized as follows: we introduce in Section 2 the n-particle system with mean filed interaction, whose dynamics follow a couple of SDEs with jumps. In Section 3, we study the existence of a solution to the corresponding McKean–Vlasov limit equation. Section 4 establishes the propagation of chaos under a proper metric, gives an estimate of the speed of convergence and proves the uniqueness of the solution to the McKean–Vlasov limit equation.

2 Model

2.1 Jump diffusion model

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual conditions. We consider an interacting system involving n SDEs which is described as follows: for $i = 1, \cdots, n$,

$$\mathrm{d}X_t^i = f(t, X_t)\mathrm{d}t + g(t, X_t)\mathrm{d}W_t^i + h(t, X_{t-}^i)\mathrm{d}N_t^i, \quad X_0^i \in \mathbb{R}^m, \quad (1)$$

where $X_t = (X_t^1, \cdots, X_t^n)$ for $t \in [0, T]$ and $W^i = (W_t^i)_{t\in[0,T]}$ for $i = 1, \cdots, n$ are n-independent (d-dimensional) Brownian motions, and for each $i = 1, \cdots, n$, $N^i = (N_t^i)_{t\in[0,T]}$ is a double stochastic Poisson processes with the state-dependent intensity given by the positive functions $\lambda_i(t, X_{t-})$ for $t \in [0, T]$. In other words, it holds that

$$M_t^i := N_t^i - \int_0^t \lambda_i(s, X_{s-})\mathrm{d}s, \quad t \in [0, T], \quad (2)$$

is a $(\mathbb{P}, \mathbb{F})$-martingale.

We impose the following assumption on the coefficients (f, g, h) of the system (1), so that this system admits a unique strong solution $X = (X_t)_{t\in[0,T]}$ which takes value on $\mathbb{R}^{mn}$.

$(\mathbf{A}_{fgh})$ The coefficients $f : [0, T] \times \mathbb{R}^{mn} \to \mathbb{R}^m$, $g : [0, T] \times \mathbb{R}^{mn} \to \mathbb{R}^{md}$ and $h : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ satisfy the uniformly locally Lipschitz condition. Namely, for any $R > 0$, there exists a constant $L_R > 0$ such that for all $t \in [0, T]$ and $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n) \in B_R(0) := \{x \in \mathbb{R}^{mn}; |x| \leqslant R\}$,

$$|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| + |h(t, x_i) - h(t, y_i)| \leqslant L_R|x - y|,$$
$$\forall i = 1, \cdots, n,$$

where $|\cdot|$ denotes the Euclidean norm.

Under a truncation argument, the condition $(\mathbf{A}_{fgh})$ can only guarantee the existence and uniqueness of local (strong) solu-

tions of (1). In order to establish the global solution, we further impose the condition on the existence of a Lyapunov function associated to the system (1) and give the following lemma:

Lemma 1. Let the assumption $(\mathbf{A}_{fgh})$ holds, also suppose the following:

$(\mathbf{A}_{Lya})$ There exists a function $V : \mathbb{R}^{mn} \to \mathbb{R}_+ := (0, +\infty)$ satisfying that

(i) Let $q_R := \inf_{|x|>R} V(x)$, then $\lim_{R\to\infty} q_R = +\infty$;

(ii) $\mathbb{E}[V(X_0)] < \infty$;

(iii) There exists a constant $C > 0$, such that for all $t \in [0, T]$,

$$\mathbb{E}[V(X_{t\wedge\tau_R})] \leqslant \mathbb{E}[V(X_0)] + C \int_0^t (1 + \mathbb{E}[V(X_{s\wedge\tau_R})])\,\mathrm{d}s,$$

where $\tau_R := \inf\{t \in [0, T); |X_t| \geqslant R\}$, $\inf \emptyset := T$.

Then there exists a unique strong solution $X = (X_t)_{t\in[0,T]}$ of the system (1) with an initial condition satisfying $\mathbb{E}[|X_0|^2] < \infty$.

Proof. By applying Theorem 9.1 of chapter IV of Ref. [23], the condition $(\mathbf{A}_{fgh})$ guarantees the existence and uniqueness of local (strong) solutions to (1). In other words, we have the solution of (1) on the time interval $[0, t \wedge \tau_R]$ for all $t \in [0, T]$ and $R > 0$. Thus, it is enough to prove that $\tau_R \to T$ as $R \to \infty$, $\mathbb{P}$-a.s. Note that the condition (iii) of $(\mathbf{A}_{Lya})$ yields that for all $t \in [0, T]$ and $R > 0$, $\mathbb{E}[V(X_{t\wedge\tau_R})] \leqslant e^{Ct}\{1 + \mathbb{E}[V(X_0)]\}$. Therefore, we have for all $t \in [0, T]$,

$$\mathbb{P}(\tau_R < t) \leqslant \frac{1}{q_R}\mathbb{E}[\mathbb{1}_{\tau_R < t}V(X_{\tau_R})] = $$
$$\frac{1}{q_R}\mathbb{E}[\mathbb{1}_{\tau_R < t}V(X_{t\wedge\tau_R})] \leqslant \frac{1}{q_R}e^{Ct}(1 + \mathbb{E}[V(X_0)]), \quad (3)$$

and hence $\mathbb{P}(\tau_R < t) \to 0$ as $R \to \infty$ by the condition (i) of $(\mathbf{A}_{Lya})$, which gives the desired result.

We next give a condition on the growth of the coefficients of the system (1):

$(\mathbf{A}_{grow})$ The coefficients, $f : [0, T] \times \mathbb{R}^{mn} \to \mathbb{R}^m$, $g : [0, T] \times \mathbb{R}^{mn} \to \mathbb{R}^{md}$, $h : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$, and $\lambda_i : [0, T] \times \mathbb{R}^{mn} \to \mathbb{R}_+$, satisfy the following growth condition: there exists a constant $K > 0$ such that for all $t \in [0, T]$ and $x = (x_1, \cdots, x_n) \in \mathbb{R}^{mn}$,

$$2\sum_{i=1}^n x_i^\top f(t, x) + \mathrm{tr}[gg^\top(t, x)] + $$
$$\sum_{i=1}^n \lambda_i(t, x)(|h(t, x_i)|^2 + 2x_i^\top h(t, x_i)) \leqslant K(1 + |x|^2).$$

It is not difficult to verify that if $\mathbb{E}[|X_0|^2] < \infty$, then $V(x) = |x|^2$ for $x \in \mathbb{R}^{mn}$ satisfies the conditions (i)–(iii) of $(\mathbf{A}_{Lya})$ under the assumption $(\mathbf{A}_{grow})$. Then, Lemma 1 yields that:

Lemma 2. Let assumptions $(\mathbf{A}_{fgh})$ and $(\mathbf{A}_{grow})$ hold. Then there exists a unique strong solution $X = (X_t)_{t\in[0,T]}$ of the system (1) with an initial condition satisfying $\mathbb{E}[|X_0|^2] < \infty$.

2.2 Mean field model

The particle system considered in this paper is given by the following mean field model:

$$\mathrm{d}X_t^i = \overline{f}\left(t, X_t^i, \frac{1}{n}\sum_{j=1}^{n}\alpha(X_t^i, X_t^j)\right)\mathrm{d}t +$$
$$\overline{g}\left(t, X_t^i, \frac{1}{n}\sum_{j=1}^{n}\beta(X_t^i, X_t^j)\right)\mathrm{d}W_t^i +$$
$$h(t, X_{t-}^i)\mathrm{d}N_t^i, \tag{4}$$

for $i = 1, \cdots, n$, where $\alpha: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $\beta: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, $\overline{f}: [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$, $\overline{g}: [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{md}$ and $h: [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ are all measurable functions.

Let us define the following functions: For $x = (x^1, \cdots, x^n) \in \mathbb{R}^{mn}$,

$$f_i(t, x) := \overline{f}\left(t, x^i, \frac{1}{n}\sum_{j=1}^{n}\alpha(x^i, x^j)\right),$$
$$g_i(t, x) := \overline{g}\left(t, x^i, \frac{1}{n}\sum_{j=1}^{n}\beta(x^i, x^j)\right). \tag{5}$$

If (f_i, g_i, h) above satisfies assumptions $(\mathbf{A}_{fgh})$ and $(\mathbf{A}_{grow})$, then by Lemma 2 the system (4) admits a unique strong solution $X = (X^1, \cdots, X^n)$ provided that $\mathbb{E}[|X_0|^2] < \infty$.

Example 1. Ref. [19] address the model of the network of FitzHugh-Nagumo neurons with simple maximum conductance variation, and such model is also investigated by Refs. [24–26]. However, the aforementioned studies predominantly concentrate on research conducted within the framework of diffusion state equation models. We next provide a verification that the network model with one population in Ref. [19] falls into our jump-diffusion system (4). To this purpose, let $m = 3$ and $d = 3$. For $i = 1, \cdots, n$ and $x^i = (x_1^i, x_2^i, x_3^i) \in \mathbb{R} \times \mathbb{R} \times [0,1]$, define the coefficients as $\alpha(x^i, x^j) := -J(x_1^i - V_{rev})x_3^j$ and $\beta(x^i, x^j) := -\sigma(x_1^i - V_{rev})x_3^j$, where $J, V_{rev}, \sigma \in \mathbb{R}_+$. For any $(t, x^i, z) \in [0,T] \times \mathbb{R}^3 \times \mathbb{R}$, we introduce the following functions:

$$\overline{f}(t, x^i, z) = \begin{bmatrix} x_1^i - \frac{(x_1^i)^3}{3} - x_2^i + I_t + z \\ c(x_1^i + a - bx_2^i) \\ \frac{a_rT}{1+e^{-\theta(x_1^i - V)}}(1 - x_3^i) - a_d x_3^i \end{bmatrix}, \tag{6}$$

$$\overline{g}(t, x^i, z) = \begin{bmatrix} \sigma_{ext} & 0 & z \\ 0 & 0 & 0 \\ 0 & 0 & \chi(x_3^i)\sqrt{\frac{a_rT}{1+e^{-\theta(x_1^i - V)}}(1 - x_3^i) + a_d x_3^i} \end{bmatrix}, \tag{7}$$

where $\chi(x_3^i) = 0.1e^{-\frac{1}{2(1-(2x_3^i - 1)^2)}}$ and $I: [0,T] \to \mathbb{R}$ is a deterministic bounded function. In addition, V, a, b, c, a_r, a_d and σ_{ext} are all positive constants.

Recall the functions f, g defined by (5) with $\overline{f}, \overline{g}$ and α, β given as above. Note that $\chi: \mathbb{R} \to \mathbb{R}_+$ is Lipschitz continuous. Then, it is not difficult to verify that f, g satisfy assumptions $(\mathbf{A}_{fgh})$ and $(\mathbf{A}_{grow})$ with $h \equiv 0$. Let us further impose the following condition:

$(\mathbf{A}_{h\lambda})$ The intensity function $\lambda_i: [0,T] \times \mathbb{R}^{mn} \to \mathbb{R}^m$ of the ith state and the jump function $h: [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ satisfy that:

(i) For any $R > 0$, there exists a constant $L_R > 0$ which is independent of t, such that for all $t \in [0,T]$ and $x = (x^1, \cdots, x^n)$, $y = (y^1, \cdots, y^n) \in B_R(0) := \{x \in \mathbb{R}^{mn}; |x| \leqslant R\}$ (where $x^i, y^i \in \mathbb{R}^m$),

$$|h(t, x^i) - h(t, y^i)| \leqslant L_R |x^i - y^i|, \quad \forall i = 1, \cdots, n;$$

(ii) For $i = 1, \cdots, n$, $\lambda_i(t, x) = p_i\psi(t, \overline{x})$, where $\psi: [0,T] \times \mathbb{R}^m \to \mathbb{R}_+$ is measurable and $\overline{x} := \frac{1}{n}\sum_{i=1}^{n}x^i$.

(iii) There exists a constant $K > 0$ such that for all $t \in [0,T]$ and $x = (x^1, \cdots, x^n) \in \mathbb{R}^{mn}$ with $x^i \in \mathbb{R}^m$,

$$\sum_{i=1}^{n} p_i\psi(t, \overline{x})(|h(t, x^i)|^2 + 2x^{i\top}h(t, x^i)) \leqslant K(1 + |x|^2).$$

Therefore, we have established a jump-diffusion model of the network of FitzHugh-Nagumo neurons with simple maximum conductance variation as an extension to the one in Ref. [19], where the jump process models an intense response of the neuron when it faces an external environmental impact, allowing for a more realistic representation of spiking activity and the integration of multiple inputs within a neuron.

3 Mean field analysis

In the following sections, we only consider the case that the intensity function of the i-th state is given by the form $\lambda_i(t, x) = p_i\psi(t, \overline{x})$ where $p_i \in \mathbb{R}_+$ and $\psi: [0,T] \times \mathbb{R}^m \to \mathbb{R}_+$ is a bounded measurable function with upper bound C_ψ. We introduce the empirical measure-value process as

$$\mu_t^n(\mathrm{d}p, \mathrm{d}x) := \frac{1}{n}\sum_{i=1}^{n}\delta_{(p_i, X_t^i)}(\mathrm{d}p, \mathrm{d}x), \tag{8}$$

where δ denotes the Dirac delta measure, $(p, x) \in E := \mathbb{R}_+ \times \mathbb{R}^m$ and $\mu_0^n(\mathrm{d}p, \mathrm{d}x_0)$ is the empirical measure related to the initial state. We propose the following assumption:

$(\mathbf{A}_{ini})$ There exists $p \in \mathbb{R}$, such that $p_i \to p$ as $i \to \infty$. The sequence of r.v.s $(X_0^i)_{i \geqslant 1}$ is i.i.d. and satisfies that $\mathbb{E}[|X_0^i|^2] < \infty$ for all $i \geqslant 1$.

Let $\mathcal{P}(E)$ be the set of all Borel probability measures on E. Under assumption $(\mathbf{A}_{ini})$, we know that there exists a probability measure $\rho_0 \in \mathcal{P}(E)$ such that $\mu_0^n \Rightarrow \rho_0$ in the weak sense.

Assume $\phi \in C^2(E)$ is a real-valued twice continuously differentiable function, let $\partial_x\phi := (\partial_{x_i}\phi)_{i=1,\cdots,n}^\top$ be the gradient of ϕ and $\partial_x^2\phi$ be the corresponding Hessian matrix. By (8), it holds for $s \in [0,T]$ that

$$\langle \mu_s^n, \alpha(x, \cdot)\rangle = \frac{1}{n}\sum_{j=1}^{n}\alpha(x, X_s^j), \quad \langle \mu_s^n, \beta(x, \cdot)\rangle = \frac{1}{n}\sum_{j=1}^{n}\beta(x, X_s^j). \tag{9}$$

Then, for $i = 1, \cdots, n$ and $t \in [0,T]$, Itô's formula yields that

$$\phi(p_i, X_t^i) = \phi(p_i, X_0^i) + \int_0^t \partial_x\phi(p_i, X_s^i)^\top \overline{f}(s, X_s^i, \langle \mu_s^n, \alpha(X_s^i, \cdot)\rangle)\mathrm{d}s +$$
$$\frac{1}{2}\int_0^t \mathrm{tr}\left[\partial_x^2\phi(p_i, X_s^i)\overline{g}\overline{g}^\top(s, X_s^i, \langle \mu_s^n, \beta(X_s^i, \cdot)\rangle)\right]\mathrm{d}s +$$
$$\int_0^t p_i(\phi(p_i, X_s^i + h(s, X_s^i)) - \phi(p_i, X_s^i))\psi(s, \overline{X}_s)\mathrm{d}s +$$
$$\int_0^t \partial_x\phi(p_i, X_s^i)^\top \overline{g}(s, X_s^i, \langle \mu_s^n, \beta(X_s^i, \cdot)\rangle)\mathrm{d}W_s^i +$$
$$\int_0^t (\phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - \phi(p_i, X_{s-}^i))\mathrm{d}M_s^i. \tag{10}$$

Define the following operators as: for $(t, p, x) \in [0,T] \times E$ and $\nu \in \mathcal{P}(E)$,

$$
\begin{cases}
\mathcal{L}_c^{t,\nu} \phi(p,x) := \\
\quad \partial_x \phi(p,x)^{\top} \overline{f}(t,x,\langle \nu, \alpha(x,\cdot) \rangle) + \\
\quad \frac{1}{2} \mathrm{tr}\left[\partial_x^2 \phi(p,x) \overline{gg}^{\top}(t,x,\langle \nu, \beta(x,\cdot) \rangle)\right], \\
\\
\mathcal{L}_j^{t,\nu} \phi(p,x) := \\
\quad p\left(\phi(p, x+h(t,x)) - \phi(p,x)\right)\psi(t,\langle \nu, I \rangle),
\end{cases}
\tag{11}
$$

where $I(p,x) = x$. Therefore, it holds that

$$
\begin{aligned}
\langle \mu_t^n, \phi \rangle = {}& \langle \mu_0^n, \phi \rangle + \int_0^t \left\langle \mu_s^n, \mathcal{L}_c^{s,\mu_s^n} \phi + \mathcal{L}_j^{s,\mu_s^n} \phi \right\rangle \mathrm{d}s \\
& + \frac{1}{n} \sum_{i=1}^n \int_0^t \partial_x \phi(p_i, X_s^i) \overline{g}(s, X_s^i, \langle \mu_s^n, \beta(X_s^i, \cdot) \rangle) \mathrm{d}W_s^i \\
& + \frac{1}{n} \sum_{i=1}^n \int_0^t \left(\phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - \phi(p_i, X_{s-}^i)\right) \mathrm{d}M_s^i.
\end{aligned}
\tag{12}
$$

Taking formally $n \to \infty$ on both side of (12) shows that any limit point of μ^n should satisfy the following FPK equation:

$$
\langle \mu_t, \phi \rangle = \langle \mu_0, \phi \rangle + \int_0^t \left\langle \mu_s, \mathcal{L}_c^{s,\mu_s} \phi + \mathcal{L}_j^{s,\mu_s} \phi \right\rangle \mathrm{d}s.
\tag{13}
$$

We next establish a solution of the above FPK equation. Prior to do this, we introduce the following state equation given by

$$
\begin{cases}
\mu_t(\mathrm{d}p, \mathrm{d}x) := \int_E \mathbb{E}\left[\delta_{X_t^{p,x_0,\mu}}(\mathrm{d}x)\right] \rho_0(\mathrm{d}p, \mathrm{d}x_0), \quad \text{on } \mathcal{B}(E), \\
X_t^{p,x_0,\mu} = x_0 + \int_0^t \overline{f}(s, X_s^{p,x_0,\mu}, \langle \mu_s, \alpha(X_s^{p,x_0,\mu}, \cdot) \rangle) \mathrm{d}s + \\
\qquad \int_0^t \overline{g}(s, X_s^{p,x_0,\mu}, \langle \mu_s, \beta(X_s^{p,x_0,\mu}, \cdot) \rangle) \mathrm{d}W_s + \\
\qquad \int_0^t h(s, X_{s-}^{p,x_0,\mu}) \mathrm{d}N_s,
\end{cases}
\tag{14}
$$

where $N = (N_t)_{t \in [0,T]}$ is a double stochastic Poisson process with intensity process $p\psi(t, \langle \mu_t, I \rangle)$, namely, the stochastic process $M_t := N_t - \int_0^t p\psi(t, \langle \mu_s, I \rangle) \mathrm{d}s$ for $t \in [0,T]$ is a martingale.

For $(s,y) \in [0,T] \times \mathbb{R}^m$, define

$$
\begin{cases}
\tilde{\alpha}(s,y) := \int_E \mathbb{E}[\alpha(y, X_s^{p,x_0,\mu})] \rho_0(\mathrm{d}p, \mathrm{d}x_0)\, (= \langle \mu_s, \alpha(y, \cdot) \rangle), \\
\tilde{\beta}(s,y) := \int_E \mathbb{E}[\beta(y, X_s^{p,x_0,\mu})] \rho_0(\mathrm{d}p, \mathrm{d}x_0)\, (= \langle \mu_s, \beta(y, \cdot) \rangle),
\end{cases}
\tag{15}
$$

then Eq. (14) can be rewritten as

$$
\begin{cases}
X_0^{p,x_0,\mu} = x_0, \\
\mathrm{d}X_t^{p,x_0,\mu} = \quad \overline{f}(t, X_t^{p,x_0,\mu}, \tilde{\alpha}(t, X_t^{p,x_0,\mu})) \mathrm{d}t + \\
\qquad\qquad \overline{g}(t, X_t^{p,x_0}, \tilde{\beta}(t, X_t^{p,x_0,\mu})) \mathrm{d}W_t + \\
\qquad\qquad h(t, X_{t-}^{p,x_0,\mu}) \mathrm{d}N_t.
\end{cases}
\tag{16}
$$

where the corresponding intensity of N_t is $p\psi(t, \mathbb{E}[X_t^{p,x_0,\mu}])$.

To establish the well-posedness of equation (16), we give the corresponding assumptions:

$(\mathbf{A}_{\overline{f}\overline{g}h})$ The coefficients $\overline{f} : [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$, $\overline{g} : [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{md}$ and $h : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ satisfy the uniformly locally Lipschitz condition. Namely, for any $R > 0$, there exists a constant $\overline{L}_R > 0$ such that for all $t \in [0,T]$, $x_1, x_2 \in B_R(0) := \{x \in \mathbb{R}^m; |x| \leqslant R\}$ and $y_1, y_2 \in \mathbb{R}$,

$$
|\overline{f}(t, x_1, y_1) - \overline{f}(t, x_2, y_2)| + |\overline{g}(t, x_1, y_1) - \overline{g}(t, x_2, y_2)| +
$$

$$
|h(t, x_1) - h(t, x_2)| \leqslant \overline{L}_R(|x_1 - x_2| + |y_1 - y_2|).
$$

$(\mathbf{A}_{\alpha\beta})$ The coefficients $\alpha : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ and $\beta : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ satisfy the uniformly global Lipschitz condition and a growth condition. Namely, there exists a constant $J > 0$ such that for all $x_1, x_2, y_1, y_2 \in \mathbb{R}^m$,

$$
|\alpha(x_1, y_1) - \alpha(x_2, y_2)| + |\beta(x_1, y_1) - \beta(x_2, y_2)| \leqslant J(|x_1 - x_2| + |y_1 - y_2|)
$$

$$
|\alpha(x_1, y_1)| + |\beta(x_1, y_1)| \leqslant J(1 + |x_1| + |y_1|).
$$

$(\mathbf{A}'_{\mathrm{grow}})$ The coefficients $\overline{f} : [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$, $\overline{g} : [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{md}$ and $h : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ satisfy the following growth condition: there exists a constant $\overline{K} > 0$ such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}^m$,

$$
2x^{\top} \overline{f}(t, x, y) + \mathrm{tr}\left[\overline{gg}^{\top}(t, x, y)\right] \leqslant \overline{K}(1 + |x|^2 + |y|^2),
$$

$$
|h(t, x)|^2 + 2x^{\top} h(t, x) \leqslant \overline{K}(1 + |x|^2).
$$

Lemma 3. Let assumptions $(\mathbf{A}_{\overline{f}\overline{g}h})$, $(\mathbf{A}_{\alpha\beta})$ and $(\mathbf{A}'_{\mathrm{grow}})$ hold. Then there exists a unique strong solution $X^{p,x_0} = (X_t^{p,x_0})_{t \in [0,T]}$ of the equation (16) with an initial condition $(p, x_0) \in E$.

Proof. By combining the two assumptions $(\mathbf{A}_{\overline{f}\overline{g}h})$ and $(\mathbf{A}_{\alpha\beta})$, we find that the coefficients of the drift term and diffusion term of (16) actually satisfy the same locally Lipschitz condition as in assumption $(\mathbf{A}_{fgh})$, hence (16) admits a unique local strong solution.

Subsequently, we show that this solution is actually a global one. Let $V(x) = |x|^2$, for all $x \in \mathbb{R}^m$. It is not hard to check $V(x)$ satisfies (i) and (ii) of assumption $(\mathbf{A}_{Lya})$. For (iii), an application of Itô's formula leads to

$$
\begin{aligned}
V(X_{t \wedge \tau_R}^{p,x_0,\mu}) = {}& x_0^2 + \int_0^{t \wedge \tau_R} 2(X_s^{p,x_0,\mu})^{\top} \overline{f}(s, X_s^{p,x_0,\mu}, \tilde{\alpha}(s, X_s^{p,x_0,\mu})) \mathrm{d}s + \\
& \int_0^{t \wedge \tau_R} 2(X_s^{p,x_0,\mu})^{\top} \overline{g}(s, X_s^{p,x_0,\mu}, \tilde{\beta}(s, X_s^{p,x_0,\mu})) \mathrm{d}W_s + \\
& \int_0^{t \wedge \tau_R} \mathrm{tr}\left[\overline{gg}^{\top}(s, X_s^{p,x_0,\mu}, \tilde{\beta}(s, X_s^{p,x_0,\mu}))\right] \mathrm{d}s + \\
& \int_0^{t \wedge \tau_R} \left((X_s^{p,x_0,\mu} + h(s, X_s^{p,x_0,\mu}))^2 - (X_s^{p,x_0,\mu})^2\right) \mathrm{d}N_s.
\end{aligned}
\tag{17}
$$

Taking expectation on both sides of Eq. (17), by assumption $(\mathbf{A}'_{\mathrm{grow}})$ and $(\mathbf{A}_{\alpha\beta})$ it holds that for some positive constant $C = C_{\overline{f},\overline{g},h,p,\psi,\alpha,\beta} > 0$,

$$
\begin{aligned}
\mathbb{E}[V(X_{t\wedge\tau_R}^{p,x_0,\mu})] =& x_0^2 + \mathbb{E}\left[\int_0^{t\wedge\tau_R} 2(X_s^{p,x_0,\mu})^\top \overline{f}\left(s, X_s^{p,x_0,\mu}, \mathbb{E}[\alpha(y, X_s^{p,x_0,\mu})]\big|_{y=X_s^{p,x_0,\mu}}\right)\mathrm{d}s + \right.\\
& \mathbb{E}\left[\int_0^{t\wedge\tau_R} \overline{gg}^\top\left(s, X_s^{p,x_0,\mu}, \mathbb{E}[\beta(y, X_s^{p,x_0,\mu})]\big|_{y=X_s^{p,x_0,\mu}}\right)\mathrm{d}s\right] + \\
& \mathbb{E}\left[\int_0^{t\wedge\tau_R} p\psi(s, \mathbb{E}[X_s^{p,x_0,\mu}])\left(|h(s, X_s^{p,x_0,\mu})|^2 + 2(X_s^{p,x_0,\mu})^\top h(s, X_s^{p,x_0,\mu})\right)\mathrm{d}s\right] \leqslant \\
& \mathbb{E}[(X_0^{p,x_0,\mu})^2] + C\mathbb{E}\left[\int_0^{t\wedge\tau_R}\left(1 + |X_s^{p,x_0,\mu}|^2 + \left|\mathbb{E}[\alpha(y, X_s^{p,x_0,\mu})]\big|_{y=X_s^{p,x_0,\mu}}\right|^2 + \right.\right.\\
& \left.\left|\mathbb{E}[\beta(y, X_s^{p,x_0,\mu})]\big|_{y=X_s^{p,x_0,\mu}}\right|^2\right)\mathrm{d}s\right] \leqslant \\
& \mathbb{E}[(X_0^{p,x_0,\mu})^2] + C\mathbb{E}\left[\int_0^{t\wedge\tau_R}\left(1 + |X_s^{p,x_0,\mu}|^2 + \left|\mathbb{E}[|X_s^{p,x_0,\mu}|]\right|^2\right)\mathrm{d}s\right] \leqslant \\
& \mathbb{E}[(X_0^{p,x_0,\mu})^2] + C\int_0^t\left(\mathbb{E}[|X_{s\wedge\tau_R}^{p,x_0,\mu}|^2] + 1\right)\mathrm{d}s,
\end{aligned}
\tag{18}
$$

which implies that $V(x)$ also satisfies (iii). Then some similar estimates as in the proof of Lemma 1 yield that the equation (16) with an initial condition $(p, x_0) \in E$ has a unique (global) strong solution.

By Lemma 3, the system (16) admits a unique strong solution. Equivalently, we have proved the strong existence and uniqueness of $\mu = (\mu_t)_{t\in[0,T]}$ defined in the system (14). Now we show that this measure is indeed a solution of the corresponding FPK equation.

Theorem 4. Let the assumptions $(\mathbf{A}_{ini})$, $(\mathbf{A}_{\overline{f}\overline{g}h})$, $(\mathbf{A}_{\alpha\beta})$ and $(\mathbf{A}'_{grow})$ hold. Then $\mu = (\mu_t)_{t\in[0,T]}$ defined in (14) is a solution of the FPK equation (13).

Proof. For any $\phi \in C^2(E)$, by Eq. (14) and Itô's formula, it holds that

$$
\begin{aligned}
\phi(p, X_t^{p,x_0}) =& \phi(p, X_0^{p,x_0}) + \int_0^t \mathscr{L}_c^{s,\mu_s}\phi(p, X_s^{p,x_0})\mathrm{d}s + \\
& \frac{1}{2}\int_0^t \partial_x\phi(p, X_s^{p,x_0})\overline{g}(s, X_s^{p,x_0}, \langle\mu_s, \beta(X_s^{p,x_0}, \cdot)\rangle)\mathrm{d}W_s + \\
& \int_0^t\left(\phi(p, X_{s-}^{p,x_0} + h(s, X_{s-}^{p,x_0})) - \phi(p, X_{s-}^{p,x_0})\right)\mathrm{d}N_s.
\end{aligned}
\tag{19}
$$

Taking expectations on both sides of this equation to kill the martingale term and integrating with respect to ρ_0 on E, we have

$$
\begin{aligned}
& \int_E \mathbb{E}[\phi(p, X_t^{p,x_0})]\rho_0(\mathrm{d}p, \mathrm{d}x_0) = \\
& \int_E \mathbb{E}[\phi(p, X_0^{p,x_0})]\rho_0(\mathrm{d}p, \mathrm{d}x_0) + \\
& \int_0^t \int_E \mathbb{E}[(\mathscr{L}_c^{s,\mu_s} + \mathscr{L}_j^{s,\mu_s})\phi(p, X_s^{p,x_0})]\rho_0(\mathrm{d}p, \mathrm{d}x_0)]\mathrm{d}s = \\
& \int_E \phi(p, x_0)\rho_0(\mathrm{d}p, \mathrm{d}x_0) + \\
& \int_0^t \int_E \mathbb{E}[(\mathscr{L}_c^{s,\mu_s} + \mathscr{L}_j^{s,\mu_s})\phi(p, X_s^{p,x_0})]\rho_0(\mathrm{d}p, \mathrm{d}x_0)]\mathrm{d}s.
\end{aligned}
\tag{20}
$$

This shows that μ satisfies the FPK equation (13), that is

$$
\langle\mu_t, \phi\rangle = \langle\mu_0, \phi\rangle + \int_0^t\left\langle\mu_s, \mathscr{L}_c^{s,\mu_s}\phi + \mathscr{L}_j^{s,\mu_s}\phi\right\rangle\mathrm{d}s.
\tag{21}
$$

Thus, we complete the proof.

4 Propagation of chaos

From Theorem 4 we know that there exists at least one solution to the FPK equation (13). We fix once and for all one such solution μ_t. Next we establish the propagation of chaos of the FPK equation (13) (which means a convergence result

of the empirical measure of the particle system to the solution of the FPK equation (13) under a suitable distance) and thereby show that such solution of (13) is actually unique.

4.1 Propagator of FPK equation

We first give the definition of propagator corresponding to Eq. (13).

Definition 1. For $(p, x) \in E$ and $0 \leqslant t \leqslant u \leqslant T$, the propagator corresponding to the probability solution $\mu = (\mu_t)_{t\in[0,T]}$ of Eq. (13) is defined as

$$
P_{t,u}\phi(p, x) := \mathbb{E}[\phi(p, X_u^{t,p,x,\mu})],
\tag{22}
$$

where $\phi \in C_b^2(E)$ is bounded and twice continuously differentiable, and the process $X^{t,p,x,\mu} = (X_s^{t,p,x,\mu})_{s\in[t,T]}$ is the unique solution of the following SDE:

$$
\begin{aligned}
X_s^{t,p,x,\mu} =& x + \int_t^s \overline{f}(r, X_r^{t,p,x,\mu}, \langle\mu_r, \alpha(X_r^{t,p,x,\mu}, \cdot)\rangle)\mathrm{d}r + \\
& \int_t^s \overline{g}(r, X_r^{t,p,x,\mu}, \langle\mu_r, \beta(X_r^{t,p,x,\mu}, \cdot)\rangle)\mathrm{d}W_r + \int_t^s h(r, X_{r-}^{t,p,x,\mu})\mathrm{d}N_r,
\end{aligned}
\tag{23}
$$

where $N = (N_t)_{t\in[0,T]}$ is a double stochastic Poisson process with intensity process $p\psi(t, \langle\mu_t, I\rangle)$, namely, the stochastic process $M_t := N_t - \int_0^t p\psi(s, \langle\mu_s, I\rangle)\mathrm{d}s$ for $t \in [0, T]$ is a martingale.

The following two lemmas give important properties of the propagator, which will be useful for estimating the convergence rate of μ^n to μ.

Lemma 5. Let the assumptions $(\mathbf{A}_{ini})$, $(\mathbf{A}_{\overline{f}\overline{g}h})$, $(\mathbf{A}_{\alpha\beta})$ and $(\mathbf{A}'_{grow})$ hold, suppose further that $\overline{f}$, $\overline{g}$ is twice continuously differentiable, also suppose the following:

$(\mathbf{A}_h)$ The coefficient $h : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ is bounded, and there exist constants L_h, $L'_h > 0$ such that for all $t \in [0, T]$ and x_1, $x_2 \in \mathbb{R}^m$,

$$
L'_h|x_1 - x_2| \leqslant |h(t, x_1) - h(t, x_2)| \leqslant L_h|x_1 - x_2|.
$$

Then the propagator in Definition 1 satisfies that for all $\phi \in C_b^2(E)$, $0 \leqslant t \leqslant u \leqslant T$ and $(p, x) \in E$,

$$
\begin{cases}
\partial_t P_{t,u}\phi(p, x) + \left(\mathscr{L}_c^{t,\mu_t} + \mathscr{L}_j^{t,\mu_t}\right)P_{t,u}\phi(p, x) = 0, \\
P_{u,u}\phi(p, x) = \phi(p, x).
\end{cases}
\tag{24}
$$

where $\mathscr{L}_c^{t,\mu_t}$ and $\mathscr{L}_j^{t,\mu_t}$ are defined as in (11). Moreover, $P_{\cdot,u}\phi(p, \cdot) \in C^{1,2}([0, u] \times \mathbb{R}^m)$.

Proof. We first consider the following parabolic equation with respect to $\theta : [0, u] \times \mathbb{R}^m \to \mathbb{R}$ for fixed p:

$$\begin{cases} \partial_t \theta(t, x) + \left(\mathcal{L}_c^{t, \mu_t} + \mathcal{L}_j^{t, \mu_t} \right) \theta(t, x) = 0, \\ \theta(u, x) = \phi(p, x). \end{cases} \tag{25}$$

This is a type of second order integro-differential parabolic equation with Cauchy boundary. By Theorem 3.1 of chapter II of Ref. [27], we know that (25) has a unique classical solution $\theta(t, x) \in C^{1,2}([0, u] \times \mathbb{R}^m)$.

Next, for all $(t, x) \in [0, u] \times \mathbb{R}^m$, by applying Itô's formula to $\theta(u, X_u^{t,p,x,\mu})$, we obtain

$$\begin{aligned} \theta(t, x) = {} & \theta(u, X_u^{t,p,x,\mu}) - \int_t^u \left(\partial_s + \mathcal{L}_c^{s,\mu_s} + \mathcal{L}_j^{s,\mu_s} \right) \theta(s, X_s^{t,p,x,\mu}) \mathrm{d}s - \\ & \int_t^u \partial_x \theta(s, X_s^{t,p,x,\mu}) \overline{g}(s, X_s^{t,p,x,\mu}, \langle \mu_s, \beta(X_s^{t,p,x,\mu}, \cdot) \rangle) \mathrm{d}W_s - \\ & \int_t^u \left(\theta(s, X_{s-}^{t,p,x,\mu} + h(s, X_{s-}^{t,p,x,\mu})) - \theta(s, X_{s-}^{t,p,x,\mu}) \right) \mathrm{d}M_s. \end{aligned} \tag{26}$$

By taking expectations on both sides of Eq. (26), we deduce that

$$\begin{aligned} \theta(t, x) = {} & \mathbb{E} \left[\theta(u, X_u^{t,p,x,\mu}) - \int_t^u \left(\partial_s + \mathcal{L}_c^{s,\mu_s} + \mathcal{L}_j^{s,\mu_s} \right) \theta(s, X_s^{t,p,x,\mu}) \mathrm{d}s \right] = \\ & \mathbb{E} \left[\theta(u, X_u^{t,p,x,\mu}) \right] = \mathbb{E} \left[\phi(p, X_u^{t,p,x,\mu}) \right] = P_{t,u} \phi(p, x). \end{aligned} \tag{27}$$

This yields that $P_{\cdot,u} \phi(p, \cdot) \in C^{1,2}([0, u] \times \mathbb{R}^m)$ satisfies the PDE (24). Thus, we complete the proof.

Lemma 6. Let conditions of Lemma 5 hold. Then for all $0 \leqslant t \leqslant u \leqslant T$,

$$\partial_t \langle \mu_t, P_{t,u} \phi \rangle = 0. \tag{28}$$

Proof. It follows from Theorem 4, Lemma 5, Eq. (14), and an application of Itô's formula to $P_{t,u} \phi(p, X_t^{0,p,x_0,\mu})$ that

$$\begin{aligned} \langle \mu_t, P_{t,u} \phi \rangle = {} & \int_{\mathbb{R}^m} \mathbb{E} \left[P_{t,u} \phi(p, X_t^{0,p,x_0,\mu}) \right] \rho_0(\mathrm{d}p, \mathrm{d}x_0) = \\ & \int_{\mathbb{R}^m} \mathbb{E} \Big[P_{0,u} \phi(p, X_0^{0,p,x_0,\mu}) + \int_0^t (\partial_s + \mathcal{L}_c^{s,\mu_s} + \mathcal{L}_j^{s,\mu_s}) P_{s,u} \phi(p, X_s^{0,p,x_0,\mu}) \mathrm{d}s + \\ & \int_0^t \partial_x P_{s,u} \phi(p, X_s^{0,p,x_0,\mu}) \overline{g}(s, X_s^{0,p,x_0,\mu}, \langle \mu_s, \beta(X_s^{0,p,x_0,\mu}, \cdot) \rangle) \mathrm{d}W_s + \\ & \int_0^t (P_{s,u} \phi(p, X_{s-}^{0,p,x_0,\mu} + h(s, X_{s-}^{0,p,x_0,\mu})) - P_{s,u} \phi(p, X_{s-}^{0,p,x_0,\mu})) \mathrm{d}M_s \Big] \rho_0(\mathrm{d}p, \mathrm{d}x_0) = \\ & \int_{\mathbb{R}^m} \mathbb{E} \left[P_{0,u} \phi(p, X_0^{0,p,x_0,\mu}) \right] \rho_0(\mathrm{d}p, \mathrm{d}x_0). \end{aligned} \tag{29}$$

It appears that $\langle \mu_t, P_{t,u} \phi \rangle$ is independent of t, which indicates the desired result.

The propagator defined by Eq. (22) can help us establish the following relation satisfied by $\mu_t^n - \mu_t$ for any $t \in [0, T]$:

Lemma 7. Let conditions of Lemma 5 hold. Then for any $t \in [0, T]$, it holds that

$$\begin{aligned} \langle \mu_t^n - \mu_t, \phi \rangle = {} & \langle \mu_0^n - \mu_0, P_{0,t} \phi \rangle + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^t \partial_x P_{s,t} \phi(p_i, X_s^i) \overline{g}\left(s, X_s^i, \frac{1}{n} \sum_{j=1}^n \beta(X_s^i, X_s^j)\right) \mathrm{d}W_s^i + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^t (P_{s,t} \phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - P_{s,t} \phi(p_i, X_{s-}^i)) \mathrm{d}M_s^i. \end{aligned} \tag{30}$$

Proof. Recall the state process of our particle system $X^i = (X_t^i)_{t \in [0,T]}$ defined by (4) for $i \geqslant 1$. By applying Itô's formula to $P_{r,t} \phi(p_i, X_t^i)$ for $r \in [0, t]$, it follows from Lemma 5 that

$$\begin{aligned} & P_{r,t} \phi(p_i, X_t^i) = \\ & P_{0,t} \phi(p_i, X_0^i) + \int_0^r (\partial_s + \mathcal{L}_c^{s,\mu_s} + \mathcal{L}_j^{s,\mu_s}) P_{s,t} \phi(p_i, X_s^i) \mathrm{d}s + \\ & \int_0^r \partial_x P_{s,t} \phi(p_i, X_s^i) \overline{g}\left(s, X_s^i, \frac{1}{n} \sum_{j=1}^n \beta(X_s^i, X_s^j)\right) \mathrm{d}W_s^i + \\ & \int_0^r (P_{s,t} \phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - P_{s,t} \phi(p_i, X_{s-}^i)) \mathrm{d}M_s^i = \\ & P_{0,t} \phi(p_i, X_0^i) + \int_0^r \partial_x P_{s,t} \phi(p_i, X_s^i) \overline{g}\left(s, X_s^i, \frac{1}{n} \sum_{j=1}^n \beta(X_s^i, X_s^j)\right) \mathrm{d}W_s^i + \\ & \int_0^r (P_{s,t} \phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - P_{s,t} \phi(p_i, X_{s-}^i)) \mathrm{d}M_s^i, \end{aligned} \tag{31}$$

then in view of the definition of μ^n, we have

$$\begin{aligned} \langle \mu_r^n, P_{r,t} \phi \rangle = {} & \frac{1}{n} \sum_{i=1}^n P_{r,t} \phi(p_i, X_r^i) = \langle \mu_0^n, P_{0,t} \phi \rangle + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^r \partial_x P_{s,t} \phi(p_i, X_s^i) \overline{g}\left(s, X_s^i, \frac{1}{n} \sum_{j=1}^n \beta(X_s^i, X_s^j)\right) \mathrm{d}W_s^i + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^r (P_{s,t} \phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - P_{s,T} \phi(p_i, X_{s-}^i)) \mathrm{d}M_s^i. \end{aligned} \tag{32}$$

Note that Lemma 6 implies

$$\langle \mu_r, P_{r,t} \phi \rangle = \langle \mu_0, P_{0,t} \phi \rangle, \text{ for all } r \in [0, t]. \tag{33}$$

From (32) and (33), we deduce that

$$\begin{aligned} \langle \mu_r^n - \mu_r, P_{r,t} \phi \rangle = {} & \langle \mu_0^n - \mu_0, P_{0,t} \phi \rangle + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^r \partial_x P_{s,t} \phi(p_i, X_s^i) \overline{g}\left(s, X_s^i, \frac{1}{n} \sum_{j=1}^n \beta(X_s^i, X_s^j)\right) \mathrm{d}W_s^i + \\ & \frac{1}{n} \sum_{i=1}^n \int_0^r (P_{s,T} \phi(p_i, X_{s-}^i + h(s, X_{s-}^i)) - P_{s,T} \phi(p_i, X_{s-}^i)) \mathrm{d}M_s^i. \end{aligned} \tag{34}$$

In Eq. (34), let $r = t$ and note that $P_{t,t} \phi = \phi$ by virtue of Lemma 5, then we get the desired result.

Next we give the estimate of the gradient $\partial_x P_{s,t} \phi(p, x)$. To this purpose, we impose the following one-side Lipschitz condition:

($\mathbf{A}_{one}$) The coefficient $\overline{f} : [0, T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^m$ satisfies a one-side Lipschitz condition. Namely, there exists a constant $L_{\overline{f}}$ such that for all $t \in [0, T]$, $x_1, x_2 \in \mathbb{R}^m$ and $y_1, y_2 \in \mathbb{R}$,

$$< x_1 - x_2, \overline{f}(t, x_1, y_1) - \overline{f}(t, x_2, y_2) > \leqslant L_{\overline{f}}(|x_1 - x_2|^2 + |y_1 - y_2|^2).$$

Lemma 8. Let conditions of Lemma 5 and ($\mathbf{A}_{one}$) hold, suppose the function ϕ satisfies that:

(i) $\phi \in C_b^2(E)$;

(ii) ϕ is Lipschitz continuous and

$$\|\phi\|_{\text{Lip}} = \sup_{(p_1, x_1) \neq (p_2, x_2) \in E} \frac{|\phi(p_1, x_1) - \phi(p_2, x_2)|}{|(p_1 - p_2, x_1 - x_2)|} \leqslant 1.$$

Then for all $0 \leqslant s \leqslant t \leqslant T$ and $(p, x) \in E$:

$$|P_{s,t}\phi(p, x_1) - P_{s,t}\phi(p, x_2)|^2 = \left|\mathbb{E}\left[\phi(p, X_t^1) - \phi(p, X_t^2)\right]\right|^2 \leqslant \mathbb{E}\left[|X_t^1 - X_t^2|^2\right].$$

An application of Itô's formula to $|X_t^1 - X_t^2|^2$ yields that
$$\mathbb{E}\left[|X_t^1 - X_t^2|^2\right] =$$

$$|X_s^1 - X_s^2|^2 + \mathbb{E}\left[\int_s^t \left\langle X_r^1 - X_r^2, \overline{f}(r, X_r^1, \langle \mu_r, \beta(X_r^1, \cdot)\rangle) - \overline{f}(r, X_r^2, \langle \mu_r, \beta(X_r^2, \cdot)\rangle)\right\rangle dr\right] +$$

$$\mathbb{E}\left[\int_s^t \left(\left|(X_r^1 - X_r^2) + (h(r, X_{r-}^1) - h(r, X_{r-}^2))\right|^2 - |X_r^1 - X_r^2|^2\right) p\psi(r, \langle \mu_r, I\rangle) dr\right] +$$

$$\mathbb{E}\left[\int_s^t \left\langle X_r^1 - X_r^2, \overline{g}(r, X_r^1, \langle \mu_r, \beta(X_r^1, \cdot)\rangle) - \overline{g}(r, X_r^2, \langle \mu_r, \beta(X_r^2, \cdot)\rangle)\right\rangle dW_r\right] +$$

$$\mathbb{E}\left[\int_s^t \left(\left|(X_r^1 - X_r^2) + (h(r, X_{r-}^1) - h(r, X_{r-}^2))\right|^2 - |X_r^1 - X_r^2|^2\right) dM_r\right] =$$

$$|x_1 - x_2|^2 + \mathbb{E}\left[\int_s^t \left\langle X_r^1 - X_r^2, \overline{f}(r, X_r^1, \langle \mu_r, \beta(X_r^1, \cdot)\rangle) - \overline{f}(r, X_r^2, \langle \mu_r, \beta(X_r^2, \cdot)\rangle)\right\rangle dr\right] +$$

$$\mathbb{E}\left[\int_s^t \left(\left|(X_r^1 - X_r^2) + (h(r, X_{r-}^1) - h(r, X_{r-}^2))\right|^2 - |X_r^1 - X_r^2|^2\right) p\psi(r, \langle \mu_r, I\rangle) dr\right]. \tag{37}$$

By assumptions ($\mathbf{A}_{one}$) and ($\mathbf{A}_{\alpha\beta}$), we have the following estimate:

$$\mathbb{E}\left[\int_s^t \left\langle X_r^1 - X_r^2, \overline{f}(r, X_r^1, \langle \mu_r, \beta(X_r^1, \cdot)\rangle) - \overline{f}(r, X_r^2, \langle \mu_r, \beta(X_r^2, \cdot)\rangle)\right\rangle dr\right] \leqslant$$

$$L\mathbb{E}\left[\int_s^t |X_r^1 - X_r^2|^2 + \left|\langle \mu_r, \beta(X_r^1, \cdot)\rangle - \langle \mu_r, \beta(X_r^2, \cdot)\rangle\right|^2 dr\right] \leqslant \tag{38}$$

$$L_{\overline{f}}(1 + K^2)\mathbb{E}\left[\int_s^t |X_r^1 - X_r^2|^2 dr\right].$$

By assumption ($\mathbf{A}_h$) and noting that $|\psi(\cdot, \cdot)| \leqslant C_\psi$, it follows that

$$\mathbb{E}\left[\int_s^t \left(\left|(X_r^1 - X_r^2) + (h(r, X_{r-}^1) - h(r, X_{r-}^2))\right|^2 - |X_r^1 - X_r^2|^2\right) p\psi(r, \langle \mu_r, I\rangle) dr\right] \leqslant$$

$$C_\psi p\mathbb{E}\left[\int_s^t \left(|X_r^1 - X_r^2|^2 + 2\left|h(r, X_{r-}^1) - h(r, X_{r-}^2)\right|^2\right) dr\right] \leqslant \tag{39}$$

$$2C_\psi(1 + L_h^2)p\mathbb{E}\left[\int_s^t |X_r^1 - X_r^2|^2 dr\right].$$

We deduce from (37), (38) and (39) that there exists a constant $C = C_{\overline{f}, h, K, \psi}$ such that

$$\mathbb{E}\left[|X_t^1 - X_t^2|^2\right] \leqslant |x_1 - x_2|^2 + Cp\int_s^t \mathbb{E}\left[|X_r^1 - X_r^2|^2\right] dr. \tag{40}$$

Then, Gronwall's lemma yields that

$$\mathbb{E}\left[|X_t^{s,p,x_1,\mu} - X_t^{s,p,x_2,\mu}|^2\right] \leqslant |x_1 - x_2|^2 e^{Cp(t-s)}. \tag{41}$$

Therefore, we have

$$|\partial_x P_{s,t}\phi(p, x_1)| = \lim_{x_2 \to x_1} \frac{|P_{s,t}\phi(p, x_1) - P_{s,t}\phi(p, x_2)|}{|x_1 - x_2|} \leqslant$$

$$\sqrt{\frac{\mathbb{E}\left[|X_t^1 - X_t^2|^2\right]}{|x_1 - x_2|^2}} \leqslant e^{\frac{1}{2}Cp(t-s)}. \tag{42}$$

This completes the proof of the lemma.

4.2 Metric and convergence of μ^n to μ

Next we establish a convergence result to the solution of our

$$|\partial_x P_{s,t}\phi(p, x)| \leqslant e^{Cp(t-s)}, \tag{35}$$

where $C = C_{\overline{f}, h, K, \psi} > 0$ is a constant.

Remark 1. Based on the proof below, condition (ii) in Lemma 8 can be weaken as ϕ is only Lipschitz continuous with respect to x with Lipschitz coefficient no more that 1. In other words, $|\phi(p, x_1) - \phi(p, x_2)| \leqslant |x_1 - x_2|$ for all $p > 0$ and $x_1, x_2 \in R^m$.

Proof. For simplicity, we use X_r^1 and X_r^2 instead of $X_r^{s,p,x_1,\mu}$ and $X_r^{s,p,x_2,\mu}$ in this proof. Note that $\|\phi\|_{\text{Lip}} \leqslant 1$, then in the view of (22) and (23), for all $x_1, x_2 \in \mathbb{R}^m$ and $p \in \mathbb{R}_+$, it follows that

$$\tag{36}$$

FPK function. To this end, we need an appropriate metric between μ^n and μ. Let $\mathcal{P}(E)$ be the set of all finite measures ν on $\mathcal{B}(E)$ such that $\nu(E) \leqslant 1$. To make a compact space, we add an extra point $\star$ to E, and denote that $E_\star := E \cup \{\star\}$. Let E be topologized by a topology $\mathcal{T}$, we then define a topology $\mathcal{T}^\star$ for $E_\star$ as follows:

(i) $\mathcal{T} \subset \mathcal{T}^\star$;

(ii) For each compact set $C \subset E$, define $U_C \in \mathcal{T}^\star$ by $U_C := (E \setminus C) \cup \{\star\}$.

We give a bijection $\iota : \mathcal{P}(E) \to \mathcal{P}(E_\star)$ by

$$(\iota\nu)(A) := \nu(A \cap E) + (1 - \nu(E))\delta_{\{\star\}}(A) \tag{43}$$

for all $A \in \mathcal{B}(E_\star)$ and $\nu \in \mathcal{P}(E)$, where δ denotes the Dirac delta measure. Then we define the integral of $\nu \in \mathcal{P}(E)$ with respect to a measurable function $\phi : E_\star \to \mathbb{R}$ as:

$$\int_{E_\star} \phi(x)(\iota\nu)(dx) = \int_E \phi(x)\nu(dx) + \phi(\star)(1 - \nu(E)). \tag{44}$$

For the parameter $q \geqslant 2$, we establish the metric $d_{q,T}$ between $\mu = (\mu_t)_{t \in [0,T]}$ and $\mu^n = (\mu_t^n)_{t \in [0,T]}$ as follows (see Ref. [24]):

$$d_{q,T}(\mu^n,\mu) := \sup_{t \in [0,T]} d_{\mathrm{BL}}(\mu_t^n,\mu_t). \qquad (45)$$

The d_{BL} here is defined as

$$d_{\mathrm{BL}}(\mu_t^n,\mu_t) = \sup_{\phi \in \mathcal{R}_1} \mathbb{E}\left[\left|\int_{E_\star} \phi(x)(\mu_t^n - \mu_t)(\mathrm{d}x)\right|^q\right]^{\frac{1}{q}}, \qquad (46)$$

where $\mathcal{R}_1$ is the set of functions $\phi : E_\star \to \mathbb{R}$ that are bounded, twice continuously differentiable and Lipschitz continuous satisfying that $\|\phi\|_\infty + \|\phi\|_{\mathrm{Lip}} \leqslant 1$, where $\|\phi\|_\infty := \inf\{C > 0 : |\phi(x)| \leqslant C, \ \forall x \in E_\star\}$.

The metric $d(\cdot,\cdot)$ on $E_\star$ is defined as in Ref. [28], that is: fix $x_0 \in E$ and let $l(x) := (1 + |x - x_0|)^{-1}$ for $x \in E$, and

$$d(x,y) := \begin{cases} |x - y| \wedge |l(x) - l(y)|, & x,y \in E; \\ l(x), & x \in E, y = \star; \\ 0, & x = y = \star. \end{cases}$$

Then we have

$$\|\phi\|_{\mathrm{Lip}} = \sup_{x \neq y, \ x,y \in E_\star} \frac{|\phi(x) - \phi(y)|}{d(x,y)}, \qquad (47)$$

hence for any $\phi \in \mathcal{R}_1$, it holds for all $x,y \in E$ that

$$|\phi(x) - \phi(y)| \leqslant d(x,y) = |x - y| \wedge |l(x) - l(y)| \leqslant |x - y|. \qquad (48)$$

Namely, the constraint of ϕ on E is also bounded Lipschitz continuous with $\|\phi\|_{\mathrm{Lip}} \leqslant 1$.

In view of (44), we have

$$\int_{E_\star} \phi(x)(\mu_t^n - \mu_t)(\mathrm{d}x) =$$
$$\int_E \phi(x)(\mu_t^n - \mu_t)(\mathrm{d}x) + \phi(\star)((1 - \mu_t^n(E)) - (1 - \mu_t(E))) =$$
$$\int_E \phi(x)(\mu_t^n - \mu_t)(\mathrm{d}x) =$$
$$\langle \mu_t^n - \mu_t, \phi \rangle. \qquad (49)$$

Therefore,

$$d_{q,T}(\mu^n,\mu) = \sup_{t \in [0,T]} \sup_{\phi \in \mathcal{R}_1} \mathbb{E}\left[\left|\langle \mu_t^n - \mu_t, \phi \rangle\right|^q\right]^{\frac{1}{q}}. \qquad (50)$$

To obtain the convergence results, we need the following assumption:

$(\mathbf{A}_{\overline{g}h})$ The coefficients $\overline{g} : [0,T] \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{md}$ and $h : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ satisfy a growth condition. Namely, there exists a constant K such that for all $t \in [0,T]$, $x \in \mathbb{R}^m$ and $y \in \mathbb{R}$,

$$|\overline{g}(t,x,y)|^2 \leqslant K(1 + |x|^2 + |y|^2), |h(t,x)|^2 \leqslant K(1 + |x|^2).$$

Theorem 9. Let conditions of Lemma 8 and $(\mathbf{A}_{\overline{g}h})$ hold, also suppose that $\mathbb{E}[|X_0^i|^q] < \infty$ for all $q \geqslant 2$ and $i \geqslant 1$. Then for fixed $T > 0$ and parameters $\kappa > q$ and $m > 0$, there exists a constant $C_q = C_{q,T,\overline{f},\overline{g},h,\psi,\{p_i\}_{i \geqslant 1},\kappa,m} > 0$ which is independent of n, such that for all $n \geqslant 1$:

$$d_{q,T}(\mu^n,\mu) \leqslant C_q\left(\gamma(\kappa,q,m,n) + \frac{1}{n^{1 - \frac{1}{q}}}\right), \qquad (51)$$

where $d_{q,T}(\cdot,\cdot)$ is defined by (45), and

$$\gamma(\kappa,q,m,n) := \begin{cases} n^{-\frac{1}{2}} + n^{-\frac{\kappa-q}{\kappa}}, & q > \dfrac{m}{2}, \kappa \neq 2q; \\[2mm] n^{-\frac{1}{2}}\ln(1+n) + n^{-\frac{\kappa-q}{\kappa}}, & q = \dfrac{m}{2}, \kappa \neq 2q; \\[2mm] n^{-\frac{q}{m}} + n^{-\frac{\kappa-q}{\kappa}}, & q < \dfrac{m}{2}, \kappa \neq \dfrac{m}{m-q}. \end{cases} \qquad (52)$$

Proof. By Lemma 7, it holds for all $\phi \in \mathcal{R}_1$ and $0 \leqslant t \leqslant T$ that

$$\langle \mu_t^n - \mu_t, \phi \rangle = \langle \mu_0^n - \mu_0, P_{0,t}\phi \rangle +$$
$$\frac{1}{n}\sum_{i=1}^n \int_0^t \partial_x P_{s,t}\phi(p_i,X_s^i)\overline{g}\left(s,X_s^i,\frac{1}{n}\sum_{j=1}^n \beta(X_s^i,X_s^j)\right)\mathrm{d}W_s^i +$$
$$\frac{1}{n}\sum_{i=1}^n \int_0^t (P_{s,t}\phi(p_i,X_{s-}^i + h(s,X_{s-}^i)) - P_{s,t}\phi(p_i,X_{s-}^i))\mathrm{d}M_s^i =:$$
$$I_1(t) + I_2(t) + I_3(t). \qquad (53)$$

Then we have that

$$\mathbb{E}\left[\left|\langle \mu_t^n - \mu_t, \phi \rangle\right|^q\right] = \mathbb{E}[|I_1(t) + I_2(t) + I_3(t)|^q] \leqslant$$
$$C_q(\mathbb{E}[|I_1(t)|^q] + \mathbb{E}[|I_2(t)|^q] + \mathbb{E}[|I_3(t)|^q]). \qquad (54)$$

In what follows, we will give the estimate of $\mathbb{E}[|I_1(t)|^q]$, $\mathbb{E}[|I_2(t)|^q]$ and $\mathbb{E}[|I_3(t)|^q]$. One is reminded that during the following proof, the constant C_q can be different from line to line.

For the first term, by using Theorem 2 in Ref. [22], we have

$$\mathbb{E}[|I_1(t)|^q] \leqslant C_{q,\kappa,m}\gamma(\kappa,q,m,n), \qquad (55)$$

where $\gamma(\kappa,q,m,n)$ is defined as in (52). (For this estimate, Ref. [29] can also be referenced.)

For the second term, noting that $\phi \in \mathcal{R}_1$, by Burkholder–Davis–Gundy inequality, Lemma 8 and assumption $(\mathbf{A}_{\overline{g}h})$, it holds that

$$\mathbb{E}[|I_2(t)|^q] = \mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^n \int_0^t \partial_x P_{s,t}\phi(p_i,X_s^i)\overline{g}\left(s,X_s^i,\frac{1}{n}\sum_{j=1}^n \beta(X_s^i,X_s^j)\right)\mathrm{d}W_s^i\right|^q\right] \leqslant$$

$$\mathbb{E}\left[\sup_{t' \in [0,t]}\left|\frac{1}{n}\sum_{i=1}^n \int_0^{t'} \partial_x P_{s,t'}\phi(p_i,X_s^i)\overline{g}\left(s,X_s^i,\frac{1}{n}\sum_{j=1}^n \beta(X_s^i,X_s^j)\right)\mathrm{d}W_s^i\right|^q\right] \leqslant$$

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n \int_0^t |\partial_x P_{s,t}\phi(p_i,X_s^i)|^2 \left|\overline{g}\left(s,X_s^i,\frac{1}{n}\sum_{j=1}^n \beta(X_s^i,X_s^j)\right)\right|^2 \mathrm{d}s\right)^{\frac{q}{2}}\right] \leqslant$$

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}e^{Cp_{i}t}\left|\overline{g}\left(s,X_{s}^{i},\frac{1}{n}\sum_{j=1}^{n}\beta(X_{s}^{i},X_{s}^{j})\right)\right|^{2}ds\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\mathbb{E}\left[\left(\sum_{i=1}^{n}\int_{0}^{t}\left|\overline{g}\left(s,X_{s}^{i},\frac{1}{n}\sum_{j=1}^{n}\beta(X_{s}^{i},X_{s}^{j})\right)\right|^{2}ds\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}\left|\overline{g}\left(s,X_{s}^{i},\frac{1}{n}\sum_{j=1}^{n}\beta(X_{s}^{i},X_{s}^{j})\right)\right|^{2}ds\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}\left(1+|X_{s}^{i}|^{2}+\frac{1}{n}\sum_{j=1}^{n}|X_{s}^{j}|^{2}\right)ds\right)^{\frac{q}{2}}\right]. \tag{56}$$

Here we have used the convergence of $(p_{i})_{i\geqslant 1}$, which leads to the existence of a constant $C'>0$ such that $e^{Cp_{i}}<C'$ for all $i\geqslant 1$. In addition, since $\frac{q}{2}\geqslant 1$, applying Jensen's inequality twice to the convex function $a(x)=x^{\frac{q}{2}}$ leads to

$$\left(\int_{0}^{t}\left(1+|X_{s}^{i}|^{2}+\frac{1}{n}\sum_{j=1}^{n}|X_{s}^{j}|^{2}\right)ds\right)^{\frac{q}{2}}=3^{\frac{q}{2}}\left(\frac{t+\int_{0}^{t}|X_{s}^{i}|^{2}ds+\left(\frac{1}{n}\sum_{j=1}^{n}\int_{0}^{t}|X_{s}^{j}|^{2}ds\right)}{3}\right)^{\frac{q}{2}}\leqslant$$

$$3^{\frac{q}{2}-1}\left(t^{\frac{q}{2}}+\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}+\left(\frac{\sum_{j=1}^{n}\int_{0}^{t}|X_{s}^{j}|^{2}ds}{n}\right)^{\frac{q}{2}}\right)\leqslant 3^{\frac{q}{2}-1}\left(t^{\frac{q}{2}}+\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}+\frac{1}{n}\sum_{j=1}^{n}\left(\int_{0}^{t}|X_{s}^{j}|^{2}ds\right)^{\frac{q}{2}}\right). \tag{57}$$

Summing over $i=1,2,\cdots,n$, multiplying by $\frac{C_{q}}{n^{q}}$ and taking expectation, it follows that

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}\left(1+|X_{s}^{i}|^{2}+\frac{1}{n}\sum_{j=1}^{n}|X_{s}^{j}|^{2}\right)ds\right)^{\frac{q}{2}}\right]\leqslant\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[3^{\frac{q}{2}-1}\left(t^{\frac{q}{2}}+\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}+\frac{1}{n}\sum_{j=1}^{n}\left(\int_{0}^{t}|X_{s}^{j}|^{2}ds\right)^{\frac{q}{2}}\right)\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[t^{\frac{q}{2}}+2\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]\leqslant\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[1+\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]. \tag{58}$$

Hence in view of (56) and (58), we have

$$\mathbb{E}[|I_{2}(t)|^{q}]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}\left(1+|X_{s}^{i}|^{2}+\frac{1}{n}\sum_{j=1}^{n}|X_{s}^{j}|^{2}\right)ds\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[1+\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]= \tag{59}$$

$$\frac{C_{q}}{n^{q-1}}\left(1+\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]\right).$$

By Minkowski's integral inequality[30] for $\frac{q}{2}\geqslant 1$, we know that

$$\mathbb{E}\left[\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]\leqslant\left(\int_{0}^{t}\left(\mathbb{E}[|X_{s}^{i}|^{q}]\right)^{\frac{2}{q}}ds\right)^{\frac{q}{2}}. \tag{60}$$

Meanwhile, through an application of Gronwall's lemma to-gether with assumption ($\mathbf{A}_{ini}$), one can verify that there exists a constant $C_{T,q}>0$ such that

$$\mathbb{E}[|X_{s}^{i}|^{q}]\leqslant C_{T,q}(1+\mathbb{E}[|X_{0}^{i}|^{q}])=C_{T,q}(1+\mathbb{E}[|X_{0}^{1}|^{q}]), \tag{61}$$

see Ref. [31, Theorem 67] for details. Therefore, (59), (60) and (61) yield that

$$\mathbb{E}[|I_{2}(t)|^{q}]\leqslant\frac{C_{q}}{n^{q-1}}\left(1+\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}|X_{s}^{i}|^{2}ds\right)^{\frac{q}{2}}\right]\right)\leqslant$$

$$\frac{C_{q}}{n^{q-1}}(1+\mathbb{E}[|X_{0}^{1}|^{q}]). \tag{62}$$

For the third term, using the mean value theorem to $P_{s,t}\phi(p_{i},\cdot)$ with $\xi_{s-}^{i}\in[0,h(t,X_{s-}^{i})]$ for all $s\in(0,t)$, with the help of Lemma 5, Lemma 8, assumption ($\mathbf{A}_{\overline{g}h}$), the boundedness of $(p_{i})_{i\geqslant 1}$ and ψ, and by some analogous methods as in the previous one, it holds that

$$\mathbb{E}[|I_{3}(t)|^{q}]=\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}P_{s,t}\phi(p_{i},X_{s-}^{i}+h(s,X_{s-}^{i}))-P_{s,t}\phi(p_{i},X_{s-}^{i})dM_{s}^{i}\right|^{q}\right]=$$

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}\partial_{x}P_{s,t}\phi(p_{i},X_{s-}^{i}+\xi_{s-}^{i})h(t,X_{s-}^{i})dM_{s}^{i}\right|^{q}\right]\leqslant$$

$$\mathbb{E}\left[\sup_{t'\in[0,t]}\left|\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t'}\partial_{x}P_{s,t'}\phi(p_{i},X_{s-}^{i}+\xi_{s-}^{i})h(s,X_{s-}^{i})dM_{s}^{i}\right|^{q}\right]\leqslant$$

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}(\partial_{x}P_{s,t}\phi(p_{i},X_{s-}^{i}+\xi_{s-}^{i}))^{2}h(s,X_{s-}^{i})^{2}\mathrm{d}\langle M^{i}\rangle_{s}\right|^{\frac{q}{2}}\right]\leqslant$$

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}\left|\partial_{x}P_{s,t}\phi(p_{i},X_{s-}^{i}+\xi_{s-}^{i})\right|^{2}\left|h(s,X_{s-}^{i})\right|^{2}p_{i}\psi\left(s,\frac{1}{n}\sum_{j=1}^{n}X_{s}^{j}\right)\mathrm{d}s\right)^{\frac{q}{2}}\right]\leqslant$$

$$C_{q}\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t}\left|h(s,X_{s-}^{i})\right|^{2}\mathrm{d}s\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}\left|h(s,X_{s-}^{i})\right|^{2}\mathrm{d}s\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}(1+|X_{s-}^{i}|^{2})\mathrm{d}s\right)^{\frac{q}{2}}\right]\leqslant$$

$$\frac{C_{q}}{n^{q}}\sum_{i=1}^{n}\mathbb{E}\left[1+\left(\int_{0}^{t}|X_{s}^{i}|^{2}\mathrm{d}s\right)^{\frac{q}{2}}\right]=$$

$$\frac{C_{q}}{n^{q-1}}\left(1+\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\left(\int_{0}^{t}|X_{s}^{i}|^{2}\mathrm{d}s\right)^{\frac{q}{2}}\right]\right)\leqslant$$

$$\frac{C_{q}}{n^{q-1}}\left(1+\mathbb{E}\left[\left|X_{0}^{1}\right|^{q}\right]\right). \tag{63}$$

Note that $\mathbb{E}\left[\left|X_{0}^{1}\right|^{q}\right]<\infty$, then in view of (45), (46), (55), (62) and (63), we complete the proof.

Conflict of interest

The author declares that he has no conflict of interest.

Biography

Zeqian Li   is a postgraduate student under the tutelage of Prof. Lijun Bo at the School of Mathematical Sciences, University of Science and Technology of China. His research mainly focuses on mean-field analysis.

References

[1] Dehnen W, Read J I. *N*-body simulations of gravitational dynamics. *The European Physical Journal Plus*, **2011**, *126*: 55.

[2] Sirignano J, Spiliopoulos K. Mean field analysis of neural networks: A central limit theorem. *Stochastic Processes and Their Applications*, **2020**, *130* (3): 1820–1852.

[3] Bolley F, Canizo J A, Carrillo J A. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Mathematical Models and Methods in Applied Sciences*, **2011**, *21* (11): 2179–2210.

[4] Bender M, Heenen P-H, Reinhard P-G. Selfconsistent mean-field models for nuclear structure. *Reviews of Modern Physics*, **2003**, *75* (1): 121–180.

[5] Touboul J. Propagation of chaos in neural fields. *Annals of Applied Probability*, **2014**, *24* (3): 1298–1327.

[6] Delarue F, Inglis J, Rubenthaler S, et al. Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Processes and Their Applications*, **2015**, *125* (6): 2451–2492.

[7] Bo L, Capponi A. Systemic risk in interbanking networks. *SIAM Journal on Financial Mathematics*, **2015**, *6* (1): 386–424.

[8] Liu W, Song Y, Zhai J, et al. Large and moderate deviation principles for McKean–Vlasov SDEs with jumps. *Potential Analysis*, **2023**, *59* (3): 1141–1190.

[9] Guillin A, Liu W, Wu L, et al. Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems. *The Annals of Applied Probability*, **2022**, *32* (3): 1590–1614.

[10] Kac M. Foundations of kinetic theory. In: Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Contributions to Astronomy and Physics. Berkeley, Los Angeles, USA: University of California Press, **1956**: 171–197.

[11] Sznitman A S. Topics in propagation of chaos. In: Ecole d'Eté de Probabilités de Saint-Flour XIX —1989. Berlin, Heidelberg: Springer, **1991**: 165–251.

[12] McKean H P. A class of Markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences*, **1966**, *56* (6): 1907–1911.

[13] Mishura Y, Veretennikov A. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*, **2020**, *103*: 59–101.

[14] Lacker D. On a strong form of propagation of chaos for McKean–Vlasov equations. *Electronic Communications in Probability*, **2018**, *23*: 1–11.

[15] Liu W, Wu L, Zhang C. Long-time behaviors of mean-field interacting particle systems related to McKean–Vlasov equations. *Communications in Mathematical Physics*, **2021**, *387* (1): 179–214.

[16] Andreis L, Dai Pra P, Fischer M. McKean–Vlasov limit for interacting systems with simultaneous jumps. *Stochastic Analysis and Applications*, **2018**, *36* (6): 960–995.

[17] Erny X. Well-posedness and propagation of chaos for McKean–Vlasov equations with jumps and locally Lipschitz coefficients. *Stochastic Processes and Their Applications*, **2022**, *150*: 192–214.

[18] Mehri S, Scheutzow M, Stannat W, et al. Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *The Annals of Applied Probability*, **2020**, *30* (1): 175–207.

[19] Baladron J, Fasoli D, Faugeras O, et al. Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons. *The Journal of Mathematical Neuroscience*, **2012**, *2*: 10.

[20] Bossy M, Faugeras O, Talay D. Clarification and complement to "Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons". *The Journal of Mathematical Neuroscience*, **2015**, *5*: 19.

[21] Xu L. Uniqueness and propagation of chaos for the Boltzmann equation with moderately soft potentials. *The Annals of Applied*

Probability, **2018**, *28* (2): 1136–1189.

[22] Bo L, Liao H. Probabilistic analysis of replicator–mutator equations. *Advances in Applied Probability*, **2022**, *54* (1): 167–201.

[23] Ikeda N, Watanabe S. Stochastic Differential Equations and Diffusion Processes. Amsterdam: Elsevier, **2014**.

[24] Luçon E, Stannat W. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, **2014**, *24* (5): 1946–1993.

[25] Bard Ermentrout G, Terman D H. Mathematical Foundations of Neuroscience. New York: Springer, **2010**.

[26] Hocquet A, Vogler A. Optimal control of mean field equations with monotone coefficients and applications in neuroscience. *Applied Mathematics & Optimization*, **2021**, *84* (2): 1925–1968.

[27] Garroni M G, Menaldi J L. Green Functions for Second Order Parabolic Integro-Differential Problems. New York: Chapman & Hall/CRC, **1992**.

[28] Mandelkern M. Metrization of the one-point compactification. *Proceedings of the American Mathematical Society*, **1989**, *107* (4): 1111–1115.

[29] Fournier N, Guillin A. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, **2015**, *162*: 707–738.

[30] Hardy G H, Littlewood J E, Pólya G. Inequalities. Cambridge, UK: Cambridge University Press, **1952**.

[31] Protter P E. Stochastic Integration and Differential Equations. Berlin: Springer, **2013**.