


# AN EXPLICIT MILSTEIN-TYPE SCHEME FOR INTERACTING PARTICLE SYSTEMS AND MCKEAN-VLASOV SDES WITH COMMON NOISE AND NON-DIFFERENTIABLE DRIFT COEFFICIENTS

BY SANI BISWAS<sup>1,a</sup>, CHAMAN KUMAR<sup>1,b</sup>, NEELIMA<sup>2,c</sup>, GONÇALO DOS REIS<sup>3,d</sup>  AND CHRISTOPH REISINGER<sup>4,e</sup>

<sup>1</sup>Department of Mathematics, India Institute of Technology Roorkee, <sup>a</sup>[sbiswas2@ma.iitr.ac.in](mailto:sbiswas2@ma.iitr.ac.in), <sup>b</sup>[chaman.kumar@ma.iitr.ac.in](mailto:chaman.kumar@ma.iitr.ac.in)

<sup>2</sup>Department of Mathematics, Delhi University, <sup>c</sup>[neelima\\_maths@ramjas.du.ac.in](mailto:neelima_maths@ramjas.du.ac.in)

<sup>3</sup>School of Mathematics, Edinburgh University, <sup>d</sup>[G.dosReis@ed.ac.uk](mailto:G.dosReis@ed.ac.uk)

<sup>4</sup>Mathematical Institute, Oxford University, <sup>e</sup>[christoph.reisinger@maths.ox.ac.uk](mailto:christoph.reisinger@maths.ox.ac.uk)

We propose an explicit drift-randomised Milstein scheme for both McKean–Vlasov stochastic differential equations and associated high-dimensional interacting particle systems with common noise. By using a drift randomisation step in space and measure, we establish the scheme’s strong convergence rate of 1 under reduced regularity assumptions on the drift coefficient: no classical (Euclidean) derivatives in space or measure derivatives (e.g., Lions/Fréchet) are required. The main result is established by enriching the concepts of bistability and consistency of numerical schemes used previously for standard SDE. We introduce certain Spijker-type norms (and associated Banach spaces) to deal with the interaction of particles present in the stochastic systems being analysed. A discussion of the scheme’s complexity is provided.

**1. Introduction.** For a given  $T > 0$ , consider the following stochastic differential equation (SDE) of McKean–Vlasov-type and with common noise:

$$(1) \quad \begin{aligned} X_t = X_0 &+ \int_0^t b(s, X_s, \mathcal{L}^1(X_s)) ds + \sum_{\ell=1}^{m_1} \int_0^t \sigma_1^\ell(s, X_s, \mathcal{L}^1(X_s)) dW_s^\ell \\ &+ \sum_{\ell=1}^{m_0} \int_0^t \sigma_0^\ell(s, X_s, \mathcal{L}^1(X_s)) dW_s^{0,\ell}, \end{aligned}$$

almost surely for all  $t \in [0, T]$ , where  $W := \{W_t\}_{t \in [0, T]}$  and  $W^0 := \{W_t^0\}_{t \in [0, T]}$  are, respectively,  $m_1$  and  $m_0$ -dimensional independent Wiener processes and  $\{\mathcal{L}^1(X_t)\}_{t \in [0, T]}$  denotes the stochastic flow of conditional marginal laws of  $X := \{X_t\}_{t \in [0, T]}$  given  $W^0$ . The initial value  $X_0$  is an  $\mathcal{F}^0$ -measurable random variable, independent of  $W$  and  $W^0$ . The McKean–Vlasov SDE (1) can be viewed as an infinite-dimensional system of particles with  $W$  representing the randomness inherent in the individual particle and  $W^0$  the randomness common to all the particles. When  $\sigma_0 \equiv 0$ , the particles are governed by only one source of randomness,  $W$ , and the stochastic flow  $\{\mathcal{L}^1(X_t)\}_{t \in [0, T]}$  becomes a deterministic one. Notice that McKean–Vlasov SDEs are different from standard SDEs due to the dependence of the coefficients on the (conditional) marginal law  $\mathcal{L}^1(X_t)$  of  $X_t$  given  $W^0$ , which brings additional difficulties.

Due to their wide applications in areas such as Finance, mathematical neuroscience and biology, machine learning and physics—animal swarming, cell movement induced by chemotaxis, opinion dynamics, particle movement in porous media and electrical battery modelling,

Received August 2022; revised June 2023.

*MSC2020 subject classifications.* Primary 65C30, 60H35; secondary 65C05, 65C35.

*Key words and phrases.* Stochastic interacting particle systems, McKean–Vlasov equations, common noise, Milstein scheme, nondifferentiable drift, drift randomisation, bistability.

self-assembly of particles and dynamical density functional theory (see, e.g., [2, 8, 9, 13, 14, 16, 26, 28, 29, 32–34])—McKean–Vlasov equations and associated interacting particle systems, with or without common noise, addressed via stochastic systems or associated Fokker–Plank equations (e.g., [21, 22, 39]) have gained immense popularity.

As in the case of SDEs, explicit solutions of McKean–Vlasov SDEs are typically not available, which necessitates the development of numerical schemes to approximate them. The numerical approximation of McKean–Vlasov SDEs can be carried out in two steps, as explained below.

- As a first step one builds the so-called *interacting particle system*,  $\{X_t^{i,N}\}_{i \in \{1, \dots, N\}}$ , where one replaces the (conditional) marginal law appearing in the coefficients of (1) by the empirical law obtained from the particles. Concretely, taking  $N$  i.i.d. copies  $\{W^i\}_{i \in \{1, \dots, N\}}$  of  $W$  and  $\{X_0^i\}_{i \in \{1, \dots, N\}}$  of  $X_0$ ,

one defines the interacting particle system associated with the above McKean–Vlasov SDE by

$$(2) \quad \begin{aligned} X_t^{i,N} = & X_0^i + \int_0^t b(s, X_s^{i,N}, \mu_s^{X,N}) ds + \sum_{\ell=1}^{m_1} \int_0^t \sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) dW_s^{i,\ell} \\ & + \sum_{\ell=1}^{m_0} \int_0^t \sigma_0^\ell(s, X_s^{i,N}, \mu_s^{X,N}) dW_s^{0,\ell}, \end{aligned}$$

almost surely for any  $t \in [0, T]$  and  $i \in \{1, \dots, N\}$ , where

$$\mu_t^{X,N}(\cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}(\cdot)$$

is the empirical measure of  $N$  particles. Subsequently, one needs to show, roughly put, that  $\text{Law}(X_t^{i,N})$  for some  $i$  (fixed) converges to  $\text{Law}(X_t)$  of (1) as  $N \rightarrow \infty$ ; see the seminal work by Sznitman [53] (and Proposition 2.5 below).

- In the second step of approximation, the temporal discretization of the interacting particle system is performed to obtain fully implementable numerical schemes for the McKean–Vlasov SDEs such as Euler-type schemes and Milstein-type schemes. The main difficulty is to show that all estimates are independent of the number of particles  $N$  (see Theorem 3.3 and Corollary 3.4 below). The direct application of results from classic SDE theory do not deliver this independence.

*It is critical to note that the results of this manuscript double for either the numerical approximation of McKean–Vlasov SDEs if one’s starting point is (1), or for stand-alone systems of interacting particle SDE systems if one’s starting point is (2).*

Results on the strong well-posedness and propagation of chaos for McKean–Vlasov SDEs are being extensively researched and we cannot possibly do justice to that growing body of literature. Nonetheless, we mention some of the milestones and more recent results fitting thematically with our manuscript: for PoC results, one starts from Sznitman’s seminal work [53] to the monographs [13, 14] and the recent developments by [23, 40, 41] to mention a few—overall, there are still gaps in the existing PoC results, especially across dimensions. Concretely, in [20], Section 3.5, the PoC rates across the dimension  $d$  is estimated numerically for a diffusion  $\sigma$  of polynomial growth and a drift  $b$  of polynomial growth (a setting outside the scope of this work) and the rates estimated are better than the rates found in any theoretical results at present. For a focus on well-posedness, one starts from McKean’s seminal work [44], to again [13, 14] and the recent developments [6, 17, 20, 25, 38, 45, 46] (and references therein).

The second step of numerical approximation of McKean–Vlasov SDEs saw relatively little development after the first Euler-type scheme was proposed and analysed (in a weak sense) in [10], but experience rapid advances in a string of recent papers, [3–5, 18, 19, 24, 37, 38, 42, 50]. In particular, [24] proposed an Euler-type numerical scheme for the interacting particle system associated with the McKean–Vlasov SDE, and its strong convergence is investigated when the drift coefficient grows super-linearly in the state variable. More precisely, the drift is assumed to be one-sided Lipschitz continuous in the state variable and Lipschitz continuous in the measure variable, while the diffusion coefficient is assumed to be Lipschitz continuous in both the state and measure variables. The rate of strong convergence of the scheme is shown to be equal to  $1/2$ . In [4, 37], a Milstein-type scheme for the interacting particle system associated with the McKean–Vlasov SDEs is proposed using the notion of Lions derivatives, introduced by P.-L. Lions in his lectures at the Collège de France and presented in [12], and its strong convergence is shown with rate 1. The drift coefficient is assumed to satisfy a one-sided Lipschitz condition in the state variable and a polynomial Lipschitz condition, and the diffusion coefficient to satisfy Lipschitz condition; both are assumed to be Lipschitz continuous in the measure variable. Furthermore, only the diffusion coefficient is required to be once differentiable (in both state and measure variables). In [4], the authors additionally require second order differentiability of the coefficients. In [38], an Euler-type scheme and a Milstein-type scheme are developed for the interacting particle system connected with McKean–Vlasov SDEs with common noise, where all the coefficients are allowed to grow super-linearly in the state variable.

In this article, we develop a Milstein-type scheme for the interacting particle system corresponding to the McKean–Vlasov SDE without assuming the differentiability of the drift coefficient (in space or measure component), which is therefore more relaxed than the corresponding results in [4, 37, 38]. It is out of the scope of this work to lift the differentiability conditions on the diffusion coefficient and hence those assumptions match those already existing in the most recent literature. The relaxation of the regularity requirement of the drift coefficient is achieved by a certain randomisation strategy that needs to be applied to *both state and measure components*: the technical developments necessary to deal with this difficulty are the second contribution of this manuscript. In the case of SDEs (when the coefficients do not depend on the law of the solution process), the technique of randomisation has been studied in [36] (also [7, 35] and more recently in [47, 49]) to construct a Milstein-type scheme without assuming the first order differentiability of the drift coefficient. As the coefficients in our settings depend on the law of the solution process as well, we require a two-fold randomisation—one with respect to the state variable and the other with respect to the measure variable. For this, we use a uniform random variable to generate a random point in each sub-interval of the time mesh and the Euler scheme is used to obtain values of the particles' states at these random points, which are then used in the drift coefficient of the Milstein scheme, both for the state and empirical measure. The precise details of this randomisation can be found in Section 3. Critically, the technique developed in [36] for the analysis cannot be used directly in our settings and a novel approach is required. We observe the following.

- The interacting particle system associated with the McKean–Vlasov SDE can be treated as an  $(\mathbb{R}^d)^N$ -dimensional SDE and thus the results of [36] could be applied directly. However, all estimates would depend on  $N$  and hence “explode” as  $N$  tends to infinity. This implies that to establish our results a new tool must be developed in order to show the independence on  $N$ .
- We propose a new notion of bistability and consistency of the numerical scheme that is appropriate for the context of high-dimensional interacting particle systems. Inspired by [36] we propose suitable stochastic Spijker norms capable of dealing with the interaction component of the particles. Further details can be found in Section 5.

- A discussion on the practicalities of implementing our scheme is given in Section 3.1.1 which includes a critical view on complexity and the consequence of having common-noise.
- In the simplest version possible of (1), three drift functions are well within the scope of our work are (linear interaction, convolution-functionals and linear interaction kernels)

$$b(s, X_s, \mathcal{L}^1(X_s)) = f_1(X_s) + f_2(\tilde{\mathbb{E}}^1(g(X_s))),$$

$$\hat{b}(s, X_s, \mathcal{L}^1(X_s)) = f_1(X_s) + \int_{\mathbb{R}^d} g(X_s - y) \mathcal{L}^1(X_s)(dy),$$

$$\tilde{b}(s, X_s, \mathcal{L}^1(X_s)) = f_1(X_s) + \int_{\mathbb{R}^d} \tilde{g}(X_s, y) \mathcal{L}^1(X_s)(dy),$$

for any real-valued functions  $f_1, f_2, g, \tilde{g}$  that satisfy a standard Lipschitz condition in space (but are not differentiable), for example,  $g(x) = -|x|$  and more complex examples for  $g, \tilde{g}$  can be found in [13, 14, 16, 31, 33] and  $\tilde{\mathbb{E}}^1$  represents the conditional expectation given the common noise  $W^0$ . The 2nd drift example,  $\hat{b}$ , corresponds to the usual convolution operator fairly common in modelling with McKean–Vlasov SDE and associated interacting particle systems (with or without common noise) [1, 15, 31, 33]. Last, we point the reader to Example 3.15 in [1] that intuitively highlights why  $\hat{b}$  is Lipschitz in the Wasserstein metric but is not Lions differentiable.

*Organization.* The main framework of the McKean–Vlasov equation and the interacting particle system including well-posedness and propagation of chaos is given in Section 2. The numerical scheme focusing on the approximation of the interacting particle system is found in Section 3 as are the main convergence results. All proofs are given in Section 5.

**1.1. Notation.** Both the Euclidean norm on  $\mathbb{R}^d$  and the standard matrix norm on  $\mathbb{R}^{d \times m}$  are denoted by  $|\cdot|$ . The notation  $\delta_x$  stands for the Dirac measure centred at  $x \in \mathbb{R}^d$ . We use the same notation  $a^\ell$  to denote the  $\ell$ th column of a matrix  $a \in \mathbb{R}^{d \times m}$  and the  $\ell$ th element of a vector  $a \in \mathbb{R}^d$ .  $\mathcal{B}(\chi)$  stands for the Borel  $\sigma$ -algebra on a topological space  $\chi$ . Further,  $\mathcal{P}_2(\mathbb{R}^d)$  denotes the space of all probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  having finite second moment and equipped with the  $\mathcal{L}^2$ -Wasserstein metric given by

$$\mathcal{W}_2(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left[ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) \right]^{1/2},$$

for any  $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ , where  $\Pi(\mu, \nu)$  denotes the set of couplings of  $\mu$  and  $\nu$ . Clearly,  $\mathcal{P}_2(\mathbb{R}^d)$  is a Polish space under this metric. We use  $\mathcal{L}^p(\Omega)$  to denote the Banach space of all  $\mathbb{R}^d$ -valued random variables  $Y$  defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and satisfies  $\|Y\|_{\mathcal{L}^p(\Omega)} := [\mathbb{E}|Y|^p]^{1/p} < \infty$ , where  $\mathbb{E}$  stands for expectation with respect to  $\mathbb{P}$ . Similarly, we use  $\mathcal{L}^p([0, T] \times \Omega)$  to denote the Banach space of processes  $Y : [0, T] \times \Omega \mapsto \mathbb{R}^d$  having  $\|Y\|_{\mathcal{L}^p([0, T] \times \Omega)} = [\int_0^T \|Y(s)\|_{\mathcal{L}^p(\Omega)}^p ds]^{1/p} < \infty$ . Also,  $\mathcal{C}^\alpha([0, T], \mathcal{L}^p(\Omega))$  stands for the space of all  $\alpha$ -Hölder continuous functions  $Y : [0, T] \mapsto \mathcal{L}^p(\Omega)$  with the following norm:

$$(3) \quad \|Y\|_{\mathcal{C}^\alpha([0, T], \mathcal{L}^p(\Omega))} = \sup_{t \in [0, T]} \|Y(t)\|_{\mathcal{L}^p(\Omega)} + \sup_{\substack{t, t' \in [0, T] \\ t \neq t'}} \frac{\|Y(t) - Y(t')\|_{\mathcal{L}^p(\Omega)}}{|t - t'|^\alpha}.$$

For a function  $f : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ ,  $\partial_x f : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}^d$  is the derivative of  $f$  with respect to the space variable and  $\partial_\mu f : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the Lions derivative with respect to the measure variable.  $I_A$  stands for the indicator function of a set  $A$  and  $\tilde{\mathbb{N}} = \mathbb{N} \cup \{0\}$ . The constants that appear in the paper vary from line to line, will depend on the problems data, for instance  $T, m_0, m_1$ , etc., but critically are independent of the number of particle  $N$  and the schemes timestep  $h$  (given below).

**2. McKean–Vlasov stochastic differential equations and the interacting particle systems.** Let  $T > 0$  be a fixed constant. Consider probability spaces  $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$  and  $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$  equipped with filtrations  $\mathbb{F}^1 := \{\mathcal{F}_t^1\}_{t \in [0, T]}$  and  $\mathbb{F}^0 := \{\mathcal{F}_t^0\}_{t \in [0, T]}$ , respectively. The filtrations  $\mathbb{F}^1$  and  $\mathbb{F}^0$  satisfy the usual conditions, that is, they are complete and right continuous. Assume that  $W := \{W_t \in \mathbb{R}^{m_1}\}_{t \in [0, T]}$  and  $W^0 := \{W_t^0 \in \mathbb{R}^{m_0}\}_{t \in [0, T]}$  are independent Brownian motions defined on  $(\Omega^1, \mathcal{F}^1, \mathbb{F}^1, \mathbb{P}^1)$  and  $(\Omega^0, \mathcal{F}^0, \mathbb{F}^0, \mathbb{P}^0)$ , respectively. In what follows, the interacting particles are governed by i.i.d. copies of  $W$  and  $W^0$  represents the noise common to all the particles. Let us define a product probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ , where  $\tilde{\Omega} := \Omega^1 \times \Omega^0$ ,  $(\tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  is the completion of  $(\mathcal{F}^1 \otimes \mathcal{F}^0, \mathbb{P}^1 \otimes \mathbb{P}^0)$  and  $\tilde{\mathbb{F}} := \{\tilde{\mathcal{F}}_t\}_{t \in [0, T]}$  is the completion and right-continuous augmentation of  $\{\mathcal{F}_t^1 \otimes \mathcal{F}_t^0\}_{t \in [0, T]}$ . The expectation with respect to  $\tilde{\mathbb{P}}$  is denoted by  $\tilde{\mathbb{E}}$ .

Let  $b : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}^d$ ,  $\sigma_1 : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}^{d \times m_1}$  and  $\sigma_0 : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}^{d \times m_0}$  be  $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{P}_2(\mathbb{R}^d))$ -measurable functions.

In this article, we consider the following  $\mathbb{R}^d$ -valued McKean–Vlasov stochastic differential equations (SDEs) with common noise defined on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ :

$$(4) \quad \begin{aligned} X_t = X_0 &+ \int_0^t b(s, X_s, \mathcal{L}^1(X_s)) ds + \sum_{\ell=1}^{m_1} \int_0^t \sigma_1^\ell(s, X_s, \mathcal{L}^1(X_s)) dW_s^\ell \\ &+ \sum_{\ell=1}^{m_0} \int_0^t \sigma_0^\ell(s, X_s, \mathcal{L}^1(X_s)) dW_s^{0, \ell}, \end{aligned}$$

almost surely for all  $t \in [0, T]$ , where  $\{\mathcal{L}^1(X_t)\}_{t \in [0, T]}$  is the flow of conditional laws of  $X_t$  given  $W^0$  and  $X_0$ . A priori, it is not certain that the flow of conditional marginals  $\{\mathcal{L}^1(X_t)\}_{t \in [0, T]}$  is an  $\mathbb{F}^0$ -adapted continuous process. However, due to Lemma 2.5 in [14], when the McKean–Vlasov SDE (4) has a unique  $\tilde{\mathbb{F}}$ -adapted continuous solution with uniformly bounded second moment, then  $\{\mathcal{L}^1(X_t)\}_{t \in [0, T]}$  is an  $\mathbb{F}^0$ -adapted continuous process.

We make the following assumptions.

ASSUMPTION H-1.  $X_0 \in \mathcal{L}^{\tilde{\mathbb{P}}}(\tilde{\Omega})$  for some  $\bar{p} \geq 2$ .

ASSUMPTION H-2. There exists a constant  $L > 0$  such that

$$|b(t, x, \mu) - b(t, x', \mu')| + \sum_{u=0}^1 |\sigma_u(t, x, \mu) - \sigma_u(t, x', \mu')| \leq L\{|x - x'| + \mathcal{W}_2(\mu, \mu')\},$$

for all  $t \in [0, T]$ ,  $x, x' \in \mathbb{R}^d$  and  $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ .

ASSUMPTION H-3. There exists a constant  $L > 0$  such that

$$|b(t, x, \mu) - b(t', x, \mu)| \leq L\{1 + |x| + \mathcal{W}_2(\mu, \delta_0)\}|t - t'|^{1/2},$$

$$\sum_{u=0}^1 |\sigma_u(t, x, \mu) - \sigma_u(t', x, \mu)| \leq L\{1 + |x| + \mathcal{W}_2(\mu, \delta_0)\}|t - t'|,$$

for all  $t, t' \in [0, T]$ ,  $x \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ .

REMARK 2.1. From Assumptions H-2 and H-3, for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ ,

$$|b(t, x, \mu)| + \sum_{u=0}^1 |\sigma_u(t, x, \mu)| \leq \bar{L}\{1 + |x| + \mathcal{W}_2(\mu, \delta_0)\},$$

where  $\bar{L} = \max\{L, LT, L\sqrt{T}, b(0, 0, \delta_0), \sigma_0(0, 0, \delta_0), \sigma_1(0, 0, \delta_0)\}$ .

The proof of the following proposition can be found in [14, 38] and Appendix A.1.

**PROPOSITION 2.2** (Well-posedness and moment bounds). *If Assumptions H-1 with  $\bar{p} \geq 2$ , H-2 and H-3 are satisfied, then the McKean–Vlasov SDE (4) has a unique  $\tilde{\mathbb{F}}$ -adapted solution  $\{X_t\}_{t \in [0, T]}$  and*

$$\left\| \sup_{t \in [0, T]} |X_t| \right\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \leq C_1 (1 + \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}}),$$

for  $C_1 := 4^{\bar{p}-1} \max\{1, 3^{\bar{p}-1} \bar{L}^{\bar{p}}(T^{\bar{p}} + 2(\frac{\bar{p}^3 T}{2(\bar{p}-1)})^{\bar{p}/2})\} \exp(12^{\bar{p}-1} \bar{L}^{\bar{p}}(2T^{\bar{p}} + 4(\frac{\bar{p}^3 T}{2(\bar{p}-1)})^{\bar{p}/2}))$ .

To introduce the interacting particle system connected with the McKean–Vlasov SDEs (4), let us consider  $N \in \mathbb{N}$  i.i.d. copies of  $W$  and  $X_0$ , denoted by  $W^i$  and  $X_0^i$  for  $i \in \{1, \dots, N\}$ , respectively. Define the following system of equations:

$$(5) \quad \begin{aligned} X_t^i &= X_0^i + \int_0^t b(s, X_s^i, \mathcal{L}^1(X_s^i)) ds + \sum_{\ell=1}^{m_1} \int_0^t \sigma_1^\ell(s, X_s^i, \mathcal{L}^1(X_s^i)) dW_s^{i, \ell} \\ &\quad + \sum_{\ell=1}^{m_0} \int_0^t \sigma_0^\ell(s, X_s^i, \mathcal{L}^1(X_s^i)) dW_s^{0, \ell}, \end{aligned}$$

almost surely for any  $t \in [0, T]$  and  $i \in \{1, \dots, N\}$ . Notice that due to Proposition 2.11 in [14],  $\mathbb{P}^0(\mathcal{L}^1(X_t^1) = \mathcal{L}^1(X_t^i))$  for all  $t \in [0, T] = 1$ . We remark that the proof of Proposition 2.11 in [14], which establishes the result in a more restrictive setting, uses (only) the well-posedness of the system (4) and Theorem 1.33 from [14] (Yamada Watanabe theorem) and thus particles have the same law under our settings also. The system (5) is popularly known as the *conditional noninteracting particle system*. On approximating  $\mathcal{L}^1(X_t^1)$  by the empirical measure of the states  $\{X_t^i\}_{i \in \{1, \dots, N\}}$  of  $N$  particles at time  $t$ , and denoted  $\mu_t^{X, N}$ , one obtains the following *system of interacting particles*:

$$(6) \quad \begin{aligned} X_t^{i, N} &= X_0^i + \int_0^t b(s, X_s^{i, N}, \mu_s^{X, N}) ds + \sum_{\ell=1}^{m_1} \int_0^t \sigma_1^\ell(s, X_s^{i, N}, \mu_s^{X, N}) dW_s^{i, \ell} \\ &\quad + \sum_{\ell=1}^{m_0} \int_0^t \sigma_0^\ell(s, X_s^{i, N}, \mu_s^{X, N}) dW_s^{0, \ell}, \end{aligned}$$

almost surely for any  $t \in [0, T]$  and  $i \in \{1, \dots, N\}$ , where

$$\mu_t^{X, N}(\cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i, N}}(\cdot),$$

is the empirical measure of  $N$  particles.

**REMARK 2.3.** The system (6) can be understood as an  $\mathbb{R}^{d \times N}$ -dimensional SDE and thus its well-posedness and moment stability up to order  $\bar{p}$  follow from [30] under Assumptions H-1 (with  $\bar{p} \geq 2$ ), H-2 and H-3. In other words, the interacting particle system (6) connected with the McKean–Vlasov SDE (4) has a unique strong solution  $\{X_t^{i, N}\}_{t \in [0, T]}$  adapted to the filtration  $\{\tilde{\mathcal{F}}_t\}_{t \in [0, T]}$  and one can show that

$$\max_{i \in \{1, \dots, N\}} \left\| \sup_{t \in [0, T]} |X_t^{i, N}| \right\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \leq C_1 \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right),$$

where  $C_1$  is the same constant as in Proposition 2.2. Note that the RHS of the moment estimate above is *critically* independent of  $N$ —this follows from [38] but does not from [30].



The following lemma gives the time-regularity of the interacting particle system (6) and its proof is given in Appendix A.2.

LEMMA 2.4. *Let Assumptions H-1 with  $\bar{p} \geq 2$ , H-2 and H-3 hold. Then,*

$$\max_{i \in \{1, \dots, N\}} \|X_t^{i,N} - X_{t'}^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \leq C_2 |t - t'|^{\bar{p}/2} \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}}\right),$$

for all  $t > t' \in [0, T]$  and  $N \in \mathbb{N}$ , where  $C_2 := (1 + 2C_1)9^{\bar{p}-1}L^{\bar{p}}(T^{\bar{p}/2} + 2(\frac{\bar{p}(\bar{p}-1)}{2})^{\bar{p}/2})$  and the constant  $C_1$  is defined in Proposition 2.2.

The convergence of the interacting particle system (6) to the noninteracting particle system (5) is popularly known in the literature as the *propagation of chaos* and is stated in the following proposition (see Theorem 2.12 in [14] for details or Appendix A.3). For this, let us define the empirical measure of the noninteracting particle system (5) as

$$\mu_t^X(\cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}(\cdot)$$

almost surely for any  $t \in [0, T]$  and  $N \in \mathbb{N}$ . Further, use Theorem 5.8 in [13] and Proposition 2.2 to obtain the following estimate:

$$(7) \qquad \|W_2(\mathcal{L}^1(X_t^1), \mu_t^X)\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \leq C_3 \begin{cases} N^{-1/2} & \text{if } d < 4, \\ N^{-1/2} \log_2 N & \text{if } d = 4, \\ N^{-2/d} & \text{if } d > 4, \end{cases}$$

for any  $t \in [0, T]$  and  $N \in \mathbb{N}$ , where  $C_3 := C_3(d, \bar{p}, \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}})$  is a positive constant.

PROPOSITION 2.5 (Propagation of chaos). *Let Assumptions H-1 with  $\bar{p} > 4$ , H-2 and H-3 hold. Then,*

$$\max_{i \in \{1, \dots, N\}} \left\| \sup_{t \in [0, T]} \|X_t^i - X_t^{i,N}\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \right\| \leq C_4 \begin{cases} N^{-1/2} & \text{if } d < 4, \\ N^{-1/2} \log_2 N & \text{if } d = 4, \\ N^{-2/d} & \text{if } d > 4, \end{cases}$$

where  $C_4 := 4(3T + 24)TL^2e^{6(3T+24)TL^2}C_3(d, \bar{p}, \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}})$ , where the constant  $C_3$  appears in (7).

The proof of this result can be found in Appendix A.3.

All in all, the PoC rates here are in line with the PoC results found in the numerics for McKean Vlasov SDE literature [3–5, 18, 19, 24, 25, 37, 38, 42, 50]. There are improved results obtaining rate of  $1/N$  and  $1/N^2$  (e.g., [23, 40]) that hold under stricter conditions that do not fit the scope of the work presented in the manuscript. We have added text to this effect in the main body of the paper.

**3. Main result.** The drift-randomised Milstein scheme was originally proposed and analysed for standard SDEs with nondifferentiable drift coefficient in [36]. In the case of McKean–Vlasov SDEs, for which the coefficients depend on the law of the solution process as well, one needs additional randomisation of the drift coefficient with respect to the measure component and a way to deal with the implications from interacting particles. The associated difficulties are tackled in this paper.

**3.1. The scheme.** In order to propose the randomised Milstein scheme for the interacting particle system (6) connected with the McKean–Vlasov SDEs (5), the map  $(x, \mu) \mapsto$

$(\sigma_1(t, x, \mu), \sigma_0(t, x, \mu))$  is assumed to be continuously differentiable for every  $t \in [0, T]$ . The notion of Lions derivative (see Appendix A.4 for details) is used to differentiate these functions with respect to the measure component. For  $u \in \{0, 1\}$ , let us define  $d \times d$  matrices  $\Lambda_{\sigma_u \sigma_1}$  and  $\Lambda_{\sigma_u \sigma_0}$  whose  $\ell$ th column are given by,

$$(8) \quad \begin{aligned} \Lambda_{\sigma_u \sigma_1}^\ell(t, s, x^i, \mu) &:= \sum_{\ell_1=1}^{m_1} \int_s^t \partial_x \sigma_u^\ell(s, x^i, \mu) \sigma_1^{\ell_1}(s, x^i, \mu) dW_r^{i, \ell_1} \\ &\quad + \frac{1}{N} \sum_{k=1}^N \sum_{\ell_1=1}^{m_1} \int_s^t \partial_\mu \sigma_u^\ell(s, x^i, \mu, x^k) \sigma_1^{\ell_1}(s, x^k, \mu) dW_r^{k, \ell_1}, \end{aligned}$$

$$(9) \quad \begin{aligned} \Lambda_{\sigma_u \sigma_0}^\ell(t, s, x^i, \mu) &:= \sum_{\ell_1=1}^{m_0} \int_s^t \partial_x \sigma_u^\ell(s, x^i, \mu) \sigma_0^{\ell_1}(s, x^i, \mu) dW_r^{0, \ell_1} \\ &\quad + \frac{1}{N} \sum_{k=1}^N \sum_{\ell_1=1}^{m_0} \int_s^t \partial_\mu \sigma_u^\ell(s, x^i, \mu, x^k) \sigma_0^{\ell_1}(s, x^k, \mu) dW_r^{0, \ell_1}, \end{aligned}$$

for all  $s, t \in [0, T]$ ,  $\ell \in \{1, \dots, m_u\}$ ,  $x^i \in \mathbb{R}^d$ ,  $i \in \{1, \dots, N\}$ ,  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ . Also, define the  $d$ -dimensional vectors  $\tilde{\sigma}_1^\ell$  and  $\tilde{\sigma}_0^{\ell_1}$  for any  $\ell \in \{1, \dots, m_1\}$  and  $\ell_1 \in \{1, \dots, m_0\}$  as

$$(10) \quad \left. \begin{aligned} \tilde{\sigma}_1^\ell(t, s, x^i, \mu) &:= \sigma_1^\ell(s, x^i, \mu) + \Lambda_{\sigma_1 \sigma_1}^\ell(t, s, x^i, \mu) + \Lambda_{\sigma_1 \sigma_0}^\ell(t, s, x^i, \mu) \\ \tilde{\sigma}_0^{\ell_1}(t, s, x^i, \mu) &:= \sigma_0^{\ell_1}(s, x^i, \mu) + \Lambda_{\sigma_0 \sigma_1}^{\ell_1}(t, s, x^i, \mu) + \Lambda_{\sigma_0 \sigma_0}^{\ell_1}(t, s, x^i, \mu) \end{aligned} \right\}$$

for all  $i \in \{1, \dots, n_h\}$ .

Now, consider a sequence  $\eta := \{\eta_j\}_{j \in \mathbb{N}}$  of i.i.d. standard uniformly distributed random variables defined on a probability space  $(\Omega^\eta, \mathcal{F}^\eta, \mathbb{P}^\eta)$ , equipped with the natural filtration  $\mathbb{F}^\eta := \{\mathcal{F}_j^\eta\}_{j \in \mathbb{N}}$  of  $\{\eta_j\}_{j \in \mathbb{N}}$ . Let  $\mathbb{E}^\eta$  stand for the expectation with respect to  $\mathbb{P}^\eta$ . The random variables  $\{\eta_j\}_{j \in \mathbb{N}}$  are assumed to be independent of  $W, W^0, W^i$  and  $X_0^i$  for all  $i \in \{1, \dots, N\}$ . Now consider the general nonequidistant temporal grid  $\mathcal{Q}_h$  of  $[0, T]$  with  $n_h$  subintervals,

$$(11) \quad \mathcal{Q}_h = \{(t_0, t_1, \dots, t_{n_h}) : 0 = t_0 < t_1 < \dots < t_{n_h} = T\},$$

with  $h_j := (t_j - t_{j-1}) > 0$  for  $j \in \{1, \dots, n_h\}$ ,  $n_h \in \mathbb{N}$  and  $h := \max_{j \in \{1, \dots, n_h\}} h_j \leq \min(1, T)$ . Now, consider a new probability space  $(\Omega, \mathcal{F}, \mathbb{P}) = (\tilde{\Omega} \times \Omega^\eta, \tilde{\mathcal{F}} \otimes \mathcal{F}^\eta, \tilde{\mathbb{P}} \otimes \mathbb{P}^\eta)$  equipped with a filtration  $\mathbb{F} := \{\mathcal{F}_j^h\}_{j \in \{0, \dots, n_h\}}$ ; where  $\mathcal{F}_j^h = \tilde{\mathcal{F}}_{t_j} \otimes \mathcal{F}_j^\eta$ . We will denote the expectation with respect to  $\mathbb{P}$  by  $\mathbb{E}$ .

The drift-randomised Milstein scheme for the interacting particle system (6) of the McKean–Vlasov SDE (5) is given by

$$(12) \quad \begin{aligned} X_{j, \eta}^{i, N, h} &= X_{j-1}^{i, N, h} + \eta_j h_j b(t_{j-1}, X_{j-1}^{i, N, h}, \mu_{j-1}^{X, N, h}) \\ &\quad + \sum_{\ell=1}^{m_1} \sigma_1^\ell(t_{j-1}, X_{j-1}^{i, N, h}, \mu_{j-1}^{X, N, h}) \int_{t_{j-1}}^{t_{j-1} + \eta_j h_j} dW_s^{i, \ell} \\ &\quad + \sum_{\ell=1}^{m_0} \sigma_0^\ell(t_{j-1}, X_{j-1}^{i, N, h}, \mu_{j-1}^{X, N, h}) \int_{t_{j-1}}^{t_{j-1} + \eta_j h_j} dW_s^{0, \ell}, \end{aligned}$$



$$\begin{aligned}
 X_j^{i,N,h} &= X_{j-1}^{i,N,h} + h_j b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}) \\
 &+ \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^{i,\ell} \\
 &+ \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^{0,\ell},
 \end{aligned}
 \tag{13}$$

almost surely for all  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$  with the initial value  $X_0^{i,N,h} = X_0^i$ , where the empirical measures  $\mu_{j-1}^{X,N,h}$  and  $\mu_{j,\eta}^{X,N,h}$  are defined as

$$\mu_{j-1}^{X,N,h}(\cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_{j-1}^{i,N,h}}(\cdot) \quad \text{and} \quad \mu_{j,\eta}^{X,N,h}(\cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_{j,\eta}^{i,N,h}}(\cdot).
 \tag{14}$$

**REMARK 3.1** (Comparison to prior art: classical case). When  $\sigma_0 \equiv 0$ , the conditional law  $\mathcal{L}^1(\cdot)$  becomes the unconditional law of the solution process. If either  $\mathcal{L}^1(\cdot)$  is known or  $b$  and  $\sigma_1$  do not depend on  $\mathcal{L}^1(\cdot)$ , then the McKean–Vlasov SDE (4) becomes a standard SDE. In such a case, the randomised Milstein scheme (13) considered here reduces to the one considered in [36]. Indeed, the terms involving measure derivatives in (8) and hence in (10), (12) and (13) vanish in this case. Additional terms that appear in (12) and (13) are due to the dependence of the coefficients  $b$ ,  $\sigma_1$  and  $\sigma_0$  on the law of the solution process (measure variable).

**3.1.1. Practical implementation.** In this section, we comment on implementation issues, specifically the sampling of the random time mesh and simulation of the Lévy area. The latter results explicitly from the presence of common noise and is not specific to our randomisation method, the former is an essential aspect of the scheme.

We first note that in equations (12) and (13) the same uniform random variables  $\eta_1, \dots, \eta_{n_h}$  are used for each particle in the system to identify random points in each sub-interval of the temporal grid  $\rho_h$ . This approach is similar to the adaptive time-stepping Euler scheme of [50] where the same random points (arising due to adaptive step-sizes) are used for each particle of the system. A refined version uses different time meshes for individual particles, but common, uniform timesteps for the definition of the empirical measure. In our setting, for each realisation of the common noise, a single path of random timesteps is simulated, so that the computational effort of the randomisation is comparable to the simulation of a discrete Brownian path, but negligible compared to the simulation of iterated stochastic integrals as required for the Milstein scheme, which we discuss further now.

If the following, commutative conditions are imposed on the diffusion coefficients:

$$\partial_x \sigma_1^\ell(t, x, \mu) \sigma_1^{\ell_1}(t, x, \mu) = \partial_x \sigma_1^{\ell_1}(t, x, \mu) \sigma_1^\ell(t, x, \mu),$$

for any  $\ell, \ell_1 = 1, \dots, m_1$ ,  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , then

$$\begin{aligned}
 &\sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_1} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{i,\ell_1} dW_s^{i,\ell} \\
 &= \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{i,\ell} dW_s^{i,\ell} \\
 &+ \sum_{\ell=1}^{m_1} \sum_{\ell_1 > \ell}^{m_1} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^s (\partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\
 &\quad - \partial_x \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})) \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{i,\ell_1} dW_s^{i,\ell}
 \end{aligned}$$

$$\begin{aligned}
& \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{i,\ell_1} dW_s^{i,\ell} \\
& + \partial_x \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{i,\ell} dW_s^{i,\ell_1} \\
& = \sum_{\ell=1}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \frac{1}{2} ((\Delta W^{i,\ell})^2 - h_j) \\
& \quad + \sum_{\ell=1}^{m_1} \sum_{\ell_1 > \ell}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \Delta W^{i,\ell} \Delta W^{i,\ell_1} \\
& = \frac{1}{2} \sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\
& \quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) (\Delta W^{i,\ell} \Delta W^{i,\ell_1} - h_j I_{\{\ell=\ell_1\}}),
\end{aligned}$$

almost surely for any  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ , and thus one can write the third term on the right-hand side of (13) as

$$\begin{aligned}
& \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^{i,\ell} \\
& = \sum_{\ell=1}^{m_1} \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \Delta W^{i,\ell} \\
& \quad + \frac{1}{2} \sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\
& \quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) (\Delta W^{i,\ell} \Delta W^{i,\ell_1} - h_j I_{\{\ell=\ell_1\}}) \\
& \quad + \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \frac{1}{N} \sum_{k=1}^N \sum_{\ell_1=1}^{m_1} \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j-1}^{k,N,h}) \\
& \quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{k,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{k,\ell_1} dW_s^{i,\ell} \\
& \quad + \sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_0} \int_{t_{j-1}}^{t_j} \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\
& \quad \times \sigma_0^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{0,\ell_1} dW_s^{i,\ell} \\
& \quad + \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \frac{1}{N} \sum_{k=1}^N \sum_{\ell_1=1}^{m_0} \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j-1}^{k,N,h}) \\
& \quad \times \sigma_0^{\ell_1}(t_{j-1}, X_{j-1}^{k,N,h}, \mu_{j-1}^{X,N,h}) dW_r^{0,\ell_1} dW_s^{i,\ell}.
\end{aligned}$$

In the above, the last three terms require the approximation of the Lévy area, which can be done with the help of the techniques developed in [54]. A similar conclusion holds for the fourth term on the right-hand side of (13). Furthermore, the terms involving Lions derivatives are of order  $\mathcal{O}(1/N)$ , as shown in [4], Proof of Proposition 2.3, for the regular case, and hence can be ignored when  $N$  is large. In addition, if  $\sigma_0 \equiv 0$ , that is, the common noise term is not present, the fourth term on the right-hand side of the above equation can also be dropped and

thus we have

$$\begin{aligned} & \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^{i,\ell} \\ &= \sum_{\ell=1}^{m_1} \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \Delta W^{i,\ell} \\ & \quad + \frac{1}{2} \sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\ & \quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) (\Delta W^{i,\ell} \Delta W^{i,\ell_1} - h_j I_{\{\ell=\ell_1\}}), \end{aligned}$$

which leads to a fully implementable randomised Milstein scheme. Summing up, when  $N$  is large and  $\sigma_0 \equiv 0$ , the randomised Milstein scheme can be reduced to

$$\begin{aligned} X_{j,\eta}^{i,N,h} &= X_{j-1}^{i,N,h} + \eta_j h_j b(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\ & \quad + \sum_{\ell=1}^{m_1} \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) (W_{t_{j-1}+\eta_j h_j}^{i,\ell} - W_{t_{j-1}}^{i,\ell}), \\ X_j^{i,N,h} &= X_{j-1}^{i,N,h} + h_j b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}) \\ & \quad + \sum_{\ell=1}^{m_1} \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \Delta W^{i,\ell} \\ & \quad + \frac{1}{2} \sum_{\ell=1}^{m_1} \sum_{\ell_1=1}^{m_1} \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \\ & \quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) (\Delta W^{i,\ell} \Delta W^{i,\ell_1} - h_j I_{\{\ell=\ell_1\}}), \end{aligned}$$

almost surely for any  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ .

**3.2. The main convergence result and its assumptions.** In order to investigate the rate of convergence of the randomised Milstein scheme (13), we make the following additional assumptions.

**ASSUMPTION H-4.** There exists a constant  $L > 0$  such that

$$|\partial_x \sigma_u^\ell(t, x, \mu) - \partial_x \sigma_u^\ell(t, x', \mu')| \leq L\{|x - x'| + \mathcal{W}_2(\mu, \mu')\},$$

$$|\partial_\mu \sigma_u^\ell(t, x, \mu, y) - \partial_\mu \sigma_u^\ell(t, x', \mu', y')| \leq L\{|x - x'| + |y - y'| + \mathcal{W}_2(\mu, \mu')\},$$

for all  $u \in \{0, 1\}$ ,  $\ell \in \{1, \dots, m_u\}$ ,  $t \in [0, T]$ ,  $x, x', y, y' \in \mathbb{R}^d$  and  $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ .

**ASSUMPTION H-5.** There exists a constant  $L > 0$  such that

$$|\partial_x \sigma_u^\ell(t, x, \mu) \sigma_v^{\ell_1}(t, x, \mu) - \partial_x \sigma_u^\ell(t, x', \mu') \sigma_v^{\ell_1}(t, x', \mu')| \leq L\{|x - x'| + \mathcal{W}_2(\mu, \mu')\},$$

$$\begin{aligned} |\partial_\mu \sigma_u^\ell(t, x, \mu, y) \sigma_v^{\ell_1}(t, y, \mu) - \partial_\mu \sigma_u^\ell(t, x', \mu', y') \sigma_v^{\ell_1}(t, y', \mu')| &\leq L\{|x - x'| + |y - y'| \\ & \quad + \mathcal{W}_2(\mu, \mu')\}, \end{aligned}$$

for all  $u, v \in \{0, 1\}$ ,  $\ell \in \{1, \dots, m_u\}$ ,  $\ell_1 \in \{1, \dots, m_v\}$ ,  $t \in [0, T]$ ,  $x, x', y, y' \in \mathbb{R}^d$  and  $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ .

REMARK 3.2. Due to Assumption H-2, we have

$$|\partial_x \sigma_u^\ell(t, x, \mu)| + |\partial_\mu \sigma_u^\ell(t, x, \mu, y)| \leq L,$$

for all  $u \in \{0, 1\}$ ,  $\ell \in \{1, \dots, m_u\}$ ,  $t \in [0, T]$ ,  $x, y \in \mathbb{R}^d$  and  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ .

Below, we state the main result of this paper containing the error rate for the approximation of the scheme to the interacting particle system (6). For notational simplicity, we use  $X_j^{i,N}$  to represent  $X_{t_j}^{i,N}$  for any  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . This result is proved in Section 5.4.

THEOREM 3.3. Let Assumptions H-1 with  $\bar{p} \geq 4$ , H-2 to H-5 hold. Then, the drift-randomised Milstein scheme (13) converges in the strong sense to the true solution of the interacting particle system (6) with order 1. Concretely, for  $q = \bar{p}/2$  and  $h \leq \min(1, T)$  we have

$$\max_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{0, \dots, n_h\}} |X_j^{i,N} - X_j^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \leq C_7 C_{10} h,$$

where the positive constants  $C_7$  and  $C_{10}$  appear in Proposition 5.5 and Proposition 5.9, respectively.

A direct combination of Proposition 2.5 and Theorem 3.3 delivers the control on the error for the numerical approximation of the McKean–Vlasov equation SDE (4) (and (5)). For convenience of notation,  $X_j^i$  is used to denote  $X_{t_j}^i$  for any  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ .

COROLLARY 3.4. Let assumptions of Theorem 3.3 hold with  $\bar{p} > 4$ . Then, the drift randomized Milstein scheme (13) converges in the strong sense to the true solution of the McKean–Vlasov SDE (4),

$$\max_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{0, \dots, n_h\}} |X_j^i - X_j^{i,N,h}| \right\|_{\mathcal{L}^2(\Omega)} \leq \begin{cases} \sqrt{C_4} N^{-1/4} + C_7 C_{10} h & \text{if } d < 4 \\ \sqrt{C_4} N^{-1/4} \sqrt{\log_2 N} + C_7 C_{10} h & \text{if } d = 4 \\ \sqrt{C_4} N^{-1/d} + C_7 C_{10} h & \text{if } d > 4 \end{cases}$$

with  $h \leq \min\{1, T\}$  and  $C_4 := 2(3T + 24)TL^2 e^{4(3T+24)TL^2} C_3(d, \bar{p}, \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}})$  where the constant  $C_3$  appears in (7), and  $C_7$  and  $C_{10}$  come from Proposition 5.5 and Proposition 5.9 respectively.

REMARK 3.5 (Application to optimal control and machine learning). In an  $n$ -player stochastic differential game with an  $(\mathbb{R}^d)^n$ -valued state process  $X = (X^1, \dots, X^n)$ , agent  $i$  chooses a control process  $(\alpha_t^i)$  with values in an action space  $A$  so as to minimize some target functional (see, e.g., the framework considered in [23], (1.1)). For a Markovian control  $\alpha_t^i = \alpha^i(t, X_t)$ , the resulting dynamics can be described by a (controlled) McKean–Vlasov SDE,

$$(15) \quad dX_t^i = \widehat{b}(X_t^i, m_{X_t}^n, \alpha^i(t, X_t)) dt + \sigma dB_t^i + \sigma_0 dW_t.$$

Throughout,  $W$  and  $B^1, \dots, B^n$  are independent Wiener processes, and we write

$$m_x^n = \frac{1}{n} \sum_{k=1}^n \delta_{x_k}$$

to denote the empirical measure of a vector  $\mathbf{x} = (x_1, \dots, x_n)$  in  $(\mathbb{R}^d)^n$ . A similar situation arises in mean-field control, where a central agent chooses the same feedback control  $\alpha(t, X_t)$  for each agent so as to minimise their (the central agent's) objective.

In both these situations, for a nondifferentiable control of (Markovian) type,  $\alpha(t, X_t)$ , appearing in  $\widehat{b}$  and not appearing in  $\sigma$  or  $\sigma_0$  (as Assumption H-4 and H-5 require differentiability), then our approximation scheme will be applicable to the simulation of the controlled (mean-field) SDE and still produce an approximation of strong order 1 as long as one can establish sufficient regularity of  $\alpha$  such that Assumption H-2 and H-3 holds for the modified drift  $b$

$$[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, x, \mu) \mapsto b(t, x, \mu) := \widehat{b}(t, x, \mu, \alpha(t, x)).$$

This exact same argument would work for controlled SDE in classical settings where the control is nondifferentiable, for example, when (adapting from (15))

$$dX_t = \widehat{b}(t, X_t, \alpha(t, X_t)) dt + \sigma(t, X_t) dW_t.$$

In situations when  $\alpha$  is a random field our theory would not apply directly (Kruse et al.'s [7, 35, 36] or [47, 49] as well). It might be possible to address space-measure mean-field controls  $\alpha(t, X_t, \mu_t)$  as in [48] but [48] also shows that proving regularity properties for  $\alpha$  in its measure component is involved.

Last, our method also fits into a situation where machine learning is applied, via tools like reinforcement learning or policy iteration, to solve the optimal control problem. The requirement is a suitable choice of control/policy iteration class that would ensure Assumption H-2 and H-3 holds. A popular choice, especially in the moderate- to high-dimensional context, are deep neural networks, and a commonly used activation function therein is a ReLU, which makes the resulting parametric ansatz function Lipschitz but not everywhere differentiable; see [51] for applications of such a policy gradient method to nonsmooth mean-field control, and to [52] for a proof that the resulting feedback control remains uniformly Lipschitz over the iterations, in a setting with controlled drift but without mean-field interaction.

**4. Moment bound.** In this section, we assume throughout that the conditions of Theorem 3.3 are in force. Here, we establish moment bounds for the scheme (13), but before proving it (Lemma 4.3), we state and prove the following auxiliary result.

LEMMA 4.1. *Let Assumptions H-2 and H-3 be satisfied. For some  $p \geq 2$ , if  $X_{j-1}^{i,N,h} \in \mathcal{L}^p(\Omega)$  for any  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ , then*

$$X_{j,\eta}^{i,N,h} \in \mathcal{L}^p(\Omega) \quad \text{and} \quad X_j^{i,N,h} \in \mathcal{L}^p(\Omega),$$

for all  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ .

PROOF. Notice that for every  $i \in \{1, \dots, N\}$ ,  $X_{j-1}^{i,N,h}$  is  $\mathcal{F}_{j-1}^h$ -measurable and Assumption H-2 gives the continuity of  $b$  and  $\sigma_u$  for  $u \in [0, 1]$ , which in turn implies that  $X_{j,\eta}^{i,N,h}$  defined in (12) is  $\mathcal{F}_j^h$ -measurable for any  $j \in \{1, \dots, n_h\}$ . Also, continuity of  $\partial_x \sigma_u^\ell$  and  $\partial_\mu \sigma_u^\ell$  for  $u \in [0, 1]$  and  $\ell \in \{1, 2, \dots, m_u\}$  implies  $X_j^{i,N,h}$  in (13) is  $\mathcal{F}_j^h$ -measurable for any  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ .

As  $X_{j-1}^{i,N,h} \in \mathcal{L}^p(\Omega)$ , we have from Remark 2.1 and Minkowski's inequality that

$$\begin{aligned} (16) \quad \|b(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} &\leq \bar{L} \{1 + \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} + \|\mathcal{W}_2(\mu_{j-1}^{X,N,h}, \delta_0)\|_{\mathcal{L}^p(\Omega)}\} \\ &\leq \bar{L} \left\{1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)}\right\} < \infty, \end{aligned}$$

where the last inequality is obtained by using

$$\|\mathcal{W}_2(\mu_{j-1}^{X,N,h}, \delta_0)\|_{\mathcal{L}^p(\Omega)} \leq \frac{1}{N} \sum_{i=1}^N \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \leq \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)}.$$

Similarly, we get for  $u \in \{0, 1\}$ ,

$$(17) \quad \|\sigma_u(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \leq \bar{L} \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \right\} < \infty,$$

which along with (12) further implies

$$\begin{aligned} \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} &\leq \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} + h_j \|b(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\quad + h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} \sum_{u=0}^1 \|\sigma_u(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} < \infty, \end{aligned}$$

for any  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ . Thus, on using Remark 2.1, one also obtains

$$\begin{aligned} (18) \quad &\|b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\leq \bar{L} \left\{ 1 + \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} + \|\mathcal{W}_2(\mu_{j,\eta}^{X,N,h}, \delta_0)\|_{\mathcal{L}^p(\Omega)} \right\} \\ &\leq \bar{L} \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \right\} < \infty, \end{aligned}$$

for all  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ . Moreover, recall (8) and use Remarks 2.1 and 3.2 along with (16) and (17) to obtain the following:

$$\begin{aligned} (19) \quad &\sum_{\ell=1}^{m_u} \|\Lambda_{\sigma_u \sigma_1}^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\leq h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} \\ &\quad \times \sum_{\ell=1}^{m_u} \sum_{\ell_1=1}^{m_1} \|\partial_x \sigma_u^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\quad + h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} \sum_{\ell=1}^{m_u} \frac{1}{N} \sum_{k=1}^N \sum_{\ell_1=1}^{m_1} \|\partial_\mu \sigma_u^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j-1}^{k,N,h}) \\ &\quad \times \sigma_1^{\ell_1}(t_{j-1}, X_{j-1}^{k,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\leq 2h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} m_u m_1 \bar{L}^2 \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \right\} < \infty, \end{aligned}$$

for all  $u \in \{0, 1\}$ ,  $s \in [t_{j-1}, t_j]$ ,  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . Similarly,

$$\begin{aligned} (20) \quad &\sum_{\ell=1}^{m_u} \|\Lambda_{\sigma_u \sigma_0}^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\ &\leq 2h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} m_u m_0 \bar{L}^2 \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \right\} < \infty, \end{aligned}$$

for all  $u \in \{0, 1\}$ ,  $s \in [t_{j-1}, t_j]$ ,  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ .



Recalling the expressions of  $\tilde{\sigma}_1$  and  $\tilde{\sigma}_0$  from (10) and then applying Remark 2.1 along with equations (20) and (19), we have

$$\begin{aligned}
 & \sum_{u=0}^1 \sum_{\ell=1}^{m_u} \|\tilde{\sigma}_u^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\
 & \leq \sum_{u=0}^1 \sum_{\ell=1}^{m_u} \|\sigma_u^\ell(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\
 & \quad + \sum_{u=0}^1 \sum_{\ell=1}^{m_1} \|\Lambda_{\sigma_1 \sigma_u}^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\
 & \quad + \sum_{u=0}^1 \sum_{\ell=1}^{m_0} \|\Lambda_{\sigma_0 \sigma_u}^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\
 & \leq \left( \bar{L}(m_0 + m_1) + 2h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} (m_0 + m_1)^2 \bar{L}^2 \right) \\
 & \quad \times \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} \right\} < \infty,
 \end{aligned}
 \tag{21}$$

for all  $i \in \{1, \dots, N\}$ ,  $s \in [t_{j-1}, t_j]$ ,  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . Thus, by (13) and Theorem 7.1 in [43],

$$\begin{aligned}
 \|X_j^{i,N,h}\|_{\mathcal{L}^p(\Omega)} & \leq \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^p(\Omega)} + h_j \|b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h})\|_{\mathcal{L}^p(\Omega)} \\
 & \quad + h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} \\
 & \quad \times \sum_{u=0}^1 \sum_{\ell=1}^{m_u} \left( \int_{t_{j-1}}^{t_j} \|\tilde{\sigma}_u^\ell(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^p(\Omega)}^p ds \right)^{1/p},
 \end{aligned}$$

for all  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ , which on using (18) and (21) completes the proof.  $\square$

As a consequence of the above lemma, we obtain the following corollary.

**COROLLARY 4.2.** *Let Assumptions H-1 with  $\bar{p} \geq 2$ , H-2 and H-3 hold. Then,  $X_j^{i,N,h} \in \mathcal{L}^{\bar{p}}(\Omega)$  for all  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ .*

**PROOF.** First, let us recall the randomised Milstein scheme given in (13). For  $j = 1$ , the result holds due to Assumption H-1 and Lemma 4.1. Assume now that the result is true for  $j = k$  for some  $k \in \{1, \dots, n_h\}$ , that is,  $X_k^{i,N,h} \in \mathcal{L}^{\bar{p}}(\Omega)$  for all  $i \in \{1, \dots, N\}$ . Then, Lemma 4.1 yields  $X_{k+1}^{i,N,h} \in \mathcal{L}^{\bar{p}}(\Omega)$ . An inductive argument complete the proof.  $\square$

The following lemma gives the moment bound of the randomised Milstein scheme (13).

**LEMMA 4.3.** *Let Assumptions H-1 with  $\bar{p} \geq 2$ , H-2 and H-3 hold. Then, for any time grid  $q_h$  (11) with  $h \leq \min(1, T)$ ,*

$$\sup_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{1, \dots, n_h\}} |X_j^{i,N,h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \leq C_5 \left( 1 + \sup_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right),$$

where

$$C_5 := \max \left\{ 2, 2\bar{L}T + 4\bar{L}^2T + 8\bar{L}^2T \left( \frac{p(p-1)}{2} \right)^{1/2} + 2C_0\sqrt{T} \right\} e^{4T(\bar{L} + 2\bar{L}^2 + 4\bar{L}^2(\frac{\bar{p}(\bar{p}-1)}{2})^{1/2} + C_0^2)}$$

with  $C_0 := \bar{C}_{\bar{p}}(\frac{\bar{p}(\bar{p}-1)}{2})^{1/2}(\bar{L}(m_0 + m_1) + 2h_j^{1/2}(\frac{\bar{p}(\bar{p}-1)}{2})^{1/2}(m_0 + m_1)^2\bar{L}^2)$  and  $\bar{C}_{\bar{p}}$  defined in Lemma A.1.

PROOF. Recall (12) and use Minkowski's inequality, Remark 2.1 and  $h \leq 1$  to obtain the following:

$$\begin{aligned} & \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \leq \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} + h_j \|\eta_j b(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + h_j^{1/2} \left( \frac{p(p-1)}{2} \right)^{1/2} \sum_{u=0}^1 \|\sigma_u(t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ (22) \quad & \leq \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + (\bar{L} + 2\bar{L} \left( \frac{p(p-1)}{2} \right)^{1/2}) (1 + \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} + \|\mathcal{W}_2(\mu_{j-1}^{X,N,h}, \delta_0)\|_{\mathcal{L}^{\bar{p}}(\Omega)}) \\ & \leq \bar{L} + 2\bar{L} \left( \frac{p(p-1)}{2} \right)^{1/2} \\ & \quad + \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{p(p-1)}{2} \right)^{1/2} \right) \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)}, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ ,  $N \in \mathbb{N}$  and  $j \in \{1, \dots, n_h\}$ . Due to (13) and Minkowski's inequality, for any  $\bar{k} \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ ,

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} h_j b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}) \right| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^i \right| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^0 \right| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)}, \end{aligned}$$

which on application of Lemma A.1 and Remark 2.1 yields

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} + \sum_{j=1}^{\bar{k}} h_j \|b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h})\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \bar{C}_{\bar{p}} \left\| \left( \sum_{j=1}^{\bar{k}} \left| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^i \right|^2 \right)^{1/2} \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \bar{C}_{\bar{p}} \left\| \left( \sum_{j=1}^{\bar{k}} \left| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^0 \right|^2 \right)^{1/2} \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \end{aligned}$$

$$\begin{aligned}
&\leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\
&\quad + \sum_{j=1}^{\bar{k}} h_j \bar{L} \{1 + \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} + \|\mathcal{W}_2(\mu_{j,\eta}^{X,N,h}, \delta_0)\|_{\mathcal{L}^{\bar{p}}(\Omega)}\} \\
&\quad + \bar{C}_{\bar{p}} \left\| \sum_{j=1}^{\bar{k}} \left| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^i \right|^2 \right\|_{\mathcal{L}^{\bar{p}/2}(\Omega)}^{1/2} \\
&\quad + \bar{C}_{\bar{p}} \left\| \sum_{j=1}^{\bar{k}} \left| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^0 \right|^2 \right\|_{\mathcal{L}^{\bar{p}/2}(\Omega)}^{1/2} \\
&\leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} + \sum_{j=1}^{\bar{k}} h_j \bar{L} \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j,\eta}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right\} \\
&\quad + \bar{C}_{\bar{p}} \left( \sum_{j=1}^{\bar{k}} \left\| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^i \right\|_{\mathcal{L}^{\bar{p}}(\Omega)}^2 \right)^{1/2} \\
&\quad + \bar{C}_{\bar{p}} \left( \sum_{j=1}^{\bar{k}} \left\| \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}) dW_s^0 \right\|_{\mathcal{L}^{\bar{p}}(\Omega)}^2 \right)^{1/2},
\end{aligned}$$

for any  $i \in \{1, \dots, N\}$  and  $\bar{k} \in \{1, \dots, n_h\}$ . Furthermore, the application of (22) and Theorem 7.1 in [43] give

$$\begin{aligned}
&\left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\
&\leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\
&\quad + \sum_{j=1}^{\bar{k}} h_j \bar{L} \left\{ 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right. \\
(23) \quad &\quad \left. + 2 \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right\} \\
&\quad + \bar{C}_{\bar{p}} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \\
&\quad \times \sum_{u=0}^1 \left( \sum_{j=1}^{\bar{k}} h_j^{(\bar{p}-2)/\bar{p}} \left( \int_{t_{j-1}}^{t_j} \|\tilde{\sigma}_u(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^{\bar{p}}(\Omega)}^{\bar{p}} ds \right)^{2/\bar{p}} \right)^{1/2}
\end{aligned}$$

for any  $\bar{k} \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ .

Also, by adapting an argument similar to the one used in (21), we have

$$\begin{aligned}
&\sum_{u=0}^1 \|\tilde{\sigma}_u(s, t_{j-1}, X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h})\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\
&\leq \left( \bar{L}(m_0 + m_1) + 2h_j^{1/2} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} (m_0 + m_1)^2 \bar{L}^2 \right) \\
&\quad \times \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i,N,h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right\},
\end{aligned}$$

for any  $s \in [t_{j-1}, t_j]$ ,  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . On substituting the above in equation (23), one obtains

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \leq \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \bar{L} \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) \sum_{j=1}^{\bar{k}} h_j \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right\} \\ & \quad + C_0 \left( \sum_{j=1}^{\bar{k}} h_j \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right\}^2 \right)^{1/2}, \end{aligned}$$

where  $C_0 := \bar{C}_{\bar{p}} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} (\bar{L}(m_0 + m_1) + 2h_j^{1/2} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} (m_0 + m_1)^2 \bar{L}^2)$ , which on using Young's inequality yields

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \leq \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \bar{L}T \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) + C_0\sqrt{T} \\ & \quad + 2\bar{L} \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + 2C_0 \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)}^{1/2} \left( \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \right)^{1/2} \\ & \leq \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} + \bar{L}T \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) + C_0\sqrt{T} \\ & \quad + 2\bar{L} \left( 1 + 2\bar{L} + 4\bar{L} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + \frac{1}{2} \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} + 2C_0^2 \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)}. \end{aligned}$$

Thus, we have

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \leq 2 \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} \\ & \quad + 2 \left( \bar{L}T + 2\bar{L}^2T + 4\bar{L}^2T \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} \right) + C_0\sqrt{T} \\ & \quad + 4 \left( \bar{L} + 2\bar{L}^2 + 4\bar{L}^2 \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} + C_0^2 \right) \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|X_{j-1}^{i, N, h}\|_{\mathcal{L}^{\bar{p}}(\Omega)} < \infty, \end{aligned}$$

for any  $\bar{k} \in \{1, \dots, n_h\}$ . The application of the discrete Grönwall inequality (see Lemma A.2) yields

$$\begin{aligned} \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, n_h\}} |X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^{\bar{p}}(\Omega)} &\leq \left( 2 \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\Omega)} + 2\bar{L}T + 4\bar{L}^2T \right. \\ &\quad \left. + 8\bar{L}^2T \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{1/2} + 2C_0\sqrt{T} \right) \\ &\quad \times e^{4T(\bar{L}+2\bar{L}^2+4\bar{L}^2(\frac{\bar{p}(\bar{p}-1)}{2})^{1/2}+C_0^2)}. \end{aligned}$$

This completes the proof.  $\square$

**5. Proof of main result.** The proof mechanism we use builds from the notions of “bistability” and “consistency” introduced in [7] (and further explored in [35, 36]). We introduce a notion of bistability and consistency for the numerical scheme (13) of the interacting particle system (6) (and associated with the McKean–Vlasov SDE (4)) by choosing suitable norms and spaces. This choice of Banach spaces and norms is designed to capture the underlying key feature of the systems being analysed, namely, that we deal with interacting particle systems—and to the best of our knowledge, are new.

Throughout this section and in line with the statement of Theorem 3.3, we take  $q = \bar{p}/2$ , where  $\bar{p}$  comes from Assumption H–1 and assumed to satisfy  $\bar{p} \geq 4$ . We next introduce the required notation and definitions to prove our main result Theorem 3.3.

**5.1. Quantities of interest, norms, Banach spaces and residuals.** For the time grid  $\varrho_h$  given in (11), define the Banach spaces  $(\mathcal{G}_q^h, \|\cdot\|_{\mathcal{G}_q^h})$ ,  $(\mathcal{G}_{S,q}^h, \|\cdot\|_{\mathcal{G}_{S,q}^h})$  of stochastic grid processes  $Y^h := \{Y_j^{i, N, h} \in \mathcal{L}^q(\Omega, \mathcal{F}_j^h, \mathbb{P}; \mathbb{R}^d); j \in \{0, 1, \dots, n_h\} \text{ and } i \in \{1, \dots, N\}\}$  as

$$\|Y^h\|_{\mathcal{G}_q^h} := \max_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{0, \dots, n_h\}} |Y_j^{i, N, h}| \right\|_{\mathcal{L}^q(\Omega)} < \infty,$$

and

$$(24) \quad \|Y^h\|_{\mathcal{G}_{S,q}^h} := \max_{i \in \{1, \dots, N\}} \|Y_0^{i, N, h}\|_{\mathcal{L}^q(\Omega)} + \max_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{1, \dots, n_h\}} \left\| \sum_{k=1}^j Y_k^{i, N, h} \right\| \right\|_{\mathcal{L}^q(\Omega)} < \infty,$$

respectively. Also, define

$$\begin{aligned} (25) \quad \Gamma_j^h(Y_{j-1}^{i, N, h}, \mu_{j-1}^{Y, N, h}, Y_{j, \eta}^{i, N, h}, \mu_{j, \eta}^{Y, N, h}, \eta_j) &:= h_j b(t_{j-1} + \eta_j h_j, Y_{j, \eta}^{i, N, h}, \mu_{j, \eta}^{Y, N, h}) \\ &\quad + \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i, N, h}, \mu_{j-1}^{Y, N, h}) dW_s^{i, \ell} \\ &\quad + \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_0^\ell(s, t_{j-1}, Y_{j-1}^{i, N, h}, \mu_{j-1}^{Y, N, h}) dW_s^{0, \ell}, \end{aligned}$$

almost surely for any  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ , where  $Y_{j, \eta}^{i, N, h}$  is defined using (12), and the empirical measures  $\mu_{j-1}^{Y, N, h}$  and  $\mu_{j, \eta}^{Y, N, h}$  are defined using (14).

Define  $\mathcal{R}[Y^h] \in \mathcal{G}_{S,q}^h$  for  $q \geq 2$  as the collection of the pointwise residuals  $\mathcal{R}_j^{i, N}[Y^h]$  (associated to executing the scheme with  $Y^h$ ) by

$$\begin{aligned} (26) \quad \mathcal{R}_0^{i, N}[Y^h] &= Y_0^{i, N, h} - X_0^{i, N, h}, \\ \mathcal{R}_j^{i, N}[Y^h] &= Y_j^{i, N, h} - Y_{j-1}^{i, N, h} - \Gamma_j^h(Y_{j-1}^{i, N, h}, \mu_{j-1}^{Y, N, h}, Y_{j, \eta}^{i, N, h}, \mu_{j, \eta}^{Y, N, h}, \eta_j), \\ &\quad j \in \{1, \dots, n_h\} \end{aligned}$$

almost surely for any  $i \in \{1, \dots, N\}$ .

*Randomised quadrature rule for stochastic processes.* In this small section we discuss the randomised quadrature rule for stochastic processes developed in [36]. For this, let us recall the sequence of i.i.d. uniform random variables  $\eta := \{\eta_j\}_{j \in \mathbb{N}}$  and the temporal grid  $\varrho_h$  from Section 3.1. Consider a stochastic process  $V : [0, T] \times \tilde{\Omega} \mapsto \mathbb{R}^d$  on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$  satisfying  $V \in \mathcal{L}^p([0, T] \times \tilde{\Omega})$  for  $p \geq 2$ . For each  $\bar{n} \in \{1, \dots, n_h\}$ ,  $\int_0^{t_{\bar{n}}} V(s) ds$  is approximated by the randomised Riemann sum

$$(27) \quad \Theta_{\bar{n}, \eta}^h[V] = \sum_{j=1}^{\bar{n}} h_j V(t_{j-1} + \eta_j h_j),$$

which is a random variable on  $(\Omega, \mathcal{F}, \mathbb{P})$  and an unbiased estimator of  $\int_0^{t_{\bar{n}}} V(s) ds$ , that is,  $\mathbb{E}^\eta \Theta_{\bar{n}, \eta}^h[V] = \int_0^{t_{\bar{n}}} V(s) ds \in \mathcal{L}^p(\tilde{\Omega})$ . Moreover, due to Theorem 4.1 in [36], we have

$$\left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \Theta_{\bar{n}, \eta}^h[V] - \int_0^{t_{\bar{n}}} V(s) ds \right| \right\|_{\mathcal{L}^p(\Omega)} \leq 2\bar{C}_p T^{(p-2)/(2p)} \|V\|_{\mathcal{L}^p([0, T] \times \tilde{\Omega})} h^{1/2}.$$

Additionally, if  $V \in \mathcal{C}^\alpha([0, T], \mathcal{L}^p(\tilde{\Omega}))$  for some  $\alpha \in (0, 1]$ , then

$$(28) \quad \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \Theta_{\bar{n}, \eta}^h[V] - \int_0^{t_{\bar{n}}} V(s) ds \right| \right\|_{\mathcal{L}^p(\Omega)} \leq \bar{C}_p \sqrt{T} \|V\|_{\mathcal{C}^\alpha([0, T], \mathcal{L}^p(\tilde{\Omega}))} h^{\alpha+1/2},$$

where  $\bar{C}_p$  is defined in Lemma A.1.

**5.2. Bistability of the scheme.** We now specify the notion of a scheme's bistability and show that the proposed randomised Milstein scheme (13) of the interacting particle system (6) is bistable (see Proposition 5.5 below). For this, recall the scheme from (13) and define  $X^h := \{X_j^{i, N, h} : j \in \{0, \dots, n_h\}, i \in \{1, \dots, N\}\}$ . Clearly,  $X^h \in \mathcal{G}_q^h$  due to Lemma 4.3. Also, recall the definition of residuals  $\mathcal{R}[Y^h]$  from (26) for some set  $Y^h \in \mathcal{G}_q^h$ .

**DEFINITION 5.1 (Stochastic bistability).** The randomised Milstein scheme (13) associated to the interacting particle system (6) is called stochastically bistable if there exist constants  $C_6, C_7 > 0$ , independent of  $h$  and  $N \in \mathbb{N}$ , such that for any arbitrary time grid  $\varrho_h$  as given in (11) and for any  $Y^h \in \mathcal{G}_q^h$ , the following holds:

$$C_6 \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S, q}^h} \leq \|Y^h - X^h\|_{\mathcal{G}_q^h} \leq C_7 \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S, q}^h},$$

for  $q \geq 2$ .

Let us first establish some useful lemmas.

**LEMMA 5.2.** *Let Assumptions H-2 and H-3 hold. Then, for any  $q \geq 2$  and  $Y^h, Z^h \in \mathcal{G}_q^h$ ,*

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} h_j [b(t_{j-1} + \eta_j h_j, Y_{j, \eta}^{i, N, h}, \mu_{j, \eta}^{Y, N, h}) - b(t_{j-1} + \eta_j h_j, Z_{j, \eta}^{i, N, h}, \mu_{j, \eta}^{Z, N, h})] \right| \right\|_{\mathcal{L}^q(\Omega)} \\ & \leq C_8 \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i, N, h} - Z_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^q(\Omega)}, \end{aligned}$$

for any  $\bar{k} \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$  where  $C_8 := 2L + 4L^2T + 8L^2\sqrt{Tq(q-1)/2}$ .



PROOF. Notice that all the terms in the statement of this lemma are well-defined due to Remark 2.1. By using (12) with  $Y^h, Z^h \in \mathcal{G}_q^h$  along with Minkowski's inequality and Assumption H-2, we have

$$\begin{aligned}
 & \|Y_{j,\eta}^{i,N,h} - Z_{j,\eta}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} \\
 & \leq \|Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} \\
 & \quad + h_j \| \eta_j (b(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - b(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})) \|_{\mathcal{L}^q(\Omega)} \\
 & \quad + \sqrt{h_j} \left( \frac{q(q-1)}{2} \right)^{1/2} \sum_{u=0}^1 \|\sqrt{\eta_j} (\sigma_u(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \\
 (29) \quad & - \sigma_u(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}))\|_{\mathcal{L}^q(\Omega)} \\
 & \leq \|Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} + \left( LT + 2L\sqrt{T} \left( \frac{q(q-1)}{2} \right)^{1/2} \right) \\
 & \quad \times \{ \|Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{W}_2(\mu_{j-1}^{Y,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \} \\
 & \leq \left( 1 + 2LT + 4L\sqrt{T} \left( \frac{q(q-1)}{2} \right)^{1/2} \right) \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - Z_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)},
 \end{aligned}$$

where the last inequality is obtained by using

$$\mathcal{W}_2(\mu_{j-1}^{Y,N,h}, \mu_{j-1}^{Z,N,h}) \leq \left( \frac{1}{N} \sum_{i=1}^N |Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}|^2 \right)^{1/2},$$

for all  $i \in \{1, \dots, N\}$ ,  $N \in \mathbb{N}$  and  $j \in \{1, \dots, n_h\}$ .

Now, the application Assumption H-2 yields

$$\begin{aligned}
 & \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} h_j [b(t_{j-1} + \eta_j h_j, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}) - b(t_{j-1} + \eta_j h_j, Z_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Z,N,h})] \right| \right\|_{\mathcal{L}^q(\Omega)} \\
 & \leq L \sum_{j=1}^{\bar{k}} h_j \{ \|Y_{j,\eta}^{i,N,h} - Z_{j,\eta}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{W}_2(\mu_{j,\eta}^{Y,N,h}, \mu_{j,\eta}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \} \\
 & \leq 2L \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \|Y_{j,\eta}^{i,N,h} - Z_{j,\eta}^{i,N,h}\|_{\mathcal{L}^q(\Omega)},
 \end{aligned}$$

and then one uses (29) to complete the proof.  $\square$

LEMMA 5.3. Let Assumptions H-2, H-3, H-5 hold. Then, for any  $q \geq 2$  and  $Y^h, Z^h \in \mathcal{G}_q^h$ ,

$$\begin{aligned}
 & \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
 & \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \right| \right\|_{\mathcal{L}^q(\Omega)} \\
 & \quad + \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_0^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
 & \quad \left. \left. - \tilde{\sigma}_0^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \right| \right\|_{\mathcal{L}^q(\Omega)}
 \end{aligned}$$

$$\begin{aligned}
& -\tilde{\sigma}_0^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{0,\ell} \Big\|_{\mathcal{L}^q(\Omega)} \\
& \leq C_9 \left( \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - Z_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)}^2 \right)^{1/2},
\end{aligned}$$

for all  $i \in \{1, \dots, N\}$  and  $\bar{k} \in \{1, \dots, n_h\}$ , where  $C_9 := (m_0 + m_1)\bar{C}_q\sqrt{q(q-1)/2}(2L + 5(m_0 + m_1)L\sqrt{Tq(q-1)/2})$  and  $\bar{C}_q$  is given in Lemma A.1.

PROOF. Notice that

$$\left\{ \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \right\}_{\bar{n}},$$

with  $\bar{n} \in \{0, \dots, n_h\}$  is an  $\{\mathcal{F}_{\bar{n}}^h\}_{\bar{n} \in \{0, \dots, n_h\}}$ -adapted standard martingale. Thus, on using Lemma A.1 and Minkowski's inequality, one obtains, for all  $i \in \{1, \dots, N\}$  and  $\bar{k} \in \{1, \dots, n_h\}$ ,

$$\begin{aligned}
& \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
& \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \right\|_{\mathcal{L}^q(\Omega)} \right\| \\
& \leq \bar{C}_q \left\| \sum_{j=1}^{\bar{k}} \left| \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
& \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \right\|_{\mathcal{L}^{q/2}(\Omega)}^2 \right\|^{1/2} \\
(30) \quad & \leq \bar{C}_q \left( \sum_{j=1}^{\bar{k}} \left\| \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
& \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}) dW_s^{i,\ell} \right\|_{\mathcal{L}^q(\Omega)}^2 \right)^{1/2} \\
& \leq \sqrt{m_1} \bar{C}_q \left( \frac{q(q-1)}{2} \right)^{1/2} \left( \sum_{j=1}^{\bar{k}} h_j^{(q-2)/q} \sum_{\ell=1}^{m_1} \left[ \int_{t_{j-1}}^{t_j} \|\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \right. \right. \\
& \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)}^q ds \right]^{2/q} \right)^{1/2},
\end{aligned}$$

where the last inequality is obtained due to Theorem 7.1 in [43].

Now, recall (8), (9) and (10) and use Assumption H-2 to get the following:

$$\begin{aligned}
& \|\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \\
& \leq \|\sigma_1^\ell(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \sigma_1^\ell(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)}
\end{aligned}$$

$$\begin{aligned}
& + \sum_{u=0}^1 \|\Lambda_{\sigma_1 \sigma_u}^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \Lambda_{\sigma_1 \sigma_u}^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \\
& \leq L \{ \|Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{W}_2(\mu_{j-1}^{Y,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \} \\
& \quad + h_j^{1/2} \left( \frac{q(q-1)}{2} \right)^{1/2} \sum_{u=0}^1 \sum_{\ell_1=1}^{m_u} \|\partial_x \sigma_1^\ell(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \sigma_u^{\ell_1}(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) \\
& \quad - \partial_x \sigma_1^\ell(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}) \sigma_u^{\ell_1}(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \\
& \quad + h_j^{1/2} \left( \frac{q(q-1)}{2} \right)^{1/2} \frac{1}{N} \sum_{u=0}^1 \sum_{\ell_1=1}^{m_u} \sum_{k=1}^N \|\partial_\mu \sigma_1^\ell(t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j-1}^{k,N,h}) \\
& \quad \times \sigma_u^{\ell_1}(t_{j-1}, Y_{j-1}^{k,N,h}, \mu_{j-1}^{Y,N,h}) - \partial_\mu \sigma_1^\ell(t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}, Z_{j-1}^{k,N,h}) \\
& \quad \times \sigma_u^{\ell_1}(t_{j-1}, Z_{j-1}^{k,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)},
\end{aligned}$$

which due to Assumption H-5 yields

$$\begin{aligned}
& \|\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \\
& \leq \left( L + 2(m_0 + m_1)L \left( \frac{Tq(q-1)}{2} \right)^{1/2} \right) \\
& \quad \times \{ \|Y_{j-1}^{i,N,h} - Z_{j-1}^{i,N,h}\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{W}_2(\mu_{j-1}^{Y,N,h}, \mu_{j-1}^{Z,N,h})\|_{\mathcal{L}^q(\Omega)} \} \\
& \quad + (m_0 + m_1)L \left( \frac{Tq(q-1)}{2} \right)^{1/2} \frac{1}{N} \sum_{k=1}^N \|Y_{j-1}^{k,N,h} - Z_{j-1}^{k,N,h}\|_{\mathcal{L}^q(\Omega)} \\
& \leq \left( 2L + 5(m_0 + m_1)L \left( \frac{Tq(q-1)}{2} \right)^{1/2} \right) \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - Z_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)},
\end{aligned}$$

for any  $s \in [t_{j-1}, t_j]$ ,  $j \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ . The proof of the first part of the lemma is completed by substituting the above in (30). A bound for the terms involving  $\tilde{\sigma}_0$  follows by similar arguments.  $\square$

As a consequence of Lemmas 5.2 and 5.3, we obtain the following corollary.

**COROLLARY 5.4.** *If Assumptions H-2, H-3 and H-5 are satisfied. Then, for any  $q \geq 2$  and  $Y^h, Z^h \in \mathcal{G}_q^h$ ,*

$$\begin{aligned}
& \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} [\Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) \right. \right. \\
& \quad \left. \left. - \Gamma_j^h(Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}, Z_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Z,N,h}, \eta_j)] \right| \right\|_{\mathcal{L}^q(\Omega)} \\
& \leq C_8 \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - Z_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\
& \quad + C_9 \left( \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - Z_{\bar{n}}^{i,N,h}|^2 \right\|_{\mathcal{L}^q(\Omega)} \right)^{1/2} \\
& \leq (C_8 T + C_9 \sqrt{T}) \|Y^h - Z^h\|_{\mathcal{G}_q^h},
\end{aligned}$$

for all  $\bar{k} \in \{1, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ , where the constants  $C_8$  and  $C_9$  appear in Lemmas 5.2 and 5.3, respectively.

PROOF. Recall  $\Gamma_j^h$  from (25) and write

$$\begin{aligned} & \Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) - \Gamma_j^h(Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h}, Z_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Z,N,h}, \eta_j) \\ &= h_j [b(t_{j-1} + \eta_j h_j, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}) - b(t_{j-1} + \eta_j h_j, Z_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Z,N,h})] \\ &+ \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_1^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{i,\ell} \\ &+ \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} [\tilde{\sigma}_0^\ell(s, t_{j-1}, Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}) - \tilde{\sigma}_0^\ell(s, t_{j-1}, Z_{j-1}^{i,N,h}, \mu_{j-1}^{Z,N,h})] dW_s^{0,\ell}. \end{aligned}$$

Then, on using Lemmas 5.2 and 5.3, we get the required result.  $\square$

We now prove the bistability of the scheme (13) in the following proposition.

PROPOSITION 5.5. Let Assumptions H-1 with  $\bar{p} \geq 4$  and set  $q = \bar{p}/2 \geq 2$ , H-2, H-3 and H-5 hold. Then, for any  $Y^h, X^h \in \mathcal{G}_q^h$ , we have

$$C_6 \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h} \leq \|Y^h - X^h\|_{\mathcal{G}_q^h} \leq C_7 \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h},$$

in other words, the randomised Milstein scheme given in (13) is stochastically bistable in the sense of Definition 5.1 with  $C_6 := \frac{1}{3+C_8T+C_9\sqrt{T}}$  and  $C_7 := 2e^{(2C_8+C_9^2)T}$ , where the constants  $C_8$  and  $C_9$  appear in Lemmas 5.2 and 5.3, respectively.

PROOF. Recall Equations (13), (25) and (26) to write

$$\begin{aligned} & X_{\bar{n}}^{i,N,h} - X_0^{i,N,h} - \sum_{j=1}^{\bar{n}} \Gamma_j^h(X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}, \eta_j) = 0, \\ (31) \quad & \sum_{j=1}^{\bar{n}} \mathcal{R}_j^{i,N}[Y^h] = Y_{\bar{n}}^{i,N,h} - Y_0^{i,N,h} - \sum_{j=1}^{\bar{n}} \Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) \\ &= (Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}) - (Y_0^{i,N,h} - X_0^{i,N,h}) \\ &\quad - \sum_{j=1}^{\bar{n}} (\Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) \\ &\quad - \Gamma_j^h(X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}, \eta_j)), \end{aligned}$$

for any  $\bar{n} \in \{1, \dots, n_h\}$ ,  $i \in \{1, \dots, N\}$  and then use the Spijker norm (24) to get the following:

$$\begin{aligned} \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h} &= \max_{i \in \{1, \dots, N\}} \|\mathcal{R}_0^{i,N}[Y^h]\|_{\mathcal{L}^q(\Omega)} + \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} \mathcal{R}_j^{i,N}[Y^h] \right| \right\|_{\mathcal{L}^q(\Omega)} \\ &\leq 2 \max_{i \in \{1, \dots, N\}} \|Y_0^{i,N,h} - X_0^{i,N,h}\|_{\mathcal{L}^q(\Omega)} \\ &\quad + \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \end{aligned}$$

$$+ \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} (\Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) - \Gamma_j^h(X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}, \eta_j)) \right| \right\|_{\mathcal{L}^q(\Omega)},$$

which on using Corollary 5.4 yields

$$(32) \quad \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h} \leq (3 + C_8 T + C_9 \sqrt{T}) \|Y^h - X^h\|_{\mathcal{G}_q^h}.$$

Further, by rearranging the terms of (31) and using Minkowski's inequality, one obtains

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \leq \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} (\Gamma_j^h(Y_{j-1}^{i,N,h}, \mu_{j-1}^{Y,N,h}, Y_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{Y,N,h}, \eta_j) - \Gamma_j^h(X_{j-1}^{i,N,h}, \mu_{j-1}^{X,N,h}, X_{j,\eta}^{i,N,h}, \mu_{j,\eta}^{X,N,h}, \eta_j)) \right| \right\|_{\mathcal{L}^q(\Omega)} \\ & \quad + \max_{i \in \{1, \dots, N\}} \|\mathcal{R}_0^{i,N}[Y^h]\|_{\mathcal{L}^q(\Omega)} + \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{1, \dots, \bar{k}\}} \left| \sum_{j=1}^{\bar{n}} \mathcal{R}_j^{i,N}[Y^h] \right| \right\|_{\mathcal{L}^q(\Omega)}, \end{aligned}$$

and then application of Corollary 5.4 gives the following:

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \leq C_8 \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \quad + C_9 \left( \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)}^2 \right)^{1/2} + \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h} \\ & \leq C_8 \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \quad + \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)}^{1/2} \\ & \quad \times C_9 \left( \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \right)^{1/2} + \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h}, \end{aligned}$$

which due to Young's inequality yields

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \leq \frac{1}{2} \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \\ & \quad + \left( C_8 + \frac{C_9^2}{2} \right) \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i,N,h} - X_{\bar{n}}^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S,q}^h}. \end{aligned}$$

This further implies

$$\begin{aligned} & \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i, N, h} - X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^q(\Omega)} \leq 2 \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S, q}^h} \\ & + (2C_8 + C_9^2) \sum_{j=1}^{\bar{k}} h_j \max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, j-1\}} |Y_{\bar{n}}^{i, N, h} - X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^q(\Omega)}. \end{aligned}$$

Due to Lemma A.2, we get

$$\max_{i \in \{1, \dots, N\}} \left\| \max_{\bar{n} \in \{0, \dots, \bar{k}\}} |Y_{\bar{n}}^{i, N, h} - X_{\bar{n}}^{i, N, h}| \right\|_{\mathcal{L}^q(\Omega)} \leq 2e^{(2C_8 + C_9^2)T} \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S, q}^h},$$

for any  $\bar{k} \in \{0, \dots, n_h\}$ , which further implies

$$\|Y^h - X^h\|_{\mathcal{G}_q^h} \leq 2e^{(2C_8 + C_9^2)T} \|\mathcal{R}[Y^h]\|_{\mathcal{G}_{S, q}^h},$$

and the proof is completed by combining the above with (32).  $\square$

**5.3. Consistency of the scheme.** Recall the temporal grid  $\varrho_h$  from (11) and set the values of the interacting particle system (6) over the grid points of  $\varrho_h$  as the set  $X^{\varrho_h} := \{X_j^{i, N}\}_{j \in \{0, \dots, n_h\}}$  where  $X_j^{i, N} = X_{t_j}^{i, N}$  for any  $j \in \{0, \dots, n_h\}$  and  $i \in \{1, \dots, N\}$ . Notice that  $X^{\varrho_h}$  is different from  $X^h$ , where the later stands for the randomised Milstein scheme (13). Further,  $\mathcal{R}[X^{\varrho_h}]$  is defined using  $X^{\varrho_h}$  in (26). Next, we introduce the notion of a scheme's consistency and establish that the randomised Milstein scheme (13) is consistent (see below Proposition 5.9).

**DEFINITION 5.6 (Consistency).** The randomised Milstein scheme (13) for the interacting particle system (6) is called consistent of order  $\gamma > 0$  if there exists a constant  $C_{10} > 0$ , independent of  $h$  and  $N \in \mathbb{N}$ , such that

$$\|\mathcal{R}[X^{\varrho_h}]\|_{\mathcal{G}_{S, q}^h} \leq C_{10} h^\gamma,$$

for  $q \geq 2$ .

In the context of the measure dependent drift coefficient, we obtain the following randomised quadrature rule.

**COROLLARY 5.7.** Let Assumptions H-1 with  $\bar{p} \geq 4$  and set  $q = \bar{p}/2 \geq 2$ , H-2, H-3 hold. Then, for all  $i \in \{1, \dots, N\}$ , the operator  $\Theta_{\bar{n}, \eta}^h[\cdot]$  defined in (27) and applied to  $b$  satisfies

$$\left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \Theta_{\bar{n}, \eta}^h[b] - \int_0^{t_{\bar{n}}} b(s, X_s^{i, N}, \mu_s^{X, N}) ds \right| \right\|_{\mathcal{L}^q(\Omega)} \leq C_{11} h,$$

where  $C_{11} := \sqrt{T} \bar{C}_q (2\bar{L} + 4\bar{L} C_1^{1/\bar{p}} + 2\bar{L} C_2^{1/\bar{p}}) (1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}})^{1/\bar{p}}$  and the positive constants  $C_1, C_2, \bar{C}_q$  appear in Proposition 2.2, Lemma 2.4 and Lemma A.1, respectively.

**PROOF.** We take  $V(s) = b(s, X_s^{i, N}, \mu_s^{X, N})$  for any  $s \in [0, T]$  and  $i \in \{1, \dots, N\}$ , which is defined on  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ . Notice that Remarks 2.1 and 2.3 along with Hölder's inequality (as



$q = \bar{p}/2$  imply

$$\begin{aligned}
 \|V\|_{\mathcal{L}^q([0,T] \times \tilde{\Omega})} &= \left( \int_0^T \|b(s, X_s^{i,N}, \mu_s^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})}^q ds \right)^{1/q} \\
 &\leq \bar{L} \left( \int_0^T (1 + \|X_s^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} + \|\mathcal{W}_2(\mu_s^{X,N}, \delta_0)\|_{\mathcal{L}^q(\tilde{\Omega})})^q ds \right)^{1/q} \\
 &\leq \bar{L} \left( \int_0^T \left(1 + 2 \max_{i \in \{1, \dots, N\}} \|X_s^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})}\right)^q ds \right)^{1/q} \\
 &\leq \bar{L} \left( \int_0^T \left(1 + 2 \max_{i \in \{1, \dots, N\}} \|X_s^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}\right)^q ds \right)^{1/q} \\
 &\leq \bar{L} \left( \int_0^T \left(1 + 2C_1^{1/\bar{p}} \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}\right)^{1/\bar{p}}\right)^q ds \right)^{1/q} \\
 &\leq \bar{L}(T)^{1/q} \left(1 + 2C_1^{1/\bar{p}} \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}\right)^{1/\bar{p}}\right) < \infty.
 \end{aligned}$$

Also, due to Assumptions H-2, H-3 and Lemma 2.4, for  $t, t' \in [0, T]$ ,

$$\begin{aligned}
 \|V(t) - V(t')\|_{\mathcal{L}^q(\tilde{\Omega})} &= \|b(t, X_t^{i,N}, \mu_t^{X,N}) - b(t', X_{t'}^{i,N}, \mu_{t'}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\leq L \{ \|X_t^{i,N} - X_{t'}^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\quad + \|\mathcal{W}_2(\mu_t^{X,N}, \mu_{t'}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\quad + |t - t'|^{1/2} (1 + \|X_t^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} + \|\mathcal{W}_2(\mu_t^{X,N}, \delta_0)\|_{\mathcal{L}^q(\tilde{\Omega})}) \} \\
 (33) \quad &\leq 2L \max_{i \in \{1, \dots, N\}} \|X_t^{i,N} - X_{t'}^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\quad + |t - t'|^{1/2} L \left\{ 1 + 2C_1^{1/\bar{p}} \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}\right)^{1/\bar{p}} \right\} \\
 &\leq \left\{ 2L(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}\right)^{1/\bar{p}} + L \right\} |t - t'|^{1/2},
 \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ . Due to (3),  $\|V\|_{\mathcal{C}^{1/2}([0,T], \mathcal{L}^q(\tilde{\Omega}))} \leq 2\bar{L} + \bar{L}(4C_1^{1/\bar{p}} + 2C_2^{1/\bar{p}})(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})})^{1/\bar{p}}$ , as  $L \leq \bar{L}$ . Then, (28) completes the proof.  $\square$

LEMMA 5.8. Let Assumptions H-1 with  $\bar{p} \geq 4$  and set  $q = \bar{p}/2 \geq 2$ , H-2 to H-4 hold. Let  $h \leq \min(1, T)$ . Then,

$$\begin{aligned}
 &\left\| \max_{\tilde{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\tilde{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{i,\ell} \right| \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\quad + \left\| \max_{\tilde{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\tilde{n}} \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} [\sigma_0^\ell(s, X_s^{i,N}, \mu_s^{X,N}) \right. \right. \\
 &\quad \left. \left. - \tilde{\sigma}_0^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{0,\ell} \right| \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\
 &\leq C_{12}h,
 \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ , where

$$C_{12} := (m_0 + m_1) \bar{C}_q \sqrt{Tq(q-1)/2} \left\{ L + 2\bar{L}^2 + 4L^2 \sqrt{q(q-1)/2} + (2(L + 2\bar{L}^2)C_1^{1/\bar{p}} + 8L^2 \sqrt{q(q-1)2(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) + 4LC_2^{2/\bar{p}}}) \left(1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{2/\bar{p}}\right) \right\}$$

and the positive constants  $C_1$ ,  $C_2$  and  $\bar{C}_q$  appear in Proposition 2.2, Lemma 2.4 and Lemma A.1, respectively.

PROOF. For notational simplicity, define

$$\zeta^\ell(s) := \sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})$$

for all  $\ell \in \{1, \dots, m_1\}$ ,  $i \in \{1, \dots, N\}$ ,  $j \in \{1, \dots, n_h\}$  and  $s \in [t_{j-1}, t_j]$ . Notice that

$$\left\{ \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \zeta^\ell(s) dW_s^{i,\ell} \right\}_{\bar{n} \in \{0, \dots, n_h\}}$$

is an  $\{\tilde{\mathcal{F}}_{t_{\bar{n}}}\}_{\bar{n} \in \{0, \dots, n_h\}}$ -adapted standard martingale. Thus, by using Lemma A.1 and Theorem 7.1 in [43], we get

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{0, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \zeta^\ell(s) dW_s^{i,\ell} \right| \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\ & \leq \bar{C}_q \left\| \left( \sum_{j=1}^{n_h} \left| \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \zeta^\ell(s) dW_s^{i,\ell} \right|^2 \right)^{1/2} \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\ (34) \quad & = \bar{C}_q \left\| \sum_{j=1}^{n_h} \left| \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} \zeta^\ell(s) dW_s^{i,\ell} \right|^2 \right\|_{\mathcal{L}^{q/2}(\tilde{\Omega})}^{1/2} \\ & \leq \sqrt{m_1} \bar{C}_q \left( \sum_{j=1}^{n_h} \sum_{\ell=1}^{m_1} \left\| \int_{t_{j-1}}^{t_j} \zeta^\ell(s) dW_s^{i,\ell} \right\|_{\mathcal{L}^q(\tilde{\Omega})}^2 \right)^{1/2} \\ & \leq \sqrt{m_1} \bar{C}_q \left( \frac{q(q-1)}{2} \right)^{1/2} \left( \sum_{j=1}^{n_h} \sum_{\ell=1}^{m_1} h_j^{(q-2)/q} \left( \int_{t_{j-1}}^{t_j} \|\zeta^\ell(s)\|_{\mathcal{L}^q(\tilde{\Omega})}^q ds \right)^{2/q} \right)^{1/2}, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ . Now, recall  $\tilde{\sigma}_1^\ell$  from (10) to write the following for any  $s \in [t_{j-1}, t_j]$ :

$$\begin{aligned} \zeta^\ell(s) &= [\sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_1^\ell(t_{j-1}, X_s^{i,N}, \mu_s^{X,N})] + \left[ \sigma_1^\ell(t_{j-1}, X_s^{i,N}, \mu_s^{X,N}) \right. \\ & \quad - \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) - \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})(X_s^{i,N} - X_{j-1}^{i,N}) \\ & \quad \left. - \frac{1}{N} \sum_{k=1}^N \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N})(X_s^{k,N} - X_{j-1}^{k,N}) \right] \\ & \quad + \left[ \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) \right. \\ & \quad \left. \times (X_s^{i,N} - X_{j-1}^{i,N}) + \frac{1}{N} \sum_{k=1}^N \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N})(X_s^{k,N} - X_{j-1}^{k,N}) \right] \end{aligned}$$

$$\left. -\Lambda_{\sigma_1\sigma_1}^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) - \Lambda_{\sigma_1\sigma_0}^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) \right],$$

which on the application of Assumption H-3, Lemma A.3 along with (6) and the values of  $\Lambda_{\sigma_1\sigma_1}^\ell, \Lambda_{\sigma_1\sigma_0}^\ell$  from (8), (9) yields

$$\begin{aligned} \zeta^\ell(s) &\leq L(1 + |X_s^{i,N}| + \mathcal{W}_2(\mu_s^{X,N}, \delta_0))|s - t_{j-1}| + \frac{3L}{2}|X_s^{i,N} - X_{j-1}^{i,N}|^2 \\ &\quad + \frac{5L}{2} \frac{1}{N} \sum_{k=1}^N |X_s^{k,N} - X_{j-1}^{k,N}|^2 + \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) b(r, X_r^{i,N}, \mu_r^{X,N}) dr \\ &\quad + \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) (\sigma_1(r, X_r^{i,N}, \mu_r^{X,N}) - \sigma_1(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})) dW_r^i \\ &\quad + \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) (\sigma_0(r, X_r^{i,N}, \mu_r^{X,N}) - \sigma_0(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})) dW_r^0 \\ &\quad + \frac{1}{N} \sum_{k=1}^N \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N}) b(r, X_r^{k,N}, \mu_r^{X,N}) dr \\ &\quad + \frac{1}{N} \sum_{k=1}^N \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N}) \\ &\quad \times (\sigma_1(r, X_r^{k,N}, \mu_r^{X,N}) - \sigma_1(t_{j-1}, X_{j-1}^{k,N}, \mu_{j-1}^{X,N})) dW_r^k \\ &\quad + \frac{1}{N} \sum_{k=1}^N \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N}) \\ &\quad \times (\sigma_0(r, X_r^{k,N}, \mu_r^{X,N}) - \sigma_0(t_{j-1}, X_{j-1}^{k,N}, \mu_{j-1}^{X,N})) dW_r^0, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ ,  $j \in \{1, \dots, n_h\}$  and  $\ell \in \{1, \dots, m_1\}$ . This further implies, due to Remarks 2.1 and 3.2,

$$\begin{aligned} \|\zeta^\ell(s)\|_{\mathcal{L}^q(\tilde{\Omega})} &\leq L \left( 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_s^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \right) h_j + 4L \max_{i \in \{1, \dots, N\}} \|X_s^{i,N} - X_{j-1}^{i,N}\|_{\mathcal{L}^{2q}(\tilde{\Omega})}^2 \\ &\quad + h_j^{(q-1)/q} 2\bar{L}^2 \left( \int_{t_{j-1}}^s \left( 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_r^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \right)^q dr \right)^{1/q} \\ &\quad + \left\| \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) (\sigma_1(r, X_r^{i,N}, \mu_r^{X,N}) \right. \\ &\quad \left. - \sigma_1(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})) dW_r^i \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\ &\quad + \left\| \int_{t_{j-1}}^s \partial_x \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}) (\sigma_0(r, X_r^{i,N}, \mu_r^{X,N}) \right. \\ &\quad \left. - \sigma_0(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})) dW_r^0 \right\|_{\mathcal{L}^q(\tilde{\Omega})} \\ &\quad + \left\| \frac{1}{N} \sum_{k=1}^N \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N}) \right. \end{aligned}$$

$$\begin{aligned}
& \times (\sigma_1(r, X_r^{k,N}, \mu_r^{X,N}) - \sigma_1(t_{j-1}, X_{j-1}^{k,N}, \mu_{j-1}^{X,N})) dW_r^k \Big\|_{\mathcal{L}^q(\tilde{\Omega})} \\
& + \left\| \frac{1}{N} \sum_{k=1}^N \int_{t_{j-1}}^s \partial_\mu \sigma_1^\ell(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j-1}^{k,N}) \right. \\
& \times (\sigma_0(r, X_r^{k,N}, \mu_r^{X,N}) - \sigma_0(t_{j-1}, X_{j-1}^{k,N}, \mu_{j-1}^{X,N})) dW_r^0 \Big\|_{\mathcal{L}^q(\tilde{\Omega})}.
\end{aligned}$$

Moreover, the application of Hölder's inequality (as  $q = \bar{p}/2$ ), Remarks 2.3 and 3.2, Lemma 2.4 and Theorem 7.1 in [43] yield

$$\begin{aligned}
\|\zeta^\ell(s)\|_{\mathcal{L}^q(\tilde{\Omega})} & \leq \left\{ (L + 2\bar{L}^2) \left( 1 + 2C_1^{1/\bar{p}} \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{1/\bar{p}} \right) \right. \\
& + 4LC_2^{2/\bar{p}} \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{2/\bar{p}} \Big\} h_j \\
& + \left( \frac{q(q-1)}{2} \right)^{1/2} L h_j^{(q-2)/2q} \\
(35) \quad & \times \sum_{u=0}^1 \left( \int_{t_{j-1}}^s \|\sigma_u(r, X_r^{i,N}, \mu_r^{X,N}) - \sigma_u(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})}^q dr \right)^{1/q} \\
& + \left( \frac{q(q-1)}{2} \right)^{1/2} L h_j^{(q-2)/2q} \sum_{u=0}^1 \frac{1}{N} \sum_{k=1}^N \left( \int_{t_{j-1}}^s \|\sigma_u(r, X_r^{k,N}, \mu_r^{X,N}) \right. \\
& \left. \left. - \sigma_u(t_{j-1}, X_{j-1}^{k,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})}^q dr \right)^{1/q},
\end{aligned}$$

for any  $s \in [t_{j-1}, t_j]$ ,  $i \in \{1, \dots, N\}$ ,  $j \in \{1, \dots, n_h\}$  and  $l \in \{1, \dots, m_1\}$ . Further, by using Assumptions H-2 and H-3, Remark 2.3, Hölder's inequality, Lemma 2.4 and  $h \leq 1$ , one obtains

$$\begin{aligned}
& \|\sigma_u(r, X_r^{i,N}, \mu_r^{X,N}) - \sigma_u(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})}^q \\
& \leq L \{ \|X_r^{i,N} - X_{j-1}^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \\
& + \|\mathcal{W}_2(\mu_r^{X,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\tilde{\Omega})} \\
& + |r - t_{j-1}| (1 + \|X_{j-1}^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} + \|\mathcal{W}_2(\mu_{j-1}^{X,N}, \delta_0)\|_{\mathcal{L}^q(\tilde{\Omega})}) \} \\
(36) \quad & \leq 2L \max_{i \in \{1, \dots, N\}} \|X_r^{i,N} - X_{j-1}^{i,N}\|_{\mathcal{L}^q(\tilde{\Omega})} \\
& + |r - t_{j-1}| L \left\{ 1 + 2C_1^{1/\bar{p}} \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{1/\bar{p}} \right\} \\
& \leq \left\{ 2L(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{1/\bar{p}} + L \right\} h_j^{1/2},
\end{aligned}$$

for all  $i \in \{1, \dots, N\}$ ,  $j \in \{1, \dots, n_h\}$ ,  $u \in \{0, 1\}$  and  $r \in [t_{j-1}, t_j]$ . On substituting (36) in (35), one obtains

$$\|\zeta^\ell(s)\|_{\mathcal{L}^q(\tilde{\Omega})} \leq \left\{ L + 2\bar{L}^2 + 4L^2 \left( \frac{q(q-1)}{2} \right)^{1/2} + \left( 2(L + 2\bar{L}^2)C_1^{1/\bar{p}} + 8L^2(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) \right) \right\} h_j^{1/2}$$

$$\times \left( \frac{q(q-1)}{2} \right)^{1/2} + 4LC_2^{2/\bar{p}} \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{2/\bar{p}} \Big\} h_j,$$

which, substituting in (34), gives the first estimate. The second estimate can be proved by similar arguments.  $\square$

**PROPOSITION 5.9.** *Let Assumptions H-1 with  $\bar{p} \geq 4$  and set  $q = \bar{p}/2 \geq 2$ , H-2 to H-4 hold. Then, the drift-randomised Milstein scheme (13) is consistent of order 1, that is,*

$$\|\mathcal{R}[X^{\mathcal{Q}h}]\|_{\mathcal{G}_{S,q}^h} \leq C_{10}h,$$

where

$$C_{10} := 2L(1 + 2\sqrt{q(q-1)/2}) \left\{ L + 2L(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) \right. \\ \left. \times \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right)^{1/\bar{p}} \right\} + C_{11} + C_{12}$$

and the positive constants  $C_1$ ,  $C_2$ ,  $C_{11}$  and  $C_{12}$  appear in Proposition 2.2, Lemma 2.4, Corollary 5.7 and Lemma 5.8, respectively.

**PROOF.** Let us recall the residual  $\mathcal{R}_j^{i,N}[X^{\mathcal{Q}h}]$  from (26), the interacting particle system from (6) and  $\Gamma_j^h$  from (25) to write  $\mathcal{R}_0^{i,N}[X^{\mathcal{Q}h}] = 0$  and

$$\begin{aligned} \mathcal{R}_j^{i,N}[X^{\mathcal{Q}h}] &= X_j^{i,N} - X_{j-1}^{i,N} - \Gamma_j^h(X_{j-1}^{i,N}, \mu_{j-1}^{X,N}, X_{j,\eta}^{i,N}, \mu_{j,\eta}^{X,N}, \eta_j) \\ &= \int_{t_{j-1}}^{t_j} [b(s, X_s^{i,N}, \mu_s^{X,N}) - b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N}, \mu_{j,\eta}^{X,N})] ds \\ (37) \quad &+ \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) - \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{i,\ell} \\ &+ \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} [\sigma_0^\ell(s, X_s^{i,N}, \mu_s^{X,N}) - \tilde{\sigma}_0^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{0,\ell}, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . For any  $\bar{n} \in \{1, \dots, n_h\}$ , from (27),

$$\Theta_{\bar{n},\eta}^h[b] = \sum_{j=1}^{\bar{n}} h_j b(t_{j-1} + \eta_j h_j, X_{t_{j-1} + \eta_j h_j}^{i,N}, \mu_{t_{j-1} + \eta_j h_j}^{X,N}),$$

which along with Assumption H-2 yields

$$\begin{aligned} &\sum_{j=1}^{\bar{n}} \int_{t_{j-1}}^{t_j} [b(s, X_s^{i,N}, \mu_s^{X,N}) - b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N}, \mu_{j,\eta}^{X,N})] ds \\ &= \sum_{j=1}^{\bar{n}} \int_{t_{j-1}}^{t_j} [b(s, X_s^{i,N}, \mu_s^{X,N}) - b(t_{j-1} + \eta_j h_j, X_{t_{j-1} + \eta_j h_j}^{i,N}, \mu_{t_{j-1} + \eta_j h_j}^{X,N})] ds \\ &\quad + \sum_{j=1}^{\bar{n}} \int_{t_{j-1}}^{t_j} [b(t_{j-1} + \eta_j h_j, X_{t_{j-1} + \eta_j h_j}^{i,N}, \mu_{t_{j-1} + \eta_j h_j}^{X,N}) - b(t_{j-1} + \eta_j h_j, X_{j,\eta}^{i,N}, \mu_{j,\eta}^{X,N})] ds \\ &\leq \int_0^{t_{\bar{n}}} b(s, X_s^{i,N}, \mu_s^{X,N}) ds - \Theta_{\bar{n},\eta}^h[b] \end{aligned}$$

$$+ L \sum_{j=1}^{\bar{n}} h_j [ |X_{t_{j-1}+\eta_j h_j}^{i,N} - X_{j,\eta}^{i,N}| + |\mathcal{W}_2(\mu_{t_{j-1}+\eta_j h_j}^{X,N}, \mu_{j,\eta}^{X,N})| ],$$

and thus from (37) one obtains

$$\begin{aligned} & \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} \mathcal{R}_j^{i,N} [X^{\mathcal{Q}_h}] \right| \right\|_{\mathcal{L}^q(\Omega)} \\ & \leq \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \Theta_{\bar{n}, \eta}^h [b] - \int_0^{t_{\bar{n}}} b(s, X_s^{i,N}, \mu_s^{X,N}) ds \right| \right\|_{\mathcal{L}^q(\Omega)} \\ & \quad + L \sum_{j=1}^{n_h} h_j [ \|X_{t_{j-1}+\eta_j h_j}^{i,N} - X_{j,\eta}^{i,N}\|_{\mathcal{L}^q(\Omega)} + \|\mathcal{W}_2(\mu_{t_{j-1}+\eta_j h_j}^{X,N}, \mu_{j,\eta}^{X,N})\|_{\mathcal{L}^q(\Omega)}] \\ (38) \quad & + \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_1} \int_{t_{j-1}}^{t_j} [\sigma_1^\ell(s, X_s^{i,N}, \mu_s^{X,N}) \right. \right. \\ & \quad \left. \left. - \tilde{\sigma}_1^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{i,\ell} \right| \right\|_{\mathcal{L}^q(\Omega)} \\ & + \left\| \max_{\bar{n} \in \{1, \dots, n_h\}} \left| \sum_{j=1}^{\bar{n}} \sum_{\ell=1}^{m_0} \int_{t_{j-1}}^{t_j} [\sigma_0^\ell(s, X_s^{i,N}, \mu_s^{X,N}) \right. \right. \\ & \quad \left. \left. - \tilde{\sigma}_0^\ell(s, t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})] dW_s^{0,\ell} \right| \right\|_{\mathcal{L}^q(\Omega)}, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$ . Furthermore, Theorem 7.1 in [43] yields

$$\begin{aligned} & \|X_{t_{j-1}+\eta_j h_j}^{i,N} - X_{j,\eta}^{i,N}\|_{\mathcal{L}^q(\Omega)} \\ & \leq h_j^{(q-1)/q} \left( \int_{t_{j-1}}^{t_j} \|b(s, X_s^{i,N}, \mu_s^{X,N}) - b(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\Omega)}^q ds \right)^{1/q} \\ & \quad + \left( \frac{q(q-1)}{2} \right)^{1/2} h_j^{(q-2)/2q} \\ & \quad \times \sum_{u=0}^1 \left( \int_{t_{j-1}}^{t_j} \|\sigma_u(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_u(t_{j-1}, X_{j-1}^{i,N}, \mu_{j-1}^{X,N})\|_{\mathcal{L}^q(\Omega)}^q ds \right)^{1/q}, \end{aligned}$$

which on using equations (33) and (36),

$$\begin{aligned} & \|X_{t_{j-1}+\eta_j h_j}^{i,N} - X_{j,\eta}^{i,N}\|_{\mathcal{L}^q(\Omega)} \leq \left( 1 + 2 \left( \frac{q(q-1)}{2} \right)^{1/2} \right) \{ L + 2L(C_1^{1/\bar{p}} + C_2^{1/\bar{p}}) \\ (39) \quad & \times \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^i\|_{\mathcal{L}^{\bar{p}}(\bar{\Omega})}^{\bar{p}} \right)^{1/\bar{p}} \} h_j, \end{aligned}$$

for all  $i \in \{1, \dots, N\}$  and  $j \in \{1, \dots, n_h\}$ . Also,

$$(40) \quad \|\mathcal{W}_2(\mu_{t_{j-1}+\eta_j h_j}^{X,N}, \mu_{j,\eta}^{X,N})\|_{\mathcal{L}^q(\Omega)} \leq \max_{i \in \{1, \dots, N\}} \|X_{t_{j-1}+\eta_j h_j}^{i,N} - X_{j,\eta}^{i,N}\|_{\mathcal{L}^q(\Omega)},$$

for all  $j \in \{1, \dots, n_h\}$ . The proof is completed by substituting (39) and (40) in (38) and using Corollary 5.7 and Lemma 5.8.  $\square$

5.4. *Rate of convergence of the scheme.* After proving bistability and consistency of the scheme (13), the main result, Theorem 3.3, follows immediately.

PROOF OF THEOREM 3.3. Set  $Y^h$  in Proposition 5.5 as  $Y^h = X^{\varrho h} = \{X_j^{i,N}\}_{j \in \{0, \dots, n_h\}} \in \mathcal{G}_q^h$  (as introduced in Section 5.3) to obtain,

$$\|X^{\varrho h} - X^h\|_{\mathcal{G}_q^h} = \max_{i \in \{1, \dots, N\}} \left\| \max_{j \in \{0, 1, \dots, n_h\}} |X_j^{i,N} - X_j^{i,N,h}| \right\|_{\mathcal{L}^q(\Omega)} \leq C_7 \|\mathcal{R}[X^{\varrho h}]\|_{\mathcal{G}_{S,q}^h},$$

which on using Proposition 5.9 completes the proof.  $\square$

## APPENDIX: AUXILIARY RESULTS

The following lemma is the discrete Burkholder–Davis–Gundy inequality, see [11].

LEMMA A.1 (Discrete Burkholder–Davis–Gundy inequality). *Let  $\{M_n\}_{n \in \bar{\mathbb{N}}}$  be a discrete-time martingale on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with respect to the filtration  $\{\mathcal{F}_n\}_{n \in \bar{\mathbb{N}}}$ . Then, there exist constants  $\underline{C}_p, \bar{C}_p > 0$  such that,*

$$\underline{C}_p \| [M]_n^{1/2} \|_{\mathcal{L}^p(\Omega)} \leq \left\| \max_{j \in \{0, 1, \dots, n\}} |M_j| \right\|_{\mathcal{L}^p(\Omega)} \leq \bar{C}_p \| [M]_n^{1/2} \|_{\mathcal{L}^p(\Omega)},$$

for any  $p \geq 1$  where  $\{[M]_n\}_{n \in \bar{\mathbb{N}}}$  is the quadratic variation process of  $\{M_n\}_{n \in \bar{\mathbb{N}}}$ .

The discrete version of Grönwall’s inequality is given below, see Proposition 4.1 in [27].

LEMMA A.2 (Discrete Grönwall inequality). *Let  $\{y_n\}_{n \in \mathbb{N}}$  and  $\{z_n\}_{n \in \mathbb{N}}$  be sequences of nonnegative real numbers satisfying  $y_n \leq a + \sum_{j=1}^{n-1} z_j y_j$  for all  $n \in \mathbb{N}$  where  $a > 0$  is a constant. Then,  $y_n \leq a \exp(\sum_{j=1}^{n-1} z_j)$  for all  $n \in \mathbb{N}$ .*

**A.1. Proof of Proposition 2.2.** Let  $t' \in [0, T]$ . Using (4) and Hölder’s inequality, we get

$$\begin{aligned} \sup_{t \in [0, t']} |X_t|^{\bar{p}} &\leq 4^{\bar{p}-1} \left\{ |X_0|^{\bar{p}} + T^{\bar{p}-1} \int_0^{t'} |b(s, X_s, \mathcal{L}^1(X_s))|^{\bar{p}} ds \right. \\ &\quad + \sup_{t \in [0, t']} \left| \int_0^t \sigma_1(s, X_s, \mathcal{L}^1(X_s)) dW_s \right|^{\bar{p}} \\ &\quad \left. + \sup_{t \in [0, t']} \left| \int_0^t \sigma_0(s, X_s, \mathcal{L}^1(X_s)) dW_s^0 \right|^{\bar{p}} \right\}, \end{aligned}$$

which on using the Burkholder–Gundy–Davis inequality (Theorem 7.2 in [43]) and Remark 2.1 yields

$$\begin{aligned} \left\| \sup_{t \in [0, t']} |X_t| \right\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} &\leq 4^{\bar{p}-1} \left\{ \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + T^{\bar{p}-1} \int_0^{t'} \|b(s, X_s, \mathcal{L}^1(X_s))\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} ds \right. \\ &\quad \left. + \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{(\bar{p}-2)/2} \sum_{u=0}^1 \int_0^{t'} \|\sigma_u(s, X_s, \mathcal{L}^1(X_s))\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} ds \right\} \\ &\leq 4^{\bar{p}-1} \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + 4^{\bar{p}-1} 3^{\bar{p}-1} \bar{L}^{\bar{p}} \left( T^{\bar{p}-1} + 2 \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{(\bar{p}-2)/2} \right) \\ &\quad \times \int_0^{t'} \{1 + \|X_s\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + \|\mathcal{W}_2(\mathcal{L}^1(X_s), \delta_0)\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}}\} ds \end{aligned}$$



$$\begin{aligned}
&\leq 4^{\bar{p}-1} \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + 12^{\bar{p}-1} \bar{L}^{\bar{p}} \left( T^{\bar{p}} + 2 \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{\bar{p}/2} \right) \\
&\quad + 12^{\bar{p}-1} \bar{L}^{\bar{p}} \left( T^{\bar{p}-1} + 2 \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{(\bar{p}-2)/2} \right) \\
&\quad \times 2 \int_0^{t'} \left\| \sup_{r \in [0, s]} |X_r| \right\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} ds
\end{aligned}$$

for all  $t' \in [0, T]$ . Then, Grönwall's inequality gives

$$\begin{aligned}
\left\| \sup_{t \in [0, T]} |X_t| \right\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} &\leq \left( 4^{\bar{p}-1} \|X_0\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + 12^{\bar{p}-1} \bar{L}^{\bar{p}} \left( T^{\bar{p}} + 2 \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{\bar{p}/2} \right) \right) \\
&\quad \times \exp \left( 12^{\bar{p}-1} \bar{L}^{\bar{p}} \left( T^{\bar{p}} + 2 \left( \frac{\bar{p}^3}{2(\bar{p}-1)} \right)^{\bar{p}/2} T^{\bar{p}/2} \right) 2 \right),
\end{aligned}$$

which completes the proof.  $\square$

**A.2. Proof of Lemma 2.4.** Recall equation (6) and use Theorem 7.1 in [43] along with Remark 2.1 to get the following:

$$\begin{aligned}
\|X_t^{i,N} - X_{t'}^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} &\leq 3^{\bar{p}-1} (t - t')^{\bar{p}-1} \int_{t'}^t \|b(s, X_s^{i,N}, \mu_s^{X,N})\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} ds \\
&\quad + 3^{\bar{p}-1} \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{\bar{p}/2} (t - t')^{(\bar{p}-2)/2} \\
&\quad \times \sum_{u=0}^1 \int_{t'}^t \|\sigma_u(s, X_s^{i,N}, \mu_s^{X,N})\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} ds \\
&\leq 9^{\bar{p}-1} L^{\bar{p}} \left( (t - t')^{\bar{p}-1} + 2 \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{\bar{p}/2} (t - t')^{(\bar{p}-2)/2} \right) \\
&\quad \times \int_{t'}^t (1 + \|X_s^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} + \|\mathcal{W}_2(\mu_s^{X,N}, \delta_0)\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}}) ds,
\end{aligned}$$

for all  $t > t' \in [0, T]$  and  $i \in \{1, \dots, N\}$ , which due to Remark 2.3 yields,

$$\begin{aligned}
\|X_t^{i,N} - X_{t'}^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} &\leq 9^{\bar{p}-1} L^{\bar{p}} \left( (t - t')^{\bar{p}-1} + 2 \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{\bar{p}/2} (t - t')^{(\bar{p}-2)/2} \right) \\
&\quad \times \int_{t'}^t \left\{ 1 + 2 \max_{i \in \{1, \dots, N\}} \|X_s^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right\} ds \\
&\leq 9^{\bar{p}-1} L^{\bar{p}} \left( (t - t')^{\bar{p}/2} + 2 \left( \frac{\bar{p}(\bar{p}-1)}{2} \right)^{\bar{p}/2} \right) \\
&\quad \times (t - t')^{\bar{p}/2} \left( 1 + 2C_1 \left( 1 + \max_{i \in \{1, \dots, N\}} \|X_0^{i,N}\|_{\mathcal{L}^{\bar{p}}(\tilde{\Omega})}^{\bar{p}} \right) \right)
\end{aligned}$$

and thus the proof is completed.  $\square$

**A.3. Proof of Proposition 2.5.** By using (5) and (6) along with Hölder's inequality, one obtains

$$\sup_{t \in [0, T]} |X_t^i - X_t^{i,N}|^2 \leq 3 \left\{ T \int_0^T |b(s, X_s^i, \mathcal{L}^1(X_s^1)) - b(s, X_s^{i,N}, \mu_s^{X,N})|^2 ds \right.$$

$$\begin{aligned}
 & + \sup_{t \in [0, T]} \left| \int_0^t (\sigma_1(s, X_s^i, \mathcal{L}^1(X_s^1)) - \sigma_1(s, X_s^{i, N}, \mu_s^{X, N})) dW_s \right|^2 \\
 & + \sup_{t \in [0, T]} \left| \int_0^t (\sigma_0(s, X_s^i, \mathcal{L}^1(X_s^1)) - \sigma_0(s, X_s^{i, N}, \mu_s^{X, N})) dW_s^0 \right|^2 \Big\},
 \end{aligned}$$

which on using the Martingale inequality (Theorem 7.2 in [43]) and Assumption H-2 yields

$$\begin{aligned}
 & \left\| \sup_{t \in [0, t']} |X_t^i - X_t^{i, N}| \right\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \\
 & \leq 3T \int_0^{t'} \|b(s, X_s^i, \mathcal{L}^1(X_s^1)) - b(s, X_s^{i, N}, \mu_s^{X, N})\|_{\mathcal{L}^2(\tilde{\Omega})}^2 ds \\
 & \quad + 12 \sum_{u=0}^1 \int_0^{t'} \|\sigma_u(s, X_s^i, \mathcal{L}^1(X_s^1)) - \sigma_u(s, X_s^{i, N}, \mu_s^{X, N})\|_{\mathcal{L}^2(\tilde{\Omega})}^2 ds \\
 & \leq 2(3T + 24)L^2 \int_0^{t'} \{ \|X_s^i - X_s^{i, N}\|_{\mathcal{L}^2(\tilde{\Omega})}^2 + \|\mathcal{W}_2(\mathcal{L}^1(X_s^1), \mu_s^{X, N})\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \} ds \\
 & \leq 2(3T + 24)L^2 \int_0^{t'} \{ \|X_s^i - X_s^{i, N}\|_{\mathcal{L}^2(\tilde{\Omega})}^2 + 2\|\mathcal{W}_2(\mathcal{L}^1(X_s^1), \mu_s^X)\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \\
 & \quad + 2\|\mathcal{W}_2(\mu_s^X, \mu_s^{X, N})\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \} ds \\
 & \leq 2(3T + 24)L^2 \int_0^{t'} \left\{ 3 \max_{i \in \{1, \dots, N\}} \|X_s^i - X_s^{i, N}\|_{\mathcal{L}^2(\tilde{\Omega})}^2 + 2\|\mathcal{W}_2(\mathcal{L}^1(X_s^1), \mu_s^X)\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \right\} ds
 \end{aligned}$$

for any  $t' \in [0, T]$  and  $N \in \mathbb{N}$ . An application of Grönwall's inequality yields

$$\begin{aligned}
 & \max_{i \in \{1, \dots, N\}} \left\| \sup_{t \in [0, T]} |X_t^i - X_t^{i, N}| \right\|_{\mathcal{L}^2(\tilde{\Omega})}^2 \\
 & \leq 4(3T + 24)L^2 e^{6(3T+24)TL^2} \int_0^T \|\mathcal{W}_2(\mathcal{L}^1(X_s^1), \mu_s^X)\|_{\mathcal{L}^2(\tilde{\Omega})}^2 ds,
 \end{aligned}$$

for any  $N \in \mathbb{N}$ . The proof is completed by using (7).  $\square$

**A.4. Lions derivative and particle projection function inequalities.** Given a function  $f : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$  and  $\nu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ , we say that  $f$  is Lions differentiable at  $\nu_0$  if we can find an atomless, Polish probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$  and a random variable  $Y_0 \in \mathcal{L}^2(\tilde{\Omega})$  with law  $\mathcal{L}^1(Y_0) := \tilde{P} \circ Y_0^{-1} = \nu_0$  such that the “lift” function  $F : \mathcal{L}^2(\tilde{\Omega}) \rightarrow \mathbb{R}$  defined by  $F(Z) := f(\mathcal{L}^1(Z))$  has Fréchet derivative  $F'[Y_0]$  at  $Y_0 \in \mathcal{L}^2(\tilde{\Omega})$ . By Riesz representation theorem, we can find  $DF(Y_0) \in \mathcal{L}^2(\tilde{\Omega})$  such that  $F'[Y_0](Z) = \tilde{E}\langle DF[Y_0], Z \rangle$  for all  $Z \in \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d)$ . Further, Theorem 6.5 (structure of the gradient) in [12] guarantees the existence of a function  $\partial_\mu f(\nu_0) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , independent of the choice of the probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$  and the random variable  $Y_0$ , satisfying  $\int_{\mathbb{R}^d} |\partial_\mu f(\nu_0)(x)|^2 \nu_0(dx) < \infty$  such that  $DF(Y_0) = \partial_\mu f(\nu_0)(Y_0)$ . Then, we call  $\partial_\mu f(\nu_0)$  as *Lions’ derivative* of  $f$  at  $\nu_0 = \mathcal{L}^1(Y_0)$ .

By using similar arguments used in Lemma 4.2 in [37], one can prove the following lemma.

**LEMMA A.3.** *Let  $g$  be a real valued function defined on  $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$  such that derivative with respect to the state variable  $\partial_x g(\cdot, \cdot, \cdot)$  and with respect to the measure variable  $\partial_\mu g(\cdot, \cdot, \cdot)$  satisfy Lipschitz condition uniformly in time, that is, there exists a constant*

$L > 0$  such that,

$$|\partial_x g(t, x, \mu) - \partial_x g(t, \bar{x}, \bar{\mu})| \leq L\{|x - \bar{x}| + \mathcal{W}_2(\mu, \bar{\mu})\},$$

$$|\partial_\mu g(t, x, \mu, y) - \partial_\mu g(t, \bar{x}, \bar{\mu}, \bar{y})| \leq L\{|x - \bar{x}| + \mathcal{W}_2(\mu, \bar{\mu}) + |y - \bar{y}|\},$$

for all  $t \in [0, T]$ ,  $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$  and  $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ . Then, for all  $t \in [0, T]$ ,  $x^i, \bar{x}^i \in \mathbb{R}^d$  and  $i \in \{1, \dots, N\}$ , the following holds:

$$\left| g\left(t, x^i, \frac{1}{N} \sum_{k=1}^N \delta_{x^k}\right) - g\left(t, \bar{x}^i, \frac{1}{N} \sum_{k=1}^N \delta_{\bar{x}^k}\right) - \partial_x g\left(t, \bar{x}^i, \frac{1}{N} \sum_{k=1}^N \delta_{\bar{x}^k}\right)(x^i - \bar{x}^i) - \frac{1}{N} \sum_{k=1}^N \partial_\mu g\left(t, \bar{x}^i, \frac{1}{N} \sum_{k=1}^N \delta_{\bar{x}^k}, \bar{x}^k\right)(x^k - \bar{x}^k) \right| \leq \frac{3L}{2}|x^i - \bar{x}^i|^2 + \frac{5L}{2} \frac{1}{N} \sum_{k=1}^N |x^k - \bar{x}^k|^2.$$

**Acknowledgments.** The authors would like to thank the anonymous referees whose comments improved the quality of this paper.

**Funding.** G. dos Reis acknowledges support from the *Fundação para a Ciência e a Tecnologia* (Portuguese Foundation for Science and Technology) through the project UIDB/00297/2020 and UIDP/00297/2020 (Center for Mathematics and Applications).

## REFERENCES

- [1] ADAMS, D., DOS REIS, G., RAVAILLE, R., SALKELD, W. and TUGAUT, J. (2022). Large deviations and exit-times for reflected McKean–Vlasov equations with self-stabilising terms and superlinear drifts. *Stochastic Process. Appl.* **146** 264–310. [MR4374937](#) <https://doi.org/10.1016/j.spa.2021.12.017>
- [2] BALADRON, J., FASOLI, D., FAUGERAS, O. and TOUBOUL, J. (2012). Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons. *J. Math. Neurosci.* **2** Art. 10, 50. [MR2974499](#) <https://doi.org/10.1186/2190-8567-2-10>
- [3] BAO, J. and HUANG, X. (2022). Approximations of McKean–Vlasov stochastic differential equations with irregular coefficients. *J. Theoret. Probab.* **35** 1187–1215. [MR4414415](#)
- [4] BAO, J., REISINGER, C., REN, P. and STOCKINGER, W. (2021). First-order convergence of Milstein schemes for McKean–Vlasov equations and interacting particle systems. *Proc. R. Soc. A* **477** Paper No. 20200258, 27. [MR4212406](#)
- [5] BAO, J., REISINGER, C., REN, P. and STOCKINGER, W. (2023). Milstein schemes and antithetic multilevel Monte Carlo sampling for delay McKean–Vlasov equations and interacting particle systems. *IMA J. Numer. Anal.* drad064. <https://doi.org/10.1093/imanum/drad064>.
- [6] BAUER, M., MEYER-BRANDIS, T. and PROSKE, F. (2018). Strong solutions of mean-field stochastic differential equations with irregular drift. *Electron. J. Probab.* **23** 1–35. [MR3896869](#) <https://doi.org/10.1214/18-EJP259>
- [7] BEYN, W.-J. and KRUSE, R. (2010). Two-sided error estimates for the stochastic theta method. *Discrete Contin. Dyn. Syst. Ser. B* **14** 389–407. [MR2660864](#) <https://doi.org/10.3934/dcdsb.2010.14.389>
- [8] BOLLEY, F., CAÑIZO, J. A. and CARRILLO, J. A. (2011). Stochastic mean-field limit: Non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.* **21** 2179–2210. [MR2860672](#) <https://doi.org/10.1142/S0218202511005702>
- [9] BOSSY, M., FAUGERAS, O. and TALAY, D. (2015). Clarification and complement to “Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons”. *J. Math. Neurosci.* **5** Art. 19, 23. [MR3392551](#) <https://doi.org/10.1186/s13408-015-0031-8>
- [10] BOSSY, M. and TALAY, D. (1997). A stochastic particle method for the McKean–Vlasov and the Burgers equation. *Math. Comp.* **66** 157–192. [MR1370849](#) <https://doi.org/10.1090/S0025-5718-97-00776-X>
- [11] BURKHOLDER, D. L. (1966). Martingale transforms. *Ann. Math. Stat.* **37** 1494–1504. [MR0208647](#) <https://doi.org/10.1214/aoms/1177699141>
- [12] CARDALIAGUET, P. (2013). Notes on Mean Field Games, notes from P. L. Lions lectures at Collège de France. <https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf>.
- [13] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. I: Mean Field FBSDEs, Control, and Games. Probability Theory and Stochastic Modelling* **83**. Springer, Cham. [MR3752669](#)

- [14] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. II: Mean Field Games with Common Noise and Master Equations. Probability Theory and Stochastic Modelling* **84**. Springer, Cham. MR3753660
- [15] CARRILLO, J. A., CHOI, Y.-P. and TSE, O. (2019). Convergence to equilibrium in Wasserstein distance for damped Euler equations with interaction forces. *Comm. Math. Phys.* **365** 329–361. MR3900833 <https://doi.org/10.1007/s00220-018-3276-8>
- [16] CARRILLO, J. A., FORNASIER, M., TOSCANI, G. and VECIL, F. (2010). Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences. Model. Simul. Sci. Eng. Technol.* 297–336. Birkhäuser, Boston, MA. MR2744704 [https://doi.org/10.1007/978-0-8176-4946-3\\_12](https://doi.org/10.1007/978-0-8176-4946-3_12)
- [17] CHAUDRU DE RAYNAL, P. E. (2020). Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.* **130** 79–107. MR4035024 <https://doi.org/10.1016/j.spa.2019.01.006>
- [18] CHEN, X. and DOS REIS, G. (2022). A flexible split-step scheme for solving McKean–Vlasov stochastic differential equations. *Appl. Math. Comput.* **427** Paper No. 127180, 23. MR4413221 <https://doi.org/10.1016/j.amc.2022.127180>
- [19] CHEN, X. and DOS REIS, G. (2023). Euler simulation of interacting particle systems and McKean–Vlasov SDEs with fully super-linear growth drifts in space and interaction. *IMA J. Numer. Anal.* drad022. <https://doi.org/10.1093/imanum/drad022>
- [20] CHEN, X., DOS REIS, G. and STOCKINGER, W. (2023). Wellposedness, exponential ergodicity and numerical approximation of fully super-linear McKean–Vlasov SDEs and associated particle systems. arXiv preprint. Available at [arXiv:2302.05133](https://arxiv.org/abs/2302.05133).
- [21] COGHI, M. and GESS, B. (2019). Stochastic nonlinear Fokker–Planck equations. *Nonlinear Anal.* **187** 259–278. MR3954095 <https://doi.org/10.1016/j.na.2019.05.003>
- [22] COGHI, M. and NILSSEN, T. (2021). Rough nonlocal diffusions. *Stochastic Process. Appl.* **141** 1–56. MR4293767 <https://doi.org/10.1016/j.spa.2021.07.002>
- [23] DELARUE, F., LACKER, D. and RAMANAN, K. (2019). From the master equation to mean field game limit theory: A central limit theorem. *Electron. J. Probab.* **24** 1–54. MR3954791 <https://doi.org/10.1214/19-EJP298>
- [24] DOS REIS, G., ENGELHARDT, S. and SMITH, G. (2022). Simulation of McKean–Vlasov SDEs with super-linear growth. *IMA J. Numer. Anal.* **42** 874–922. MR4367675 <https://doi.org/10.1093/imanum/draa099>
- [25] DOS REIS, G., SALKELD, W. and TUGAUT, J. (2019). Freidlin–Wentzell LDP in path space for McKean–Vlasov equations and the functional iterated logarithm law. *Ann. Appl. Probab.* **29** 1487–1540. MR3914550 <https://doi.org/10.1214/18-AAP1416>
- [26] DREYER, W., GABERŠČEK, M., GUHLKE, C., HUTH, R. and JAMNIK, J. (2011). Phase transition in a rechargeable lithium battery. *European J. Appl. Math.* **22** 267–290. MR2795141 <https://doi.org/10.1017/S0956792511000052>
- [27] EMMRICH, E. (1999). *Discrete Versions of Grönwall’s Lemma and Their Application to the Numerical Analysis of Parabolic Problems*. TU, Berlin, FB Mathematik, Preprint.
- [28] GODDARD, B. D., PAVLIOTIS, G. A. and KALLIADASIS, S. (2012). The overdamped limit of dynamic density functional theory: Rigorous results. *Multiscale Model. Simul.* **10** 633–663. MR2968853 <https://doi.org/10.1137/110844659>
- [29] GUHLKE, C., GAJEWSKI, P., MAURELLI, M., FRIZ, P. K. and DREYER, W. (2018). Stochastic many-particle model for LFP electrodes. *Contin. Mech. Thermodyn.* **30** 593–628. MR3783112 <https://doi.org/10.1007/s00161-018-0629-7>
- [30] GYÖNGY, I. and KRYLOV, N. V. (1980). On stochastic equations with respect to semimartingales. I. *Stochastics* **4** 1–21. MR0587426 <https://doi.org/10.1080/03610918008833154>
- [31] HARANG, F. A. and MAYORCAS, A. (2023). Pathwise regularisation of singular interacting particle systems and their mean field limits. *Stochastic Process. Appl.* **159** 499–540. MR4555283 <https://doi.org/10.1016/j.spa.2023.02.005>
- [32] HOLM, D. D. and PUTKARADZE, V. (2006). Formation of clumps and patches in self-aggregation of finite-size particles. *Phys. D* **220** 183–196. MR2256869 <https://doi.org/10.1016/j.physd.2006.07.010>
- [33] JIN, S., LI, L. and LIU, J.-G. (2020). Random batch methods (RBM) for interacting particle systems. *J. Comput. Phys.* **400** 108877, 30. MR4019084 <https://doi.org/10.1016/j.jcp.2019.108877>
- [34] KOLOKOLNIKOV, T., CARRILLO, J. A., BERTOZZI, A., FETEAU, R. and LEWIS, M. (2013). Emergent behaviour in multi-particle systems with non-local interactions [Editorial]. *Phys. D* **260** 1–4. MR3143990 <https://doi.org/10.1016/j.physd.2013.06.011>
- [35] KRUSE, R. (2012). Characterization of bistability for stochastic multistep methods. *BIT* **52** 109–140. MR2891656 <https://doi.org/10.1007/s10543-011-0341-5>

- [36] KRUSE, R. and WU, Y. (2019). A randomized Milstein method for stochastic differential equations with non-differentiable drift coefficients. *Discrete Contin. Dyn. Syst. Ser. B* **24** 3475–3502. [MR3986241](#) <https://doi.org/10.3934/dcdsb.2018253>
- [37] KUMAR, C. and NEELIMA (2021). On explicit Milstein-type scheme for McKean–Vlasov stochastic differential equations with super-linear drift coefficient. *Electron. J. Probab.* **26** 1–32. [MR4302574](#) <https://doi.org/10.1214/21-ejp676>
- [38] KUMAR, C., NEELIMA, REISINGER, C. and STOCKINGER, W. (2022). Well-posedness and tamed schemes for McKean–Vlasov equations with common noise. *Ann. Appl. Probab.* **32** 3283–3330. [MR4497846](#) <https://doi.org/10.1214/21-aap1760>
- [39] KURTZ, T. G. and XIONG, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* **83** 103–126. [MR1705602](#) [https://doi.org/10.1016/S0304-4149\(99\)00024-1](https://doi.org/10.1016/S0304-4149(99)00024-1)
- [40] 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>
- [41] LACKER, D. and LE FLEM, 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>
- [42] LI, Y., MAO, X., SONG, Q., WU, F. and YIN, G. (2023). Strong convergence of Euler–Maruyama schemes for McKean–Vlasov stochastic differential equations under local Lipschitz conditions of state variables. *IMA J. Numer. Anal.* **43** 1001–1035. [MR4568438](#) <https://doi.org/10.1093/imanum/drab107>
- [43] MAO, X. (2008). *Stochastic Differential Equations and Applications*, 2nd ed. Horwood, Chichester. [MR2380366](#) <https://doi.org/10.1533/9780857099402>
- [44] MCKEAN, H. P. JR. (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>
- [45] MEHRI, S., SCHEUTZOW, M., STANNAT, W. and ZANGENEH, B. Z. (2020). Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.* **30** 175–207. [MR4068309](#) <https://doi.org/10.1214/19-AAP1499>
- [46] MISHURA, Y. and VERETENNIKOV, A. (2020). Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory Probab. Math. Statist.* **103** 59–101. [MR4421344](#) <https://doi.org/10.1090/tpms/1135>
- [47] MORKISZ, P. and PRZYBYŁOWICZ, P. (2021). Randomized derivative-free Milstein algorithm for efficient approximation of solutions of SDEs under noisy information. *J. Comput. Appl. Math.* **383** Paper No. 113112, 22. [MR4131887](#) <https://doi.org/10.1016/j.cam.2020.113112>
- [48] POSSAMAÏ, D. and TANGPI, L. (2021). Non-asymptotic convergence rates for mean-field games: Weak formulation and McKean–Vlasov BSDEs. arXiv preprint. Available at [arXiv:2105.00484](#).
- [49] PRZYBYŁOWICZ, P., SCHWARZ, V. and SZÖLGYENYI, M. (2024). Randomized Milstein algorithm for approximation of solutions of jump-diffusion SDEs. *J. Comput. Appl. Math.* **440** Paper No. 115631, 21. [MR4660616](#) <https://doi.org/10.1016/j.cam.2023.115631>
- [50] REISINGER, C. and STOCKINGER, W. (2022). An adaptive Euler–Maruyama scheme for McKean–Vlasov SDEs with super-linear growth and application to the mean-field FitzHugh–Nagumo model. *J. Comput. Appl. Math.* **400** Paper No. 113725, 23. [MR4293705](#) <https://doi.org/10.1016/j.cam.2021.113725>
- [51] REISINGER, C., STOCKINGER, W. and ZHANG, Y. (2021). A fast iterative PDE-based algorithm for feedback controls of nonsmooth mean-field control problems. arXiv preprint. Available at [arXiv:2108.06740](#).
- [52] REISINGER, C., STOCKINGER, W. and ZHANG, Y. (2023). Linear convergence of a policy gradient method for some finite horizon continuous time control problems. *SIAM J. Control Optim.* **61** 3526–3558. [MR4679972](#) <https://doi.org/10.1137/22M1492180>
- [53] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D’Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. [MR1108185](#) <https://doi.org/10.1007/BFb0085169>
- [54] WIKTORSSON, M. (2001). Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions. *Ann. Appl. Probab.* **11** 470–487. [MR1843055](#) <https://doi.org/10.1214/aoap/1015345301>