# Projected particle methods for solving McKean-Vlasov stochastic differential equations

Denis Belomestny[1] and John Schoenmakers [2]

April 24, 2018

**Abstract**

We propose a novel projection-based particle method for solving McKean-Vlasov stochastic differential equations. Our approach is based on a projection-type estimation of the marginal density of the solution in each time step. The projection-based particle method leads in many situation to a significant reduction of numerical complexity compared to the widely used kernel density estimation algorithms. We derive strong convergence rates and rates of density estimation. The convergence analysis, particularly in the case of linearly growing coefficients, turns out to be rather challenging and requires some new type of averaging technique. This case is exemplified by explicit solutions to a class of McKean-Vlasov equations with affine drift. The performance of the proposed algorithm is illustrated by several numerical examples.

*Keywords:* Mckean-Vlasov equations, particle systems, projection estimators, explicit solutions.
*2010 MSC:* 60H10, 60K35.

## 1 Introduction

Nonlinear Markov processes are stochastic processes whose transition functions may depend not only on the current state of the process but also on the current distribution of the process. These processes were introduced by McKean [12] to model plasma dynamics. Later nonlinear Markov processes were studied by a number of authors; we mention here the books of Kolokoltsov [9] and Sznitman [16]. These processes arise naturally in the study of the limit behavior of a large number of weakly interacting Markov processes and have a wide range of applications, including financial mathematics, population dynamics, and neuroscience (see, e.g., [4] and the references therein).

Let $[0, T]$ be a finite time interval and $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete probability space, where a standard $m$-dimensional Brownian motion $W$ is defined. We consider a class of McKean-Vlasov SDEs, i.e. stochastic differential equation (SDE) whose drift and diffusion coefficients may depend on the current distribution of

[1]Duisburg-Essen University, Thea-Leymann-Str. 9, D-45127 Essen, Germany, denis.belomestny@uni-due.de
[2]Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39, 10117 Berlin, Germany, schoenma@wias-berlin.de

the process of the form:

$$\begin{cases} X_t & = \xi + \int_0^t \int_{\mathbb{R}^d} a(X_s, y)\mu_s(dy)ds + \int_0^t \int_{\mathbb{R}^d} b(X_s, y)\mu_s(dy)dW_s \\ \mu_t & = \mathrm{Law}(X_t), \quad t \in [0, T], \end{cases} \tag{1}$$

where $X_0 = \xi$ is an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^d$, $a : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$. If the functions $a$ and $b$ are smooth with uniformly bounded derivatives and the random variable $\xi$ has finite moments of any order, then (see [1]) there is a unique strong solution of (1) such that for all $p > 1$,

$$\mathrm{E}\left[\sup_{s\le T} |X_s|^p\right] \le \infty. \tag{2}$$

In the sequel we assume that there exists a unique strong solution of (1) such that (2) holds and refer to [5] for more general sufficient conditions for this.

Assume that $d = 1$ and for any $t \ge 0$, the measure $\mu_t(du)$ possesses a bounded density $\mu_t(u)$. Then the family of these densities satisfies a nonlinear Fokker-Planck equation of the form

$$\begin{aligned} \frac{\partial \mu_t(x)}{\partial t} &= -\frac{\partial}{\partial x}\left(\left(\int a(x, y)\mu_t(y)\, dy\right) \mu_t(x)\right) \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\left(\int b(x, y)\mu_t(y)\, dy\right)^2 \mu_t(x)\right), \end{aligned} \tag{3}$$

which can be seen as an analogue of a well-known linear Fokker-Planck equation in the case of linear stochastic differential equations. In Section 4.1 we will show that if the drift $a$ is affine in $x$, and the diffusion coefficient $b$ is independent of $x$, then the system (1), and hence (3), has an explicit solution. These solutions, apart from being interesting in their own right, also provide explicit cases of an explosive behavior.

The theory of the propagation of chaos developed in [16], states that (1) is a limiting equation of the system of stochastic interacting particles (samples) with the following dynamics

$$X_t^{i,N} = \xi^i + \int_0^t \int_{\mathbb{R}^d} a(X_s^{i,N}, y)\mu_s^N(dy)\, ds + \int_0^t \int_{\mathbb{R}^d} b(X_s^{i,N}, y)\mu_s^N(dy)\, dW_s^i \tag{4}$$

for $i = 1, \dots, N$, where $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$, $\xi^i$, $i = 1, \dots, N$, are i.i.d copies of $\xi$, distributed according the law $\mu_0$, and $W^i$, $i = 1, ..., N$, are independent copies of $W$. In fact it can be shown, under sufficient regularity conditions on the coefficients, that convergence in law for empirical measures on the path space holds, i.e., $\mu^N = \{\mu_t^N : t \in [0, T]\} \to \mu$, $N \to \infty$, see [13].

Despite the numerous branches of research on stochastic particle systems, results on numerical approximations of McKean-Vlasov-SDEs are very sparse. The authors in [1] proposed to use the Euler scheme with time-step $h = T/L$, that for $l = 0, \dots, L-1$, yields

$$\bar{X}_{t_{l+1}}^{i,N} = \bar{X}_{t_l}^{i,N} + \frac{1}{N}\sum_{j=1}^N a(\bar{X}_{t_l}^{i,N}, \bar{X}_{t_l}^{j,N})\, h + \frac{1}{N}\sum_{j=1}^N b(\bar{X}_{t_l}^{i,N}, \bar{X}_{t_l}^{j,N})\, \Delta_{l+1} W^i \tag{5}$$

for $i = 1, \ldots, N$, $t_l = hl$, and $\Delta_{l+1} W^i = W^i_{h(l+1)} - W^i_{hl}$ (see also [2] for more general MVSDEs and [8] for Gauss-quadrature based approach). Implementation of the above algorithm requires usually $N^2 \times L$ operations in every step of the Euler scheme. By using the algorithm presented here one can significantly reduce the complexity of the particle simulation especially if the coefficients of the corresponding McKean-Vlasov SDE are smooth enough.

The contribution of this paper is twofold. On the one hand, we propose a new approximation methodology based on a projection-type estimation of the marginal densities of (1). This methodology often leads to numerically more efficient algorithms than the kernel-type approximation algorithms, as they can profit from a global smoothness of coefficients $a$,$b$ and the corresponding marginal densities. On the other hand, we present a comprehensive convergence analysis of the proposed algorithms in the case of possibly linearly growing (in $x$) coefficients $a$ and $b$. To the best of our knowledge, no stability analysis of MVSDEs under this linear growth assumption was done before. In fact such analysis is rather challenging and requires a special type of averaging technique. And, last but not least, we study a general class of MVSDEs with affine drift and derive their explicit solutions, to the best of our knowledge, for the first time.

The paper is organized as follows. In Section 2 we present the idea of our projected particle method. Section 3 is devoted to the convergence analysis of the projected particle method. In particular, in Section 3.1 we derive the convergence rates for the corresponding projected density estimate. Section 4 presents a thorough study of affine MVSDEs. Numerical examples for affine and convolution-type MVSDEs are presented in Section 5. All proofs are collected in Section 6.

## 2 Projected particle method

Let $w : \mathbb{R}^d \to \mathbb{R}_+$ be some weight function with $w > 0$, such that

$$a(x, \cdot), b(x, \cdot) \in L_2(\mathbb{R}^d, w) \quad \text{for any } x \in \mathbb{R}^d.$$

Let further $(\varphi_k,\ k = 0, 1, 2, ..)$ be a total orthonormal system in $L_2(\mathbb{R}^d, w)$. The corresponding (generalized) Fourier coefficients of the functions $a(x, \cdot)$ and $b(x, \cdot)$ are given by

$$\alpha_k(x) := \int a(x, u)\varphi_k(u)\, w(u)\, du \in \mathbb{R}^d, \tag{6}$$
$$\beta_k(x) := \int b(x, u)\varphi_k(u)\, w(u)\, du \in \mathbb{R}^{d\times m}$$

and the following series representation holds

$$a(x, \cdot) = \sum_{k=0}^{\infty} \alpha_k(x)\varphi_k(\cdot) \quad \text{and} \quad b(x, \cdot) = \sum_{k=0}^{\infty} \beta_k(x)\varphi_k(\cdot), \quad x \in \mathbb{R}^d,$$

in $L_2(\mathbb{R}^d, w)$. Further it is assumed that each function $\varphi_k$ is bounded so that the functions

$$\gamma_k(s) := \mathrm{E}\left[\varphi_k(X_s)\right] \tag{7}$$

are well defined. Now let us fix some natural number $K > 0$ and consider a *projected particle approximation* for (1)

$$X_t^{i,K,N} = \xi^i + \int_0^t \sum_{k=0}^K \gamma_k^N(s)\alpha_k(X_s^{i,K,N})\,ds + \int_0^t \sum_{k=0}^K \gamma_k^N(s)\beta_k(X_s^{i,K,N})\,dW_s^i \quad (8)$$

for $i = 1, \ldots, N$, where

$$\gamma_k^N(s) := \frac{1}{N}\sum_{j=1}^N \varphi_k(X_s^{j,K,N}) \quad (9)$$

can be regarded as an approximation to (7). The projected system (8), with (9), is heuristically motivated by assuming that for any $s \geq 0$ the measure $\mu_s(du)$ possesses a density $\mu_s(u)$ that, in view of

$$\gamma_k(s) = \int \mu_s(u)\varphi_k(u)du$$

(cf. (7)), formally satisfies

$$\mu_s(u) = \sum_{k=0}^\infty \gamma_k(s)\varphi_k(u)w(u). \quad (10)$$

Then, we (formally) have the expansion

$$\int a(x,u)\mu_s(u)du = \sum_{k=0}^\infty \alpha_k(x)\gamma_k(s),$$

and this motivates the drift term in (8). For the diffusion term in (8) an analogue motivation applies. In order to solve (8) we may consider, for any fixed $L > 0$, an Euler-type approximation,

$$\begin{aligned}\bar{X}_t^{i,K,N} &= \bar{X}_{\eta(t)}^{i,K,N} + \sum_{k=0}^K \gamma_k^N\left(\eta(t)\right)\,\alpha_k\left(\bar{X}_{\eta(t)}^{i,K,N}\right)(t-\eta(t)) \quad (11)\\ &+ \sum_{k=0}^K \gamma_k^N\left(\eta(t)\right)\,\beta_k\left(\bar{X}_{\eta(t)}^{i,K,N}\right)(W_t^i - W_{\eta(t)})\end{aligned}$$

for $i = 1, \ldots, N$, and $h = T/L$, where $\eta(t) := lh$ for $t \in [lh, (l+1)h)$, $l = 1, \ldots, L$. Note that in order to generate a discretized particle system $(\bar{X}_{hl}^{i,K,N})$, $i = 1, \ldots, N$, $l = 1, \ldots, L$, we need to perform (up to a constant depending on the dimension) $NLK$ operations. This should be compared to $N^2L$ operations in (5). Thus if $K$ is much smaller than $N$, we get a significant cost reduction. Of course, this complexity analysis implicitly assumes that the generalized Fourier coefficients $\alpha_k(x)$ and $\beta_k(x)$ are known in closed form or can be cheaply computed. For more details in this respect see Remark 1 below.

**Remark 1** *Many well known McKean-Vlasov type models used in physics and engineering are constructed and formulated via certain Fourier type expansions of the respective drift and/or diffusion coefficients. For example, in the famous*

*Kuramoto-Shinomoto-Sakaguchi model (see e.g. [4], eq. (5.214)) or in the coupled Brownian phase oscillators (see [10]) the mean field potential is given by its Fourier series, which entails a similar expansion for the coefficient* $a(x,u)$ *(*$b(x,u)$ *is constant). Let us also mention a classical work of [3], where a known power series expansion for the coefficients of a nonlinear Fokker-Planck equation is assumed. From another point of view, since the basis* $(\varphi_k)$ *with the corresponding weight* $w$ *can, in principle, be chosen freely, it is natural to assume they can be chosen such that the coefficients* $\alpha_k(x)$ *and* $\beta_k(x)$ *can be computed in closed form. In this respect, let us give some further examples. If for any* $x$*,* $a(x,\cdot)$ *is a linear combination of functions of the form:*

$$q_1(u_1)\cdots q_d(u_d)$$

*where each* $q_i:\mathbb{R}\to\mathbb{R}$ *is a polynomial with coefficients possibly depending on* $x$*, then* $(\varphi_k)$ *may be taken to be Hermite functions in* $\mathbb{R}^d$*, i.e.*

$$\varphi_\alpha(u)=H_{\alpha_1}(u_1)\cdots H_{\alpha_d}(u_d)e^{-|u|^2/2},\quad \alpha=(\alpha_1,\ldots,\alpha_d).$$

*The latter situation appears for instance in the popular interaction case with* $a(x,u)=A(x-u)$*, where the function* $A$ *has a given representation*

$$A(z)=\sum_\alpha c_\alpha z_1^{\alpha_1}\cdot\ldots\cdot z_d^{\alpha_d},\quad z\in\mathbb{R}^d,\quad \alpha\in\mathbb{N}_0^d.$$

*As another example, note that the Fourier coefficients of any function of the form*

$$u\to u_1^{\alpha_1}\cdot\ldots\cdot u_d^{\alpha_d}e^{-\frac{|u-c|^2}{\sigma}},\quad c\in\mathbb{R}^d,\quad \alpha\in\mathbb{N}_0^d,\ \sigma>0,$$

*with respect to the Hermite basis above can be expressed in closed form. One so could also consider* $a(x,u)$*,* $b(x,u)$ *of the form*

$$\sum_{r=1}^R q_r(x)u^{\alpha_r}e^{-|u-c_r(x)|^2/\sigma_r(x)},$$

*with free to choose* $q_r(x)\in\mathbb{R}$*,* $c_r(x)\in\mathbb{R}^d$*,* $\sigma_r(x)\in\mathbb{R}_+$*,* $\alpha_r\in\mathbb{N}_0^d$*,* $R\in\mathbb{N}$*.*

# 3 Convergence analysis

In this section we first study the convergence of the approximated particle system (8) to the solution of the original system (1). As a first obvious but important observation, we note that the distribution of the triple $\left(X_s^{j,K,N},X_s^{K,N},X_s^j\right)$ with $X_s^{K,N}:=\left(X_s^{1,K,N},\ldots,X_s^{N,K,N}\right)$ does not depend on $j$, and therefore we can write

$$\left(X^{j,K,N},X^{K,N},X^j\right)\overset{\text{distr.}}{=}\left(X^{\cdot,K,N},X^{K,N},X^\cdot\right)\quad\text{for } j=1,...,N.\tag{12}$$

For ease of notation, henceforth we denote with $|\cdot|:=|\cdot|_{\text{dim}}$ for a generic dimension dim the standard Euclidian norm in $\mathbb{R}^{\text{dim}}$. Let us make the following assumptions.

**(AF)** The basis functions $(\varphi_k)$ fulfil

$$|\varphi_k(z) - \varphi_k(z')| \leq L_{k,\varphi} |z - z'|, \quad |\varphi_k(z)| \leq D_{k,\varphi}, \quad k = 0, 1, \ldots$$

for all $z, z' \in \mathbb{R}^d$ and some constants $L_{k,\varphi}, D_{k,\varphi} > 0$.

**(AC)** The functions $\alpha_k(x)$, $\beta_k(x)$, $k = 0, 1, 2, \ldots$ satisfy

$$|\alpha_k(x)| \leq A_{k,\alpha}(1 + |x|) \quad \text{with} \quad (A_{k,\alpha})_{k=0,1,\ldots} \in l_2,$$
$$\sum_{k=0}^{\infty} D_{k,\varphi} A_{k,\alpha} \leq D_\varphi A_\alpha, \quad \text{and} \quad \sum_{k=0}^{\infty} L_{k,\varphi} A_{k,\alpha} \leq L_\varphi A_\alpha,$$
$$|\beta_k(x)| \leq A_{k,\beta}(1 + |x|) \quad \text{with} \quad (A_{k,\beta})_{k=0,1,\ldots} \in l_2,$$
$$\sum_{k=0}^{\infty} D_{k,\varphi} A_{k,\beta} \leq D_\varphi A_\beta, \quad \text{and} \quad \sum_{k=0}^{\infty} L_{k,\varphi} A_{k,\beta} \leq L_\varphi A_\beta,$$

for some constants $A_\alpha$, $A_\beta$, $D_\varphi$, and $L_\varphi > 0$, and further

$$\sup_{x,x' \in \mathbb{R}^d,\ x \neq x'} \frac{|\alpha_k(x) - \alpha_k(x')|}{|x - x'|} \leq B_{k,\alpha} \quad \text{with} \quad \sum_{k=0}^{\infty} D_{k,\varphi} B_{k,\alpha} \leq D_\varphi B_\alpha,$$
$$\sup_{x,x' \in \mathbb{R}^d,\ x \neq x'} \frac{|\beta_k(x) - \beta_k(x')|}{|x - x'|} \leq B_{k,\beta} \quad \text{with} \quad \sum_{k=0}^{\infty} D_{k,\varphi} B_{k,\beta} \leq D_\varphi B_\beta,$$

for some $B_\alpha$, $B_\beta > 0$.

**(AM$_p$)** For some $p > 0$ the initial distribution $\mu_0$ possesses a finite absolute moment of order $p$.

In the sequel, for any random variable $\xi \in \mathbb{R}^{\dim}$ on $(\Omega, \mathcal{F}, \mathrm{P})$ we shall use $\|\xi\|_p$ for the norm of $|\xi|$ in $L_p(\Omega)$. The following bound on the strong error can be proved.

**Theorem 1** *For $p \geq 2$, it holds under assumptions (AC), (AF) and (AM$_p$) that*

$$\left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_p \lesssim N^{-1/2} + \sum_{k=K+1}^{\infty} A_{k,\alpha} \|\gamma_k\|_{L_p[0,T]}$$
$$+ \sum_{k=K+1}^{\infty} A_{k,\beta} \|\gamma_k\|_{L_p[0,T]}, \tag{13}$$

*where $\lesssim$ stands for an inequality with some (hidden) positive finite constant depending only on $A_\alpha, A_\beta, B_\alpha, B_\beta, D_\varphi, L_\varphi$, $p$, and $T$.*

**Remark 2** *For $1 \leq p' \leq 2$, we simply have*

$$\left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_{p'} \leq \left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_p \tag{14}$$

*for any $p \geq 2$.*

The next theorem, on the convergence of the Euler approximation (11) to the projected system (8), can be proved along the same lines as the proof of Theorem 1.

**Theorem 2** *For $p \geq 2$, it holds under assumptions (AC), (AF) and ($AM_p$) that for any natural $K, N$*

$$\left\| \sup_{0 \leq r \leq T} \left| \bar{X}_r^{\cdot,K,N} - X_r^{\cdot,K,N} \right| \right\|_p \lesssim \sqrt{h},$$

*where $\lesssim$ stands for an inequality with some (hidden) positive finite constant depending only on $A_\alpha, A_\beta, B_\alpha, B_\beta, D_\varphi, L_\varphi$, $p$ and $T$.*

**Discussion** The bound (13) is proved under rather general assumptions on the coefficients $a(x,y)$ and $b(x,y)$. In particular, we allow for linear growth of these coefficients in $x$. This makes the proof of the bound in Theorem 1 rather challenging, since we need to avoid an explosion. In order to overcome this problem, we employ a kind of averaging technique which, being combined with the symmetry of the particle distribution and the existence of moments (see Section 7.1), gives the desired bound. Note that for this we have to assume existence and uniqueness of a strong solution of the original MVSDE (1). Funaki [5] proved existence and uniqueness under (essentially) global Lipschitz condition. However, one should be able to extend his results by exploiting a kind of one sided Lipschitz condition like in [6] or [7].

The bound (13) consists of stochastic and approximation errors. While the first error is of order $1/\sqrt{N}$, the second one depends on $K$ and the properties of the coefficients $a(x,y)$ and $b(x,y)$. If these coefficients are smooth in the sense that their generalized Fourier coefficients $(\alpha_k)$ and $(\beta_k)$ decay fast, then the approximation error can be made small even for medium values of $K$.

**Example 1** *The (normalized) Hermite polynomial of order $j$ is given, for $j \geq 0$, by*

$$\overline{H}_j(x) = c_j (-1)^j e^{x^2} \frac{d^j}{dx^j}(e^{-x^2}), \quad c_j = \left(2^j j! \sqrt{\pi}\right)^{-1/2}.$$

*These polynomials satisfy: $\int_{\mathbb{R}} \overline{H}_j(x) \overline{H}_\ell(x) e^{-x^2} dx = \delta_{j,\ell}$ and, as a consequence,*

$$\varphi_k(u) = \overline{H}_k(u) e^{-u^2/2}, \quad k = 1, 2, \ldots, \tag{15}$$

*is a total orthonormal system in $L_2\left(\mathbb{R}^d\right)$ (i.e. here $w = 1$). Moreover, $(\varphi_k)_{k \geq 0}$ fulfil the assumption (AF) with $D_{k,\varphi}$ and $L_{k,\varphi}$ being uniformly bounded in $k$, see, e.g. [15], p. 242. Now let us suppose that $a(x,\cdot), b(x,\cdot) \in L_2(\mathbb{R}^d)$ for any $x \in \mathbb{R}$, and discuss the assumptions (AC).*

**Lemma 1** *Suppose that for any $x \in \mathbb{R}$, the functions (in $u$)*

$$\widetilde{a}(x,u) := \frac{a(x,u)}{\sqrt{1+x^2}}, \quad \widetilde{b}(x,u) := \frac{b(x,u)}{\sqrt{1+x^2}}$$

*admit derivatives in $u$ up to order $s > 2$ such that the functions (in $u$)*

$$u^\ell \partial_u^m \widetilde{a}(x,u), \quad u^\ell \partial_u^m \widetilde{b}(x,u), \quad 0 \leq l + m \leq s$$

*are bounded and belong to $L_1(\mathbb{R})$ (uniformly in $x$) together with their first derivatives in $x$. Then the assumption (AC) is satisfied and*

$$\left\| \sup_{0\le r\le T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_p \lesssim K^{1-s/2} + N^{-1/2}, \tag{16}$$

*as $K, N \to \infty$.*

**Proof.** We have

$$\alpha_k(x) = \sqrt{1+x^2} \int \widetilde{a}(x,u)\overline{H}_k(u)\, e^{-u^2/2}\, du = \sqrt{1+x^2}\, \widetilde{\alpha}_k(x),$$

$$\beta_k(x) = \sqrt{1+x^2} \int \widetilde{b}(x,u)\overline{H}_k(u)\, e^{-u^2/2}\, du = \sqrt{1+x^2}\, \widetilde{\beta}_k(x).$$

The identity

$$(2k+2)^{1/2}\overline{H}_k(x) = \overline{H}'_{k+1}(z)$$

and the integration-by-parts formula imply

$$\begin{aligned} \widetilde{\alpha}_k(x) = {} & \left. \frac{\widetilde{a}(x,u)e^{-u^2/2}\overline{H}_{k+1}(u)}{(2k+2)^{1/2}} \right|_{-\infty}^{\infty} \\ & - \frac{1}{(2k+2)^{1/2}} \int_{-\infty}^{\infty} \left[ \frac{\partial \widetilde{a}(x,u)}{\partial u} - u\widetilde{a}(x,u) \right] \overline{H}_{k+1}(u) e^{-u^2/2}\, du. \end{aligned}$$

Note that $|\overline{H}_k(u)|\, e^{-u^2/2} \le 1$ uniformly in $u$ and $k$ (see, e.g. [15], p. 242.) Hence if $\widetilde{a}(x,u)$ is bounded and

$$\int \left| \frac{\partial \widetilde{a}(x,u)}{\partial u} - u\widetilde{a}(x,u) \right| du$$

is bounded uniformly in $x$, then $\widetilde{\alpha}_k(x) = O\left(k^{-1/2}\right)$ uniformly in $x$. The second integration-by-parts shows that $\widetilde{\alpha}_k(x) = O\left(k^{-1}\right)$, provided the functions

$$u^2 \cdot \widetilde{a}(x,u), \frac{\partial^2 \widetilde{a}(x,u)}{\partial u^2}, u \cdot \frac{\partial \widetilde{a}(x,u)}{\partial u}$$

are integrable on $\mathbb{R}$ with their $L_1(\mathbb{R})$ norms uniformly bounded in $x$. Integrating by parts further, we derive the desired statement. ■

**Remark 3** *As a rule, one chooses $N$ and $K$ such that the errors in (16) are balanced, that is $N^{1/(s-2)} \sim K$, yielding a proportional reduction of computational cost of order $N \cdot K/N^2 \sim N^{-(s-3)/(s-2)}$. Alternatively we can compare the complexity, that is the computational cost for achieving a prescribed accuracy $\varepsilon$, for of the Euler schemes (5) and (11). It is not difficult to see that, after incorporating the path-wise time discretization error, the standard Euler scheme (5) has complexity of order $\varepsilon^{-6}$, while the projected one (11) has complexity of order $\varepsilon^{-(4s-6)/(s-2)}$ which is significantly smaller when $s > 3$. Moreover, in [1] conditions are formulated, guaranteeing that all measures $\mu_t$, $t \ge 0$, possess infinitely smooth exponentially decaying densities. In this case we can additionally profit from the decay of the generalized Fourier coefficients $(\gamma_k)$ such that the convergence rates in* (13) *give rise to a proportional reduction of computational cost approaching $N^{-1}$, corresponding to a complexity of order $\varepsilon^{-4}$ (modulo some logarithmic term) for the method (11).*

## 3.1 Density estimation

Let us now discuss the estimation of the densities $\mu_t$, $t \geq 0$. Let us assume that the formal relationship (10) holds in the sense that

$$\frac{\mu_s}{w} = \sum_{k=0}^{\infty} \gamma_k(s)\varphi_k$$

in $L_2\left(\mathbb{R}^d, w\right)$, i.e. $\mu_s^2/w \in L_1\left(\mathbb{R}^d\right)$. Fix some $t > 0$, $K_{\text{test}} \in \mathbb{N}$ and set

$$\widehat{\mu}_t^{K_{\text{test}},K,N}(x) := \sum_{k=1}^{K_{\text{test}}} \gamma_k^N(t)\varphi_k(x)w(x)$$

with $\gamma_k^N(t) := \frac{1}{N}\sum_{i=1}^N \varphi_k(X_t^{i,K,N})$, $k = 1, \ldots, K_{\text{test}}$. We obviously have

$$\begin{aligned}\mathrm{E}\int |\widehat{\mu}_t^{K_{\text{test}},K,N}(x) - \mu_t(x)|^2 w^{-1}(x)\,dx = &\sum_{k=1}^{K_{\text{test}}} \mathrm{E}\left[|\gamma_k^N(t) - \gamma_k(t)|^2\right] \\ &+ \sum_{k=K_{\text{test}}+1}^{\infty} |\gamma_k(t)|^2,\end{aligned}$$

where (due to (AF))

$$\begin{aligned}\mathrm{E}\left[|\gamma_k^N(t) - \gamma_k(t)|^2\right] &= \mathrm{E}\left[\left|\frac{1}{N}\sum_{j=1}^N \varphi_k(X_t^{j,K,N}) - \mathrm{E}\left[\varphi_k(X_t^{\cdot})\right]\right|^2\right] \\ &\leq 2\mathrm{E}\left[\left|\frac{1}{N}\sum_{j=1}^N \left(\varphi_k(X_t^{j,K,N}) - \varphi_k(X_t^j)\right)\right|^2\right] \\ &+ 2\mathrm{E}\left[\left|\frac{1}{N}\sum_{j=1}^N \left(\varphi_k(X_t^j) - \mathrm{E}\left[\varphi_k(X_t^j)\right]\right)\right|^2\right] \\ &\leq 2L_{k,\varphi}^2 \mathrm{E}\left[\left|X_t^{\cdot,K,N} - X_t^{\cdot}\right|^2\right] + \frac{2}{N}\mathrm{Var}\left[\varphi_k(X_t)\right],\end{aligned}$$

since the $X^j$ are independent. Theorem 1 now implies

$$\begin{aligned}\left(\mathrm{E}\int |\widehat{\mu}_t^{K_{\text{test}},K,N}(x) - \mu_t(x)|^2\, w^{-1}(x)\,dx\right)^{1/2} &\lesssim \left(\frac{1}{N}\sum_{k=1}^{K_{\text{test}}} (L_{k,\varphi}^2 + D_{k,\varphi}^2)\right)^{1/2} \\ + \left(\sum_{k=1}^{K_{\text{test}}} L_{k,\varphi}^2\right)^{1/2} &\left(\sum_{k=K+1}^{\infty} (A_{k,\alpha} + A_{k,\beta})\,\|\gamma_k\|_{L_p[0,T]}\right) \\ &+ \left(\sum_{k=K_{\text{test}}+1}^{\infty} |\gamma_k(t)|^2\right)^{1/2}. \quad (17)\end{aligned}$$

The last term always converges to zero as $K_{\text{test}} \to \infty$, since $\mu_t/w \in L_2(\mathbb{R}^d, w)$. The first term can be controlled for any fixed $K_{\text{test}}$ by taking $N$ large enough. Finally, for any fixed $K_{\text{test}}$, the second term can be made small by taking $K$ large enough and using the condition (AC).

# 4 Specific models

## 4.1 Generalized Shimizu-Yamada Models

Inspired by the work of Shimizu and Yamada [14], [17] and [11], we consider one-dimensional McKean-Vlasov equations of the form (1) with

$$a(x,u) := a^0(u) + a^1(u)x, \quad b(x,u) := b(u).$$

This class of models allows for a linear dependence of drift on the distribution of $X$ through $\mathrm{E}\left[a^0(X_t)\right]$ and $\mathrm{E}\left[a^1(X_t)\right]$. Let us define for polynomially bounded and measurable functions $a^j$ and $b$ the generalized Gauss transforms,

$$H_{a^j}(p,q) := \frac{1}{\sqrt{2\pi q}} \int a^j(u) e^{-\frac{(p-u)^2}{2q}} du, \quad j = 0,1,$$
$$H_b(p,q) := \frac{1}{\sqrt{2\pi q}} \int b(u) e^{-\frac{(p-u)^2}{2q}} du, \quad p \in \mathbb{R}, \quad q > 0.$$

Let moreover $a^j$ and $b$ be such that the partial derivatives,

$$\partial_p H_{a^j}(p,q),\ \partial_q H_{a^j}(p,q) \quad j = 0,1, \quad \text{and} \quad \partial_p H_b(p,q),\ \partial_q H_b(p,q),$$
$$\text{extend continuously to any } (p,q) \in \mathbb{R} \times \mathbb{R}_{\geq 0}. \tag{18}$$

It is not difficult to see that (18) holds if $a^j$ and $b$ are entire functions for which the coefficients of their power series around $u = 0$ decay fast enough to zero (which is trivially satisfied for any polynomial). A complete characterization of $a^j$ and $b$ such that (18) holds, is connected with analytic vectors for semigroups related to the heat kernel and considered beyond the scope of this paper however.

**Theorem 3** *Let $a^j$ and $b$ satisfy (18). (i) Then the following system of ODEs*

$$G'_t = H_b^2(A_t, G_t) + 2H_{a^1}(A_t, G_t)\, G_t \tag{19}$$
$$A'_t = H_{a^0}(A_t, G_t) + H_{a^1}(A_t, G_t)\, A_t, \quad (A_0, G_0) = (x_0, 0),$$

*has for $0 \leq t < t_\infty \leq \infty$, i.e. up to some possibly finite exploding time $t_\infty$, a unique solution $(A_t, G_t) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$. (ii) The McKean-Vlasov SDE*

$$dX_t = \left(\mathrm{E}\left[a^0(X_t)\right] + X_t\, \mathrm{E}\left[a^1(X_t)\right]\right) dt + \mathrm{E}\left[b(X_t)\right] dW_t, \quad X_0 = x_0 \tag{20}$$

*is then equivalent to*

$$dX_t = \left(H_{a^0}(A_t, G_t) + H_{a^1}(A_t, G_t)\, X_t\right) dt + H_b(A_t, G_t)\, dW_t, \quad X_0 = x_0, \tag{21}$$

*and has explicit solution,*

$$X_t = x_0 e^{\int_0^t H_{a^1}(A_s, G_s) ds} + \int_0^t H_{a^0}(A_s, G_s)\, e^{\int_s^t H_{a^1}(A_r, G_r) dr} ds \tag{22}$$
$$+ \int_0^t H_b(A_s, G_s)\, e^{\int_s^t H_{a^1}(A_r, G_r) dr} dW_s, \quad 0 \leq t < t_\infty \leq \infty.$$

*Note: the Wiener integral in (22) can be interpreted by an ordinary integral after partial integration, due to the smoothness of the (deterministic) integrand.*

## 4.2 Affine structures

Let us consider affine functions

$$
\begin{aligned}
a^0(u) &= a_0^0 + a_1^0 u, \\
a^1(u) &= a_0^1 + a_1^1 u, \\
b(u) &= b_0 + b_1 u.
\end{aligned}
$$

Then for $c \equiv a^0$, $c \equiv a^1$, and $c \equiv b$, respectively, we have

$$
\begin{aligned}
H_c(p,q) &= \frac{1}{\sqrt{2\pi q}} \int c(u) e^{-\frac{(p-u)^2}{2q}} du \\
&= \frac{1}{\sqrt{2\pi q}} \int c_0 e^{-\frac{(p-u)^2}{2q}} du + \frac{1}{\sqrt{2\pi q}} \int c_1 u e^{-\frac{(p-u)^2}{2q}} du \\
&= c_0 + c_1 p
\end{aligned}
$$

with $c(u) = c_0 + c_1 u$. In particular, the $H_c(p,q)$ do not depend on $q$, and so (19) simplifies to

$$
A_t' = a_0^0 + \left(a_1^0 + a_0^1\right) A_t + a_1^1 A_t^2, \quad A_0 = x_0. \tag{23}
$$

We first consider the case $a_1^1 = 0$, then (23) reads $A_t' = a_0^0 + \left(a_1^0 + a_0^1\right) A_t$ with solution

$$
\begin{aligned}
A_t &= \left(x_0 + \frac{a_0^0}{a_1^0 + a_0^1}\right) e^{\left(a_1^0 + a_0^1\right)t} - \frac{a_0^0}{a_1^0 + a_0^1} \quad \text{if} \quad a_1^0 + a_0^1 \neq 0, \quad \text{and} \qquad (24) \\
A_t &= x_0 + a_0^0 t \quad \text{if} \quad a_1^0 + a_0^1 = 0.
\end{aligned}
$$

For the case $a_1^1 \neq 0$ the solution (checked by Mathematica) is as follows. If $D := \left(a_1^0 + a_0^1\right)^2 - 4a_0^0 a_1^1 < 0$, $a_1^1 \neq 0$,

$$
A_t = -\frac{\left(a_1^0 + a_0^1\right)}{2a_1^1} + \frac{\sqrt{-D}}{2a_1^1} \tan\left[\frac{1}{2}\sqrt{-D}t + \arctan\left[\frac{a_1^0 + a_0^1 + 2a_1^1 x_0}{\sqrt{-D}}\right]\right]. \tag{25}
$$

If $D > 0$, $a_1^1 \neq 0$,

$$
A_t = \frac{1}{2}\left(\sqrt{D} - a_1^0 - a_0^1\right)/a_1^1 + \frac{x_0 - \frac{1}{2}\left(\sqrt{D} - a_1^0 - a_0^1\right)/a_1^1}{1 + \frac{1}{2}\left(\sqrt{D} + a_1^0 + a_0^1 + 2a_1^1 x_0\right)\left(e^{-\sqrt{D}t} - 1\right)/\sqrt{D}}. \tag{26}
$$

If $D = 0$, $a_1^1 \neq 0$,

$$
A_t = -\frac{a_1^0 + a_0^1}{2a_1^1} + \frac{1}{a_1^1} \frac{a_1^0 + a_0^1 + 2a_1^1 x_0}{2 - \left(a_1^0 + a_0^1 + 2a_1^1 x_0\right) t}. \tag{27}
$$

As a result, the McKean-Vlasov SDE

$$
dX_t = (a_0^0 + a_1^0 A_t + \left(a_0^1 + a_1^1 A_t\right) X_t)dt + \left(b_0 + b_1 A_t\right) dW_t
$$

has the following (unique) solution

$$
\begin{aligned}
X_t &= x_0 e^{\int_0^t \left(a_0^1 + a_1^1 A_s\right) ds} + \int_0^t \left(a_0^0 + a_1^0 A_s\right) e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} ds \\
&\quad + \int_0^t \left(b_0 + b_1 A_s\right) e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} dW_s,
\end{aligned} \tag{28}
$$

where $A_t$ is given by (24), (25), (26), or (27).

**Example 2** *By taking in Section 4.2*

$$a(x,u) = a_1^0 u + a_0^1 x, \quad b(x,u) = b_0, \quad a_1^0 + a_0^1 < 0,$$

*we get essentially the Shimizu-Yamada model. From (24) we then have*

$$A_t = x_0 e^{\left(a_1^0 + a_0^1\right)t},$$

*and from (28) we then get the explicit solution*

$$X_t = x_0 e^{\left(a_1^0 + a_0^1\right)t} + \int_0^t b_0 e^{a_0^1 (t-s)} dW_s$$

*which is Gaussian with mean* $x_0 e^{\left(a_1^0 + a_0^1\right)t}$ *and variance* $b_0^2 \frac{e^{2a_0^1 t} - 1}{2a_0^1}$*, and which is consistent with the terminology in ([4], Section 3.10), where* $a_1^0 + a_0^1 = -\gamma$ *and* $a_0^1 = -\gamma - \kappa$.

**Example 3** *By taking in Section 4.2*

$$a(x,u) = \left(a_0^1 + a_1^1 u\right) x, \quad b(x,u) = b_0,$$

*we straightforwardly get from (26),*

$$A_t = \frac{x_0 e^{a_0^1 t}}{1 - \frac{a_1^1}{a_0^1} x_0 \left(e^{a_0^1 t} - 1\right)}, \tag{29}$$

*and*

$$X_t = x_0 e^{\int_0^t \left(a_0^1 + a_1^1 A_s\right) ds} + \int_0^t b_0 e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} dW_s, \tag{30}$$

*respectively. Plugging (29) into (30) then yields*

$$X_t = \frac{x_0 e^{a_0^1 t}}{1 - \frac{a_1^1}{a_0^1} x_0 \left(e^{a_0^1 t} - 1\right)} + \frac{b_0 e^{a_0^1 t}}{1 - \frac{a_1^1}{a_0^1} x_0 \left(e^{a_0^1 t} - 1\right)} \Gamma_t$$

*with Gaussian* $\Gamma_t = \int_0^t \left(1 - \frac{a_1^1}{a_0^1} x_0 \left(e^{a_0^1 s} - 1\right)\right) e^{-a_0^1 s} dW_s$*. In particular, if* $a_0^1 = 0$ *we get*

$$A_t = \frac{x_0}{1 - a_1^1 x_0 t},$$

*and*

$$X_t = \frac{x_0}{1 - a_1^1 x_0 t} + b_0 \int_0^t \frac{1 - a_1^1 x_0 s}{1 - a_1^1 x_0 t} dW_s.$$

**Remark 4** *From Example 3 it is clear that if* $a_1^1 \neq 0$*, the affine McKean-Vlasov solution may explode in finite time. This is not surprising since in this case the derivative* $\partial_u a(x,u)$ *is unbounded and so the main results in [1] do not apply. On the other hand, it is easy to check that for the case* $a_1^1 \neq 0$*, the affine solutions in Section 4.2 are non-exploding whenever,*

$$D \geq 0 \text{ and } \sqrt{D} \geq a_1^0 + a_0^1 + 2a_1^1 x_0.$$

*That is, in the case* $D \geq 0$*,* $a_1^1 \neq 0$*, it is always possible to choose* $x_0$ *such that the solution does or does not explode.*

### 4.3 Kuramoto-Shinomoto-Sakaguchi type models

In the Kuramoto-Shinomoto-Sakaguchi model the nonlinear one-dimensional Fokker-Planck equation (3) is considered in the domain $(t,x) \in (0,\infty) \times (0,2\pi)$, where $b=1$, $a(x,y) = a(x-y) = -\frac{d}{dx} U_{MF}(x-y)$ with

$$U_{MF}(z) = -\sum_{n=1}^{\infty} c_n \cos(nz)$$

and the process starts in $u$ at time zero, for some fixed $u \in (0,2\pi)$, see for details [4] (Sect. 5.3.2). Thus $a$ is a $2\pi$-periodic function related to a $2\pi$-periodic potential. Let us consider the corresponding McKean-Vlasov SDE

$$\begin{cases} X_t &= u + \int_0^t \int_{\mathbb{R}} a(X_s - y)\mu_s(dy)ds + W_t \\ \mu_t &= \mathrm{Law}(X_t), \quad t \in [0,T], \end{cases} \tag{31}$$

and define the integer valued function $k(x) := \max\{j \in \mathbb{Z} : 2\pi j \le x\}$. Obviously, the process

$$Y_t := X_t - 2\pi k(X_t) \tag{32}$$

has state space $[0,2\pi)$. Let $\rho_t(x;u) = \rho_t(x)$ be the density of $Y_t$, which is concentrated on $(0,2\pi)$. Note that for any $2\pi$-periodic function $f$ we have by (32) that

$$\int_0^{2\pi} f(x)\rho_t(x)dx = \mathrm{E}\left[f(Y_t)\right] = \mathrm{E}\left[f(X_t)\right] = \int_{-\infty}^{\infty} f(x)\mu_t(x)dx,$$

and for any test function $g$ with support in $(0,2\pi)$ it holds that,

$$\begin{aligned} \int_0^{2\pi} g(x)\rho_t(x)dx &= \mathrm{E}\left[g(Y_t)\right] = \mathrm{E}\left[g(X_t - 2\pi k(X_t))\right] \\ &= \sum_{j\in\mathbb{Z}} \int_{2\pi j}^{2\pi(j+1)} g(x - 2\pi j)\mu_t(x)dx \\ &= \int_0^{2\pi} g(z) \sum_{j\in\mathbb{Z}} \mu_t(z + 2\pi j)dz, \end{aligned}$$

that is

$$\rho_t(z) = \sum_{j\in\mathbb{Z}} \mu_t(z + 2\pi j) \tag{33}$$

for $z \in (0,2\pi)$. Thus, in particular,

$$\int_{\mathbb{R}} a(x-y)\mu_t(y)dy = \int_0^{2\pi} a(x-y)\rho_t(y)dy. \tag{34}$$

an (31) is equivalent to

$$\begin{aligned} X_t &= u + \int_0^t \int_0^{2\pi} a(X_s - y)\rho_s(y)dyds + W_t \\ \rho_t &= \mathrm{Law}(Y_t), \quad t \in [0,T], \end{aligned}$$

(see (32)). Note that by using (33) and (34) it straightforwardly follows that $\rho_t(x) = \rho_t(x;u)$ satisfies (3) in the above context. Instead of taking the scalar product in $L_2(\mathbb{R}^d, w)$, we now consider the scalar product in $L_2([0,2\pi))$, i.e. $w \equiv 1$, and take for $(\varphi_k)$ the standard (total) orthonormal trigonometric basis consisting of the $2\pi$-periodic functions $(2\pi)^{-1/2}$, $\pi^{-1/2}\cos(my)$ and $\pi^{-1/2}\sin(my)$, $m = 1, 2, \ldots$ suitably ordered. Thus, by defining

$$\gamma_k(t) = \int_0^{2\pi} \rho_t(y)\varphi_k(y)\,dy, \quad \alpha_k(x) = \int_0^{2\pi} a(x-y)\varphi_k(y)\,dy,$$

one has

$$\rho_t(y) = \sum_{k=0}^{\infty} \gamma_k(t)\varphi_k(y) \quad \text{and} \quad \int_{\mathbb{R}} a(x-y)\mu_t(y)dy = \sum_{k=0}^{\infty} \alpha_k(x)\gamma_k(t),$$

due to (34). That is (8) reads,

$$X_t^{i,K,N} = u + \int_0^t \sum_{k=0}^{K} \gamma_k^N(s)\alpha_k(X_s^{i,K,N})\,ds + W_t^i$$

with $\gamma_k^N$ as in (9). Next we may follow (11) for the corresponding Euler scheme. Finally, the estimator for the density $\rho_t$ reads

$$\widehat{\rho}_t^{K_{\text{test}},K,N}(y) := \sum_{k=1}^{K_{\text{test}}} \gamma_k^N(t)\varphi_k(y) \tag{35}$$

(cf. the estimator for $\mu_t$ in Section 3.1).

# 5 Numerical test cases

## 5.1 Affine MVSDE models

Let us now test the numerical performance of the projected particle approach for the processes discussed in Section 4.1. Consider the situation where

$$a^0(u) = (1+u^M)\exp(-u^2/2), \quad a^1(u) = \rho\exp(-u^2/2), \quad b(x,u) \equiv \sigma$$

for some $M > 0$, $\rho \geq 0$ and $\sigma > 0$. Then, by using the Hermite functions (15) with $w \equiv 1$ and the well-known identity

$$u^M = \frac{1}{2^M} \sum_{m=0}^{\lfloor \frac{M}{2} \rfloor} \frac{M!}{m!(M-2m)!} H_{M-2m}(u),$$

we derive straightforwardly,

$$\begin{aligned}\alpha_k(x) &= \int (1+u^M)\exp(-u^2)\overline{H}_k(u)\,du + \rho x \cdot \int \exp(-u^2)\overline{H}_k(u)\,du \\ &= \begin{cases} 0 & \text{if} \quad k > M \text{ or } k \text{ is uneven} \\ \frac{\pi^{1/4}}{2^{M-k/2}} \frac{M!}{\left(\frac{M-k}{2}\right)!\sqrt{k!}} & \text{if} \quad 0 \leq k \leq M, \quad \text{s.t.} \quad M-k \quad \text{is even} \end{cases} \\ &+ (1+\rho x) \cdot \begin{cases} \pi^{1/4} & \text{if} \quad k = 0, \\ 0 & \text{if} \quad k > 0. \end{cases}\end{aligned}$$

On the other hand, by some algebra we get

$$H_{a^0}(p,q) = \frac{1}{\sqrt{2\pi q}} \int (1+u^M) e^{-u^2/2} e^{-\frac{(p-u)^2}{2q}}\, du = \frac{1}{\sqrt{1+q}} e^{-\frac{p^2}{2(1+q)}}$$
$$+ \frac{1}{\sqrt{2\pi(1+q)}} e^{-\frac{p^2}{2(1+q)}} \int \left( \sqrt{\frac{q}{1+q}} y + \frac{p}{1+q} \right)^M e^{-y^2/2} dy,$$

and

$$H_{a^1}(p,q) = \frac{\rho}{\sqrt{1+q}} e^{-\frac{p^2}{2(1+q)}}.$$

The explicit solution of the MVSDE

$$dX_t = \left( \mathsf{E}\left[(1+X_t^M)\exp(-X_t^2/2)\right] + \rho X_t \mathsf{E}\left[\exp(-X_t^2/2)\right] \right) dt + \sigma dW_t \quad (36)$$

is given by (19) and (22). Hence the density of $X_t$ is normal with mean

$$x_0 e^{\int_0^t H_{a^1}(A_s,G_s)ds} + \int_0^t H_{a^0}(A_s,G_s)\, e^{\int_s^t H_{a^1}(A_r,G_r)dr}\, ds$$

and variance

$$\sigma^2 \int_0^t e^{2\int_s^t H_{a^1}(A_r,G_r)dr}\, ds.$$

In our numerical example we take $M = 2$, $\rho = -1$, $\sigma = 1$ and $x_0 = 0$. Our aim is to approximate the normal density of $X_1$ by using our projected particle method based on Hermite basis. To this end, we first simulate $N$ paths of the process $\bar{X}^{i,K,N}$, defined in (11) with a time step $h = 0.02$. Since $M = 2$, the case $K = 2$ corresponds to a perfect approximation of the integral $\int_{\mathbb{R}^d} a(x,y)\mu_t(y)\, dy = \sum_{k=0}^{2} \alpha_k(x)\gamma_k(t)$. Next using the obtained sample $\bar{X}_1^{1,K,N}, \ldots, \bar{X}_1^{N,K,N}$, we construct projection estimates for the density of $X_1$ by using Hermite basis functions of order $K_{\text{test}} \in \{1, 2, \ldots, 10\}$. The mean (0.727) and the variance (0.487) of the true normal density are approximated by solving the ODE system (19) using Euler method with time step 0.0001. The Figure 1 shows the box plots of $L_2$-distance between $\mu_1$ and $\widehat{\mu}_t^{K_{\text{test}},K,N}$ for $K \in \{1, 2\}$ based on 50 different replications of the process $\bar{X}^{i,K,N}$. As can be seen, the choice of $K_{\text{test}}$ is crucial and depends on $K$ and $N$. It also should be stressed that the truncation error dominates the statistical one already for medium sample sizes. An optimal balance between $K$, $N$ and $K_{\text{test}}$ can be found by analyzing the right hand side of (17) under various assumptions on the coefficients $(A_{k,\alpha})$, $(A_{k,\beta})$ and $(\gamma_k)$.

## 5.2 Convolution-type MVSDE models

Consider the MVSDE of the form:

$$dX_t = \mathrm{E}_{X'}\left[Q(X_t - X_t')\right] dt + \sigma\, dW_t, \quad t \in [0,1], \quad X_0 \sim \mathcal{N}(0,1),$$

i.e. of the form (1) with $a(x,y) = Q(x-y)$, $b(x,y) = \sigma$ and $\mu_0(x) = (1/\sqrt{2\pi})e^{-x^2/2}$. Let us again use the Hermite basis to approximate the density of $X_t$ for any

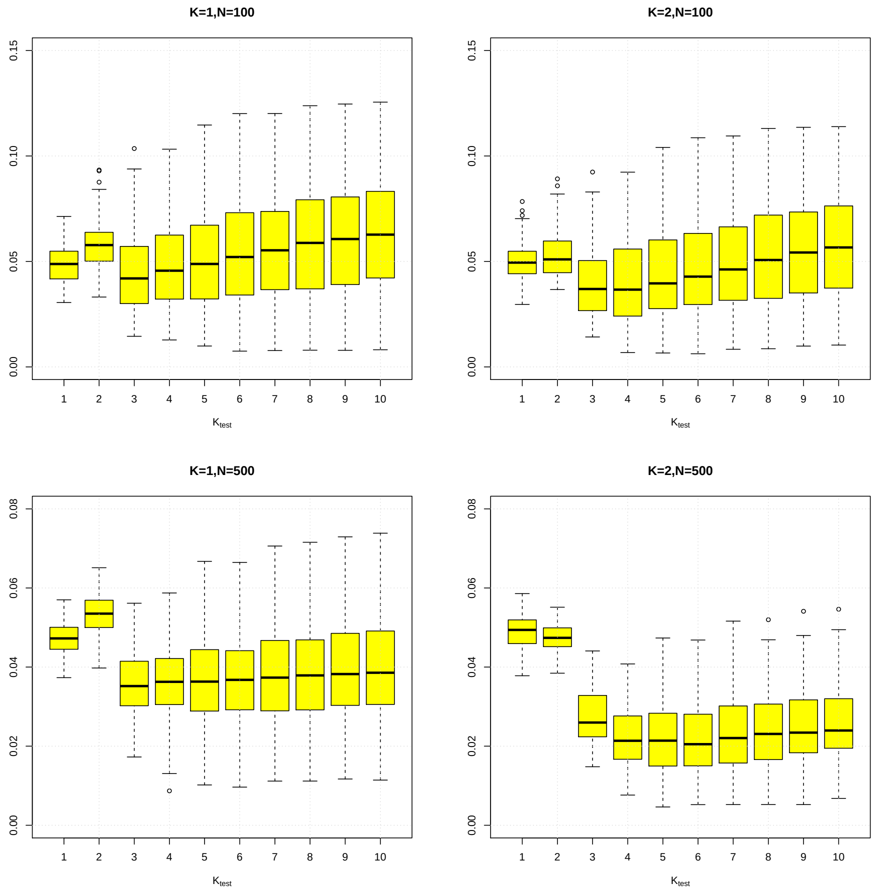


Figure 1: Box plots of $L_2$-distances between the true normal density of $X_1$ with mean 0.727 and variance 0.487 in the model (36) and its estimates obtained by using $N \in \{100, 500\}$ paths, Hermite basis functions up to order $K \in \{1, 2\}$ to approximate coefficients $\alpha_k$ and Hermite basis functions up to order $K_{\text{test}} \in \{1, 2, \ldots, 10\}$ for projection based density estimation.

$t \in [0,1]$. In the case $Q(x) = e^{-x^2/2}$, we explicitly derive via repeated integration by parts

$$\int_{\mathbb{R}} e^{-(x-y)^2/2-x^2/2} H_n(x)\, dx = \frac{e^{-y^2/4}}{2} \int e^{-(z-y)^2/4} H_n(z/2)\, dz$$
$$= \sqrt{\pi}\, \frac{e^{-y^2/4}}{2} \left(\frac{1}{2}\right)^{n-1} (2y)^n.$$

As a result

$$\alpha_n(y) = \int e^{-(x-y)^2/2-x^2/2} \overline{H}_n(x)\, dx = \pi^{1/4} \left(\frac{1}{2}\right)^{n/2} \frac{y^n}{\sqrt{n!}} e^{-y^2/4},$$

where $\overline{H}_n$ stands for the normalized Hermite polynomial of order $n$. We take $\sigma = 0.1$. Using the Euler scheme (5) with time step $h = 1/L = 0.01$, we first simulate $N = 500$ paths of the time discretized process $\bar{X}^{\cdot,N}$. Next, by means of the closed form expressions for $\alpha_n$, we generate $N$ paths of the projected approximating process $\bar{X}^{\cdot,K,N}$, $K \in \{1, \ldots, 20\}$ (see (11)), using the same Wiener increments as for $\bar{X}^{\cdot,N}$, so that the approximations $\bar{X}^{\cdot,N}$ and $\bar{X}^{\cdot,K,N}$ are coupled. Finally, we compute the strong approximation error

$$E_{N,K} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\bar{X}_1^{i,K,N} - \bar{X}_1^{i,N}\right)^2}$$

of the projective system relative to the system (5) and record times needed to compute approximations $\bar{X}_1^{\cdot,N}$ and $\bar{X}_1^{\cdot,K,N}$, respectively. Figure 2 shows the (natural) logarithm of $E_{N,K}$ versus the logarithm of the corresponding (relative) computational time gain defined as (comp. time due to (5) $-$ comp. time due to (11))/comp. time due to (5), for values $K \in \{1, \ldots, 20\}$. As can be seen, the relation between logarithmic strong error and logarithmic computational time gain can be well approximated by a linear function. On the right-hand side of Figure 2 we depict the projection estimate for the density of $X_1$ corresponding to $K = 20$. Note that we compare two particle systems (projected and non projected ones) for a fixed $N$ and are mainly interested in the dependence of their strong distance on $K$. In fact, the choice of $N$ doesn't have much influence on $E_{N,K}$, provided $N$ is large enough.

# 6 Proofs

## 6.1 Proof of Theorem 1

Let us introduce

$$\mathfrak{a}_{K,N}(x,y) := \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} \alpha_k(x) \varphi_k(y^j) = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} \varphi_k(y^j) \int a(x,u) \varphi_k(u) w(u) du,$$

$$\mathfrak{b}_{K,N}(x,y) := \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} \beta_k(x) \varphi_k(y^j) = \frac{1}{N} \sum_{j=1}^{N} \sum_{k=1}^{K} \varphi_k(y^j) \int b(x,u) \varphi_k(u) w(u) du,$$

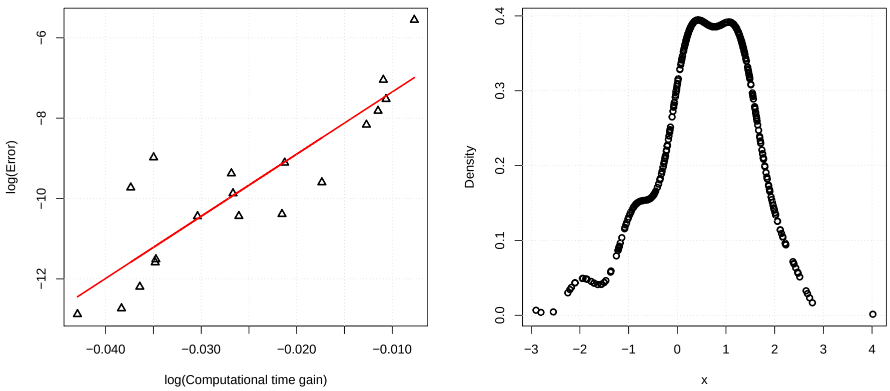


Figure 2: Left: Strong error $E_{500,K}$ between the solution of projected (see (11)) and non-projected (see (5)) time-discretized particle systems versus the difference (gain) in computational time. Right: Estimated density of $X_1$ using 21 basis functions.

and

$$\mathfrak{a}_s(x) := \int_{\mathbb{R}^d} a(x,u)\mu_s(du)ds,$$
$$\mathfrak{b}_s(x) := \int_{\mathbb{R}^d} b(x,u)\mu_s(du)ds$$

for any $x \in \mathbb{R}^d$, $y \in \mathbb{R}^{d\times N}$. We so have that

$$\Delta_t^i := X_t^{i,K,N} - X_t^i = \int_0^t \left(\mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{a}_s(X_s^i)\right) ds + \int_0^t \left(\mathfrak{b}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{b}_s(X_s^i)\right) dW_s^i,$$

where $W^i$, $i = 1, ..., N$, are i.i.d. copies of the $m$-dimensional Wiener process $W$. Hence,

$$\left|\Delta_t^i\right|^p \leq 2^{p-1}t^{p-1} \int_0^t \left|\mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{a}_s(X_s^i)\right|^p ds + 2^{p-1}d^{p-1}\sum_{q=1}^{d}\left|\int_0^t \left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right) dW_s^i\right|^p, \tag{37}$$

where

$$\mathfrak{b}_{K,N}^q := \left(\mathfrak{b}_{K,N}^{q,1}, ..., \mathfrak{b}_{K,N}^{q,m}\right), \quad q = 1, ..., d, \tag{38}$$

denote the rows of the $\mathbb{R}^{d\times m}$ valued $\mathfrak{b}_{K,N}$, and so we have with

$$\overline{\Delta_t^p} := \frac{1}{N}\sum_{i=1}^{N} \sup_{s\in[0,t]} \left|\Delta_s^i\right|^p$$

the bound

$$\overline{\Delta_t^p} \le 2^{p-1} t^{p-1} \frac{1}{N} \sum_{i=1}^N \int_0^t \left| \mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{a}_s(X_s^i) \right|^p ds$$
$$+ 2^{p-1} d^{p-1} \sum_{q=1}^d \frac{1}{N} \sum_{i=1}^N \sup_{s\in[0,t]} \left| \int_0^s \left( \mathfrak{b}_{K,N}^q(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i) \right) dW_s^i \right|^p$$
$$=: 2^{p-1} t^{p-1} \,\mathrm{Term}_1 + 2^{p-1} d^{p-1} \,\mathrm{Term}_2. \tag{39}$$

Assumption (AC) implies

$$\begin{aligned} |\mathfrak{a}_{K,N}(x,y) - \mathfrak{a}_{K,N}(x',y')| &= \left| \frac{1}{N} \sum_{j=1}^N \sum_{k=1}^K \left( \alpha_k(x)\varphi_k(y_j) - \alpha_k(x')\varphi_k(y_j') \right) \right| \\ &\le \frac{1}{N} \sum_{j=1}^N \sum_{k=1}^K |\alpha_k(x) - \alpha_k(x')| \left|\varphi_k(y_j')\right| \\ &+ \frac{1}{N} \sum_{j=1}^N \sum_{k=1}^K |\alpha_k(x)| \left|\varphi_k(y_j) - \varphi_k(y_j')\right| \\ &\le |x - x'| \, D_\varphi B_\alpha + \frac{L_\varphi A_\alpha}{N} (1 + |x|) \sum_{j=1}^N \left| y_j - y_j' \right|. \end{aligned} \tag{40}$$

Hence

$$\begin{aligned} |\mathfrak{a}_{K,N}(x,y) - \mathfrak{a}_{K,N}(x',y')|^p &\le 2^{p-1} \left| x - x' \right|^p D_\varphi^p B_\alpha^p \\ &+ 2^{p-1} L_\varphi^p A_\alpha^p (1 + |x|)^p \frac{1}{N} \sum_{j=1}^N \left| y_j - y_j' \right|^p. \end{aligned}$$

So it holds that

$$\begin{aligned} \left| \mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) \right|^p &\le 2^{p-1} D_\varphi^p B_\alpha^p |\Delta_s^i|^p \\ &+ 2^{p-1} L_\varphi^p A_\alpha^p (1 + \left| X_s^i \right|)^p \frac{1}{N} \sum_{j=1}^N |\Delta_s^j|^p, \end{aligned}$$

and then it follows that, with regard to $\mathrm{Term}_1$,

$$\begin{aligned} \mathrm{E}\left[\mathrm{Term}_1\right] &\le 2^{2p-2} D_\varphi^p B_\alpha^p \int_0^t \mathrm{E}\left[\overline{\Delta_s^p}\right] ds \\ &+ 2^{2p-2} L_\varphi^p A_\alpha^p \int_0^t \mathrm{E}\left[ \overline{\Delta_s^p} \cdot \frac{1}{N} \sum_{i=1}^N (1 + \left| X_s^i \right|)^p \right] ds \\ &+ 2^{p-1} \frac{1}{N} \sum_{i=1}^N \int_0^t \mathrm{E}\left[ \left| \mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i) \right|^p \right] ds. \end{aligned} \tag{41}$$

Let us now consider the middle term. Set

$$\zeta_{s,N} := \frac{1}{N} \sum_{i=1}^N (1 + \left| X_s^i \right|)^p - \frac{1}{N} \sum_{i=1}^N \mathrm{E}\left[ (1 + \left| X_s^i \right|)^p \right]$$

so that

$$\mathrm{E}\left[\overline{\Delta_s^p}\cdot\frac{1}{N}\sum_{i=1}^{N}(1+\left|X_s^i\right|)^p\right]=\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left[(1+\left|X_s^i\right|)^p\right]\cdot\mathrm{E}\left[\overline{\Delta_s^p}\right]$$
$$+\,\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\right]. \tag{42}$$

For arbitrary but fixed $\theta>0$, it holds that

$$\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\right]=\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\,1_{\{\zeta_{s,N}\le\theta\}}\right]+\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\,1_{\{\zeta_{s,N}>\theta\}}\right], \tag{43}$$

where on the one hand

$$\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\,1_{\{\zeta_{s,N}\le\theta\}}\right]\le\theta\mathrm{E}\left[\overline{\Delta_s^p}\right] \tag{44}$$

and on the other

$$\mathrm{E}\left[\zeta_{s,N}\cdot\overline{\Delta_s^p}\,1_{\{\zeta_{s,N}>\theta\}}\right]\le\sqrt{\mathrm{E}\left[\zeta_{s,N}^2 1_{\{\zeta_{s,N}>\theta\}}\right]}\sqrt{\mathrm{E}\left[\left(\overline{\Delta_s^p}\right)^2\right]}. \tag{45}$$

Due to (2) we have that for any $\eta>0$, there exists $C_{\theta,\eta}>0$ such that

$$\mathrm{E}\left[\zeta_{s,N}^2 1_{\{\zeta_{s,N}>\theta\}}\right]=\frac{1}{N}\mathrm{E}\left[\left(\sqrt{N}\zeta_{s,N}\right)^2 1_{\left\{\sqrt{N}\zeta_{s,N}>\theta\sqrt{N}\right\}}\right]\le\frac{C_{\theta,\eta}^2}{N^{\eta+1}},\qquad 0\le s\le T,$$

for $N$ large enough and

$$\begin{aligned}\mathrm{E}\left[\left(\overline{\Delta_s^p}\right)^2\right]&\le\mathrm{E}\left[\frac{1}{N}\sum_{j=1}^{N}\sup_{r\in[0,T]}\left|\Delta_r^j\right|^{2p}\right]=\mathrm{E}\left[\sup_{r\in[0,T]}\left|\Delta_r^{\cdot}\right|^{2p}\right]\\&=\mathrm{E}\left[\sup_{r\in[0,T]}\left|X_r^{\cdot,K,N}-X_r^{\cdot}\right|^{2p}\right]\\&\le 2^{2p-1}\mathrm{E}\left[\sup_{r\in[0,T]}\left|X_r^{\cdot,K,N}\right|^{2p}\right]+2^{2p-1}\mathrm{E}\left[\sup_{r\in[0,T]}\left|X_r^{\cdot}\right|^{2p}\right]\\&\le D_1+D_2=:D^2,\end{aligned} \tag{46}$$

where due to Theorem 4, Appendix 7,

$$2^{2p-1}\mathrm{E}\left[\sup_{r\in[0,T]}\left|X_r^{\cdot,K,N}\right|^{2p}\right]\le D_1<\infty\quad\text{uniform in } N \text{ and } K,$$

and

$$D_2:=2^{2p-1}\mathrm{E}\left[\sup_{r\in[0,T]}\left|X_r^{\cdot}\right|^{2p}\right]<\infty$$

due to (2). Thus, by combining (42)–(46), one has

$$\mathrm{E}\left[\overline{\Delta_s^p}\cdot\frac{1}{N}\sum_{i=1}^{N}(1+\left|X_s^i\right|)^p\right]\le F_1^p\cdot\mathrm{E}\left[\overline{\Delta_s^p}\right]+\frac{F_2}{N^{p/2+1/2}}$$

with $F_1 := \theta^{1/p} + \sup_{0\le s\le T} \|1+|X_s|\|_p$ and $F_2 := C_{\theta,p}D$, where we have taken $\eta = p$. Set now

$$H(s) := \mathrm{E}\left[\overline{\Delta_s^p}\right],$$

then the estimate (41) (cf. (39)) reads

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{a}_s(X_s^i)\right|^p\right] ds \\
&\le \left(2^{2p-2}D_\varphi^p B_\alpha^p + 2^{2p-2}L_\varphi^p A_\alpha^p F_1^p\right)\int_0^t H(s)ds + 2^{2p-2}L_\varphi^p A_\alpha^p \frac{F_2}{N^{p/2+1/2}}t \\
&\quad +2^{p-1}\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i,X_s) - \mathfrak{a}_s(X_s^i)\right|^p\right] ds. \qquad (47)
\end{aligned}$$

Regarding the term $\mathrm{Term}_2$ we call upon the Burkholder-Davis-Gundy's inequality which states that for any $p \ge 1$,

$$\begin{aligned}
&\left\|\sup_{s\in[0,t]}\left|\int_0^s \left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right) dW_s^i\right|\right\|_p \\
&\le C_p\left(\mathrm{E}\left[\left(\int_0^t \left|\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right)\right|^2 ds\right)^{p/2}\right]\right)^{1/p}.
\end{aligned}$$

This implies that for $p \ge 2$,

$$\begin{aligned}
&\mathrm{E}\sup_{s\in[0,t]}\left|\int_0^s \left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right) dW_s^i\right|^p \qquad (48)\\
&\le C_p^p\mathrm{E}\left[\left(\int_0^t \left|\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right)\right|^2 ds\right)^{p/2}\right] \\
&\le C_p^p t^{p/2-1}\mathrm{E}\left[\int_0^t \left|\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right)\right|^p ds\right] \\
&\le C_p^p t^{p/2-1}\mathrm{E}\left[\int_0^t \left|\left(\mathfrak{b}_{K,N}(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s(X_s^i)\right)\right|^p ds\right].
\end{aligned}$$

Now, completely analogue to the derivation of (47), we get

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^{i,K,N},X_s^{K,N}) - \mathfrak{b}_s(X_s^i)\right|^p\right] ds \\
&\le \left(2^{2p-2}D_\varphi^p B_\beta^p + 2^{2p-2}L_\varphi^p A_\beta^p F_1\right)\int_0^t H(s)ds + 2^{2p-2}L_\varphi^p A_\beta^p \frac{F_2}{N^{p/2+1/2}}t \\
&\quad +2^{p-1}\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^i,X_s) - \mathfrak{b}_s(X_s^i)\right|^p\right] ds. \qquad (49)
\end{aligned}$$

Now by taking expectations on both sides of (39) and gathering all together, we arrive at

$$
\begin{aligned}
H(t) \leq & \left(D_\varphi^p B_\alpha^p T^{p-1} + L_\varphi^p A_\alpha^p F_1^p T^{p-1}\right. \\
& \left. + C_p^p D_\varphi^p B_\beta^p d^p T^{p/2-1} + C_p^p L_\varphi^p A_\beta^p d^p T^{p/2-1} F_1^p\right) 2^{3p-3} \int_0^t H(s) ds \\
& + 2^{3p-3} \left( L_\varphi^p A_\alpha^p T^p + d^p C_p^p L_\varphi^p A_\beta^p T^{p/2} \right) \frac{F_2}{N^{p/2+1/2}} \qquad (50) \\
& + 2^{2p-2} T^{p-1} \frac{1}{N} \sum_{i=1}^N \int_0^t \mathrm{E}\left[ \left| \mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i) \right|^p \right] ds \\
& + 2^{2p-2} d^p C_p^p T^{p/2-1} \frac{1}{N} \sum_{i=1}^N \int_0^t \mathrm{E}\left[ \left| \mathfrak{b}_{K,N}(X_s^i, X_s) - \mathfrak{b}_s(X_s^i) \right|^p \right] ds.
\end{aligned}
$$

We next proceed with explicit estimates for the last two terms above. Let us write

$$
\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i) = \sum_{k=1}^K \alpha_k(X_s^i) \sum_{j=1}^N \frac{1}{N} \left( \varphi_k(X_s^j) - \gamma_k(s) \right) - \sum_{k=K+1}^\infty \alpha_k(X_s^i) \gamma_k(s),
$$

then we have by the Minkowski inequality,

$$
\begin{aligned}
\left\| \mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i) \right\|_p \leq & \sum_{k=1}^K \left\| \alpha_k(X_s^i) \frac{1}{N} \sum_{j=1}^N \xi_k^j \right\|_p \\
& + \sum_{k=K+1}^\infty \left\| \alpha_k(X_s^i) \gamma_k(s) \right\|_p,
\end{aligned}
$$

where $\xi_k^j := \varphi_k(X_s^j) - \gamma_k(s)$, $j = 1, \ldots, N$, have mean zero. Let us now observe that

$$
\begin{aligned}
\mathrm{E}\left[ \left| \sum_{j=1}^N \xi_k^j \right|^p \middle| X^i \right] &= \mathrm{E}\left[ \left| \xi_k^i + \sum_{j \neq i}^N \xi_k^j \right|^p \middle| X^i \right] \\
&\leq 2^{p-1} \mathrm{E}\left[ \left| \xi_k^i \right|^p + \left| \sum_{j \neq i}^N \xi_k^j \right|^p \middle| X^i \right] \\
&\leq 2^{2p-1} D_{k,\varphi}^p + 2^{p-1} \mathrm{E}\left[ \left| \sum_{j \neq i}^N \xi_k^j \right|^p \right]
\end{aligned}
$$

using (7). For $p \geq 2$, it follows from the Rosenthal's inequality that,

$$
\mathrm{E}\left[ \left| \sum_{j \neq i}^N \xi_k^j \right|^p \right] \leq C_p^{(1)} \left( \left( \sum_{j \neq i}^N \mathrm{E} \left| \xi_k^j \right|^2 \right)^{p/2} + \sum_{j \neq i}^N \mathrm{E} \left| \xi_k^j \right|^p \right)
$$

for a constant $C_p^{(1)}$ only depending on $p$, and, in fact, for $p = 2$ we have simply,

$$
\mathrm{E}\left[ \left| \sum_{j \neq i}^N \xi_k^j \right|^p \right] = \sum_{j \neq i}^N \mathrm{E} \left| \xi_k^j \right|^2.
$$

Thus, for $p \geq 2$,

$$\mathrm{E}\left[\left|\frac{1}{N}\sum_{j=1}^{N}\xi_k^j\right|^p \middle| X_s^i\right] \leq \frac{2^{2p-1}D_{k,\varphi}^p}{N^p} + \frac{2^{p-1}C_p^{(1)}}{N^p}\left(\left(\sum_{j\neq i}^{N}\mathrm{E}\left|\xi_k^j\right|^2\right)^{p/2} + \sum_{j\neq i}^{N}\mathrm{E}\left|\xi_k^j\right|^p\right)$$
$$\leq \frac{2^{2p-1}D_{k,\varphi}^p}{N^p} + \frac{2^{2p-1}C_p^{(1)}D_{k,\varphi}^p}{N^{p/2}} + \frac{2^{2p-1}C_p^{(1)}D_{k,\varphi}^p}{N^{p-1}}$$
$$\leq \frac{\left(C_p^{(2)}\right)^p D_{k,\varphi}^p}{N^{p/2}} \quad \text{for } N > N_p \text{ and some constants } C_p^{(2)} > 0, N_p > 0.$$

So for any $p \geq 2$,

$$\left\|\alpha_k(X_s^i)\frac{1}{N}\sum_{j=1}^{N}\xi_k^j\right\|_p^p \leq A_{k,\alpha}^p \mathrm{E}\left[(1+|X_s^i|)^p\,\mathrm{E}\left[\left|\frac{1}{N}\sum_{j=1}^{N}\xi_k^j\right|^p \middle| X_s^i\right]\right]$$
$$\leq A_{k,\alpha}^p D_{k,\varphi}^p \frac{\left(C_p^{(2)}\right)^p}{N^{p/2}}\mathrm{E}\left[(1+|X_s|)^p\right],$$

hence

$$\left\|\alpha_k(X_s^i)\frac{1}{N}\sum_{j=1}^{N}\xi_k^j\right\|_p \leq C_p^{(2)}A_{k,\alpha}D_{k,\varphi}F_3N^{-1/2} \quad \text{with} \quad F_3 := \sup_{0\leq s\leq T}\left\|1+|X_s|\right\|_p,$$

and further

$$\sum_{k=K+1}^{\infty}\left\|\alpha_k(X_s^i)\gamma_k(s)\right\|_p \leq F_3\sum_{k=K+1}^{\infty}A_{k,\alpha}\left|\gamma_k(s)\right|.$$

We thus obtain,

$$\left\|\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i)\right\|_p \leq C_p^{(2)}A_\alpha D_\varphi F_3 N^{-1/2} + F_3\sum_{k=K+1}^{\infty}A_{k,\alpha}\left|\gamma_k(s)\right|,$$

that is,

$$\mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i)\right|^p\right] \leq 2^{p-1}\left(C_p^{(2)}\right)^p A_\alpha^p D_\varphi^p F_3^p N^{-p/2} + 2^{p-1}F_3^p\left(\sum_{k=K+1}^{\infty}A_{k,\alpha}\left|\gamma_k(s)\right|\right)^p. \tag{51}$$

Analogously we get

$$\mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^i, X_s) - \mathfrak{b}_s(X_s^i)\right|^p\right] \leq 2^{p-1}\left(C_p^{(2)}\right)^p A_\beta^p D_\varphi^p F_3^p N^{-p/2} + 2^{p-1}F_3^p\left(\sum_{k=K+1}^{\infty}A_{k,\beta}\left|\gamma_k(s)\right|\right)^p. \tag{52}$$

Now, combining the estimates (51) and (52) with (50), yields for $0 \leq t \leq T$,

$$
\begin{aligned}
H(t) \leq & \left(C_{p,\varphi,X}T^{p-1} + D_{p,\varphi,X}d^pT^{p/2-1}\right)\int_0^t H(s)ds \\
& + \left(E_{p,\varphi,X}T^p + F_{p,\varphi,X}d^pT^{p/2} + O(N^{-1/2})\right)N^{-p/2} \\
& + G_{p,\varphi,X}T^{p-1}\int_0^T \left(\sum_{k=K+1}^{\infty} A_{k,\alpha}\left|\gamma_k(s)\right|\right)^p ds \\
& + H_{p,\varphi,X}d^pT^{p/2-1}\int_0^T \left(\sum_{k=K+1}^{\infty} A_{k,\beta}\left|\gamma_k(s)\right|\right)^p ds
\end{aligned}
$$

with abbreviations

$$
\begin{aligned}
C_{p,\varphi,X} &= 2^{3p-3}D_\varphi^pB_\alpha^p + 2^{3p-3}L_\varphi^pA_\alpha^pF_1^p \\
D_{p,\varphi,X} &= 2^{3p-3}C_p^pD_\varphi^pB_\beta^p + 2^{3p-3}C_p^pL_\varphi^pA_\beta^pF_1^p \\
E_{p,\varphi,X} &= 2^{3p-3}\left(C_p^{(2)}\right)^p A_\alpha^pD_\varphi^pF_3^p \\
F_{p,\varphi,X} &= 2^{3p-3}C_p^p\left(C_p^{(2)}\right)^p A_\beta^pD_\varphi^pF_3^p \\
G_{p,\varphi,X} &= 2^{3p-3}F_3^p \\
H_{p,\varphi,X} &= 2^{3p-3}C_p^pF_3^p.
\end{aligned}
$$

Finally, the statement of the theorem follows from Gronwall's lemma by raising the resulting inequality to the power $1/p$, then using that $\left(\sum_{i=1}^q |a_i|^p\right)^{1/p} \leq \sum_{i=1}^q |a_i|$ for arbitrary $a_i \in \mathbb{R}$, $p, q \in \mathbb{N}$, a Minkowski type inequality, and the observation that

$$
\mathrm{E}\left[\overline{\Delta_T^p}\right] = \frac{1}{N}\sum_{i=1}^N \mathrm{E}\left[\sup_{s\in[0,T]}\left|\Delta_s^i\right|^p\right] = \mathrm{E}\left[\sup_{s\in[0,T]}\left|\Delta_s^\cdot\right|^p\right].
$$

## 6.2 Proof of Theorem 3

(i): Under the assumption (18) the functions $H_{a^j}$ and $H_b$ are locally Lipschitz in $\mathbb{R}\times\mathbb{R}_{\geq 0}$ and, obviously, their extensions $(p,q)\to H_{a^j}(p,|q|)$, $H_b(p,|q|)$ are locally Lipschitz in $\mathbb{R}\times\mathbb{R}$. Thus, by standard ODE theory, there exists a unique solution to the system

$$
\begin{aligned}
G_t' &= H_b^2\left(A_t,|G_t|\right) + 2H_{a^1}\left(A_t,|G_t|\right)G_t \\
A_t' &= H_{a^0}\left(A_t,|G_t|\right) + H_{a^1}\left(A_t,|G_t|\right)A_t, \quad (A_0,G_0) = (x_0,0), \quad 0\leq t< t_\infty \leq \infty,
\end{aligned}
$$

for some possibly finite explosion time $t_\infty$. Then it can be straightforwardly checked that this (unique) solution can be represented as

$$
\begin{aligned}
G_t &= \int_0^t H_b^2\left(A_s,|G_s|\right)e^{2\int_s^t H_{a^1}(A_r,|G_r|)dr}ds \qquad (53)\\
A_t &= e^{\int_0^t H_{a^1}(A_s,|G_s|)ds}x_0 + \int_0^t H_{a^0}\left(A_s,|G_s|\right)e^{\int_s^t H_{a^1}(A_r,|G_r|)dr}ds, \ 0\leq t<t_\infty,
\end{aligned}
$$

whence in particular $G_t \geq 0$ for $0 \leq t < t_\infty$. This proves (i).

(ii): By straightforward differentiating with respect to $t$, it follows that (22) is a solution to (21). Let us abbreviate in (22)

$$\mathfrak{a}_t^0 \equiv H_{a^0}\left(A_t, G_t\right), \quad \mathfrak{a}_t^1 \equiv H_{a^1}\left(A_t, G_t\right), \quad \mathfrak{b}_t \equiv H_b\left(A_t, G_t\right), \quad 0 \leq t < t_\infty.$$

The characteristic function of $X_t$ in (22) then takes the form

$$\varphi_t(v) = \exp\left[\mathrm{i}v \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds - \frac{1}{2}v^2 \int_0^t \mathfrak{b}_s^2 e^{2\int_s^t \mathfrak{a}_r^1 dr} ds + \mathrm{i}v e^{\int_0^t \mathfrak{a}_s^1 ds} x_0\right]. \quad (54)$$

Since

$$\frac{e^{-\frac{(p-u)^2}{2q}}}{\sqrt{2\pi q}} = \frac{1}{2\pi}\int e^{-\mathrm{i}vu} \exp\left[\mathrm{i}vp - v^2 q/2\right] dv,$$

we have for $j = 0, 1$,

$$H_{a^j}(p, q) = \frac{1}{2\pi}\int a^j(u) du \int \exp\left[\mathrm{i}vp - v^2 q/2\right] e^{-\mathrm{i}vu} dv.$$

It then follows that

$$\begin{aligned} & H_{a^j}(e^{\int_0^t \mathfrak{a}_s^1 ds} x_0 + \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds, \int_0^t \left(\mathfrak{b}_s^0\right)^2 e^{2\int_s^t \mathfrak{a}_r^1 dr} ds) \\ & = \frac{1}{2\pi}\int a^j(u) du \int \varphi_t(v) e^{-\mathrm{i}vu} dv \\ & = \int a^j(u)\mu_t(u) du = \mathrm{E}\left[a^j(X_t)\right], \quad j = 0, 1, \end{aligned} \quad (55)$$

with $\mu_t$ being the density of $X_t$, and similarly,

$$H_b(e^{\int_0^t \mathfrak{a}_s^1 ds} x_0 + \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds, \int_0^t \left(\mathfrak{b}_s^0\right)^2 e^{2\int_s^t \mathfrak{a}_r^1 dr} ds) = \mathrm{E}\left[b(X_t)\right]. \quad (56)$$

On the other hand, in view of (53) and the fact that $G \geq 0$, one has

$$\begin{aligned} \int_0^t \left(\mathfrak{b}_s^0\right)^2 e^{2\int_s^t \mathfrak{a}_r^1 dr} ds &= G_t \\ e^{\int_0^t \mathfrak{a}_s^1 ds} x_0 + \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds &= A_t, \end{aligned} \quad (57)$$

that is, by (55), (56), and (57), we obtain (20) from (21).

# 7 Appendix

## 7.1 Existence of moments

**Theorem 4** *Fix some $p \geq 2$ and suppose that $\mathrm{E}[|X_0|^p] < \infty$. Then it holds under assumptions (AC) and (AF),*

$$\left\| \sup_{s\in[0,T]} \left|X_s^{\cdot,K,N}\right| \right\|_p < \infty,$$

*uniformly in $K$ and $N$.*

**Proof.** Fix some $i \in \{1, \ldots, N\}$ and for every $R > 0$ introduce the stopping time

$$\tau_{i,R} = \inf\left\{t \in [0,T] \,:\, \left|X_t^{i,K,N} - X_0^i\right| > R\right\}.$$

We obviously have

$$\sup_{t\in[0,T]} \left|X_{t\wedge\tau_{i,R}}^{i,K,N}\right| \le R + \left|X_0^i\right|$$

so that the non-decreasing function $f_R(t) := \left\|\sup_{s\in[0,t]}\left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right|\right\|_p$, $t \in [0,T]$, is bounded by $R + \left\|X_0^i\right\|_p$. On the other hand

$$\begin{aligned}
\sup_{s\in[0,t]} \left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right| &\le \left|X_0^i\right| + \int_0^{t\wedge\tau_{i,R}} \left|\mathfrak{a}_{K,N}(X_r^{i,K,N}, X_r^{K,N})\right| dr \\
&\quad + \sup_{s\in[0,t]}\left|\int_0^{s\wedge\tau_{i,R}} \mathfrak{b}_{K,N}(X_r^{i,K,N}, X_r^{K,N})\, dW_r^i\right| \\
&\le \left|X_0^i\right| + \int_0^{t\wedge\tau_{i,R}} \left|\mathfrak{a}_{K,N}(X_r^{i,K,N}, X_r^{K,N})\right| dr \\
&\quad + \sum_{q=1}^{d}\sup_{s\in[0,t]}\left|\int_0^{s\wedge\tau_{i,R}} \mathfrak{b}_{K,N}^q(X_r^{i,K,N}, X_r^{K,N})\, dW_r^i\right|
\end{aligned}$$

(cf. (38)). It then follows from the Minkowski and BDG inequality that

$$\begin{aligned}
f_R(t) &\le \|X_0\|_p + \int_0^t \left\|1_{\{s\le\tau_{i,R}\}}\mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N})\right\|_p ds \\
&\quad + dC_p^{BDG}\left\|\sqrt{\int_0^{t\wedge\tau_{i,R}}\left|\mathfrak{b}_{K,N}(X_s^{i,K,N}, X_s^{K,N})\right|^2 ds}\right\|_p \\
&\le \|X_0\|_p + A_\alpha D_\varphi \int_0^t \left\|\left(1 + \left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right|\right)\right\|_p ds \\
&\quad + A_\beta D_\varphi d C_p^{BDG}\left\|\sqrt{\int_0^t\left|\left(1 + \left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right|\right)\right|^2 ds}\right\|_p \\
&\le \|X_0\|_p + A_\alpha D_\varphi \int_0^t \left(1 + \left\|\left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right|\right\|_p\right) ds \\
&\quad + A_\beta D_\varphi d C_p^{BDG}\left(\sqrt{t} + \left(\int_0^t \left\|\left|X_{s\wedge\tau_{i,R}}^{i,K,N}\right|^2\right\|_{p/2} ds\right)^{1/2}\right)
\end{aligned}$$

again by the Minkowski inequality ($p \ge 2$). Consequently, the function $f_R$ satisfies

$$f_R(t) \le \|X_0\|_p + A_\alpha D_\varphi \int_0^t (1 + f_R(s))\, ds + A_\beta D_\varphi d C_p^{BDG}\left(\sqrt{t} + \left(\int_0^t f_R^2(s)\, ds\right)^{1/2}\right),$$

that is,

$$f_R(t) \leq \|X_0\|_p + A_\alpha D_\varphi t + A_\beta D_\varphi d C_p^{BDG} \sqrt{t}$$
$$+ A_\alpha D_\varphi \int_0^t f_R(s)\, ds + A_\beta D_\varphi d C_p^{BDG} \left( \int_0^t f_R^2(s)\, ds \right)^{1/2}.$$

By Lemma 2 (see Appendix) it follows that

$$\left\| \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p \leq 2e^{\left(2A_\alpha D_\varphi + A_\beta^2 D_\varphi^2 d^2 \left(C_p^{BDG}\right)^2\right)T} \times \tag{58}$$
$$\left( \|X_0\|_p + A_\alpha D_\varphi T + A_\beta D_\varphi d C_p^{BDG} \sqrt{T} \right).$$

Now note that the stopping times $\tau_{i,R}$ are non-decreasing in $R$, and thus converges non-decreasingly to $\tau_{i,\infty}$ say, with $\tau_{i,\infty} \in [0,T] \cup \{\infty\}$. Thus,

$$R \to \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right|$$

is nondecreasing with

$$\lim_{R\uparrow\infty} \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| = \begin{cases} \sup_{s\in[0,T]} \left| X^{i,K,N}_s \right| & \text{on } \{\tau_{i,\infty} = \infty\} \\ \infty & \text{on } \{\tau_{i,\infty} \leq T\} \end{cases}. \tag{59}$$

Indeed, on the set $\{\tau_{i,\infty} \leq T\}$ we have for any $R > 0$, $\left| X^{i,K,N}_{\tau_{i,R}} - X^i_0 \right| \geq R$ with $\tau_{i,R} \leq T$, so that

$$\sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \geq \left| X^{i,K,N}_{\tau_{i,R}} \right| \geq \left| X^{i,K,N}_{\tau_{i,R}} \right| \geq R - \left| X^i_0 \right|.$$

The Fatou lemma (59) implies (with $0 := \infty \cdot 0$),

$$\left\| \lim_{R\uparrow\infty} 1_{\{\tau_{i,\infty}\leq T\}} \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p = \infty \cdot P\left(\{\tau_{i,\infty} \leq T\}\right)$$
$$\leq \liminf_R \left\| 1_{\{\tau_{i,\infty}\leq T\}} \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p$$
$$\leq \liminf_R \left\| \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p < \infty,$$

because of (58). So $P\left(\{\tau_{i,\infty} \leq T\}\right) = 0$, i.e. $\tau_{i,\infty} = \infty$ almost surely. Again by the Fatou lemma, (59) then implies

$$\left\| \sup_{s\in[0,T]} \left| X^{i,K,N}_s \right| \right\|_p \leq \liminf_R \left\| \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p < \infty,$$

uniformly in $K$ and $N$, because of (58) again. ■

The following lemma is consequence of Gronwall's theorem.

**Lemma 2** *Let $f : [0,T] \to \mathbb{R}_+$ and $\psi : [0,T] \to \mathbb{R}_+$ be two non-negative non-decreasing functions satisfying*

$$f(t) \le A \int_0^t f(s)\,ds + B \left( \int_0^t f^2(s)\,ds \right)^{1/2} + \psi(t), \quad t \in [0,T], \tag{60}$$

*where $A, B$ are two positive real constants. Then*

$$f(t) \le 2e^{\left(2A+B^2\right)t}\,\psi(t), \quad t \in [0,T].$$

**Proof.** It follows from the elementary inequality $\sqrt{xy} \le \frac{1}{2}\left(x/B + By\right)$, $x, y \ge 0$,$B > 0$, that

$$\left( \int_0^t f^2(s)\,ds \right)^{1/2} \le \left( f(t) \int_0^t f(s)\,ds \right)^{1/2} \le \frac{f(t)}{2B} + \frac{B}{2} \int_0^t f(s)\,ds.$$

Plugging this into (60) yields

$$f(t) \le (2A + B^2) \int_0^t f(s)\,ds + 2\psi(t).$$

Now the standard Gronwall inequality yields the desired result. ■

# References


[1] Fabio Antonelli, Arturo Kohatsu-Higa, et al. Rate of convergence of a particle method to the solution of the McKean–Vlasov equation. *The Annals of Applied Probability*, 12(2):423–476, 2002.

[2] Mireille Bossy and Denis Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Mathematics of Computation of the American Mathematical Society*, 66(217):157–192, 1997.

[3] AN Drozdov and M Morillo. Expansion for the moments of a nonlinear stochastic model. *Physical review letters*, 77(16):3280, 1996.

[4] Till Daniel Frank. *Nonlinear Fokker-Planck equations.* Springer Series in Synergetics. Springer-Verlag, Berlin, 2005. Fundamentals and applications.

[5] Tadahisa Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3):331–348, 1984.

[6] Desmond J Higham, Xuerong Mao, and Andrew M Stuart. Strong convergence of euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis*, 40(3):1041–1063, 2002.

[7] Martin Hutzenthaler, Arnulf Jentzen, and Peter E Kloeden. Strong convergence of an explicit numerical method for sdes with nonglobally lipschitz continuous coefficients. *The Annals of Applied Probability*, pages 1611–1641, 2012.

[8] Peter Kloeden and Tony Shardlow. Gauss-quadrature method for one-dimensional mean-field SDEs. *SIAM Journal on Scientific Computing*, 39(6):A2784–A2807, 2017.

[9] Vassili N Kolokoltsov. *Nonlinear Markov processes and kinetic equations*, volume 182. Cambridge University Press, 2010.

[10] Marcin Kostur, J Łuczka, and L Schimansky-Geier. Nonequilibrium coupled brownian phase oscillators. *Physical Review E*, 65(5):051115, 2002.

[11] Chi-Fai Lo and CH Hui. A simple analytical model for dynamics of time-varying target leverage ratios. *The European Physical Journal B*, 85(3):1–6, 2012.

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

[13] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996.

[14] Hiroshi Shimizu and Takenori Yamada. Phenomenological equations of motion of muscular contraction. *Progress of theoretical physics*, 47(1):350–351, 1972.

[15] G Szegö. Orthogonal polynomials 4th edn (providence, ri: American mathematical society). 1975.

[16] Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX - 1989*, pages 165–251. Springer, 1991.

[17] DS Zhang, GW Wei, DJ Kouri, and DK Hoffman. Numerical method for the nonlinear Fokker-Planck equation. *Physical review E*, 56(1):1197, 1997.