Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Projected particle methods for solving McKean-Vlasov equations

Denis Belomestny,[1] John G. M. Schoenmakers[2]

submitted: November 29, 2016

[1] Duisburg-Essen University
Mathematics Department
Thea-Leymann-Str. 9
45127 Essen
Germany
E-Mail: denis.belomestny@uni-due.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: John.Schoenmakers@wias-berlin.de

No. 2341
Berlin 2016

2010 *Mathematics Subject Classification.* 60H10, 60K35.

Key words and phrases. Mckean-Vlasov equations, particle systems, projection estimators, explicit solutions.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We study a novel projection-based particle method to the solution of the corresponding McKean-Vlasov equation. Our approach is based on the projection-type estimation of the marginal density of the solution in each time step. The projection-based particle method can profit from additional smoothness of the underlying density and leads in many situation to a significant reduction of numerical complexity compared to kernel density estimation algorithms. We derive strong convergence rates and rates of density estimation. The case of linearly growing coefficients of the McKean-Vlasov equation 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.

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 [6] to model plasma dynamics. Later nonlinear Markov processes were studied by a number of authors; we mention here the books of Kolokoltsov [4] and Sznitman [9]. 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., [3] 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 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 $X_0 \sim \xi$ 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 $b : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^m$. If the functions a and b are smooth with uniformly bounded derivatives and the density of μ_0 satisfies

$$\mu_0(x) \lesssim \exp\left(-\rho_0 |x|^{\rho_1}\right), \quad |x| \to \infty$$

for some $\rho_0 > 0$, $\rho_1 > 0$, then (see [1]) there is a unique strong solution of (1) such that for all $p > 1$,

$$\mathrm{E}\left(\sup_{s\leq T} |X_s|^p\right) \leq \infty. \tag{2}$$

Assume that $d = 1$ and for any $t \geq 0$, 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 the linear Fokker-Planck equation in the SDE case. In Section 4.1 we will show that if the drift a is moreover affine in x, and the diffusion coefficient b independent of x, 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 explosion, hence where the assumptions of [1] are (partially) violated.

The theory of the propagation of chaos developed in [9], 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 $i = 1, \ldots, N$, where $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$, ξ^i, $i = 1, \ldots, N$, are i.i.d copies of ξ, distributed according the law μ_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 [7].

Despite the numerous branches of research on stochastic particle systems, results on numerical approximations of McKean-Vlasov-SDEs are very sparse. Authors in [1] proposed to use the Euler scheme with time-step $h = T/L$, that for $l = 0, \ldots, L-1$, 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 $\Delta_{l+1}W^i = W_{h(l+1)}^i - W_{hl}^i$. 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.

2 Projected particle method

Assume that for any $t \geq 0$, the measure $\mu_t(du)$ possesses a bounded density $\mu_t(u)$. Let $(\varphi_k,\, k = 0, 1, 2, ..)$ be a total orthonormal system in $L_2(\mathbb{R}^d)$. We can formally write

$$\mu_t(u) = \sum_{k=0}^{\infty} \gamma_k(t)\varphi_k(u),$$

where the sequence

$$\gamma_k(t) := \int_{\mathbb{R}^d} \mu_t(u)\varphi_k(u)\, du = \mathrm{E}\left[\varphi_k(X_t)\right] \tag{6}$$

converges in l_2 for any fixed $t \in [0, T]$. Let us introduce the functions

$$\begin{aligned}\alpha_k(x) &:= \int a(x,u)\varphi_k(u)\, du \in \mathbb{R}^d,\\ \beta_k(x) &:= \int b(x,u)\varphi_k(u)\, du \in \mathbb{R}^d \times \mathbb{R}^m\end{aligned}$$

so that

$$\begin{aligned}\int_{\mathbb{R}^d} a(x,y)\mu_t(dy) &= \sum_{k=0}^{\infty} \alpha_k(x)\gamma_k(t),\\ \int_{\mathbb{R}^d} b(x,y)\mu_t(dy) &= \sum_{k=0}^{\infty} \beta_k(x)\gamma_k(t).\end{aligned} \tag{7}$$

Fix some natural number $K > 0$ and consider a *projected particle approximation* for (1)

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

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

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

Consequently, we can consider an Euler-type approximation for (8)

$$\begin{aligned}\bar{X}_t^{i,K,N} = \bar{X}_{\eta(t)}^{i,K,N} &+ \sum_{k=0}^{K} \gamma_k^N\left(\eta(t)\right)\, \alpha_k\big(\bar{X}_{\eta(t)}^{i,K,N}\big)\, (t - \eta(t))\\ &+ \sum_{k=0}^{K} \gamma_k^N\left(\eta(t)\right)\, \beta_k\big(\bar{X}_{\eta(t)}^{i,K,N}\big)\, (W_t^i - W_{\eta(t)})\end{aligned} \tag{9}$$

for $i = 1, \ldots, N$, and $h = T/L$, where $\eta(t) := lh$ for $t \in [lh, (l+1)h)$. Note that in order to generate a discretized particle system $(\bar{X}_{hl}^{i,N})$, $i = 1, \ldots, N$, $l = 1, \ldots, L$, we need to perform (up to a constant depending on the dimension) NKL 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.

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

$$\left(X^{j,K,N}, X^{K,N}, X^j\right) \overset{\text{distr.}}{=} \left(X^{\cdot,K,N}, X^{K,N}, X^{\cdot}\right) \quad \text{for } j = 1, ..., N. \tag{10}$$

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

(AF) The basis functions (φ_k) fulfil

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

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

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

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d} |\alpha_k(x)| \leq A_{k,\alpha}(1 + |x|) \quad \text{with} \\
&\qquad \sum_{k=0}^{\infty} A_{k,\alpha} \leq A_\alpha \quad \text{and} \quad \sum_{k=0}^{\infty} L_{k,\varphi} A_{k,\alpha} \leq L_\varphi A_\alpha, \\
&\sup_{x\in\mathbb{R}^d} |\beta_k(x)| \leq A_{k,\beta}(1 + |x|) \quad \text{with} \\
&\qquad \sum_{k=0}^{\infty} A_{k,\beta} \leq A_\beta, \quad \text{and} \quad \sum_{k=0}^{\infty} L_{k,\varphi} A_{k,\beta} \leq L_\varphi A_\beta,
\end{aligned}$$

for some constant $L_\varphi > 0$, and further

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

(AM) The density of μ_0 satisfies

$$\mu_0(x) \lesssim \exp\left(-\rho_0 |x|^{\rho_1}\right), \quad |x| \to \infty$$

for some $\rho_0 > 0$ and $\rho_1 > 0$.

Note that, if the sequence $(L_{k,\varphi})_{k=0,1,\ldots}$ in (AF) is bounded, one may take $L_\varphi = \sup_{k\geq 0} L_{k,\varphi}$ in (AC). Henceforth, 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 can be proved.

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

$$\left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_p \lesssim N^{-1/2} + \sum_{k=K+1}^{\infty} A_{k,\alpha} \left\| \gamma_k \right\|_{L_p[0,T]} + \sum_{k=K+1}^{\infty} A_{k,\beta} \left\| \gamma_k \right\|_{L_p[0,T]}, \tag{11}$$

where $\lesssim$ stands for an inequality with some positive finite constant depending on $A_\alpha, A_\beta, B_\alpha, B_\beta, D_\varphi, L_\varphi, \rho_0, \rho_1,$ p, and T.

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

$$\left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_{p'} \leq \left\| \sup_{0 \leq r \leq T} \left| X_r^{\cdot,K,N} - X_r^{\cdot} \right| \right\|_p \tag{12}$$

for any $p \geq 2$.

The next theorem, on the convergence of the Euler approximation (9) 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) that*

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

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

Let us now discuss the estimation of the densities μ_t, $t \geq 0$. Fix some $t > 0$ and set

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

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

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

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

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

The last term always converges to zero, since μ_t is bounded and hence $\mu_t \in L_2(\mathbb{R}^d)$. The first term can be controlled for any K by taking N large enough. However, in order to ensure that the middle term goes to zero for $K \to \infty$ we need in addition to (AC), (AF), and (AM) some supplementary assumptions.

Discussion The bound (11) is proved under rather general assumptions on the coefficients $a(x,y)$ and $b(x,y)$. So we allow for a linear growth of these coefficients in x. This makes the proof of the bound rather challenging, since we need to avoid the explosion of coefficients while using the Gronwall's lemma. 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 6.1), gives the desired bound. The bound (11) 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 Fourier coefficients (α_k) and (β_k) decay fast, then the approximation error can be made small even for medium values of K.

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

$$H_j(x) = (-1)^j e^{x^2}\frac{d^j}{dx^j}(e^{-x^2}).$$

Hermite polynomials are orthogonal with respect to the weight function e^{-x^2} and satisfy: $\int_{\mathbb{R}} H_j(x)H_\ell(x)e^{-x^2}dx = 2^j j!\sqrt{\pi}\delta_{j,\ell}$. The Hermite function of order j is given by:

$$\varphi_j(x) = c_j H_j(x)e^{-x^2/2}, \quad c_j = \left(2^j j!\sqrt{\pi}\right)^{-1/2}. \tag{13}$$

The sequence $(\varphi_j, j \geq 0)$ is an orthonormal basis of $L_2(\mathbb{R})$. The density μ_t can be developed in the Hermite basis $\mu_t(u) = \sum_{j\geq 0}\gamma_j(t)\varphi_j(u)$ where $\gamma_j(t) = \int_{\mathbb{R}}\mu_t(x)\varphi_j(x)dx = \langle\mu_t, \varphi_j\rangle$. This leads to a collection of projection estimators $\hat{\mu}_{K,N} = \sum_{j=0}^{K}\gamma_j^N(t)\varphi_j$, where $\gamma_j^N = N^{-1}\sum_{i=1}^{N}\phi_j(X_t^i)$ is the empirical estimator of $\gamma_j(t)$. Let us discuss the assumptions (AC) and (AF). Using Theorem 34 in [2] and the definition of α_k and β_k, we can derive the following result.

Theorem 3. *Suppose that for any $u \in \mathbb{R}$, the coefficients $a(x,u)$ and $b(x,u)$ admit derivatives in u up to order $s > 2$ such that all functions*

$$\begin{gathered} a(x,u), \partial_u a(x,u), \ldots, \partial_u^s a(x,u), u^{s-\ell}\partial_u^\ell a(x,u) \\ b(x,u), \partial_u b(x,u), \ldots, \partial_u^s b(x,u), u^{s-\ell}\partial_u^\ell b(x,u), \end{gathered}$$

$\ell = 0, \ldots, s-1$, belong to $L_2(\mathbb{R})$ (in u) together with their first derivatives in x. Then the assumptions (AC) and (AF) are satisfied and

$$\left\| \sup_{0\leq r\leq T}\left|X_r^{\cdot,K,N} - X_r^{\cdot}\right| \right\|_p \lesssim K^{1-s/2} + N^{-1/2}, \tag{14}$$

as $K, N \to \infty$.

Remark 2. *As a rule, one chooses N and K such that the errors in (14) are balanced, that is $N^{1/(s-2)} \sim K$, yielding a proportional reduction of computational cost of order of $N \cdot K/N^2 \sim N^{-(s-3)/(s-2)}$. In [1] conditions are formulated, guaranteeing that all measures μ_t, $t \geq 0$, possess smooth exponentially decaying densities. In this case we can additionally profit from the decay of the Fourier coefficients (γ_k) such that the convergence rates in* (11) *give rise to a proportional reduction of computational cost approaching N^{-1}. Generally, the smoother the density μ_t is, the faster is the decay rate of its Fourier coefficients $\gamma_k(t)$.*

4 Specific models

4.1 Generalised Shimizu-Yamada Models

Inspired by the work of Shimizu and Yamada [8], [10] and [5], we consider the 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 $\mathrm{E}\left[a^1(X_t)\right]$.

Theorem 4. *Define*

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

and let (A_t, G_t) *be a solution of the following system of ODEs*

$$\begin{aligned} G_t' &= H_b^2(A_t, G_t) + 2H_{a^1}(A_t, G_t) G_t \\ A_t' &= H_{a^0}(A_t, G_t) + H_{a^1}(A_t, G_t) A_t, \quad (G_0, A_0) = (0, x_0). \end{aligned} \tag{15}$$

Then the McKean-Vlasov equation

$$dX_t = \left(\mathrm{E}\left[a^0(X_t)\right] + X_t \mathrm{E}\left[a^1(X_t)\right]\right) dt + \mathrm{E}\left[b(X_t)\right] dW_t, \quad X_0 = x_0 \tag{16}$$

with bounded functions $a^i : \mathbb{R} \to \mathbb{R}$, $i = 0, 1$, *has an explicit solution of the form:*

$$\begin{aligned} X_t &= x_0 e^{\int_0^t H_{a^1}(A_s, G_s) ds} + \int_0^t H_{a^0}(A_s, G_s) e^{\int_s^t H_{a^1}(A_r, G_r) dr} ds \\ &+ \int_0^t H_b(A_s, G_s) e^{\int_s^t H_{a^1}(A_r, G_r) dr} dW_s. \end{aligned} \tag{17}$$

Proof. It can be straightforwardly checked that for arbitrary continuous and deterministic functions $\mathfrak{a}_t^0, \mathfrak{a}_t^1$, and $\mathfrak{b}_t$ on $[0, T]$, the solution of the SDE

$$dX_t = \left(\mathfrak{a}_t^0 + \mathfrak{a}_t^1 X_t\right) dt + \mathfrak{b}_t dW_t, \quad X_0 = x_0, \quad 0 \le t \le T,$$

reads explicitly,

$$X_t = x_0 e^{\int_0^t \mathfrak{a}_s^1 ds} + \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds + \int_0^t \mathfrak{b}_s e^{\int_s^t \mathfrak{a}_r^1 dr} dW_s,$$

and thus the characteristic function of X_t takes the form

$$\varphi_t(v) = \exp\left[\mathrm{i}v \int_0^t \mathfrak{a}_s^0 e^{\int_s^t \mathfrak{a}_r^1 dr} ds - \frac{1}{2} v^2 \int_0^t \mathfrak{b}_s^2 e^{2\int_s^t \mathfrak{a}_r^1 dr} ds + \mathrm{i}v e^{\int_0^t \mathfrak{a}_s^1 ds} x_0\right]. \tag{18}$$

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 $j = 0, 1$,

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

Now let

$$\mathfrak{a}_t^j = \mathrm{E}\left[a^j(X_t)\right], \quad j = 0, 1, \quad \text{and} \quad \mathfrak{b}_t = \mathrm{E}\left[b(X_t)\right].$$

It then follows that

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

and similarly,

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

By next introducing

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

the system (15) follows straightforwardly. Conversely, it is easy to see that a solution to (15) yields an explicit solution (17) to (16). □

Example 2. *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

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

with $c(u) = c_0 + c_1 u$. In particular, the $H_c(p,q)$ do not depend on q, and so (15) simplifies to

$$
A_t' = a_0^0 + \left(a_1^0 + a_0^1\right)A_t + a_1^1 A_t^2, \quad A_0 = x_0. \tag{19}
$$

The solution (can be checked via Mathematica) is given by

$$
A_t = -\frac{(a_1^0 + a_0^1)}{2a_1^1} + \frac{\sqrt{-D}}{2a_1^1}\tan\left[\frac{1}{2}\sqrt{-D}t + \arctan\left[\frac{a_1^0 + a_0^1 + 2a_1^1 x_0}{\sqrt{-D}}\right]\right] \tag{20}
$$

if $D := \left(a_1^0 + a_0^1\right)^2 - 4a_0^0 a_1^1 < 0$*, and by*

$$A_t = \frac{\sqrt{D} - a_1^0 - a_0^1}{2a_1^1} - \frac{\sqrt{D}}{a_1^1} \frac{1}{1 + \frac{\sqrt{D} + a_1^0 + a_0^1 + 2a_1^1 x_0}{\sqrt{D} - a_1^0 - a_0^1 - 2a_1^1 x_0} e^{\,-\sqrt{D}t}} \tag{21}$$

if $D := \left(a_1^0 + a_0^1\right)^2 - 4a_0^0 a_1^1 \geq 0$*. Consequently, the corresponding McKean-Vlasov SDE reads*

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

with explicit solution

$$X_t = x_0 e^{\int_0^t \left(a_0^1 + a_1^1 A_s\right) ds} + \int_0^t \left(a_0^0 + a_1^0 A_s\right) e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} ds$$
$$+ \int_0^t \left(b_0 + b_1 A_s\right) e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} dW_s, \tag{22}$$

where A_t *is given by (20) or (21). Moreover, it is also possible to give closed form expressions for the mean and variance of (22), but omitted here since these expressions are rather long.*

Example 3. *By taking in Example 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 Shimizo-Yamada model. With $\sqrt{D} := -a_1^0 - a_0^1$*, and taking the limit for* $a_1^1 \to 0$ *in (21) we get after some trivial manipulations*

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

which of course can also be found directly from (19) by taking $a_0^0 = a_1^1 = 0$*. From (22) we then get straightforwardly 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 ([3], Section 3.10), where* $a_1^0 + a_0^1 = -\gamma$ *and* $a_0^1 = -\gamma - \kappa$*.*

Example 4. *By taking in Example 2*

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

we straightforwardly get from (21),

$$A_t = \frac{x_0 e^{\,a_0^1 t}}{1 - \frac{a_1^1}{a_0^1} x_0 \left(e^{\,a_0^1 t} - 1\right)}, \tag{23}$$

and

$$X_t = x_0 e^{\int_0^t \left(a_0^1 + a_1^1 A_s\right) ds} + \int_0^t b_0 e^{\int_s^t \left(a_0^1 + a_1^1 A_r\right) dr} dW_s, \tag{24}$$

respectively. Plugging (23) into (24) 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 solution

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

Remark 3. *From Example 4 it is clear that if* $a_1^1 \neq 0$*, the McKean-Vlasov solution may explode in finite time. On the other hand, this is not surprising since when* $a_1^1 \neq 0$ *the derivative* $\partial_u a(x, u)$ *is unbounded and so the main results in [1] do not apply.*

5 Proofs

5.1 Proof of Theorem 1

Let us introduce

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

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

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,

$$\left|\Delta_t^i\right|^p \leq 2^{p-1}t^{p-1}\int_0^t \left|\mathfrak{a}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{a}_s(X_s^i)\right|^p ds \\ + 2^{p-1}d^{p-1}\sum_{q=1}^d \left|\int_0^t \left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{b}_s^q(X_s^i)\right) dW_s^i\right|^p, \tag{25}$$

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

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

Assumption (AC) implies

$$\begin{aligned}\left|\mathfrak{a}_{K,N}(x,y) - \mathfrak{a}_{K,N}(x',y')\right| &= \left|\frac{1}{N}\sum_{j=1}^N\sum_{k=1}^K \left(\alpha_k(x)\varphi_k(y_j) - \alpha_k(x')\varphi_k(y_j')\right)\right| \\ &\leq \frac{1}{N}\sum_{j=1}^N\sum_{k=1}^K \left|\alpha_k(x) - \alpha_k(x')\right|\left|\varphi_k(y_j')\right| \\ &+ \frac{1}{N}\sum_{j=1}^N\sum_{k=1}^K \left|\alpha_k(x)\right|\left|\varphi_k(y_j) - \varphi_k(y_j')\right| \\ &\leq \left|x - x'\right| D_\varphi B_{\alpha,\varphi} + \frac{L_\varphi A_\alpha}{N}(1 + |x'|)\sum_{j=1}^N \left|y_j - y_j'\right|. \end{aligned} \tag{26}$$

Hence

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

So it holds that

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

and then it follows that, with regard to Term_1,

$$\begin{aligned}\text{Term}_1 \leq\ & 2^{2p-2} D_\varphi^p B_\alpha^p \int_0^t \mathrm{E}\left[\overline{\Delta_s^p}\right] ds \\ & + 2^{2p-2} L_\varphi^p A_\alpha^p \int_0^t \mathrm{E}\left[\overline{\Delta_s^p} \cdot \frac{1}{N}\sum_{j=1}^N (1+\left|X_s^i\right|)^p\right] ds \\ & + 2^{p-1}\frac{1}{N}\sum_{i=1}^N \int_0^t \mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i)\right|^p\right] ds. \end{aligned} \tag{27}$$

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

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

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

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

where on the one hand

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

and on the other

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

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

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

for N large enough and

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

where due to Theorem 5

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

Thus, finally,

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

with $F_1:=\theta^{1/p}+\sup_{0\le s\le T}\left\|1+|X_s|\right\|_p$ 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 (27) reads

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^N\int_0^t\mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^{i,K,N},X_s^{K,N})\;-\mathfrak{a}_s(X_s^i)\right|^p\right]ds\\ &\le\left(2^{2p-2}D_\varphi^pB_\alpha^p+2^{2p-2}L_\varphi^pA_\alpha^pF_1^p\right)\int_0^t H(s)ds+2^{2p-2}L_\varphi^pA_\alpha^p\frac{F_2}{N^{p/2+1/2}}t\\ &\quad+2^{p-1}\frac{1}{N}\sum_{i=1}^N\int_0^t\mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i,X_s)\;-\mathfrak{a}_s(X_s^i)\right|^p\right]ds.\end{aligned}\tag{28}$$

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

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

This implies that for $p\ge 2$,

$$\begin{aligned}&\mathrm{E}\sup_{s\in[0,t]}\left|\int_0^s\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N})\;-\mathfrak{b}_s^q(X_s^i)\right)dW_s^i\right|^p\\ &\le C_p^p\mathrm{E}\left[\left(\int_0^t\left|\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N})\;-\mathfrak{b}_s^q(X_s^i)\right)\right|^2ds\right)^{p/2}\right]\\ &\le C_p^pt^{p/2-1}\mathrm{E}\left[\int_0^t\left|\left(\mathfrak{b}_{K,N}^q(X_s^{i,K,N},X_s^{K,N})\;-\mathfrak{b}_s^q(X_s^i)\right)\right|^pds\right]\\ &\le C_p^pt^{p/2-1}\mathrm{E}\left[\int_0^t\left|\left(\mathfrak{b}_{K,N}(X_s^{i,K,N},X_s^{K,N})\;-\mathfrak{b}_s(X_s^i)\right)\right|^pds\right].\end{aligned}\tag{29}$$

Now, completely analogue to the derivation of (28), we get

$$\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^{i,K,N}, X_s^{K,N}) - \mathfrak{b}_s(X_s^i)\right|^p\right] ds$$
$$\leq \left(2^{2p-2}D_\varphi^p B_\beta^p + 2^{2p-2}L_\varphi^p A_\beta^p F_1\right)\int_0^t H(s)ds + 2^{2p-2}L_\varphi^p A_\beta^p \frac{F_2}{N^{p/2+1/2}}t$$
$$+2^{p-1}\frac{1}{N}\sum_{i=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^i, X_s) - \mathfrak{b}_s(X_s^i)\right|^p\right] ds. \tag{30}$$

Now by gathering all together and taking expectations, we arrive at

$$H(t) \leq \left(D_\varphi^p B_\alpha^p T^{p-1} + L_\varphi^p A_\alpha^p F_1^p T^{p-1}\right.$$
$$\left.+C_p^p D_\varphi^p B_\beta^p d^p T^{p/2-1} + C_p^p L_\varphi^p A_\beta^p d^p T^{p/2-1} F_1^p\right) 2^{3p-3}\int_0^t H(s)ds$$
$$+2^{3p-3}\left(L_\varphi^p A_\alpha^p T^p + d^p C_p^p L_\varphi^p A_\beta^p T^{p/2}\right)\frac{F_2}{N^{p/2+1/2}} \tag{31}$$
$$+2^{2p-2}T^{p-1}\frac{1}{N}\sum_{j=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i)\right|^p\right] ds$$
$$+2^{2p-2}d^p C_p^p T^{p/2-1}\frac{1}{N}\sum_{j=1}^{N}\int_0^t \mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^i, X_s) - \mathfrak{b}_s(X_s^i)\right|^p\right] ds.$$

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,

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

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

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

using (6). 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 $p \geq 2$,

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

So for any $p \geq 2$,

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

hence

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

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,

$$\begin{aligned}
\mathrm{E}\left[\left|\mathfrak{a}_{K,N}(X_s^i, X_s) - \mathfrak{a}_s(X_s^i)\right|^p\right] &\leq 2^{p-1}\left(C_p^{(2)}\right)^p A_\alpha^p D_\varphi^p F_3^p N^{-p/2} \\
&\quad + 2^{p-1}F_3^p\left(\sum_{k=K+1}^{\infty} A_{k,\alpha}\left|\gamma_k(s)\right|\right)^p. \qquad (32)
\end{aligned}$$

Analogously we get

$$\mathrm{E}\left[\left|\mathfrak{b}_{K,N}(X_s^i, X_s) - \mathfrak{b}_s(X_s^i)\right|^p\right] \leq 2^{p-1}\left(C_p^{(2)}\right)^p A_\beta^p D_\varphi^p F_3^p N^{-p/2} \\ + 2^{p-1} F_3^p \left(\sum_{k=K+1}^{\infty} A_{k,\beta} \left|\gamma_k(s)\right|\right)^p. \tag{33}$$

Now, combining the estimates (32) and (33) with (31), yields for $0 \leq t \leq T$,

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

with abbreviations

$$\begin{aligned} C_{p,\varphi,X} &= 2^{3p-3} D_\varphi^p B_\alpha^p + 2^{3p-3} L_\varphi^p A_\alpha^p F_1^p \\ D_{p,\varphi,X} &= 2^{3p-3} C_p^p D_\varphi^p B_\beta^p + 2^{3p-3} C_p^p L_\varphi^p A_\beta^p F_1^p \\ E_{p,\varphi,X} &= 2^{3p-3} \left(C_p^{(2)}\right)^p A_\alpha^p D_\varphi^p F_3^p \\ F_{p,\varphi,X} &= 2^{3p-3} C_p^p \left(C_p^{(2)}\right)^p A_\beta^p D_\varphi^p F_3^p \\ G_{p,\varphi,X} &= 2^{3p-3} F_3^p \\ H_{p,\varphi,X} &= 2^{3p-3} C_p^p F_3^p. \end{aligned}$$

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 Appendix

6.1 Existence of moments

Theorem 5. *Fix some $p \geq 2$ and suppose that $\mathrm{E}[|X_0|^p] < \infty$. Then it holds under assumptions (AC) and (AF),*

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

Proof. Fix some $i \in \{1, \dots, N\}$ 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

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

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 $R + \|X_0^i\|_p$. On the other hand

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

It follows from the Minkowski and BDG inequality that

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

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

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

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 1 (see Appendix) it follows that

$$\left\| \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \right\|_p \leq 2e^{\left(2A_\alpha D_\varphi + A_\beta^2 D_\varphi^2 d^2 \left(C_p^{BDG}\right)^2\right)T} \times \\ \left(\|X_0\|_p + A_\alpha D_\varphi T + A_\beta D_\varphi d C_p^{BDG}\sqrt{T}\right). \tag{34}$$

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

$$\lim_{R\uparrow\infty} \sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| = \begin{cases} \sup_{s\in[0,T]} \left| X^{i,K,N}_{s} \right| & \text{on } \{\tau_{i,\infty} = \infty\} \\ \infty & \text{on } \{\tau_{i,\infty} \leq T\} \end{cases}. \tag{35}$$

Indeed, on the set $\{\tau_{i,\infty} \leq T\}$ we have for any $R > 0$, $\left| X^{i,K,N}_{\tau_{i,R}} - X_0^i \right| \geq R$ with $\tau_{i,R} \leq T$, so that

$$\sup_{s\in[0,T]} \left| X^{i,K,N}_{s\wedge\tau_{i,R}} \right| \geq \left| X^{i,K,N}_{\tau_{i,R}} \right| \geq \left| X^{i,K,N}_{\tau_{i,R}} \right| \geq R - \left| X_0^i \right|.$$

The Fatou lemma (35) implies (with $0 := \infty \cdot 0$),

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

because of (34). So $P\left(\{\tau_{i,\infty} \leq T\}\right) = 0$, i.e. $\tau_\infty = \infty$ almost surely. Again by the Fatou lemma, (35) then implies

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

because of (34) again. □

The following lemma is consequence of Gronwall's theorem.

Lemma 1. *Let* $f : [0,T] \to \mathbb{R}_+$ *and* $\psi : [0,T] \to \mathbb{R}_+$ *be two non-negative non-decreasing functions satisfying*

$$f(t) \leq A \int_0^t f(s)\, ds + B \left(\int_0^t f^2(s)\, ds \right)^{1/2} + \psi(t), \quad t \in [0,T], \tag{36}$$

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

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

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

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

Plugging this into (36) yields

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

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] John P Boyd. *Chebyshev and Fourier spectral methods*. Courier Corporation, 2001.

[3] TD Frank. Stochastic feedback, nonlinear families of markov processes, and nonlinear fokker–planck equations. *Physica A: Statistical Mechanics and its Applications*, 331(3):391–408, 2004.

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

[5] 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.

[6] 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.

[7] 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.

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

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

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