

Bias behaviour and antithetic sampling in mean-field particle approximations of SDEs nonlinear in the sense of McKean

O. Bencheikh and B. Jourdain*

November 12, 2021

Abstract

In this paper, we prove that the weak error between a stochastic differential equation with nonlinearity in the sense of McKean given by moments and its approximation by the Euler discretization with time-step h of a system of N interacting particles is $\mathcal{O}(N^{-1} + h)$. We provide numerical experiments confirming this behaviour and showing that it extends to more general mean-field interaction and study the efficiency of the antithetic sampling technique on the same examples.

1 Introduction

According to [16], the strong rate of convergence of particle approximations of McKean-Vlasov Stochastic differential equations with Lipschitz coefficients is $\mathcal{O}(N^{-1/2})$ when N denotes the number of particles. This rate is driven by the statistical error and one may wonder whether the bias vanishes quicker. A parallel can be drawn with the time discretization of standard stochastic differential equations where, for Lipschitz coefficients, the strong rate of convergence of the explicit Euler-Maruyama scheme is $\mathcal{O}(\sqrt{h})$ [12] with $h > 0$ denoting the time-step. Using the Feynman-Kac partial differential equation associated with the stochastic differential equation, Talay and Tubaro [17] checked that, for smooth coefficients, the weak error behaves in $\mathcal{O}(h)$ and can even be expanded in powers of h . In the context of particles interacting through jumps, the $\mathcal{O}(N^{-1})$ behaviour of the bias is known. According to [6], for particle approximations of generalized Boltzmann equations, the total variation distance between the law of the path of a particle and the one of the limiting nonlinear Boltzmann process behaves in $\mathcal{O}(N^{-1})$. For Feynman-Kac particle models, expansions in powers of $1/N$ are obtained in [4].

The interest in the bias introduced by particle approximations is motivated by numerical efficiency. Indeed, the numerical experiments performed in Section 3 show a general $\mathcal{O}(N^{-1})$ behaviour of the bias, even in models with not so smooth coefficients. Under this behaviour, simulating \sqrt{N} independent copies of the system with \sqrt{N} particles leads to the same order of error (bias with the same order as the $\mathcal{O}(N^{-1/4} \times N^{-1/4}) = \mathcal{O}(N^{-1/2})$ statistical error) as one simulation of the system with N particles (bias smaller than the $\mathcal{O}(N^{-1/2})$ statistical error). And the former approach is less expensive than the latter as soon as the computational cost of the particle system grows more than linearly with the number of particles. The behaviour of the bias is also of interest in order to adapt to the number of particles the multilevel Monte Carlo method introduced by Giles [7] in the context of time discretization of standard stochastic differential equations. In [10], Haji-Ali and Tempone combine both discretizations through the Multi-index Monte Carlo method. In this perspective, another interesting question is the possibility to take advantage of the antithetic sampling technique introduced in [9] to reduce the variance (see [3, 1] or [8] p57 for the use of this technique in nested Monte Carlo computations). Does the variance of the difference between the empirical mean of the system with $2N$ particles driven by the i.i.d. couples $(X_0^i, W^i)_{1 \leq i \leq 2N}$ (X_0^i and W^i respectively denote the initial condition and the Brownian motion of the i -th particle) and the mean of the empirical means of the two independent systems with N particles respectively driven by $(X_0^i, W^i)_{1 \leq i \leq N}$ and $(X_0^i, W^i)_{N+1 \leq i \leq 2N}$ converge quicker to 0 than $\mathcal{O}(N^{-1})$?

In this paper, using the Feynman-Kac partial differential equation associated with the limiting nonlinear process and its moments, we prove the respective $\mathcal{O}(N^{-1})$ and $\mathcal{O}(N^{-1} + h)$ behaviour of the bias for systems of diffusive

*Université Paris-Est, Cermics (ENPC), INRIA, F-77455 Marne-la-Vallée, France. E-mails : benjamin.jourdain@enpc.fr, oumaïma.bencheikh@enpc.fr

particles interacting through moments and their Euler discretization with time step h . Of course, the computational cost of such systems is linear in the number N of particles and the above numerical motivation is not valid. Nevertheless this result can be seen as a first step before addressing more general interactions which could necessitate more advanced tools like the master equation for mean-field games introduced by Lions in his seminal lectures at *Collège de France* and studied in [2] from a probabilistic point of view. Theorem 3.2 [13] is proved using the master equation and implies the $\mathcal{O}(N^{-1})$ behaviour of the bias for one-dimensional stochastic differential equations with general interaction in the drift coefficient but no interaction in the diffusion coefficient. We provide numerical experiments showing that the $\mathcal{O}(N^{-1})$ behaviour holds in more general situations including ones with non smooth coefficients. Last, on the same examples, we check that the antithetic variance does not generally decrease quicker than $\mathcal{O}(N^{-1})$.

2 Estimation of the bias for systems of particles interacting through moments

Let $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^p$ be Lipchitz continuous and $\sigma = (\sigma_j^l)_{1 \leq j \leq n, 1 \leq l \leq d} : [0, T] \times \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$, $b = (b_j)_{1 \leq j \leq n} : [0, T] \times \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be Lipschitz continuous in their $p + n$ last components uniformly in their first component and such that $\sup_{t \in [0, T]} (|\sigma(t, 0, 0)| + |b(t, 0, 0)|) < \infty$. We consider the stochastic differential equation in dimension n nonlinear in the sense of McKean

$$(2.1) \quad X_t = X_0 + \int_0^t \sigma(s, \mathbb{E}[\alpha(X_s)], X_s) dW_s + \int_0^t b(s, \mathbb{E}[\alpha(X_s)], X_s) ds, \quad t \in [0, T]$$

where X_0 is some square integrable \mathbb{R}^n -valued random variable independent from the d -dimensional Brownian motion $(W_t)_{t \geq 0}$. The drift and diffusion coefficient at time s depend on the law of X_s through the moments $\mathbb{E}[\alpha(X_s)]$. Since σ may only depend on a subset of coordinates of this expectation on b on another subset, moments in drift and diffusion can differ as well as their respective dimensions. By a solution, we mean a continuous process $(X_t)_{t \in [0, T]}$ adapted to the filtration generated by the Brownian motion W and X_0 such that $\sup_{t \in [0, T]} \mathbb{E}[|\alpha(X_t)|] < \infty$ and the above equation holds with the integrals with respect to dW_s and ds making sense. Notice that for any solution, $x \mapsto (\sigma(s, \mathbb{E}[\alpha(X_s)], x), b(s, \mathbb{E}[\alpha(X_s)], x))$ has affine growth uniformly for $s \in [0, T]$. With the square integrability of X_0 , this implies that $\mathbb{E} \left[\sup_{t \in [0, T]} |X_t|^2 \right] + \sup_{0 \leq s < t \leq T} \frac{\mathbb{E}[|X_t - X_s|^2]}{t - s} < \infty$ so that, by Lipschitz continuity of α , $t \mapsto \mathbb{E}[\alpha(X_t)]$ is Hölder continuous with exponent $1/2$ on $[0, T]$.

The particle approximation of the SDE nonlinear in the sense of McKean (2.1) is given by the system with mean-field interaction

$$(2.2) \quad X_t^{i, N} = X_0^i + \int_0^t \sigma \left(s, \frac{1}{N} \sum_{j=1}^N \alpha(X_s^{j, N}), X_s^{i, N} \right) dW_s^i + \int_0^t b \left(s, \frac{1}{N} \sum_{j=1}^N \alpha(X_s^{j, N}), X_s^{i, N} \right) ds, \quad 1 \leq i \leq N, \quad t \in [0, T],$$

with $((X_0^i, W^i))_{i \geq 1}$ i.i.d. copies of (X_0, W) . By the Lipschitz assumptions, existence and trajectorial uniqueness hold for this $N \times n$ dimensional equation. The Yamada-Watanabe theorem ensures weak uniqueness and therefore exchangeability of the particles $((X_t^{i, N})_{t \in [0, T]})_{1 \leq i \leq N}$. Let us also introduce the Euler discretizations with time-step $h > 0$ of the SDE (2.1) and the particle system :

$$(2.3) \quad \begin{aligned} X_t^h &= X_0 + \int_0^t \sigma(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^h)], X_{\tau_s^h}^h) dW_s + \int_0^t b(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^h)], X_{\tau_s^h}^h) ds, \quad t \in [0, T] \text{ where } \tau_s^h = \lfloor s/h \rfloor h, \\ X_t^{i, N, h} &= X_0^i + \int_0^t \sigma \left(\tau_s^h, \frac{1}{N} \sum_{j=1}^N \alpha(X_{\tau_s^h}^{j, N, h}), X_{\tau_s^h}^{i, N, h} \right) dW_s^i + \int_0^t b \left(\tau_s^h, \frac{1}{N} \sum_{j=1}^N \alpha(X_{\tau_s^h}^{j, N, h}), X_{\tau_s^h}^{i, N, h} \right) ds, \quad 1 \leq i \leq N, \quad t \in [0, T]. \end{aligned}$$

It is natural and convenient to consider that $\tau_s^0 = s$, $(X_t^0)_{t \in [0, T]} = (X_t)_{t \in [0, T]}$ and $(X_t^{i, N, 0})_{t \in [0, T], 1 \leq i \leq N} = (X_t^{i, N})_{t \in [0, T], 1 \leq i \leq N}$ and we use these notations in what follows.

Reasoning like in the laboratory example in [16] or in Theorem 1.3 [11], one easily checks the following result.

Proposition 2.1. *Strong existence and trajectorial uniqueness hold for the SDE nonlinear in the sense of McKean (2.1) and its Euler discretization (2.3). Moreover $\sup_{h \geq 0} \mathbb{E} \left[\sup_{t \in [0, T]} |X_t^h|^2 \right] \leq C(1 + \mathbb{E}[|X_0|^2])$ where the finite constant C does not depend on X_0 . Last, if for $i \in \mathbb{N}^*$, $(X_t^{i, h})_{t \in [0, T]}$ denotes the process obtained by replacement of (X_0, W) by (X_0^i, W^i) in (2.3), then*

$$\exists C < \infty, \forall N \in \mathbb{N}^*, \sup_{h \geq 0} \max_{1 \leq i \leq N} \mathbb{E} \left[\sup_{t \in [0, T]} |X_t^{i, N, h} - X_t^{i, h}|^2 \right] \leq C \frac{1 + \mathbb{E}[|X_0|^2]}{N}.$$

If for $i \in \mathbb{N}^*$, $(X_t^i)_{t \in [0, T]}$ denotes the process obtained by replacement of (X_0, W) by (X_0^i, W^i) in (2.1), this implies the following estimation of the bias introduced by the particle discretization: for any function $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ Lipschitz with constant L_ψ and in particular for each coordinate of α ,

$$\begin{aligned} \forall s \in [0, T], |\mathbb{E}[\psi(X_s^{1, N})] - \mathbb{E}[\psi(X_s)]| &= |\mathbb{E}[\psi(X_s^{1, N})] - \mathbb{E}[\psi(X_s^1)]| \leq L_\psi \mathbb{E}[|X_s^{1, N} - X_s^1|] \\ &\leq L_\psi \left(\sup_{h \geq 0} \mathbb{E} \left[\sup_{t \in [0, T]} |X_t^{1, N, h} - X_t^{1, h}|^2 \right] \right)^{1/2} \leq \sqrt{C} L_\psi \frac{\sqrt{1 + \mathbb{E}[|X_0|^2]}}{\sqrt{N}}. \end{aligned}$$

The first inequality is crude since it prevents cancelations in average and one may wonder whether the bias converges faster to 0 as $N \rightarrow \infty$. Under additional regularity, the answer is positive, which is our main result.

Theorem 2.2. *Assume that*

- σ is Lipschitz continuous in its $p + n$ last components uniformly in its first component and such that $\sup_{t \in [0, T]} (|\sigma(t, 0, 0)| + |b(t, 0, 0)|) < \infty$,
- α, ψ are two times continuously differentiable with bounded derivatives up to the order two and Lipschitz continuous second order derivatives,
- $a = \sigma \sigma^*$ and b are globally Lipschitz continuous, continuously differentiable with respect to their variables with index in $\{2, \dots, 1 + p\}$ with derivatives Lipschitz continuous with respect to their $p + n$ last variables uniformly in their first variable,
- there exists $\tilde{d} \in \mathbb{N}^*$, $\tilde{\sigma} : [0, T] \times \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times \tilde{d}}$ such that for all $(t, y, x) \in [0, T] \times \mathbb{R}^p \times \mathbb{R}^n$, $a(t, y, x) = \tilde{\sigma} \tilde{\sigma}^*(t, y, x)$ and $\tilde{\sigma}, b$ are continuous and two times continuously differentiable with respect to their n last variables with bounded derivatives up to the order two and second order derivatives Lipschitz continuous with respect to their n last variables uniformly in their $1 + p$ first variables,

Then

$$\exists C < \infty, \forall h \geq 0, \forall N \in \mathbb{N}^*, \sup_{t \in [0, T]} |\mathbb{E}[\psi(X_t^{1, N, h})] - \mathbb{E}[\psi(X_t)]| \leq C \left(\frac{1}{N} + h \right).$$

The idea of the proof is, like in [17], to use the Feynman-Kac partial differential equation associated with the nonlinear SDE (2.1) to first check that the estimation holds for the coordinates of α before concluding that it holds for each test function ψ . The following proposition ensures existence and smoothness to this Feynman-Kac PDE which only depends on σ through $a = \sigma \sigma^*$. In its statement and its proof based on a stochastic flow approach, we take advantage of the flexibility given by the choice of a square root of a which explains the last assumption in Theorem 2.2. Let $(\tilde{W}_t)_{t \geq 0}$ be a \tilde{d} -dimensional Brownian motion with coordinates $(\tilde{W}_t^i)_{1 \leq i \leq \tilde{d}}$ and for $(s, x) \in [0, T] \times \mathbb{R}^n$, $(\tilde{X}_t^{s, x})_{t \in [s, T]}$ denote the solution to

$$(2.4) \quad \tilde{X}_t^{s, x} = x + \int_s^t \tilde{\sigma}(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s, x}) d\tilde{W}_r + \int_s^t b(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s, x}) dr, \quad t \in [s, T]$$

where the coefficients depend on $\mathbb{E}[\alpha(X_r)]$ and not $\mathbb{E}[\alpha(\tilde{X}_r^{s, x})]$.

Proposition 2.3. Let $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ be two times continuously differentiable with bounded derivatives up to the order two and Lipschitz continuous second order derivatives. Under the last assumption in Theorem 2.2, for each $t \in [0, T]$, the function $[0, t] \times \mathbb{R}^n \ni (s, x) \mapsto u_{t,\psi}(s, x) := \mathbb{E}[\psi(\tilde{X}_t^{s,x})]$ is once (resp. twice) continuously differentiable with respect to the time (resp. space) variable s (resp. x) and such that $u_{t,\psi}$ together with its spatial derivatives up to the order two are Lipschitz continuous in x uniformly in $0 \leq s \leq t \leq T$. Moreover, $u_{t,\psi}$ is a classical solution to the Feynman-Kac PDE

$$(2.5) \quad \begin{cases} \partial_s u_{t,\psi}(s, x) + \frac{1}{2} \sum_{j,k=1}^n a_{jk}(s, \mathbb{E}[\alpha(X_s)], x) \partial_{jk} u_{t,\psi}(s, x) + \sum_{j=1}^n b_j(s, \mathbb{E}[\alpha(X_s)], x) \partial_j u_{t,\psi}(s, x) = 0, & (s, x) \in [0, t] \times \mathbb{R}^n \\ u_{t,\psi}(t, x) = \psi(x), & x \in \mathbb{R}^n \end{cases}$$

Here ∂_j and ∂_{jk} denote the partial derivative with respect to the j -th coordinate of x and with respect to its j -th and k -th coordinates.

Proof. Even if this result seems to be well-known, we could not find its proof in the literature. Therefore, we are going to give a sketch of its proof. For notational simplicity, we set $\tilde{\sigma}^0 = b$ and $\tilde{W}_t^0 = t$.

Let for $j \in \{1, \dots, n\}$, e_j denote the j -th vector of the canonical basis of \mathbb{R}^n and for $l \in \{0, \dots, \tilde{d}\}$, $\partial \tilde{\sigma}^l = (\partial_{1+p+k} \tilde{\sigma}_j^l)_{1 \leq j, k \leq n}$ where ∂_{1+p+k} is the partial derivative with respect to the $(1+p+k)$ -th variable. Let also for $l \in \{0, \dots, \tilde{d}\}$ and $y, z \in \mathbb{R}^n$, $\partial^2 \tilde{\sigma}^l yz \in \mathbb{R}^n$ be defined by $(\partial^2 \tilde{\sigma}^l yz)_j = \sum_{k,m=1}^n (\partial_{1+p+k} \partial_{1+p+m} \tilde{\sigma}_j^l) y_k z_m$ for $j \in \{1, \dots, n\}$. Since $x \mapsto [\tilde{\sigma}, b](r, \mathbb{E}[\alpha(X_r)], x)$ has affine growth uniformly in $r \in [0, T]$, standard moment estimations ensure that

$$\forall q \geq 1, \exists C < \infty, \forall (s, x) \in [0, T] \times \mathbb{R}^n, \mathbb{E} \left[\sup_{r \in [s, T]} |\tilde{X}_r^{s,x}|^q \right] \leq C(1 + |x|^q).$$

By [15], Theorem 3.3 p223 and its proof, for $r \in [s, T]$, $x \mapsto \tilde{X}_r^{s,x}$ is twice continuously differentiable \mathbb{P} -almost surely with derivatives satisfying for $j, k \in \{1, \dots, n\}$

$$\begin{aligned} d\tilde{X}_r^{s,x} &= \sum_{l=0}^{\tilde{d}} \partial \tilde{\sigma}^l(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s,x}) \partial_j \tilde{X}_r^{s,x} d\tilde{W}_r^l, \quad r \in [s, T], \quad \partial_j \tilde{X}_s^{s,x} = e_j \\ d\partial_{jk} \tilde{X}_r^{s,x} &= \sum_{l=0}^{\tilde{d}} \left(\partial \tilde{\sigma}^l(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s,x}) \partial_{jk} \tilde{X}_r^{s,x} + \partial^2 \tilde{\sigma}^l(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s,x}) \partial_j \tilde{X}_r^{s,x} \partial_k \tilde{X}_r^{s,x} \right) d\tilde{W}_r^l, \quad r \in [s, T], \quad \partial_{jk} \tilde{X}_s^{s,x} = 0. \end{aligned}$$

Moreover, for any $q \in [1, \infty)$,

$$\sup_{0 \leq s \leq r \leq T} \sup_{x \neq y \in \mathbb{R}^n} \sum_{j=1}^n \mathbb{E} \left[|\partial_j \tilde{X}_r^{s,x}|^q + \sum_{k=1}^n \left(|\partial_{jk} \tilde{X}_r^{s,x}|^q + \frac{|\partial_{jk} \tilde{X}_r^{s,x} - \partial_{jk} \tilde{X}_r^{s,y}|^q}{|x - y|^q} \right) \right] < \infty$$

and $\exists C < \infty, \forall x \in \mathbb{R}^n, \forall 0 \leq s \leq \tilde{s} \leq r \leq T$,

$$\mathbb{E} \left[|\tilde{X}_r^{s,x} - \tilde{X}_r^{\tilde{s},x}|^q + \sum_{j=1}^n \left(|\partial_j \tilde{X}_r^{s,x} - \partial_j \tilde{X}_r^{\tilde{s},x}|^q + \sum_{k=1}^n |\partial_{jk} \tilde{X}_r^{s,x} - \partial_{jk} \tilde{X}_r^{\tilde{s},x}|^q \right) \right] \leq C(1 + |x|^q)(\tilde{s} - s)^{q/2}.$$

These properties ensure that $x \mapsto u_{t,\psi}(s, x)$ is two times continuously differentiable with derivatives $\partial_j u_{t,\psi}(s, x) = \mathbb{E}[\nabla \psi(\tilde{X}_t^{s,x}) \cdot \partial_j \tilde{X}_t^{s,x}]$ and $\partial_{jk} u_{t,\psi}(s, x) = \mathbb{E}[\nabla \psi(\tilde{X}_t^{s,x}) \cdot \partial_{jk} \tilde{X}_t^{s,x} + \partial_k \tilde{X}_t^{s,x} \cdot \nabla^2 \psi(\tilde{X}_t^{s,x}) \partial_j \tilde{X}_t^{s,x}]$ bounded and Lipschitz continuous in x uniformly in $0 \leq s \leq t \leq T$. Moreover, $u_{t,\psi}$ and its spatial derivatives up to the order two are continuous in the time variable. Let us now check that $u_{t,\psi}$ satisfies the Feynman-Kac PDE (2.5). Let $s \in (0, t]$ and $\varepsilon \in (0, s]$. By the flow property stated in Theorem 3.3 [15], $u_{t,\psi}(s - \varepsilon, x) = \mathbb{E}[u_{t,\psi}(s, \tilde{X}_s^{s-\varepsilon, x})]$. On the other hand, by Taylor expansion,

$$\begin{aligned} u_{t,\psi}(s, \tilde{X}_s^{s-\varepsilon, x}) - u_{t,\psi}(s, x) &= \nabla_x u_{t,\psi}(s, x) \cdot \left(\int_{s-\varepsilon}^s \tilde{\sigma}(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon, x}) d\tilde{W}_r + \int_{s-\varepsilon}^s b(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon, x}) dr \right) \\ &\quad + \frac{1}{2} \sum_{j,k=1}^n \partial_{jk} u_{t,\psi}(s, x) \left(\sum_{l=0}^{\tilde{d}} \int_{s-\varepsilon}^s \tilde{\sigma}_j^l(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon, x}) d\tilde{W}_r^l \right) \left(\sum_{m=0}^{\tilde{d}} \int_{s-\varepsilon}^s \tilde{\sigma}_k^m(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon, x}) d\tilde{W}_r^m \right) + R_\varepsilon \end{aligned}$$

where R_ε is some random reminder such that $|R_\varepsilon| \leq C|\tilde{X}_r^{s-\varepsilon,x} - x|^3$ with C a deterministic finite constant. Since $\mathbb{E} \left[\int_{s-\varepsilon}^s \tilde{\sigma}(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) d\tilde{W}_r \right] = 0$ and, by Itô's isometry,

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{l=1}^{\tilde{d}} \int_{s-\varepsilon}^s \tilde{\sigma}_j^l(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) d\tilde{W}_r^l \right) \left(\sum_{m=1}^{\tilde{d}} \int_{s-\varepsilon}^s \tilde{\sigma}_k^m(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) d\tilde{W}_r^m \right) \right] \\ = \mathbb{E} \left[\int_{s-\varepsilon}^s a_{jk}(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) dr \right], \end{aligned}$$

we deduce that

$$\begin{aligned} & \left| \frac{u_{t,\psi}(s, x) - u_{t,\psi}(s - \varepsilon, x)}{\varepsilon} + \nabla_x u_{t,\psi}(s, x) \cdot b(s, \mathbb{E}[\alpha(X_s)], x) + \frac{1}{2} \sum_{j,k=1}^n \partial_{jk} u_{t,\psi}(s, x) a_{jk}(s, \mathbb{E}[\alpha(X_s)], x) \right| \\ & \leq C \left(\left| b(s, \mathbb{E}[\alpha(X_s)], x) - \frac{1}{\varepsilon} \mathbb{E} \left[\int_{s-\varepsilon}^s b(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) dr \right] \right| + \left| a(s, \mathbb{E}[\alpha(X_s)], x) - \frac{1}{\varepsilon} \mathbb{E} \left[\int_{s-\varepsilon}^s a(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) dr \right] \right| \right) \\ & + \frac{\mathbb{E}[|\tilde{X}_s^{s-\varepsilon,x} - x|^3]}{\varepsilon} + \frac{1}{\varepsilon} \sum_{j,k=1}^n \mathbb{E} \left[\left| \int_{s-\varepsilon}^s b_j(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) dr \sum_{m=0}^{\tilde{d}} \int_{s-\varepsilon}^s \tilde{\sigma}_k^m(r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x}) d\tilde{W}_r^m \right| \right]. \end{aligned}$$

By continuity and uniform integrability of $r \mapsto [\tilde{\sigma}, b](r, \mathbb{E}[\alpha(X_r)], \tilde{X}_r^{s-\varepsilon,x})$ on $[s-\varepsilon, T]$, the two first terms in the right-hand side converge to 0 as $\varepsilon \rightarrow 0$. Moreover the expectations in the two last terms are smaller than $C\varepsilon^{3/2}$. Taking into account the continuity with respect to s of $\nabla_x u_{t,\psi}(s, x) \cdot b(s, \mathbb{E}[\alpha(X_s)], x) + \frac{1}{2} \sum_{j,k=1}^n \partial_{jk} u_{t,\psi}(s, x) a_{jk}(s, \mathbb{E}[\alpha(X_s)], x)$, we conclude that $u_{t,\psi}$ is a classical solution of the Feynman-Kac PDE (2.5). \square

We are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. Let $t \in [0, T]$. Applying Itô's formula and taking into account the Feynman-Kac PDE (2.5), we obtain that

$$\begin{aligned} u_{t,\psi}(t, X_t^1) - u_{t,\psi}(t, X_t^{1,N,h}) &= u_{t,\psi}(0, X_0^1) - u_{t,\psi}(0, X_0^1) + \int_0^t \nabla_x u_{t,\psi}(s, X_s^1) \cdot \sigma(s, \mathbb{E}[\alpha(X_s)], X_s^1) dW_s^1 \\ &\quad - \int_0^t \nabla_x u_{t,\psi}(s, X_s^{1,N,h}) \cdot \sigma \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) dW_s^1 \\ &\quad + \int_0^t \nabla_x u_{t,\psi}(s, X_s^{1,N,h}) \cdot \left(b(s, \mathbb{E}[\alpha(X_s)], X_s^{1,N,h}) - b \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) \right) ds \\ &\quad + \frac{1}{2} \sum_{j,k=1}^n \int_0^t \partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(a_{jk}(s, \mathbb{E}[\alpha(X_s)], X_s^{1,N,h}) - a_{jk} \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) \right) ds. \end{aligned}$$

Integrability deduced from Propositions 2.1 and 2.3 and the properties of σ ensure that the expectations of the stochastic integrals vanish. Therefore, setting for $s \in [0, t]$

$$\begin{aligned} e^\psi(s) &:= \mathbb{E} \left[\nabla_x u_{t,\psi}(s, X_s^{1,N,h}) \cdot \left(b(s, \mathbb{E}[\alpha(X_s)], X_s^{1,N,h}) - b \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) \right) \right] \\ e_{jk}^\psi(s) &:= \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(a_{jk}(s, \mathbb{E}[\alpha(X_s)], X_s^{1,N,h}) - a_{jk} \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) \right) \right], \quad 1 \leq j, k \leq n, \end{aligned}$$

and using that $u_{t,\psi}(t, \cdot) = \psi(\cdot)$, we obtain

$$(2.6) \quad \mathbb{E}[\psi(X_t)] - \mathbb{E}[\psi(X_t^{1,N,h})] = \int_0^t \left(e^\psi(s) + \frac{1}{2} \sum_{j,k=1}^n e_{jk}^\psi(s) \right) ds.$$

Let us now estimate $e_{jk}^\psi(s)$:

$$\begin{aligned}
e_{jk}^\psi(s) = & \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(a_{jk}(s, \mathbb{E}[\alpha(X_s)], X_s^{1,N,h}) - a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_s^{1,N,h}) \right) \right] \\
& + \mathbb{E} \left[\left(\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) - \partial_{jk} u_{t,\psi}(s, X_{\tau_s^h}^{1,N,h}) \right) \left(a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_s^{1,N,h}) - a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_{\tau_s^h}^{1,N,h}) \right) \right] \\
& + \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_{\tau_s^h}^{1,N,h}) \left(a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_s^{1,N,h}) - a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_{\tau_s^h}^{1,N,h}) \right) \right] \\
& + \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_{\tau_s^h}^{1,N,h}) - a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h}) \right) \right] \\
& + \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h}) - a_{jk} \left(\tau_s^h, \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) \right) \right].
\end{aligned}$$

Let us respectively denote by $e_{jk}^{1,\psi}(s)$, $e_{jk}^{2,\psi}(s)$, $e_{jk}^{3,\psi}(s)$, $e_{jk}^{4,\psi}(s)$ and $\bar{e}_{jk}^\psi(s)$ the five terms in the right-hand side. Since, by Proposition 2.3, $\partial_{jk} u_{t,\psi}$ is bounded by a finite constant $M_\psi^{(2)}$ not depending in t and a is Lipschitz continuous with constant L_a , $|e_{jk}^{1,\psi}(s)| \leq M_\psi^{(2)} L_a ((s - \tau_s^h) + |\mathbb{E}[\alpha(X_s)] - \mathbb{E}[\alpha(X_{\tau_s^h})]|)$. Since α is \mathcal{C}^2 with bounded derivatives and Lipschitz continuous second order derivatives, for $0 \leq r \leq s \leq T$, computing $\alpha(X_s) - \alpha(X_r)$ by Itô's formula, taking expectations and remarking that the expectation of the stochastic integral vanishes, we obtain the existence of a finite constant C not depending on r and s such that $|\mathbb{E}[\alpha(X_s)] - \mathbb{E}[\alpha(X_r)]| \leq C(s - r)$. Hence $|e_{jk}^{1,\psi}(s)| \leq M_\psi^{(2)} L_a(1 + C)(s - \tau_s^h) \leq M_\psi^{(2)} L_a(1 + C)h$.

Since, by Proposition 2.3, $\partial_{jk} u_{t,\psi}$ is Lipschitz continuous in space with constant $L_\psi^{(2)}$ not depending on t and a_{jk} is Lipschitz continuous in space with constant L_a we have $|e_{jk}^{2,\psi}(s)| \leq L_\psi^{(2)} L_a \mathbb{E}[|X_s^{1,N,h} - X_{\tau_s^h}^{1,N,h}|^2]$. Combining the fact that $\sup_{N \geq 1, h \geq 0, s \in [0, T]} \mathbb{E}[|X_s^{1,N,h}|^2] < \infty$ deduced from Proposition 2.1, the Lipschitz continuity of α and the affine growth of σ, b in their $p + n$ last variables uniform in their first variable, one easily checks that there is a finite constant C not depending on N and h such that for all $0 \leq r \leq s \leq T$, $\mathbb{E}[|X_s^{1,N,h} - X_r^{1,N,h}|^2] \leq C(s - r)$ and deduce that $|e_{jk}^{2,\psi}(s)| \leq L_\psi^{(2)} L_a C h$.

Remarking that the first and second order derivatives of a_{jk} with respect to its n last variables have affine growth under our assumptions, computing $a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_s^{1,N,h}) - a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h})], X_{\tau_s^h}^{1,N,h})$ by Itô's formula, multiplying by the random variable $\partial_{jk} u_{t,\psi}(s, X_{\tau_s^h}^{1,N,h})$ bounded by $M_\psi^{(2)}$, taking expectations and remarking that the contribution of the stochastic integral vanishes, we obtain that $|e_{jk}^{3,\psi}(s)| \leq C h$ with C not depending on N, s, t, h . Last, $|e_{jk}^{4,\psi}(s)| \leq M_\psi^{(2)} L_a |\mathbb{E}[\alpha(X_{\tau_s^h})] - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})]|$. Hence there is a finite constant C not depending on N, s, t, h such that

$$(2.7) \quad \forall 0 \leq s \leq t \leq T, |e_{jk}^\psi(s)| \leq C \left(h + \left| \mathbb{E}[\alpha(X_{\tau_s^h})] - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] \right| \right) + |\bar{e}_{jk}^\psi(s)|.$$

Let us now estimate $|\bar{e}_{jk}^\psi(s)|$. Denoting by $\nabla_2 a_{jk}$ the partial gradient of a with respect to its variables with indices in $\{2, \dots, 1 + p\}$, we have

$$\begin{aligned}
\bar{e}_{jk}^\psi(s) = & \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \left(\mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] - \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}) \right) \right. \\
& \cdot \int_0^1 \left\{ \nabla_2 a_{jk} \left(\tau_s^h, v \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] + \frac{1-v}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}), X_{\tau_s^h}^{1,N,h} \right) - \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h}) \right\} dv \Bigg] \\
& + \mathbb{E} \left[\left(\mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] - \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}) \right) \cdot \partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h}) \right].
\end{aligned}$$

Let us respectively denote by $e_{jk}^{5,\psi}(s)$ and $e_{jk}^{6,\psi}(s)$ the two terms in the right-hand side. By Proposition 2.3, $\partial_{jk} u_{t,\psi}$ is bounded by $M_\psi^{(2)}$ and Lipschitz continuous in space with constant $L_\psi^{(2)}$ with $(M_\psi^{(2)}, L_\psi^{(2)})$ not depending on t .

By assumption $\nabla_2 a_{jk}$ is bounded by $M_a^{(2)}$ and Lipschitz continuous in its $p + n$ last variables with constant $L_a^{(2)}$. Therefore,

$$\begin{aligned} \frac{1}{M_\psi^{(2)} L_a^{(2)}} |e_{jk}^{5,\psi}(s)| &\leq \frac{1}{2} \mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}) - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] \right|^2 \right] \\ &\leq \mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N (\alpha(X_{\tau_s^h}^{i,N,h}) - \alpha(X_{\tau_s^h}^{i,h})) - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h}) - \alpha(X_{\tau_s^h}^{1,h})] \right|^2 \right] + \mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,h}) - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,h})] \right|^2 \right]. \end{aligned}$$

The second term in the right-hand side of the inequality is the variance of the empirical mean of i.i.d. random variables. It is therefore equal to $\frac{1}{N} \left(\mathbb{E}[|\alpha(X_{\tau_s^h}^{1,h})|^2] - |\mathbb{E}[\alpha(X_{\tau_s^h}^{1,h})]|^2 \right)$. The first term is also a variance and we upper-bound it by the corresponding second order moment, which is not greater than $\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\left| \alpha(X_{\tau_s^h}^{i,h}) - \alpha(X_{\tau_s^h}^{i,N,h}) \right|^2 \right]$ according to Jensen's inequality for the empirical mean. Therefore, denoting by L_α the Lipschitz constant of α , we have

(2.8)

$$|e_{jk}^{5,\psi}(s)| \leq M_\psi^{(2)} L_a^{(2)} \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\left| \alpha(X_{\tau_s^h}^{i,h}) - \alpha(X_{\tau_s^h}^{i,N,h}) \right|^2 \right] + \frac{1}{N} \left(\mathbb{E}[|\alpha(X_{\tau_s^h}^{1,h})|^2] - |\mathbb{E}[\alpha(X_{\tau_s^h}^{1,h})]|^2 \right) \right) \leq L_\alpha^2 M_\psi^{(2)} L_a^{(2)} \frac{C}{N},$$

with C finite and not depending on N, s, t according to Proposition 2.1. Since $\mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] - \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h})$ is centered and by exchangeability of $(X_{\tau_s^h}^{i,N,h})_{1 \leq i \leq N}$, we may replace $\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h})$ by $\frac{1}{N} \sum_{i=1}^N \partial_{jk} u_{t,\psi}(s, X_s^{i,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{i,N,h}) - \mathbb{E}[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h})]$ in the expectation defining $e_{jk}^{6,\psi}(s)$. With the Cauchy-Schwarz inequality, we deduce that

$$\begin{aligned} |e_{jk}^{6,\psi}(s)| &\leq \left(\mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N \alpha(X_{\tau_s^h}^{i,N,h}) - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] \right|^2 \right] \right)^{1/2} \left(\mathbb{E} \left[\left| \frac{1}{N} \sum_{i=1}^N \partial_{jk} u_{t,\psi}(s, X_s^{i,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{i,N,h}) \right. \right. \right. \\ &\quad \left. \left. \left. - \mathbb{E} \left[\partial_{jk} u_{t,\psi}(s, X_s^{1,N,h}) \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], X_{\tau_s^h}^{1,N,h}) \right] \right|^2 \right] \right)^{1/2}. \end{aligned}$$

Since $\mathbb{R}^n \ni x \mapsto \partial_{jk} u_{t,\psi}(s, x)$ (resp. $\mathbb{R}^n \ni x \mapsto \nabla_2 a_{jk}(\tau_s^h, \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})], x)$) is bounded by $M_\psi^{(2)}$ (resp. $M_a^{(2)}$) and Lipschitz continuous with constant $L_\psi^{(2)}$ (resp. $L_a^{(2)}$), reasoning like in the derivation of (2.8), we obtain that the second factor in the right-hand side is smaller than $(M_\psi^{(2)} L_a^{(2)} + M_a^{(2)} L_\psi^{(2)}) \frac{C}{\sqrt{N}}$ so that $|e_{jk}^{6,\psi}(s)| \leq L_\alpha (M_\psi^{(2)} L_a^{(2)} + M_a^{(2)} L_\psi^{(2)}) \frac{C}{N}$ with C finite and not depending on N, s, t, ψ . With (2.7), we deduce that

$$\forall 0 \leq s \leq t \leq T, |e_{jk}^\psi(s)| \leq C \left(\frac{1}{N} + h + \left| \mathbb{E}[\alpha(X_{\tau_s^h})] - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] \right| \right).$$

Estimating $e^\psi(s)$ in a similar way and using (2.6), we deduce the existence of a finite constant C not depending on N and h such that

$$(2.9) \quad \forall t \in [0, T], \sup_{s \in [0, t]} |\mathbb{E}[\psi(X_s)] - \mathbb{E}[\psi(X_s^{1,N,h})]| \leq C \left(\frac{1}{N} + h \right) + C \int_0^t \left| \mathbb{E}[\alpha(X_{\tau_s^h})] - \mathbb{E}[\alpha(X_{\tau_s^h}^{1,N,h})] \right| ds.$$

Summing for $j \in \{1, \dots, p\}$ this inequality applied with ψ equal to the j -th coordinate of α , remarking that Proposition 2.1 and the Lipschitz continuity of α ensure that $\sup_{t \in [0, T]} \left| \mathbb{E}[\alpha(X_t^{i,N,h})] - \mathbb{E}[\alpha(X_t)] \right| < \infty$ and applying Gronwall's lemma, we deduce that $\sup_{t \in [0, T]} \left| \mathbb{E}[\alpha(X_t)] - \mathbb{E}[\alpha(X_t^{1,N,h})] \right| \leq C \left(\frac{1}{N} + h \right)$. We conclude by combining this estimation and (2.9). \square

Remark 2.4. • A careful look at the proof shows that the Lipschitz continuity of a and b in the time variable is not needed to obtain that $\sup_{t \in [0, T]} |\mathbb{E}[\psi(X_t^{1, N})] - \mathbb{E}[\psi(X_t)]| \leq \frac{C}{N}$. Indeed, this property is only used to estimate $e_{jk}^{1, \psi}(s)$ which vanishes when $h = 0$ since $\tau_s^0 = s$. Moreover, if a and b are only Hölder continuous with exponent $\alpha \in (0, 1)$ in the time variable, then the above estimation deteriorates to $\sup_{t \in [0, T]} |\mathbb{E}[\psi(X_t^{1, N, h})] - \mathbb{E}[\psi(X_t)]| \leq C \left(\frac{1}{N} + h^\alpha \right)$.

- On the other hand, when there is no nonlinearity in the sense of McKean in the SDE (2.1) ($p = 0$), we obtain that the weak order of convergence of the Euler scheme is 1 under mere global Lipschitz continuity of a, b , continuity of σ, b and C^2 regularity in space of σ, b, ψ with bounded and Lipschitz derivatives. The key step is that the decomposition of $e_{jk}^{2, \psi}(s) + e_{jk}^{3, \psi}(s)$ avoids to differentiate the solution to the Feynman-Kac partial differential equation more than two times in space and only requires Lipschitz continuity of the second order spatial derivatives.

3 Numerical Experiments

We conduct two types of numerical tests. First, we estimate the bias using regular Monte-Carlo for examples of one dimensional ($n = 1$) mean-field SDEs taken from [14] to provide numerical evidence of the $\mathcal{O}(N^{-1})$ behaviour of the bias for a fixed value of the time step h . Then we present the antithetic sampling results on these same examples.

The code for running these experiments has a number of iterations as an input parameter. This latter is useful to observe the behaviour of the bias when increasing the number of particles. Therefore, we give an initial number of particles that we multiply by two from an iteration to the other. Except for the polynomial drift and the plane rotator examples where it is respectively equal to 8 and 4, the number of iterations chosen is equal to five and the initial number of particles is twenty so that the final number of particles is 320. The simulation is done with $5 \cdot 10^6$ runs except for the Plane rotator example where we push further the number of Monte Carlo runs up to 4.9×10^8 .

We also define the precision as half the width of the 95% confidence interval of the empirical mean i.e. $Precision = 1.96 \times \sqrt{\frac{Variance}{number\ of\ runs}}$ where $Variance$ denotes the empirical variance over the runs of the empirical mean over the particles.

In order to measure the relevance of the antithetic sampling technique for variance reduction, we compute the variance of the difference between the empirical mean of the system with $2N$ particles and the mean of the empirical means of the two independent systems with N particles:

$$(3.1) \quad \text{Var} \left[\frac{1}{2N} \sum_{i=1}^{2N} \psi(X_T^{i, 2N}) - \frac{1}{2N} \sum_{i=1}^N \left(\psi(X_T^{i, N}) + \psi(Y_T^{i, N}) \right) \right].$$

Here $(Y_t^{i, N})_{1 \leq i \leq N}$ is the system with N particles driven by $(X_0^i, W^i)_{N+1 \leq i \leq 2N}$.

3.1 Generalised Ornstein-Uhlenbeck process

Model

We consider the following generalization of the Ornstein-Uhlenbeck SDE to a linear mean-field SDE:

$$dX_t = \left[\gamma X_t + \beta \mathbb{E}[X_t] \right] dt + v dW_t, \text{ with } X(0) = x_0$$

for parameters $\gamma, \beta, v \in \mathbb{R}$ and initial data $x_0 \in \mathbb{R}$.

The functions $\alpha(x) = x$, $b(t, y, x) = \gamma x + \beta y$ and $\sigma(t, y, x) = v$ therefore satisfy the hypotheses of regularity of Theorem 2.2.

The first and second moments of X_t are respectively given by $\mathbb{E}[X_t] = x_0 \exp((\gamma + \beta)t)$ and $\mathbb{E}[X_t^2] = x_0^2 \exp(2(\gamma + \beta)t) + \frac{v^2}{2\gamma} (\exp(2\gamma t) - 1)$.

The associated particle approximation of the SDE is given by the following system:

$$dX_t^{i, N} = \left[\gamma X_t^{i, N} + \frac{\beta}{N} \sum_{j=1}^N X_t^{j, N} \right] dt + v dW_t^i, \quad 1 \leq i \leq N \text{ with } X_0^{i, N} = x_0$$

where $\left((W_t^i)_{t \geq 0}\right)_{1 \leq i \leq N}$ are independent Brownian motions.

Because of the linearity of the drift coefficient, it is possible to obtain closed form expressions for $\mathbb{E}[X_t^{1,N,h}]$ and $\mathbb{E}[(X_t^{1,N,h})^2]$ and deduce the asymptotic behaviour of the biases of the first and second order moments as $N \rightarrow \infty$ and $h \rightarrow 0$.

Let $h = T/K$ for some $K \in \mathbb{N}^*$. For $k \in \{0, \dots, K-1\}$ and $i \in \{1, \dots, N\}$, we have

$$X_{(k+1)h}^{i,N,h} = (1 + \gamma h)X_{kh}^{i,N,h} + \beta h \bar{X}_{kh}^{N,h} + v(W_{(k+1)h}^i - W_{kh}^i) \text{ with } \bar{X}_{kh}^{N,h} = \frac{1}{N} \sum_{j=1}^N X_{kh}^{j,N,h}.$$

Hence $\mathbb{E}[X_{kh}^{i,N,h}] = (1 + (\gamma + \beta)h)^k x_0$ so that

$$\mathbb{E}[X_T] - \mathbb{E}[X_T^{i,N,h}] = \frac{1}{2}(\gamma + \beta)^2 T e^{(\gamma + \beta)T} h x_0 + \mathcal{O}(h^2) \text{ as } h \rightarrow 0.$$

To compute $\mathbb{E}[(X_{kh}^{1,N,h})^2]$, we remark that for $k \in \{0, \dots, K-1\}$

$$\begin{aligned} \mathbb{E}[(X_{(k+1)h}^{1,N,h})^2] &= (1 + \gamma h)^2 \mathbb{E}[(X_{kh}^{1,N,h})^2] + (2(1 + \gamma h) + \beta h) \beta h \left(\frac{1}{N} \mathbb{E}[(X_{kh}^{1,N,h})^2] + \frac{N-1}{N} \mathbb{E}[X_{kh}^{1,N,h} X_{kh}^{2,N,h}] \right) + v^2 h \\ \mathbb{E}[X_{(k+1)h}^{1,N,h} X_{(k+1)h}^{2,N,h}] &= (1 + \gamma h)^2 \mathbb{E}[X_{kh}^{1,N,h} X_{kh}^{2,N,h}] + (2(1 + \gamma h) + \beta h) \beta h \left(\frac{1}{N} \mathbb{E}[(X_{kh}^{1,N,h})^2] + \frac{N-1}{N} \mathbb{E}[X_{kh}^{1,N,h} X_{kh}^{2,N,h}] \right) \end{aligned}$$

Since $X_0^{i,N,h} = x_0$ for all $i \in \{1, \dots, N\}$, subtracting the two last equations, we obtain

$$\forall k \in \{1, \dots, K\}, \mathbb{E}[(X_{kh}^{1,N,h})^2] - \mathbb{E}[X_{kh}^{1,N,h} X_{kh}^{2,N,h}] = \frac{(1 + \gamma h)^{2k} - 1}{2\gamma + \gamma^2 h} \times v^2$$

and deduce that

$$\mathbb{E}[(X_{(k+1)h}^{1,N,h})^2] = (1 + (\gamma + \beta)h)^2 \mathbb{E}[(X_{kh}^{1,N,h})^2] + (2(1 + \gamma h) + \beta h) \beta \frac{1-N}{N} \times \frac{(1 + \gamma h)^{2k} - 1}{2\gamma + \gamma^2 h} \times v^2 h + v^2 h.$$

We conclude that

$$\begin{aligned} \mathbb{E}[(X_{kh}^{1,N,h})^2] &= (1 + (\gamma + \beta)h)^{2k} x_0^2 + \frac{(1 + (\gamma + \beta)h)^{2k} - 1}{2(\gamma + \beta) + (\gamma + \beta)^2 h} \left(1 + \frac{N-1}{N} \times \frac{2\beta + (2\beta\gamma + \beta^2)h}{2\gamma + \gamma^2 h} \right) v^2 \\ &\quad + \frac{1-N}{N} \times \frac{(1 + (\gamma + \beta)h)^{2k} - (1 + \gamma h)^{2k}}{2\gamma + \gamma^2 h} \times v^2 \\ &= (1 + (\gamma + \beta)h)^{2k} x_0^2 + \frac{N-1}{N} \times \frac{(1 + \gamma h)^{2k} - 1}{2\gamma + \gamma^2 h} \times v^2 + \frac{1}{N} \times \frac{(1 + (\gamma + \beta)h)^{2k} - 1}{2(\gamma + \beta) + (\gamma + \beta)^2 h} \times v^2 \end{aligned}$$

so that $\mathbb{E}[(X_T)^2] - \mathbb{E}[(X_T^{1,N,h})^2] = \left((\gamma + \beta)^2 T e^{2(\gamma + \beta)T} x_0^2 + \frac{e^{2\gamma T} - 1 + 2\gamma T e^{2\gamma T}}{4} v^2 \right) h + \left(\frac{e^{2\gamma T} - 1}{2\gamma} + \frac{1 - e^{2(\gamma + \beta)T}}{2(\gamma + \beta)} \right) \frac{v^2}{N} + \mathcal{O}(h^2 + \frac{h}{N} + \frac{1}{N^2})$ as $h \rightarrow 0$ and $N \rightarrow \infty$. Moreover,

$$\mathbb{E}[(X_T^h)^2] - \mathbb{E}[(X_T^{1,N,h})^2] = \frac{1}{N} \left(\frac{(1 + \gamma h)^{2k} - 1}{2\gamma + \gamma^2 h} + \frac{1 - (1 + (\gamma + \beta)h)^{2k}}{2(\gamma + \beta) + (\gamma + \beta)^2 h} \right) \times v^2.$$

The bias of the time discretized second order moment is exactly of order 1 in $\frac{1}{N}$.

Results

In order to illustrate the behaviour of the first and second order moments, we compute the difference between the closed-form discretized moments and the estimated moments and expect the difference to be null up to the statistical error. The results are shown in Tables 1 and 2 below where we denote by "Difference" that entity. As a test case, we use this model with $\gamma = \frac{1}{2}$, $\beta = \frac{4}{5}$, $v^2 = \frac{1}{2}$ and $x_0 = 1$.

Concerning the first order moment, we proved above that the bias does not depend on the number N of particles, which is what is observe in the first row of the table below.

Nb. particles	20	40	80	160	320
Estimated first moment	1.34862	1.34861	1.34869	1.34866	1.34866
Difference	0.00003	0.00004	-0.00004	-0.00001	-0.00001
Precision	0.00016	0.00011	0.00008	0.00006	0.00004

Table 1: Generalised Ornstein-Uhlenbeck SDE: Comparison of the estimated first moments with the closed-form discretized value 1.34865 as well as the associated precision when increasing the number of particles for a number of 5.10^6 runs, 50 time steps and a time horizon $T=1$.

As for the second order moment, from the third row we observe that the estimation fits the closed-form discretized value which confirms the behaviour of the bias of order 1 in $\frac{1}{N}$.

Nb. particles	20	40	80	160	320
Closed-form discretized second moment	2.15552	2.14648	2.14195	2.13969	2.13856
Estimated second moment	2.15531	2.14655	2.14205	2.13970	2.13859
Difference	0.00021	-0.00007	-0.00010	-0.00001	-0.00003
Precision	0.00045	0.00032	0.00022	0.00016	0.00011

Table 2: Generalised Ornstein-Uhlenbeck SDE: Comparison of the estimated second moments with their closed-form discretized values as well as the associated precision when increasing the number of particles for a number of 5.10^6 runs, 50 time steps and a time horizon $T=1$.

Concerning the antithetic sampling of the first order moment, using the linearity of the model, one exactly checks that for $k \in \{0, \dots, K-1\}$, $\bar{X}_{kh}^{2N,h} = \frac{1}{2} (\bar{X}_{kh}^{N,h} + \bar{Y}_{kh}^{N,h})$ where $\bar{Y}_{kh}^{N,h}$ denotes the empirical mean of the discretized system driven by $(W^i)_{N+1 \leq i \leq 2N}$. The first row of Table 3 confirms that the variance (3.1) for $\psi(x) = x$ is null.

For the second order moment, we compute the variance (3.1) for $\psi(x) = x^2$ and observe the ratio of decrease $\frac{\text{Variance}(N/2)}{\text{Variance}(N)}$ when increasing the number of particles. The results are shown in the third and forth rows of Table 3.

Nb. particles	20	40	80	160	320
Variance for the first moment	7.82256e-32	9.52151e-32	1.52652e-31	2.79857e-31	5.39704e-31
Precision	9.88245e-35	1.18553e-34	1.89269e-34	3.47065e-34	6.69575e-34
Variance for the second moment	0.000766641	0.000191654	4.77727e-05	1.19847e-05	2.99537e-06
Ratio of decrease	\times	4.00013	4.01179	3.98613	4.00109
Precision	2.4997e-06	6.27208e-07	1.55659e-07	3.9365e-08	9.75389e-09

Table 3: Generalised Ornstein-Uhlenbeck SDE: Evolution of the antithetic variance for both $\psi(x) = x$ and $\psi(x) = x^2$ with its associated precision when increasing the number of particles for a number of 5.10^6 runs, 50 time steps and a time horizon of 1.

For the second order moment, the ratio of successive variances is around 4 which means that the variance of the antithetic estimator is roughly proportional to N^{-2} . The antithetic sampling technique therefore shows an important improvement for this diffusion.

3.2 Plane Rotator

Model

The following SDE refers to a model for coupled oscillators in the presence of noise, also known as the Kuramoto model:

$$\begin{aligned} dX_t &= \left[K \int_{\mathbb{R}} \sin(y - X_t) dP_t^\mu(y) - \sin(X_t) \right] dt + \sqrt{2k_B T} dW_t \\ &= \left[K \cos(X_t) \int_{\mathbb{R}} \sin(y) dP_t^\mu - K \sin(X_t) \int_{\mathbb{R}} \cos(y) dP_t^\mu - \sin(X_t) \right] dt + \sqrt{2k_B T} dW_t \end{aligned}$$

where X_0 is distributed according to μ , P_t^μ denotes the distribution of X_t , $K > 0$ a coupling parameter, k_B the Boltzmann constant and T the temperature.

The functions $\sigma(x) = \sqrt{2k_B T}$ and $\alpha(x) = \begin{bmatrix} \sin(x) \\ \cos(x) \end{bmatrix}$ satisfy the hypotheses of Theorem 2.2. One may also find a function $b(t, y, x)$ coinciding with $K(\cos(x)y_1 - \sin(x)y_2) - \sin(x)$ on $[0, T] \times [-1, 1]^2 \times \mathbb{R}$ which satisfies the hypotheses even if $(x, y) \rightarrow \cos(x)y_1 - \sin(x)y_2$ is not Lipschitz continuous.

The particle system has the following dynamics:

$$dX_t^{i,N} = \left[K \left(\cos(X_t^{i,N}) \frac{1}{N} \sum_{j=1}^N \sin(X_t^{j,N}) - \sin(X_t^{i,N}) \frac{1}{N} \sum_{j=1}^N \cos(X_t^{j,N}) - \sin(X_t^{i,N}) \right) \right] dt + \sqrt{2k_B T} dW_t^i$$

where $\left((W_t^i)_{t \geq 0} \right)_{1 \leq i \leq N}$ are independent Brownian motions.

Results

We use this model with $K = 1$, $k_B T = \frac{1}{8}$ and initial distribution $\mu = \mathcal{N}(\frac{\pi}{4}, \frac{3\pi}{4})$ as a test case. The reference value was computed for 2.5×10^8 Monte Carlo runs, 1000 particles and the same input parameters as for the general estimation of the bias (time horizon $T = 1$ and time step $h = 50$). The value obtained is 0.737576. The results are shown in Table 4.

Nb. particles	20	40	80	160
First moment error	0.000725	0.000355	0.000175	0.000067
Ratio of decrease	\times	2.04225	2.02857	2.61194
Precision	4.79156e-05	3.38946e-05	2.39723e-05	1.69533e-05

Table 4: Plane Rotator SDE: Evolution of the first order moment errors as well as the associated precision when increasing the number of particles for a number of 4.9×10^8 runs, 50 time steps and a time horizon $T=1$.

These results are consistent with a bias proportional to N^{-1} .

Table 5 exposes the results obtained for the antithetic variance.

Nb. particles	20	40	80	160
Variance	0.000119023	3.12526e-05	8.01086e-06	2.0277e-06
Ratio of decrease	\times	3.80842	3.90128	3.95071
Precision	2.51877e-08	6.75418e-09	1.75258e-09	4.4622e-10

Table 5: Plane Rotator SDE: Evolution of the antithetic variance and its associated precision when increasing the number of particles for a number of 4.9×10^8 runs, 50 time steps and a time horizon $T=1$.

The antithetic sampling method is therefore relevant in this example since the ratio of decrease is close to four which means that the variance of the antithetic estimation is roughly proportional to N^{-2} . This behaviour has also been observed in Section 3.2.2 [10].

3.3 Polynomial Drift

Model

Let us consider the following mean-field SDE:

$$dX_t = \left[\gamma X_t + \mathbb{E}[X_t] - X_t \mathbb{E}[X_t^2] \right] dt + X_t dW_t \quad \text{with } X(0) = x_0$$

for a certain parameter $\gamma \in \mathbb{R}$ and initial data $x_0 \in \mathbb{R}$.

The function $\sigma(t, y, x) = x$ satisfy the hypotheses of regularity. However, the functions $b(t, y, x) = \gamma x + y_1 - xy_2$ and $\alpha(x) = \begin{bmatrix} x \\ x^2 \end{bmatrix}$ are not Lipschitz.

From the evolution of the Euler discretization $X_{(k+1)h}^h = X_{kh}^h \left((1 + \gamma h - h \mathbb{E}[(X_{kh}^h)^2]) \right) + h \mathbb{E}[X_{kh}^h] + X_{kh}^h (W_{(k+1)h} - W_{kh})$, we deduce that:

$$\mathbb{E}[X_{(k+1)h}^h] = \left(1 + (\gamma + 1)h - \mathbb{E}[(X_{kh}^h)^2]h \right) \mathbb{E}[X_{kh}^h]$$

$$\mathbb{E}[(X_{(k+1)h}^h)^2] = \left(1 + \gamma h - \mathbb{E}[(X_{kh}^h)^2]h \right)^2 \mathbb{E}[(X_{kh}^h)^2] + 2 \left(1 + \gamma h - \mathbb{E}[(X_{kh}^h)^2]h \right) h \mathbb{E}^2[X_{kh}^h] + \mathbb{E}[(X_{kh}^h)^2]h + \mathbb{E}^2[X_{kh}^h]h^2$$

And we solve this system of inductive equations numerically to obtain reference values of the first and second order moments.

The idea here, once again, is to approximate $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ by their empirical means in order to define the dynamics of the particle system:

$$dX_t^{i,N} = \left[\gamma X_t^{i,N} + \frac{1}{N} \sum_{j=1}^N X_t^{j,N} - X_t^{i,N} \frac{1}{N} \sum_{j=1}^N (X_t^{j,N})^2 \right] dt + X_t^{i,N} dW_t^i$$

where $\left((W_t^i)_{t \geq 0} \right)_{1 \leq i \leq N}$ are independent Brownian motions.

Results

We use this model with $\gamma = 2$ and $x = 1$ as a test case. The reference values obtained for the time discretized first and second moments are respectively 1.3845 and 3.13743. For this example, we push the iterations further until 8 so that the final number of particles is 2560. We denote by "Ratio of decrease 1" and "Ratio of decrease 2" the respective entities $\frac{\text{First moment error}(N/2)}{\text{First moment error}(N)}$ and $\frac{\text{Second moment error}(N/2)}{\text{Second moment error}(N)}$. The results are shown in Tables 6 and 7.

Nb. particles	40	80	160	320	640	1280	2560
First moment error	-0.02597	-0.01561	-0.00898	-0.00492	-0.00261	-0.00138	-0.00071
Ratio of decrease 1	×	1.66368	1.73831	1.82520	1.88506	1.89130	1.94366
Precision	9.348e-05	7.402e-05	5.779e-05	4.431e-05	3.330e-05	2.459e-05	1.789e-05

Table 6: Polynomial Drift SDE: Evolution of the first order moment errors as well as the associated precision when increasing the number of particles for a number of 5.10^6 runs, 50 time steps and a time horizon $T=1$

Nb. particles	40	80	160	320	640	1280	2560
Second moment error	0.06025	0.03575	0.02077	0.01157	0.00637	0.00333	0.00171
Ratio of decrease 2	×	1.68531	1.72123	1.79516	1.81633	1.91291	1.94737
Precision	0.0005018	0.0003940	0.0003029	0.0002297	0.0001707	0.0001251	9.063e-05

Table 7: Polynomial Drift SDE: Evolution of the second order moment errors as well as the associated precision when increasing the number of particles for a number of $5 \cdot 10^6$ runs, 50 time steps and a time horizon $T=1$

We observe that the ratios of decrease of both the first and the second order moment errors seem to grow as the number of particle increases. It confirms the behaviour of the bias in $\mathcal{O}(\frac{1}{N})$ of Theorem 2.2 and even tends towards a behaviour of order 1 in $\frac{1}{N}$ when the number of particle is large.

In order to measure the relevance of the antithetic sampling, we compute the variance (3.1) for the first and second order moments. We denote by "Ratio of decrease V1" and "Ratio of decrease V2" the respective entities $\frac{\text{Variance for first moment error}(N/2)}{\text{Variance for first moment error}(N)}$ and $\frac{\text{Variance for Second moment error}(N/2)}{\text{Variance for Second moment error}(N)}$. The results are shown in Tables 8 and 9.

Nb. particles	40	80	160	320	640	1280	2560
Variance for the first moment	0.0008910	0.0004406	0.0001930	7.474e-05	2.584e-05	8.000e-06	2.262e-06
Ratio of decrease V1	×	2.023	2.283	2.582	2.892	3.230	3.537
Precision	3.621e-06	2.260e-06	1.312e-06	6.650e-07	3.278e-07	1.294e-07	7.131e-08

Table 8: Polynomial Drift SDE: Evolution of the antithetic variance for $\psi(x) = x$ with its associated precision when increasing the number of particles for a number of $5 \cdot 10^6$ runs, 50 time steps and a time horizon $T=1$.

Nb. particles	40	80	160	320	640	1280	2560
Variance for the second moment	0.02380	0.01078	0.00440	0.00170	0.00058	0.00018	5.282e-05
Ratio of decrease V2	×	2.2074	2.45094	2.60197	2.89779	3.20462	3.44723
Precision	0.000113	6.672e-05	3.597e-05	2.415e-05	1.125e-05	5.042e-06	2.862e-06

Table 9: Polynomial Drift SDE: Evolution of the antithetic variance for $\psi(x) = x^2$ with its associated precision when increasing the number of particles for a number of $5 \cdot 10^6$ runs, 50 time steps and a time horizon $T=1$.

The results exposed in Tables 8 and 9 both show that when increasing the number of particles, the ratios of decrease grow gradually. Therefore, the antithetic sampling method may be relevant when simulating with a large number of particles.

3.4 Viscous Burgers equation

Model

Let us consider the following mean-field SDE for a parameter $v > 0$:

$$dX_t = \bar{F}_t(X_t)dt + v dW_t$$

where $\bar{F}_t(x) = \mathbb{P}(X_t \geq x)$.

Using the Fokker-Planck equation satisfied by the density of X_t , we show that $\bar{F}_t(x)$ is solution to the viscous Burgers equation:

$$\partial_t V(t, x) = \frac{v^2}{2} \partial_{xx} V(t, x) - V(t, x) \partial_x V(t, x)$$

We now suppose that the initial condition is $X(0) = 0$ so that $\bar{F}_0(x) = \mathbf{1}_{\{x \leq 0\}}$. Then the Cole-Hopf transformation gives:

$$\bar{F}_t(x) = \frac{\mathcal{N}\left(\frac{t-x}{v\sqrt{t}}\right)}{\exp\left(\frac{2x-t}{2v^2}\right)\mathcal{N}\left(\frac{x}{v\sqrt{t}}\right) + \mathcal{N}\left(\frac{t-x}{v\sqrt{t}}\right)}$$

where $\mathcal{N}(x) = \int_{-\infty}^x \exp(-\frac{y^2}{2}) \frac{dy}{\sqrt{2\pi}}$.

The function $\sigma(t, y, x) = v$ and $b(t, y, x) = y$ are regular enough to satisfy the hypotheses of Theorem 2.2. However, this type of example is not interacting through moments but through a kernel and even a discontinuous one.

We approximate $\bar{F}_t(x)$ by its associated empirical mean $\bar{F}_t^N(x) := \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{X_t^{j,N} \geq x\}}$ calculated upon N particles which leads to the following dynamics for the particle system:

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{X_t^{j,N} \geq X_t^{i,N}\}} dt + v dW_t^i$$

where $\left((W_t^i)_{t \geq 0}\right)_{1 \leq i \leq N}$ are independent Brownian motions.

Results

We use this model with $v = \frac{1}{4}$ and $x_0 = 0$ as a test case. We estimate the solution $\bar{F}_1(\frac{1}{2})$ of the viscous Burgers equation. Since for $t = 1$ and $x = \frac{1}{2}$, $t - x = x$ and $2x - t = 0$, one has $\bar{F}_1(\frac{1}{2}) = \frac{1}{2}$. The results are shown in Table 10.

Nb. particles	20	40	80	160	320
Solution Error	0.0141425	0.0070329	0.00352185	0.00174521	0.000870359
Ratio of decrease	×	2.01091	1.99693	2.01802	2.00515
Precision	9.84069e-05	6.94588e-05	4.90673e-05	3.46603e-05	2.45057e-05

Table 10: Viscous Burgers equation: Evolution of the solution errors as well as the associated precision when increasing the number of particles for a number of $5 \cdot 10^6$ runs, 500 time steps and a time horizon $T=1$.

We observe that the ratio of decrease of the solution error defined by $\frac{\text{Solution error}(N/2)}{\text{Solution error}(N)}$ is consistent with an error proportional to N^{-1} confirming a bias behaviour of order 1 in $\frac{1}{N}$.

As for the antithetic sampling, we compute the corresponding variance (3.1) as well as the related precision and obtain:

Nb. particles	20	40	80	160	320
Variance of the solution	0.00286895	0.00109648	0.000413156	0.000155082	5.89916e-05
Ratio of decrease	×	2.61651	2.65392	2.66411	2.62889
Precision	3.98911e-06	1.49543e-06	5.52988e-07	2.04566e-07	7.65869e-08

Table 11: Viscous Burgers Equation: Evolution of the antithetic variance and its associated precision when increasing the number of particles for a number of $5 \cdot 10^6$ runs, 500 time steps and a time horizon $T=1$.

From Table 11, we observe that the ratio of decrease of the variance $\frac{\text{Variance of solution}(N/2)}{\text{Variance of solution}(N)}$ is around 2.64 which is slightly greater than two. Therefore, there is a slight gain in using the antithetic sampling for this type of diffusion.

References

- [1] K. Bujok, B. Hambly and C. Reisinger, Multilevel Simulation of Bernoulli Random Variables with Applications to Basket Credit Derivatives, *Methodol. Comput. Appl. Probab.* 17:579-604, 2015.
- [2] J.-F. Chassagneux, D. Crisan and F. Delarue, A Probabilistic approach to classical solutions of the master equation for large population equilibria, Preprint arXiv:1411.3009
- [3] N. Chen, P. Glynn and Y. Liu, Estimating expectations of functionals of conditional expectations via multilevel nested simulation, Presentation at MCQMC 2012.
- [4] P. Del Moral, F. Patras and S. Rubenthaler, Tree based functional expansions for Feynman-Kac particle models, *Ann. Appl. Probab.* 10(2):778-825, 2009.
- [5] A. Friedman. *Stochastic Differential Equations and Applications*, vol 1. Academic Press, 1975.
- [6] C. Graham and S. Méléard, Stochastic Particle Approximations for Generalized Boltzmann Models and Convergence Estimates, *Ann. Probab.* 25(1):115-132, 1997.
- [7] M. -B. Giles, Multilevel Monte Carlo Path Simulation, *Oper. Res.* 56(3):607-617, 2008.
- [8] M. Giles, Multilevel Monte Carlo methods, *Acta Numerica* 24:259-328, 2015
- [9] A.-L. Haji-Ali, Pedestrian flow in the mean-field limit, MSc thesis, KAUST, 2012.
- [10] A.-L. Haji-Ali and R. Tempone, Multilevel and Multi-index Monte Carlo methods for the McKean-Vlasov equation, *Stat. Comput.*, Published online : 12 september 2017.
- [11] B. Jourdain, S. Méléard and W. Woyczynski, Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab. Math. Stat.* 4:1-29, 2008.
- [12] S. Kanagawa, On the rate of convergence for Maruyama's approximate solutions of stochastic differential equations. *Yokohama Math. J.* 36(1):79-86, 1988.
- [13] V. Kolokoltsov and M. Troeva, On the mean field games with common noise and the McKean-Vlasov SPDEs, Preprint arXiv:1506.04594
- [14] P. Kloeden and T. Shardlow, Gauss-quadrature method for one-dimensional mean-field SDEs. Preprint arXiv:1608.06741, 2016.
- [15] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms. *Ecole d'été de Probabilités de Saint-Flour XII - 1982. Lecture Notes in Mathematics.* Springer Verlag, 1984, pp. 144-305.
- [16] A.-S. Snitzman, Topics in propagation of chaos. *Ecole d'été de Probabilités de Saint-Flour XIX - 1989. Lecture Notes in Mathematics.* Springer Berlin Heidelberg, 1991, pp. 165-251.
- [17] D. Talay and L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* 8(4):483-509, 1990.