



# Particle system algorithm and chaos propagation related to non-conservative McKean type stochastic differential equations

Anthony Le Cavil, Nadia Oudjane, Francesco Russo

## ► To cite this version:

Anthony Le Cavil, Nadia Oudjane, Francesco Russo. Particle system algorithm and chaos propagation related to non-conservative McKean type stochastic differential equations. 2016. hal-01241704v2

HAL Id: hal-01241704

<https://hal-ensta-paris.archives-ouvertes.fr//hal-01241704v2>

Submitted on 2 Aug 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Particle system algorithm and chaos propagation related to non-conservative McKean type stochastic differential equations.

ANTHONY LE CAVIL <sup>\*</sup>, NADIA OUDJANE <sup>†</sup> AND FRANCESCO RUSSO <sup>‡</sup>

August 2nd 2016

## Abstract

We discuss numerical aspects related to a new class of nonlinear Stochastic Differential Equations in the sense of McKean, which are supposed to represent non conservative nonlinear Partial Differential equations (PDEs). We propose an original interacting particle system for which we discuss the propagation of chaos. We consider a time-discretized approximation of this particle system to which we associate a random function which is proved to converge to a solution of a regularized version of a nonlinear PDE.

**Key words and phrases:** Chaos propagation; Nonlinear Partial Differential Equations; McKean type Non-linear Stochastic Differential Equations; Particle systems; Probabilistic representation of PDEs.

**2010 AMS-classification:** 65C05; 65C35; 68U20; 60H10; 60H30; 60J60; 58J35

## 1 Introduction

Stochastic differential equations of various types are very useful to investigate nonlinear partial differential equations (PDEs) at the theoretical and numerical level. From a theoretical point of view, they constitute probabilistic tools to study the analytic properties of the equation. Moreover they provide a microscopic interpretation of physical phenomena macroscopically drawn by a nonlinear PDE. From a numerical point of view, such representations allow for extended Monte Carlo type methods, which are potentially less sensitive to the dimension of the state space.

Let us consider  $d, p \in \mathbb{N}^*$ . Let  $\Phi : [0, T] \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^{d \times p}$ ,  $g : [0, T] \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$ ,  $\Lambda : [0, T] \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ , be Borel bounded functions,  $K : \mathbb{R}^d \rightarrow \mathbb{R}$  be a smooth mollifier in  $\mathbb{R}^d$  and  $\zeta_0$  be a probability on  $\mathbb{R}^d$ . When it is absolutely continuous  $v_0$  will denote its density so that  $\zeta_0(dx) = v_0(x)dx$ . The main motivation of this work is the simulation of solutions to PDEs of the form

$$\begin{cases} \partial_t v = \sum_{i,j=1}^d \partial_{ij}^2 ((\Phi \Phi^t)_{i,j}(t, x, v)v) - \operatorname{div}(g(t, x, v)v) + \Lambda(t, x, v) \\ v(0, dx) = v_0(dx), \end{cases} \quad (1.1)$$

---

<sup>\*</sup>ENSTA-ParisTech, Université Paris-Saclay. Unité de Mathématiques Appliquées (UMA). E-mail: anthony.lecavil@ensta-paristech.fr

<sup>†</sup>EDF R&D, and FiME (Laboratoire de Finance des Marchés de l'Energie (Dauphine, CREST, EDF R&D) www.fime-lab.org). E-mail:nadia.oudjane@edf.fr

<sup>‡</sup>ENSTA-ParisTech, Université Paris-Saclay. Unité de Mathématiques Appliquées (UMA). E-mail:francesco.russo@ensta-paristech.fr

through probabilistic numerical methods. Examples of nonlinear and nonconservative PDEs that are of that form arise in hydrodynamics and biological modeling. For instance one model related to underground water flows is known in the literature as the Richards equation

$$\begin{cases} \partial_t v = \Delta(\beta(v)) + \operatorname{div}(\alpha(v)) + \phi(x) \\ v(0, \cdot) = v_0, \end{cases} \quad (1.2)$$

where  $\beta : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\alpha : \mathbb{R} \rightarrow \mathbb{R}^d$  and  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ . Another example concerns biological mechanisms as migration of biological species or the evolution of a tumor growth. Such equations can be schematically written as

$$\begin{cases} \partial_t v = \Delta\beta(v) + f(v) \\ v(0, \cdot) = v_0, \end{cases} \quad (1.3)$$

where  $\beta : \mathbb{R} \rightarrow \mathbb{R}$  is bounded, monotone and  $f : \mathbb{R} \rightarrow \mathbb{R}$ . This family of PDEs is sometimes called Porous Media type Equation with proliferation, due to the presence of the term  $f$  that characterizes a proliferation phenomena and the term  $\Delta\beta(v)$  delineates a porous media effect. In particular, for  $\beta(v) = v^2$  and  $f(v) = v(1 - v)$ , this type of equation appears in the modeling of tumors.

The present paper focuses on numerical aspects of a specific forward probabilistic representation initiated in [20], relying on nonlinear SDEs in the sense of McKean [21]. In [20], we have introduced and studied a generalized regularized McKean type nonlinear stochastic differential equation (NLSDE) of the form

$$\begin{cases} Y_t = Y_0 + \int_0^t \Phi(s, Y_s, u(s, Y_s)) dW_s + \int_0^t g(s, Y_s, u(s, Y_s)) ds, & \text{with } Y_0 \sim \zeta_0, \\ u(t, y) = \mathbb{E}[K(y - Y_t) \exp \left\{ \int_0^t \Lambda(s, Y_s, u(s, Y_s)) ds \right\}], & \text{for any } t \in [0, T], \end{cases} \quad (1.4)$$

where the solution is the couple process-function  $(Y, u)$ . The novelty with respect to classical McKean type equations consists in the form of the second equation, where, for each  $t > 0$ , in the classical case ( $\Lambda = 0$ )  $u(t, \cdot)$  was explicitly given by the marginal law of  $Y_t$ . The present paper aims at proposing and implementing a stochastic particle algorithm to approximate (1.4) and investigating carefully its convergence properties.

(1.4) is the probabilistic representation of the partial integro-differential equation (PIDE)

$$\begin{cases} \partial_t \bar{v} = \frac{1}{2} \sum_{i,j=1}^d \partial_{i,j}^2 ((\Phi\Phi^t)_{i,j}(t, x, K * \bar{v})\bar{v}) - \operatorname{div}(g(t, x, K * \bar{v})\bar{v}) + \Lambda(t, x, K * \bar{v})\bar{v}, \\ \bar{v}(0, x) = v_0, \end{cases} \quad (1.5)$$

in the sense that, given a solution  $(Y, u)$  of (1.4), there is a solution  $\bar{v}$  of (1.5) in the sense of distributions, such that  $u = K * \bar{v} := \int_{\mathbb{R}^d} K(\cdot - y)\bar{v}(y)dy$ . This follows, for instance, by a simple application of Itô's formula, as explained in Theorems 6.1 and 6.2, Section 6 in [20]. Ideally our interest is devoted to (1.4) when the smoothing kernel  $K$  reduces to a Dirac measure at zero. To reach that scope, one would need to replace in previous equation  $K$  into  $K_\varepsilon$ , where  $K_\varepsilon$  converges to the Dirac measure and to analyze the convergence of the corresponding solutions. However, such a theoretical analysis is out of the scope of this paper, but it will be investigated numerically via simulations reported at the end.

In fact, in the literature appear several probabilistic representations, with the objective of simulating numerically the corresponding PDE. One method which has been largely investigated for approximating solutions of time evolutionary PDEs is the method of forward-backward SDEs (FBSDEs). FBSDEs were initially developed in [23], see also [22] for a survey and [24] for a recent monograph on the subject. The idea is to express the PDE solution  $v(t, \cdot)$  at time  $t$  as the expectation of a functional of a so called forward diffusion

process  $X$ , starting at time  $t$ . Based on that idea, many judicious numerical schemes have been proposed by [9, 14]. However, all those rely on computing recursively conditional expectation functions which is known to be a difficult task in high dimension. Besides, the FBSDE approach is *blind* in the sense that the forward process  $X$  is not ensured to explore the most relevant space regions to approximate efficiently the solution of the FBSDE of interest. On the theoretical side, the FBSDE representation of fully nonlinear PDEs still requires complex developments and is the subject of active research (see for instance [10]). Branching diffusion processes provide alternative probabilistic representation of semi-linear PDEs, involving a specific form of non-linearity on the zero order term. This type of approach has been recently extended in [15, 16] to a more general class of non-linearities on the zero order term, with the so-called *marked branching process*. One of the main advantages of this approach compared to FBSDEs is that it does not involve any regression computation to calculate conditional expectations. A third class of numerical approximation schemes relies on McKean type representations. In the time continuous framework, classical McKean representations are restricted to the conservative case ( $\Lambda = 0$ ). Relevant contributions at the algorithmic level are [7, 8, 6, 4], and the survey paper [28]. In the case  $\Lambda = 0$  with  $g = 0$ , but with  $\Phi$  possibly discontinuous, some empirical implementations were conducted in [2, 3] in the one-dimensional and multi-dimensional case respectively, in order to predict the large time qualitative behavior of the solution of the corresponding PDE.

In the present paper we extend this type of McKean based numerical schemes to the case of non-conservative PDEs ( $\Lambda \neq 0$ ). An interesting aspect of this approach is that it is potentially able to represent fully nonlinear PDEs, by considering a more general class of functions  $\Lambda$  which may depend non-linearly not only on  $u$  but on its space derivatives up to the second order. This more general setting will be focused in a future work. In the discrete-time framework, Feynman-Kac formula and various types of related particle approximation schemes were extensively analyzed in the reference books of Del Moral [12] and [13] but without considering the specific case of a time continuous system (1.4) coupled with a weighting function  $\Lambda$  which depends nonlinearly on  $u$ .

By (3.3) we introduce an interacting particle system associated to (1.4). Indeed we replace one single McKean type stochastic differential equation with unknown process  $Y$ , with a system of  $N$  ordinary stochastic differential equations, whose solution consists in a system of particles  $\xi = (\xi^{j,N})$ , replacing the law of the process  $Y$  by the empirical mean law  $S^N(\xi) := \sum_{j=1}^N \delta_{\xi^{j,N}}$ .

In Theorem 4.2 we prove the convergence of the time-discretized particle system under Lipschitz type assumptions on the coefficients  $\Phi$ ,  $g$  and  $\Lambda$ , obtaining an explicit rate. The mentioned rate is based on the contribution of two effects. First, the particle approximation error between the solution  $u$  of (1.4) and the approximation  $u^{S^N(\xi)}$ , solution of

$$u_t^{S^N(\xi)}(y) = \frac{1}{N} \sum_{j=1}^N K(y - \xi_t^{j,N}) \exp \left\{ \int_0^t \Lambda(s, \xi_s^{j,N}, u^{S^N(\xi)}(s, \xi_s^{j,N})) ds \right\}, \quad (1.6)$$

which is evaluated in Theorem 3.1. The second effect is the time discretization error, established in Proposition 4.1. The errors are evaluated in the  $L^p$ ,  $p = 2, +\infty$  mean distance, in terms of the number  $N$  of particles and the time discretization step. One significant consequence of Theorem 3.1 is Corollary 3.2 which states the chaos propagation of the interacting particle system.

We emphasize that the proof of Theorem 3.1 relies on Proposition 3.3, whose formula (3.12) allows to control the particle approximation error without use of exchangeability assumptions on the particle system, see Remark 3.4.

The paper is organized as follows. After this introduction, we formulate the basic assumptions valid

along the paper and recall important results proved in [20] and used in the sequel. The evaluation of the particle approximation error is discussed in Section 3. Section 4 focuses on the convergence of the time-discretized particle system. Finally in Section 5 we provide numerical simulations illustrating the performances of the interacting particle system in approximating the limit PDE (i.e. when the smoothing kernel  $K$  reduces to a Dirac measure at zero), in a specific case where the solution is explicitly known.

## 2 Notations and assumptions

Let us consider  $\mathcal{C}^d := \mathcal{C}([0, T], \mathbb{R}^d)$  metrized by the supremum norm  $\|\cdot\|_\infty$ , equipped with its Borel  $\sigma$ -field  $\mathcal{B}(\mathcal{C}^d) = \sigma(X_t, t \geq 0)$  (and  $\mathcal{B}_t(\mathcal{C}^d) := \sigma(X_u, 0 \leq u \leq t)$  the canonical filtration) and endowed with the topology of uniform convergence.  $X$  will be the canonical process on  $\mathcal{C}^d$  and  $\mathcal{P}_r(\mathcal{C}^d)$  the set of Borel probability measures on  $\mathcal{C}^d$  admitting a moment of order  $r \geq 0$ . For  $r = 0$ ,  $\mathcal{P}(\mathcal{C}^d) := \mathcal{P}_0(\mathcal{C}^d)$  is naturally the Polish space (with respect to the weak convergence topology) of Borel probability measures on  $\mathcal{C}^d$  naturally equipped with its Borel  $\sigma$ -field  $\mathcal{B}(\mathcal{P}(\mathcal{C}^d))$ . When  $d = 1$ , we often omit it and we simply set  $\mathcal{C} := \mathcal{C}^1$ .

We recall that the Wasserstein distance of order  $r$  and respectively the *modified Wasserstein distance of order  $r$*  for  $r \geq 1$ , between  $m$  and  $m'$  in  $\mathcal{P}_r(\mathcal{C}^d)$ , denoted by  $W_T^r(m, m')$  (and resp.  $\tilde{W}_T^r(m, m')$ ) are such that

$$(W_t^r(m, m'))^r := \inf_{\mu \in \Pi(m, m')} \left\{ \int_{\mathcal{C}^d \times \mathcal{C}^d} \sup_{0 \leq s \leq t} |X_s(\omega) - X_s(\omega')|^r d\mu(\omega, \omega') \right\}, \quad t \in [0, T], \quad (2.1)$$

$$(\tilde{W}_t^r(m, m'))^r := \inf_{\mu \in \tilde{\Pi}(m, m')} \left\{ \int_{\mathcal{C}^d \times \mathcal{C}^d} \sup_{0 \leq s \leq t} |X_s(\omega) - X_s(\omega')|^r \wedge 1 d\mu(\omega, \omega') \right\}, \quad t \in [0, T], \quad (2.2)$$

where  $\Pi(m, m')$  (resp.  $\tilde{\Pi}(m, m')$ ) denotes the set of Borel probability measures in  $\mathcal{P}(\mathcal{C}^d \times \mathcal{C}^d)$  with fixed marginals  $m$  and  $m'$  belonging to  $\mathcal{P}_r(\mathcal{C}^d)$  (resp.  $\mathcal{P}(\mathcal{C}^d)$ ). In this paper we will use very frequently the Wasserstein distances of order 2. For that reason, we will simply set  $W_t := W_t^2$  (resp.  $\tilde{W}_t := \tilde{W}_t^2$ ).

Given  $N \in \mathbb{N}^*$ ,  $l \in \mathcal{C}^d$ ,  $l^1, \dots, l^N \in \mathcal{C}^d$ , a significant role in this paper will be played by the Borel measures on  $\mathcal{C}^d$  given by  $\delta_l$  and  $\frac{1}{N} \sum_{j=1}^N \delta_{l^j}$ .

**Remark 2.1.** Given  $l^1, \dots, l^N, \tilde{l}^1, \dots, \tilde{l}^N \in \mathcal{C}^d$ , by definition of the Wasserstein distance we have, for all  $t \in [0, T]$

$$W_t \left( \frac{1}{N} \sum_{j=1}^N \delta_{l^j}, \frac{1}{N} \sum_{j=1}^N \delta_{\tilde{l}^j} \right) \leq \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq s \leq t} |l_s^j - \tilde{l}_s^j|^2.$$

In this paper  $\mathcal{C}_b(\mathcal{C}^d)$  denotes the space of bounded, continuous real-valued functions on  $\mathcal{C}^d$ , for which the supremum norm is denoted by  $\|\cdot\|_\infty$ .  $\mathbb{R}^d$  is equipped with the scalar product  $\cdot$  and  $|x|$  stands for the induced Euclidean norm for  $x \in \mathbb{R}^d$ . Given two reals  $a, b$  ( $d = 1$ ) we will denote in the sequel  $a \wedge b := \min(a, b)$  and  $a \vee b := \max(a, b)$ .

$\mathcal{M}_f(\mathbb{R}^d)$  is the space of finite, Borel measures on  $\mathbb{R}^d$ .  $\mathcal{S}(\mathbb{R}^d)$  is the space of Schwartz fast decreasing test functions and  $\mathcal{S}'(\mathbb{R}^d)$  is its dual.  $\mathcal{C}_b(\mathbb{R}^d)$  is the space of bounded, continuous functions on  $\mathbb{R}^d$ ,  $\mathcal{C}_0^\infty(\mathbb{R}^d)$  is the space of smooth functions with compact support.  $\mathcal{C}_b^\infty(\mathbb{R}^d)$  is the space of bounded and smooth functions.  $\mathcal{C}_0(\mathbb{R}^d)$  represents the space of continuous functions with compact support in  $\mathbb{R}^d$ .  $W^{r,p}(\mathbb{R}^d)$  is the Sobolev space of order  $r \in \mathbb{N}$  in  $(L^p(\mathbb{R}^d), \|\cdot\|_p)$ , with  $1 \leq p \leq \infty$ .

$\mathcal{F}(\cdot) : f \in \mathcal{S}(\mathbb{R}^d) \mapsto \mathcal{F}(f) \in \mathcal{S}'(\mathbb{R}^d)$  will denote the Fourier transform on the classical Schwartz space  $\mathcal{S}(\mathbb{R}^d)$  such that for all  $\xi \in \mathbb{R}^d$ ,

$$\mathcal{F}(f)(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dx.$$

We will designate in the same manner the corresponding Fourier transform on  $\mathcal{S}'(\mathbb{R}^d)$ .

For any Polish space  $E$ , we will designate by  $\mathcal{B}(E)$  its Borel  $\sigma$ -field. It is well-known that  $\mathcal{P}(E)$  is also a Polish space with respect to the weak convergence topology, whose Borel  $\sigma$ -field will be denoted by  $\mathcal{B}(\mathcal{P}(E))$  (see Proposition 7.20 and Proposition 7.23, Section 7.4 Chapter 7 in [5]).

Let  $(\Omega, \mathcal{F})$  be a measured space. A map  $\eta : (\Omega, \mathcal{F}) \longrightarrow (\mathcal{P}(E), \mathcal{B}(\mathcal{P}(E)))$  will be called **random probability** (or **random probability kernel**) if it is measurable. We will indicate by  $\mathcal{P}^\Omega(E)$  the space of random probabilities.

**Remark 2.2.** Let  $\eta : (\Omega, \mathcal{F}) \longrightarrow (\mathcal{P}(E), \mathcal{B}(\mathcal{P}(E)))$ .  $\eta$  is a random probability if and only if the two following conditions hold:

- for each  $\bar{\omega} \in \Omega$ ,  $\eta_{\bar{\omega}} \in \mathcal{P}(E)$ ,
- for all Borel set  $A \in \mathcal{B}(\mathcal{P}(E))$ ,  $\bar{\omega} \mapsto \eta_{\bar{\omega}}(A)$  is  $\mathcal{F}$ -measurable.

This was highlighted in Remark 3.20 in [11] (see also Proposition 7.25 in [5]).

**Remark 2.3.** Given  $\mathbb{R}^d$ -valued continuous processes  $Y^1, \dots, Y^n$ , the application  $\frac{1}{N} \sum_{j=1}^N \delta_{Y^j}$  is a random probability on  $\mathcal{P}(\mathcal{C}^d)$ . In fact  $\delta_{Y^j}$ ,  $1 \leq j \leq N$  is a random probability by Remark 2.2.

In this article, the following assumptions will be used.

**Assumption 1.** 1.  $\Phi$  and  $g$  Borel functions defined on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}$  taking values respectively in  $\mathbb{R}^{d \times p}$  (space of  $d \times p$  matrices) and  $\mathbb{R}^d$  that are Lipschitz w.r.t. space variables: there exist finite positive reals  $L_\Phi$  and  $L_g$  such that for any  $(t, y, y', z, z') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}$ , we have

$$|\Phi(t, y', z') - \Phi(t, y, z)| \leq L_\Phi(|z' - z| + |y' - y|) \quad \text{and} \quad |g(t, y', z') - g(t, y, z)| \leq L_g(|z' - z| + |y' - y|).$$

2.  $\Lambda$  is a Borel real valued function defined on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}$  Lipschitz w.r.t. the space variables: there exists a finite positive real,  $L_\Lambda$  such that for any  $(t, y, y', z, z') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}$ , we have

$$|\Lambda(t, y, z) - \Lambda(t, y', z')| \leq L_\Lambda(|y' - y| + |z' - z|).$$

3.  $\Lambda$  is supposed to be uniformly bounded: there exist a finite positive real  $M_\Lambda$  such that, for any  $(t, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$

$$|\Lambda(t, y, z)| \leq M_\Lambda.$$

4.  $K : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is integrable, Lipschitz, bounded and whose integral is 1: there exist finite positive reals  $M_K$  and  $L_K$  such that for any  $(y, y') \in \mathbb{R}^d \times \mathbb{R}^d$

$$|K(y)| \leq M_K, \quad |K(y') - K(y)| \leq L_K|y' - y| \quad \text{and} \quad \int_{\mathbb{R}^d} K(x)dx = 1.$$

5.  $\zeta_0$  is a fixed Borel probability measure on  $\mathbb{R}^d$  admitting a second order moment.

6. The functions  $s \in [0, T] \mapsto \Phi(s, 0, 0)$  and  $s \in [0, T] \mapsto g(s, 0, 0)$  are bounded.  $m_\Phi$  (resp.  $m_g$ ) will denote the supremum  $\sup_{s \in [0, T]} |\Phi(s, 0, 0)|$  (resp.  $\sup_{s \in [0, T]} |g(s, 0, 0)|$ ).

Given a finite signed Borel measure  $\gamma$  on  $\mathbb{R}^d$ ,  $K * \gamma$  will denote the convolution function  $x \mapsto \gamma(K(x - \cdot))$ . In particular if  $\gamma$  is absolutely continuous with density  $\dot{\gamma}$ , then  $(K * \gamma)(x) = \int_{\mathbb{R}^d} K(x - y)\dot{\gamma}(y)dy$ .

To simplify we introduce the following notations.

- $V : [0, T] \times \mathcal{C}^d \times \mathcal{C} \rightarrow \mathbb{R}$  defined for any pair of functions  $y \in \mathcal{C}^d$  and  $z \in \mathcal{C}$ , by

$$V_t(y, z) := \exp \left( \int_0^t \Lambda(s, y_s, z_s) ds \right) \quad \text{for any } t \in [0, T]. \quad (2.3)$$

- The real valued process  $Z$  such that  $Z_s = u(s, Y_s)$ , for any  $s \in [0, T]$ , will often be denoted by  $u(Y)$ .

With these new notations, the second equation in (1.4) can be rewritten as

$$\nu_t(\varphi) = \mathbb{E}[(\check{K} * \varphi)(Y_t)V_t(Y, u(Y))], \quad \text{for any } \varphi \in \mathcal{C}_b(\mathbb{R}^d, \mathbb{R}), \quad (2.4)$$

where  $u(t, \cdot) = \frac{d\nu_t}{dx}$  and  $\check{K}(x) := K(-x)$ .

**Remark 2.4.** Under Assumption 1. 3.(b),  $\Lambda$  is bounded. Consequently

$$0 \leq V_t(y, z) \leq e^{tM_\Lambda}, \quad \text{for any } (t, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}. \quad (2.5)$$

Under Assumption 1. 2.  $\Lambda$  is Lipschitz. Then  $V$  inherits in some sense this property. Indeed, observe that for any  $(a, b) \in \mathbb{R}^2$ ,

$$e^b - e^a = (b - a) \int_0^1 e^{\alpha b + (1-\alpha)a} d\alpha \leq e^{\sup(a,b)} |b - a|. \quad (2.6)$$

Then for any continuous functions  $y, y' \in \mathcal{C}^d = \mathcal{C}([0, T], \mathbb{R}^d)$ , and  $z, z' \in \mathcal{C}([0, T], \mathbb{R})$ , taking  $a = \int_0^t \Lambda(s, y_s, z_s) ds$  and  $b = \int_0^t \Lambda(s, y'_s, z'_s) ds$  in the above equality yields

$$\begin{aligned} |V_t(y', z') - V_t(y, z)| &\leq e^{tM_\Lambda} \int_0^t |\Lambda(s, y'_s, z'_s) - \Lambda(s, y_s, z_s)| ds \\ &\leq e^{tM_\Lambda} L_\Lambda \int_0^t (|y'_s - y_s| + |z'_s - z_s|) ds. \end{aligned} \quad (2.7)$$

In Section 4, Assumption 1. will be replaced by what follows.

**Assumption 2.** All items of Assumption 1. are in force excepted 1. and 2. which are replaced by the following.

1. There exist positive reals  $L_\Phi, L_g$  such that, for any  $(t, t', y, y', z, z') \in [0, T]^2 \times (\mathbb{R}^d)^2 \times \mathbb{R}^2$ ,

$$|\Phi(t, y, z) - \Phi(t', y', z')| \leq L_\Phi (|t - t'|^{\frac{1}{2}} + |y - y'| + |z - z'|),$$

$$|g(t, y, z) - g(t', y', z')| \leq L_g (|t - t'|^{\frac{1}{2}} + |y - y'| + |z - z'|).$$

2. There exists a positive real  $L_\Lambda$  such that, for any  $(t, t', y, y', z, z') \in [0, T]^2 \times (\mathbb{R}^d)^2 \times \mathbb{R}^2$ ,

$$|\Lambda(t, y, z) - \Lambda(t', y', z')| \leq L_\Lambda (|t - t'|^{\frac{1}{2}} + |y - y'| + |z - z'|).$$

We end this section by recalling important results established in our companion paper [20], for which Assumption 1. is supposed to be satisfied. Let us first remark that the second equation of (1.4) can be rewritten as

$$u^m(t, y) = \int_{\mathcal{C}^d} K(y - \omega_t) \exp \left\{ \int_0^t \Lambda(s, \omega_s, u^m(s, \omega_s)) \right\} dm(\omega), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \quad (2.8)$$

with  $m = m_Y$  being the law of the process  $Y$  on the canonical space  $\mathcal{C}^d$ .

Indeed for every  $m \in \mathcal{P}(\mathcal{C}^d)$ , Theorem 3.1 of [20] shows that equation (2.8) is well-posed and so it properly defines a function  $u^m$ . The lemma below, established in Proposition 3.3 of [20], states stability results on the function  $(m, t, y) \mapsto u^m(t, y)$ .

**Proposition 2.5.** *We assume the validity of items 2., 3 and 4. of Assumption 1.*

*The following assertions hold.*

1. *For any couple of probabilities  $(m, m') \in \mathcal{P}_2(\mathcal{C}^d) \times \mathcal{P}_2(\mathcal{C}^d)$ , for all  $(t, y, y') \in [0, T] \times \mathcal{C}^d \times \mathcal{C}^d$ , we have*

$$|u^m(t, y) - u^{m'}(t, y')|^2 \leq C_{K,\Lambda}(t) [|y - y'|^2 + |W_t(m, m')|^2], \quad (2.9)$$

*where  $C_{K,\Lambda}(t) := 2C'_{K,\Lambda}(t)(t+2)(1 + e^{2tC'_{K,\Lambda}(t)})$  with  $C'_{K,\Lambda}(t) = 2e^{2tM_\Lambda}(L_K^2 + 2M_K^2L_\Lambda^2t)$ . In particular the functions  $C_{K,\Lambda}$  only depend on  $M_K, L_K, M_\Lambda, L_\Lambda$  and  $t$  and is increasing with  $t$ .*

2. *For any  $(m, m') \in \mathcal{P}(\mathcal{C}^d) \times \mathcal{P}(\mathcal{C}^d)$ , for all  $(t, y, y') \in [0, T] \times \mathcal{C}^d \times \mathcal{C}^d$ , we have*

$$|u^m(t, y) - u^{m'}(t, y')|^2 \leq \mathfrak{C}_{K,\Lambda}(t) [|y - y'|^2 + |\widetilde{W}_t(m, m')|^2], \quad (2.10)$$

*where  $\mathfrak{C}_{K,\Lambda}(t) := 2e^{2tM_\Lambda}(\max(L_K, 2M_K)^2 + 2M_K^2 \max(L_\Lambda, 2M_\Lambda)^2t)$ .*

3. *The function  $(m, t, x) \mapsto u^m(t, x)$  is continuous on  $\mathcal{P}(\mathcal{C}^d) \times [0, T] \times \mathbb{R}^d$ , where  $\mathcal{P}(\mathcal{C}^d)$  is endowed with the topology of weak convergence.*

4. *Suppose that  $K \in W^{1,2}(\mathbb{R}^d)$ . Then for any  $(m, m') \in \mathcal{P}_2(\mathcal{C}^d) \times \mathcal{P}_2(\mathcal{C}^d)$ ,  $t \in [0, T]$*

$$\|u^m(t, \cdot) - u^{m'}(t, \cdot)\|_2^2 \leq \tilde{C}_{K,\Lambda}(t)(1 + 2tC_{K,\Lambda}(t))|W_t(m, m')|^2, \quad (2.11)$$

*where  $C_{K,\Lambda}(t) := 2C'_{K,\Lambda}(t)(t+2)(1 + e^{2tC'_{K,\Lambda}(t)})$  with  $C'_{K,\Lambda}(t) = 2e^{2tM_\Lambda}(L_K^2 + 2M_K^2L_\Lambda^2t)$  and  $\tilde{C}_{K,\Lambda}(t) := 2e^{2tM_\Lambda}(2M_KL_\Lambda^2t(t+1) + \|\nabla K\|_2^2)$ ,  $\|\cdot\|_2$  being the standard  $L^2(\mathbb{R}^d)$  or  $L^2(\mathbb{R}^d, \mathbb{R}^d)$ -norms.*

*In particular the functions  $t \mapsto C'_{K,\Lambda}(t)$  and  $t \mapsto C_{K,\Lambda}(t)$  only depend on  $M_K, L_K, M_\Lambda, L_\Lambda$  and are increasing with respect to  $t$ .*

5. *Suppose that  $\mathcal{F}(K) \in L^1(\mathbb{R}^d)$ . Then there exists a constant  $\bar{C}_{K,\Lambda}(t) > 0$  (depending only on  $t, M_\Lambda, L_\Lambda, \|\mathcal{F}(K)\|_1$ ) such that for any random probability  $\eta : (\Omega, \mathcal{F}) \rightarrow (\mathcal{P}_2(\mathcal{C}^d), \mathcal{B}(\mathcal{P}(\mathcal{C}^d)))$ , for all  $(t, m) \in [0, T] \times \mathcal{P}(\mathcal{C}^d)$*

$$\mathbb{E}[\|u^\eta(t, \cdot) - u^m(t, \cdot)\|_\infty^2] \leq \bar{C}_{K,\Lambda}(t) \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle \eta - m, \varphi \rangle|^2], \quad (2.12)$$

*where we recall that  $\mathcal{P}(\mathcal{C}^d)$  is endowed with the topology of weak convergence. We remark that the expectation in both sides of (2.12) is taken w.r.t. the randomness of the random probability  $\eta$ .*

**Remark 2.6.** *The map  $d_2^\Omega : (\nu, \mu) \mapsto \sqrt{\sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle \nu - \mu, \varphi \rangle|^2]}$  defines a (homogeneous) distance on  $\mathcal{P}^\Omega(\mathcal{C}^d)$ .*

The lemma below was proved in Lemma 7.1 in [20].

**Lemma 2.7.** *Let  $r : [0, T] \mapsto [0, T]$  be a non-decreasing function such that  $r(s) \leq s$  for any  $s \in [0, T]$  and  $Y_0$  be a random variable admitting  $\zeta_0$  as law.*

*Let  $\mathcal{U} : (t, y) \in [0, T] \times \mathcal{C}^d \rightarrow \mathbb{R}$  (respectively  $\mathcal{U}' : (t, y) \in [0, T] \times \mathcal{C}^d \rightarrow \mathbb{R}$ ), be a given Borel function such that for all  $t \in [0, T]$ , there is a Borel map  $\mathcal{U}_t : \mathcal{C}([0, t], \mathbb{R}^d) \rightarrow \mathbb{R}$  (resp.  $\mathcal{U}'_t : \mathcal{C}([0, t], \mathbb{R}^d) \rightarrow \mathbb{R}$ ) such that  $\mathcal{U}(t, \cdot) = \mathcal{U}_t(\cdot)$  (resp.  $\mathcal{U}'(t, \cdot) = \mathcal{U}'_t(\cdot)$ ).*

*Then the following two assertions hold.*



1. Consider  $Y$  (resp.  $Y'$ ) a solution of the following SDE for  $v = \mathcal{U}$  (resp.  $v = \mathcal{U}'$ ):

$$Y_t = Y_0 + \int_0^t \Phi(r(s), Y_{r(s)}, v(r(s), Y_{\wedge r(s)})) dW_s + \int_0^t g(r(s), Y_{r(s)}, v(r(s), Y_{\wedge r(s)})) ds, \quad \text{for any } t \in [0, T], \quad (2.13)$$

where, we emphasize that for all  $\theta \in [0, T]$ ,  $Z_{\cdot \wedge \theta} := \{Z_u, 0 \leq u \leq \theta\} \in \mathcal{C}([0, \theta], \mathbb{R}^d)$  for any continuous process  $Z$ . For any  $a \in [0, T]$ , we have

$$\mathbb{E}[\sup_{t \leq a} |Y'_t - Y_t|^2] \leq C_{\Phi, g}(T) \mathbb{E} \left[ \int_0^a |\mathcal{U}(r(t), Y_{\wedge r(t)}) - \mathcal{U}'(r(t), Y'_{\wedge r(t)})|^2 dt \right], \quad (2.14)$$

where  $C_{\Phi, g}(T) = 12(4L_{\Phi}^2 + TL_g^2)e^{12T(4L_{\Phi}^2 + TL_g^2)}$ .

2. Suppose moreover that  $\Phi$  and  $g$  are  $\frac{1}{2}$ -Hölder w.r.t. the time and Lipschitz w.r.t. the space variables i.e. there exist some positive constants  $L_{\Phi}$  and  $L_g$  such that for any  $(t, t', y, y', z, z') \in [0, T]^2 \times \mathbb{R}^{2d} \times \mathbb{R}^2$

$$\begin{cases} |\Phi(t, y, z) - \Phi(t', y', z')| \leq L_{\Phi}(|t - t'|^{\frac{1}{2}} + |y - y'| + |z - z'|) \\ |g(t, y, z) - g(t', y', z')| \leq L_g(|t - t'|^{\frac{1}{2}} + |y - y'| + |z - z'|). \end{cases} \quad (2.15)$$

Let  $r_1, r_2 : [0, T] \mapsto [0, T]$  being two non-decreasing functions verifying  $r_1(s) \leq s$  and  $r_2(s) \leq s$  for any  $s \in [0, T]$ . Let  $Y$  (resp.  $Y'$ ) be a solution of (2.13) for  $v = \mathcal{U}$  and  $r = r_1$  (resp.  $v = \mathcal{U}'$  and  $r = r_2$ ). Then for any  $a \in [0, T]$ , the following inequality holds:

$$\begin{aligned} \mathbb{E}[\sup_{t \leq a} |Y'_t - Y_t|^2] &\leq C_{\Phi, g}(T) \left( \|r_1 - r_2\|_{L^1([0, T])} + \int_0^a \mathbb{E}[|Y'_{r_1(t)} - Y'_{r_2(t)}|^2] dt \right. \\ &\quad \left. + \mathbb{E} \left[ \int_0^a |\mathcal{U}(r_1(t), Y_{\wedge r_1(t)}) - \mathcal{U}'(r_2(t), Y'_{\wedge r_2(t)})|^2 dt \right] \right), \end{aligned} \quad (2.16)$$

The theorem below was the object of Theorem 3.9 in [20].

**Theorem 2.8.** *Under Assumption 1, the McKean type SDE (1.4) admits strong existence and pathwise uniqueness.*

For a precise formulation of the notion of existence and uniqueness for the McKean type equation (1.4) we refer to Definition 2.6 of [20].

We finally recall an important non-anticipating property of the map  $(m, t, x) \mapsto u^m(t, x)$ , stated in [20].

**Definition 2.9.** *Let us fix  $t \in [0, T]$ . Given a non-negative Borel measure  $m$  on  $(\mathcal{C}^d, \mathcal{B}(\mathcal{C}^d))$ . From now on,  $m_t$  will denote the (unique) induced measure on  $(\mathcal{C}_t^d, \mathcal{B}(\mathcal{C}_t^d))$  (with  $\mathcal{C}_t^d := \mathcal{C}([0, t], \mathbb{R}^d)$ ) defined by*

$$\int_{\mathcal{C}_t^d} F(\phi) m_t(d\phi) = \int_{\mathcal{C}^d} F(\phi|_{[0, t]}) m(d\phi),$$

where  $F : \mathcal{C}_t^d \rightarrow \mathbb{R}$  is bounded and continuous.

**Remark 2.10.** *Let  $t \in [0, T]$ ,  $m = \delta_{\xi}$ ,  $\xi \in \mathcal{C}^d$ . The induced measure  $m_t$ , on  $\mathcal{C}_t^d$ , is  $\delta_{(\xi_r | 0 \leq r \leq t)}$ .*

For each  $t \in [0, T]$ , the same construction as the one carried on in Theorem 3.1 in [20] allows us to define the unique solution to

$$u^{m_t}(s, y) = \int_{\mathcal{C}_t^d} K(y - X_s(\omega)) \exp \left( \int_0^s \Lambda(r, X_r(\omega), u^{m_t}(r, X_r(\omega))) dr \right) m_t(d\omega) \quad \forall s \in [0, t]. \quad (2.17)$$

The proposition and corollary below were the object of Proposition 3.7 and Corollary 3.8 in [20].

**Proposition 2.11.** *Under Assumption 1, we have*

$$\forall (s, y) \in [0, t] \times \mathbb{R}^d, u^m(s, y) = u^{m^t}(s, y).$$

**Corollary 2.12.** *Let  $N \in \mathbb{N}$ ,  $\xi^1, \dots, \xi^i, \dots, \xi^N$  be  $(\mathcal{G}_t)$ -adapted continuous processes, where  $\mathcal{G}$  is a filtration (defined on some probability space) fulfilling the usual conditions. Let  $m(d\omega) = \frac{1}{N} \sum_{i=1}^N \delta_{\xi^i}(d\omega)$ . Then,  $(u^m(t, y))$  is a  $(\mathcal{G}_t)$ -adapted random field, i.e. for any  $(t, y) \in [0, T] \times \mathbb{R}^d$ , the process is  $(\mathcal{G}_t)$ -adapted.*

### 3 Particle systems approximation and propagation of chaos

In this section, we introduce an interacting particle system  $\xi = (\xi^{i,N})_{i=1, \dots, N}$  whose empirical law will be shown to converge to the law of the solution  $Y$  of the McKean type equation (1.4). A consequence of the so called *propagation of chaos* which describes the asymptotic independence of the components of  $\xi$  when the size  $N$  of the particle system goes to  $\infty$ . That property was introduced in [21] and further developed and popularized by [27]. The convergence of  $(\xi^{i,N})_{i=1, \dots, N}$  induces a natural approximation of  $u$ , solution of (1.4).

We suppose here the validity of Assumption 1. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a fixed probability space, and  $(W^i)_{i=1, \dots, N}$  be a sequence of independent  $\mathbb{R}^p$ -valued Brownian motions. Let  $(Y_0^i)_{i=1, \dots, N}$  be i.i.d. r.v. according to  $\zeta_0$ . We consider  $\mathbf{Y} := (Y^i)_{i=1, \dots, N}$  the sequence of processes such that  $(Y^i, u^{m^i})$  are solutions to

$$\begin{cases} Y_t^i = Y_0^i + \int_0^t \Phi(s, Y_s^i, u_s^{m^i}(Y_s^i)) dW_s^i + \int_0^t g(s, Y_s^i, u_s^{m^i}(Y_s^i)) ds \\ u_t^{m^i}(y) = \mathbb{E} \left[ K(y - Y_t^i) V_t(Y^i, u^{m^i}(Y^i)) \right], \quad \text{with } m^i := \mathcal{L}(Y^i), \end{cases} \quad (3.1)$$

recalling that  $V_t(Y^i, u^{m^i}(Y^i)) = \exp \left( \int_0^t \Lambda_s(Y_s^i, u_s^{m^i}(Y_s^i)) ds \right)$ . The existence and uniqueness of the solution of each equation is ensured by Theorem 2.8. We recall that the map  $(m, t, y) \mapsto u^m(t, y)$  fulfills the regularity properties given at the second and third item of Proposition 2.5.

Obviously the processes  $(Y^i)_{i=1, \dots, N}$  are independent. They are also identically distributed since Theorem 2.8 also states uniqueness in law.

So we can define  $m^0 := m^i$  the common distribution of the processes  $Y^i, i = 1, \dots, N$ , which is of course the law of the process  $Y$ , such that  $(Y, u)$  is a solution of (1.4).

From now on,  $\mathcal{C}^{dN}$  will denote  $(\mathcal{C}^d)^N$ , which is obviously isomorphic to  $\mathcal{C}([0, T], \mathbb{R}^{dN})$ . For every  $\bar{\xi} \in \mathcal{C}^{dN}$  we will denote

$$S^N(\bar{\xi}) := \frac{1}{N} \sum_{i=1}^N \delta_{\bar{\xi}^i, N}. \quad (3.2)$$

The function  $(t, x) \mapsto u_t^{S^N(\bar{\xi})}(x)$  is obtained by composition of  $m \mapsto u_t^m(x)$  (defined in (2.8)) with  $m = S^N(\bar{\xi})$ .

Now let us introduce the system of equations

$$\begin{cases} \xi_t^{i,N} = \xi_0^{i,N} + \int_0^t \Phi(s, \xi_s^{i,N}, u_s^{S^N(\xi)}(\xi_s^{i,N})) dW_s^i + \int_0^t g(s, \xi_s^{i,N}, u_s^{S^N(\xi)}(\xi_s^{i,N})) ds \\ \xi_0^{i,N} = Y_0^i \\ u_t^{S^N(\xi)}(y) = \frac{1}{N} \sum_{j=1}^N K(y - \xi_t^{j,N}) V_t(\xi^{j,N}, u^{S^N(\xi)}(\xi^{j,N})). \end{cases} \quad (3.3)$$

Conformally with (3.2), we consider the empirical (random) measure  $S^N(\mathbf{Y}) = \frac{1}{N} \sum_{i=1}^N \delta_{Y^i}$  related to  $\mathbf{Y} := (Y^i)_{i=1, \dots, N}$ , where we recall that for each  $i \in \{1, \dots, N\}$ ,  $Y^i$  is solution of (3.1). We observe that by Remark

2.3,  $S^N(\xi)$  and  $S^N(\mathbf{Y})$  are measurable maps from  $(\Omega, \mathcal{F})$  to  $(\mathcal{P}(\mathcal{C}^d), \mathcal{B}(\mathcal{P}(\mathcal{C}^d)))$ ; moreover  $S^N(\xi), S^N(\mathbf{Y}) \in \mathcal{P}_2(\mathcal{C}^d)$   $\mathbb{P}$ -a.s. A solution  $\xi := (\xi^{i,N})_{i=1, \dots, N}$  of (3.3) is called **interacting particle system**.

The first line of (3.3) is in fact a path-dependent stochastic differential equation. We claim that its coefficients are measurable. Indeed, the map  $(t, \bar{\xi}) \mapsto (S^N(\bar{\xi}), t, \bar{\xi}_t^i)$  being continuous from  $([0, T] \times \mathcal{C}^{dN}, \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathcal{C}^{dN}))$  to  $(\mathcal{P}(\mathcal{C}^d) \times [0, T] \times \mathbb{R}^d, \mathcal{B}(\mathcal{P}(\mathcal{C}^d)) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}^d))$  for all  $i \in \{1, \dots, N\}$ , by composition with the continuous map  $(m, t, y) \mapsto u^m(t, y)$  (see Proposition 2.5 3.) we deduce the continuity of  $(t, \bar{\xi}) \mapsto (u_t^{S^N(\bar{\xi})}(\bar{\xi}_t^i))_{i=1, \dots, N}$ , and so the measurability from  $([0, T] \times \mathcal{C}^{dN}, \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathcal{C}^{dN}))$  to  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . In the sequel, for simplicity we set  $\bar{\xi}_{r \leq s} := (\bar{\xi}_{r \leq s}^i)_{1 \leq i \leq N}$ . We remark that, by Proposition 2.11 and Remark 2.10, we have

$$\left( u_s^{S^N(\bar{\xi})}(\bar{\xi}_s^i) \right)_{i=1, \dots, N} = \left( u_s^{S^N(\bar{\xi}_{r \leq s})}(\bar{\xi}_s^i) \right)_{i=1, \dots, N}, \quad (3.4)$$

for any  $s \in [0, T]$ ,  $\bar{\xi} \in \mathcal{C}^{dN}$  and so stochastic integrands of (3.3) are adapted (so progressively measurable being continuous in time) and so the corresponding Itô integral makes sense. We discuss below the well-posedness of (3.3).

The fact that (3.3) has a unique (strong) solution  $(\xi^{i,N})_{i=1, \dots, N}$  holds true because of the following arguments.

1.  $\Phi$  and  $g$  are Lipschitz. Moreover the map  $\bar{\xi}_{r \leq s} \mapsto \left( u_s^{S^N(\bar{\xi}_{r \leq s})}(\bar{\xi}_s^i) \right)_{i=1, \dots, N}$  is Lipschitz.

Indeed, for given  $(\xi_{r \leq s}, \eta_{r \leq s}) \in \mathcal{C}^{dN} \times \mathcal{C}^{dN}$ ,  $s \in [0, T]$ , by using successively inequality (2.9) of Proposition 2.5 and Remark 2.1, for all  $i \in \{1, \dots, N\}$  we have

$$\begin{aligned} |u_s^{S^N(\xi_{r \leq s})}(\xi_s^i) - u_s^{S^N(\eta_{r \leq s})}(\eta_s^i)| &\leq \sqrt{C_{K, \Lambda}(T)} \left( |\xi_s^i - \eta_s^i| + \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq r \leq s} |\xi_r^j - \eta_r^j| \right) \\ &\leq 2\sqrt{C_{K, \Lambda}(T)} \max_{j=1, \dots, N} \sup_{0 \leq r \leq s} |\xi_r^j - \eta_r^j|. \end{aligned} \quad (3.5)$$

Finally the functions

$$\begin{aligned} \bar{\xi}_{r \leq s} &\mapsto \left( \Phi(s, \bar{\xi}_s^i, u_s^{S^N(\bar{\xi}_{r \leq s})}(\bar{\xi}_s^i)) \right)_{i=1, \dots, N} \\ \bar{\xi}_{r \leq s} &\mapsto \left( g(s, \bar{\xi}_s^i, u_s^{S^N(\bar{\xi}_{r \leq s})}(\bar{\xi}_s^i)) \right)_{i=1, \dots, N} \end{aligned}$$

are uniformly Lipschitz and bounded.

2. A classical argument of well-posedness for systems of path-dependent stochastic differential equations with Lipschitz dependence on the sup-norm of the path, see Chapter V, Section 2.11, Theorem 11.2 page 128 in [25].

After the preceding introductory considerations, we can state and prove the main theorem of the section.

**Theorem 3.1.** *Let us suppose the validity of Assumption 1. Let  $N$  be a fixed positive integer. Let  $(Y^i)_{i=1, \dots, N}$  (resp.  $(\xi^{i,N})_{i=1, \dots, N}$ ) be the solution of (3.1) (resp. (3.3)), let  $m^0$  as defined after (3.1). The following assertions hold.*

1. *If  $\mathcal{F}(K) \in L^1(\mathbb{R}^d)$ , there is a positive constant  $C$  only depending on  $L_\Phi, L_g, M_K, M_\Lambda, L_K, L_\Lambda, T, \|\mathcal{F}(K)\|_1$ , such that, for all  $i = 1, \dots, N$  and  $t \in [0, T]$ ,*

$$\mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_\infty^2] \leq \frac{C}{N} \quad (3.6)$$

$$\mathbb{E}[\sup_{0 \leq s \leq t} |\xi_s^{i,N} - Y_s^i|^2] \leq \frac{C}{N}. \quad (3.7)$$

2. If  $K$  belongs to  $W^{1,2}(\mathbb{R}^d)$ , there is a positive constant  $C$  only depending on  $L_\Phi, L_g, M_K, M_\Lambda, L_K, L_\Lambda, T$  and  $\|\nabla K\|_2$ , such that, for all  $t \in [0, T]$ ,

$$\mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_2^2] \leq \frac{C}{N}. \quad (3.8)$$

Before proving Theorem 3.1, we remark that the propagation of chaos follows easily.

**Corollary 3.2.** *Under Assumption 1, the propagation of chaos holds for the interacting particle system  $(\xi^{i,N})_{i \in \mathbb{N}}$ .*

*Proof.* We prove here that Theorem 3.1 implies the propagation of chaos.

Indeed, for all  $k \in \mathbb{N}^*$ , (3.7) implies

$$(\xi^{1,N} - Y^1, \xi^{2,N} - Y^2, \dots, \xi^{k,N} - Y^k) \xrightarrow[N \rightarrow +\infty]{L^2(\Omega, \mathcal{F}, \mathbb{P})} 0,$$

which implies in particular the convergence in law of the vector  $(\xi^{1,N}, \xi^{2,N}, \dots, \xi^{k,N})$  to  $(Y^1, Y^2, \dots, Y^k)$ . Consequently, since  $(Y^i)_{i=1, \dots, k}$  are i.i.d. according to  $m^0$

$$(\xi^{1,N}, \xi^{2,N}, \dots, \xi^{k,N}) \text{ converges in law to } (m^0)^{\otimes k} \text{ when } N \rightarrow +\infty. \quad (3.9)$$

□

The validity of (3.6) and (3.7) will be the consequence of the significant more general proposition below.

**Proposition 3.3.** *Let us suppose the validity of Assumption 1. Let  $N$  be a fixed positive integer. Let  $(W^{i,N})_{i=1, \dots, N}$  be a family of  $p$ -dimensional standard Brownian motions (not necessarily independent). Let  $(Y_0^i)_{i=1, \dots, N}$  be the family of i.i.d. r.v. initializing the system (3.1). We consider the processes  $(\bar{Y}^{i,N})_{i=1, \dots, N}$ , such that for each  $i \in \{1, \dots, N\}$ ,  $\bar{Y}^{i,N}$  is the unique strong solution of*

$$\begin{cases} \bar{Y}_t^{i,N} = Y_0^i + \int_0^t \Phi(s, \bar{Y}_s^{i,N}, u_s^{m^{i,N}}(\bar{Y}_s^{i,N})) dW_s^{i,N} + \int_0^t g(s, \bar{Y}_s^{i,N}, u_s^{m^{i,N}}(\bar{Y}_s^{i,N})) ds, & \text{for all } t \in [0, T] \\ u_t^{m^{i,N}}(y) = \mathbb{E} \left[ K(y - \bar{Y}_t^{i,N}) V_t(\bar{Y}^{i,N}, u^{m^{i,N}}(\bar{Y}^{i,N})) \right], & \text{with } m^{i,N} := \mathcal{L}(\bar{Y}^{i,N}), \end{cases} \quad (3.10)$$

recalling that  $V_t(Y^{i,N}, u^{m^{i,N}}(Y^{i,N})) = \exp \left( \int_0^t \Lambda(s, Y_s^{i,N}, u_s^{m^{i,N}}(Y_s^{i,N})) ds \right)$ .

Let us consider now the system of equations (3.3), where the processes  $W^i$  are replaced by  $W^{i,N}$ , i.e.

$$\begin{cases} \xi_t^{i,N} = \xi_0^{i,N} + \int_0^t \Phi(s, \xi_s^{i,N}, u_s^{S^N(\xi)}(\xi_s^{i,N})) dW_s^{i,N} + \int_0^t g(s, \xi_s^{i,N}, u_s^{S^N(\xi)}(\xi_s^{i,N})) ds \\ \xi_0^{i,N} = Y_0^i \\ u_t^{S^N(\xi)}(y) = \frac{1}{N} \sum_{j=1}^N K(y - \xi_t^{j,N}) V_t(\xi^{j,N}, u^{S^N(\xi)}(\xi^{j,N})). \end{cases} \quad (3.11)$$

Then the following assertions hold.

1. For any  $i = 1, \dots, N$ ,  $(\bar{Y}_t^{i,N})_{t \in [0, T]}$  have the same law  $m^{i,N} = m^0$ , where  $m^0$  is the common law of processes  $(Y^i)_{i=1, \dots, N}$  defined by the system (3.1).
2. Equation (3.11) admits a unique strong solution.
3. Suppose moreover that  $\mathcal{F}(K)$  is in  $L^1(\mathbb{R}^d)$ . Then there is a positive constant  $C$  only depending on  $L_\Phi, L_g, M_K, M_\Lambda, L_K, L_\Lambda, T$  and  $\|\mathcal{F}(K)\|_1$  such that, for all  $t \in [0, T]$ ,

$$\sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] + \mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_\infty^2] \leq C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|S^N(\bar{\mathbf{Y}}) - m^0, \varphi|^2], \quad (3.12)$$

with again  $S^N(\bar{\mathbf{Y}}) := \frac{1}{N} \sum_{j=1}^N \delta_{\bar{Y}^{j,N}}$ .

**Remark 3.4.** The convergence of the numerical approximation  $u_t^{S^N(\xi)}$  to  $u_t^{m^0}$  only requires the convergence of  $d_2^\Omega(S^N(\bar{Y}), m^0)$  to 0, where the distance  $d_2^\Omega$  has been defined at Remark 2.6. This holds if, for each  $N$ ,  $\bar{Y}^{i,N}$ ,  $i = 1, \dots, N$  are independent; however, this is only a sufficient condition.

This gives the opportunity to define new numerical schemes for which the convergence of the empirical measure  $S^N(\bar{Y})$  is verified without i.i.d. particles. Let us consider  $(\bar{Y}^{i,N})_{i=1, \dots, N}$  (resp.  $(\xi^{i,N})_{i=1, \dots, N}$ ) solutions of (3.10) (resp. (3.11)). Observe that for any real valued test function in  $C_b(\mathcal{C}^d)$

$$\mathbb{E}[\langle S^N(\bar{Y}) - m^0, \varphi \rangle^2] = \frac{\sigma_\varphi^2}{N} \left(1 + \frac{2}{N} \sum_{i < j} \rho_\varphi^{i,j}\right),$$

where  $\sigma_\varphi := \sqrt{\text{Var}(\varphi(\bar{Y}^{1,N}))}$  and  $\rho_\varphi^{i,j} := \frac{\mathbb{E}[\varphi(\bar{Y}^{i,N})\varphi(\bar{Y}^{j,N})] - \mathbb{E}[\varphi(\bar{Y}^{i,N})]\mathbb{E}[\varphi(\bar{Y}^{j,N})]}{\sigma_\varphi^2}$ .

In the specific case where  $(W^{i,N})_{i=1, \dots, N}$  are independent Brownian motions then  $\rho_\varphi^{i,j} = 0$  for any bounded  $\varphi \in C_b(\mathcal{C}^d)$  and

$$\sup_{\substack{\varphi \in C_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2] \leq \frac{1}{N}. \quad (3.13)$$

With our error bound one can naturally investigate antithetic variables approaches to improve the interacting particle system convergence. Let us consider  $N = 2N'$  and take  $(W^{i,N})_{i=1, \dots, N'}$  as  $N'$  iid Brownian motions, then for the rest of the particles, for any  $j = N' + 1, N' + 2, \dots, 2N'$ , set  $W^{j,N} = -W^{j-N',N}$ . In this situation, we obtain

$$\mathbb{E}[\langle S^N(\bar{Y}) - m^0, \varphi \rangle^2] = \frac{\sigma_\varphi^2}{N} (1 + \rho_\varphi^{1,1+N'}).$$

So, even in this case, the rate of convergence of  $u_t^{S^N(\xi)}$  to  $u_t^{m^0}$  is still of order  $1/\sqrt{N}$ .

If moreover one has  $\sup_{\substack{\varphi \in C_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \rho_\varphi^{1,1+N'} \leq 0$ , the variance will also be reduced with respect to the case of independent Brownian motions, see (3.13).

*Proof of Proposition 3.3.* Let us fix  $t \in [0, T]$ . In this proof,  $C := C(\Phi, g, \Lambda, K, T)$  is a real positive constant, which may change from line to line.

Equation (3.10) has  $N$  blocks, numbered by  $1 \leq i \leq N$ . Theorem 2.8 gives uniqueness in law for each block equation, which implies that for any  $i = 1, \dots, N$ ,  $m^{i,N} = m^0$  and proves the first item.

Concerning item 2., i.e. the strong existence and pathwise uniqueness of (3.11), the same argument as for the well-statement of (3.3) operates. The only difference consists in the fact that the Brownian motions may be correlated. A very close proof to the one of Theorem 11.2 page 128 in [25] works: the main argument is the multidimensional BDG inequality, see e.g. Problem 3.29 of [18].

We discuss now item 3. proving inequality (3.12). On the one hand, since the map  $(t, \bar{\xi}) \in [0, T] \times \mathcal{C}^{dN} \mapsto (u_t^{S^N(\bar{\xi})}(\bar{\xi}_t^i))_{i=1, \dots, N}$  is measurable and satisfies the non-anticipative property (3.4), the first assertion of Lemma 2.7 gives for all  $i \in \{1, \dots, N\}$

$$\begin{aligned} \mathbb{E}\left[\sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2\right] &\leq C \mathbb{E}\left[\int_0^t |u_s^{S^N(\xi)}(\xi_s^{i,N}) - u_s^{m^0}(\bar{Y}_s^{i,N})|^2 ds\right] \\ &\leq C \int_0^t \mathbb{E}[|u_s^{S^N(\xi)}(\xi_s^{i,N}) - u_s^{m^0}(\xi_s^{i,N})|^2] ds + \int_0^t \mathbb{E}[|u_s^{m^0}(\xi_s^{i,N}) - u_s^{m^0}(\bar{Y}_s^{i,N})|^2] ds \\ &\leq C \int_0^t \left( \mathbb{E}[\|u_s^{S^N(\xi)} - u_s^{m^0}\|_\infty^2] + \mathbb{E}\left[\sup_{0 \leq r \leq s} |\xi_r^{i,N} - \bar{Y}_r^{i,N}|^2\right] \right) ds, \quad \text{by (2.9),} \end{aligned} \quad (3.14)$$

which implies

$$\sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] \leq C \int_0^t \left( \mathbb{E}[\|u_s^{S^N(\xi)} - u_s^{m^0}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq r \leq s} |\xi_r^{i,N} - \bar{Y}_r^{i,N}|^2 \right] \right) ds. \quad (3.15)$$

We use inequalities (2.9) for  $m = S^N(\xi)(\bar{\omega})$  and  $m' = S^N(\bar{Y})(\bar{\omega})$ , where  $\bar{\omega}$  is a random realization in  $\Omega$  and (2.12) (with the random probability  $\eta = S^N(\bar{Y})$  and  $m = m^0$ ) in item 5. of Proposition 2.5. This yields

$$\begin{aligned} \mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_\infty^2] &\leq 2\mathbb{E} \left[ \|u_t^{S^N(\xi)} - u_t^{S^N(\bar{Y})}\|_\infty^2 \right] + 2\mathbb{E}[\|u_t^{S^N(\bar{Y})} - u_t^{m^0}\|_\infty^2] \\ &\leq 2C\mathbb{E}[|W_t(S^N(\xi), S^N(\bar{Y}))|^2] + 2C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2] \\ &\leq \frac{2C}{N} \sum_{i=1}^N \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] + C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2] \\ &\leq 2C \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] \\ &\quad + C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2], \end{aligned} \quad (3.16)$$

where the third inequality follows from Remark 2.1.

Let us introduce the non-negative function  $G$  defined on  $[0, T]$  by

$$G(t) := \mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right].$$

From inequalities (3.15) and (3.16) that are valid for all  $t \in [0, T]$ , we obtain

$$\begin{aligned} G(t) &\leq (2C + 1) \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] + C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2] \\ &\leq C \int_0^t \left( \mathbb{E}[\|u_s^{S^N(\xi)} - u_s^{m^0}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq r \leq s} |\xi_r^{i,N} - \bar{Y}_r^{i,N}|^2 \right] \right) ds \\ &\quad + C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2] \\ &\leq C \int_0^t G(s) ds + C \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2]. \end{aligned} \quad (3.17)$$

By Gronwall's lemma, for all  $t \in [0, T]$ , we obtain

$$\mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{0 \leq s \leq t} |\xi_s^{i,N} - \bar{Y}_s^{i,N}|^2 \right] \leq C e^{Ct} \sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2]. \quad (3.18)$$

This concludes the proof of Proposition 3.3.  $\square$

From now on, we prove Theorem 3.1,

*Proof of Theorem 3.1.* As we have mentioned above we will apply Proposition 3.3 setting for all  $i \in \{1, \dots, N\}$ ,  $W^{i,N} := W^i$ . Pathwise uniqueness of systems (3.1) and (3.10) implies  $\bar{Y}^{i,N} = Y^i$  for all  $i \in \{1, \dots, N\}$ . Taking into account (3.12) in Proposition 3.3, in order to establish inequalities (3.6) and (3.7), we need to bound the quantity  $\sup_{\substack{\varphi \in \mathcal{C}_b(\mathcal{C}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\bar{Y}) - m^0, \varphi \rangle|^2]$ . This is possible via (3.13) in Remark 3.4, since  $(Y^i)_{i=1, \dots, N}$  are

i.i.d. according to  $m^0$ . This concludes the proof of item 1.

It remains now to prove (3.8) in item 2. First, the inequality

$$\mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{m^0}\|_2^2] \leq 2\mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{S^N(\mathbf{Y})}\|_2^2] + 2\mathbb{E}[\|u_t^{S^N(\mathbf{Y})} - u_t^{m^0}\|_2^2], \quad (3.19)$$

holds for all  $t \in [0, T]$ . Using inequality (2.11) of Proposition 2.5, for all  $t \in [0, T]$ , for  $m = S^N(\xi)$ ,  $m' = S^N(\mathbf{Y})$ , we get

$$\begin{aligned} \mathbb{E}[\|u_t^{S^N(\xi)} - u_t^{S^N(\mathbf{Y})}\|_2^2] &\leq C\mathbb{E}[W_t(S^N(\xi), S^N(\mathbf{Y}))^2] \\ &\leq C\frac{1}{N}\sum_{j=1}^N\mathbb{E}[\sup_{0 \leq r \leq t} |\xi_r^{j,N} - Y_r^j|^2] \\ &\leq \frac{C}{N}, \end{aligned} \quad (3.20)$$

where the latter inequality is obtained through (3.7). The second term of the r.h.s. in (3.19) needs more computations. Let us fix  $i \in \{1, \dots, N\}$ . First,

$$\mathbb{E}[\|u_t^{S^N(\mathbf{Y})} - u_t^{m^0}\|_2^2] \leq 2(\mathbb{E}[\|A_t\|_2^2] + \mathbb{E}[\|B_t\|_2^2]), \quad (3.21)$$

where, for all  $t \in [0, T]$

$$\begin{cases} A_t(x) := \frac{1}{N}\sum_{j=1}^N K(x - Y_t^j) [V_t(Y^j, u^{S^N(\mathbf{Y})}(Y^j)) - V_t(Y^j, u^{m^0}(Y^j))] \\ B_t(x) := \frac{1}{N}\sum_{j=1}^N K(x - Y_t^j) V_t(Y^j, u^{m^0}(Y^j)) - \mathbb{E}[K(x - Y_t^1) V_t(Y^1, u^{m^0}(Y^1))], \end{cases} \quad (3.22)$$

where we recall that  $m^0$  is the common law of all the processes  $Y^i$ ,  $1 \leq i \leq N$ .

To simplify notations, we set  $P_j(t, x) := K(x - Y_t^j) V_t(Y^j, u^{m^0}(Y^j)) - \mathbb{E}[K(x - Y_t^1) V_t(Y^1, u^{m^0}(Y^1))]$  for all  $j \in \{1, \dots, N\}$ ,  $x \in \mathbb{R}^d$  and  $t \in [0, T]$ .

We observe that for all  $x \in \mathbb{R}^d$ ,  $t \in [0, T]$ ,  $(P_j(t, x))_{j=1, \dots, N}$  are i.i.d. centered r.v. Hence,

$$\mathbb{E}[B_t(x)^2] = \frac{1}{N}\mathbb{E}[P_1^2(t, x)] = \frac{1}{N}\text{Var}(P_1(t, x)) \leq \frac{1}{N}\mathbb{E}[K^2(x - Y_t^1) V_t^2(Y^1, u^{m^0}(Y^1))] \leq \frac{M_K e^{2tM_\Lambda}}{N}\mathbb{E}[K(x - Y_t^1)].$$

By integrating each side of the inequality above w.r.t.  $x \in \mathbb{R}^d$ , we obtain

$$\mathbb{E}\left[\int_{\mathbb{R}^d} |B_t(x)|^2 dx\right] = \int_{\mathbb{R}^d} \mathbb{E}[|B_t(x)|^2] dx \leq \frac{M_K e^{2tM_\Lambda}}{N}, \quad (3.23)$$

where we have used that  $\|K\|_1 = 1$ .

Concerning  $A_t(x)$ , we write

$$\begin{aligned} |A_t(x)|^2 &\leq \frac{1}{N}\sum_{j=1}^N K(x - Y_t^j)^2 [V_t(Y^j, u^{S^N(\mathbf{Y})}(Y^j)) - V_t(Y^j, u^{m^0}(Y^j))]^2 \\ &= \frac{1}{N}\sum_{j=1}^N K(x - Y_t^j) K(x - Y_t^j) [V_t(Y^j, u^{S^N(\mathbf{Y})}(Y^j)) - V_t(Y^j, u^{m^0}(Y^j))]^2 \\ &\leq \frac{M_K T}{N} e^{2tM_\Lambda} L_\Lambda^2 \sum_{j=1}^N K(x - Y_t^j) \int_0^t |u_r^{S^N(\mathbf{Y})}(Y^j) - u_r^{m^0}(Y^j)|^2 dr \\ &\leq \frac{M_K T}{N} e^{2tM_\Lambda} L_\Lambda^2 \sum_{j=1}^N K(x - Y_t^j) \int_0^t \|u_r^{S^N(\mathbf{Y})} - u_r^{m^0}\|_\infty^2 dr, \end{aligned} \quad (3.24)$$

where the third inequality comes from (2.7). Integrating w.r.t.  $x \in \mathbb{R}^d$  and taking expectation on each side of the above inequality gives us, for all  $t \in [0, T]$ ,

$$\begin{aligned} \mathbb{E}\left[\int_{\mathbb{R}^d} |A_t(x)|^2 dx\right] &\leq M_K T e^{2tM_\Lambda} L_\Lambda^2 \int_0^t \mathbb{E}[\|u_r^{S^N(\mathbf{Y})} - u_r^{m^0}\|_\infty^2] dr \\ &\leq M_K T^2 e^{2tM_\Lambda} L_\Lambda^2 C \sup_{\substack{\varphi \in C_b(\mathbb{R}^d) \\ \|\varphi\|_\infty \leq 1}} \mathbb{E}[|\langle S^N(\mathbf{Y}) - m^0, \varphi \rangle|^2] \\ &\leq \frac{M_K T^2 e^{2tM_\Lambda} L_\Lambda^2 C}{N}, \end{aligned} \quad (3.25)$$

where we have used (2.12) of Proposition 2.5 for the second inequality above and (3.13) for the latter one. To conclude, it is enough to replace (3.23), (3.25) in (3.21), and inject (3.20), (3.21) into (3.19).  $\square$

## 4 Particle algorithm

### 4.1 Time discretization of the particle system

In this section Assumption 2. is in force. Let  $(Y_0^i)_{i=1, \dots, N}$  be i.i.d. r.v. distributed according to  $\zeta_0$ . In the sequel, we are interested in discretizing the interacting particle system (3.3).  $(\xi^{i,N}, 1 \leq i \leq N)$  will denote again the corresponding solution. Let us consider a regular time grid  $0 = t_0 \leq \dots \leq t_k = k\delta t \leq \dots \leq t_n = T$ , with  $\delta t = T/n$ . We introduce the continuous  $\mathbb{R}^{dN}$ -valued process  $(\tilde{\xi}_t)_{t \in [0, T]}$  and the family of nonnegative functions  $(\tilde{u}_t)_{t \in [0, T]}$  defined on  $\mathbb{R}^d$  constructively such that

$$\begin{cases} \tilde{\xi}_t^{i,N} = \tilde{\xi}_0^{i,N} + \int_0^t \Phi(r(s), \tilde{\xi}_{r(s)}^{i,N}, \tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N})) dW_s^i + \int_0^t g(r(s), \tilde{\xi}_{r(s)}^{i,N}, \tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N})) ds \\ \tilde{\xi}_0^{i,N} = Y_0^i \\ \tilde{u}_t(y) = \frac{1}{N} \sum_{j=1}^N K(y - \tilde{\xi}_t^{j,N}) \exp \left\{ \int_0^t \Lambda(r(s), \tilde{\xi}_{r(s)}^{j,N}, \tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{j,N})) ds \right\}, \text{ for any } t \in ]0, T], \\ \tilde{u}_0 = K * \zeta_0, \end{cases} \quad (4.1)$$

where  $r : s \in [0, T] \mapsto r(s) \in \{t_0, \dots, t_n\}$  is the piecewise constant function such that  $r(s) = t_k$  when  $s \in [t_k, t_{k+1}[$ . We can observe that  $(\tilde{\xi}^{i,N})_{i=1, \dots, N}$  is an adapted and continuous process. The interacting particle system  $(\tilde{\xi}^{i,N})_{i=1, \dots, N}$  can be simulated perfectly at the discrete instants  $(t_k)_{k=0, \dots, n}$  via independent standard and centered Gaussian random variables. We will show that this interacting particle system provides an approximation to the solution  $(\xi^{i,N})_{i=1, \dots, N}$ , of system (3.3), which converges at a rate bounded by  $\sqrt{\delta t}$ , up to a multiplicative constant.

**Proposition 4.1.** *Let us suppose the validity of Assumption 2. The time discretized particle system (4.1) converges to the original particle system (3.3). More precisely, for all  $t \in [0, T]$ , the following estimates hold:*

$$\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E} \left[ \sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2 \right] \leq C \delta t, \quad (4.2)$$

where  $C$  is a finite positive constant only depending on  $M_K, M_\Lambda, L_K, L_\Lambda, m_\Phi, m_g, T$ .

If we assume moreover that  $K \in W^{1,2}(\mathbb{R}^d)$ , then

$$\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_2^2] \leq C \delta t, \quad t \in [0, T], \quad (4.3)$$

where  $C$  is a finite positive constant only depending on  $M_K, M_\Lambda, m_\Phi, m_g, L_K, L_\Lambda, T$  and  $\|\nabla K\|_2$ .

The left-hand side of (4.3) is generally known, as Mean Integrated Squared Error (MISE).



The result below states the convergence of  $\tilde{u}_t$  to  $u_t^{m^0}$  when  $\delta t \rightarrow 0$  and  $N \rightarrow +\infty$ , with an explicit rate of convergence.

**Theorem 4.2.** *We suppose Assumption 2. We indicate by  $m^0$  the law of  $Y$ , where  $(Y, u)$  is the solution of (1.4). The time discretized particle system (4.1) converges to the solution of (1.4). More precisely, we have the following. We suppose  $\mathcal{F}(K) \in L^1$  (resp.  $K \in W^{1,2}(\mathbb{R}^d)$ ). There exists a real constant  $C > 0$  such that for all  $t \in [0, T]$ ,*

$$\mathbb{E}[\|u_t^{m^0} - \tilde{u}_t\|_\infty^2] \leq C(\delta t + \frac{1}{N}), \quad (4.4)$$

(respectively

$$\mathbb{E}[\|u_t^{m^0} - \tilde{u}_t\|_2^2] \leq C(\delta t + \frac{1}{N})). \quad (4.5)$$

**Remark 4.3.** *When  $\Lambda = 0$  and  $\Phi$  and  $g$  are infinitely differentiable with all derivatives being bounded, Corollary 1.1 of [19] states that, for fixed smooth test function with polynomial growth  $\varphi$ , one has*

$$\mathbb{E}(|\langle S^N(\tilde{\xi}_t) - m_t^0, \varphi \rangle|) \leq C_\varphi(\frac{1}{\sqrt{N}} + \delta t), \quad \text{where again } S^N(\tilde{\xi}_t) := \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{\xi}_t^{i,N}}. \quad (4.6)$$

This leads reasonably to the conjecture that the rate in (4.4) is not optimal and it could be replaced by  $(\delta t)^2 + \frac{1}{N}$ . This intuition will be confirmed by numerical simulations in Section 5.

*Proof.* We first observe that for all  $t \in [0, T]$ , for  $p = \infty, 2$

$$\mathbb{E}[\|u_t^{m^0} - \tilde{u}_t\|_p^2] \leq 2\mathbb{E}[\|u_t^{m^0} - u_t^{S^N(\xi)}\|_p^2] + 2\mathbb{E}[\|u_t^{S^N(\xi)} - \tilde{u}_t\|_p^2]. \quad (4.7)$$

The first term in the r.h.s. of (4.7) is bounded by  $\frac{C}{N}$  using Theorem 3.1, inequality (3.6) (respectively (3.8)). The second term of the same inequality is controlled by  $C\delta t$ , through Proposition 4.1, inequality (4.2) (resp. (4.3)).  $\square$

The proof of Proposition 4.1 relies on similar techniques used to prove Theorem 3.1. The idea is first to estimate through Lemma 4.4 the perturbation error due to the time discretization scheme of the SDE and of the integral appearing in the exponential weight in system (4.1). Later the propagation of this error through the dynamical system (3.3) will be controlled via Gronwall's lemma. Lemma 4.4 below will be proved in the Appendix.

**Lemma 4.4.** *Let us suppose the validity of Assumption 2. There exists a finite constant  $C > 0$  only depending on  $T, M_K, m_\Phi, m_g, L_K, L_\Phi, L_g$  and  $M_\Lambda, L_\Lambda$  such that for any  $t \in [0, T]$ ,*

$$\mathbb{E}[|\tilde{\xi}_{r(t)}^{i,N} - \xi_t^{i,N}|^2] \leq C\delta t \quad (4.8)$$

$$\mathbb{E}[\|\tilde{u}_{r(t)} - \tilde{u}_t\|_\infty^2] \leq C\delta t \quad (4.9)$$

$$\mathbb{E}[\|\tilde{u}_{r(t)} - u_t^{S^N(\tilde{\xi})}\|_\infty^2] \leq C\delta t. \quad (4.10)$$

*Proof of Proposition 4.1.* All along this proof,  $C$  will denote a positive constant that only depends on  $T, M_K, m_\Phi, m_g, L_K, L_\Phi, L_g$  and  $M_\Lambda, L_\Lambda$  and that can change from line to line. Let us fix  $t \in [0, T]$ .

- We begin by considering inequality (4.2). We first fix  $1 \leq i \leq N$ . By (4.9) and (4.10) in Lemma 4.4 and (2.9) in Proposition 2.5, we obtain

$$\begin{aligned} \mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_\infty^2] &\leq \mathbb{E}\left[\left(\|\tilde{u}_t - \tilde{u}_{r(t)}\|_\infty + \|\tilde{u}_{r(t)} - u_t^{S^N(\tilde{\xi})}\|_\infty + \|u_t^{S^N(\tilde{\xi})} - u_t^{S^N(\xi)}\|_\infty\right)^2\right] \\ &\leq 3(\mathbb{E}[\|\tilde{u}_t - \tilde{u}_{r(t)}\|_\infty^2] + \mathbb{E}[\|\tilde{u}_{r(t)} - u_t^{S^N(\tilde{\xi})}\|_\infty^2] + \mathbb{E}[\|u_t^{S^N(\tilde{\xi})} - u_t^{S^N(\xi)}\|_\infty^2]) \\ &\leq C\delta t + C\mathbb{E}[|W_t(S^N(\tilde{\xi}), S^N(\xi))|^2] \\ &\leq C\delta t + C \sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2], \end{aligned} \quad (4.11)$$

where the function  $u^{S^N(\tilde{\xi})}$  makes sense since  $\tilde{\xi}$  has almost surely continuous trajectories and so  $S^N(\tilde{\xi})$  is a random probability in  $\mathcal{P}(\mathcal{C}^d)$ .

Besides, by the second assertion of Lemma 2.7, setting  $Y' := \tilde{\xi}^{i,N}$ ,  $r_1(s) = r(s)$  and  $Y := \xi^{i,N}$ ,  $r_2(s) = s$ , we get

$$\mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] \leq C \mathbb{E} \left[ \int_0^t |\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\xi)}(\xi_s^{i,N})|^2 ds \right] + C \int_0^t \mathbb{E} \left[ |\tilde{\xi}_{r(s)}^{i,N} - \xi_s^{i,N}|^2 \right] ds + C \delta t. \quad (4.12)$$

Concerning the first term in the r.h.s. of (4.12), we have for all  $s \in [0, T]$

$$\begin{aligned} |\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\xi)}(\xi_s^{i,N})|^2 &\leq 2|\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\xi)}(\tilde{\xi}_{r(s)}^{i,N})|^2 + 2|u_s^{S^N(\xi)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\xi)}(\xi_s^{i,N})|^2 \\ &\leq 2\|\tilde{u}_{r(s)} - u_s^{S^N(\xi)}\|_\infty^2 + 2C|\tilde{\xi}_{r(s)}^{i,N} - \xi_s^{i,N}|^2, \end{aligned} \quad (4.13)$$

where the second inequality above follows by (2.9) in Proposition 2.5, setting  $m = m' = S^N(\xi)$ . Consequently, by (4.12)

$$\begin{aligned} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] &\leq C \left\{ \mathbb{E} \left[ \int_0^t \|\tilde{u}_{r(s)} - u_s^{S^N(\xi)}\|_\infty^2 ds \right] + \int_0^t \mathbb{E} \left[ |\tilde{\xi}_{r(s)}^{i,N} - \xi_s^{i,N}|^2 \right] ds + \delta t \right\} \\ &\leq C \left\{ \mathbb{E} \left[ \int_0^t \|\tilde{u}_{r(s)} - \tilde{u}_s\|_\infty^2 ds \right] + \mathbb{E} \left[ \int_0^t \|\tilde{u}_s - u_s^{S^N(\xi)}\|_\infty^2 ds \right] \right. \\ &\quad \left. + \mathbb{E} \left[ \int_0^t |\tilde{\xi}_{r(s)}^{i,N} - \tilde{\xi}_s^{i,N}|^2 ds \right] + \mathbb{E} \left[ \int_0^t |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2 ds \right] + \delta t \right\}. \end{aligned} \quad (4.14)$$

Using inequalities (4.8) and (4.9) in Lemma 4.4, for all  $t \in [0, T]$ , we obtain

$$\sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] \leq C \delta t + C \int_0^t \left[ \mathbb{E}[\|\tilde{u}_s - u_s^{S^N(\xi)}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E}[\sup_{\theta \leq s} |\tilde{\xi}_\theta^{i,N} - \xi_\theta^{i,N}|^2] \right] ds. \quad (4.15)$$

Gathering the latter inequality together with (4.11) yields

$$\begin{aligned} \mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_\infty^2] + \sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] &\leq C \delta t + 2C \sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] \\ &\leq C \delta t \\ &\quad + C \int_0^t \left[ \mathbb{E}[\|\tilde{u}_s - u_s^{S^N(\xi)}\|_\infty^2] \right. \\ &\quad \left. + \sup_{i=1, \dots, N} \mathbb{E}[\sup_{\theta \leq s} |\tilde{\xi}_\theta^{i,N} - \xi_\theta^{i,N}|^2] \right] ds. \end{aligned} \quad (4.16)$$

Applying Gronwall's lemma to the function

$$t \mapsto \sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] + \mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_\infty^2]$$

ends the proof of (4.2).

- We focus now on (4.3). First we observe that

$$\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\xi)}\|_2^2] \leq 2\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\tilde{\xi})}\|_2^2] + 2\mathbb{E}[\|u_t^{S^N(\tilde{\xi})} - u_t^{S^N(\xi)}\|_2^2]. \quad (4.17)$$

Using successively item 4. of Proposition 2.5, Remark 2.1 and inequality (4.2), we can bound the second term on the r.h.s. of (4.17) as follows:

$$\begin{aligned} \mathbb{E}[\|u_t^{S^N(\tilde{\xi})} - u_t^{S^N(\xi)}\|_2^2] &\leq C \mathbb{E}[|W_t(S^N(\tilde{\xi}), S^N(\xi))|^2] \\ &\leq C \sup_{i=1, \dots, N} \mathbb{E}[\sup_{s \leq t} |\tilde{\xi}_s^{i,N} - \xi_s^{i,N}|^2] \\ &\leq C \delta t. \end{aligned} \quad (4.18)$$

To simplify the notations, we introduce the real valued random variables

$$V_t^i := e^{\int_0^t \Lambda(s, \tilde{\xi}_s^{i,N}, u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_s^{i,N})) ds} \quad \text{and} \quad \tilde{V}_t^i := e^{\int_0^t \Lambda(r(s), \tilde{\xi}_{r(s)}^{i,N}, \tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N})) ds}, \quad (4.19)$$

defined for any  $i = 1, \dots, N$  and  $t \in [0, T]$ .

Concerning the first term on the r.h.s. of (4.17), inequality (6.2) of Lemma 6.1 in the Appendix gives for all  $y \in \mathbb{R}^d$

$$|\tilde{u}_t(y) - u_t^{S^N(\tilde{\xi})}(y)|^2 \leq \frac{M_K}{N} \sum_{i=1}^N K(y - \tilde{\xi}_t^{i,N}) |\tilde{V}_t^i - V_t^i|^2. \quad (4.20)$$

Integrating the inequality (4.20) with respect to  $y$ , yields

$$\|\tilde{u}_t - u_t^{S^N(\tilde{\xi})}\|_2^2 = \int_{\mathbb{R}^d} |\tilde{u}_t(y) - u_t^{S^N(\tilde{\xi})}(y)|^2 dy \leq \frac{M_K}{N} \sum_{i=1}^N |\tilde{V}_t^i - V_t^i|^2,$$

which, in turn, implies

$$\mathbb{E} \left[ \|\tilde{u}_t - u_t^{S^N(\tilde{\xi})}\|_2^2 \right] \leq \frac{M_K}{N} \sum_{i=1}^N \mathbb{E} \left[ |\tilde{V}_t^i - V_t^i|^2 \right]. \quad (4.21)$$

Using successively item 1. of Lemma 6.1 and inequality (4.8) of Lemma 4.4, for all  $i \in \{1, \dots, N\}$ , we obtain

$$\begin{aligned} \mathbb{E}[|\tilde{V}_t^i - V_t^i|^2] &\leq C\delta t + C\mathbb{E} \left[ \int_0^t |\tilde{\xi}_{r(s)}^{i,N} - \tilde{\xi}_s^{i,N}|^2 ds \right] + C\mathbb{E} \left[ \int_0^t |\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_s^{i,N})|^2 ds \right] \\ &\leq C\delta t + C\mathbb{E} \left[ \int_0^t |\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_s^{i,N})|^2 ds \right] \\ &\leq C\delta t + C\mathbb{E} \left[ \int_0^t |\tilde{u}_{r(s)}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_{r(s)}^{i,N})|^2 ds \right] \\ &\quad + C\mathbb{E} \left[ \int_0^t |u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_{r(s)}^{i,N}) - u_s^{S^N(\tilde{\xi})}(\tilde{\xi}_s^{i,N})|^2 ds \right] \\ &\leq C\delta t + C \int_0^t \left[ \mathbb{E}[\|\tilde{u}_{r(s)} - u_s^{S^N(\tilde{\xi})}\|_\infty^2] + \mathbb{E}[|\tilde{\xi}_{r(s)}^{i,N} - \tilde{\xi}_s^{i,N}|^2] \right] ds \\ &\leq C\delta t + C \int_0^t \mathbb{E}[\|\tilde{u}_{r(s)} - u_s^{S^N(\tilde{\xi})}\|_\infty^2] ds, \end{aligned} \quad (4.22)$$

where the fourth inequality above follows from Proposition 2.5, see (2.9). Consequently using (4.22) and inequality (4.10) of Lemma 4.4, (4.21) becomes

$$\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\tilde{\xi})}\|_2^2] \leq \frac{C}{N} \sum_{i=1}^N \mathbb{E}[|\tilde{V}_t^i - V_t^i|^2] \underbrace{\leq}_{(4.22)} C\delta t + C \int_0^t \mathbb{E}[\|\tilde{u}_{r(s)} - u_s^{S^N(\tilde{\xi})}\|_\infty^2] \underbrace{\leq}_{(4.10)} C\delta t. \quad (4.23)$$

Finally, injecting (4.23) and (4.18) in (4.17) yields

$$\mathbb{E}[\|\tilde{u}_t - u_t^{S^N(\tilde{\xi})}\|_2^2] \leq C\delta t,$$

which ends the proof of Proposition 4.1.  $\square$

## 4.2 Algorithm description

In this section, we describe precisely the algorithm relying on the time-discretization (4.1) of the interacting particle system (3.3). Let  $v_0$  be the law density of  $Y_0$  where  $Y$  is the solution of (1.4). In the sequel, we will make use of the same notations as in previous section. In particular,  $0 = t_0 \leq \dots \leq t_k = k\delta t \leq \dots \leq t_n = T$  is a regular time grid with  $\delta t = T/n$ . We consider a real-valued function  $K : \mathbb{R}^d \rightarrow \mathbb{R}$  being a mollifier depending on some bandwidth parameter  $\varepsilon$ .

**Initialization** for  $k = 0$ .

1. Generate  $(\tilde{\xi}_{t_0}^{i,N})_{i=1,\dots,N}$  i.i.d.  $\sim v_0(x)dx$ ;
2. set  $G_0^i := 1, i = 1, \dots, N$ ;
3. set  $\tilde{u}_{t_0}(\cdot) := (K * v_0)(\cdot)$ ;

**Iterations** for  $k = 0, \dots, n-1$ .

- Independently for each particle  $\tilde{\xi}_{t_k}^{j,N}$  for  $j = 1, \dots, N$ ,

$$\tilde{\xi}_{t_{k+1}}^{j,N} = \tilde{\xi}_{t_k}^{j,N} + \Phi(t_k, \tilde{\xi}_{t_k}^{j,N}, \tilde{u}_{t_k}(\tilde{\xi}_{t_k}^{j,N}))\sqrt{\delta t}\epsilon_{k+1}^j + g(t_k, \tilde{\xi}_{t_k}^{j,N}, \tilde{u}_{t_k}(\tilde{\xi}_{t_k}^{j,N}))\delta t,$$

where  $(\epsilon_k^j)_{j=1,\dots,N,k=1,\dots,n}$  is a sequence of i.i.d centered and standard Gaussian variables;

- set for  $j = 1, \dots, N$ ,

$$G_{k+1}^j := G_k^j \times \exp\left(\Lambda(t_k, \tilde{\xi}_{t_k}^{j,N}, \tilde{u}_{t_k}(\tilde{\xi}_{t_k}^{j,N}))\delta t\right);$$

- set

$$\tilde{u}_{t_{k+1}}(\cdot) = \frac{1}{N} \sum_{j=1}^N G_{k+1}^j \times K(\cdot - \tilde{\xi}_{t_{k+1}}^{j,N}).$$

**Remark 4.5.** For a fixed  $k \in \{0, \dots, n-1\}$ , we observe that the simulation of the  $j$ -th particle  $\tilde{\xi}_{t_{k+1}}^{j,N}$  at time  $t_{k+1}$  involves the whole particle system through the evaluation of  $\tilde{u}_{t_k}(\tilde{\xi}_{t_k}^{j,N})$ , which implies a complexity of the algorithm of order  $nN^2$ .

## 5 Numerical results

### 5.1 Preliminary considerations

One motivating issue of this section is how the interacting particle system  $\xi := (\xi^{i,N,\varepsilon})$  defined in (3.3) with  $K = K^\varepsilon, K^\varepsilon(x) := \frac{1}{\varepsilon^d}\phi^d(\frac{x}{\varepsilon})$  for some mollifier  $\phi^d$ , can be used to approach the solution  $v$  of the PDE

$$\begin{cases} \partial_t v = \frac{1}{2} \sum_{i,j=1}^d \partial_{ij}^2 ((\Phi\Phi^t)_{i,j}(t, x, v)) - \operatorname{div}(g(t, x, v)v) + \Lambda(t, x, v)v \\ v(0, x) = v_0, \end{cases} \quad (5.24)$$

to which we can reasonably expect that (1.5) converges when  $K^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \delta$ .

Two significant parameters, i.e.  $\varepsilon \rightarrow 0, N \rightarrow +\infty$ , intervene. We expect to approximate  $v$  by  $u^{\varepsilon,N}$ , which is the solution of the linking equation (1.6), associated with the empirical measure  $m = S^N(\xi)$ . To this

purpose, we want to control empirically the Mean Integrated Squared Error (MISE) between the solution  $v$  of (5.24) and the particle approximation  $u^{\varepsilon, N}$ , i.e. for  $t \in [0, T]$ ,

$$\mathbb{E}[\|u_t^{\varepsilon, N} - v_t\|_2^2] \leq 2\mathbb{E}[\|u_t^{\varepsilon, N} - u_t^\varepsilon\|_2^2] + 2\|u_t^\varepsilon - v_t\|_2^2, \quad (5.25)$$

where  $u^\varepsilon = u^{m^0}$  with  $K = K^\varepsilon, m^0$  being the common law of processes  $Y^i, 1 \leq i \leq N$  in (3.1). Even though the second term in the r.h.s. of (5.25) does not explicitly involve the number of particles  $N$ , the first term crucially depends on both parameters  $\varepsilon, N$ . The behavior of the first term relies on the propagation of chaos. This phenomenon has been observed in Corollary 3.2, which is a consequence of Theorem 3.1, for a fixed  $\varepsilon > 0$ , when  $N \rightarrow +\infty$ . According to Theorem 3.1, the first error term on the r.h.s. of the above inequality can be bounded by  $\frac{C(\varepsilon)}{N}$ .

Concerning the second error term, no result is available but we expect that it converges to zero when  $\varepsilon \rightarrow 0$ . To control the MISE, it remains to determine a relation  $N \mapsto \varepsilon(N)$  such that

$$\varepsilon(N) \xrightarrow{N \rightarrow +\infty} 0 \quad \text{and} \quad \frac{C(\varepsilon(N))}{N} \xrightarrow{N \rightarrow +\infty} 0. \quad (5.26)$$

When the coefficients  $\Phi, g$  and the initial condition are smooth with  $\Phi$  non-degenerate and  $\Lambda \equiv 0$  (i.e. in the conservative case), Theorem 2.7 of [17] gives a description of such a relation.

In our empirical analysis, we have concentrated on a test case, for which we have an explicit solution.

We first illustrate the chaos propagation for fixed  $\varepsilon > 0$ , i.e. the result of Theorem 3.1. On the other hand, we give an empirical insight concerning the following:

- the asymptotic behavior of the second error term in inequality (5.25) for  $\varepsilon \rightarrow 0$ ;
- the tradeoff  $N \mapsto \varepsilon(N)$  verifying (5.26).

Moreover, the simulations reveal two behaviors regarding the chaos propagation intensity.

## 5.2 The target PDE

We describe now the test case. For a given triple  $(m, \mu, A) \in ]1, \infty[ \times \mathbb{R}^d \times \mathbb{R}^{d \times d}$  we consider the following nonlinear PDE of the form (5.24):

$$\begin{cases} \partial_t v &= \frac{1}{2} \sum_{i,j=1}^d \partial_{i,j}^2 (v(\Phi\Phi^t)_{i,j}(t, x, v)) - \operatorname{div}(vg(t, x, v)) + v\Lambda(t, x, v), \\ v(0, x) &= B_m(2, x)f_{\mu, A}(x) \quad \text{for all } x \in \mathbb{R}^d, \end{cases} \quad (5.27)$$

where the functions  $\Phi, g, \Lambda$  defined on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}$  are such that

$$\Phi(t, x, z) = f_{\mu, A}^{\frac{1-m}{2}}(x) z^{\frac{m-1}{2}} I_d, \quad (5.28)$$

$I_d$  denoting the identity matrix in  $\mathbb{R}^{d \times d}$ ,

$$g(t, x, z) = f_{\mu, A}^{1-m}(x) z^{m-1} \frac{A + A^t}{2} (x - \mu), \quad \text{and} \quad \Lambda(t, x, z) = f_{\mu, A}^{1-m}(x) z^{m-1} \operatorname{Tr} \left( \frac{A + A^t}{2} \right). \quad (5.29)$$

Here  $f_{\mu, A} : \mathbb{R}^d \rightarrow \mathbb{R}$  is given by

$$f_{\mu, A}(x) = C e^{-\frac{1}{2} \langle x - \mu, A(x - \mu) \rangle}, \quad \text{normalized by} \quad C = \left[ \int_{\mathbb{R}^d} B_m(2, x) e^{-\frac{1}{2} \langle x - \mu, A(x - \mu) \rangle} \right]^{-1} \quad (5.30)$$

and  $B_m$  is the  $d$ -dimensional Barenblatt-Pattle density associated to  $m > 1$ , i.e.

$$B_m(t, x) = \frac{1}{2} (D - \kappa t^{-2\beta} |x|)_+^{\frac{1}{m-1}} t^{-\alpha}, \quad (5.31)$$

with  $\alpha = \frac{d}{(m-1)d+2}$ ,  $\beta = \frac{\alpha}{d}$ ,  $\kappa = \frac{m-1}{m}\beta$  and  $D = [2\kappa^{-\frac{d}{2}} \frac{\pi^{\frac{d}{2}} \Gamma(\frac{m-1}{m})}{\Gamma(\frac{d}{2} + \frac{m-1}{m})}]^{\frac{2(1-m)}{2+d(m-1)}}$ .

In the specific case where  $A$  is the zero matrix of  $\mathbb{R}^{d \times d}$ , then  $f_{\mu, A} \equiv 1$ ;  $g \equiv 0$  and  $\Lambda \equiv 0$ . Hence, we recover the conservative porous media equation, whose explicit solution is

$$v(t, x) = B_m(t + 2, x), \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d,$$

see [1]. For general values of  $A \in \mathbb{R}^{d \times d}$ , extended calculations produce the following explicit solution

$$v(0, \cdot) = v_0(\cdot) \quad \text{and} \quad v(t, x) = B_m(t + 2, x) f_{\mu, A}(x), \quad \text{for all } (t, x) \in [0, T] \times \mathbb{R}^d, \quad (5.32)$$

of (5.27), which is non conservative.

### 5.3 Details of the implementation

Once fixed the number  $N$  of particles, we have run  $M = 100$  i.i.d. particle systems producing  $(u^{\varepsilon, N, i})_{i=1, \dots, M}$ , which are  $M$  i.i.d. realizations of  $u^{\varepsilon, N}$  introduced just after (5.24). The MISE is then approximated by the Monte Carlo approximation

$$\mathbb{E}[\|u_t^{\varepsilon, N} - v_t\|_2^2] \approx \frac{1}{MQ} \sum_{i=1}^M \sum_{j=1}^Q |u_t^{\varepsilon, N, i}(X^j) - v_t(X^j)|^2 v^{-1}(0, X^j), \quad \text{for all } t \in [0, T], \quad (5.33)$$

where  $(X^j)_{j=1, \dots, Q=1000}$  are i.i.d  $\mathbb{R}^d$ -valued random variables with common density  $v(0, \cdot)$ . In our simulation, we have chosen  $T = 1$ ,  $m = 3/2$ ,  $\mu = 0$  and  $A = \frac{2}{3}I_d$ .  $K^\varepsilon = \frac{1}{\varepsilon^d} \phi^d(\frac{\cdot}{\varepsilon})$  with  $\phi^d$  being the standard and centered Gaussian density.

In this subsection, we fix the dimension to  $d = 5$ . We have run a discretized version of the interacting particle system with Euler scheme mesh  $kT/n$  with  $n = 10$ . Notice that this discretization error is neglected in the present analysis.

Our simulations show that the approximation error presents two types of behavior depending on the number  $N$  of particles with respect to the regularization parameter  $\varepsilon$ .

1. For large values of  $N$ , we visualize a *chaos propagation behavior* for which the error estimates are similar to the ones provided by the density estimation theory [26] corresponding to the classical framework of independent samples.
2. For small values of  $N$  appears a *transient behavior* for which the bias and variance errors cannot be easily described.

Observe that the Mean Integrated Squared Error  $\text{MISE}_t(\varepsilon, N) := \mathbb{E}[\|u_t^{\varepsilon, N} - v_t\|_2^2]$  can be decomposed as the sum of the variance  $V_t(\varepsilon, N)$  and squared bias  $B_t^2(\varepsilon, N)$  as follows:

$$\begin{aligned} \text{MISE}_t(\varepsilon, N) &= V_t(\varepsilon, N) + B_t^2(\varepsilon, N) \\ &= \mathbb{E} \left[ \|u_t^{\varepsilon, N} - \mathbb{E}[u_t^{\varepsilon, N}]\|_2^2 \right] + \|\mathbb{E}[u_t^{\varepsilon, N}] - v_t\|_2^2. \end{aligned} \quad (5.34)$$

For  $N$  large enough, according to Corollary 3.2, one expects that the propagation of chaos holds. Then the particle system  $(\xi^{i, N})_{i=1, \dots, N}$  (solution of (3.3)) is close to an i.i.d. system with common law  $m^0$ . We observe

that, in the specific case where the weighting function  $\Lambda$  does not depend on the density  $u$ , for  $t \in [0, T]$ , we have

$$\begin{aligned}\mathbb{E}[u_t^{\varepsilon, N}] &= \frac{1}{N} \mathbb{E} \left[ \sum_{j=1}^N K^\varepsilon(\cdot - \xi_t^{j, N}) \exp \left\{ \int_0^t \Lambda(r(s), \xi_{r(s)}^{j, N}) ds \right\} \right], \\ &= \mathbb{E} [K^\varepsilon(\cdot - Y_t^1) V_t(Y^1)] \\ &= u_t^\varepsilon.\end{aligned}\tag{5.35}$$

Therefore, under the chaos propagation behavior, the approximations below hold for the variance and the squared bias:

$$V_t(\varepsilon, N) \approx \mathbb{E} \left[ \|u_t^{\varepsilon, N} - u_t^\varepsilon\|_2^2 \right] \quad \text{and} \quad B_t^2(\varepsilon, N) \approx \|u_t^\varepsilon - v_t\|_2^2.\tag{5.36}$$

We recall that the relation  $u^\varepsilon = K^\varepsilon * v^\varepsilon$  comes from Theorem 6.1 of [20], where  $v^\varepsilon$  is solution of (1.5) with  $K = K^\varepsilon$ .

On Figure 1, we have reported the estimated variance error  $V_t(\varepsilon, N)$  as a function of the particle number  $N$ , (on the left graph) and as a function of the regularization parameter  $\varepsilon$ , (on the right graph), for  $t = T = 1$  and  $d = 5$ . We have used for this a similar Monte Carlo approximation as (5.33).

That figure shows that, when the number of particles is *large enough*, the variance error behaves precisely as in the classical case of density estimation encountered in [26], i.e., vanishing at a rate  $\frac{1}{N\varepsilon^d}$ , see relation (4.10), Chapter 4., Section 4.3.1. This is in particular illustrated by the log-log graphs, showing almost linear curve, when  $N$  is sufficiently large. In particular we observe the following.

- On the left graph,  $\log(V_t(\varepsilon, N)) \approx a - \alpha \log N$  with slope  $\alpha = 1$ ;
- On the right graph,  $\log V_t(\varepsilon, N) \approx b - \beta \log \varepsilon$  with slope  $\beta = 5 = d$ .

It seems that the threshold  $N$  after which appears the linear behavior (compatible with the propagation of chaos situation corresponding to asymptotic-i.i.d. particles) decreases when  $\varepsilon$  grows. In other words, when  $\varepsilon$  is large, less particles  $N$  are needed to give evidence to the chaotic behavior.

This phenomenon can be probably explained by analyzing the particle system dynamics. Indeed, at each time step, the interaction between the particles is due to the empirical estimation of  $u^\varepsilon = K^\varepsilon * v^\varepsilon$  based on the particle system. Intuitively, the more accurate the approximation  $u^{\varepsilon, N}$  of  $u^\varepsilon$  is, the less strong the interaction between particles will be. In the limiting case when  $u^{\varepsilon, N} = u^\varepsilon$ , the interaction disappears.

Now observe that at time step 0, the particle system  $(\xi_0^{i, N})$  is i.i.d. according to  $v_0(\cdot)$ , so that the estimation of  $(K^\varepsilon * v^\varepsilon)(0, \cdot)$  provided by (4.1) reduces to the classical density estimation approach, see [26] as mentioned above. In that classical framework, we emphasize that, for larger values of  $\varepsilon$ , the number of particles, needed to achieve a given density estimation accuracy, is smaller. Hence, one can imagine that for larger values of  $\varepsilon$  less particles will be needed to obtain a quasi-i.i.d particle system at time step 1,  $(\xi_{t_1}^{i, N})$ . We can then reasonably presume that this initial error propagates along the time steps.

On Figure 2, we have reported the estimated squared bias error,  $B_t^2(\varepsilon, N)$ , as a function of the regularization parameter,  $\varepsilon$ , for different values of the particle number  $N$ , for  $t = T = 1$  and  $d = 5$ .

One can observe that, similarly to the classical i.i.d. case, (see relation (4.9) in Chapter 4., Section 4.3.1 in [26]), for  $N$  large enough, the bias error does not depend on  $N$  and can be approximated by  $a\varepsilon^4$ , for some constant  $a > 0$ . This is in fact coherent with the bias approximation (5.36), developed in the specific case where the weighting function  $\Lambda$  does not depend on the density. Assuming the validity of approximation (5.36) and of the previous empirical observation implies that one can bound the error between the

solution,  $v^\varepsilon$ , of the regularized PDE of the form (1.5) (with  $K = K^\varepsilon$ ) associated to (5.27), and the solution,  $v$ , of the limit (non regularized) PDE (5.27) as follows

$$\begin{aligned} \|v_t^\varepsilon - v_t\|_2^2 &\leq 2\|v_t^\varepsilon - u_t^\varepsilon\|_2^2 + 2\|u_t^\varepsilon - v_t\|_2^2 \\ &\leq 2\|v_t^\varepsilon - K^\varepsilon * v_t^\varepsilon\|_2^2 + 2\|u_t^\varepsilon - v_t\|_2^2 \\ &\leq 2(a' + a)\varepsilon^4. \end{aligned} \tag{5.37}$$

Indeed, at least, the first term in the second line can be easily bounded, supposing that  $v_t^\varepsilon$  has a bounded second derivative. This constitutes an empirical proof of the fact that  $v^\varepsilon$  converges to  $v$ .

As observed in the variance error graphs, the threshold  $N$ , above which the propagation of chaos behavior is observed decreases with  $\varepsilon$ . Indeed, for  $\varepsilon > 0.6$  we observe a chaotic behavior of the bias error, starting from  $N \geq 500$ , whereas for  $\varepsilon \in [0.4, 0.6]$ , this chaotic behavior appears only for  $N \geq 5000$ .

For small values of  $\varepsilon \leq 0.6$ , the bias highly depends on  $N$  for any  $N \leq 10^4$ ; moreover that dependence becomes less relevant when  $N$  increases. This is probably due to the combination of two effects: the lack of chaos propagation phenomenon and the fact that the coefficient  $\Lambda$  depends on  $u$ , so that (5.35) does not hold in that context.

Taking into account both the bias and the variance error in the MISE (5.34), the choice of  $\varepsilon$  has to be carefully optimized w.r.t. the number of particles:  $\varepsilon$  going to zero together with  $N$  going to infinity at a judicious relative rate seem to ensure the convergence of the estimated MISE to zero. This kind of tradeoff is standard in density estimation theory and was already investigated theoretically in the context of forward interacting particle systems related to conservative regularized nonlinear PDE in [17]. Extending this type of theoretical analysis to our non conservative framework is beyond the scope of the present paper.

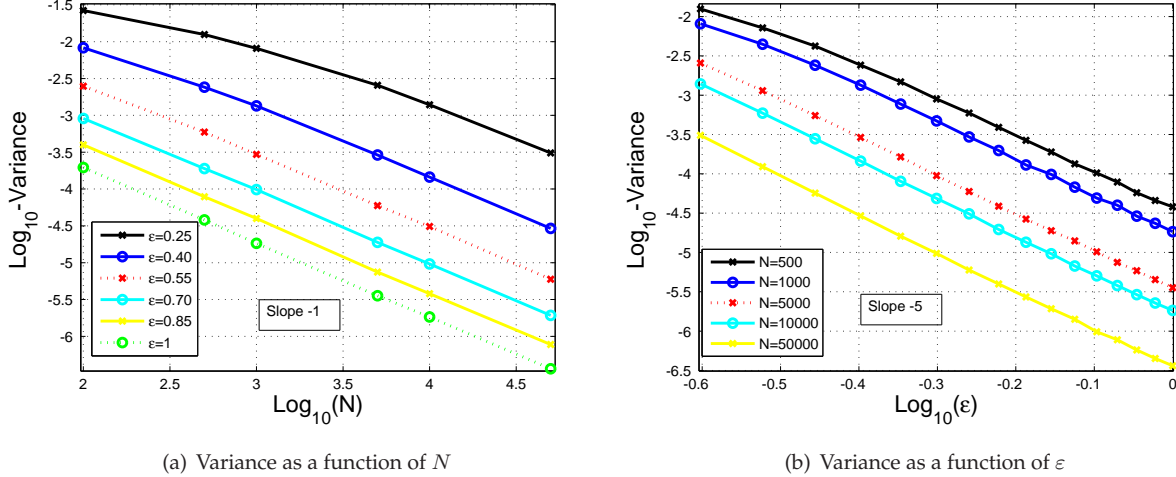


Figure 1: Variance error as a function of the number of particles,  $N$ , and the mollifier window width,  $\varepsilon$ , for dimension  $d = 5$  at the final time step  $T = 1$ .

### 5.3.1 Time discretization error

In this subsection, we are interested in analysing via numerical simulations the time discretization error w.r.t. to  $\delta t = T/n$ . As announced in Remark 4.3, we suspect that the rate in (4.4) is not optimal and that the



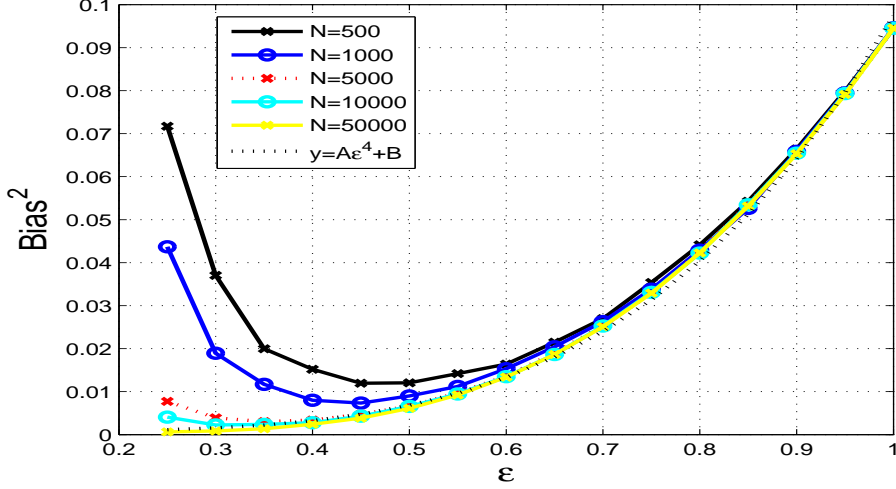


Figure 2: Bias error as a function of the mollifier window width,  $\varepsilon$ , for dimension  $d = 5$  at the final time step  $T = 1$ .

MISE error induced by the time discretization is of order  $1/n^2$  instead of  $1/n$ .

Let  $\tilde{u}_T^{\varepsilon, N, n}$  denote the particle approximation obtained by scheme (4.1) with a number of particles,  $N$ , a regularization parameter,  $\varepsilon$ , and a number of time steps,  $n$ . In order to focus on the time discretization error apart from the particle approximation and the regularization error (related to  $N$  and  $\varepsilon$ ), we have considered errors of the type  $\mathbb{E}[\|\tilde{u}_T^{\varepsilon, N, n} - \tilde{u}_T^{\varepsilon, N, n_0}\|_2^2]$  for different numbers of time steps  $n < n_0$  where  $n_0$  is supposed to be a large number of time steps. More precisely, we have decomposed this error into a variance and a squared bias term as

$$\begin{aligned} \mathbb{E}[\|\tilde{u}_T^{\varepsilon, N, n} - \tilde{u}_T^{\varepsilon, N, n_0}\|_2^2] &= \underbrace{\mathbb{E}[\|\tilde{u}_T^{\varepsilon, N, n} - \mathbb{E}[\tilde{u}_T^{\varepsilon, N, n}]\|_2^2] + \mathbb{E}[\|\tilde{u}_T^{\varepsilon, N, n_0} - \mathbb{E}[\tilde{u}_T^{\varepsilon, N, n_0}]\|_2^2]}_{\text{Variance}} \\ &+ \underbrace{\|\mathbb{E}[\tilde{u}_T^{\varepsilon, N, n}] - \mathbb{E}[\tilde{u}_T^{\varepsilon, N, n_0}]\|_2^2}_{\text{Bias}^2}, \end{aligned} \quad (5.38)$$

if  $u_T^{\varepsilon, N, n}$  and  $u_T^{\varepsilon, N, n_0}$  are independent.

On Figure 3, we have reported the Monte Carlo estimation (according to (5.33), with  $Q = 1000$  runs) of the above variance and squared bias terms in a log-log scale in order to diagnose the expected rate of convergence  $1/n^2$  via a straight line with slope  $-2$ . All the parameters are similar to the simulations performed in previous subsection excepted for the dimension  $d = 1$ ,  $N = 5000$  and  $n_0$  is set to 1000 time steps. One can observe that the variance term (in dashed lines) seems not to depend on the number of time steps  $n$  whereas the squared bias term decreases as expected at a rate close to  $1/n^2$ .

## 6 Appendix

In this appendix, we present the proof of Lemma 4.4. We first proceed with the proof of some intermediary inequalities.

**Lemma 6.1.** *We suppose Assumption 1. Let  $N \in \mathbb{N}^*$ . Let  $(\xi^{i, N})_{i=1, \dots, N}$  be (a solution of) the interacting particle*