McKean Feynman-Kac probabilistic representations of non-linear partial differential equations

Lucas Izydorczyk, Nadia Oudjane, Francesco Russo

McKean Feynman-Kac probabilistic representations of non-linear partial differential equations

Lucas Izydorczyk [*], Nadia Oudjane [†] and Francesco Russo [‡]

December 2019

Abstract

This paper presents a partial state of the art about the topic of representation of generalized Fokker-Planck Partial Differential Equations (PDEs) by solutions of McKean Feynman-Kac Equations (MFKEs) that generalize the notion of McKean Stochastic Differential Equations (MSDEs). While MSDEs can be related to non-linear Fokker-Planck PDEs, MFKEs can be related to non-conservative non-linear PDEs. Motivations come from modeling issues but also from numerical approximation issues in computing the solution of a PDE, arising for instance in the context of stochastic control. MFKEs also appear naturally in representing final value problems related to backward Fokker-Planck equations.

Key words and phrases: backward diffusion; McKean stochastic differential equation; probabilistic representation of PDEs; time reversed diffusion; HJB equation; Feynman-Kac measures.

2010 AMS-classification: 60H10; 60H30; 60J60; 65C05; 65C35; 35K58.

1 Introduction and motivations

1.1 General considerations

The idea of the present article is to focus on models which have a double macroscopic-microscopic face in the form of *perturbation* of a so called Fokker-Planck type equation that we call *generalized* Fokker-Planck equation. Our ambition is driven by two main reasons.

1. A *modeling* reason: the idea is to observe both from a macroscopic-microscopic point of view phenomena arising from physics, biology, chemistry or complex systems.

2. A *numerical simulation* reason: to provide Monte-Carlo suitable algorithms to approach PDEs.

[*]ENSTA-Paris, Institut Polytechnique de Paris. Unité de Mathématiques Appliquées (UMA). E-mail: `lucas.izydorczyk@ensta-paris.fr`

[†]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`

[‡]ENSTA Paris, Institut Polytechnique de Paris. Unité de Mathématiques Appliquées (UMA). E-mail:`francesco.russo@ensta-paris.fr`.

The *target macroscopic* Fokker-Planck equation is

$$
\begin{cases}
\partial_t u &= \frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij}^2 \left((\sigma\sigma^\top)_{i,j}(t,x,u)u \right) - div\left(b(t,x,u,\nabla u)u \right) \\
&\quad + \Lambda(t,x,u,\nabla u)u\,, \quad \text{for } t \in]0,T]\,, \\
u(0,\cdot) &= \mathbf{u_0},
\end{cases}
\tag{1.1}
$$

where $\mathbf{u_0}$ is a Borel probability measure $\sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to M_{d,p}(\mathbb{R})$, $b : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, $\Lambda : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and ∇ denotes the gradient operator. The initial condition in (1.1) means that for every continuous bounded real function φ we have $\int \varphi(x)u(t,x)dx \to \int \varphi(x)\mathbf{u_0}(dx)$ when $t \to 0$. When $\mathbf{u_0}$ admits a density, we denote it by u_0. The unknown function $u :]0,T] \times \mathbb{R}^d \to \mathbb{R}$ is supposed to run in $L^1(\mathbb{R}^d)$ considered as a subset of the space of finite Radon measures $\mathcal{M}(\mathbb{R}^d)$. The idea consists in finding a probabilistic representation via the solution of a *Stochastic Differential Equation (SDE)* whose coefficients do not depend only on time and the position of the *particle* but also on its probability law. The *target microscopic equation* we have in mind is

$$
\begin{cases}
Y_t = Y_0 + \int_0^t \sigma\left(s, Y_s, u(s, Y_s)\right)dW_s + \int_0^t b\left(s, Y_s, u(s, Y_s)\right)ds \\
Y_0 \sim \mathbf{u_0} \\
\int \varphi(x)u(t,x)dx = \mathbb{E}\left[\varphi(Y_t) \, \exp\left\{ \int_0^t \Lambda\left(s, Y_s, u(s, Y_s), \nabla u(s, Y_s)\right)ds \right\} \right], \quad \text{for } t \in]0,T]\,,
\end{cases}
\tag{1.2}
$$

for any continuous bounded real valued test function φ. Sometimes we denominate the third line equation of (1.2) the *linking equation*. When $\Lambda = 0$, in equation (1.2), the linking equation simply says that $u(t,\cdot)$ coincides with the density of the marginal distribution $\mathcal{L}(Y_t)$. In this specific case, equation (1.2) reduces to a *McKean Stochastic Differential Equation (MSDE)*, which is in general an SDE whose coefficients, at time t, depend, not only on (t, Y_t), but also on the marginal law $\mathcal{L}(Y_t)$. With more general functions Λ, the role of the linking equation is more intricate since the whole history of the process $(Y_s)_{0 \le s \le t}$ is involved. This fairly general type of equations will be called *McKean Feynman-Kac Equation (MFKE)* to emphasize the fact that $u(t,x)dx$ now corresponds to a non-conservative Feynman-Kac measure.

An interesting feature of MSDEs (so when $\Lambda = 0$) is that the law of the process Y can often be characterized as the limiting empirical distribution of a large number of interacting particles, whose dynamics are described by a coupled system of classical SDEs. When the number of particles grows to infinity, the particles behave closely to a system of independent copies of Y. This constitutes the so called *propagation of chaos* phenomenon, already observed in the literature when the drift and diffusion coefficients are Lipschitz dependent on the solution marginal law, with respect to the Wasserstein metric, see e.g. [41, 51, 52, 63, 54]. Propagation of chaos is a common phenomenon arising in many physical contexts, see for instance [1] concerning Nelson stochastic mechanics.

When $\Lambda = 0$, equation (1.1) is a *non-linear* Fokker-Planck equation, it is conservative and it is known that, under mild assumptions, it describes the dynamics of the marginal probability densities, $u(t,\cdot)$, of the process Y. This correspondence between PDE (1.1) with MSDE (1.2) and interacting particles have extensive interesting applications. In physics, biology or economics, it is a way to relate a microscopic model involving interacting particles to a macroscopic model involving the dynamics of the underlying density. Numerically, this correspondence motivates Monte-Carlo approximation schemes for PDEs. In particular, [19] has contributed to develop stochastic particle methods in the spirit of McKean to provide original numerical schemes approaching a PDE related to Burgers equation providing also the rate of convergence.

Below we list some situations of particular interest where such correspondence holds.

1.2 Some motivating examples

Burgers equation

We fix $d = p = 1$ and let $\nu > 0$ and u_0 be a probability density on $\mathbb{R}$. We consider two equivalent specific cases of (1.1). The first $\sigma \equiv \nu$, $b \equiv 0$, $\Lambda(t,x,u,z) = z$. The second $\sigma \equiv \nu$, $b(t,x,u) = \frac{u}{2}, \Lambda = 0$. Both instantiations correspond to the the *viscid Burgers equation* in dimension $d = 1$, given by

$$\begin{cases} \partial_t u = \frac{\nu^2}{2}\partial_{xx}u - u\partial_x u, & (t,x) \in [0,T] \times \mathbb{R}, \\ u(0,\cdot) = \mathbf{u_0} . \end{cases} \tag{1.3}$$

Generalized Burgers-Huxley equation

We fix $d = p = 1$ and let $\nu > 0$ and u_0 be a probability density on $\mathbb{R}$. We consider the particular cases of (1.1) where $\sigma \equiv \nu$, $b(t,x,u) = \alpha\frac{u^n}{n+1}$, $\Lambda(t,x,u) = \beta(1-u^n)(u^n - \gamma)$, with fixed reals α, β, γ and a non-negative integer n. This instantiation corresponds to a natural extension of Burgers equation called *Generalized Burgers-Huxley equation* or *Burgers-Fisher equation* which is of great importance to represent non-linear phenomena in various fields such as biology [2, 55], physiology [42] and physics [68]. These equations have the particular interest to describe the interaction between the reaction mechanisms, convection effect, and diffusion transport. Those are non-linear and non-conservative PDEs of the form

$$\begin{cases} \partial_t u = \frac{\nu}{2}\partial_{xx}u - \alpha u^n\partial_x u + \beta u(1-u^n)(u^n - \gamma), & (t,x) \in [0,T] \times \mathbb{R}, \\ u(0,\cdot) = \mathbf{u_0}. \end{cases} \tag{1.4}$$

Fokker-Planck equation with terminal condition

The present example does not properly integrate the framework of (1.1). In terms of application, we are interested by inverse problems that can be formulated by a PDE with terminal condition

$$\begin{cases} \partial_t u & = & \frac{1}{2}\sum_{i,j=1}^{d} \partial_{ij}^2\left((\sigma\sigma^t)_{i,j}(t,x)u\right) - div\left(b(t,x)u\right) \\ & & +\Lambda(t,x)u\,, \quad \text{for } t \in]0,T[\,, \\ u(T,\cdot) & = & \mathbf{u_T}, \end{cases} \tag{1.5}$$

where $\mathbf{u_T}$ is a prescribed probability measure. Solving that equation by analytical means constitutes a delicate task. A probabilistic representation may help for studying well-posedness or providing numerical schemes.

Backward simulation of diffusions is a subject of active research in various domains of physical sciences and engineering, as heat conduction [12], material science [60] or hydrology [3]. In particular, *hydraulic inversion* is interested in inverting a diffusion phenomenon representing the concentration of a pollutant to identify the pollution source location when the final concentration profile is observed. The problem is in general ill-posed because either the solution is not unique or the solution is not stable. For this type of problem, the existence is ensured by the fact that the observed contaminant has necessarily originated from some place at a given time (as soon as the model is correct). To correct the lack of well-posedness two regularization procedures have been proposed in the literature: the first one relies on the notion of quasi-solution, introduced by Tikhonov [65], the second one on the method of quasi-reversibility, introduced by Lattes and Lions, [44]. Besides well-posedness, a second crucial issue consists in providing a numerical approximating scheme to the backward diffusion equation. A probabilistic representation of (1.5) via the time-reversal of a diffusion could show those issues under a new light.

The stochastic Fokker-Planck with multiplicative noise

We fix $p = d$, $\sigma(t, x, u) = \Phi(u)Id_d$, where $\Phi : \mathbb{R} \rightarrow \mathbb{R}$ and $b = \Lambda \equiv 0$. Typical examples are the case of classical porous media type equation (resp. fast diffusion equation), when $\Phi(u) = u^q$, $1 \leq q$ (resp. $0 < q < 1$). The (singular) case $\Phi(u) = \gamma H(u - e_c)$, H being the Heavisise function and e_c a given threshold in $\mathbb{R}$, appears in the science of complex systems, more precisely in the so called *self-organized criticality*, see e.g. [4, 23, 5].

$$\begin{cases} \partial_t u & = & \frac{\gamma}{2}\Delta(H(u - e_c)u) \\ u(0, \cdot) & = & \mathbf{u_0}. \end{cases} \tag{1.6}$$

The phenomenon of *self-organized criticality* often is described in two scale phases: a fast dynamics (of *avalanch type*) described by the PDE (1.6) and a slower motion of *sand storming* modeled by the addition of a supplementary stochastic noise $\Lambda(t, x; \omega)$. In that case the *target macroscopic* equation is

$$\begin{cases} \partial_t u & = & \frac{\gamma}{2}\Delta(H(u - e_c)u) + \Lambda(t, x; \omega)u \\ u(0, \cdot) & = & \mathbf{u_0}, \end{cases} \tag{1.7}$$

where $\Lambda(t, x; \omega)$ is a quenched realization of a space-time coloured (ideally white) noise. The SPDE will be represented by a MSDE in random environment, see Section 6.

1.3 Structure of the paper

In the rest of the paper, to simplify notations, most of the results are stated in the one-dimensional setting. The generalization to the multi-dimensional case is straightforward.

The paper is organized as follows. Next section presents a brief review of basic situations where Fokker-Planck equations can be represented by MSDEs which in turn can be represented by interacting particles systems. Section 3, considers the case of generalized Fokker-Planck equations in the sense of (1.1) with a non-zero term Λ allowing to take into account non-conservative PDEs including a large class of semi-linear PDEs. Section 5 highlights the correspondence between MFKEs and MSDEs with jumps which paves the way to a great variety of numerical approximations schemes for non-linear PDEs. Section 4 is devoted to a particular inverse problem which consists in modeling backwardly in time the evolution of a Fokker-Planck equation with a given terminal condition. This problem can be related to a time-reversed SDE which in turn can be represented by a MSDE. In Section 6 we analyze the well-posedness of generalized Fokker-Planck equation where the term Λ in (1.1) may involve an exogenous noise resulting in a Stochastic non-linear PDE. Finally, in Section 7, we consider a stochastic control problem for which the associated Hamilton-Jacobi-Bellman equation can be represented by a MFKE.

2 McKean representations of non linear Fokker-Planck equations

In this section, we recall some standard situations where a Fokker-Planck PDE can be represented by an SDE which in turn can be approached by an interacting particles system.

2.1 Probabilistic representation of linear Fokker-Planck equations

Suppose there exists a solution $(Y_t)_{t\in[0,T]}$ (in law) to the SDE

$$\begin{cases} Y_t = Y_0 + \int_0^t \sigma(s, Y_s)dW_s + \int_0^t b(s, Y_s)ds, t \in [0, T], \\ Y_0 \sim \mathbf{u_0}, \end{cases} \tag{2.1}$$

where W is a real valued Brownian motion on $[0, T]$ and $\mathbf{u_0}$ is a probability measure on $\mathbb{R}$. A direct application of Itô formula shows that the marginal probability laws $(\mu(t, \cdot) := \mathcal{L}(Y_t))_{t\in[0,T]}$ generate a distributional solution of the linear Fokker-Planck PDE

$$\begin{cases} \partial_t \mu = \frac{1}{2}\partial_{xx}^2(\sigma^2(t,x)\mu) - \partial_x(b(t,x)\mu) \\ \mu(0, dx) = \mathbf{u_0}(dx). \end{cases} \tag{2.2}$$

This naturally suggests a Monte Carlo algorithm to approximate the above linear PDE, consisting in simulating N i.i.d. particles $(\xi^i)_{i=1,\cdots N}$ with N i.i.d. Brownian motions $(W^i)_{i=1,\cdots N}$ i.e.

$$\begin{cases} \xi_t^i = \xi_0^i + \int_0^t \sigma(s, \xi_s^i)dW_s^i + \int_0^t b(s, \xi_s^i)ds \\ \xi_0^i \quad \text{i.i.d.} \sim \mathbf{u_0} \\ \mu_t^N = \frac{1}{N}\sum_{j=1}^N \delta_{\xi_t^j}. \end{cases} \tag{2.3}$$

Then the law of large numbers provides the convergence of the empirical approximation $\mu_t^N \xrightarrow[N\to\infty]{} \mu(t, \cdot)$, the solution of the Fokker-Planck equation (2.2).

2.2 McKean probabilistic representation of non-linear Fokker-Planck equation

We consider the non-linear SDE in the sense of McKean (MSDE)

$$\begin{cases} Y_t = Y_0 + \int_0^t \sigma\Big(s, Y_s, (K*\mu)(s, Y_s)\Big)dW_s + \int_0^t b\Big(s, Y_s, (K*\mu)(s, Y_s)\Big)ds \\ Y_0 \sim \mathbf{u_0} \\ \mu(t, \cdot) \text{ is the probability law of } Y_t, t \in [0, T], \end{cases} \tag{2.4}$$

whose solution is a couple (Y, μ). Here σ, b are Lipschitz, $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ denotes a Lipschitz continuous convolution kernel such that $(K * \mu)(t, y) := \int K(y, z)\mu(t, dz)$ for any $y \in \mathbb{R}$. We emphasize that this type of regularized dependence of the drift and diffusion coefficients on μ is essentially different (and in general easier to handle) from a pointwise dependence where the coefficients b or σ may depend on the value of the marginal density at the current particle position $\frac{d\mu}{dx}(s, Y_s)$. This regularized or non-local dependence on the time-marginals $\mu(t, \cdot)$ is a particular case of the framework when the diffusion and drift coefficients are Lipschitz with respect to $\mu(t, \cdot)$ according to the the Wasserstein metric.

Again, by Itô formula, given a solution (Y, μ) of (2.4), μ solves the non-local non-linear PDE

$$\begin{cases} \partial_t\mu = \frac{1}{2}\partial_{xx}^2\Big(\sigma^2(t, x, K*\mu)\mu\Big) - \partial_x\Big(b(t, x, K*\mu)\mu\Big) \\ \mu(0, dx) = \mathbf{u_0}(dx), \end{cases} \tag{2.5}$$

in the sense of distributions. In this setting, the well-posedness of (2.4) relies on a fixed point argument in the space of trajectories under the Wasserstein metric, see e.g. [63], at least in the case when the diffusion term does not depend on the law. We will denominate this situation as the *traditional* setting.

Deriving a Monte-Carlo approximation scheme from this probabilistic representation already becomes more tricky since it can no more rely on independent particles but should involve an interacting particles system as initially proposed in [41, 63]. Consider N interacting particles $(\xi^{i,N})_{i=1,\cdots N}$ with N i.i.d. Brownian motions (W^i), i.e.

$$
\begin{cases}
\xi_t^{i,N} = \xi_0^{i,N} + \int_0^t \sigma\Big(s, \xi_s^{i,N}, (K * \mu_s^N)(\xi_s^{i,N})\Big) dW_s^i + \int_0^t b\Big(s, \xi_s^{i,N}, (K * \mu_s^N)(\xi_s^{i,N})\Big) ds \\
\xi_0^{i,N} \text{ i.i.d. } \sim \mathbf{u_0} \\
\mu_t^N = \dfrac{1}{N}\sum_{j=1}^N \delta_{\xi_t^{j,N}},
\end{cases}
\tag{2.6}
$$

with $(K * \mu_t^N)(y) = \frac{1}{N}\sum_{j=1}^N K(y, \xi_t^{j,N})$. The above system defines a so-called *weakly interacting* particles system, as pointed out in [56]. This terminology underlines the fact that any particle interacts with the rest of the population with a vanishing impact of order $1/N$. In this setting, at least when the diffusion coefficient does not depend ton the law, [63] proves the so called *chaos propagation* which means that $(\xi_t^{i,N})_{i=1,\cdots N}$ asymptotically behaves as an i.i.d. sample according to $\mu(t, \cdot)$ as the number of particles N grows to infinity, where μ is the solution of the regularized non-linear PDE (2.5). This in particular implies the convergence of the empirical measures $\mu_t^N \xrightarrow[N\to\infty]{} \mu(t, \cdot)$ with the rate $C/\sqrt{N}$ inherited from the law of large numbers.

As already announced, the case where the coefficients depend pointwisely on the density law $u(t, \cdot)$ of $\mu(t, \cdot)$, $t > 0$, is far more singular. Indeed the dependence of the coefficients on the law of Y is no more continuous with respect to the Wasserstein metric. In this context, well-posedness results rely generally on analytical methods. One important contribution in this direction is reported in [39], where strong existence and pathwise uniqueness are established when the diffusion coefficient σ and the drift b exhibit pointwise dependence on u but are assumed to satisfy strong smoothness assumptions together with the initial condition. In this case, the solution u is a classical solution of the PDE

$$
\begin{cases}
\partial_t u = \dfrac{1}{2}\partial_{xx}^2\Big(\sigma^2(t, x, u(t, x))u\Big) - \partial_x\Big(b(t, x, u(t, x))u\Big) \\
u(0, x) = \mathbf{u_0}(dx),
\end{cases}
\tag{2.7}
$$

which is formally derived from (2.5) setting $K(x, y) = \delta_0(x - y)$. Let us fix K^ε being a mollifier (depending on a window-width parameter ε), such that $K^\varepsilon(x, y) = \dfrac{1}{\varepsilon^d}\phi(\dfrac{x-y}{\varepsilon}) \xrightarrow[\varepsilon\to 0]{} \delta_0(x - dy)$. As in (2.6), we consider the N interacting particles $(\xi^{i,N})_{i=1,\cdots N}$ solving

$$
\begin{cases}
\xi_t^{i,N} = \xi_0^{i,N} + \int_0^t \sigma\Big(s, \xi_s^{i,N}, u_s^{N,\varepsilon}(\xi_s^{i,N})\Big) dW_s^i + \int_0^t b\Big(s, \xi_s^{i,N}, u_s^{N,\varepsilon}(\xi_s^{i,N})\Big) ds \\
\xi_0^{i,N} \text{ i.i.d. } \sim \mathbf{u_0} \\
u_t^{N,\varepsilon} = \dfrac{1}{N}\sum_{j=1}^N K^\varepsilon(\cdot, \xi_t^{j,N}).
\end{cases}
\tag{2.8}
$$

Under the smooth assumptions on b, σ, u_0 mentioned before and non-degeneracy of σ, [39] proved the convergence of the regularized particle approximation $u_t^{N,\varepsilon}$ to the solution u of the pointwise non-linear PDE (2.7) as soon as $\varepsilon(N) \xrightarrow[N\to\infty]{} 0$ slowly enough. According to [56], the system (2.8) defines a so-called *moderately interacting* particle system with $u_t^{N,\varepsilon}(x) = \frac{1}{N\varepsilon^d}\sum_{j=1}^N \phi(\frac{x-\xi_t^{j,N}}{\varepsilon})$. Indeed as the window width of the kernel, ε, goes to zero, the number of particles that significantly impact a single one is of order $N\varepsilon^d$ with a strength of interaction of order $\frac{1}{N\varepsilon^d}$. In contrast, when ε is fixed, we recover the weakly interacting situation in which case the strength of interaction of each particle is of order $\frac{1}{N}$ which is smaller than $\frac{1}{N\varepsilon^d}$.

In this case of moderate interaction, the propagation of chaos occurs with a slower rate than $C/\sqrt{N}$ and depends exponentially on the space dimension. [39] constitutes an extension of the weak propagation of chaos of moderately interacting particles proved in [56] for the limited case of identity diffusion matrix.

The peculiar case where the drift vanishes and the diffusion coefficient $\sigma(u(t, Y_t))$ has a pointwise dependence on the law density $u(t, \cdot)$ of Y_t has been more particularly studied in [15] for classical porous media type equations and [17, 8, 14, 13, 7] who obtain well-posedness results for measurable and possibly singular functions σ. In that case the solution u of the associated PDE (1.1), is understood in the sense of distributions.

3 McKean Feynman-Kac representations for non-conservative and non-linear PDEs

The idea of generalizing MSDEs to MFKEs (1.2) was originally introduced in the sequence of papers [46, 45, 47], with an earlier contribution in [10], where $\Lambda(t, x, u, \nabla u) = \xi_t(x)$, ξ being the sample of a Gaussian noise random field, white in time and regular in space, see Section 6. The goal was to provide some probabilistic representation for non-conservative non-linear PDEs (1.1) by introducing some exponential weights defining Feynman-Kac measures instead of probability measures. An interesting aspect of this strategy is that it is potentially able to represent an extended class of second order non-linear PDEs. One particularity of MFKE equations is that the probabilistic representation involves the past of the process (via the exponential weights). In this context, it is worth to quote the recent paper [38] which proposes a probabilistic representation, which also includes a dependence on the past, in relation with Keller-Segel model with application to chemotaxis.

It is important to consider carefully the two major features differentiating the MFKE (1.2) from the traditional setting of MSDEs. To recover the traditional setting one has to do the following.

1. First, one has to put $\Lambda = 0$ in the third line equation of (1.2) Then $u(t, \cdot)$ is explicitly given by the third line equation of (1.2) and reduces to the density of the marginal distribution, $\mathcal{L}(Y_t)$. When $\Lambda \neq 0$, the relation between $u(t, \cdot)$ and the process Y is more complex. Indeed, not only does Λ embed an additional non-linearity with respect to u, but it also involves the whole past trajectory $(Y_s)_{0 \leq s \leq t}$ of the process Y.

2. Secondly, one has to replace the pointwise dependence $b(s, Y_s, u(s, Y_s))$ in equation (1.2) with a mollified dependence $b(s, Y_s, \int_{\mathbb{R}^d} K(Y_s - y)u(s, y)dy)$, where the dependence with respect to $u(s, \cdot)$ is Wasserstein continuous. Here $K : \mathbb{R} \to \mathbb{R}$ is a convolution kernel.

One interesting aspect of probabilistic representation (1.2) is that it naturally yields numerical approximation schemes involving weighted interacting particle systems. More precisely, we consider N interacting particles $(\xi^{i,N})_{i=1,\cdots N}$ with N i.i.d. Brownian motions $(W^i)_{i=1,\cdots N}$, i.e.

$$\begin{cases} \xi_t^{i,N} = \xi_0^{i,N} + \int_0^t \sigma\Big(s, \xi_s^{i,N}, u_s^{N,\varepsilon}(\xi_s^{i,N})\Big)dW_s^i + \int_0^t b\Big(s, \xi_s^{i,N}, u_s^{N,\varepsilon}(\xi_s^{i,N})\Big)ds \\ \xi_0^{i,N} \text{ i.i.d. } \sim \mathbf{u_0} \\ u_t^{N,\varepsilon}(\xi_t^i) = \sum_{j=1}^N \omega_t^{j,N} K^\varepsilon(\xi_t^{i,N} - \xi_t^{j,N}) \,, \end{cases} \qquad (3.1)$$

where the mollifier K^ε is such that $K^\varepsilon(x) = \frac{1}{\varepsilon^d}\phi(\frac{x}{\varepsilon}) \xrightarrow[\varepsilon \to 0]{} \delta_0$ and the weights $\omega_t^{j,N}$ for $j = 1, \cdots, N$ verify

$$
\begin{aligned}
\omega_t^{j,N} &:= \exp\left\{\int_0^t \Lambda\left(r, \xi_r^{j,N}, u_r^{\varepsilon,N}(\xi_r^{j,N}), \nabla u_r^{\varepsilon,N}(\xi_r^{j,N})\right) dr\right\} \\
&= \omega_s^{j,N} \exp\left\{\int_s^t \Lambda\left(r, \xi_r^{j,N}, u_r^{\varepsilon,N}(\xi_r^{j,N}), \nabla u_r^{\varepsilon,N}(\xi_r^{j,N})\right) dr\right\}.
\end{aligned}
$$

[48, 47] consider the case of pointwise semilinear PDEs of the form

$$
\begin{cases}
\partial_t u = \frac{1}{2}\partial_{xx}^2(\sigma^2(t,x)u) - \partial_x b(t,x)u) + \Lambda(t,x,u,\nabla u)u \\
u(0,x) = u_0(x),
\end{cases}
\tag{3.2}
$$

for which the target probabilistic representation is

$$
\begin{cases}
Y_t = Y_0 + \int_0^t \sigma\left(s, Y_s\right)dW_s + \int_0^t b\left(s, Y_s\right)ds \\
Y_0 \sim \mathbf{u_0} \\
\int \varphi(x)u_t(x)dx := \mathbb{E}\left[\varphi(Y_t) \exp\left\{\int_0^t \Lambda\left(s, Y_s, u_s(Y_s), \nabla u_s(Y_s)\right)ds\right\}\right].
\end{cases}
\tag{3.3}
$$

We set

$$
L_t f := \frac{1}{2}\sigma^2(t,x)f''(x) + b(t,x)f'(x), t \in]0,T[, \quad \text{for any } f \in C^2(\mathbb{R}).
\tag{3.4}
$$

Let us consider the family of Markov transition functions $P(s, x_0, t, \cdot)$ associated with (L_t), see [48]. We recall that if X is a processs solving the first line of (3.1) with $X_s \equiv x_0 \in \mathbb{R}$, then $\int_{\mathbb{R}} P(s, x_0, t, x)f(x)dx = \mathbb{E}(f(X_t)), t \geq s$, for every bounded Borel function $f : \mathbb{R} \to \mathbb{R}$. $u : [0,T] \times \mathbb{R} \to \mathbb{R}$ will be called **mild solution** of (3.2) (related to (L_t)) if for all $\varphi \in C_0^\infty(\mathbb{R})$, $t \in [0,T]$,

$$
\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x)u(t,x)dx &= \int_{\mathbb{R}^d} \varphi(x)\int_{\mathbb{R}^d} \mathbf{u_0}(dx_0)P(0,x_0,t,dx) \\
&+ \int_{[0,t]\times\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \varphi(x)P(s,x_0,t,dx)\right)\Lambda(s,x_0,u(s,x_0),\nabla u(s,x_0))u(s,x_0)dx_0ds.
\end{aligned}
$$

The following theorem states conditions ensuring equivalence between (3.3) and (3.2) together with the convergence of the related particle approximation (3.1).

Theorem 3.1. *We suppose that σ and b are Lipschitz with linear growth and Λ is bounded.*

1. *Let $u : [0,T] \times \mathbb{R} \to \mathbb{R} \in L^1([0,T]; W^{1,1}(\mathbb{R}^d))$. u is a mild solution of PDE (3.2) if and only if u verifies (3.3).*

2. *Suppose that $\sigma \geq c > 0$ and Λ is uniformly Lipschitz w.r.t. to u and ∇u. There is a unique mild solution in $L^1([0,T]; W^{1,1}(\mathbb{R}) \cap L^\infty([0,T] \times \mathbb{R})$ of (3.2), therefore also of (3.3).*

3. *Under the same assumption of item 2., the particle approximation $u^{N,\varepsilon}$ (3.1) converges in $L^1([0,T]; W^{1,1}(\mathbb{R})$ to the solution of (3.2) as $N \to \infty$ and $\varepsilon(N) \to 0$ slowly enough.*

Item 1. was the object of Theorem 3.5 in [48]. Item 2. (resp. item 3.) was treated in Theorem 3.6 (resp. Corollary 3.5) in [48].

Remark 3.2. *The error induced by the discrete time approximation of the particle system was evaluated in [47].*

[49] considers the case where b is replaced by $b + b_1$ where b is only supposed bounded Borel, without regularity assumption on the space variable. In particular they treat the pointwise semilinear PDEs of the form

$$
\begin{cases}
\partial_t u = \frac{1}{2}\partial_{xx}^2(\sigma^2(t,x)u) - \partial_x\left((b(t,x) + b_1(t,x,u))u\right) + \Lambda(t,x,u)u \\
u(0,x) = u_0(x),
\end{cases}
\tag{3.5}
$$

8

for which the target probabilistic representation is

$$\begin{cases} Y_t = Y_0 + \int_0^t \sigma(s, Y_s) dW_s + \int_0^t \left[b(s, Y_s) + b_1\Big(s, Y_s, u(s, Y_s)\Big) \right] ds \\ Y_0 \sim \mathbf{u_0} \\ \int \varphi(x) u_t(x) dx := \mathbb{E}\left[\varphi(Y_t) \, \exp\left\{ \int_0^t \Lambda\big(s, Y_s, u_s(Y_s)\big) ds \right\} \right]. \end{cases} \tag{3.6}$$

The following theorem states conditions ensuring equivalence between (3.6) and (3.5) together with well-posedness conditions for both equations.

Theorem 3.3. *We formulate the following assumptions.*

1. *The PDE in the sense of distributions $\partial u = L_t^* u_t$ admits as unique solution $u \equiv 0$, where L_t was defined in (3.4).*

2. *b is bounded measurable and σ is continuous $\sigma \geq c > 0$ for some constant $c > 0$.*

3. *$b_1, \Lambda : [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is uniformly bounded, Lipschitz with respect to the third argument.*

4. *The family of Markov transition functions associated with (L_t), are of the form $P(s, x_0, t, dx) = p(s, x_0, t, x) dx$, i.e. they admit measurable densities p.*

5. *The first order partial derivatives of the map $x_0 \mapsto p(s, x_0, t, x)$ exist in the distributional sense.*

6. *For almost all $0 \leq s < t \leq T$ and $x_0, x \in \mathbb{R}$ there are constants $C_u, c_u > 0$ such that*

$$p(s, x_0, t, x) \leq C_u q(s, x_0, t, x) \quad and \quad |\partial_{x_0} p(s, x_0, t, x)| \leq C_u \frac{1}{\sqrt{t-s}} q(s, x_0, t, x) , \tag{3.7}$$

where $q(s, x_0, t, x) := \left(\frac{c_u(t-s)}{\pi} \right)^{\frac{1}{2}} e^{-c_u \frac{|x-x_0|^2}{t-s}}$ is a Gaussian probability density.

The following results hold.

1. *Let $u \in (L^1 \cap L^\infty)([0, T] \times \mathbb{R})$. u is a solution of PDE (3.5) in the sense of distributions if and only if u verifies (3.6) for a solution Y in the sense of probability laws.*

2. *There is a unique solution $u \in (L^1 \cap L^\infty)([0, T] \times \mathbb{R})$ in the sense of distributions of PDE (3.5) (and therefore of (3.6)).*

The result 1. (resp. result 2.) was the object of Theorem 12. (resp. Proposition 16., Theorems 13., 22.) of [49].

Remark 3.4. *Under more restrictive assumptions on b, item 3. of Theorem 13. in [49] states the well-posedness of (3.6)) in the sense of strong existence and pathwise uniqueness.*

[46] and [45] studied a mollified version of (1.1), whose probabilistic representation falls into the Wasserstein continuous traditional setting mentioned above. Following the spirit of [63], a fixed point argument was carried out in the general case in [46] to prove well-posedness of

$$\begin{cases} Y_t = Y_0 + \int_0^t \sigma\Big(s, Y_s, K * u_s(Y_s)\Big) dW_s + \int_0^t b\Big(s, Y_s, K * u_s(Y_s)\Big) ds \\ Y_0 \sim \mathbf{u_0} \\ (K * u_t)(x) := \mathbb{E}\left[K(x - Y_t) \, \exp\left\{ \int_0^t \Lambda\big(s, Y_s, K * u_s(Y_s)\big) ds \right\} \right], \end{cases} \tag{3.8}$$

9

where $K : \mathbb{R} \to \mathbb{R}$ is a mollified kernel. We remark that if (Y, u) is a solution of (3.8), then u is a solution (in the sense of distribution) of

$$\begin{cases} \partial_t u = \frac{1}{2}\partial_{xx}^2(\sigma^2(t, x, K * u)u) - \partial_x(b(t, x, K * u)u) + \Lambda(t, x, K * u)u \\ u(0, x) = u_0(x). \end{cases} \tag{3.9}$$

Remark 3.5. *1. Existence and uniqueness results (in the strong sense and in the sense of probability laws) for the MFKE (3.8) are established under various technical assumptions, see [45].*

2. Chaos propagation for the interacting particle system (3.1) providing an approximation to the regularized PDE (3.9), as $N \to \infty$ (for fixed K), [46].

4 McKean representation of a Fokker-Planck equation with terminal condition

Let us consider the PDE with terminal condition (1.5) and $\Lambda = 0$.

$$\begin{cases} \partial_t u & = & \frac{1}{2}\sum_{i,j=1}^{d} \partial_{ij}^2 \left((\sigma\sigma^t)_{i,j}(t, x)u \right) - div\,(b(t, x)u) \\ u(T, dx) & = & \mathbf{u_T}(dx), \end{cases} \tag{4.1}$$

where $\mathbf{u_T}$ is a given Borel probability measure. In the present section we assume the following.

Assumption 1. *Suppose that (4.1) admits uniqueness, i.e. that there is at most one solution of (4.1).*

Remark 4.1. *Different classes of sufficient conditions for that are provided in [37].*

A natural representation of (4.1) is the following MSDE, where β is a Brownian motion.

$$\begin{cases} Y_t = \xi - \int_0^t \tilde{b}\,(s, Y_s; v_s)\,ds + \int_0^t \sigma\,(T - s, Y_s)\,d\beta_s, t \in [0, T] \\ \int_{\mathbb{R}^d} v_t(x)\varphi(x)dx = \mathbb{E}(\varphi(Y_t)), t \in [0, T] \\ \xi \sim \mathbf{u_T}, \end{cases} \tag{4.2}$$

where $\tilde{b}(s, y; v_s) = (\tilde{b}^1(s, y; v_s), \ldots, \tilde{b}^d(s, y; v_s))$ is defined as

$$\tilde{b}\,(s, y; v_s) := \left[\frac{div_y\left(\sigma\sigma_{j\cdot}^t\,(T - s, y)\,v_s\,(y) \right)}{v_s\,(y)} \right]_{j \in [\![1, d]\!]} - b\,(T - s, y)\,. \tag{4.3}$$

For $d = 1$ previous expression gives

$$\tilde{b}(s, y; v_s) := \frac{\left(\sigma^2(T - s, \cdot)v_s \right)'}{v_s}(y) - b\,(T - s, y)\,. \tag{4.4}$$

Remark 4.2. (4.2) *is in particular fulfilled if Y is the time reversal process $\hat{X}_t := X_{T-t}$ of a diffusion satisfying the SDE*

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s)ds + \sigma(s, X_s)dW_s, t \in [0, T] \\ X_0 \sim \mathbf{u_0} \in \mathcal{P}(\mathbb{R}). \end{cases} \tag{4.5}$$

This happens under locally Lipschitz conditions on σ and b and minimal regularity conditions on the law density p_t of X_t. Indeed in [34], the authors prove that

$$\hat{X}_t = X_T + \int_0^t \tilde{b}\left(s, \hat{X}_s; p_{T-s}\right) ds + \int_0^t \sigma\left(T-s, \hat{X}_s\right) d\beta_s, \ t \in [0, T], \tag{4.6}$$

where $\tilde{b}$ is defined in (4.3) and p_t is the density of X_t. We emphasize that the main difference between (4.2) and (4.6) is that in the first equation the solution is a couple (Y, v), in the second one, a solution is just Y, p being exogeneously defined by (4.5).

We observe now that a solution (Y, v) of (4.2) provides a solution u of (4.1). This justifies indeed the terminology of probabilistic representation.

Proposition 4.3.   1. Let (Y, v) be a solution of (4.2). Then $u(t, \cdot) := v(T - t, \cdot), t \in [0, T])$, is a solution of (4.1) with terminal value $\mathbf{u_T}$.

    2. If (4.1) admits at most one solution, then there is at most one v such that (Y, v) solves (4.2).

Proof. In order to avoid technicalities which complicate the task of the reader we express the proof for $d = 1$. We prove 1. since 2. is an immediate consequence of 1.

Let $\phi \in \mathcal{C}^\infty(\mathbb{R})$ with compact support and $t \in [0, T]$. Itô formula gives

$$\mathbb{E}\left[\phi\left(Y_{T-t}\right)\right] - \int_{\mathbb{R}^d} \phi(y)\, \mathbf{u_T}(dy) = \int_0^{T-t} \mathbb{E}\left[\tilde{b}\left(s, Y_s; v_s\right)\phi'\left(Y_s\right) + \frac{1}{2}\left(\sigma^2\left(T-s, Y_s\right)\phi''\left(Y_s\right)\right)\right] ds.$$

Fixing $s \in [0, T]$, we have

$$
\begin{aligned}
\mathbb{E}\left[\tilde{b}(s, Y_s; v_s)\phi'(Y_s)\right] &= \int_{\mathbb{R}} (\sigma^2(T-s, \cdot)v_s)'(y)\phi'(y)dy - \int_{\mathbb{R}} b(T-s, y)\phi'(y)v_s(y)dy \\
&= -\int_{\mathbb{R}} (\sigma^2)(T-s, y)\phi''(y)v_s(y)dy - \int_{\mathbb{R}} b(T-s, y), \phi'(y)v_s(y)dy.
\end{aligned}
$$

Hence, we have the identity

$$\mathbb{E}\left[\phi\left(Y_{T-t}\right)\right] = \int_{\mathbb{R}} \phi(y)\, \mathbf{u_T}(dy) - \int_0^{T-t} \int_{\mathbb{R}} L_{T-s}\phi(y)\, v_s(y)\, dy ds.$$

Applying the change of variable $t \mapsto T - t$, we finally obtain the identity

$$\int_{\mathbb{R}} \phi(y)\, v_{T-t}(y)\, dy = \int_{\mathbb{R}} \phi(y)\, \mathbf{u_T}(dy) - \int_t^T \int_{\mathbb{R}^d} L_s\phi(y)\, v_{T-s}(y)\, dy ds.$$

This means that $t \mapsto u_t$ is a solution of (1.2) with terminal value $\mathbf{u_T}$. $\qquad \square$

Remark 4.4. *Precise discussions on existence and uniqueness of (4.2) are provided in [37]. In particular we have the following.*

    1. *There is at most one solution (in law) (Y, v) of (4.2) such that v is locally bounded in $[0, T[\times \mathbb{R}^d$.*

    2. *There is at most one strong solution (Y, v) of (4.2) such that v is locally Lipschitz in $[0, T[\times \mathbb{R}^d$.*

Item 1. is a consequence of Theorem 10.1.3 of [62]. Item 2. is a consequence of usual pathwise uniqueness arguments for SDEs.

5 Probabilistic representation with jumps for non-conservative PDEs

In this section, we outline the link between non-conservative PDEs and non-linear jump diffusions. This kind of representation was emphasized in [27, 24] to design interacting jump particles systems to approximate time-dependent Feynman-Kac measures. For simplicity, we present this correspondence in the simple case of the non-conservative linear PDE (1.1) when the coefficients do not depend on the solution, see (1.1). However, the same ideas could be extended to the non-linear case where the coefficients σ, b, Λ may depend on the PDE solution.

Let us consider the SDE

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 \sim \mathbf{u_0}, \end{cases} \tag{5.1}$$

where W is a one-dimensional Brownian motion. Assume that (5.1) admits a (weak) solution. Let Λ be a bounded and negative function defined on $[0, T] \times \mathbb{R}$. For any $t \in [0, T]$, we define the measure, $\gamma(t, \cdot)$ such that for any real-valued Borel measurable test function φ

$$\int \gamma(t, dx)\varphi(x) = \mathbb{E}\left[\varphi(X_t) \exp\left(\int_0^t \Lambda(s, X_s)ds\right)\right] . \tag{5.2}$$

We recall that by Section 3 we know that γ is a solution (in the distributional sense) of the linear and non-conservative PDE

$$\begin{cases} \partial_t\gamma = \frac{1}{2}\partial_{xx}^2(\sigma^2(t, x)\gamma) - \partial_x(b(t, x)\gamma) + \Lambda(t, x)\gamma \\ \gamma(0, \cdot) = \mathbf{u_0}. \end{cases} \tag{5.3}$$

Remark 5.1. *If uniqueness of distributional solutions of* (5.3) *holds, then γ defined by (5.1,5.2) is the unique solution of* (5.3).

Let $\gamma(t, \cdot)$ be a solution of (5.2) which for each t is a positive measure. We introduce the family of probability measures $(\eta(t, \cdot))_{t \in [0,T]}$, obtained by normalizing $\gamma(t, \cdot)$, such that for any real valued bounded and measurable test function φ we have

$$\int \eta(t, dx)\varphi(x) := \frac{\int \gamma(t, dx)\varphi(x)}{\int \gamma(t, dx)} . \tag{5.4}$$

By simple differentiation of the above ratio and using the fact that γ satisfies (5.3), we obtain that η is a solution in the distributional sense of the integro-differential PDE

$$\begin{cases} \partial_t\eta = \frac{1}{2}\partial_{xx}(\sigma^2(t, x)\eta) - \partial_x(b(t, x)\eta) + \left(\Lambda(t, x) - \int \eta(t, dx)\Lambda(t, x)\right)\eta \\ \eta_0 = \mathbf{u_0} . \end{cases} \tag{5.5}$$

Besides one can express $\gamma(t, \cdot)$ as a function of $(\eta(s, \cdot))_{s \in [0,t]}$. Indeed, since γ solves the linear PDE (5.3) then in particular approaching the constant test function 1, yields

$$\partial_t \int \gamma(t, dx) = \int \gamma(t, dx)\Lambda(t, x) = \int \gamma(t, dx) \int \eta(t, dx)\Lambda(t, x) ,$$

which gives $\int \gamma(t, dx) = \exp\left(\int_0^t \int \eta(s, dx)\Lambda(s, x)ds\right)$. Then by definition (5.4) of η,

$$\gamma(t, \cdot) = \left(\int \gamma(t, dx)\right)\eta(t, \cdot) = \exp\left(\int_0^t \int \eta(s, dx)\Lambda(s, x)ds\right)\eta(t, \cdot) . \tag{5.6}$$

We already know that for any solution γ of (5.3) one can build a solution η to (5.5) according to relation (5.4). Conversely, for any solution η of (5.5), by similar manipulations one can build a solution γ of (5.3) according to (5.6). Hence well-posedness of (5.3) is equivalent to well-posedness of (5.5).

We propose now an alternative probabilistic representation to (5.1) and (5.2) of (5.3). Let us introduce the non-linear jump diffusion Y (if it exists), which evolves between two jumps according to the diffusion dynamics (5.1) and jumps at exponential times with intensity $-\Lambda_t(Y_t) \geq 0$ to a new point independent of the current position and distributed according to the current law, $\mathcal{L}(Y_t)$. More specifically, we consider a process Y solution of the following non-linear (in the sense of McKean) SDE with jumps

$$
\begin{cases}
dY_t = b(t, Y_{t-})dt + \sigma(t, Y_{t-})dW_t + \int_{\mathbb{R}} x\mathbf{1}_{|x|>1} J_t(\mu_{t-}, Y_{t-}, dx)dt + \int_{\mathbb{R}} x\mathbf{1}_{|x|\leq 1}(J_t - \bar{J}_t)(\mu_{t-}, Y_{t-}, dx)dt \\
Y_0 \sim \mathbf{u_0} \\
\mu_{t-} = \mathcal{L}(Y_{t-})\,,
\end{cases}
$$
(5.7)

where J denotes the jump measure and $\bar{J}$ is the associated predictable compensator such that for any probability measure ν on $\mathbb{R}$

$$
\bar{J}_t(\nu, y, d(y'-y)) = -\Lambda_t(y)\nu(dy')\,, \quad \text{for any} \quad y\,, y' \in \mathbb{R}\,.
$$

Note that well-posedness analysis of the above equation constitutes a difficult task. In particular, [40] analyzes well-posedeness and particle approximations of some types of non-linear jump diffusions. However, contrarily to (5.7), the nonlinearity considered in [40] is concentrated on the diffusion matrix (assumed to be Lipschitz in the time-marginals of the process w.r.t. Wasserstein metric) and does not involve the jump measure which is assumed to be given.

Assume that MSKE (5.7) admits a weak solution. By application of Itô formula, we observe that the marginals of Y are distributional solutions of (5.5). Indeed, for any real valued test function in $\mathcal{C}_0^\infty(\mathbb{R})$

$$
\begin{aligned}
\mathbb{E}[\varphi(Y_t)] &= \mathbb{E}[\varphi(Y_0)] \\
&+ \int_0^t \mathbb{E}\left[b(s, Y_{s-})\varphi'(Y_{s-}) + \frac{1}{2}\sigma^2(s, Y_{s-})\varphi''(Y_{s-})\right] ds \\
&+ \int_0^t \mathbb{E}\left[\int \varphi(Y_{s-} + x)\bar{J}_s(\mu_{s-}, Y_{s-}, dx)\right] ds \\
&- \int_0^t \mathbb{E}\left[\varphi(Y_{s-})\bar{J}_s(\mu_{s-}, Y_{s-}, \mathbb{R})\right] ds\,.
\end{aligned}
$$
(5.8)

Conclusion 5.2. *Suppose that (5.3) admits a unique distributional solution γ; let η defined by (5.4). Suppose the existence of a (weak) solution X (resp. Y) of (5.1) (resp. (5.7)).*

1. *η is the unique solution (in the sense of distributions) of (5.5). Moreover $\int_{\mathbb{R}} \varphi(x)\eta(t, dx) = \mathbb{E}[\varphi(Y_t)], t \geq 0.$*

2. *We obtain the following identities for γ and η:*

$$
\begin{aligned}
\int \gamma(t, dx)\varphi(x) &= \mathbb{E}\left[\varphi(X_t)\exp\left(\int_0^t \Lambda(s, X_s)ds\right)\right] \\
&= \exp\left(\int_0^t \int \eta(s, dx)\Lambda(s, x)ds\right)\eta_t(\varphi) \\
&= \exp\left(\int_0^t \mathbb{E}[\Lambda(s, Y_s)]ds\right)\mathbb{E}[\varphi(Y_t)].
\end{aligned}
$$
(5.9)

Using the above identities allows to design discrete time interacting particles systems with geometric interacting jump processes. In particular, in [26] the authors provide non asymptotic bias and variance theorems w.r.t. the time step and the size of the system, allowing to numerically approximate the time-dependent family of Feynman-Kac measures γ.

6 McKean SDEs in random environment

6.1 The (S)PDE and the basic idea

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space. We consider a progressively measurable random field $(\xi(t, x))$. We want to discuss probabilistic representations of

$$\begin{cases} \partial_t u &= \frac{1}{2}\Delta(\beta(u)) + \partial_t \xi(t, x) u(t, x), \quad \text{with} \quad \beta(u) = \sigma^2(u)u. \\ u(0, \cdot) &= \mathbf{u_0}. \end{cases} \tag{6.1}$$

Suppose for a moment that ξ has random realizations which are smooth in time so that

$$\partial_t \xi(t, x) = \Lambda(t, x; \omega). \tag{6.2}$$

Under some regularity assumptions on Λ, (6.1) can be observed as a randomization of a particular case of the PDE (1.1). For each random realization $\omega \in \Omega$, the natural (double) probabilistic representation is

$$\begin{cases} Y_t = Y_0 + \int_0^t \sigma\Big(u(s, Y_s)\Big) dW_s \\ Y_0 \sim \mathbf{u_0} \\ \int \varphi(x) u(t, x) dx = \mathbb{E}^\omega \left[\varphi(Y_t) \exp\left\{ \int_0^t \Lambda\big(s, Y_s; \omega)\big) ds \right\} \right], \quad \text{for } t \in [0, T], \end{cases} \tag{6.3}$$

where $\mathbb{E}^\omega$ denotes the expectation with frozen ω. However the assumption (6.2) is not realistic and we are interested in $\partial_t \xi$ being a white noise in time. Let $N \in \mathbb{N}^*$. Let $B^1, \ldots, B^N$ be N independent $(\mathcal{F}_t)$-Brownian motions, $e^1, \ldots, e^N$ be functions in $C_b^2(\mathbb{R})$. In particular they are H^{-1}-multiplier, i.e. the maps $\varphi \to \varphi e^i$ are continuous in H^{-1}.

We define the random field $\xi(t, x) = \sum_{i=0}^N e^i(x) B_t^i$, where $B_t^0 \equiv t$ and we consider the SPDE (6.1) in the sense of distributions, i.e.

$$\int_\mathbb{R} \varphi(x) u(t, x) dx = \int_\mathbb{R} \varphi(x) \mathbf{u_0}(dx) + \frac{1}{2} \int_0^t \int_\mathbb{R} \varphi''(x) \sigma^2(u(s, x)) ds dx + \int_0^t \int_\mathbb{R} \varphi(x) u(s, x) \xi(ds, x) dx, \tag{6.4}$$

where the latter stochastic integral is intended in the Itô sense.

6.2 Well-posedness of the SPDE

The theorem below contains results taken from [9, 61].

Theorem 6.1. *Suppose that β is Lipschitz.*

- *Suppose that $u_0 \in L^2(\mathbb{R})$. There is a solution to equation* (6.1).

- *Assume further that β is non-degenerate, i.e. $\beta(r) \geq ar^2$, $r \in \mathbb{R}$, where $a > 0$. Then, there is a solution u to* (6.1) *for any probability $\mathbf{u_0}(dx)$ (even in $H^{-1}(\mathbb{R})$).*

- *There is at most one solution in the class of random fields u such that $\int_{[0,T]\times\mathbb{R}} u^2(t,x)dtdx < \infty$ a.s.*

Remark 6.2.   • *Previous result extends to the case of an infinite number of modes e^i and for $d \geq 1$.*

- *We remark that the $\partial_t \xi(t,x)$ is a coloured noise (in space). The case of space-time white noise seems very difficult to treat.*

6.3   McKean equation in random environment

Given a local martingale M, $\mathcal{E}(M)$ denotes the Doléans exponential of M i.e. $\exp(M_t - \frac{1}{2}[M]_t), t \geq 0$. We say that a filtered probability space $(\Omega_0, \mathcal{G}, (\mathcal{G}_t), Q)$ is a **suitable enlarged space** of $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, if the following holds.

1. There is a measurable space $(\Omega_1, \mathcal{H})$ with $\Omega_0 = \Omega \times \Omega_1$, $\mathcal{G} = \mathcal{F} \otimes \mathcal{H}$ and a random kernel $(\omega, H) \mapsto \mathbb{Q}^\omega(H)$ defined on $\Omega \times \mathcal{H} \to [0,1]$ such that the probability $\mathbb{Q}$ on $(\Omega_0, \mathcal{G})$ is defined by $d\mathbb{Q}(\omega, \omega_1) = d\mathbb{P}(\omega)\mathbb{Q}^\omega(\omega_1)$.

2. The processes $B^1, \ldots, B^N$ are $(\mathcal{G}_t)$- Brownian motions where $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}$.

Definition 6.3. *We say that the non-linear doubly-stochastic diffusion*

$$
\begin{cases}
Y_t & = & Y_0 + \int_0^t \Phi(u(s, Y_s))dW_s, \\
\int \varphi(x)u(t,x)dx & = & \mathbb{E}^{\mathbb{Q}^\omega}\left(\varphi(Y_t(\omega, \cdot))\mathcal{E}_t\left(\int_0 \xi(ds, Y_s)(\omega, \cdot)\right)\right), \\
\xi - \mathrm{Law}(Y_0) & = & \mathbf{u_0}(dx),
\end{cases}
\tag{6.5}
$$

*admits **weak existence** on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ if there is a suitably enlarged probability space $(\Omega_0, \mathcal{G}, (\mathcal{G}_t), \mathbb{Q})$ an $(\mathcal{G}_t)$-Brownian motion W such that (6.5) is verified. The couple (Y, u) will be called **weak solution** of (6.5).*

Remark 6.4.   • *We remark that the second line in (6.5) represents a sort of ξ-**marginal weighted law** of Y_t.*

- *Let (Y, u) be a solution to (6.5). Then u is a solution to (6.1).*

- *Such representation allows to show that $u(t,x) \geq 0$, $d\mathbb{P}dtdx$ a.e. and, at least if $e^0 = 0$, $\mathbb{E}\left(\int_{\mathbb{R}} u(t,x)dx\right) = 1$, so that the conservativity is maintained at the expectation level.*

Definition 6.5. *Let two measurable random fields $u^i : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}, i = 1,2$ on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$, and Y^i, on a suitable extended probability space $(\Omega_0^i, \mathcal{G}^i, (\mathcal{G}_t^i), \mathbb{Q}^i), i = 1,2$, such that (Y^i, u^i) are (weak) solutions of (6.5) on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. If we always have that $(Y^1, B^1, \ldots, B^N)$ and $(Y^2, B^1, \ldots, B^N)$ have the same law, then we say that (6.5) admits **weak uniqueness** (on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$).*

Theorem 6.6. *Under the assumption of Theorem 6.1 equation (6.5) admits (weak) existence and uniqueness on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$.*

7   McKean representation of stochastic control problems

7.1   Stochastic control problems and non-linear Partial Differential Equations

Let us briefly recall the link between stochastic control and non-linear PDEs given by the Hamilton-Jacobi-Bellman (HJB) equation. We refer for instance to [66, 59, 30] for more details. Consider a *state process*

$(X_s^{t_0,x,\alpha})_{t_0 \leq s \leq T}$ on $[t_0, T] \times \mathbb{R}^d$ solution to the controlled SDE

$$\begin{cases} dX_s^{t_0,x,\alpha} &= b\big(s, X_s^{t_0,x,\alpha}, \alpha(s, X_s^{t_0,x,\alpha})\big)ds + \sigma\big(s, X_s^{t_0,x,\alpha}, \alpha(s, X_s^{t_0,x,\alpha})\big)dW_s \\ X_{t_0}^{t_0,x,\alpha} &= x \, , \end{cases} \tag{7.1}$$

where W denotes the Brownian motion on $[t_0, T] \times \mathbb{R}^d$, and $\alpha(s, X_s^{t_0,x,\alpha})$ represents *Markovian* control in the sense that the control at time t is supposed here to depend on t and on the current value of the state process:

$$\alpha \in \mathcal{A}_{t_0, T} := \big\{ \alpha : (t, x) \in [t_0, T] \times \mathbb{R}^d \mapsto \alpha(t, x) \in A \subset \mathbb{R}^k \big\} \, , \tag{7.2}$$

A being a subset of $\mathbb{R}^k$. For a given initial time and state $(t_0, x) \in [0, T] \times \mathbb{R}^d$, we are interested in maximizing, over the Markovian controls $\alpha \in \mathcal{A}_{t_0, T}$, the criteria

$$J(t_0, x, \alpha) := \mathbb{E}\left[g(X_T^{t_0,x,\alpha}) + \int_{t_0}^T f\big(s, X_s^{t_0,x,\alpha}, \alpha(s, X_s^{t_0,x,\alpha})\big)ds \right] \, . \tag{7.3}$$

In the above criteria, the function f is called the *running gain* whereas g is called the *terminal gain*.

Remark 7.1. *At first glance, the set of control processes of the form $\alpha_t = \alpha(t, X_t)$ defined in (7.2) may appear too restrictive compared to a larger set of non-anticipative controls (α_t) which may depend on all the past history of the state process (X_t). However, in the framework of Markov control problems (for which the state process $(X_t^{t_0,x,a})$ is Markov, as soon as the control is fixed to a deterministic value $\alpha_t = a \in A$, for all $t \in [t_0, T]$), it is well-known that the optimal control process (α_t) lies in the set of Markovian controls verifying $\alpha_t = \alpha(t, X_t)$. Hence, considering controls of the particular form (7.2) is done here without loss of generality.*

To tackle this finite horizon stochastic control problem, the usual approach consists in introducing the associated value (or *Bellman*) function $v : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}$ representing the maximum gain one can expect, starting from time t at state x, i.e.

$$v(t, x) := \sup_{\alpha \in \mathcal{A}_{t, T}} J(t, x, \alpha) \, , \quad \text{for } t \in [t_0, T] \, . \tag{7.4}$$

Note that the terminal condition is known, which fixes $v(T, x) = g(x)$, whereas the initial condition $v(t_0, x)$ corresponds to the solution of the original minimization problem. The value function is then proved to verify the *Dynamic Programming Principle (DPP)* which consists in the backward induction

$$v(t, x) = \sup_{\alpha \in \mathcal{A}_{t, \tau}} \mathbb{E}\Big[\int_t^\tau f\big(s, X_s^{t,x,\alpha}, \alpha(s, X_s^{t,x,\alpha})\big)ds + v(\tau, X_\tau^{t,x,\alpha}) \Big] \, , \quad \text{for any stopping time } \tau \in]t, T] \, . \tag{7.5}$$

Under continuity assumptions on b, σ, f, g, using DPP together with Itô formula allows to characterize v as a viscosity solution of the HJB equation

$$\begin{cases} v(T, x) = g(x) \\ \partial_t v(t, x) + H(t, x, \nabla v(t, x), \nabla^2 v(t, x)) = 0 \, , \end{cases} \tag{7.6}$$

where ∇ and ∇^2 denote the gradient and the Hessian operators and the so-called, *Hamiltonian*, H denotes the real valued function defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{S}^d$ ($\mathcal{S}^d$ denoting the set of symmetric matrices in $\mathbb{R}^{d \times d}$), such that

$$H(t, x, \delta, \gamma) := \sup_{a \in A} \left\{ f(t, x, a) + b(t, x, a)^\top \delta(t, x) + \frac{1}{2} Tr[\sigma\sigma'(t, x, a)\gamma(t, x)] \right\} \, . \tag{7.7}$$

16

Note that (7.6) is a non-linear PDE because of the nonlinearity in the Hamiltonian induced by the supremum operator. Besides, assuming that, for all $(t, x) \in [t_0, T] \times \mathbb{R}^d$, the supremum in (7.7) is attained at a unique maximizer, then the optimal control α^* is directly obtained as a function of the Bellman funtion and its derivatives, i.e.

$$\alpha^*(t, x) = \arg \max_{a \in A} \left\{ f(t, x, a) + b(t, x, a)^\top \nabla v(t, x) + \frac{1}{2} Tr[\sigma \sigma^\top(t, x, a) \nabla^2 v(t, x)] \right\} . \tag{7.8}$$

Except in some very concrete cases such as the Linear Quadratic Gaussian (LQG) setting (where the states dynamics involve an affine drift with Gaussian noise and the cost is quadratic both w.r.t. the control and the state), there is no explicit solution to stochastic control problems. To numerically approximate the solution of equation (7.6), several approaches have been proposed, mainly differing in the way the value function v is interpreted. Indeed, as pointed out, v can be viewed either as the solution to the control problem (7.4), or as a (viscosity) solution of the non-linear PDE (7.6).

1. When v is defined as the solution to the control problem (7.4), a natural approach consists in dis-cretizing the time continuous control problem and apply the time discrete Dynamic Programming Principle [16]. Then the problem consists in maximizing over the controls, backwardly in time, the conditional expectation of the value function related to (7.5). The maximization at time step t_k can be done via a parametrization of the control $x \mapsto \alpha_{t_k}^\theta(x)$ via a parameter θ so that parametric optimiza-tion methods such as the stochastic gradient algorithm could be applied to maximize the expectation over θ. It remains to approximate the conditional expectations by numerical methods such as PDE, Fourier, Monte Carlo, Quantization or lattice methods... A great variety of numerical approximation schemes have been developed in the specific Bermudan option valuation test-bed [22, 50, 6, 67, 25, 21]. Alternatively, one can use Markov chain approximation method [43] which consists in a time-space discretization designed to obtain a proper Markov chain.

2. In the second approach we recall that v is viewed as the solution of (7.6). The problem amounts then to discretize a non-linear PDE. Then one can rely on numerical analysis methods (e.g. finite differences, or finite elements) and use monotone approximation schemes in the sense of Barles and Sougani-dis [11] to build converging approximation schemes, e.g. [18, 32]. This type of approach is in general limited to state space dimension lower than 4. To tackle higher dimensional problems, one approach consists in converting the PDE into a probabilistic setting in order to apply Monte Carlo types algo-rithms. To that end, various kinds of probabilistic representations of non-linear PDEs are available. Forward Backward Stochastic Differential Equations (FBSDE) were introduced in [58] as probabilistic representations of semi-linear PDEs. Then various types of numerical schemes for FBSDE have been developed. They mainly differ in the approach of evaluating conditional expectations: [20] (resp. [33], [28, 57]) use kernel (resp. regression, quantization) methods. Recently, important progresses have been done performing machine learning techniques, see e.g. [29, 36]. Branching processes [53, 35] can also provide probabilistic representations of semi-linear PDEs via Feynman-Kac formula. Non-linear SDEs in the sense of McKean [52] are another approach that constitutes the subject of the present paper.

3. Other approaches take advantage of both interpretations see for instance [31] and in [64].

7.2 McKean type representation in a toy control problem example

In order to illustrate the application of MFKEs to control problems, we consider a simple example corresponding to an inventory problem, for which the Hamiltonian maximization (7.7),(7.8) is explicit. The state $(X_t)_{t\in[t_0,T]}$ denotes the stock level evolving randomly with a control of the drift α:

$$\begin{cases} dX_t^{t_0,x,\alpha} = -\alpha(t, X_t^{t_0,x,\alpha})dt + \sigma dW_t \\ J(t_0, x, \alpha) = \sup_{\alpha \in \mathcal{A}_{t_0,T}} \mathbb{E}\Big[g(X_T^{t_0,x,\alpha}) - \int_{t_0}^T \Big[\big(\alpha(t, X_t^{t_0,x,\alpha}) - D_t\big)^2 + h(X_t^{t_0,x,\alpha})\Big] dt\Big]. \end{cases}$$

Bound constraints on the storage level are implicitly forced by the penalization h. A target terminal level is indicated by the terminal gain g, supposed here to be Lebesgue integrable. The objective is then to follow a deterministic target profile $(D_t)_{t\in[0,T]}$, on a given finite horizon $[t_0, T]$. When the admissible set in which the controls take their values $A = \mathbb{R}$, one can explicitly derive the optimal control as a function of the value function derivative

$$\alpha^*(t, \cdot) = D_t + \frac{1}{2}(\partial_x v)(t, \cdot) ,$$

which yields the following HJB equation

$$\partial_t v + \frac{1}{4}(\partial_x v)^2 + D_t \partial_x v + \frac{\sigma^2}{2}\partial_{xx} v - h = 0 .$$

Reversing the time, (with $t_0 = 0$) gives $(t, x) \mapsto u(t, x) := v(T - t, x)$ solution of

$$\begin{cases} \partial_t u &= \frac{1}{4}(\partial_x u)^2 + \frac{\sigma^2}{2}\partial_{xx} u + D_t \partial_x u - h, \\ u(0, x) &= g(x). \end{cases} \tag{7.9}$$

We recover the framework of (1.1), with $\Lambda(t, x, y, z) = \frac{1}{4}\frac{|z|^2}{y} - \frac{h(x)}{y}$ and $b(t, x, y) = -D_t$. Consequently the Bellman function v can be represented via

$$\begin{cases} Y_t = Y_0 + \sigma W_t - \int_0^t D_s ds \\ Y_0 \sim \frac{g(x)dx}{\int_{\mathbb{R}} g(y)dy} \\ \int \varphi(x)v(t, x)dx = \big(\int_{\mathbb{R}} g(y)dy\big) \quad \mathbb{E}\Big[\varphi(Y_{T-t}) \exp\Big\{\int_t^T \Lambda\big(s, Y_{T-s}, v(s, Y_{T-s}), \nabla v(s, Y_{T-s})\big)ds\Big\}\Big], \\ \text{for } t \in [0, T]. \end{cases} \tag{7.10}$$

Acknowledgments

References

[1] S. Albeverio and S. Ugolini. A Doob h-transform of the Gross-Pitaevskii Hamiltonian. *J. Stat. Phys.,* 161(2):486–508, 2015.

[2] D. G. Aronson and H. F. Weinberger. Multidimensional nonlinear diffusion arising in population genetics. *Adv. Math.,* 30:33–76, 1978.

18

[3] A. C. Bagtzoglou and J. Atmadja. Marching-jury backward beam equation and quasi-reversibility methods for hydrologic inversion: Application to contaminant plume spatial distribution recovery. *Water Resources Research*, 39(2), 2003.

[4] P. Bak. *How nature works*. Copernicus, New York, 1996. The science of self-organized criticality.

[5] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality. *Phys. Rev. A (3)*, 38(1):364–374, 1988.

[6] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multidimensional American options. *Math. Finance*, 15(1):119–168, 2005.

[7] V. Barbu and M. Röckner. Probabilistic representation for solutions to nonlinear Fokker-Planck equations. *SIAM J. Math. Anal.*, 50(4):4246–4260, 2018.

[8] V. Barbu, M. Röckner, and F. Russo. Probabilistic representation for solutions of an irregular porous media type equation: the irregular degenerate case. *Probab. Theory Related Fields*, 151(1-2):1–43, 2011.

[9] V. Barbu, M. Röckner, and F. Russo. Stochastic porous media equations in $\mathbb{R}^d$. *J. Math. Pures Appl. (9)*, 103(4):1024–1052, 2015.

[10] V. Barbu, M. Röckner, and F. Russo. Doubly probabilistic representation for the stochastic porous media type equation. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(4):2043–2073, 2017.

[11] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[12] J. V. Beck, B. Blackwell, and Ch. R. St. Clair Jr. *Inverse heat conduction: Ill-posed problems*. James Beck, 1985.

[13] N. Belaribi, F. Cuvelier, and F. Russo. Probabilistic and deterministic algorithms for space multidimensional irregular porous media equation. *SPDEs: Analysis and Computations*, 1(1):3–62, 2013.

[14] N. Belaribi and F. Russo. Uniqueness for Fokker-Planck equations with measurable coefficients and applications to the fast diffusion equation. *Electron. J. Probab.*, 17:no. 84, 28, 2012.

[15] S. Benachour, P. Chassaing, B. Roynette, and P. Vallois. Processus associés à l'équation des milieux poreux. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 23(4):793–832, 1996.

[16] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978. The discrete time case.

[17] P. Blanchard, M. Röckner, and F. Russo. Probabilistic representation for solutions of an irregular porous media type equation. *Ann. Probab.*, 38(5):1870–1900, 2010.

[18] J.F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numer. Anal.*, 41(3):1008–1021, 2003.

[19] M. Bossy and D. Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 1996.

[20] B. Bouchard and N. Touzi. Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111:175–206, 2004.

[21] B. Bouchard and X. Warin. Monte-Carlo valorisation of American options: facts and new algorithms to improve existing methods. In *Numerical Methods in finance*. Springer, 2012.

[22] M. Broadie, P. Glasserman, et al. A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance*, 7:35–72, 2004.

[23] R. Cafiero, V. Loreto, L. Pietronero, A. Vespignani, and S. Zapperi. Local rigidity and self-organized criticality for avalanches. *Europhysics Letters*, 29:111–116, 1995.

[24] P. Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.

[25] P. Del Moral, P. Hu, N. Oudjane, and B. Rémillard. On the robustness of the Snell envelope. *SIAM J. Financial Math.*, 2(1):587–626, 2011.

[26] P. Del Moral, P. E. Jacob, A. Lee, L. Murray, and G. W. Peters. Feynman-Kac particle integration with geometric interacting jumps. *Stoch. Anal. Appl.*, 31(5):830–871, 2013.

[27] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV*, volume 1729 of *Lecture Notes in Math.*, pages 1–145. Springer, Berlin, 2000.

[28] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear PDEs. *Math. Comp.*, 77(261):125–158 (electronic), 2008.

[29] W. E, M. Hutzenthaler, A. Jentzen, and Th. Kruse. On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. *J. Sci. Comput.*, 79(3):1534–1571, 2019.

[30] G. Fabbri, F. Gozzi, and A. Świech. *Stochastic optimal control in infinite dimension*, volume 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017. Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore.

[31] A. Fahim, N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic PDEs. *Ann. Appl. Probab.*, 21(4):1322–1364, 2011.

[32] P.A. Forsyth and K. R. Vetzal. Numerical methods for nonlinear pdes in finance. In *Handbook of Computational Finance*, pages 503–528. Springer, 2012.

[33] E. Gobet, J-P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005.

[34] U. G. Haussmann and É. Pardoux. Time reversal of diffusions. *Ann. Probab.*, 14(4):1188–1205, 1986.

[35] P. Henry-Labordère, N. Oudjane, X. Tan, N. Touzi, and X. Warin. Branching diffusion representation of semilinear pdes and Monte Carlo approximations. *Available at http://arxiv.org/pdf/1603.01727v1.pdf*, 2016.

[36] C. Huret, H. Pham, and X. Warin. Some machine learning schemes for high-dimensional nonlinear pdes, 2019.

[37] L. Izydorczyk, N. Oudjane, F. Russo, and M. Tessitore. Probabilistic representations of PDEs with terminal condition and applications to time reversal of diffusions. *In preparation*, 2020.

[38] J-F Jabir, D. Talay, and M. Tomašević. Mean-field limit of a particle approximation of the one-dimensional parabolic-parabolic Keller-Segel model without smoothing. *Electron. Commun. Probab.*, 23:Paper No. 84, 14, 2018.

[39] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(6):727–766, 1998.

[40] B. Jourdain, S. Méléard, and W. A. Woyczynski. Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:1–29, 2008.

[41] M. Kac. *Probability and related topics in physical sciences*, volume 1957 of *With special lectures by G. E. Uhlenbeck, A. R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, Colo.* Interscience Publishers, London-New York, 1959.

[42] J. P. Keener and J. Sneyd. *Mathematical Physiology II: Systems Physiology*. Springer, New York, 2008.

[43] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1992.

[44] R. Lattès and J.-L. Lions. The method of quasi-reversibility: applications to partial differential equations. Technical report, 1969.

[45] A. Le Cavil, N. Oudjane, and F. Russo. Particle system algorithm and chaos propagation related to a non-conservative McKean type stochastic differential equations. *Stochastics and Partial Differential Equations: Analysis and Computation.*, pages 1–37, 2016.

[46] A. Le Cavil, N. Oudjane, and F. Russo. Probabilistic representation of a class of non-conservative nonlinear partial differential equations. *ALEA Lat. Am. J. Probab. Math. Stat*, 13(2):1189–1233, 2016.

[47] A. Le Cavil, N. Oudjane, and F. Russo. Monte-Carlo algorithms for a forward Feynman–Kac-type representation for semilinear nonconservative partial differential equations. *Monte Carlo Methods Appl.*, 24(1):55–70, 2018.

[48] A. Le Cavil, N. Oudjane, and F. Russo. Forward Feynman-Kac type representation for semilinear nonconservative partial differential equations. *Stochastics: an International Journal of Probability and Stochastic Processes*, to appear. First version 2016, Preprint hal-01353757.

[49] J. Lieber, N. Oudjane, and F. Russo. On the well-posedness of a class of McKean Feynman-Kac equations. *Markov Processes and Related Fields. Arxiv 1810.10205*, to appear.

[50] F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: a simple least-squares approach. *Review of Fin. Stud.*, 14:113–147, 2001.

[51] H. P. Jr. McKean. A class of Markov processes associated with nonlinear parabolic equations. In *Proc. Nat. Acad. Sci. U.S.A., 1966)*, pages 1907–1911. 1966.

[52] H. P. Jr. McKean. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.

[53] H. P. Jr. McKean. Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Comm. Pure Appl. Math.*, 28(3):323–331, 1975.

[54] S. Méléard and S. Roelly-Coppoletta. A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.*, 26(2):317–332, 1987.

[55] J. D. Murray. *Mathematical biology. I*, volume 17 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. An introduction.

[56] K. Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984.

[57] G. Pagès and A. Sagna. Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering. *Stochastic Process. Appl.*, 128(3):847–883, 2018.

[58] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lect. Notes Control Inf. Sci.*, pages 200–217. Springer, Berlin, 1992.

[59] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.

[60] M. Renardy, W. J. Hrusa, and J. A. Nohel. Mathematical problems in viscoelasticity. *New York*, 1987.

[61] M. Röckner and F. Russo. Uniqueness for a class of stochastic Fokker-Planck and porous media equations. *J. Evol. Equ.*, 17(3):1049–1062, 2017.

[62] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[63] A-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[64] X. Tan. Probabilistic numerical approximation for stochastic control problems. *preprint*, 2012.

[65] A. N. Tikhonov and V. I. Arsenin. *Solutions of ill-posed problems*, volume 14. Vh Winston, 1977.

[66] N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2013. With Chapter 13 by Agnès Tourin.

[67] J. N. Tsitsiklis and B. Van Roy. Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks*, 12(4):694–703, 2001.

[68] X. Y. Wang, Z. S. Zhu, and Y. K. Lu. Solitary wave solutions of the generalized Burgers-Huxley equation. *J. Phys. A: Math. Gen.*, 23:271–274, 1990.