# Forward Feynman-Kac type representation for semilinear nonconservative Partial Differential Equations

ANTHONY LECAVIL *, NADIA OUDJANE † AND FRANCESCO RUSSO ‡

August 12th 2016

**Abstract**

We propose a nonlinear forward Feynman-Kac type equation, which represents the solution of a non-conservative semilinear parabolic Partial Differential Equations (PDE). We show in particular existence and uniqueness in the first part of the article. The second part is devoted to the construction of a probabilistic particle algorithm and the proof of its convergence. Illustrations of the efficiency of the algorithm are provided by numerical experiments.

**Key words and phrases:** Semilinear Partial Differential Equations; Nonlinear Feynman-Kac type functional; Particle systems; Probabilistic representation of PDEs.

**2010 AMS-classification**: 60H10; 60H30; 60J60; 65C05; 65C35; 68U20; 35K58.

## 1 Introduction

In this paper, we are interested in proposing a forward probabilistic representation of the following semilinear Partial Differential Equation (PDE) on $[0,T]\times\mathbb{R}^d$

$$\left\{\begin{array}{l}\partial_t u = L_t^* u + u\Lambda(t,x,u,\nabla u)\\ u(0,\cdot)=u_0\ ,\end{array}\right. \tag{1.1}$$

where $u_0$ is a Borel probability measure on $\mathbb{R}^d$ and $L^*$ is a partial differential operator of the type

$$(L_t^*\varphi)(x)=\frac{1}{2}\sum_{i,j=1}^d \partial_{ij}^2(a_{i,j}(t,x)\varphi)(x)-\sum_{i=1}^d\partial_i(g_i(t,x)\varphi)(x),\qquad \text{for } \varphi\in\mathcal{C}_0^\infty(\mathbb{R}^d). \tag{1.2}$$

When $\Lambda=0$, equation (1.1) is reduced to the classical Fokker-Planck PDE, that has been extensively studied and for which very general existence/uniqueness results have been proved, see e.g. [3, 1]. In this

*ENSTA-ParisTech, Université Paris-Saclay. Unité de Mathématiques Appliquées (UMA). E-mail: anthony.lecavil@ensta-paristech.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-ParisTech, Université Paris-Saclay. Unité de Mathématiques Appliquées (UMA). E-mail:francesco.russo@ensta-paristech.fr

specific case, a forward probabilistic representation of (1.1) is related to the solution $Y$ of the Stochastic Differential Equation (SDE) associated with the infinitesimal generator $L$ and the initial condition $u_0$, i.e.

$$\begin{cases} Y_t = Y_0 + \int_0^t \Phi(s, Y_s) dW_s + \int_0^t g(s, Y_s) ds \\ Y_0 \sim u_0 \ , \end{cases} \tag{1.3}$$

with $\Phi\Phi^t = a$. More precisely, if (1.3) admits a solution $Y$, then the marginal laws $(u_t(dx), t \geq 0)$ of $(Y_t, t \geq 0)$ satisfy the Fokker-Planck (also called forward Kolmogorov equation) which corresponds to PDE (1.1) when $\Lambda = 0$. In this sense, the couple $(Y, u)$ is a (forward) probabilistic representation of (1.1).
In the case where $\Lambda \neq 0$, we propose a representation which is constituted by a couple $(Y, u)$, solution of the system

$$\begin{cases} Y_t = Y_0 + \int_0^t \Phi(s, Y_s) dW_s + \int_0^t g(s, Y_s) ds, \quad Y_0 \sim u_0 \\ \int_{\mathbb{R}^d} \varphi(x) u(t,x) dx = \mathbb{E}\Big[\varphi(Y_t) \exp\Big(\int_0^t \Lambda(s, Y_s, u(s, Y_s), \nabla u(s, Y_s))\Big)\Big], \quad \text{for } t \in (0, T]\,,\ \varphi \in \mathcal{C}_b(\mathbb{R}^d)\ . \end{cases} \tag{1.4}$$

The main starting point of the paper is the following. *If $(Y, u)$ is a solution of* (1.4), *then $u$ solves* (1.1) *in the sense of distributions.* This follows by a direct application of Itô formula and integration by parts.
A function $u$ solving the second line of (1.4) will be often identified as ***Feynman-Kac type representation*** of (1.1). We emphasize that a solution to equation (1.4) introduced here is a couple $(Y, u)$, where $Y$ is a process solving a ***classical SDE***, and $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ satisfies the second line equation of (1.4).

One could theoretically consider a variant of (1.4) where the functions $\Phi, g$ depend on $u$ so that the first equation of the system is a ***McKean type SDE***, but its investigation goes beyond the scope of the present paper. Indeed, a solution $(Y, u)$ of that McKean system would be a solution of a more general non-linear equation than (1.1). In [15] and [14] we have fully analyzed a regularized version of that McKean type system, where $\Phi, g$ together with $\Lambda$ also depend on the unknown function $u$, but no dependence on $\nabla u$ was considered at that level. The first paper focuses on various results on existence and uniqueness, the second one on numerical approximation schemes. Even though, the present paper does not consider any McKean type non linearity, it extends the class of nonlinearities considered with respect to (w.r.t.) $\nabla u$. Indeed, the dependence of $\Lambda$ appears to be more singular than in [15], since it involves not only $u$ but also $\nabla u$ allowing to cover a different class of semilinear PDEs of the form (1.1). Besides, the probabilistic representation (1.4) will be directly related to the target PDE (1.1), whereas [15] provides a probabilistic representation related to an integro-differential equation corresponding to a regularization of the target PDE, without any theoretical result ensuring the convergence of the regularization to the target PDE.

In addition to the theoretical interest of the probabilistic representations in general, associated numerical schemes have also been the object of extensive developments in the literature. One major approach which has been largely investigated for approximating solutions of time evolutionary PDEs is the forward-backward SDEs (FBSDEs) method. FBSDEs were initially developed in [17], see also [16] for a survey and [18] 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 [4, 8]. 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. The FBSDE representation of fully nonlinear PDEs still requires complex developments and is the subject of active research, see for instance [5]. Branching diffusion processes provide alternative probabilistic representation of semilinear PDEs, involving a specific form of non-linearity on the zero order term. This type of approach

has been extended in [9, 11] to a more general class of non-linearities on the zero order term, with the so-called *marked branching process*. One of the main advantage of this approach compared to FBSDEs is that it does not involve any regression computation to calculate conditional expectations. More recently, an extension of the branching diffusion representation to a class of semilinear PDE has been proposed in [10]. The main idea of the present paper is to investigate the forward Feynman-Kac type representation (1.4) with an extended class of nonlinearities considered through the weighting function $\Lambda$, depending both on $u$ and $\nabla u$. 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 $\Phi$, $g$ and $\Lambda$ which may depend non-linearly not only on $u$, $\nabla u$ but also on $\nabla^2 u$. This more general setting could be investigated in a future work.
As we have mentioned, the focus of the paper is on (1.1) and on (1.4) which constitutes its natural probabilistic counterpart. The natural interpretation for (1.1) is the sense of distributions (weak solution), see Definition 2.1 1. which is equivalent to the mild sense which is formulated with the help of semigroups, see Definition 2.1 2. provided there is a unique solution when $\Lambda = 0$, see Lemma 2.2. Under Lipschitz conditions on the second variable on $\Phi, g$, Theorem 3.5 makes the bridge between analysis and probability showing that (1.4) admits a solution if and only if (1.1) admits a solution. Under Assumption 2 Proposition 3.6 shows existence and uniqueness of the mild solution in $L^1([0,T], W^{1,1}(\mathbb{R}^d))$. When $\Lambda$ does not depend on the gradient Theorem 3.9 shows well-posedness in the larger space $L^1([0,T], L^1(\mathbb{R}^d))$. Summarizing, the first contribution of the paper is the well-posedness of the non-linear diffusion equation (1.4), which is naturally associated with (1.1).
As a second main contribution we propose and analyze an original Monte Carlo scheme (5.31) to approximate the solution of (1.4) and consequently also of the solution $u$ of (1.1) which constitutes an equivalent (deterministic) form. This numerical scheme relies on three approximation steps: a regularization procedure based on a kernel convolution, a space discretization based on Monte Carlo simulations of the diffusion $Y$ (1.4) and a time discretization. Each resulting error is analyzed properly and their combination allows to establish the convergence of the numerical scheme in Theorem 5.6. Section 4 analyzes the convergence and its approximation rate in $L^1$ of the solution $u$ of (1.1) via a sequence of approximation $u^\varepsilon$ which are associated to the solution of a smoothed form of (1.1) i.e. (4.2). This is the object of Theorem 4.5 and Corollary 4.7. In Section 5, we present our original particle approximation scheme whose convergence is established in Proposition 5.2 and 5.5. Section 6 is finally devoted to numerical simulations.

# 2 Preliminaries

## 2.1 Notations

Let $d \in \mathbb{N}^\star$. Let us consider $\mathcal{C}^d := \mathcal{C}([0,T], \mathbb{R}^d)$ metricized by the supremum norm $\|\cdot\|_\infty$, equipped with its Borel $\sigma-$ field $\mathcal{B}(\mathcal{C}^d)$ and endowed with the topology of uniform convergence.
Let $(E, d_E)$. $\mathcal{P}(E)$ denotes the Polish space (with respect to the weak convergence topology) of Borel probability measures on $E$ naturally equipped with its Borel $\sigma$-field $\mathcal{B}(\mathcal{P}(E))$. The reader can consult Proposition 7.20 and Proposition 7.23, Section 7.4 Chapter 7 in [2] for more exhaustive information. When $d = 1$, we often omit it and we simply note $\mathcal{C} := \mathcal{C}^1$. $\mathcal{C}_b(E)$ denotes the space of bounded, continuous real-valued functions on $E$.

In this paper, $\mathbb{R}^d$ is equipped with the Euclidean scalar product $\cdot$ and $|x|$ stands for the induced norm for $x \in \mathbb{R}^d$. The gradient operator (w.r.t. $x \in \mathbb{R}^d$) for differentiable functions defined on $\mathbb{R}^d$ is denoted by $\nabla_x$. If there is no confusion, $\nabla$ will simply denote the gradient on $\mathbb{R}^d$. $M_{d,p}(\mathbb{R})$ denotes the space of $\mathbb{R}^{d\times p}$ real

matrices equipped with the Frobenius norm (also denoted $|\cdot|$), i.e. the one induced by the scalar product $(A, B) \in M_{d,p}(\mathbb{R}^d) \times M_{d,p}(\mathbb{R}) \mapsto Tr(A^t B)$ where $A^t$ stands for the transpose matrix of $A$ and $Tr$ is the trace operator. $\mathcal{S}_d$ is the set of symmetric, non-negative definite $d \times d$ real matrices and $\mathcal{S}_d^+$ the set of strictly positive definite matrices of $\mathcal{S}_d$.
$\mathcal{M}_f(\mathbb{R}^d)$ is the space of finite Borel measures on $\mathbb{R}^d$. When it is endowed with the weak convergence topology, $\mathcal{B}(\mathcal{M}_f(\mathbb{R}^d))$ stands for its Borel $\sigma$-field. It is well-known that $(\mathcal{M}_f(\mathbb{R}^d), \|\cdot\|_{TV})$ is a Banach space, where $\|\cdot\|_{TV}$ denotes the total variation norm.
$\mathcal{S}(\mathbb{R}^d)$ 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 $\mathbb{R}^d$ and $\mathcal{C}_0^\infty(\mathbb{R}^d)$ the space of smooth functions with compact support. For any positive integers $p, k \in \mathbb{N}$, $C_b^{k,p} := C_b^{k,p}([0,T] \times \mathbb{R}^d, \mathbb{R})$ denotes the set of continuously differentiable bounded functions $[0,T] \times \mathbb{R}^d \to \mathbb{R}$ with uniformly bounded derivatives with respect to the time variable $t$ (resp. with respect to space variable $x$) up to order $k$ (resp. up to order $p$). In particular, for $k = p = 0$, $C_b^{0,0}$ coincides with the space of bounded, continuous functions also denoted by $\mathcal{C}_b$. $\mathcal{C}_b^\infty(\mathbb{R}^d)$ is the space of bounded and smooth functions. $\mathcal{C}_0(\mathbb{R}^d)$ denotes the space of continuous functions with compact support in $\mathbb{R}^d$. For $r \in \mathbb{N}$, $W^{r,p}(\mathbb{R}^d)$ is the Sobolev space of order $r$ in $(L^p(\mathbb{R}^d), ||\cdot||_p)$, with $1 \leq p \leq \infty$. $W^{1,1}_{loc}(\mathbb{R}^d)$ denotes the space of functions $f : \mathbb{R}^d \to \mathbb{R}$ such that $f$ and $\nabla f$ (existing in the weak sense) belong to $L^1_{loc}(\mathbb{R}^d)$.
For convenience we introduce the following notation.

- $V : [0,T] \times \mathcal{C}^d \times \mathcal{C} \times \mathcal{C}^d$ is defined for any functions $x \in \mathcal{C}^d$, $y \in \mathcal{C}$ and $z \in \mathcal{C}^d$, by

$$V_t(x, y, z) := \exp\left(\int_0^t \Lambda(s, x_s, y_s, z_s) ds\right) \quad \text{for any } t \in [0,T] \, . \tag{2.1}$$

The finite increments theorem gives, for all $(a, b) \in \mathbb{R}^2$,

$$\exp(a) - \exp(b) = (b - a) \int_0^1 \exp(\alpha a + (1-\alpha) b) d\alpha \, . \tag{2.2}$$

In particular, if $\Lambda$ is supposed to be bounded and Lipschitz w.r.t. to its space variables $(x, y, z)$, uniformly w.r.t. $t$, we observe that (2.2) implies for all $t \in [0,T]$, $x, x' \in \mathcal{C}^d$, $y, y' \in \mathcal{C}$, $z, z' \in \mathcal{C}^d$,

$$|V_t(x, y, z) - V_t(x', y', z')| \leq L_\Lambda e^{tM_\Lambda} \int_0^t \big(|x_s - x'_s| + |y_s - y'_s| + |z_s - z'_s|\big) ds \, , \tag{2.3}$$

$M_\Lambda$ (resp. $L_\Lambda$) denoting an upper bound of $|\Lambda|$ (resp. the Lipschitz constant of $\Lambda$), see also Assumption 2.
In the whole paper, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ will denote a filtered probability space and $W$ an $\mathbb{R}^p$-valued $(\mathcal{F}_t)$-Brownian motion.

## 2.2 Mild and Weak solutions

We first introduce the following assumption.

**Assumption 1.** *1. $\Phi$ and $g$ are functions defined on $[0,T] \times \mathbb{R}^d$ taking values in $M_{d,p}(\mathbb{R}^d)$ and $\mathbb{R}^d$. There exist $L_\Phi, L_g > 0$ such that for any $t \in [0,T]$, $(x, x') \in \mathbb{R}^d \times \mathbb{R}^d$,*

$$|\Phi(t, x) - \Phi(t, x')| \leq L_\Phi |x - x'| \, , \tag{2.4}$$

*and*

$$|g(t, x) - g(t, x')| \leq L_g |x - x'| \, . \tag{2.5}$$

2. *The functions* $s \in [0,T] \mapsto |\Phi(s,0)|$ *and* $s \in [0,T] \mapsto |g(s,0)|$ *are bounded.*

In the whole paper we will write $a = \Phi\Phi^t$; in particular $a : [0,T] \times \mathbb{R}^d \longrightarrow \mathcal{S}_d$. Through some definitions, we make here precise in which sense we will consider solutions of the PDE (1.1). We are interested in different concepts of solutions $u : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ of that semilinear PDE where, for $t \in [0,T]$, $L_t$ is given by

$$(L_t\varphi)(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x)\partial_{ij}^2\varphi(x) + \sum_{i=1}^{d} g_i(t,x)\partial_i\varphi(x), \ \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d). \tag{2.6}$$

Its "adjoint" $L_t^*$ defined in (1.2), verifies

$$\int_{\mathbb{R}^d} L_t\varphi(x)\psi(x)dx = \int_{\mathbb{R}^d} \varphi(x)L_t^*\psi(x)dx \ , \ (\varphi,\psi) \in \mathcal{C}_0^\infty(\mathbb{R}^d), t \in [0,T] \ , \tag{2.7}$$

and the initial condition $u_0$ of (1.1) has to be understood in the sense that

$$\lim_{t\to 0}\int_{\mathbb{R}^d} \varphi(x)u_t(dx) = \int_{\mathbb{R}^d} \varphi(x)u_0(dx) \ , \text{ for all } \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d) \ ,$$

since, a priori, it can be irregular and not necessarily a function.
We observe that if $\Lambda = 0$, (1.1) is the classical Fokker-Planck equation which can be understood in the sense of distributions, i.e., for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$.

$$\int_{\mathbb{R}^d} u(t,dx)\varphi(x) = \int_{\mathbb{R}^d} u_0(dx)\varphi(x) + \int_0^t\int_{\mathbb{R}^d} u(s,dx)(L_s\varphi)(x)ds \ . \tag{2.8}$$

Under Assumption 1 it is well-known (see Introduction and Section 2.2, Chapter 2 in [19]), that there exists a *good* family of probability transition $P(s,x_0,t,\cdot)$ (see Introduction and Section 2.2, Chapter 2 in [19]), for which the Fokker-Planck equation (understood in the sense of distributions) is verified, i.e.

$$\begin{cases} \partial_t P(s,x_0,t,\cdot) = L_t^* P(s,x_0,t,\cdot) \\ \lim_{t\downarrow s} P(s,x_0,t,\cdot) = \delta_{x_0}, \quad 0 \le s < t \le T, x_0 \in \mathbb{R}^d \ . \end{cases} \tag{2.9}$$

Given a random variable $Y_0$, classical theorems for SDE with Lipschitz coefficients imply strong existence and pathwise uniqueness for the SDE

$$dY_t = \Phi(t,Y_t)dW_t + g(t,Y_t)dt. \tag{2.10}$$

By the classical theory of Markov processes (see e.g. Chapter 2 in [19]), we know that the transition probability function $P$, satisfying (2.9), defines and characterizes uniquely the law of the process $Y$, provided the law $u_0$ of $Y_0$ is specified. In particular, we have the following.

1. For $t \in [0,T]$, the marginal law of $Y_t$ is given for all $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ by

$$\mathbb{E}[\varphi(Y_t)] \quad = \quad \int_{\mathbb{R}^d} u_0(dx_0)\int_{\mathbb{R}^d} \varphi(x)P(0,x_0,t,dx) \ . \tag{2.11}$$

2. For all $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ and $0 \le s < t \le T$,

$$\mathbb{E}[\varphi(Y_t)|Y_s] = \int_{\mathbb{R}^d} \varphi(x)P(s,Y_s,t,dx). \tag{2.12}$$

Let $\Lambda : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be bounded, Borel measurable, we recall the notions of **weak solution** and **mild solution** associated to (1.1).

**Definition 2.1.** *Let $u : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be a Borel function such that for every $t \in ]0,T]$, $u(t,\cdot) \in W^{1,1}_{\rm loc}(\mathbb{R}^d)$.*

1. *$u$ will be called* **weak solution** *of* (1.1) *if for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$,*

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)u(t,x)dx - \int_{\mathbb{R}^d} \varphi(x)u_0(dx) &= \int_0^t \int_{\mathbb{R}^d} u(s,x)L_s\varphi(x)dxds \\ &+ \int_0^t \int_{\mathbb{R}^d} \varphi(x)\Lambda(s,x,u(s,x),\nabla u(s,x))u(s,x)dxds \,.\end{aligned} \tag{2.13}$$

2. *$u$ will be called* **mild solution** *of* (1.1) *if for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$,*

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)u(t,x)dx &= \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} u_0(dx_0)P(0,x_0,t,dx) \\ &+ \int_{[0,t]\times\mathbb{R}^d} \Big( \int_{\mathbb{R}^d} \varphi(x)P(s,x_0,t,dx)\Big)\Lambda(s,x_0,u(s,x_0),\nabla u(s,x_0))u(s,x_0)dx_0ds \,.\end{aligned} \tag{2.14}$$

The object of the first lemma below is to show to what extent the concept of mild solution is equivalent to the weak one.

**Lemma 2.2.** *We assume there exists a unique weak solution $v$ of*

$$\begin{cases} \partial_t v = L_t^* v \\ v(0,\cdot) = 0 \,, \end{cases} \tag{2.15}$$

*where $L_t^*$ is given by* (1.2).
*Then, $u$ is a mild solution of* (1.1) *if and only if $u$ is a weak solution of* (1.1)

**Remark 2.3.** *There exist several sets of technical assumptions (see e.g. [3, 7]) leading to the uniqueness assumed in Lemma 2.2 above. In particular, under items 1., 2. and 3. of Assumption 2 stated below (which will constitutes our framework in the sequel), Theorem 4.7 in Chapter 4 of [7] ensures (classical) existence and uniqueness of the solution of* (2.15)*, see also Lemma 7.4 below.*

*Proof.* We first suppose that $u$ is a mild solution of (1.1). Taking into account that $P(s,x_0,t,\cdot)$ is a distributional solution of (2.9), classical computations show that $u$ is indeed a weak solution of (1.1).
Conversely, suppose that $u$ is a weak solution of (1.1), in the sense of Definition 2.1. We also consider

$$\begin{aligned}\bar{v}(t,dx) &:= \int_{\mathbb{R}} P(0,x_0,t,dx)u_0(dx_0) + \int_0^t ds \int_{\mathbb{R}^d} P(s,x_0,t,dx)u(s,dx_0)\Lambda(s,x_0,u(s,x_0),\nabla u(s,x_0))dx_0 \\ &= \int_{\mathbb{R}} P(0,x_0,t,dx)u_0(dx_0) + \int_0^t ds \int_{\mathbb{R}^d} P(s,x_0,t,dx)\bar{\Lambda}(u)(s,x_0)dx_0 \,,\end{aligned} \tag{2.16}$$

where $\bar{\Lambda}(u)(t,x) := u(s,x)\Lambda(s,x,u(s,x),\nabla u(s,x))$ for $(s,x) \in [0,T]\times\mathbb{R}^d$. We want to ensure that $u = \bar{v}$.
On the one hand, integrating the function $\bar{v}(t,\cdot)$ against a test function and using again that $P(s,x_0,t,\cdot)$ is a distributional solution of (2.9), we obtain that $\bar{v}$ is a weak solution of

$$\begin{cases} \partial_t \bar{v} = L_t^* \bar{v} + \bar{\Lambda}(u) \\ \bar{v}(0,\cdot) = u_0 \,, \end{cases} \tag{2.17}$$

On the other hand, $u$ being a weak solution of (1.1), it also satisfies (2.17) (in the sense of distributions). We set $v := \bar{v} - u$. It follows that $v$ and $\hat{v} := 0$ both satisfy (2.15). Uniqueness of the solution of (2.15) implies that $v = 0$, which concludes the proof. □

# 3 Feynman-Kac type representation

We suppose here the validity of Assumption 1. Let $u_0 \in \mathcal{P}(\mathbb{R}^d)$ and fix a random variable $Y_0$ distributed according to $u_0$ and consider the strong solution $Y$ of (2.10).
The aim of this section is to show how a mild solution of (1.1) can be linked with a Feynman-Kac type equation, where we recall that a solution is given by a function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ satisfying the second line equation of (1.4).

Given $\tilde{\Lambda} : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ a bounded, Borel measurable function, let us consider the measure-valued map $\mu : [0,T] \longrightarrow \mathcal{M}_f(\mathbb{R}^d)$ defined by

$$\int_{\mathbb{R}^d} \varphi(x)\mu(t,dx) = \mathbb{E}\Big[\varphi(Y_t)\exp\Big(\int_0^t \tilde{\Lambda}(s,Y_s)ds\Big)\Big], \text{ for all } \varphi \in \mathcal{C}_b(\mathbb{R}^d), t \in [0,T]\,. \tag{3.1}$$

The first proposition below shows how the map $t \mapsto \mu(t,\cdot)$ can be characterized as a solution of the linear parabolic PDE

$$\begin{cases} \partial_t v = L_t^* v + \tilde{\Lambda}(t,x)v \\ v(0,\cdot) = u_0\,. \end{cases} \tag{3.2}$$

Before stating the corresponding proposition, we introduce the notion of *measure-mild solution*.

**Definition 3.1.** *A measure-valued map $\mu : [0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ will be called* **measure-mild solution** *of* (3.2) *if for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$,*

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(x)\mu(t,dx) &= \int_{\mathbb{R}^d} \varphi(x) \int_{\mathbb{R}^d} u_0(dx_0) P(0,x_0,t,dx) \\ &+ \int_{[0,t]\times\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} P(r,x_0,t,dx)\varphi(x)\Big)\tilde{\Lambda}(r,x_0)\mu(r,dx_0)dr. \end{aligned} \tag{3.3}$$

**Remark 3.2.**

1. *Since $\mu$ is a (finite) measure valued function, by usual approximation arguments, it is not difficult to show that an equivalent formulation for Definition 2.1 can be expressed taking $\varphi$ in $\mathcal{C}_b(\mathbb{R}^d)$ instead of $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$.*

2. *Although the definition of* **mild solution** *(see item 2. of Definition of 2.1) and the one of* **measure-mild solution** *seem to be formally close, the two concepts do not make sense in the same situations. Indeed, the notion of mild-solution makes sense for PDEs with nonlinear terms of the general form $\Lambda(t,x,u,\nabla u)$, whereas a measure-mild solution can exist only for linear PDEs. However, in the case where a measure $\mu$ on $\mathbb{R}^d$, absolutely continuous w.r.t. the Lebesgue measure $dx$, is a measure-mild solution of the linear PDE* (3.2)*, its density indeed coincides with the mild solution (in the sense of item 2. of Definition 2.1) of* (3.2)*.*

**Proposition 3.3.** *Under Assumption 1 the measure-valued map $\mu$ defined by* (3.1) *is the unique measure-mild solution of*

$$\begin{cases} \partial_t v = L_t^* v + \tilde{\Lambda}(t,x)v \\ v(0,\cdot) = u_0\,, \end{cases} \tag{3.4}$$

*where the operator $L_t^*$ is defined by* (1.2).

*Proof.* We first prove that a function $\mu$ defined by (3.1) is a measure-mild solution of (3.4).
Observe that for all $t \in [0,T]$,

$$\exp\Big(\int_0^t \tilde{\Lambda}(r,Y_r)dr\Big) = 1 + \int_0^t \tilde{\Lambda}(r,Y_r)e^{\int_0^r \tilde{\Lambda}(s,Y_s)ds}dr\,. \tag{3.5}$$

From (3.1), it follows that for all test function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ and $t \in [0,T]$,

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)\mu(t,dx) &= \mathbb{E}\Big[\varphi(Y_t)\exp\Big(\int_0^t \tilde{\Lambda}(r,Y_r)dr\Big)\Big] \\ &= \mathbb{E}[\varphi(Y_t)] + \int_0^t \mathbb{E}\Big[\varphi(Y_t)\tilde{\Lambda}(r,Y_r)e^{\int_0^r \tilde{\Lambda}(s,Y_s)ds}\Big]dr \; . \end{aligned} \tag{3.6}$$

On the one hand, by (2.11), we have

$$\mathbb{E}[\varphi(Y_t)] = \int_{\mathbb{R}^d} u_0(dx_0)\int_{\mathbb{R}^d} \varphi(x)P(0,x_0,t,dx) \; , \quad \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d) \text{ and } t \in [0,T] \; . \tag{3.7}$$

On the other hand, using (2.12) yields, for $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $0 \le r \le t$,

$$\begin{aligned}\mathbb{E}\Big[\varphi(Y_t)\tilde{\Lambda}(r,Y_r)e^{\int_0^r \tilde{\Lambda}(s,Y_s)ds}\Big] &= \mathbb{E}\Big[\tilde{\Lambda}(r,Y_r)e^{\int_0^r \tilde{\Lambda}(s,Y_s)ds}\mathbb{E}\Big[\varphi(Y_t)\Big|Y_r\Big]\Big] \\ &= \mathbb{E}\Big[\Big(\tilde{\Lambda}(r,Y_r)\int_{\mathbb{R}^d}\varphi(x)P(r,Y_r,t,dx)\Big)e^{\int_0^r \tilde{\Lambda}(s,Y_s)ds}\Big] \\ &= \int_{\mathbb{R}^d}\Big(\tilde{\Lambda}(r,x_0)\int_{\mathbb{R}^d}\varphi(x)P(r,x_0,t,dx)\Big)\mu(r,dx_0) \; , \end{aligned} \tag{3.8}$$

where the third equality above comes from (3.1) applied to the bounded, measurable test function $z \mapsto \tilde{\Lambda}(r,z)\int_{\mathbb{R}^d}\varphi(x)P(r,z,t,dx)$. Injecting (3.8) and (3.7) in the right-hand side (r.h.s.) of (3.6) gives for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t \in [0,T]$,

$$\begin{aligned}\int_{\mathbb{R}^d}\varphi(x)\mu(t,dx) &= \int_{\mathbb{R}^d} u_0(dx_0)\int_{\mathbb{R}^d} P(0,x_0,t,dx)\varphi(x)dx \\ &+ \int_0^t\int_{\mathbb{R}^d}\mu(r,dx_0)\tilde{\Lambda}(r,x_0)\int_{\mathbb{R}^d}P(r,x_0,t,dx)\varphi(x)\; dr \; . \end{aligned} \tag{3.9}$$

It remains now to prove uniqueness of the measure-mild solution of (3.4).
We recall that $\mathcal{M}_f(\mathbb{R}^d)$ denotes the vector space of finite Borel measures on $\mathbb{R}^d$, that is here equipped with the total variation norm $\|\cdot\|_{TV}$. We also recall that an equivalent definition of the total variation norm is given by

$$\|\mu\|_{TV} = \sup_{\substack{\psi\in\mathcal{C}_b(\mathbb{R}^d)\\ \|\psi\|_\infty\le 1}}\Big|\int_{\mathbb{R}^d}\psi(x)\mu(dx)\Big| \; . \tag{3.10}$$

Consider $t \in [0,T]$ and let $\mu_1$, $\mu_2$ be two measure-mild solutions of PDE (3.4). We set $\nu := \mu_1 - \mu_2$. Since $\tilde{\Lambda}$ is bounded, we observe that (3.1) implies $\|\nu(t,\cdot)\|_{TV} < +\infty$. Moreover, taking into account item 1. of Remark 3.2, we have that $\nu$ satisfies,

$$\forall\, \varphi \in \mathcal{C}_b(\mathbb{R}^d),\; \int_{\mathbb{R}^d}\varphi(x)\nu(t,dx) = \int_0^t\int_{\mathbb{R}^d}\tilde{\Lambda}(r,x_0)\nu(r,dx_0)\int_{\mathbb{R}^d}\varphi(x)P(r,x_0,t,dx)\; dr \; . \tag{3.11}$$

Taking the supremum over $\varphi$ s.th. $\|\varphi\|_\infty \le 1$ in each side of (3.11), we get

$$\|\nu(t,\cdot)\|_{TV} \quad\le \sup_{(s,x)\in[0,T]\times\mathbb{R}^d}|\tilde{\Lambda}(s,x)|\int_0^t\|\nu(r,\cdot)\|_{TV}\; dr \; . \tag{3.12}$$

Gronwall's lemma implies that $\nu(t,\cdot) = 0$. Uniqueness of measure-mild solution for (3.4) follows. This ends the proof. □

The next lemma shows how a measure-mild solution of (3.4), which is a function defined on $[0,T]$ can be built by defining it recursively on each sub-interval of the form $[r, r+\tau]$. In particular, it will be used in Theorem 3.6 and Proposition 4.4. Its proof is postponed in Appendix (see Section 7.3).

**Lemma 3.4.** *Let us fix $\tau > 0$ be a real constant and $\delta := (\alpha_0 := 0 < \cdots < \alpha_k := k\tau < \cdots < \alpha_N := T)$ be a finite partition of $[0,T]$.*
*A measure-valued map $\mu : [0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ satisfies*

$$\begin{cases} \mu(0,\cdot) = u_0 \\ \mu(t,dx) = \int_{\mathbb{R}^d} P(k\tau, x_0, t, dx)\mu(k\tau, dx_0) + \int_{k\tau}^t ds \int_{\mathbb{R}^d} P(s, x_0, t, dx)\tilde{\Lambda}(s, x_0)\mu(s, dx_0) \ , \end{cases} \tag{3.13}$$

*for all $t \in [k\tau, (k+1)\tau]$ and $k \in \{0, \cdots, N-1\}$, if and only if $\mu$ is a measure-mild solution (in the sense of Definition 3.1) of* (3.4).

We now come back to the case where the bounded, Borel measurable real-valued function $\Lambda$ is defined on $[0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$. Let $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ belonging to $L^1([0,T], W^{1,1}(\mathbb{R}^d))$. In the sequel, we set $\tilde{\Lambda}^u(t,x) := \Lambda(t, x, u(t,x), \nabla u(t,x))$. $\mu^u$ will denote the measure-valued map $\mu$ defined by (3.1) with $\tilde{\Lambda} = \tilde{\Lambda}^u$, i.e.,

$$\int_{\mathbb{R}^d} \varphi(x)\mu^u(t,dx) = \mathbb{E}\Big[\varphi(Y_t)\exp\Big(\int_0^t \tilde{\Lambda}^u(s, Y_s)ds\Big)\Big], \text{ for all } \varphi \in \mathcal{C}_b(\mathbb{R}^d), t \in [0,T] \ . \tag{3.14}$$

By Proposition 3.3, it follows that $\mu^u$ is the unique measure-mild solution of the linear PDE (3.4) with $\tilde{\Lambda} = \tilde{\Lambda}^u$. (3.14) can be interpreted as a Feynman-Kac type representation for the measure-mild solution $\mu^u$ of the linear PDE (3.4), for the corresponding $\tilde{\Lambda}^u$. More generally, Theorem 3.5 below establishes such representation formula for a mild solution of the semilinear PDE (1.1).

**Theorem 3.5.** *Assume that Assumption 1 is fulfilled. We indicate by $Y$ the unique strong solution of* (2.10)*.*
*Suppose that $\Lambda : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is bounded and Borel measurable. A function $u : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ in $L^1([0,T], W^{1,1}(\mathbb{R}^d))$ is a mild solution of* (1.1) *if and only if, for all $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$, $t \in [0,T]$,*

$$\int_{\mathbb{R}^d} \varphi(x)u(t,x)dx = \mathbb{E}\Big[\varphi(Y_t)\exp\Big(\int_0^t \Lambda(s, Y_s, u(s,Y_s), \nabla u(s, Y_s))\Big)\Big] \ . \tag{3.15}$$

A function $u$ verifying (3.15) will be called a **Feynman-Kac type representation** of (1.1).

*Proof.* We first suppose that $u$ is a mild solution of (1.1). The aim is then to show that $u$ satisfies the Feynman-Kac equation (3.15).
Since $\Lambda$ is supposed to be bounded, Borel and $u$ is Borel, it is clear that

$$\tilde{\Lambda}^u(t,x) := \Lambda(t, x, u(t,x), \nabla u(t,x)) \ , \tag{3.16}$$

is bounded and Borel measurable. On the one hand, by applying Proposition 3.3 with $\tilde{\Lambda}^u$, it follows that $\mu^u$ (defined by (3.14)) is the unique measure-mild solution of

$$\begin{cases} \partial_t v = L_t^* v + \tilde{\Lambda}^u(t,x)v \\ v(0,\cdot) = u_0 \ . \end{cases} \tag{3.17}$$

On the other hand, since $u$ is supposed to be a mild solution of (1.1), $u(t,x)dx$ is also a measure-mild solution of (3.17). By uniqueness of the solution of (3.17), for all $t \in [0,T]$, we have $u(t,x)dx = \mu^u(t,dx)$.

Taking into account item 1. of Remark 3.2, it implies for all $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$, $t \in [0,T]$,

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x) u(t,x) dx &= \int_{\mathbb{R}^d} \varphi(x) \mu^u(t,dx) \\
&= \mathbb{E}\Big[\varphi(Y_t) \exp\Big(\int_0^t \tilde{\Lambda}^u(s,Y_s)\Big) ds\Big], \text{ by (3.14)} \\
&= \mathbb{E}\Big[\varphi(Y_t) \exp\Big(\int_0^t \Lambda(s,Y_s,u(s,Y_s),\nabla u(s,Y_s)) ds\Big)\Big], \text{ by (3.16)} .
\end{aligned}$$

Conversely, suppose that $u$ satisfies the Feynman-Kac equation (3.15). Recalling (3.16) and setting $\bar{\mu}(t,dx) := u(t,x)dx$ , (3.15) can be re-written

$$\int_{\mathbb{R}^d} \varphi(x) \bar{\mu}(t,dx) = \mathbb{E}\Big[\varphi(Y_t) \exp\Big(\int_0^t \tilde{\Lambda}^u(s,Y_s) ds\Big)\Big] , \quad \varphi \in \mathcal{C}_b(\mathbb{R}^d), t \in [0,T] . \tag{3.18}$$

Proposition 3.3 applied again with $\tilde{\Lambda}^u$ shows that $\bar{\mu}$ is the unique measure-mild solution of (3.17). In particular for every $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$,

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x) u(t,x) dx &= \int_{\mathbb{R}^d} \varphi(x) \bar{\mu}(t,dx) \\
&= \int_{\mathbb{R}^d} u_0(dx_0) \int_{\mathbb{R}^d} P(0,x_0,t,dx) \varphi(x) dx \\
&\quad + \int_0^t \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} P(r,x_0,t,dx)\varphi(x)\Big) \Lambda(r,x_0,u(r,x_0),\nabla u(r,x_0)) u(r,x_0) dx_0 \, dr.
\end{aligned} \tag{3.19}$$

This shows that $u$ is a mild solution of (1.1) and concludes the proof. □

We now precise more restrictive assumptions to ensure regularity properties of the transition probability function $P(s,x_0,t,dx)$ used in the sequel.

**Assumption 2.** 1. *$\Phi$ and $g$ are functions defined on $[0,T] \times \mathbb{R}^d$ taking values respectively in $M_{d,p}(\mathbb{R})$ and $\mathbb{R}^d$. There exist $\alpha \in ]0,1]$, $C_\alpha, L_\Phi, L_g > 0$, such that for any $(t,t',x,x') \in [0,T] \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^d$,*

$$|\Phi(t,x) - \Phi(t,x')| \leq C_\alpha |t-t'|^\alpha + L_\Phi |x-x'| ,$$

*and*

$$|g(t,x) - g(t',x')| \leq C_\alpha |t-t'|^\alpha + L_g |x-x'| .$$

2. *$\Phi$ and $g$ belong to $C_b^{0,3}$. In particular, $\Phi$, $g$ are uniformly bounded and $M_\Phi$ (resp. $M_g$) denote the upper bound of $|\Phi|$ (resp. $|g|$).*

3. *$\Phi$ is non-degenerate, i.e. there exists $c > 0$ such that for all $x \in \mathbb{R}^d$*

$$\inf_{s \in [0,T]} \inf_{v \in \mathbb{R}^d \setminus \{0\}} \frac{\langle v, \Phi(s,x)\Phi^t(s,x) v\rangle}{|v|^2} \geq c > 0. \tag{3.20}$$

4. *$\Lambda$ is a Borel real-valued function defined on $[0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ and Lipschitz uniformly w.r.t. $(t,x)$ i.e. there exists a finite positive real, $L_\Lambda$, such that for any $(t,x,z_1,z_1',z_2,z_2') \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^2 \times (\mathbb{R}^d)^2$, we have*

$$|\Lambda(t,x,z_1,z_2) - \Lambda(t,x,z_1',z_2')| \leq L_\Lambda (|z_1 - z_1'| + |z_2 - z_2'|) . \tag{3.21}$$

5. $\Lambda$ *is supposed to be uniformly bounded: let* $M_\Lambda$ *be an upper bound for* $|\Lambda|$.

6. $u_0$ *is a Borel probability measure on* $\mathbb{R}^d$ *admitting a bounded density (still denoted by the same letter) belonging to* $W^{1,1}(\mathbb{R}^d)$.

**Theorem 3.6.** *Under Assumption 2, there exists a unique mild solution* $u$ *of* (1.1) *in* $L^1([0,T], W^{1,1}(\mathbb{R}^d)) \cap L^\infty([0,T] \times \mathbb{R}^d, \mathbb{R})$.

The idea of the proof will be first to construct a unique "mild solution" $u^k$ of (1.1) on each subintervals of the form $[k\tau, (k+1)\tau]$ with $k \in \{0, \cdots, N-1\}$ and $\tau > 0$ a constant supposed to be fixed for the moment. This will be the object of Lemma 3.7. Secondly we will show that the function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, defined by being equal to $u^k$ on each $[k\tau, (k+1)\tau]$, is indeed a mild solution of (1.1) on $[0,T] \times \mathbb{R}^d$. This will be a consequence of Lemma 3.4. Finally, uniqueness will follow classically from Lipschitz property of $\Lambda$.

Let us fix $\phi \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. For $r \in [0, T-\tau]$, we first define a function $\widehat{u_0}$ on $[r, r+\tau] \times \mathbb{R}^d$ by setting

$$\widehat{u_0}(r,\phi)(t,x) := \int_{\mathbb{R}^d} p(r, x_0, t, x)\phi(x_0)dx_0, \ (t,x) \in [r, r+\tau] \times \mathbb{R}^d. \tag{3.22}$$

Consider now the map $\Pi : L^1([r, r+\tau], W^{1,1}(\mathbb{R}^d)) \to L^1([r, r+\tau], W^{1,1}(\mathbb{R}^d))$ given by

$$\Pi(v)(t,x) := \int_r^t ds \int_{\mathbb{R}^d} p(s, x_0, t, x)\Lambda(s, x_0, v + \widehat{u_0}(r,\phi), \nabla(v + \widehat{u_0}(r,\phi)))\big(v + \widehat{u_0}\big)(r,\phi)(s, x_0)dx_0 \ , \tag{3.23}$$

$$\Lambda(t, z, v, \nabla v) := \Lambda(t, z, v(t,z), \nabla v(t,z)) \quad \text{with} \quad (t,z) \in [0,T] \times \mathbb{R}^d \ , \tag{3.24}$$

that will also be used in the sequel.
Later, the dependence on $r, \phi$ will be omitted when it is self-explanatory. Since $\phi$ and $u_0$ belong to $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, we have

$$\|\widehat{u_0}(t,\cdot)\|_1 \leq \|\phi\|_1 \quad \text{and} \quad \|\widehat{u_0}(t,\cdot)\|_\infty \leq \|\phi\|_\infty, \text{ if } t \in [r, r+\tau] \ . \tag{3.25}$$

The lemma below establishes, under a suitable choice of $\tau > 0$, existence and uniqueness of the mild solution on $[r, r+\tau]$, with initial condition $\phi$ at time $r$, i.e. existence and uniqueness of the fixed-point for the application $\Pi$.

**Lemma 3.7.** *Assume the validity of Assumption 2. Let* $\phi \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.
*Let* $M > 0$ *such that* $M \geq \max(\|\phi\|_\infty; \|\phi\|_1)$. *Then, there is* $\tau > 0$ *only depending on* $M_\Lambda$ *and on* $C_u, c_u$ *(the constants coming from inequality* (7.17)*, only depending on* $\Phi$, $g$*) such that for any* $r \in [0, T-\tau]$, $\Pi$ *admits a unique fixed-point in* $L^1([r, r+\tau], B(0,M)) \cap B_\infty(0,M)$, *where* $B(0,M)$ *(resp.* $B_\infty(0,M)$*) denotes the centered ball in* $W^{1,1}(\mathbb{R}^d)$ *(resp.* $L^\infty([r, r+\tau] \times \mathbb{R}^d, \mathbb{R})$*) of radius* $M$.

*Proof.* We first insist on the fact that all along the proof, the dependence of $\widehat{u}_0$ w.r.t. $r, \phi$ in (3.23) will be omitted to simplify notations. Let us fix $r \in [0, T-\tau]$.
By item 1. of Lemma 7.4, the transition probabilities are absolutely continuous and $P(s, x_0, t, dx) = p(s, x_0, t, x)dx$ for some Borel function $p$. The rest of the proof relies on a fixed-point argument in the Banach space $L^1([r, r+\tau], W^{1,1}(\mathbb{R}^d))$ equipped with the norm $\|f\|_{1,1} := \int_r^{r+\tau} \|f(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}ds$ and for the map $\Pi$ (3.23). Moreover, we emphasize that $L^1([r, r+\tau], B(0,M)) \cap B_\infty(0,M)$ is complete as a closed subset of $L^1([r, r+\tau], B(0,M))$.

We first check that $\Pi\Big(L^1([r,r+\tau],B(0,M))\cap B_\infty(0,M)\Big)\subset L^1([r,r+\tau],B(0,M))\cap B_\infty(0,M)$. Let us fix $v\in L^1([r,r+\tau],B(0,M))\cap B_\infty(0,M)$. Indeed, for $t\in[r,r+\tau]$,

$$\begin{aligned}\|\Pi(v)(t,\cdot)\|_1 &= \int_{\mathbb{R}^d}|\Pi(v)(t,x)|dx\\ &\le M_\Lambda\int_r^t\big(\|v(s,\cdot)\|_1+\|\widehat{u_0}(s,\cdot)\|_1\big)ds\\ &\le 2M_\Lambda M\tau\ ,\end{aligned}\tag{3.26}$$

where we have used the fact that $x\mapsto p(s,x_0,t,x)$ is a probability density, the boundedness of $\Lambda$ and the bounds $\|v(s,\cdot)\|_1\le M$ and $\|\widehat{u_0}(s,\cdot)\|_1\le M$ for $s\in[r,r+\tau]$.
Let us fix $t\in[r,r+\tau]$. Since the transition probability function $x\mapsto p(s,x_0,t,x)$ is twice continuously differentiable for $0\le s<t\le T$ (see item 2. of Lemma 7.4) and taking into account inequality (7.17), we differentiate under the integral sign to obtain that $\nabla\Pi(v)(t,\cdot)$ exists (in the sense of distributions) and is a real-valued function such that for almost all $x\in\mathbb{R}^d$,

$$\nabla\Pi(v)(t,x) = \int_r^t ds\int_{\mathbb{R}^d}\nabla_x p(s,x_0,t,x)\big(v+\widehat{u_0}\big)(s,x_0)\Lambda(s,x_0,v+\widehat{u_0},\nabla(v+\widehat{u_0}))dx_0.\tag{3.27}$$

Integrating each side of (3.27) on $\mathbb{R}^d$ w.r.t. $dx$ and using inequality (7.17) (with $(m_1,m_2)=(0,1)$) yield

$$\begin{aligned}\|\nabla\Pi(v)(t,\cdot)\|_1 &= \int_{\mathbb{R}^d}|\nabla\Pi(v)(t,x)|dx\\ &\le M_\Lambda\int_r^t\frac{ds}{\sqrt{t-s}}\int_{\mathbb{R}^d}dx\int_{\mathbb{R}^d}C_u\frac{e^{\frac{-c_u|x-x_0|}{t-s}}}{\sqrt{(t-s)^d}}\big(|v(s,x_0)|+|\widehat{u_0}(s,x_0)|\big)dx_0\\ &= \hat{C}M_\Lambda\int_r^t\frac{ds}{\sqrt{t-s}}\int_{\mathbb{R}^d}\big(|v(s,x_0)|+|\widehat{u_0}(s,x_0)|\big)dx_0\\ &\le \hat{C}M_\Lambda\int_r^t\big(\|v(s,\cdot)\|_1+\|\widehat{u_0}(s,\cdot)\|_1\big)\frac{ds}{\sqrt{t-s}}\\ &\le 4\hat{C}M_\Lambda M\sqrt{\tau}\ ,\end{aligned}\tag{3.28}$$

with $\hat{C}:=\hat{C}(C_u,c_u)>0$ and $C_u,c_u$ are the constants coming from inequality (7.17) and only depending on $\Phi$ and $g$. Consequently, taking into account (3.26) and (3.28), we obtain,

$$\|\Pi(v)\|_{1,1}=\int_r^{r+\tau}\|\Pi(v)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}dt \le 2MM_\Lambda(\tau^2+2\hat{C}\tau\sqrt{\tau})\ .\tag{3.29}$$

Moreover using again inequality (7.17), with $(m_1,m_2)=(0,0)$, gives existence of a constant $\bar{C}:=\bar{C}(C_u,c_u)$ such that

$$\|\Pi(v)\|_\infty \le 2\bar{C}MM_\Lambda\tau\ .\tag{3.30}$$

Now, setting

$$\tau:=\min\Big(\sqrt{\frac{1}{6M_\Lambda}};\Big(\frac{1}{12\hat{C}M_\Lambda}\Big)^{\frac{2}{3}};\frac{1}{6\bar{C}M_\Lambda}\Big),\tag{3.31}$$

we have

$$2MM_\Lambda(\tau^2+2\hat{C}\tau\sqrt{\tau})\le\frac{2M}{3}\quad\text{and}\quad 2\bar{C}MM_\Lambda\tau\le\frac{M}{3}\ ,$$

which implies

$$\|\Pi(v)\|_{1,1}\le M\quad\text{and}\quad\|\Pi(v)\|_\infty\le M\ .$$

We deduce that $\Pi(v) \in L^1([r, r+\tau], B(0, M)) \cap B_\infty(0, M)$.
Let us fix $t \in [r, r+\tau]$, $v_1, v_2 \in L^1([r, r+\tau], B(0, M)) \cap B_\infty(0, M)$. $\Lambda$ being bounded and Lipschitz, the notation introduced in (3.24) and inequality (2.3) imply

$$\begin{aligned}
\|\Pi(v_1)(t,\cdot) - \Pi(v_2)(t,\cdot)\|_1 \quad &\le \quad \int_r^t ds \int_{\mathbb{R}^d} \Big| v_1(s,x_0)\Lambda(s,x_0,v_1+\widehat{u_0},\nabla(v_1+\widehat{u_0})) - v_2(s,x_0)\Lambda(s,x_0,v_2+\widehat{u_0},\nabla(v_2+\widehat{u_0}))\Big| dx_0 \\
&\quad + \int_r^t ds \int_{\mathbb{R}^d} |\widehat{u_0}(s,x_0)| \Big| \Lambda(s,x_0,v_1+\widehat{u_0},\nabla(v_1+\widehat{u_0})) - \Lambda(s,x_0,v_2+\widehat{u_0},\nabla(v_2+\widehat{u_0}))\Big| dx_0 \\
&\le \quad \int_r^t ds \Big( \int_{\mathbb{R}^d} |v_1(s,x_0) - v_2(s,x_0)| \, |\Lambda(s,x_0,v_1+\widehat{u_0},\nabla(v_1+\widehat{u_0}))| dx_0 \\
&\quad + L_\Lambda \int_r^t ds \int_{\mathbb{R}^d} \big( |\widehat{u_0}(s,x_0)| + |v_2(s,x_0)| \big) \, |v_1(s,x_0) - v_2(s,x_0)| dx_0 \\
&\quad + L_\Lambda \int_r^t ds \int_{\mathbb{R}^d} \big( |\widehat{u_0}(s,x_0)| + |v_2(s,x_0)| \big) \, |\nabla v_1(s,x_0) - \nabla v_2(s,x_0)| dx_0 \\
&\le \quad (M_\Lambda + 2ML_\Lambda) \int_r^t \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} ds \, , 
\end{aligned} \tag{3.32}$$

where we have used the fact that $\int_{\mathbb{R}^d} p(s,x_0,t,x)dx = 1, \ 0 \le s < t \le T$.
In the same way and by using inequality (7.17) with $(m_1, m_2) = (0, 1)$,

$$\begin{aligned}
\Big\| \nabla \Big( \Pi(v_1) - \Pi(v_2) \Big)(t,\cdot) \Big\|_1 \quad &\le \quad C_u (M_\Lambda + 2ML_\Lambda) \int_{\mathbb{R}^d} \int_r^t \int_{\mathbb{R}^d} \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{(t-s)^d}} e^{-c_u \frac{|x-x_0|^2}{t-s}} \Big( |v_1(s,x_0) - v_2(s,x_0)| \\
&\quad + |\nabla v_1(s,x_0) - \nabla v_2(s,x_0)| \Big) dx_0 ds dx \, . 
\end{aligned} \tag{3.33}$$

By Fubini's theorem we have

$$\Big\| \nabla \Big( \Pi(v_1) - \Pi(v_2) \Big)(t,\cdot) \Big\|_1 \quad \le \quad \tilde{C} \int_r^t \frac{1}{\sqrt{t-s}} \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} ds \, , \tag{3.34}$$

with $\tilde{C} := \tilde{C}(C_u, c_u, M_\Lambda, L_\Lambda, M)$ some positive constant. From (3.32) and (3.34), we deduce there exists a strictly positive constant $C = C(C_u, c_u, \Phi, g, \Lambda, M)$ (which may change from line to line) such that for all $t \in [r, r+\tau]$,

$$\begin{aligned}
\|\Pi(v_1)(t,\cdot) - \Pi(v_2)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \quad &\le \quad C \Big\{ \int_r^t \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} ds \\
&\quad + \int_r^t \frac{1}{\sqrt{t-s}} \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} ds \Big\} \, . 
\end{aligned} \tag{3.35}$$

Iterating the procedure once again yields

$$\begin{aligned}
\|\Pi^2(v_1)(t,\cdot) - \Pi^2(v_2)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \quad &\le \quad C \Big\{ \int_r^t \int_r^s \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} d\theta ds \\
&\quad + \int_r^t \int_r^s \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{s-\theta}} \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} d\theta ds \Big\} \, , 
\end{aligned} \tag{3.36}$$

for all $t \in [r, r+\tau]$. Interchanging the order in the second integral in the r.h.s. of (3.36), we obtain

$$\begin{aligned}
\int_r^t \int_r^s \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{s-\theta}} \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} d\theta ds \quad &= \quad \int_r^t d\theta \, \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \int_\theta^t \frac{1}{\sqrt{t-s}} \frac{1}{\sqrt{s-\theta}} ds \\
&= \quad \int_r^t d\theta \, \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \int_0^\alpha \frac{1}{\sqrt{\alpha-\omega}} \frac{1}{\sqrt{\omega}} d\omega, \\
&\le \quad 4 \int_r^t \|v_1(\theta,\cdot) - v_2(\theta,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} d\theta \, , 
\end{aligned} \tag{3.37}$$

where the latter line above comes from the fact for all $\theta > 0$, $\int_0^\theta \frac{1}{\sqrt{\theta-\omega}}\frac{1}{\sqrt{\omega}}d\omega = \int_0^1 \frac{1}{\sqrt{1-\omega}}\frac{1}{\sqrt{\omega}}d\omega = \Gamma(\frac{1}{2})$, $\Gamma$ denoting the Euler gamma function.
Injecting inequality (3.37) in (3.36), we obtain for all $t \in [r, r+\tau]$

$$\|\Pi^2(v_1)(t,\cdot) - \Pi^2(v_2)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \leq 5C\int_r^t \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}ds\ . \tag{3.38}$$

Iterating previous inequality, one obtains the following. For all $k \geq 1$, $t \in [r, r+\tau]$,

$$\|\Pi^{2k}(v_1)(t,\cdot) - \Pi^{2k}(v_2)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} \leq (5C)^k\int_r^t \frac{(t-s)^{k-1}}{(k-1)!}\|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}ds\ . \tag{3.39}$$

By induction on $k \geq 1$ (3.39) can indeed be established. Finally, by integrating each sides of (3.39) w.r.t. $dt$ and using Fubini's theorem, for $k \geq 1$, we obtain

$$\int_r^{r+\tau} \|\Pi^{2k}(v_1)(t,\cdot) - \Pi^{2k}(v_2)(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}dt \leq (5C)^k\frac{T^{k-1}}{(k-1)!}\int_r^{r+\tau} \|v_1(s,\cdot) - v_2(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)}ds\ . \tag{3.40}$$

For $k_0 \in \mathbb{N}$ large enough, $\frac{(5CT)^{k_0}}{T(k_0-1)!}$ will be strictly smaller than 1 and $\Pi^{2k_0}$ will admit a unique fixed-point by Banach fixed-point theorem. In consequence, it implies easily that $\Pi$ will also admit a unique fixed-point and this concludes the proof of Lemma 3.7. □

*Proof of Theorem 3.6.* Without restriction, we can suppose there exists $N \in \mathbb{N}$ such that $T = N\tau$, where we recall that $\tau$ is given by (3.31). Similarly to the notations used in the preceding proof, in all the sequel, we agree that for $M > 0$, $B(0, M)$ (resp. $B_\infty(0, M)$) denotes the centered ball of radius $M$ in $W^{1,1}(\mathbb{R}^d)$ (resp. in $L^\infty([0,T] \times \mathbb{R}^d, \mathbb{R})$ or in $L^\infty([r, r+\tau] \times \mathbb{R}^d, \mathbb{R})$ for $r \in [0, T-\tau]$ according to the context). The notations introduced in (3.24) will also be used in the present proof.

Indeed, for $r = 0$, $\phi = u_0$ and $M \geq \max(\|u_0\|_\infty; \|u_0\|_1)$, Lemma 3.7 implies there exists a unique function $v^0 : [0,\tau] \times \mathbb{R}^d \to \mathbb{R}$ (belonging to $L^1([0,\tau], B(0,M)) \cap B_\infty(0,M)$) such that for $(t,x) \in [0,\tau] \times \mathbb{R}^d$,

$$v^0(t,x) = \int_0^t ds\int_{\mathbb{R}^d} p(s,x_0,t,x)(v^0(s,x_0) + \widehat{u_0}^0(s,x_0))\Lambda\Big(s,x_0,v^0 + \widehat{u_0}^0, \nabla(v^0 + \widehat{u_0}^0)\Big)dx_0\ , \tag{3.41}$$

where $\widehat{u_0}^0(t,x)$ is given by (3.22) with $\phi = u_0$, i.e.

$$\widehat{u_0}^0(t,x) = \int_{\mathbb{R}^d} p(0,x_0,t,x)u_0(x_0)dx_0\ , \quad (t,x) \in [0,\tau] \times \mathbb{R}^d\ . \tag{3.42}$$

Setting $u^0 := \widehat{u_0}^0 + v^0$, i.e.

$$u^0(t,\cdot) = \int_{\mathbb{R}^d} p(0,x_0,t,\cdot)u_0(x_0)dx_0 + v^0(t,\cdot), \quad t \in [0,\tau]\ , \tag{3.43}$$

it appears that $u^0$ satisfies for all $(t,x) \in [0,\tau] \times \mathbb{R}^d$

$$u^0(t,x) = \int_{\mathbb{R}^d} p(0,x_0,t,x)u_0(x_0)dx_0 + \int_0^t ds\int_{\mathbb{R}^d} p(s,x_0,t,x)u^0(s,x_0)\Lambda(s,x_0,u^0,\nabla u^0)dx_0\ . \tag{3.44}$$

Let us fix $k \in \{1,\cdots,N-1\}$. Suppose now given a family of functions $u^1, u^2, \cdots, u^{k-1}$, where for all $j \in \{1,\cdots,k-1\}$, $u^j \in L^1([j\tau,(j+1)\tau], W^{1,1}(\mathbb{R}^d)) \cap L^\infty([j\tau,(j+1)\tau] \times \mathbb{R}^d, \mathbb{R})$ and satisfies for all $(t,x) \in [j\tau,(j+1)\tau] \times \mathbb{R}^d$,

$$u^j(t,x) = \int_{\mathbb{R}^d} p(j\tau,x_0,t,x)u^{j-1}(j\tau,x_0)dx_0 + \int_{j\tau}^t ds\int_{\mathbb{R}^d} p(s,x_0,t,x)u^j(s,x_0)\Lambda(s,x_0,u^j,\nabla u^j)dx_0\ . \tag{3.45}$$

Let us introduce

$$\begin{aligned}\widehat{u_0}^k(t,x) &:= \widehat{u_0}(u^{k-1})(t,x) \\ &= \int_{\mathbb{R}^d} p(k\tau, x_0, t, x) u^{k-1}(k\tau, x_0) dx_0 \ , \quad (t,x) \in [k\tau, (k+1)\tau] \times \mathbb{R}^d \ , \end{aligned} \tag{3.46}$$

where the second inequality comes from (3.22) with $r = k\tau$ and $\phi = u^{k-1}(k\tau, \cdot)$.
By choosing the real $M$ large enough (i.e. $M \geq \max(\|u^{k-1}(k\tau,\cdot)\|_\infty ; \|u^{k-1}(k\tau,\cdot)\|_1)$), Lemma 3.7 applied with $r = k\tau$, $\phi = u^{k-1}(k\tau, \cdot)$ implies existence and uniqueness of a function $v^k : [k\tau,(k+1)\tau] \times \mathbb{R}^d \to \mathbb{R}$ that belongs to $L^1([k\tau,(k+1)\tau], B(0,M)) \cap B_\infty(0,M)$ and satisfying

$$v^k(t,x) = \int_{k\tau}^t ds \int_{\mathbb{R}^d} p(s,x_0,t,x)(v^k(s,x_0) + \widehat{u_0}^k(s,x_0))\Lambda\big(s,x_0,v^k + \widehat{u_0}^k, \nabla(v^k + \widehat{u_0}^k)\big)dx_0 \ , \tag{3.47}$$

for all $(t,x) \in [k\tau,(k+1)\tau] \times \mathbb{R}^d$. Setting $u^k := \widehat{u_0}^k + v^k$, we have for all $(t,x) \in [k\tau,(k+1)\tau] \times \mathbb{R}^d$

$$u^k(t,x) = \int_{\mathbb{R}^d} p(k\tau,x_0,t,x)u^{k-1}(k\tau,x_0)dx_0 + \int_{k\tau}^t ds \int_{\mathbb{R}^d} p(s,x_0,t,x)u^k(s,x_0)\Lambda(s,x_0,u^k,\nabla u^k)dx_0 \ . \tag{3.48}$$

Consequently, by induction we can construct a family of functions $(u^k : [k\tau,(k+1)\tau] \times \mathbb{R}^d \to \mathbb{R})_{k=0,\cdots,N-1}$ such that for all $k \in \{0,\cdots,N-1\}$, $u^k \in L^1([k\tau,(k+1)\tau], W^{1,1}(\mathbb{R}^d)) \cap L^\infty([k\tau,(k+1)\tau] \times \mathbb{R}^d, \mathbb{R})$ and verifies (3.48).

We now consider the real-valued function $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ defined as being equal to $u^k$ (defined by (3.48)) on each interval $[k\tau,(k+1)\tau]$. Then, Lemma 3.4 applied with $\tau$ given by (3.31) and

$$\delta = (\alpha_0 := 0 < \cdots < \alpha_k := k\tau < \cdots < \alpha_N := T = N\tau) \quad , \quad \mu(t,dx) = u(t,x)dx, \tag{3.49}$$

shows that $u$ is a mild solution of (1.1) on $[0,T] \times \mathbb{R}^d$, in the sense of Definition 2.1, item 2. It now remains to ensure that $u$ is indeed the unique mild solution of (1.1) on $[0,T] \times \mathbb{R}^d$ belonging to $L^1([0,T], W^{1,1}(\mathbb{R}^d)) \cap L^\infty([0,T] \times \mathbb{R}^d, \mathbb{R})$. This follows, in a classical way, by boundedness and Lipschitz property of $\Lambda$.
Indeed, if $U$, $V$ are two mild solutions of (1.1), then very similar computations as the ones done in (3.32), (3.34), and (3.39) to obtain (3.40) give the following. There exists $C := C(\Phi, g, \Lambda, U, V) > 0$ such that

$$\int_0^T \|U(t,\cdot) - V(t,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} dt \leq (5C)^j \frac{T^{j-1}}{(j-1)!} \int_0^T \|U(s,\cdot) - V(s,\cdot)\|_{W^{1,1}(\mathbb{R}^d)} ds. \tag{3.50}$$

If we choose $j \in \mathbb{N}^\star$ large enough so that $(5C)^j \frac{T^{j-1}}{(j-1)!} < 1$, we obtain $U(t,x) = V(t,x)$ for almost all $(t,x) \in [0,T] \times \mathbb{R}^d$. This concludes the proof of Theorem 3.6. □

**Corollary 3.8.** *Under Assumption 2, there exists a unique function $u : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ satisfying the Feynman-Kac equation* (3.15)*. In particular, such $u$ coincides with the mild solution of* (1.1)*.*

In the case where the function $\Lambda$ does not depend on $\nabla u$, existence and uniqueness of a solution of (1.1) in the mild sense can be proved under weaker assumptions. This is the object of the following result.

**Theorem 3.9.** *Assume that Assumption 1 is satisfied. Let $u_0 \in \mathcal{P}(\mathbb{R}^d)$ admitting a bounded density (still denoted by the same letter). Let $Y$ the the strong solution of* (2.10) *with prescribed $Y_0$.*
*We suppose that the transition probability function $P$ (see* (2.9)*) admits a density $p$ such that $P(s,x_0,t,dx) = p(s,x_0,t,x)dx$, for all $s,t \in [0,T]$, $x_0 \in \mathbb{R}^d$. $\Lambda$ is supposed to satisfy items 4. and 5. of Assumption 2. Then, there exists a unique mild solution $u$ of* (1.1) *in $L^1([0,T], L^1(\mathbb{R}^d))$, i.e. $u$ satisfies*

$$u(t,x) = \int_{\mathbb{R}^d} p(0,x_0,t,x)u_0(x_0)dx_0 + \int_0^t \int_{\mathbb{R}^d} p(s,x_0,t,x)u(s,x_0)\Lambda(s,x_0,u(s,x_0))dx_0 \, ds, \quad (t,x) \in [0,T] \times \mathbb{R}^d \ . \tag{3.51}$$

*Proof.* Since this theorem can be proved in a very similar way as Theorem 3.6 but with simpler computations, we omit the details. □

# 4 Existence/uniqueness of the Regularized Feynman-Kac equation

In this section, we introduce a regularized version of PDE (1.1) to which we associate a regularized Feynman-Kac equation corresponding to a regularized version of (3.15). This regularization procedure constitutes a preliminary step for the construction of a particle scheme approximating (3.15). Indeed, as detailed in the next section devoted to the particle approximation, the point dependence of $\Lambda$ on $u$ and $\nabla u$ will require to derive from a discrete measure (based on the particle system) estimates of densities $u$ and their derivatives $\nabla u$, which can classically be achieved by kernel convolution.

Assumption 1 is in force. Let $u_0$ be a Borel probability measure on $\mathbb{R}^d$ and $Y_0$ a random variable distributed according to $u_0$. We consider $Y$ the strong solution of the SDE (2.10).

Let us consider $(K_\varepsilon)_{\varepsilon>0}$, a sequence of mollifiers such that

$$K_\varepsilon \xrightarrow[\varepsilon\to 0]{} \delta_0, \text{ (weakly)} \quad \text{and} \quad \forall\, \varepsilon>0, K_\varepsilon \in W^{1,1}(\mathbb{R}^d)\cap W^{1,\infty}(\mathbb{R}^d)\,. \tag{4.1}$$

Let $\Lambda : [0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$ be bounded, Borel measurable. As announced, we introduce the following integro-PDE corresponding to a regularized version of (1.1)

$$\begin{cases} \partial_t \gamma_t = L_t^* \gamma_t + \gamma_t \Lambda(t,x,K_\varepsilon * \gamma_t, \nabla K_\varepsilon * \gamma_t) \\ \gamma_0 = u_0\,. \end{cases} \tag{4.2}$$

The concept of mild solution associated to this type of equation is clarified by the following definition.

**Definition 4.1.** *A Borel measure-valued function $\gamma : [0,T] \longrightarrow \mathcal{M}_f(\mathbb{R}^d)$ will be called a* **mild solution** *of* (4.2) *if it satisfies, for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $t\in[0,T]$,*

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)\gamma(t,dx) &= \int_{\mathbb{R}^d}\varphi(x)\int_{\mathbb{R}^d} u_0(dx_0)P(0,x_0,t,dx) \\ &+ \int_{[0,t]\times\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\varphi(x)P(s,x_0,t,dx)\Big)\Lambda\big(s,x_0,(K_\varepsilon*\gamma(s,\cdot))(x_0),(\nabla K_\varepsilon*\gamma(s,\cdot))(x_0)\big)\gamma(s,dx_0)ds\,. \end{aligned} \tag{4.3}$$

Similarly as Theorem 3.5, we straightforwardly derive the following equivalence result.

**Proposition 4.2.** *Suppose that Assumption 1 and* (4.1) *are fulfilled. We indicate by $Y$ the unique strong solution of* (2.10)*. A Borel measure-valued function $\gamma^\varepsilon : [0,T] \longrightarrow \mathcal{M}_f(\mathbb{R}^d)$ is a mild solution of* (4.2) *if and only if, for all $\varphi\in\mathcal{C}_b(\mathbb{R}^d)$, $t\in[0,T]$,*

$$\int_{\mathbb{R}^d}\varphi(x)\gamma_t^\varepsilon(dx) = \mathbb{E}\Big[\varphi(Y_t)\exp\Big(\int_0^t \Lambda\big(s,Y_s,(K_\varepsilon*\gamma_s^\varepsilon)(Y_s),(\nabla K_\varepsilon*\gamma_s^\varepsilon)(Y_s)\big)\,ds\Big)\Big]\,. \tag{4.4}$$

*Proof.* The proof follows the same lines as the proof of Theorem 3.5. First assume that $\gamma^\varepsilon$ satisfies (4.4), we can show that $\gamma^\varepsilon$ is a mild solution (4.2) by imitating the first step of the proof of Proposition 3.3. Secondly, the converse is proved by applying Proposition 3.3 with $\tilde{\Lambda}(t,x) := \Lambda(t,x,(K_\varepsilon*\gamma_t^\varepsilon)(x),(\nabla K_\varepsilon*\gamma_t^\varepsilon)(x))$ and $\mu(t,dx) := \gamma_t^\varepsilon(dx)$. □

Let us now prove existence and uniqueness of a mild solution for the integro-PDE (4.2). To this end, we proceed similarly as for the proof of Theorem 3.6 using Lemma 3.7. Let $\tau > 0$ be a constant supposed to be fixed for the moment and let us fix $\varepsilon > 0$, $r \in [0, T-\tau]$. $\mathcal{B}([r, r+\tau], \mathcal{M}_f(\mathbb{R}^d))$ denotes the space of bounded, measure-valued maps, where $\mathcal{M}_f(\mathbb{R}^d)$ is equipped with the total variation norm $\|\cdot\|_{TV}$. We introduce the measure-valued application $\Pi_\varepsilon : \beta \in \mathcal{B}([r, r+\tau], \mathcal{M}_f(\mathbb{R}^d)) \longrightarrow \Pi_\varepsilon(\beta)$, defined by

$$\begin{aligned}
\Pi_\varepsilon(\beta)(t,dx) &= \int_r^t \int_{\mathbb{R}^d} P(s,x_0,t,dx)\Lambda\big(s,x_0,(K_\varepsilon * \widehat{\beta}(s,\cdot))(x_0),(\nabla K_\varepsilon * \widehat{\beta}(s,\cdot))(x_0)\big)\ \widehat{\beta}(s,dx_0)ds\ . \\
\widehat{\beta}(s,\cdot) &= \beta(s,\cdot) + \widehat{u_0}(s,\cdot)\ ,
\end{aligned} \tag{4.5}$$

where the function $\widehat{u_0}$, defined on $[r, r+\tau] \times \mathcal{M}_f(\mathbb{R}^d)$, is given by

$$\widehat{u_0}(r,\pi)(t,dx) := \int_{\mathbb{R}^d} p(r,x_0,t,dx)\pi(dx_0),\ t \in [r,r+\tau],\ \pi \in \mathcal{M}_f(\mathbb{R}^d)\ , \tag{4.6}$$

similarly to (3.22). In the sequel, the dependence of $\widehat{u}_0$ w.r.t. $r, \pi$ will be omitted when it is self-explanatory.

**Lemma 4.3.** *Assume the validity of items 4. and 5. of Assumption 2 and of* (4.1). *Let* $\pi \in B(0,M)$.
*Let us fix* $\varepsilon > 0$ *and* $M > 0$ *such that* $M \geq \|\pi\|_{TV}$. *Then, there is* $\tau > 0$ *only depending on* $M_\Lambda$ *such that for any* $r \in [0, T-\tau]$, $\Pi_\varepsilon$ *admits a unique fixed-point in* $\mathcal{B}([r, r+\tau], B(0,M))$, *where* $B(0,M)$ *denotes here the centered ball in* $(\mathcal{M}_f(\mathbb{R}^d), \|\cdot\|_{TV})$ *with radius* $M$.

*Proof.* Let us define $\tau := \frac{1}{2M_\Lambda}$. For every $\lambda \geq 0$, $\mathcal{B}([r, r+\tau], \mathcal{M}_f(\mathbb{R}^d))$ will be equipped with one of the equivalent norms

$$\|\beta\|_{TV,\lambda} := \sup_{t\in[r,r+\tau]} e^{-\lambda t}\|\beta(t,\cdot)\|_{TV}\ . \tag{4.7}$$

Recalling (4.5), where $\widehat{u_0}$ is defined by (4.6), it follows that for all $\beta \in \mathcal{B}([r, r+\tau], B(0,M))$, $t \in [r, r+\tau]$,

$$\|\Pi_\varepsilon(\beta)(t,\cdot)\|_{TV} \leq M_\Lambda \int_r^t \|\beta(s,\cdot)\|_{TV}ds + M_\Lambda M\tau \leq 2MM_\Lambda\tau \leq M\ , \tag{4.8}$$

where for the latter inequality of (4.8) we have used the definition of $\tau := \frac{1}{2M_\Lambda}$. We deduce that $\Pi(\mathcal{B}([r, r+\tau], B(0,M))) \subset \mathcal{B}([r, r+\tau], B(0,M))$.
Consider now $\beta^1, \beta^2 \in \mathcal{B}([r, r+\tau], B(0,M))$. For all $\lambda > 0$ we have

$$\begin{aligned}
\|\Pi_\varepsilon(\beta^1(t,\cdot)) - \Pi_\varepsilon(\beta^2(t,\cdot))\|_{TV} &\leq \int_0^t \|\beta^1(s,\cdot) - \beta^2(s,\cdot)\|_{TV}(L_\Lambda\|K_\varepsilon\|_\infty\|\beta^1(s,\cdot)\|_{TV} + M_\Lambda)ds \\
&\quad + L_\Lambda\|\nabla K_\varepsilon\|_\infty \int_r^t \|\beta^1(s,\cdot)\|_{TV}\|\beta^1(s,\cdot) - \beta^2(s,\cdot)\|_{TV}ds \\
&\quad + L_\Lambda(\|K_\varepsilon\|_\infty + \|\nabla K_\varepsilon\|_\infty)\int_r^t \|\widehat{u_0}(s,\cdot)\|_{TV}\|\beta^1(s,\cdot) - \beta^2(s,\cdot)\|_{TV}ds \\
&\leq C_{\varepsilon,T}\int_0^t \|\beta^1(s,\cdot) - \beta^2(s,\cdot)\|_{TV}ds \\
&\leq C_{\varepsilon,T}\int_0^t e^{s\lambda}\ \|\beta^1 - \beta^2\|_{TV,\lambda}\ ds \\
&= C_{\varepsilon,T}\|\beta^1 - \beta^2\|_{TV,\lambda}\ \frac{e^{\lambda t}-1}{\lambda}\ ,
\end{aligned} \tag{4.9}$$

with $C_{\varepsilon,T} := 2L_\Lambda M(\|K_\varepsilon\|_\infty + \|\nabla K_\varepsilon\|_\infty) + M_\Lambda$. It follows

$$\begin{aligned} \|\Pi_\varepsilon(\beta^1) - \Pi_\varepsilon(\beta^2)\|_{TV,\lambda} &= \sup_{t\in[r,r+\tau]} e^{-\lambda t}\|\Pi(\beta^1)(t,\cdot) - \Pi(\beta^2)(t,\cdot)\|_{TV} \\ &\leq C_{\varepsilon,T}\|\beta^1 - \beta^2\|_{TV,\lambda} \sup_{t\geq 0}\Big(\frac{1-e^{-\lambda t}}{\lambda}\Big) \\ &\leq \frac{C_{\varepsilon,T}}{\lambda}\|\beta^1 - \beta^2\|_{TV,\lambda} . \end{aligned} \tag{4.10}$$

Hence, taking $\lambda > C_{\varepsilon,T}$, $\Pi_\varepsilon$ is a contraction on $\mathcal{B}([r, r+\tau], B(0,M))$.
Since $\mathcal{B}\big([r, r+\tau], (\mathcal{M}_f(\mathbb{R}^d), \|\cdot\|_{TV,\lambda})\big)$ is a Banach space whose $\mathcal{B}([r, r+\tau], B(0,M))$ is a closed subset, the proof ends by a simple application of Banach fixed-point theorem. □

The next step is to show how the proposition above, with the help of Lemma 3.4, permits us to construct a mild solution of (4.2). The reasoning is similar to the one explained in the proof of Theorem 3.6. Indeed, without restriction of generality, we can suppose there exists $N \in \mathbb{N}^\star$ such that $T = N\tau$. Then, for all $k = 0, \cdots, N-1$, Lemma 4.3 applied on each interval $[k\tau, (k+1)\tau]$ (with $r = k\tau$, $\pi = \beta_\varepsilon^{k-1}(k\tau,\cdot)$ for $k \geq 1$ and $\pi = u_0$ for $k = 0$) gives existence of a family of measure-valued maps $(\beta_\varepsilon^k : [k\tau, (k+1)\tau] \to \mathcal{M}_f(\mathbb{R}^d), k = 0, \cdots, N-1)$ defined by

$$\begin{aligned} \beta_\varepsilon^k(t,dx) &= \int_{k\tau}^t \int_{\mathbb{R}^d} P(s,x_0,t,dx)\Lambda\big(s,x_0,(K_\varepsilon * \widehat{\beta}_\varepsilon^k(s,\cdot))(x_0),(\nabla K_\varepsilon * \widehat{\beta}_\varepsilon^k(s,\cdot))(x_0)\big)\ \widehat{\beta}_\varepsilon^k(s,dx_0)ds . \\ \widehat{\beta_\varepsilon^k}(s,\cdot) &= \beta_\varepsilon^k(s,\cdot) + \widehat{u_0}^k(s,\cdot) , \end{aligned} \tag{4.11}$$

where for $k = 0$, $t \in [0,\tau]$,

$$\widehat{u}_0^0(t,dx) = \int_{\mathbb{R}^d} P(0,x_0,t,dx)u_0(dx_0) , \quad \text{by (4.6) with } \pi = u_0 , \tag{4.12}$$

and for all $k \in \{1, \cdots, N\}$, $t \in [k\tau, (k+1)\tau]$,

$$\begin{aligned} \widehat{u_0}^k(t,dx) &:= \widehat{u_0}(\beta_\varepsilon^{k-1})(t,dx) \\ &= \int_{\mathbb{R}^d} P(k\tau,x_0,t,dx)\beta_\varepsilon^{k-1}(k\tau,dx_0) , \quad \text{by (4.6) with } \pi = \beta_\varepsilon^{k-1}(k\tau,\cdot) . \end{aligned} \tag{4.13}$$

We now consider the following measure-valued maps $\widehat{U_0} : [0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ and $\beta_\varepsilon : [0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ defined by their restrictions on each interval $[k\tau, (k+1)\tau]$, $k = 0, \cdots, N-1$ such that

$$\widehat{U_0}(t,x) := \widehat{u_0}^k(t,x) \quad \text{and} \quad \beta_\varepsilon(t,x) := \beta_\varepsilon^k(t,x) \quad \text{for } (t,x) \in [k\tau,(k+1)\tau] \times \mathbb{R}^d , \tag{4.14}$$

and we finally define $\gamma^\varepsilon : [0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ by

$$\gamma^\varepsilon := \widehat{U_0} + \beta_\varepsilon \text{ on } [0,T] \times \mathbb{R}^d. \tag{4.15}$$

To ensure that $\gamma^\varepsilon$ is indeed a mild solution on $[0,T] \times \mathbb{R}^d$ (in the sense of Definition 4.1) of the integro-PDE (4.2), it is enough to apply Lemma 3.4 with $\tau := \frac{1}{2M_\Lambda}$, $\mu(t,dx) := \gamma^\varepsilon(t,dx)$ and $(\alpha_k := k\tau)_{k=0,\cdots,N}$.

Previous discussion leads us to the following proposition.

**Proposition 4.4.** *Suppose the validity of Assumption 1 and items 4. and 5. of Assumption 2. Suppose also that* (4.1) *is fulfilled. Let us fix $\varepsilon > 0$ and let $\gamma^\varepsilon$ denote the map defined by* (4.15)*. The following statements hold.*

1. *$\gamma^\varepsilon$ is the unique mild solution of the integro-PDE* (4.2)*, see Definition 4.1.*

2. *$\gamma^\varepsilon$ is the unique solution to the regularized Feynman-Kac equation* (4.4).

*Proof.* The existence of a mild solution $\gamma^\varepsilon$ of (4.2) has already been proved through the discussion just above. It remains to justify uniqueness. Consider $\gamma^{\varepsilon,1}$, $\gamma^{\varepsilon,2}$ be two mild solutions of (4.3). Then, with similar computations as the ones leading to inequality (4.10), there exists a constant $\mathfrak{C} := \mathfrak{C}(M_\Lambda, L_\Lambda, \|K_\varepsilon\|_\infty, \|\nabla K_\varepsilon\|_\infty) > 0$ such that

$$\|\gamma^{\varepsilon,1} - \gamma^{\varepsilon,2}\|_{TV,\lambda} \leq \frac{\mathfrak{C}}{\lambda}\|\gamma^{\varepsilon,1} - \gamma^{\varepsilon,2}\|_{TV,\lambda} \ , \tag{4.16}$$

for all $\lambda > 0$ and where we recall that $\|\cdot\|_{TV,\lambda}$ has been defined by (4.7). Taking $\lambda > \mathfrak{C}$, uniqueness follows. This shows item 1. Item 2. follows then by Proposition 4.2. □

The theorem below states the convergence of the solution of the regularized Feynman-Kac equation (4.4) to the solution to the Feynman-Kac equation (3.15). This is equivalent to the convergence of the solution of the regularized PDE (4.2) to solution of the target PDE (1.1), when the regularization parameter $\varepsilon$ goes to zero.

**Theorem 4.5.** *Suppose the validity of Assumption 2. Suppose also that* (4.1) *is fulfilled. For any $\varepsilon > 0$, consider the real valued function $u^\varepsilon$ such that for any $t \in [0,T]$,*

$$u^\varepsilon(t,\cdot) := K_\varepsilon * \gamma_t^\varepsilon \ , \tag{4.17}$$

*where $\gamma^\varepsilon$ is the unique solution of both* (4.4) *and* (4.2)*. Then $u^\varepsilon$ converges to $u$, the unique solution of both* (3.15) *and* (1.1)*, in the sense that*

$$\|u^\varepsilon(t,\cdot) - u(t,\cdot)\|_1 + \|\nabla u^\varepsilon(t,\cdot) - \nabla u(t,\cdot)\|_1 \xrightarrow[\varepsilon \to 0]{} 0 \ , \quad \textit{for any } t \in [0,T] \ . \tag{4.18}$$

Before proving Theorem 4.5, we state and prove a preliminary lemma.

**Lemma 4.6.** *Suppose the validity of Assumption 2 and of* (4.1)*. Consider $u$ the unique solution of* (3.15)*, then for all $t \in [0,T]$*

$$u(t,x) = F_0(t,x) + \int_0^t \mathbb{E}\Big[p(s,Y_s,t,x)\Lambda(s,Y_s,u,\nabla u)e^{\int_0^s \Lambda(r,Y_r,u,\nabla u)dr}\Big]ds, \ dx \text{ a.e.} \tag{4.19}$$

*For a given $\varepsilon > 0$, consider $u^\varepsilon$ defined by* (4.17)*. Then for almost all $x \in \mathbb{R}^d$ and all $t \in [0,T]$,*

$$u^\varepsilon(t,x) = (K_\varepsilon * F_0(t,\cdot))(x) + \int_0^t \mathbb{E}\Big[(K_\varepsilon * p(s,Y_s,t,\cdot))(x)\Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^s \Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)dr}\Big]ds \ , \tag{4.20}$$

*where $F_0(t,x) := \int_{\mathbb{R}^d} p(0,x_0,t,x)u_0(x_0)dx_0$ for $t > 0, x \in \mathbb{R}^d$ and $F_0(0,\cdot) := u_0$. We remark that we have used again the notation*

$$\Lambda(s,\cdot,v,\nabla v) := \Lambda(s,\cdot,v(s,\cdot),\nabla v(s,\cdot)), \quad t \in [0,T] \ , \tag{4.21}$$

*for $v \in L^1([0,T], W^{1,1}(\mathbb{R}^d))$.*

*Proof.* Equalities (4.19) and (4.20) are proved in a very similar way, so we only provide the proof of equation (4.20).
We observe that for all $t \in [0,T]$, $v \in L^1([0,T], W^{1,1}(\mathbb{R}^d))$,

$$e^{\int_0^t \Lambda(r,Y_r,v(r,Y_r),\nabla v(r,Y_r))dr} = 1 + \int_0^t \Lambda(r,Y_r,v(r,Y_r),\nabla v(r,Y_r))e^{\int_0^r \Lambda(s,Y_s,v(s,Y_s),\nabla v(s,Y_s))ds} \ . \tag{4.22}$$

Taking into account the notation introduced in (4.21), (4.22) above implies for almost all $x \in \mathbb{R}^d$,

$$\begin{aligned}
u^\varepsilon(t,x) &= \mathbb{E}\Big[K_\varepsilon(x-Y_t)\,\exp\Big\{\int_0^t \Lambda\Big(s,Y_s,u^\varepsilon,\nabla u^\varepsilon\Big)ds\Big\}\Big] \\
&= \mathbb{E}\Big[K_\varepsilon(x-Y_t)\Big] + \int_0^t \mathbb{E}\Big[K_\varepsilon(x-Y_t)\Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^r \Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)ds}\Big]dr \\
&= \int_{\mathbb{R}^d} K_\varepsilon(x-y)\int_{\mathbb{R}^d} p(0,x_0,t,y)u_0(x_0)dx_0\; dy + \\
&\quad \int_0^t \mathbb{E}\Big[\mathbb{E}\Big[K_\varepsilon(x-Y_t)\Big|Y_r\Big]\Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^r \Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)ds}\Big]dr \\
&= (K_\varepsilon * F_0)(t,\cdot)(x) + \\
&\quad \int_0^t \mathbb{E}\Big[\Big(\int_{\mathbb{R}^d} K_\varepsilon(x-y)p(r,Y_r,t,y)dy\Big)\;\Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^r \Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)ds}\Big]dr \\
&= (K_\varepsilon * F_0)(t,\cdot)(x) + \\
&\quad \int_0^t \mathbb{E}\Big[(K_\varepsilon * p(r,Y_r,t,\cdot))(x)\;\Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^r \Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)ds}\Big]dr\,. \qquad (4.23)
\end{aligned}$$

This ends the proof. □

*Proof of Theorem 4.5.* In this proof, $C$ denotes a real constant that may change from line to line, only depending on $M_\Lambda$, $L_\Lambda$, $C_u$, $c_u$ and $\|u_0\|_\infty$, where we recall that the constants $C_u$, $c_u$ come from inequality (7.17) (and only depend on $\Phi, g$).
We first observe that for $t = 0$, the convergence of $u^\varepsilon(0,\cdot)$ (resp. $\nabla u^\varepsilon(0,\cdot)$) to $u(0,\cdot)$ (resp. $\nabla u(0,\cdot)$) in $L^1(\mathbb{R}^d)$-norm when $\varepsilon$ goes to 0 is clear. Let us fix $t \in (0,T]$.
By Lemma 4.6, for almost all $x \in \mathbb{R}^d$, we have the decomposition

$$\begin{aligned}
u^\varepsilon(t,x) - u(t,x) &= (K_\varepsilon * F_0(t,\cdot))(x) - F_0(t,x) + \\
&\quad \int_0^t \mathbb{E}\Big[\Big\{(K_\varepsilon * p(s,Y_s,t,\cdot))(x) - p(s,Y_s,t,x)\Big\}\Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^s \Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)dr}\Big]ds + \\
&\quad \int_0^t \mathbb{E}\Big[p(s,Y_s,t,x)\Big\{\Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^s \Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)dr} - \Lambda(s,Y_s,u,\nabla u)e^{\int_0^s \Lambda(r,Y_r,u,\nabla u)dr}\Big\}\Big]ds\,. \qquad (4.24)
\end{aligned}$$

By integrating the absolute value of both sides of (4.24) w.r.t. $dx$, it follows there exists a constant $C > 0$ such that

$$\begin{aligned}
\|u^\varepsilon(t,\cdot) - u(t,\cdot)\|_1 \;&\le\; C\Big\{\|K_\varepsilon * F_0 - F_0\|_1 + \int_0^t \mathbb{E}\Big[\|K_\varepsilon * p(s,Y_s,t,\cdot) - p(s,Y_s,t,\cdot)\|_1\Big]ds \\
&\quad + \int_0^t \mathbb{E}\Big[|u^\varepsilon(s,Y_s) - u(s,Y_s)| + |\nabla u^\varepsilon(s,Y_s) - \nabla u(s,Y_s)|\Big]ds \\
&\quad + \int_0^t\int_0^s \mathbb{E}\Big[|u^\varepsilon(r,Y_r) - u(r,Y_r)| + |\nabla u^\varepsilon(r,Y_r) - \nabla u(r,Y_r)|\Big]dr\;ds\Big\}\,. \qquad (4.25)
\end{aligned}$$

Moreover, denoting $p_s$ the law density of $Y_s$, by inequality (7.18) of Lemma 7.4 we get

$$\begin{aligned}
\mathbb{E}\Big[|u^\varepsilon(s,Y_s) - u(s,Y_s)|\Big] &= \int_{\mathbb{R}^d} |u^\varepsilon(s,x) - u(s,x)|p_s(x)dx \\
&\le C\,\|u_0\|_\infty \int_{\mathbb{R}^d} |u^\varepsilon(s,x) - u(s,x)|dx \\
&= C\,\|u^\varepsilon(s,\cdot) - u(s,\cdot)\|_1\,, \quad s \in [0,T]\,, \qquad (4.26)
\end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}\Big[|\nabla u^\varepsilon(s,Y_s)-\nabla u(s,Y_s)|\Big] &= \int_{\mathbb{R}^d} |\nabla u^\varepsilon(s,x)-\nabla u(s,x)|p_s(x)dx \\
&\leq C\,\|u_0\|_\infty \int_{\mathbb{R}^d} |\nabla u^\varepsilon(s,x)-\nabla u(s,x)|dx \\
&= C\,\|\nabla u^\varepsilon(s,\cdot)-\nabla u(s,\cdot)\|_1\,, \quad s\in[0,T]\,. \qquad (4.27)
\end{aligned}
$$

Injecting (4.26) and (4.27) into the r.h.s. of (4.25), it comes

$$
\begin{aligned}
\|u^\varepsilon(t,\cdot)-u(t,\cdot)\|_1 \leq\; & C\Big\{\|K_\varepsilon * F_0 - F_0\|_1 + \int_0^t \mathbb{E}\Big[\|K_\varepsilon * p(s,Y_s,t,\cdot)-p(s,Y_s,t,\cdot)\|_1\Big]ds \\
&+ \int_0^t \|u^\varepsilon(s,\cdot)-u(s,\cdot)\|_1 + \|\nabla u^\varepsilon(s,\cdot)-\nabla u(s,\cdot)\|_1 ds\Big\}\,. \qquad (4.28)
\end{aligned}
$$

Now, we need to bound $\|\nabla u^\varepsilon(t,\cdot)-\nabla u(t,\cdot)\|_1$. To this end, we can remark that for almost all $x\in\mathbb{R}^d$,

$$
\nabla u(t,x) = \nabla F_0(t,x) + \int_0^t \mathbb{E}\Big[\nabla_x p(s,Y_s,t,x)\Lambda(s,Y_s,u,\nabla u)e^{\int_0^s \Lambda(r,Y_r,u,\nabla u)dr}\Big]ds\,, \qquad (4.29)
$$

and

$$
\begin{aligned}
\nabla u^\varepsilon(t,x) =\; & (K_\varepsilon * \nabla F_0(t,\cdot))(x) \\
&+ \int_0^t \mathbb{E}\Big[(K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot))(x)\Lambda(s,Y_s,u^\varepsilon,\nabla u^\varepsilon)e^{\int_0^s \Lambda(r,Y_r,u^\varepsilon,\nabla u^\varepsilon)dr}\Big]ds\,. \qquad (4.30)
\end{aligned}
$$

These equalities follow by computing the derivative of $u(t,\cdot)$ and $u^\varepsilon(t,\cdot)$ in the sense of distributions. Taking into account (4.29) and (4.30), it is easy to see that very similar arguments as those invoked above to prove (4.28), lead to

$$
\begin{aligned}
\|\nabla u^\varepsilon(t,\cdot)-\nabla u(t,\cdot)\|_1 \leq\; & C\Big\{\|K_\varepsilon * \nabla F_0(t,\cdot) - \nabla F_0(t,\cdot)\|_1 + \int_0^t \mathbb{E}\Big[\|K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot)-\nabla_x p(s,Y_s,t,\cdot)\|_1\Big]ds \\
&+ \int_0^t \|u^\varepsilon(s,\cdot)-u(s,\cdot)\|_1 + \|\nabla u^\varepsilon(s,\cdot)-\nabla u(s,\cdot)\|_1 ds\Big\}\,. \qquad (4.31)
\end{aligned}
$$

Gathering (4.28) together with (4.31) yields

$$
\begin{aligned}
\|u^\varepsilon(t,\cdot)-u(t,\cdot)\|_1 + \|\nabla u^\varepsilon(t,\cdot)-\nabla u(t,\cdot)\|_1 \leq\; & C\,\Big\{\|K_\varepsilon * F_0(t,\cdot) - F_0(t,\cdot)\|_1 + \|K_\varepsilon * \nabla F_0(t,\cdot) - \nabla F_0(t,\cdot)\|_1 + \\
& \int_0^t \mathbb{E}\Big[\|K_\varepsilon * p(s,Y_s,t,\cdot)-p(s,Y_s,t,\cdot)\|_1\Big]ds + \\
& \int_0^t \mathbb{E}\Big[\|K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot)-\nabla_x p(s,Y_s,t,\cdot)\|_1\Big]ds\Big\} \\
&+ \int_0^t \|u^\varepsilon(s,\cdot)-u(s,\cdot)\|_1 + \|\nabla u^\varepsilon(s,\cdot)-\nabla u(s,\cdot)\|_1 ds\,.
\end{aligned}
\qquad (4.32)
$$

Applying Gronwall's lemma to the real-valued function

$$
t \mapsto \|u^\varepsilon(t,\cdot)-u(t,\cdot)\|_1 + \|\nabla u^\varepsilon(t,\cdot)-\nabla u(t,\cdot)\|_1\,,
$$

we obtain

$$
\begin{aligned}
\|u^\varepsilon(t,\cdot)-u(t,\cdot)\|_1 + \|\nabla u^\varepsilon(t,\cdot)-\nabla u(t,\cdot)\|_1 \leq\; & Ce^{CT}\,\Big\{\|K_\varepsilon * F_0(t,\cdot) - F_0(t,\cdot)\|_1 + \|K_\varepsilon * \nabla F_0(t,\cdot) - \nabla F_0(t,\cdot)\|_1 + \\
& \int_0^t \mathbb{E}\Big[\|K_\varepsilon * p(s,Y_s,t,\cdot)-p(s,Y_s,t,\cdot)\|_1\Big]ds + \\
& \int_0^t \mathbb{E}\Big[\|K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot)-\nabla_x p(s,Y_s,t,\cdot)\|_1\Big]ds\Big\}. \qquad (4.33)
\end{aligned}
$$

Since $F_0(t,\cdot)$, $\nabla F_0(t,\cdot)$, $x \mapsto p(s,x_0,t,x)$ and $x \mapsto \nabla_x p(s,x_0,t,x)$ belong to $L^1(\mathbb{R}^d)$, classical properties of convergence of the mollifiers give

$$\|K_\varepsilon * F_0(t,\cdot) - F_0(t,\cdot)\|_1 \to 0, \quad \|K_\varepsilon * \nabla F_0(t,\cdot) - \nabla F_0(t,\cdot)\|_1 \to 0, \tag{4.34}$$

and

$$\|K_\varepsilon * p(s,Y_s,t,\cdot) - p(s,Y_s,t,\cdot)\|_1 \to 0, \quad \|K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot) - \nabla_x p(s,Y_s,t,\cdot)\|_1 \to 0, \ a.s.\,. \tag{4.35}$$

Moreover, by inequality (7.17) of Lemma 7.4, there exists a constant $C := C(C_u, c_u) > 0$ such that for $0 \le s < t \le T$,

$$\|K_\varepsilon * p(s,Y_s,t,\cdot) - p(s,Y_s,t,\cdot)\|_1 + \|K_\varepsilon * \nabla_x p(s,Y_s,t,\cdot) - \nabla_x p(s,Y_s,t,\cdot)\|_1 \le 2C(1 + \frac{1}{\sqrt{t-s}})\ a.s.\,. \tag{4.36}$$

Lebesgue dominated convergence theorem then implies that the third and fourth terms in the r.h.s. of (4.33) converge to $0$ when $\varepsilon$ goes to $0$. This ends the proof. □

**Proposition 4.7.** *We assume here that $(K_\varepsilon)_{\varepsilon>0}$ is explicitly given by*

$$K_\varepsilon(x) := \frac{1}{\varepsilon^d} K(\frac{x}{\varepsilon})\,, \tag{4.37}$$

*with $K \ge 0$ satisfying*

$$\int_{\mathbb{R}^d} K(x)\,dx = 1\,, \quad \int_{\mathbb{R}^d} x\,K(x)\,dx = 0 \quad and \quad \kappa := \frac{1}{2}\int_{\mathbb{R}^d} |x|\,K(x)\,dx < \infty. \tag{4.38}$$

*Let $u^\varepsilon$ be the real-valued function defined by*(4.17)*, (such that for any $t \in [0,T]$, $u^\varepsilon(t,\cdot) := K_\varepsilon * \gamma^\varepsilon_t$) with $K_\varepsilon$ given by* (4.37)*. Under Assumption 2 and in the particular case where the function $\Lambda(t,x,u)$ does not depend on the gradient $\nabla u$, there exists a constant $C := C(\kappa, C_u, c_u) > 0$ such that, for all $t \in (0,T]$,*

$$\|u^\varepsilon(t,\cdot) - u(t,\cdot)\|_1 \le \ \varepsilon C\big(\tfrac{1}{\sqrt{t}} + 2\sqrt{t}\big), \tag{4.39}$$

*with $C_u$, $c_u$ denoting the constants given by* (7.17) *(only depending on $\Phi$, $g$).*

*Proof.* This proof is based on the same arguments as the ones used in the proof of Theorem 4.5 since in the present case, $\Lambda$ only depends on $(t,x,u)$ and not on $\nabla u$. In particular, we obtain for $t \in ]0,T]$,

$$\|u^\varepsilon(t,\cdot) - u(t,\cdot)\|_1 \le Ce^{CT}\|K_\varepsilon * F_0(t,\cdot) - F_0(t,\cdot)\|_1 + \int_0^t \mathbb{E}\Big[\|K_\varepsilon * p(s,Y_s,t,\cdot) - p(s,Y_s,t,\cdot)\|_1\Big]ds, \tag{4.40}$$

that corresponds to inequality (4.33) in the proof above, without the term containing the gradient $\nabla u$. Invoking inequality (7.17) of Lemma 7.4 and inequality (7.9) of Lemma 7.3 with $H = K$, and successively with $f = F_0(t,\cdot)$ and $f = p(s,y,t,\cdot)$ for fixed $y \in \mathbb{R}^d$, imply that

$$\|K_\varepsilon * F_0(t,\cdot) - F_0(t,\cdot)\|_1 \le \frac{\varepsilon\kappa\mathfrak{C}}{\sqrt{t}}\,, \quad 0 < t \le T, \tag{4.41}$$

and

$$\|K_\varepsilon * p(s,Y_s,t,\cdot) - p(s,Y_s,t,\cdot)\|_1 \le \frac{\varepsilon\kappa\mathfrak{C}}{\sqrt{t-s}}\,\text{, a.e,} \quad 0 \le s < t < T\,, \tag{4.42}$$

with $\mathfrak{C} = \mathfrak{C}(C_u, c_u)$. This concludes the proof of (4.39). □

# 5 Particles system algorithm

To simplify notations in the rest of the paper, $f_t$ will denote $f(t)$ where $f : [0,T] \to E$ is an $E$-valued Borel function and $(E, d_E)$ a metric space.
In previous sections, we have studied existence, uniqueness for a semilinear PDE of the form (1.1) and we have established a Feynman-Kac type representation for the corresponding solution $u$, see Theorem 3.5. The regularized form of (1.1) is the integro-PDE (4.2) for which we have established well-posedness in Proposition 4.4 1. In the sequel, we denote by $\gamma^\varepsilon$ the corresponding solution and again by $u^\varepsilon(t,x) := (K_\varepsilon * \gamma^\varepsilon_t)(x)$ (see (4.17)). We recall that $u^\varepsilon$ converges to $u$, when the regularization parameter $\varepsilon$ vanishes to $0$, see Theorem 4.5. In the present section, we propose a Monte Carlo approximation $u^{\varepsilon,N}$ of $u^\varepsilon$, providing an original numerical approximation of the semilinear PDE (1.1), when both the number of particles $N \to \infty$ and the regularization parameter $\varepsilon \to 0$ with a judicious relative rate. Let $u_0$ be a Borel probability measure on $\mathcal{P}(\mathbb{R}^d)$.

## 5.1 Convergence of the particle system

We suppose the validity of Assumption 2.
For fixed $N \in \mathbb{N}^\star$, let $(W^i)_{i=1,\cdots,N}$ be a family of independent Brownian motions and $(Y^i_0)_{i=1,\cdots,N}$ be i.i.d. random variables distributed according to $u_0(x)dx$. For any $\varepsilon > 0$, we define the measure-valued functions $(\gamma^{\varepsilon,N}_t)_{t\in[0,T]}$ such that for any $t \in [0,T]$

$$\begin{cases} \xi^i_t = \xi^i_0 + \int_0^t \Phi(s,\xi^i_s)dW^i_s + \int_0^t g(s,\xi^i_s)ds \ , \quad \text{for } i = 1,\cdots,N \ , \\ \xi^i_0 = Y^i_0 \quad \text{for } i = 1,\cdots,N \ , \\ \gamma^{\varepsilon,N}_t = \dfrac{1}{N}\displaystyle\sum_{i=1}^N V_t\big(\xi^i, (K_\varepsilon * \gamma^{\varepsilon,N})(\xi^i), (\nabla K_\varepsilon * \gamma^{\varepsilon,N})(\xi^i)\big)\delta_{\xi^i_t} \ . \end{cases} \tag{5.1}$$

where $(K_\varepsilon)_{\varepsilon>0}$ denotes a sequence of mollifiers such that for all $\varepsilon > 0$, $K_\varepsilon \in W^{1,1}(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d)$ and we recall that $V_t$ is given by (2.1). The first line of (5.1) is a $d$-dimensional classical SDE whose strong existence and pathwise uniqueness are ensured by classical theorems for Lipschitz coefficients. Moreover $\xi^i, i = 1,\cdots,N$ are i.i.d. In the following lemma, we prove by a fixed-point argument that the third line equation of (5.1) has a unique solution.

**Lemma 5.1.** *We suppose the validity of Assumption 2. Let us fix $\varepsilon > 0$ and $N \in \mathbb{N}^\star$. Consider the i.i.d. system $(\xi^i)_{i=1,\cdots,N}$ of particles, solution of the two first equations of* (5.1). *Then, there exists a unique function $\gamma^{\varepsilon,N}$ : $[0,T] \to \mathcal{M}_f(\mathbb{R}^d)$ such that for all $t \in [0,T]$, $\gamma^{\varepsilon,N}_t$ is solution of* (5.1).

*Proof.* The proof relies on a fixed-point argument applied to the map $T^{\varepsilon,N} : \mathcal{C}([0,T], \mathcal{M}_f(\mathbb{R}^d)) \to \mathcal{C}([0,T], \mathcal{M}_f(\mathbb{R}^d))$ given by

$$T^{\varepsilon,N}(\gamma)(t) : \gamma \mapsto \frac{1}{N}\sum_{i=1}^N V_t\big(\xi^i, (K_\varepsilon * \gamma)(\xi^i), (\nabla K_\varepsilon * \gamma)(\xi^i)\big)\delta_{\xi^i_t} \ . \tag{5.2}$$

In the rest of the proof, the notation $T^{\varepsilon,N}_t(\gamma)$ will denote $T^{\varepsilon,N}(\gamma)(t)$.
In order to apply the Banach fixed-point theorem, we emphasize that $\mathcal{C}\big([0,T], \mathcal{M}_f(\mathbb{R}^d)\big)$ is equipped with one of the equivalent norms $\|\cdot\|_{TV,\lambda}$, $\lambda \geq 0$, defined by

$$\|\gamma\|_{TV,\lambda} := \sup_{t\in[0,T]} e^{-\lambda t}\|\gamma(t,\cdot)\|_{TV} \ , \tag{5.3}$$

and for which $\big(\mathcal{C}([0,T],\mathcal{M}_f(\mathbb{R}^d)),\|\cdot\|_{TV,\lambda}\big)$ is still complete.
From now on, it remains to ensure that $T^{\varepsilon,N}$ is indeed a contraction with respect $\|\gamma\|_{TV,\lambda}$ for some $\lambda$. To simplify notations, we set for all $i\in\{1,\cdots,N\}$,

$$T_t^{\varepsilon,N,i}(\gamma) := V_t\big(\xi^i,(K_\varepsilon*\gamma)(\xi^i),(\nabla K_\varepsilon*\gamma)(\xi^i)\big)\ ,\quad (t,\gamma)\in[0,T]\times\mathcal{C}\big([0,T],\mathcal{M}_f(\mathbb{R}^d)\big)\ , \tag{5.4}$$

to re-write (5.2) in the form

$$T_t^{\varepsilon,N}(\gamma) = \frac{1}{N}\sum_{i=1}^N T_t^{\varepsilon,N,i}(\gamma)\delta_{\xi_t^i}\ ,\quad (t,\gamma)\in[0,T]\times\mathcal{C}\big([0,T],\mathcal{M}_f(\mathbb{R}^d)\big)\ . \tag{5.5}$$

Let $\lambda>0$. Consider now $\gamma^1,\gamma^2\in\mathcal{C}\big([0,T],\mathcal{M}_f(\mathbb{R}^d)\big)$. On the one hand, taking into account (2.1) and (2.3), for all $t\in[0,T]$, $i\in\{1,\cdots,N\}$, we have

$$\begin{aligned}
|T_t^{\varepsilon,N,i}(\gamma^1)-T_t^{\varepsilon,N,i}(\gamma^2)| &\le L_\Lambda e^{TM_\Lambda}\int_0^t\Big(|(K_\varepsilon*\gamma^1)(\xi_s^i)-(K_\varepsilon*\gamma^2)(\xi_s^i)|\\
&\quad + |(\nabla K_\varepsilon*\gamma^1)(\xi_s^i)-(\nabla K_\varepsilon*\gamma^2)(\xi_s^i)\Big)ds\\
&\le C\int_0^t\|\gamma_s^1-\gamma_s^2\|_{TV}ds\\
&\le C\int_0^t e^{s\lambda}\ \|\gamma^1-\gamma^2\|_{TV,\lambda}\ ds\\
&= C\|\gamma^1-\gamma^2\|_{TV,\lambda}\ \frac{e^{\lambda t}-1}{\lambda}\ ,
\end{aligned} \tag{5.6}$$

with $C=C(\|K_\varepsilon\|_\infty,\|\nabla K_\varepsilon\|_\infty,L_\Lambda,M_\Lambda)$. It follows that

$$\begin{aligned}
\|T^{\varepsilon,N}(\gamma^1)-T^{\varepsilon,N}(\gamma^2)\|_{TV,\lambda} &\le \frac{1}{N}\sum_{i=1}^N\|T^{\varepsilon,N,i}(\gamma^1)\delta_{\xi^i}-T^{\varepsilon,N,i}(\gamma^2)\delta_{\xi^i}\|_{TV,\lambda}\\
&\le \frac{C}{\lambda}\|\gamma^1-\gamma^2\|_{TV,\lambda}\ .
\end{aligned} \tag{5.7}$$

By taking $\lambda>C$ and invoking Banach fixed-point theorem, we end the proof. □

After the preceding preliminary considerations, we can state and prove the main result of the section.

**Proposition 5.2.** *We suppose the validity of Assumption 2. Assume that the kernel $K$ is a probability density verifying*

$$K\in W^{1,1}(\mathbb{R}^d)\cap W^{1,\infty}(\mathbb{R}^d)\ ,\quad \int_{\mathbb{R}^d}|x|^{d+1}\ K(x)dx<\infty\ ,\quad \textit{and}\quad \int_{\mathbb{R}^d}|x|^{d+1}\ |\nabla K(x)|dx<\infty\ . \tag{5.8}$$

*For $\varepsilon>0$, we suppose that $K_\varepsilon$ is explicitly given by* (4.37) *with $K$ satisfying* (5.8). *Let $u^\varepsilon$ be the real valued function defined by* (4.17), *and $u^{\varepsilon,N}$ such that for any $t\in[0,T]$,*

$$u^{\varepsilon,N}(t,\cdot) := K_\varepsilon*\gamma_t^{\varepsilon,N}\ , \tag{5.9}$$

*where $\gamma^{\varepsilon,N}$ is defined by the third line of* (5.1). *There exists a finite positive constant $C$ (depending on $L_\Lambda$, $M_\Lambda$, $K$) such that for all $t\in[0,T]$ and $N\in\mathbb{N}^*$, $\varepsilon>0$ verifying min$(N,N\varepsilon^d)>C$ we have*

$$\mathbb{E}\Big[\|u_t^{\varepsilon,N}-u_t^\varepsilon\|_1\Big]+\mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N}-\nabla u_t^\varepsilon\|_1\Big]\le\frac{C}{\sqrt{N\varepsilon^{d+4}}}e^{\frac{C}{\varepsilon^{d+1}}}\ . \tag{5.10}$$

**Remark 5.3.** *For all $t \in [0,T]$, $N \in \mathbb{N}^\star$, $\varepsilon > 0$, the following bound takes place:*

$$\begin{aligned}\mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t\|_1\Big] &\leq \mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t^\varepsilon\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t^\varepsilon\|_1\Big] \\ &\quad + \|u_t^\varepsilon - u_t\|_1 + \|\nabla u_t^\varepsilon - \nabla u_t\|_1 \\ &\leq \frac{C}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} + \|u_t^\varepsilon - u_t\|_1 + \|\nabla u_t^\varepsilon - \nabla u_t\|_1 \,.\end{aligned} \tag{5.11}$$

*where we have used Proposition 5.2 for the latter inequality. Taking into account Theorem 4.5 above, it appears clearly that the convergence of $u^{\varepsilon,N}$ (resp. $\nabla u^{\varepsilon,N}$) to $u$ (resp. $\nabla u$) will hold as soon as $\frac{1}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} \to 0$ when $\varepsilon \to 0$, $N \to +\infty$. This requires a "trade-off condition" between the speed of convergence of $N$ and $\varepsilon$. Setting $\Phi(\varepsilon) := \varepsilon^{-(d+4)} e^{\frac{2C}{\varepsilon^{d+1}}}$, the trade-off condition can be formulated as*

$$\frac{1}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} \to 0 \iff \frac{\Phi(\varepsilon)}{N} \to 0 \quad \textit{when} \quad \varepsilon \to 0,\ N \to +\infty. \tag{5.12}$$

*An example of such trade-off between $N$ and $\varepsilon$ can be given by the relation $\varepsilon(N) = (\frac{1}{log(N)})^{\frac{1}{d+4}}$.*

*Proof.* Let us fix $\varepsilon > 0$, $N \in \mathbb{N}^\star$. For any $\ell = 1, \cdots, d$, we introduce the real-valued function $G_\varepsilon^\ell$ defined on $\mathbb{R}^d$ such that

$$G_\varepsilon^\ell(x) := \frac{1}{\varepsilon^d} \frac{\partial K}{\partial x_\ell} \left(\frac{x}{\varepsilon}\right) , \quad \text{for almost all } \ x \in \mathbb{R}^d \,. \tag{5.13}$$

By (5.8), there exists a finite positive constant $C$ independent of $\varepsilon$ such that $\|G_\varepsilon^\ell\|_\infty \leq \frac{C}{\varepsilon^d}$ and $\|G_\varepsilon^\ell\|_1 = \|G_1^\ell\|_1 \leq C$. In the sequel, $C$ will always denote a finite positive constant independent of $(\varepsilon, N)$ that may change from line to line. For any $t \in [0,T]$, we introduce the random Borel measure $\tilde{\gamma}_t^{\varepsilon,N}$ on $\mathbb{R}^d$, defined by

$$\tilde{\gamma}_t^{\varepsilon,N} := \frac{1}{N} \sum_{i=1}^N V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big)\, \delta_{\xi_t^i} \,. \tag{5.14}$$

One can first decompose the error on the l.h.s of inequality (5.10) as follows

$$\begin{aligned}\mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t^\varepsilon\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t^\varepsilon\|_1\Big] &= \mathbb{E}\Big[\|K_\varepsilon * (\gamma_t^{\varepsilon,N} - \gamma_t^\varepsilon)\|_1\Big] + \frac{1}{\varepsilon}\sum_{\ell=1}^d \mathbb{E}\Big[\|G_\varepsilon^\ell * (\gamma_t^{\varepsilon,N} - \gamma_t^\varepsilon)\|_1\Big] \\ &\leq \mathbb{E}\Big[\|K_\varepsilon * (\gamma_t^{\varepsilon,N} - \tilde{\gamma}_t^{\varepsilon,N})\|_1\Big] + \frac{1}{\varepsilon}\sum_{\ell=1}^d \mathbb{E}\Big[\|G_\varepsilon^\ell * (\gamma_t^{\varepsilon,N} - \tilde{\gamma}_t^{\varepsilon,N})\|_1\Big] \\ &\quad + \mathbb{E}\Big[\|K_\varepsilon * (\tilde{\gamma}_t^{\varepsilon,N} - \gamma_t^\varepsilon)\|_1\Big] + \frac{1}{\varepsilon}\sum_{\ell=1}^d \mathbb{E}\Big[\|G_\varepsilon^\ell * (\tilde{\gamma}_t^{\varepsilon,N} - \gamma_t^\varepsilon)\|_1\Big] \\ &= \mathbb{E}\Big[\|A_t^{\varepsilon,N}\|_1\Big] + \mathbb{E}\Big[\|A_t'^{\varepsilon,N}\|_1\Big] + \mathbb{E}\Big[\|B_t^{\varepsilon,N}\|_1\Big] + \mathbb{E}\Big[\|B_t'^{\varepsilon,N}\|_1\Big] ,\end{aligned} \tag{5.15}$$

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

$$\begin{cases} A_t^{\varepsilon,N}(x) := \frac{1}{N}\sum_{i=1}^N K_\varepsilon(x - \xi_t^i)\Big[V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) - V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big)\Big] \\ A_t'^{\varepsilon,N}(x) := \frac{1}{\varepsilon}\sum_{\ell=1}^d \left|\frac{1}{N}\sum_{i=1}^N G_\varepsilon^\ell(x - \xi_t^i)\Big[V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) - V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big)\Big]\right| \\ B_t^{\varepsilon,N}(x) := \frac{1}{N}\sum_{i=1}^N K_\varepsilon(x - \xi_t^i) V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big) - \mathbb{E}\Big[K_\varepsilon(x - \xi_t^1) V_t\big(\xi^1, u^\varepsilon(\xi^1), \nabla u^\varepsilon(\xi^1)\big)\Big] \\ B_t'^{\varepsilon,N}(x) := \frac{1}{\varepsilon}\sum_{\ell=1}^d \left|\frac{1}{N}\sum_{i=1}^N G_\varepsilon^\ell(x - \xi_t^i) V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big) - \mathbb{E}\Big[G_\varepsilon^\ell(x - \xi_t^1) V_t\big(\xi^1, u^\varepsilon(\xi^1), \nabla u^\varepsilon(\xi^1)\big)\Big]\right| . \end{cases} \tag{5.16}$$

We will proceed in two steps, first bounding $\mathbb{E}\Big[\|B_t^{\varepsilon,N}\|_1\Big]$ and $\mathbb{E}\Big[\|B_t'^{\varepsilon,N}\|_1\Big]$ and then $\mathbb{E}\Big[\|A_t^{\varepsilon,N}\|_1\Big]$ and $\mathbb{E}\Big[\|A_t'^{\varepsilon,N}\|_1\Big]$.

**Step 1: Bounding $\mathbb{E}\|B_t^{\varepsilon,N}\|_1$ and $\mathbb{E}\|B_t'^{\varepsilon,N}\|_1$.** For any $i \in \{1,\cdots,N\}$ and $(t,x) \in [0,T]\times\mathbb{R}^d$ we set

$$P_i^\varepsilon(t,x) := K_\varepsilon(x-\xi_t^i)V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big) - \mathbb{E}\Big[K_\varepsilon(x-\xi_t^i)V_t\big(\xi^i, u^\varepsilon(\xi^i), \nabla u^\varepsilon(\xi^i)\big)\Big]\,. \tag{5.17}$$

Notice that for fixed $(t,x) \in [0,T]\times\mathbb{R}^d$, $(P_i^\varepsilon(t,x))_{i=1,\cdots,N}$ are i.i.d. centered square integrable random variables. Hence using Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\mathbb{E}\Big[\|B_t^{\varepsilon,N}\|_1\Big] &= \int_{\mathbb{R}^d}\mathbb{E}\Big[\Big|\frac{1}{N}\sum_{i=1}^N P_i^\varepsilon(t,x)\Big|\Big]dx \\
&\leq \int_{\mathbb{R}^d}\sqrt{\mathbb{E}\Big[\Big(\frac{1}{N}\sum_{i=1}^N (P_i^\varepsilon(t,x))\Big)^2\Big]}dx \\
&= \frac{1}{\sqrt{N}}\int_{\mathbb{R}^d}\sqrt{\mathbb{E}\Big[(P_1^\varepsilon(t,x))^2\Big]}dx\,.
\end{aligned} \tag{5.18}$$

By the boundedness assumption on $\Lambda$ (item 5. of Assumption 2),

$$\mathbb{E}[(P_1^\varepsilon(t,x))^2] \leq 4e^{2M_\Lambda T}\mathbb{E}[(K_\varepsilon(x-\xi_t^1))^2]\,,$$

which implies

$$\mathbb{E}\Big[\|B_t^{\varepsilon,N}\|_1\Big] \leq \frac{C}{\sqrt{N}}\int_{\mathbb{R}^d}\sqrt{\mathbb{E}\Big[(K_\varepsilon(x-\xi_t^1))^2\Big]}dx = \frac{C}{\sqrt{N}}\frac{\sqrt{\int_{\mathbb{R}^d}K^2(x)dx}}{\sqrt{\varepsilon^d}}\int_{\mathbb{R}^d}\sqrt{H_\varepsilon * p_t}(x)dx\,, \tag{5.19}$$

where $p_t$ is the law density of $\xi_t^1$ and $H_\varepsilon$ is the probability density on $\mathbb{R}^d$ such that for almost all $x \in \mathbb{R}^d$, $H_\varepsilon(x) := \frac{1}{\int_{\mathbb{R}^d}K^2(x)dx}\frac{1}{\varepsilon^d}K^2(\frac{x}{\varepsilon})$, which is well-defined thanks to assumption (5.8). Finally, applying Lemma 7.2 with $G = \frac{K^2}{\|K\|_2^2}$ and $f = p_t$ we obtain

$$\mathbb{E}\Big[\|B_t^{\varepsilon,N}\|_1\Big] \leq \frac{C}{\sqrt{N\varepsilon^d}}\,, \quad \text{for } \varepsilon \text{ small enough.} \tag{5.20}$$

Proceeding similarly for the term $B_t'^{\varepsilon,N}$ leads to

$$\mathbb{E}\Big[\|B_t'^{\varepsilon,N}\|_1\Big] \leq \frac{C}{\sqrt{N\varepsilon^2}}\sum_{\ell=1}^d\int_{\mathbb{R}^d}\sqrt{\mathbb{E}\Big[(G_\varepsilon^\ell(x-\xi_t^1))^2\Big]}dx = \frac{C}{\sqrt{N\varepsilon^2}}\sum_{\ell=1}^d\frac{\sqrt{\int_{\mathbb{R}^d}|\frac{\partial K}{\partial x_\ell}(x)|^2dx}}{\sqrt{\varepsilon^d}}\int_{\mathbb{R}^d}\sqrt{H_\varepsilon^\ell * p_t}(x)dx\,, \tag{5.21}$$

where $H_\varepsilon^\ell, \ell = 1,\cdots,d$ denotes the probability densities on $\mathbb{R}^d$ such that for almost all $x \in \mathbb{R}^d$, $H_\varepsilon^\ell(x) := \frac{1}{\int_{\mathbb{R}^d}|\frac{\partial K}{\partial x_\ell}(x)|^2dx}\frac{1}{\varepsilon^d}|\frac{\partial K}{\partial x_\ell}(\frac{x}{\varepsilon})|^2$. Applying again Lemma 7.2 with $G = \frac{|\frac{\partial K}{\partial x_\ell}|^2}{\|\frac{\partial K}{\partial x_\ell}\|_2^2}, \ell = 1,\cdots,d$ and $f = p_t$ we obtain

$$\mathbb{E}\Big[\|B_t'^{\varepsilon,N}\|_1\Big] \leq \frac{C}{\sqrt{N\varepsilon^{d+2}}}\,, \quad \text{for } \varepsilon \text{ small enough.} \tag{5.22}$$

**Step 2: Bounding $\mathbb{E}\|A_t^{\varepsilon,N}\|_1$ and $\mathbb{E}\|A_t'^{\varepsilon,N}\|_1$.** Recall that $A_t^{\varepsilon,N}(x) = K_\varepsilon * (\gamma_t^{\varepsilon,N} - \tilde{\gamma}_t^{\varepsilon,N})(x)$ and $A_t'^{\varepsilon,N}(x) = \frac{1}{\varepsilon}\sum_{\ell=1}^d|G_\varepsilon^\ell * (\gamma_t^{\varepsilon,N} - \tilde{\gamma}_t^{\varepsilon,N})|(x)$, which yields

$$\mathbb{E}\Big[\|A_t^{\varepsilon,N}\|_1\Big] + \mathbb{E}\Big[\|A_t'^{\varepsilon,N}\|_1\Big] \leq \frac{C}{\varepsilon}\mathbb{E}\Big[\|\gamma_t^{\varepsilon,N} - \tilde{\gamma}_t^{\varepsilon,N}\|_{TV}\Big]\,. \tag{5.23}$$

We are now interested in bounding the r.h.s. of (5.23).
Recalling (5.1), (5.14) and inequality (2.3), we have

$$
\begin{aligned}
\mathbb{E}\Big[\|\gamma_t^{\varepsilon,N}-\tilde{\gamma}_t^{\varepsilon,N}\|_{TV}\Big] &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\Big[|V_t(\xi^i,u^{\varepsilon,N},\nabla u^{\varepsilon,N})-V_t(\xi^i,u^{\varepsilon},\nabla u^{\varepsilon})|\Big] \\
&\leq C\mathbb{E}\left[\int_0^t \left(|u_s^{\varepsilon,N}-u_s^{\varepsilon}|(\xi_s^1)+|\nabla u_s^{\varepsilon,N}-\nabla u_s^{\varepsilon}|(\xi_s^1)\right)ds\right] \\
&\leq C\int_0^t \left(\mathbb{E}\Big[|K_\varepsilon*(\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N})(\xi_s^1)|\Big]+\mathbb{E}\Big[|K_\varepsilon*(\tilde{\gamma}_s^{\varepsilon,N}-\gamma_s^{\varepsilon})(\xi_s^1)|\Big]ds\right)\ , \\
&\quad +\frac{C}{\varepsilon}\sum_{\ell=1}^{d}\int_0^t \left(\mathbb{E}\Big[|G_\varepsilon^\ell*(\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N})(\xi_s^1)|\Big]+\mathbb{E}\Big[|G_\varepsilon^\ell*(\tilde{\gamma}_s^{\varepsilon,N}-\gamma_s^{\varepsilon})(\xi_s^1)|\Big]\right)ds.
\end{aligned}
\tag{5.24}
$$

By inequality (7.18) in Lemma 7.4, there exists a finite constant $C>0$ such that $\|p_s\|_\infty \leq C\|u_0\|_\infty$ for all $s\in[0,T]$. Thus using inequality (5.20), we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[|K_\varepsilon*(\tilde{\gamma}_s^{\varepsilon,N}-\gamma_s^{\varepsilon})(\xi_s^1)|\Big] \\
&\leq \frac{1}{N}\mathbb{E}\Big[\,\big|K_\varepsilon(0)V_s(\xi^1,u^\varepsilon(\xi^1),\nabla u^\varepsilon(\xi^1))-\mathbb{E}[K_\varepsilon(\xi_s^1-\xi_s^2)V_s(\xi^2,u^\varepsilon(\xi^2),\nabla u^\varepsilon(\xi^2))\,|\,\xi^2]\big|\,\Big] \\
&+\frac{N-1}{N}\frac{1}{N-1}\int_{\mathbb{R}^d}\left|\sum_{i=2}^{N}\Big[K_\varepsilon(x-\xi_s^i)V_s(\xi^i,u^\varepsilon(\xi^i),\nabla u^\varepsilon(\xi^i))-\mathbb{E}[K(x-\xi_s^i)V_s(\xi^i,u^\varepsilon(\xi^i),\nabla u^\varepsilon(\xi^i))]\Big]\right|p_s(x)dx \\
&\leq \frac{C}{N\varepsilon^d}+\frac{N-1}{N}\frac{C}{\sqrt{(N-1)\varepsilon^d}} \\
&\leq \frac{C}{\sqrt{N\varepsilon^d}} \qquad \text{(for} \quad (N \text{ and } N\varepsilon^d) \quad \text{sufficiently large)},\ s\in[0,T]\ .
\end{aligned}
\tag{5.25}
$$

Similarly we get

$$
\sum_{\ell=1}^{d}\mathbb{E}\Big[|\frac{1}{\varepsilon}G_\varepsilon^\ell*(\tilde{\gamma}_s^{\varepsilon,N}-\gamma_s^{\varepsilon})(\xi_s^1)|\Big]\leq \frac{C}{\sqrt{N\varepsilon^{d+2}}},\quad s\in[0,T]\ . \tag{5.26}
$$

Moreover, for all $s\in[0,T]$, the boundedness of $|K|$ and $|\nabla K|$ implies

$$
\mathbb{E}\Big[|K_\varepsilon*(\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N})(\xi_s^1)|\Big]+\sum_{\ell=1}^{d}\mathbb{E}\Big[|\frac{1}{\varepsilon}G_\varepsilon^\ell*(\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N})(\xi_s^1)|\Big]\leq \frac{C}{\varepsilon^{d+1}}\Big[\|\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N}\|_{TV}\Big]. \tag{5.27}
$$

Injecting inequalities (5.25) (5.26) and (5.27) into (5.24) gives

$$
\mathbb{E}\Big[\|\gamma_t^{\varepsilon,N}-\tilde{\gamma}_t^{\varepsilon,N}\|_{TV}\Big]\leq \frac{C}{\sqrt{N\varepsilon^{d+2}}}+\frac{C}{\varepsilon^{d+1}}\int_0^t\mathbb{E}\Big[\|\gamma_s^{\varepsilon,N}-\tilde{\gamma}_s^{\varepsilon,N}\|_{TV}\Big]ds\ .
$$

By Gronwall's lemma we obtain $\mathbb{E}\Big[\|\gamma_t^{\varepsilon,N}-\tilde{\gamma}_t^{\varepsilon}\|_{TV}\Big]\leq \frac{C}{\sqrt{N\varepsilon^{d+2}}}e^{\frac{C}{\varepsilon^{d+1}}}$ , which together with (5.23) completes the proof by implying the inequality

$$
\mathbb{E}[\|A_t^{\varepsilon,N}\|_1]+\mathbb{E}[\|A_t'^{\varepsilon,N}\|_1]\leq \frac{C}{\sqrt{N\varepsilon^{d+4}}}e^{\frac{C}{\varepsilon^{d+1}}}\ . \tag{5.28}
$$

□

As a straightforward consequence of Proposition 5.2 and Theorem 4.5, the corollary below follows.

**Corollary 5.4.** *Assume that the same assumptions as in Proposition 5.2 are fulfilled.*
*If* $\varepsilon \to 0$, $N \to +\infty$ *such that* $\frac{1}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} \to 0$ , *(where* $C$ *is the constant coming from Proposition 5.2) then*

$$\mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t\|_1\Big] \longrightarrow 0 \, . \tag{5.29}$$

*Proof.* Let us fix $\varepsilon > 0$, $N \in \mathbb{N}^\star$, $t \in [0,T]$. The proof is based on the bound

$$\begin{aligned} \mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t\|_1\Big] &\leq \mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t^\varepsilon\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t^\varepsilon\|_1\Big] \\ &\quad + \|u_t^\varepsilon - u_t\|_1 + \|\nabla u_t^\varepsilon - \nabla u_t\|_1 \, , \\ &\leq \frac{C}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} + \|u_t^\varepsilon - u_t\|_1 + \|\nabla u_t^\varepsilon - \nabla u_t\|_1 \, , \end{aligned} \tag{5.30}$$

where we have used Proposition 5.2 for the second inequality above. Moreover, Theorem 4.5 gives $\|u_t^\varepsilon - u_t\|_1 + \|\nabla u_t^\varepsilon - \nabla u_t\|_1 \longrightarrow 0$ for $\varepsilon \to 0$. This concludes the proof of the corollary. □

## 5.2 Time discretized scheme

In this section, we will make use of the assumptions below.

**Assumption 3.** *All items of Assumption 2 are in force excepted 1. and 4. which are replaced by the following.*

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

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

*and*

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

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

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

*We also add the following item.*

7. $(K_\varepsilon)_{\varepsilon>0}$ *denotes a sequence of mollifiers explicitly given by* $K_\varepsilon(x) := \frac{1}{\varepsilon^d} K(\frac{x}{\varepsilon})$, *where the kernel* $K$ *is a probability density belonging to* $W^{1,1}(\mathbb{R}^d) \cap W^{1,\infty}(\mathbb{R}^d)$.

We assume the validity of Assumption 3. For $n \in \mathbb{N}^\star$, we set $\delta t = T/n$ and introduce the time grid $\big(0 = t_0 < \cdots < t_k = k\delta t < \cdots < t_n = T\big)$. For any $N \in \mathbb{N}^\star$, $\varepsilon > 0$ and $n \in \mathbb{N}^*$, we define the measure-valued functions $(\bar{\gamma}_t^{\varepsilon,N,n})_{t\in[0,T]}$ such that for any $t \in [0,T]$,

$$\left\{ \begin{array}{l} \bar{\xi}_t^i = \bar{\xi}_0^i + \int_0^t \Phi(r(s), \bar{\xi}_{r(s)}^i) dW_s^i + \int_0^t g(r(s), \bar{\xi}_{r(s)}^i) ds \, , \text{ for } i = 1, \cdots, N, \\ \bar{\xi}_0^i = Y_0^i \quad \text{for } i = 1, \cdots, N \, , \\ \bar{\gamma}_t^{\varepsilon,N,n} = \dfrac{1}{N} \displaystyle\sum_{i=1}^N \bar{V}_t\big(\bar{\xi}^i, (K_\varepsilon * \bar{\gamma}^{\varepsilon,N,n})(\bar{\xi}^i), (\nabla K_\varepsilon * \bar{\gamma}^{\varepsilon,N,n})(\bar{\xi}^i)\big) \delta_{\bar{\xi}_t^i} \, , \end{array} \right. \tag{5.31}$$

where for $(t,x,y,z) \in [0,T] \times \mathcal{C}^d \times \mathcal{C} \times \mathcal{C}^d$,

$$\bar{V}_t\big(x,y,z\big) \quad := \quad \exp\Big\{ \int_0^t \Lambda(r(s), x_{r(s)}, y_{r(s)}, z_{r(s)}) \, ds \Big\} \, , \tag{5.32}$$

and $r : s \in [0,T] \mapsto r(s) \in \{t_0, \cdots, t_n\}$ is the piecewise constant function such that $r(s) = t_k$ when $s \in [t_k, t_{k+1})$. In the sequel, we will often use the notation $\bar{\gamma}_t^{\varepsilon,N}$ instead of $\bar{\gamma}_t^{\varepsilon,N,n}$ deliberately dropping the exponent $n$ to simplify the notation. The proposition below establishes the convergence of the time discretized scheme (5.31) to the continuous time version (5.1).

**Proposition 5.5.** *Suppose the validity of Assumption 3. The gradient $\nabla K$ of $K$ is also supposed to be Lipschitz with the corresponding constant $L_{\nabla K}$. For fixed parameters $\varepsilon > 0$, $N \in \mathbb{N}^\star$ and $n \in \mathbb{N}^\star$, we introduce $\bar{u}^{\varepsilon,N,n}$ such that for any $t \in [0,T]$,*

$$\bar{u}^{\varepsilon,N,n}(t,\cdot) := K_\varepsilon * \bar{\gamma}_t^{\varepsilon,N,n} \; . \tag{5.33}$$

*where $\bar{\gamma}_t^{\varepsilon,N,n}$ is defined by* (5.31). *Then*

$$\mathbb{E}\Big[\|u_t^{\varepsilon,N} - \bar{u}_t^{\varepsilon,N,n}\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla \bar{u}_t^{\varepsilon,N,n}\|_1\Big] \leq \frac{\bar{C}}{\varepsilon^{d+3}\sqrt{n}} e^{\frac{\bar{C}}{\varepsilon^{d+1}}} \; , \tag{5.34}$$

*where $\bar{C}$ is a finite, positive constant only depending on $M_\Phi$, $M_g$, $M_\Lambda$, $\|K\|_\infty$, $\|\nabla K\|_\infty$, $L_\Phi$, $L_g$, $L_\Lambda$, $L_{\nabla K}$, $T$.*

From Proposition 5.5 and Corollary 5.4 follows the result below.

**Theorem 5.6.** *Suppose the validity of Assumption 3. The gradient $\nabla K$ of $K$ is supposed to be Lipschitz with constant $L_{\nabla K}$. Suppose moreover that*

$$\int_{\mathbb{R}^d} |x|^{d+1} \; K(x)dx < \infty \; , \quad \textit{and} \quad \int_{\mathbb{R}^d} |x|^{d+1}| \; \nabla K(x)|dx < \infty \; . \tag{5.35}$$

*If $\varepsilon \to 0$, $n \to +\infty$ and $N \to +\infty$ such that*

$$\frac{1}{\sqrt{N\varepsilon^{d+4}}} e^{\frac{C}{\varepsilon^{d+1}}} \longrightarrow 0 \quad \textit{and} \quad \frac{1}{\varepsilon^{d+3}\sqrt{n}} e^{\frac{\bar{C}}{\varepsilon^{d+1}}} \longrightarrow 0 \; , \tag{5.36}$$

*where $C$ and $\bar{C}$ are constants respectively appearing in Proposition 5.2 and 5.5, then for all $t \in [0,T]$, the particle approximation $\bar{u}_t^{\varepsilon,N,n}$ defined by* (5.33) *converges to the unique solution, $u_t$, of* (1.1) *or equivalently* (3.15)*, in the sense that*

$$\mathbb{E}\Big[\|\bar{u}_t^{\varepsilon,N,n} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla \bar{u}_t^{\varepsilon,N,n} - \nabla u_t\|_1\Big] \longrightarrow 0 \; . \tag{5.37}$$

*Proof.* For all $N, n \in \mathbb{N}^\star$, $\varepsilon > 0$ and $t \in [0,T]$, we have

$$\begin{aligned} \mathbb{E}\Big[\|\bar{u}_t^{\varepsilon,N,n} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla \bar{u}_t^{\varepsilon,N,n} - \nabla u_t\|_1\Big] \quad \leq \quad & \mathbb{E}\Big[\|\bar{u}_t^{\varepsilon,N,n} - u_t^{\varepsilon,N}\|_1\Big] + \mathbb{E}\Big[\|\nabla \bar{u}_t^{\varepsilon,N,n} - \nabla u_t^{\varepsilon,N}\|_1\Big] \\ & + \mathbb{E}\Big[\|u_t^{\varepsilon,N} - u_t\|_1\Big] + \mathbb{E}\Big[\|\nabla u_t^{\varepsilon,N} - \nabla u_t\|_1\Big] \; . \end{aligned} \tag{5.38}$$

Inequality (5.34) of Proposition 5.5 and the second trade-off condition in (5.36) imply that the first two expectations in the r.h.s. of (5.38) converges to $0$.
By Corollary 5.4, we claim it is enough to state that the third and fourth expectations in the r.h.s. of (5.38) also converges to $0$. This concludes the proof. □

The proof of Proposition 5.5 above will be based on the following technical lemma, proved in the appendix.

**Lemma 5.7.** *We assume that the same assumptions as in Proposition 5.5 are fulfilled. Let $\varepsilon > 0$, $(N,n) \in (\mathbb{N}^\star)^2$ such that $\delta t = \frac{T}{n}$. Let $u^{\varepsilon,N}$ (resp. $\bar{u}^{\varepsilon,N}$) be the function defined by* (5.9) *(resp.* (5.33)*).*
*There exists a constant $C > 0$, only depending on $M_\Phi$, $M_g$, $M_\Lambda$, $\|K\|_\infty$, $\|\nabla K\|_\infty$, $L_\Lambda$, $L_{\nabla K}$ and $T$, such that for all $t \in [0,T]$, the following estimates hold.*

1. *For almost all $x, y \in \mathbb{R}^d$,*

$$|\bar{u}_t^{\varepsilon,N}(x) - \bar{u}_t^{\varepsilon,N}(y)| \leq \frac{C}{\varepsilon^{d+1}}|x-y| \quad \textit{and} \quad |\nabla \bar{u}_t^{\varepsilon,N}(x) - \nabla \bar{u}_t^{\varepsilon,N}(y)| \leq \frac{C}{\varepsilon^{d+2}}|x-y| \; . \tag{5.39}$$

2.
$$\mathbb{E}\Big[\|\bar{u}^{\varepsilon,N}_t - \bar{u}^{\varepsilon,N}_{r(t)}\|_\infty\Big] \le \frac{C\sqrt{\delta t}}{\varepsilon^{d+1}} \quad \textit{and} \quad \mathbb{E}\Big[\|\nabla\bar{u}^{\varepsilon,N}_t - \nabla\bar{u}^{\varepsilon,N}_{r(t)}\|_\infty\Big] \le \frac{C\sqrt{\delta t}}{\varepsilon^{d+2}}. \tag{5.40}$$

*Proof of Proposition 5.5.* In this proof, $C$ denotes a real positive constant (depending on $M_\Phi$, $M_g$, $M_\Lambda$, $\|K\|_\infty$, $\|\nabla K\|_\infty$, $L_\Phi$, $L_g$, $L_\Lambda$, $L_{\nabla K}$, $T$) that may change from line to line. Let us fix $\varepsilon > 0$, $N \in \mathbb{N}^\star$, $n \in \mathbb{N}^\star$.
Let us now prove inequality (5.34). To this end, $G^\ell_\varepsilon$ will again denote the real-valued functions defined on $\mathbb{R}^d$ by (5.13). It is easy to observe that there exists a constant $C > 0$ depending on $\|K\|_1, \|\frac{\partial K}{\partial x_\ell}\|_1, \ell = 1, \cdots, d$, such that
$$\|K_\varepsilon\|_1 + \sum_{\ell=1}^d \|G^\ell_\varepsilon\|_1 \le C\ , \tag{5.41}$$
and
$$\|K_\varepsilon\|_\infty + \sum_{\ell=1}^d \|G^\ell_\varepsilon\|_\infty \le \frac{C}{\varepsilon^d}\ . \tag{5.42}$$
From (5.9) and (5.33), we recall that $u^{\varepsilon,N}$ and $\bar{u}^{\varepsilon,N}$ are defined by
$$\forall\, t \in [0,T],\ u^{\varepsilon,N}_t = K_\varepsilon * \gamma^{\varepsilon,N}_t \quad \text{and} \quad \bar{u}^{\varepsilon,N}_t = K_\varepsilon * \bar{\gamma}^{\varepsilon,N}_t\ . \tag{5.43}$$
For all $t \in [0,T]$, we have
$$\begin{aligned}
\mathbb{E}\Big[\|u^{\varepsilon,N}_t - \bar{u}^{\varepsilon,N}_t\|_1\Big] + \mathbb{E}\Big[\|\nabla u^{\varepsilon,N}_t - \nabla\bar{u}^{\varepsilon,N}_t\|_1\Big] &\le \mathbb{E}\Big[\|K_\varepsilon * (\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t)\|_1\Big] + \frac{1}{\varepsilon}\sum_{l=1}^d \mathbb{E}\Big[\|G^\ell_\varepsilon * (\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t)\|_1\Big] \\
&\le \mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] + \frac{1}{\varepsilon}\sum_{\ell=1}^d \|G^\ell_\varepsilon\|_1 \mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] \\
&= \frac{C}{\varepsilon}\mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] \ \text{by (5.41)}\ . 
\end{aligned} \tag{5.44}$$
It follows, for $t \in [0,T]$,
$$\begin{aligned}
\mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] &= \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\Big|V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) - \bar{V}_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla u^{\varepsilon,N,n}(\bar{\xi}^i)\big)\Big|\Big] \\
&\le \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\Big|V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) - V_t\big(\xi^i, \bar{u}^{\varepsilon,N}(\xi^i), \nabla\bar{u}^{\varepsilon,N}(\xi^i)\big)\Big|\Big] \\
&\quad + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\Big|V_t\big(\xi^i, \bar{u}^{\varepsilon,N}(\xi^i), \nabla\bar{u}^{\varepsilon,N}(\xi^i)\big) - V_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla u^{\varepsilon,N}(\bar{\xi}^i)\big)\Big|\Big] \\
&\quad + \frac{1}{N}\sum_{i=1}^N \mathbb{E}\Big[\Big|V_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big) - \bar{V}_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla u^{\varepsilon,N}(\bar{\xi}^i)\big)\Big|\Big].
\end{aligned} \tag{5.45}$$

We are now interested in bounding each term in the r.h.s. of (5.45). Let us fix $t \in [0,T]$, $i \in \{1, \cdots, N\}$. Since $\Lambda$ is bounded and Lipschitz, inequality (2.3) implies
$$\begin{aligned}
A^{i,\varepsilon,N,n}_t &:= \mathbb{E}\Big[\Big|V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) - V_t\big(\xi^i, \bar{u}^{\varepsilon,N}(\xi^i), \nabla\bar{u}^{\varepsilon,N}(\xi^i)\big)\Big|\Big] \\
&\le e^{M_\Lambda T}\mathbb{E}\Big[\int_0^t \Big|\Lambda(s, \xi^i_s, u^{\varepsilon,N}_s(\xi^i_s), \nabla u^{\varepsilon,N}_s(\xi^i_s)) - \Lambda(s, \xi^i_s, \bar{u}^{\varepsilon,N}_s(\xi^i_s), \nabla\bar{u}^{\varepsilon,N}_s(\xi^i_s))\Big|\Big]\, ds \\
&\le e^{M_\Lambda T} L_\Lambda \int_0^t \Big\{\mathbb{E}\big[|u^{\varepsilon,N}_s(\xi^i_s) - \bar{u}^{\varepsilon,N}_s(\xi^i_s)|\big] + \mathbb{E}\big[|\nabla u^{\varepsilon,N}_s(\xi^i_s) - \nabla\bar{u}^{\varepsilon,N}_s(\xi^i_s)|\big]\Big\}\, ds\ .
\end{aligned} \tag{5.46}$$

Taking into account (5.43), for all $s \in [0,T]$, it follows

$$\begin{aligned}\mathbb{E}\big[|u_s^{\varepsilon,N}(\xi_s^i) - \bar{u}_s^{\varepsilon,N}(\xi_s^i)|\big] &= \mathbb{E}\big[|K_\varepsilon * (\gamma_s^{\varepsilon,N} - \bar{\gamma}_s^{\varepsilon,N})(\xi_s^i)|\big] \\ &\leq \frac{C}{\varepsilon^d}\mathbb{E}\Big[\|\gamma_s^{\varepsilon,N} - \bar{\gamma}_s^{\varepsilon,N}\|_{TV}\Big] , \qquad (5.47)\end{aligned}$$

where we have used inequality (5.42). In a very similar way, we also obtain

$$\begin{aligned}\mathbb{E}\big[|\nabla u_s^{\varepsilon,N}(\xi_s^i) - \nabla\bar{u}_s^{\varepsilon,N}(\xi_s^i)|\big] &= \frac{1}{\varepsilon}\sum_{\ell=1}^{d}\mathbb{E}\big[|G_\varepsilon^\ell * (\gamma_s^{\varepsilon,N} - \bar{\gamma}_s^{\varepsilon,N})(\xi_s^i)|\big] \\ &\leq \frac{C}{\varepsilon^{d+1}}\mathbb{E}\Big[\|\gamma_s^{\varepsilon,N} - \bar{\gamma}_s^{\varepsilon,N}\|_{TV}\Big] , \qquad (5.48)\end{aligned}$$

for all $s \in [0,T]$. Injecting (5.47) and (5.48) in the r.h.s. of (5.46) yields

$$A_t^{i,\varepsilon,N,n} \leq \frac{C}{\varepsilon^{d+1}}\int_0^t \mathbb{E}\Big[\|\gamma_s^{\varepsilon,N} - \bar{\gamma}_s^{\varepsilon,N}\|_{TV}\Big]\, ds . \qquad (5.49)$$

Concerning the second term in the r.h.s. of (5.45), we invoke again (2.3) to obtain

$$\begin{aligned}B_t^{i,\varepsilon,N,n} &:= \mathbb{E}\Big[\Big|V_t\big(\xi^i, \bar{u}^{\varepsilon,N}(\xi^i), \nabla\bar{u}^{\varepsilon,N}(\xi^i)\big) - V_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\Big|\Big] \\ &\leq e^{M_\Lambda T}L_\Lambda\mathbb{E}\Big[\int_0^t \Big|\Lambda(s,\xi_s^i,\bar{u}_s^{\varepsilon,N}(\xi_s^i),\nabla\bar{u}_s^{\varepsilon,N}(\xi_s^i)) - \Lambda(s,\bar{\xi}_s^i,\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i),\nabla\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i))\Big|\Big]\, ds \\ &\leq e^{M_\Lambda T}L_\Lambda\int_0^t\Big\{\mathbb{E}\Big[|\xi_s^i - \bar{\xi}_s^i|\Big] + \mathbb{E}\big[|\bar{u}_s^{\varepsilon,N}(\xi_s^i) - \bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i)|\big] + \mathbb{E}\big[|\nabla\bar{u}_s^{\varepsilon,N}(\xi_s^i) - \nabla\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i)|\big]\Big\}\, ds \\ &\leq \frac{Ce^{M_\Lambda T}L_\Lambda T\sqrt{\delta t}}{\varepsilon^{d+2}} \\ &\leq \frac{C}{\varepsilon^{d+2}\sqrt{n}}, \qquad (5.50)\end{aligned}$$

where we have used successively classical bounds of the Euler scheme (see e.g. Section 10.2, Chapter 10 in [13]) and (5.39).

Regarding the third term, similarly as for the above inequality (5.50), (2.3) yields

$$\begin{aligned}C_t^{i,\varepsilon,N,n} &:= \mathbb{E}\Big[\Big|V_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big) - \bar{V}_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\Big|\Big] \\ &\leq e^{M_\Lambda T}L_\Lambda\int_0^t\Big(|s-r(s)|^{\frac{1}{2}} + \mathbb{E}\Big[|\bar{\xi}_s^i - \bar{\xi}_{r(s)}^i|\Big] + \mathbb{E}\Big[|\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i) - \bar{u}_{r(s)}^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\Big] \\ &\quad + \mathbb{E}\Big[|\nabla\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i) - \nabla\bar{u}_{r(s)}^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\Big]\Big)ds , \qquad (5.51)\end{aligned}$$

where we have used Hölder property of $\Lambda$ w.r.t. the time variable.

Boundedness of $\Phi$, $g$ with classical Burkholder-Davis-Gundy (BDG) inequality give

$$\mathbb{E}\big[|\bar{\xi}_s^i - \bar{\xi}_{r(s)}^i|\big] \leq 2C\sqrt{\delta t} \leq \frac{C}{\sqrt{n}} , \quad s \in [0,T] . \qquad (5.52)$$

To bound the third term in the r.h.s. of (5.51), we use the following decomposition: for all $s \in [0,T]$,

$$\mathbb{E}\big[|\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i) - \bar{u}_{r(s)}^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\big] \leq \mathbb{E}\big[|\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i) - \bar{u}_s^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\big] + \mathbb{E}\big[|\bar{u}_s^{\varepsilon,N}(\bar{\xi}_{r(s)}^i) - \bar{u}_{r(s)}^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\big] . \qquad (5.53)$$

We first observe that the first inequality (5.39) gives

$$\mathbb{E}\big[|\bar{u}_s^{\varepsilon,N}(\bar{\xi}_s^i) - \bar{u}_s^{\varepsilon,N}(\bar{\xi}_{r(s)}^i)|\big] \leq \frac{C}{\varepsilon^{d+1}}\mathbb{E}\Big[|\bar{\xi}_s^i - \bar{\xi}_{r(s)}^i|\Big] \leq \frac{C\sqrt{\delta t}}{\varepsilon^{d+1}} \leq \frac{C}{\varepsilon^{d+1}\sqrt{n}} , \qquad (5.54)$$

for all $s \in [0,T]$. Invoking now the first inequality of (5.40) leads to

$$\mathbb{E}\big[|\bar{u}^{\varepsilon,N}_s(\bar{\xi}^i_{r(s)}) - \bar{u}^{\varepsilon,N}_{r(s)}(\bar{\xi}^i_{r(s)})|\big] \leq \frac{C\sqrt{\delta t}}{\varepsilon^{d+1}} \leq \frac{C}{\varepsilon^{d+1}\sqrt{n}}\ , \quad s \in [0,T]\ . \tag{5.55}$$

Injecting now (5.55) and (5.54) in (5.53) yield

$$\mathbb{E}\big[|\bar{u}^{\varepsilon,N}_s(\bar{\xi}^i_s) - \bar{u}^{\varepsilon,N}_{r(s)}(\bar{\xi}^i_{r(s)})|\big] \leq \frac{C}{\varepsilon^{d+1}\sqrt{n}}\ , \quad s \in [0,T]\ . \tag{5.56}$$

With very similar arguments as those used to obtain (5.56) (i.e. decomposition (5.53) and inequalities (5.39), (5.40)), we obtain for all $s \in [0,T]$,

$$\mathbb{E}\Big[|\nabla\bar{u}^{\varepsilon,N}_s(\bar{\xi}^i_s) - \nabla\bar{u}^{\varepsilon,N}_{r(s)}(\bar{\xi}^i_{r(s)})|\Big] \leq \frac{C\sqrt{\delta t}}{\varepsilon^{d+2}} \leq \frac{C}{\varepsilon^{d+2}\sqrt{n}}\ . \tag{5.57}$$

Gathering (5.57), (5.56) and (5.52) in (5.51) gives

$$C^{i,\varepsilon,N,n}_t \quad \leq \quad \frac{C\sqrt{\delta t}}{\varepsilon^{d+2}} \leq \frac{C}{\varepsilon^{d+2}\sqrt{n}}\ . \tag{5.58}$$

Finally, injecting (5.58), (5.50) and (5.49) in (5.45), we obtain for all $t \in [0,T]$,

$$\mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] \quad \leq \quad C\Big(\frac{1}{\varepsilon^{d+2}\sqrt{n}} + \frac{1}{\varepsilon^{d+1}}\int_0^t \mathbb{E}\Big[\|\gamma^{\varepsilon,N}_s - \bar{\gamma}^{\varepsilon,N}_s\|_{TV}\Big]\,ds\Big)\ . \tag{5.59}$$

Gronwall's lemma applied to the function $t \in [0,T] \mapsto \mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big]$ implies

$$\mathbb{E}\Big[\|\gamma^{\varepsilon,N}_t - \bar{\gamma}^{\varepsilon,N}_t\|_{TV}\Big] \leq \frac{C}{\varepsilon^{d+2}\sqrt{n}} e^{\frac{C}{\varepsilon^{d+1}}}\ , \quad t \in [0,T]\ . \tag{5.60}$$

The result follows by injecting (5.60) in (5.44). □

The particle algorithm used to simulate the dynamics (5.31) consists of the following steps.

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

1. Generate $(\bar{\xi}^i_0)_{i=1,..,N}$ i.i.d.$\sim u_0(x)dx$;
2. set $G^i_0 := 1$, $i = 1, \cdots, N$;
3. set $\bar{u}^{\varepsilon,N}_{t_0}(\cdot) := (K_\varepsilon * u_0)(\cdot)$.

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

- For $i = 1, \cdots N$, set $\bar{\xi}^i_{t_{k+1}} := \bar{\xi}^i_{t_k} + \Phi(t_k, \bar{\xi}^i_{t_k})\sqrt{\delta t}\ \epsilon^i_{k+1} + g(t_k, \bar{\xi}^i_{t_k})\delta t$ , where $(\epsilon^i_k)^{i=1,\cdots,N}_{k=1,\cdots n}$ is a sequence of i.i.d centered and standard Gaussian variables;
- for $i = 1, \cdots N$, set $G^i_{k+1} := G^i_k \times \exp\Big(\Lambda(t_k, \bar{\xi}^i_k, \bar{u}^{\varepsilon,N}_{t_k}(\bar{\xi}^i_{t_k}), \nabla\bar{u}^{\varepsilon,N}_{t_k}(\bar{\xi}^i_{t_k}))\delta t\Big)$ ;
- set $\bar{u}^{\varepsilon,N}_{t_{k+1}}(\cdot) = \dfrac{1}{N}\sum_{i=1}^{N} G^i_{k+1} \times K_\varepsilon(\cdot - \bar{\xi}^i_{t_{k+1}})$.

**Remark 5.8.** *Observe that each particle evolves independently without any interaction by contrast to the case considered in [15, 14]. However, since the evaluation of the function $\bar{u}^{\varepsilon,N}$ at any point $(t_k, \bar{\xi}^i_{t_k})$ requires to sum up $N$ terms, involving the whole particle system, the complexity of the algorithm is still of order $nN^2$.*

# 6 Numerical simulations

The aim of this section is to illustrate the performances of the original numerical scheme proposed in previous section to approximate the solution of semilinear PDEs (1.1) and inspect to what extent this approach remains valid out of Assumption 3. Following [6], we consider two types of semilinear PDEs, for which a semi-explicit expression of the solution is available: the one dimensional Burgers equation and the $d$-dimensional KPZ equation.

## 6.1 Burgers equation

Let $u_0$ be a probability density on $\mathbb{R}$ and $U_0 = \int_{-\infty}^{\cdot} u_0(y)dy$. Let us consider the viscid Burgers equation in dimension $d = 1$, given by

$$\begin{cases} \partial_t u = \frac{\nu^2}{2}\partial_{xx}u - u\partial_x u, & (t,x) \in [0,T] \times \mathbb{R}, \nu > 0 \\ u(0,\cdot) = u_0 \; . \end{cases} \tag{6.1}$$

It is well-known (see e.g. [6]) that (6.1) admits a unique classical solution if $u_0 \in L^1(\mathbb{R}^d)$. Using the so-called Cole-Hopf transformation, the solution $u$ admits the semi-explicit formula

$$u(t,x) = \frac{\mathbb{E}[u_0(x+\nu B_t)e^{-\frac{U_0(x+\nu B_t)}{\nu^2}}]}{\mathbb{E}[e^{-\frac{U_0(x+\nu B_t)}{\nu^2}}]}, \quad (t,x) \in [0,T] \times \mathbb{R}, \tag{6.2}$$

where $B$ denotes the real-valued standard Brownian motion. Integrating against test functions in space it is not difficult to show that the classical solution $u$ is also a weak solution of (1.1) with

$$\Phi = \nu, g \equiv 0, \Lambda(t,x,y,z) = z.$$

Apparently our Assumption 2 is not fulfilled, at least for what concerns $\Lambda$. However choosing $u_0$ being a bounded probability density, it is not difficult to show that there exists $M > 0$ such that $u$ is a solution of the subsidiary equation of type (1.1) with $\Phi \equiv \nu, \Lambda(t,x,y,z) := \Lambda_M(z)$ where $\Lambda_M : \mathbb{R} \to \mathbb{R}$ is a smooth bounded function such that $\Lambda_M(z) = z$ if $|z| \le M$ and $\Lambda_M(z) = 0$ if $|z| > M+1$. In this case Assumption 3 is fulfilled for the subsidiary equation.

In our numerical tests, we have implemented the time discretized particle scheme (5.31) with the following values of parameters $\Phi(t,x) := \nu$, $g(t,x) := 0$, $\Lambda(t,x,y,z) := z$, in order to approximate the solution of (6.1).

## 6.2 KPZ (deterministic) equation

This PDE makes sense in dimension $d \ge 1$. Through this second example, we want to give some empirical evidences that the convergence of $u^{\varepsilon,N}$ to $u$, when $N \to +\infty$ and $\varepsilon \to 0$, remains valid even when our assumptions are not fulfilled. Let us consider the KPZ equation

$$\begin{cases} \partial_t u = \frac{\nu^2}{2}\Delta u + |\nabla u|^2, & \text{for any } (t,x) \in [0,T] \times \mathbb{R}^d, \\ u(0,dx) = u_0(x)dx \; , \end{cases} \tag{6.3}$$

where $\Delta$ denotes as usual the Laplace operator and we recall that $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$. Using again the Cole-Hopf transformation, the solution $u$ admits the semi-explicit formula

$$u(t,x) = \log\left(\mathbb{E}\left[e^{u_0(x+\sigma B_t)}\right]\right), \tag{6.4}$$

where $B$ denotes a $\mathbb{R}^d$-valued standard Brownian motion. We suppose here that the initial condition $u_0$ is chosen strictly positive which ensures $u(t,x) \neq 0$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$. Indeed we have $e^{u(t,x)} = \mathbb{E}\big[e^{u_0(x+\sigma B_t)}\big] \geq 1 + \mathbb{E}[u_0(x+\sigma B_t)] > 1$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$. We remark that a strictly positive function $u$ is solution of (6.3) if and only if it is a solution of equation

$$\begin{cases} \partial_t u = \frac{\nu^2}{2} \Delta u + u\Lambda(t,x,u,\nabla u), \quad (t,x) \in [0,T] \times \mathbb{R}^d, \\ \Lambda(t,x,y,z) := \frac{|z|^2}{y}, \quad \text{for any } (t,x,y,z) \in [0,T] \times \mathbb{R}^d \times (0,+\infty) \times \mathbb{R}^d, \\ u(0,\cdot) = u_0 . \end{cases} \tag{6.5}$$

Notice that $\Lambda$ here is clearly not Lipschitz and then it does not satisfy Assumption 3. However, in our numerical tests, we have implemented the time discretized particle scheme (5.31) with the following choice of parameters $\Phi(t,x) := \nu$, $g(t,x) := 0$ and $\Lambda(t,x,y,z) := \frac{|z|^2}{y}$, to approximate the solution of (6.5).

## 6.3 Details of the implementation

In our figures, we have reported an approximation of the $L^1$-mean error committed by our numerical scheme (5.31) at the terminal time $T$. This error is approximated by Monte Carlo simulations as

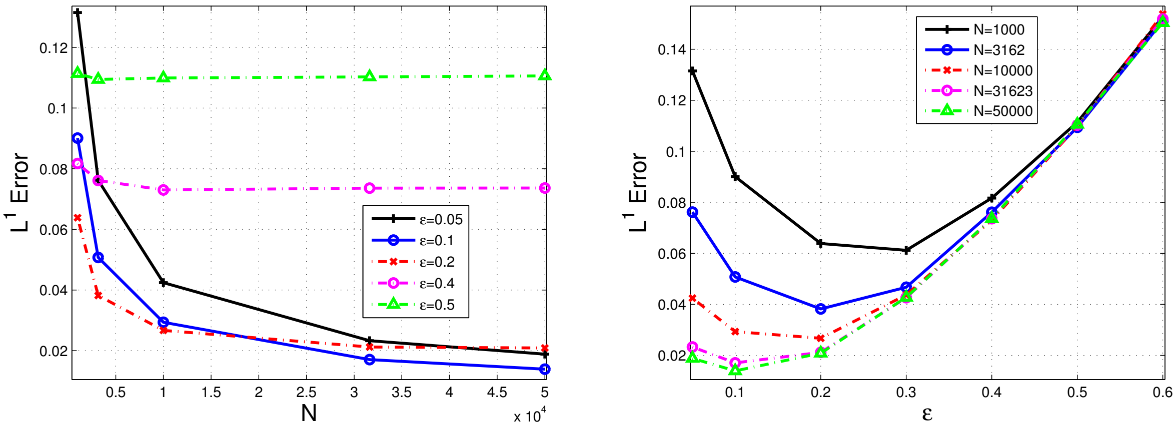
$$\mathbb{E}[\|u_T^{\varepsilon,N} - u_T\|_1] \approx \frac{1}{MQ} \sum_{i=1}^{M} \sum_{j=1}^{Q} |u_T^{\varepsilon,N,i}(X^j) - \hat{u}_T(X^j)|\; u_0^{-1}(X^j)\,, \qquad \text{where} \tag{6.6}$$

- $(u_T^{\varepsilon,N,i})_{i=1,\cdots,M=100}$ are i.i.d. estimates based on $M$ i.i.d. particle systems;
- $(X^j)_{j=1,\cdots,Q=1000}$ are i.i.d $\mathbb{R}^d$-valued random variables (independent of the particles defining $(u_T^{\varepsilon,N,i})_{i=1,\cdots,M=100}$), with common density $u_0$;
- $\hat{u}_T$ denotes a Monte Carlo estimation of the exact solution, $u_T$, with $10000$ simulations approximating the expectation formulas (6.2) for the Burgers equation and (6.4) for the KPZ equation.

The parameters of the problem in both cases (Burgers and KPZ) are $T = 0.1, \nu = 0.1$, with the centered and standard Gaussian distribution as initial condition i.e. $u_0(x)dx \sim \mathcal{N}(0, I_d)$.
Concerning the parameters of our numerical scheme, $n = 10$ time steps and $K = \phi^d$ with $\phi^d$ being the standard and centered Gaussian density on $\mathbb{R}^d$. To illustrate the trade-off condition (see (5.12)) between $N$ and $\varepsilon$, several values have been considered for the number particles $N = 1000,\ 3162,\ 10000,\ 31623,\ 50000$ and for the regularization parameter $\varepsilon = 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6$.

## 6.4 Simulations results

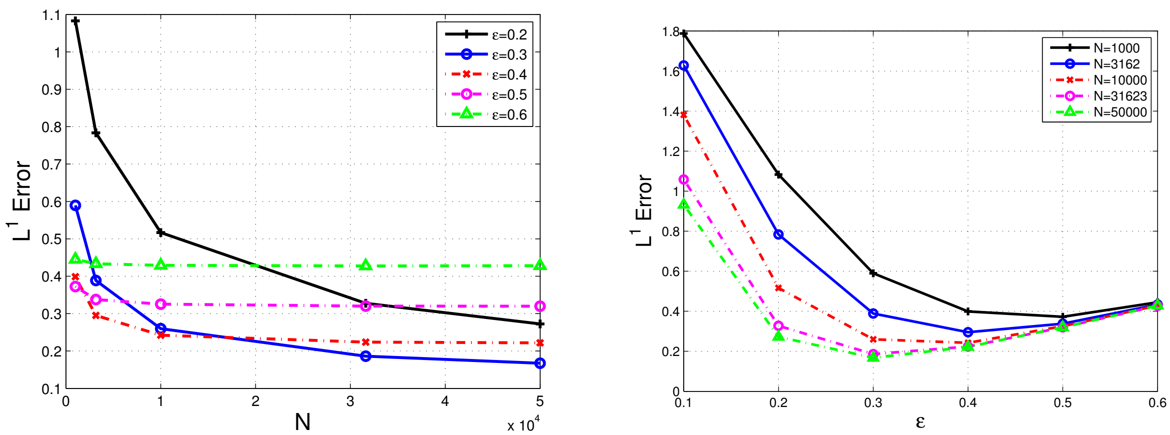
We have reported the estimated $L^1$ error committed by our approximation scheme (5.31) on Figure 1, for the Burgers equation (6.1) and on Figure 2, for the KPZ equation (6.3). Those results illustrate perfectly the tradeoff stated in (5.36) deriving from our theoretical bounds of the approximation error. Indeed, in both cases, one can observe on the left graphs that the error decreases with the number of particles, at a rate $N^{-1/2}$. However, when the regularization parameter $\varepsilon$ is large, the major part of the error is due to $\varepsilon$ so that the impact of increasing $N$ is rapidly negligible. On the right graphs, one can observe that the decrease of $\varepsilon$ to zero should be carefully adjusted to the increase of $N$ at an optimal rate which empirically corresponds to the evolution of the minimum of each curve with $N$.




(a) $L^1$ error as a function of $N$

(b) $L^1$ error as a function of $\epsilon$

Figure 1: $L^1$ error as a function of the number of particles, $N$, (on the left graph) and the mollifier window width, $\epsilon$, (on the right graph), for the Burgers equation (6.1), dimension $d = 1$.




(a) $L^1$ error as a function of $N$

(b) $L^1$ error as a function of $\epsilon$

Figure 2: $L^1$ error as a function of the number of particles, $N$, (on the left graph) and the mollifier window width, $\epsilon$, (on the right graph), for the KPZ equation (6.3), dimension $d = 5$.

# 7 Appendix

## 7.1 General inequalities

If $f$ is a probability density on $\mathbb{R}^d$, $I(f)$ denotes the quantity $I(f) := \int_{\mathbb{R}^d} |x|^{d+1} f(x)dx$.

**Lemma 7.1** (Multidimensional Carlson's inequality). *Let $f$ be a probability density on $\mathbb{R}^d$ such that $I(f) < \infty$, then*

$$\int_{\mathbb{R}^d} \sqrt{f(x)}dx \leq A_d \, I(f)^{\frac{d}{2(d+1)}} \qquad \textit{where} \qquad A_d = \left( \frac{(2\pi)^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})} \right)^{1/2} . \tag{7.1}$$

We refer to [12], where a more precise estimate is proved. From Carlson's inequality, we deduce the following lemma.

**Lemma 7.2.** *Let $G$ and $f$ be two probability densities defined on $\mathbb{R}^d$ such that*

$$I(G) < \infty\,, \quad \textit{and} \quad I(f) < \infty\,. \tag{7.2}$$

*Then for any strictly positive real $\varepsilon \le (1/I(G))^{\frac{1}{d+1}}$,*

$$\int_{\mathbb{R}^d} \sqrt{(G_\varepsilon * f)(x)}\, dx \le 2^{\frac{d}{2}}\, A_d\, [1 + I(f)] \quad \textit{where} \quad A_d = \left( \frac{(2\pi)^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})} \right)^{1/2}\,, \tag{7.3}$$

*where $G_\varepsilon(\cdot) := \frac{1}{\varepsilon^d} G(\frac{\cdot}{\varepsilon})$.*

*Proof.* By Carlson's inequality (7.1) we have

$$\int_{\mathbb{R}^d} \sqrt{(G_\varepsilon * f)(x)} dx \le A_d\, [I(G_\varepsilon * f)]^{\frac{d}{2(d+1)}}\,. \tag{7.4}$$

Then, by Minkowski's inequality,

$$\begin{aligned} [I(G_\varepsilon * f)]^{\frac{1}{d+1}} &= \Big[ \int_{\mathbb{R}^d\times\mathbb{R}^d} |x|^{d+1}\, G_\varepsilon(x-y)\, f(y) dy\, dx \Big]^{\frac{1}{d+1}} \\ &= \Big[ \int_{\mathbb{R}^d\times\mathbb{R}^d} |y + u\varepsilon|^{d+1}\, G(u)\, f(y) dy\, du \Big]^{\frac{1}{d+1}} \\ &\le I(f)^{\frac{1}{d+1}} + \varepsilon\, I(G)^{\frac{1}{d+1}}\,. \end{aligned}$$

Since $x \in \mathbb{R}^+ \mapsto x^d$ is convex, it follows

$$I(G_\varepsilon * f)^{\frac{d}{2(d+1)}} \quad \le \quad 2^{\frac{d-1}{2}}\, \big[\, [I(f)]^{\frac{d}{d+1}} + \varepsilon^d\, [I(G)]^{\frac{d}{d+1}} \big]^{\frac{1}{2}}\,.$$

Hence, as soon as $\varepsilon \le (1/I(G))^{\frac{1}{d+1}}$, we have

$$[I(G_\varepsilon * f)]^{\frac{d}{2(d+1)}} \le 2^{\frac{d}{2}}\, [1 + I(f)]\,, \tag{7.5}$$

which, owing to (7.4), concludes the proof. □

**Lemma 7.3.** *Let $H$ be a density kernel on $\mathbb{R}^d$ satisfying*

$$H \ge 0, \int_{\mathbb{R}^d} H(x)\, dx = 1\,, \quad \int_{\mathbb{R}^d} x\, H(x)\, dx = 0\,. \tag{7.6}$$

*Let $f : \mathbb{R}^d \to \mathbb{R}$ be a real-valued function. For any $\varepsilon > 0$, we consider the function $H_\varepsilon$ given by*

$$H_\varepsilon(\cdot) := \frac{1}{\varepsilon^d} H\left(\frac{\cdot}{\varepsilon}\right)\,. \tag{7.7}$$

*If $a := \frac{1}{2}\int_{\mathbb{R}^d} |x|^2\, H(x)\, dx < \infty$ (resp. $\tilde{a} := \int_{\mathbb{R}^d} |x|\, H(x)\, dx < \infty$) and $f \in W^{2,p}$ (resp. $f \in W^{1,p}$) for some integer $p \ge 1$, then for any $\varepsilon > 0$,*

$$\|H_\varepsilon * f - f\|_p \le \varepsilon^2\, a \sum_{i,j=1}^{d} \|\partial_i \partial_j f\|_p\,. \tag{7.8}$$

$$\Big(\textit{resp.} \quad \|H_\varepsilon * f - f\|_p \le \varepsilon\, \tilde{a} \sum_{i=1}^{d} \|\partial_i f\|_p\Big). \tag{7.9}$$

*Proof.* The proof is modeled on [12] and it is only written in the case $f \in W^{2,p}$. The case $f \in W^{1,p}$ follows exactly the same reasoning. We omit the corresponding details.
For $\varepsilon > 0$ and any integer $1 \leq i < j \leq d$ let us introduce the real-valued function $L_\varepsilon^{i,j}$ defined on $\mathbb{R}^d$ with values in $\bar{\mathbb{R}}_+$, associated with $H$ such that for almost all $x \in \mathbb{R}^d$,

$$L_\varepsilon^{i,j}(x) = \frac{x_i x_j}{\varepsilon^2} \int_0^1 \frac{1-t}{t^2} H_{\varepsilon t}(x)\, dt, \tag{7.10}$$

where $x_i$ is the $i$-th coordinate of $x$ and $H_t$ given by (7.7). Observe that, for any $\varepsilon > 0$, $1 \leq i < j \leq d$,

$$\|L_\varepsilon^{i,j}\|_1 = \int_{\mathbb{R}^d} |L_\varepsilon^{i,j}(x)| dx \leq a, \tag{7.11}$$

which implies that $L_\varepsilon^{i,j} < \infty$ a.e.

Developing $f$ according to the Lagrange expansion up to order two, yields, for almost all $(x, y) \in (\mathbb{R}^d)^2$,

$$f(x-y) = f(x) - \sum_{i=1}^d (\partial_i f)(x) y_i + \sum_{i,j=1}^d \int_0^1 (1-t)\partial_i \partial_j f(x - ty) y_i y_j \, dt \, .$$

Integrating this expression against $H_\varepsilon$ w.r.t. $y$ and using the symmetry of $H$, yields for almost all $x \in \mathbb{R}^d$,

$$\begin{aligned}
(H_\varepsilon * f)(x) - f(x) &= \int_{\mathbb{R}^d} [f(x-y) - f(x)]\, H_\varepsilon(y)\, dy \\
&= \sum_{i,j=1}^d \int_{\mathbb{R}^d} \int_0^1 (1-t)\partial_i\partial_j f(x-ty) y_i y_j \, dt \, H_\varepsilon(y)\, dy \\
&= \varepsilon^2 \sum_{i,j=1}^d \int_{\mathbb{R}^d} \partial_i \partial_j f(x-u) \frac{u_i u_j}{\varepsilon^2} \int_0^1 \frac{1-t}{t^2} H_{\varepsilon t}(u)\, dt\, du \\
&= \varepsilon^2 \sum_{i,j=1}^d (L_\varepsilon^{i,j} * (\partial_i \partial_j f))(x) \, ,
\end{aligned} \tag{7.12}$$

Taking the $L^p$ norm in equality (7.12), Young's inequality yields

$$\|H_\varepsilon * f - f\|_p \leq \varepsilon^2 \sum_{i,j=1}^d \|\partial_i \partial_j f\|_p \|L_\varepsilon^{i,j}\|_1 \, ,$$

which gives the result by recalling (7.11). □

## 7.2 About transition kernels

In the following lemma, we state well-known technical properties about the transition probability function of a diffusion process. All the statements below are established in [7].

**Lemma 7.4.** *We assume here the validity of Assumption 2. Consider a stochastic process $Z$, solution of the SDE*

$$Z_t = Z_0 + \int_0^t \Phi(s, Z_s) dW_s + \int_0^t g(s, Z_s) ds \, , \tag{7.13}$$

*with $Z_0$ a random variable admitting a bounded density $u_0$. $P(s, x_0, t, \Gamma)$ denotes its transition probability function, for all $(s, x_0, t, \Gamma) \in [0, T] \times \mathbb{R}^d \times [0, T] \times \mathcal{B}(\mathbb{R}^d)$. The following statements hold.*

1. *The transition probability function $P$ admits a density, i.e. there exists a Borel function $p : (s, x_0, t, x) \mapsto p(s, x_0, t, x)$ such that for all $(s, x_0, t) \in [0, T] \times \mathbb{R}^d \times [0, T]$,*

$$P(s, x_0, t, \Gamma) = \int_\Gamma p(s, x_0, t, x)dx \quad , \quad \Gamma \in \mathcal{B}(\mathbb{R}^d) \ . \tag{7.14}$$

2. *The function $p$ satisfies (in the classical sense) Kolmogorov backward $(B)$ and forward $(F)$ equations: i.e. $(s, x_0) \mapsto p(s, x_0, x, t)$ belongs to $C^{1,2}([0, t[\times\mathbb{R}^d, \mathbb{R})$ and satisfies*

$$(B) \left\{ \begin{array}{l} \partial_s p + L_s p = 0 \ , \quad 0 \leq s < t \leq T \\ p(s, x_0, t, x) \underset{s\uparrow t}{\longrightarrow} \delta_x \ , \textit{weakly} \ , \end{array} \right. \tag{7.15}$$

*where for $s \in [0, T]$, $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $L_s\varphi$ has been defined in* (2.6)*;*
*$(t, x) \mapsto p(s, x_0, t, x)$ is in $C^{1,2}(]s, T] \times \mathbb{R}^d, \mathbb{R})$ and satisfies*

$$(F) \left\{ \begin{array}{l} \partial_t p = L_t^* p \ , \quad 0 \leq s < t \leq T \\ p(s, x_0, t, x) \underset{t\downarrow s}{\longrightarrow} \delta_{x_0} \ , \textit{weakly} \ , \end{array} \right. \tag{7.16}$$

*where for $t \in [0, T]$, for $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$, $L_t^*\varphi$ has been defined by* (2.7)*.*
*In particular, $p$ is twice continuously differentiable w.r.t. $x_0$ and $x$.*

3. *There exist real constants $C_u, c_u > 0$ such that, for $0 \leq s < t \leq T$, $(x_0, x) \in \mathbb{R}^d \times \mathbb{R}^d$ and for all multi-index $m := (m_1, m_2)$ whose length $|m| := m_1 + m_2$ is less or equal to 2, we have*

$$\Big| \frac{\partial^{m_1}}{\partial z_i} \frac{\partial^{m_2}}{\partial x_j} p(s, z, t, x) \Big| \leq \frac{C_u}{(t-s)^{\frac{d+|m|}{2}}} e^{-c_u \frac{|x-z|^2}{t-s}}, \quad 0 \leq s < t \leq T, (z, x) \in (\mathbb{R}^d)^2 \ . \tag{7.17}$$

*In particular, there exists a constant $C > 0$ (only depending on $C_u$, $c_u$) such that for all $t \in [0, T]$, the law density $p_t$ of $Y_t$ satisfies*

$$\|p_t\|_\infty \leq C \|u_0\|_\infty \ , \tag{7.18}$$

*where $p_t$ is given by $\int_{\mathbb{R}^d} p(0, x_0, t, x)u_0(x_0)dx_0$.*

*Proof.* Under Assumption 2, the results are a result of Theorems 4.6, 4.7, Section 4 and Theorem 5.4, Section 5 in Chapter 6 in [7]. □

## 7.3 Proof of technicalities of Section 3

We give in this section the proof of Lemma 3.4.

*Proof of Lemma 3.4.* We only prove the direct implication since the converse follows easier with similar arguments. Without restriction of generality, we can assume that $T = N\tau$ for some integer $N \in \mathbb{N}$. The aim is to prove, for all $n \in \{1, \cdots, N\}$,

$$(H_n) \left\{ \begin{array}{l} \mu(t, dx) = \int_{\mathbb{R}^d} P(0, x_0, t, dx)u_0(dx_0) + \int_0^t ds \int_{\mathbb{R}^d} P(s, x_0, t, dx)\tilde{\Lambda}(s, x_0)\mu(s, dx_0) \ , \\ \text{for all } t \in [0, n\tau] \ . \end{array} \right. \tag{7.19}$$

We are going to proceed by induction on $n$. For $n = 1$, formula (7.19) follows from (3.13) by taking $k = 0$. We suppose now that $(H_{n-1})$ holds for some integer $n \geq 1$. Then, by taking $t = (n-1)\tau$ in the first line equation of (7.19), it follows immediately that

$$\mu((n-1)\tau, dx_0) = \int_{\mathbb{R}^d} P(0, \widetilde{x_0}, (n-1)\tau, dx_0) u_0(d\widetilde{x_0}) + \int_0^{(n-1)\tau} ds \int_{\mathbb{R}^d} P(s, \widetilde{x_0}, (n-1)\tau, dx_0) \tilde{\Lambda}(s, \widetilde{x_0}) \mu(s, d\widetilde{x_0}) \,. \tag{7.20}$$

On the other hand, since (3.13) is valid for all $t \in [(n-1)\tau, n\tau]$ by plugging $k = n-1$, we obtain

$$\mu(t, dx) = \int_{\mathbb{R}^d} P((n-1)\tau, x_0, t, dx) \mu((n-1)\tau, dx_0) + \int_{(n-1)\tau}^t ds \int_{\mathbb{R}^d} P(s, x_0, t, dx) \tilde{\Lambda}(s, x_0) \mu(s, dx_0) \,, \tag{7.21}$$

for all $t \in [(n-1)\tau, n\tau]$. Inserting (7.20) in (7.21) yields

$$\begin{aligned} \mu(t, dx) &= \int_{\mathbb{R}^d} u_0(d\widetilde{x_0}) \int_{\mathbb{R}^d} P(0, \widetilde{x_0}, (n-1)\tau, dx_0) P((n-1)\tau, x_0, t, dx) \\ &+ \int_0^{(n-1)\tau} ds \int_{\mathbb{R}^d} \mu(s, d\widetilde{x_0}) \tilde{\Lambda}(s, \widetilde{x_0}) \int_{\mathbb{R}^d} P(s, \widetilde{x_0}, (n-1)\tau, dx_0) P((n-1)\tau, x_0, t, dx) \\ &+ \int_{(n-1)\tau}^t ds \int_{\mathbb{R}^d} P(s, x_0, t, dx) \tilde{\Lambda}(s, x_0) \mu(s, dx_0) \,,\ t \in [(n-1)\tau, n\tau] \,. \end{aligned} \tag{7.22}$$

Invoking the Chapman-Kolmogorov equation satisfied by the transition probability function $P(s, x_0, t, dx)$ (see e.g. expression (2.1) in Section 2.2, Chapter 2 in [19]), we have

$$P(s, \widetilde{x_0}, t, dx) = \int_{\mathbb{R}^d} P(s, \widetilde{x_0}, \theta, dz) P(\theta, z, t, dx), \quad s < \theta < t,\ (\widetilde{x_0}, z) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{7.23}$$

Applying (7.23) with $\theta = (n-1)\tau$, it follows that for all $t \in [0, n\tau]$,

$$\begin{aligned} \mu(t, dx) &= \int_{\mathbb{R}^d} u_0(d\widetilde{x_0}) P(0, \widetilde{x_0}, t, dx) \\ &+ \int_0^t ds \int_{\mathbb{R}^d} P(s, \widetilde{x_0}, t, dx) \tilde{\Lambda}(s, \widetilde{x_0}) \mu(s, d\widetilde{x_0}) \,. \end{aligned} \tag{7.24}$$

This shows that $(H_n)$ holds.

□

### 7.4 Technicalities related to Section 5

*Proof of Lemma 5.7.* Let us fix $\varepsilon > 0, N \in \mathbb{N}^\star$, $t \in [0, T]$. We first recall that for almost all $x \in \mathbb{R}^d$,

$$\begin{cases} u_t^{\varepsilon,N}(x) = \dfrac{1}{N} \displaystyle\sum_{i=1}^N K_\varepsilon(y - \xi_t^i) V_t\big(\xi^i, u^{\varepsilon,N}(\xi^i), \nabla u^{\varepsilon,N}(\xi^i)\big) \,, \\ \bar{u}_t^{\varepsilon,N}(x) = \frac{1}{N} \sum_{i=1}^N K_\varepsilon(x - \bar{\xi}_t^i) \bar{V}_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla \bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big), \end{cases} \tag{7.25}$$

for which $V_t$ (resp. $\bar{V}_t$) is given by (2.1) (resp. (5.32)).
Let us fix $i \in \{1, \cdots, N\}$.

- *Proof of* (5.39). The proof of the two inequalities (5.39) is almost the same. That is why we only give details for the proof of the first inequality.

From the second line equation of (7.25), we have

$$\begin{aligned}|\bar{u}^{\varepsilon,N}_{r(t)}(x)-\bar{u}^{\varepsilon,N}_{r(t)}(y)| &\leq \frac{1}{N}\sum_{i=1}^{N}\big|K_\varepsilon(x-\bar{\xi}^i_{r(t)})-K_\varepsilon(y-\bar{\xi}^i_{r(t)})\big|\bar{V}_{r(t)}\big(\bar{\xi}^i,\bar{u}^{\varepsilon,N}(\bar{\xi}^i),\nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\\ &\leq \frac{e^{M_\Lambda T}}{N\varepsilon^{d+1}}\sum_{i=1}^{N}L_K|x-y|\\ &\leq \frac{e^{M_\Lambda T}L_K}{\varepsilon^{d+1}}|x-y|\,, \end{aligned} \tag{7.26}$$

where for the second step above, we have used the fact that $K$ is in particular Lipschitz. The same arguments lead also to

$$|\nabla_x\bar{u}^{\varepsilon,N}_{r(t)}(x)-\nabla_x\bar{u}^{\varepsilon,N}_{r(t)}(y)| \leq \frac{e^{M_\Lambda T}L_{\nabla K}}{\varepsilon^{d+2}}|x-y|\,, \tag{7.27}$$

which ends the proof of (5.39).

- *Proof of* (5.40). From

$$\bar{u}^{\varepsilon,N}_t(x)=\frac{1}{N}\sum_{i=1}^{N}K_\varepsilon(x-\bar{\xi}^i_t)\bar{V}_t\big(\bar{\xi}^i,\bar{u}^{\varepsilon,N}(\bar{\xi}^i),\nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\,,\quad x\in\mathbb{R}^d\,, \tag{7.28}$$

we deduce, for almost all $x\in\mathbb{R}^d$,

$$\begin{aligned}|\bar{u}^{\varepsilon,N}_t(x)-\bar{u}^{\varepsilon,N}_{r(t)}(x)| &\leq \frac{e^{M_\Lambda T}}{N}\sum_{i=1}^{N}\Big|K_\varepsilon(x-\bar{\xi}^i_t)-K_\varepsilon(x-\bar{\xi}^i_{r(t)})\Big|\\ &\quad+\frac{\|K\|_\infty}{N\varepsilon^d}\sum_{i=1}^{N}\big|\bar{V}_t\big(\bar{\xi}^i,\bar{u}^{\varepsilon,N}(\bar{\xi}^i),\nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)-\bar{V}_{r(t)}\big(\bar{\xi}^i,\bar{u}^{\varepsilon,N}(\bar{\xi}^i),\nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\big|\,.\end{aligned} \tag{7.29}$$

We remark that $K$ is Lipschitz and we deonte the related constant by $L_K$. We then obtain, for almost all $x\in\mathbb{R}^d$,

$$\begin{aligned}|\bar{u}^{\varepsilon,N}_t(x)-\bar{u}^{\varepsilon,N}_{r(t)}(x)| &\leq \frac{L_K e^{M_\Lambda T}}{N\varepsilon^{d+1}}\sum_{i=1}^{N}|\bar{\xi}^i_t-\bar{\xi}^i_{r(t)}|\\ &\quad+\frac{L_\Lambda e^{M_\Lambda T}\|K\|_\infty}{N\varepsilon^d}\sum_{i=1}^{N}\int_{r(t)}^{t}\Lambda(r(s),\bar{\xi}^i_{r(s)},\bar{u}^{\varepsilon,N}_{r(s)}(\bar{\xi}^i_{r(s)}),\nabla_x\bar{u}^{\varepsilon,N}_{r(s)}(\bar{\xi}^i_{r(s)}))ds\,,\end{aligned} \tag{7.30}$$

where the second term in (7.30) comes from inequality (2.3). Since $\Lambda$ is bounded, by taking the supremum w.r.t. $x$ and the expectation in both sides of inequality above we have

$$\begin{aligned}\mathbb{E}\Big[\|\bar{u}^{\varepsilon,N}_t-\bar{u}^{\varepsilon,N}_{r(t)}\|_\infty\Big] &\leq \frac{L_K e^{M_\Lambda T}}{N\varepsilon^{d+1}}\sum_{i=1}^{N}\mathbb{E}\Big[|\bar{\xi}^i_t-\bar{\xi}^i_{r(t)}|\Big]\\ &\quad+\frac{L_\Lambda e^{M_\Lambda T}\|K\|_\infty}{\varepsilon^d}M_\Lambda\delta t\leq\frac{C\sqrt{\delta t}}{\varepsilon^{d+1}}\,,\end{aligned} \tag{7.31}$$

where we have used the fact that $\mathbb{E}\Big[|\bar{\xi}^i_s-\bar{\xi}^i_{r(s)}|^2\Big]\leq C\delta t$, since $\Phi, g$ are bounded.
The bound of $\mathbb{E}\Big[\|\nabla_x\bar{u}^{\varepsilon,N}_t-\nabla_x\bar{u}^{\varepsilon,N}_{r(t)}\|_\infty\Big]$ is obtained by proceeding exactly in with the same way as

above, starting with

$$\frac{\partial \bar{u}_t^{\varepsilon,N}}{\partial x_\ell}(\cdot) = \frac{1}{N\varepsilon}\sum_{i=1}^{N}\frac{\partial K_\varepsilon}{\partial x_\ell}(\cdot - \bar{\xi}_t^i)\bar{V}_t\big(\bar{\xi}^i, \bar{u}^{\varepsilon,N}(\bar{\xi}^i), \nabla\bar{u}^{\varepsilon,N}(\bar{\xi}^i)\big)\ , l = 1, \cdots, d\ , \tag{7.32}$$

instead of (7.28), where $x_\ell$ denotes the $\ell$-th coordinate of $x \in \mathbb{R}^d$. It follows then

$$\mathbb{E}\Big[\|\nabla_x \bar{u}_t^{\varepsilon,N} - \bar{\nabla}_x u_{r(t)}^{\varepsilon,N}\|_\infty\Big] \quad \leq \quad \frac{C\sqrt{\delta t}}{\varepsilon^{d+2}}\ . \tag{7.33}$$

□

**ACKNOWLEDGMENTS.** Part of this work has been done during a stay of the third named author in the University of Bielefeld, SFB 701. He is grateful to Prof. Michael Röckner for the kind invitation.

# References

[1] 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.

[2] 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.

[3] Vladimir I. Bogachev, Nicolai V. Krylov, Michael Röckner, and Stanislav V. Shaposhnikov. *Fokker-Planck-Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

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

[5] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.*, 60(7):1081–1110, 2007.

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

[7] A. Friedman. *Stochastic differential equations and applications. Vol. 1*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Probability and Mathematical Statistics, Vol. 28.

[8] 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.

[9] P. Henry-Labordère. Counterparty risk valuation: A marked branching diffusion approach. *Available at SSRN: http://ssrn.com/abstract=1995503 or http://dx.doi.org/10.2139/ssrn.1995503*, 2012.

[10] 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.

[11] P. Henry-Labordère, X. Tan, and N. Touzi. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl.*, 124(2):1112–1140, 2014.

[12] L. Holmström and J. Klemelä. Asymptotic bounds for the expected $L^1$ error of a multivariate kernel density estimator. *J. Multivariate Anal.*, 42(2):245–266, 1992.

[13] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[14] A. Le Cavil, N. Oudjane, and F. Russo. Particle system algorithm and chaos propagation related to a non-conservative McKean type stochastic differential equations. *Preprint HAL. https://hal.inria.fr/hal-01241704.*, 2015.

[15] A. Le Cavil, N. Oudjane, and F. Russo. Probabilistic representation of a class of non conservative nonlinear partial differential equations. *Preprint HAL. https://hal.archives-ouvertes.fr/hal-01241701*, 2015.

[16] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 79–127. Birkhäuser Boston, Boston, MA, 1998.

[17] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[18] E. Pardoux and A. Raşcanu. *Stochastic differential equations, Backward SDEs, Partial differential equations*, volume 69. Springer, 2014.

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