

A probabilistic algorithm approximating solutions of a singular PDE of porous media type

Nadia Belaribi, François Couvelier and Francesco Russo

Abstract. The object of this paper is a one-dimensional generalized porous media equation (PDE) with possibly discontinuous coefficient β , which is well-posed as an evolution problem in $L^1(\mathbb{R})$. In some recent papers of Blanchard et alia and Barbu et alia, the solution was represented by the solution of a non-linear stochastic differential equation in law if the initial condition is a bounded integrable function. We first extend this result, at least when β is continuous and the initial condition is only integrable with some supplementary technical assumption. The main purpose of the article consists in introducing and implementing a stochastic particle algorithm to approach the solution to (PDE) which also fits in the case when β is possibly irregular, to predict some long-time behavior of the solution and in comparing with some recent numerical deterministic techniques.

Keywords. Stochastic particle algorithm, porous media equation, monotonicity, stochastic differential equations, non-parametric density estimation, kernel estimator.

AMS classification. MSC 2010: 65C05, 65C35, 82C22, 35K55, 35K65, 35R05, 60H10, 60J60, 62G07, 65M06.

Actual version: October 14th 2011

1. Introduction

The main aim of this work is to construct and implement a probabilistic algorithm which will allow us to approximate solutions of a porous media type equation with monotone irregular coefficient. Indeed, we are interested in the parabolic problem

$$\begin{cases} \partial_t u(t, x) &= \frac{1}{2} \partial_{xx}^2 \beta(u(t, x)), \quad t \in [0, +\infty[, \\ u(0, x) &= u_0(dx), \quad x \in \mathbb{R}, \end{cases} \quad (1.1)$$

in the sense of distributions, where u_0 is an initial probability measure. If u_0 has a density, we will still denote it by the same letter. We look for a solution of (1.1) with time evolution in $L^1(\mathbb{R})$. We formulate the following assumption:

Assumption(A)

- (i) $\beta : \mathbb{R} \rightarrow \mathbb{R}$ such that $\beta|_{\mathbb{R}_+}$ is monotone.
- (ii) $\beta(0) = 0$ and β continuous at zero.
- (iii) We assume the existence of $\lambda > 0$ such that $(\beta + \lambda id)(\mathbb{R}_+) = (\mathbb{R}_+)$, $id(x) \equiv x$.

A monotone function $\beta_0 : \mathbb{R} \rightarrow \mathbb{R}$ can be completed into a graph by setting $\beta_0(x) = [\beta_0(x_-), \beta_0(x_+)]$. An odd function $\beta_0 : \mathbb{R} \rightarrow \mathbb{R}$ such that $\beta|_{\mathbb{R}_+} = \beta_0|_{\mathbb{R}_+}$ produces in this way a maximal monotone graph.

In this introduction, however β and β_0 will be considered single-valued for the sake of simplicity. We leave more precise formulations (as in Proposition 2.1 and Theorem 2.8) for the body of the article.

We remark that if β fulfills Assumption(A), then the odd symmetrized β_0 fulfills the more natural

Assumption(A')

- (i) $\beta_0 : \mathbb{R} \rightarrow \mathbb{R}$ is monotone.
- (ii) $\beta_0(0) = 0$ and β_0 continuous at zero.
- (iii) We assume the existence of $\lambda > 0$ such that $(\beta_0 + \lambda id)(\mathbb{R}) = (\mathbb{R})$, $id(x) \equiv x$.

We define $\Phi : \mathbb{R} \rightarrow \mathbb{R}_+$, setting

$$\Phi(u) = \begin{cases} \sqrt{\frac{\beta_0(u)}{u}} & \text{if } u \neq 0, \\ C & \text{if } u = 0, \end{cases} \quad (1.2)$$

where $C \in [\liminf_{u \rightarrow 0^+} \Phi(u), \limsup_{u \rightarrow 0^+} \Phi(u)]$.

Note that when $\beta(u) = u \cdot |u|^{m-1}$, $m > 1$, the partial differential equation (PDE) in (1.1) is nothing else but the classical porous media equation. In this case $\Phi(u) = |u|^{\frac{m-1}{2}}$ and in particular $C = 0$.

Our main target is to analyze the case of an irregular coefficient β . Indeed, we are particularly interested in the case when β is continuous excepted for a possible jump at one positive point, say $u_c > 0$. A typical example is:

$$\beta(u) = H(u - u_c) \cdot u, \quad (1.3)$$

H being the Heaviside function and u_c will be called *critical value* or *critical threshold*.

Definition 1.1. i) We will say that the PDE in (1.1), or β is *non-degenerate* if there is a constant $c_0 > 0$ such that $\Phi \geq c_0$, on each compact of \mathbb{R}_+ .

ii) We will say that the PDE in (1.1), or β is *degenerate* if $\lim_{u \rightarrow 0^+} \Phi(u) = 0$.

Remark 1.2. i) We remark that β is non-degenerate if and only if $\liminf_{u \rightarrow 0^+} \Phi(u) > 0$.

ii) We observe that β may be neither degenerate nor non-degenerate.

Of course, β in (1.3) is degenerate. Equation (1.3) constitutes a model intervening in some self-organized criticality (often called SOC) phenomena, see [2] for a significant monograph on the subject. We mention the interesting physical paper [16], which makes reference to a system whose evolution is similar to the evolution of a "snow layer" under the influence of an "avalanche effect" which starts whenever the top of the layer is bigger than a critical value u_c .

We, in particular, refer to [9] (resp. [3]), which concentrates on the avalanche phase and therefore investigates the problem (1.1) discussing existence, uniqueness and probabilistic representation when β is non-degenerate (resp. degenerate). The authors had in mind the singular PDE in (1.1) as a macroscopic model for which they gave a microscopic view via a probabilistic representation provided by a non-linear stochastic differential equation (NLSDE); the stochastic equation is supposed to describe the evolution of a single point of the layer. The analytical assumptions formulated by the authors were Assumption(A) and the Assumption(B) below which postulates linear growth for β .

Assumption(B). There exists a constant $c > 0$ such that $|\beta(u)| \leq c|u|$.

Obviously we have,

Assumption(B'). There exists a constant $c > 0$ such that $|\beta_0(u)| \leq c|u|$.

Clearly (1.3) fulfills Assumption(B).

To the best of our knowledge the first author who considered a probabilistic representation (of the type studied in this paper) for the solutions of non linear deterministic partial differential equations was McKean [32]. However, in his case, the coefficients were smooth. From then on, the literature steadily grew and nowadays there is a vast amount of contributions to the subject. A probabilistic interpretation of (1.1) when $\beta(u) = u|u|^{m-1}$, $m > 1$ was provided in [5]. For the same β , though the method could be adapted to the case where β is Lipschitz, in [28], the author studied the evolution problem (1.1) when the initial condition and the evolution takes values in the class of probability distribution functions on \mathbb{R} . He studied both the probabilistic representation and the so-called *propagation of chaos*.

At the level of probabilistic representation, under Assumptions (A) and (B), supposing that u_0 has a bounded density, [9] (resp. [3]) proves existence and uniqueness (resp. existence) in law for the corresponding (NLSDE). In the present work we are interested in some theoretical complements, but the main purpose consists in examining numerical implementations provided by the (NLSDE), in comparison with numerical deterministic schemes appearing in one recent paper, see [18].

Let us now describe the principle of the probabilistic representation. The stochastic differential equation (in law) rendering the probabilistic representation is given by the

following (NLSDE):

$$\begin{cases} Y_t &= Y_0 + \int_0^t \Phi(u(s, Y_s)) dW_s, \\ u(t, \cdot) &= \text{Law density of } Y_t, \quad \forall t > 0, \\ u(0, \cdot) &= u_0 \quad \text{Law of } Y_0, \end{cases} \quad (1.4)$$

where W is a classical Brownian motion. The solution of that equation may be visualized as a continuous process Y on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ equipped with an $(\mathcal{F}_t)_{t \geq 0}$ -Brownian motion W .

Until now, theoretical results about well-posedness (resp. existence) for (1.4) were established when β is non-degenerate (resp. possibly degenerate) and in the case when $u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$.

Initially our aim was to produce an algorithm which allows to start even with a measure or an unbounded function as initial condition. Unfortunately, up to now, our implementation techniques do not allow to treat this case.

However, even if the present paper concentrates on numerical experiments, two theoretical contributions are performed when Φ is continuous.

- A first significant theoretical contribution is Theorem 2.9 which consists in fact in extending the probabilistic representation obtained by [3] to the case when $u_0 \in L^1(\mathbb{R})$, locally of bounded variation outside a discrete set of points, not necessarily bounded.

- A second contribution consists in showing in the non-degenerate case that the mollified version of PDE in (1.1) is in fact equivalent to its probabilistic representation, even when the initial condition u_0 is a probability measure. This is done in Theorem 3.2.

The connection between (1.4) and (1.1) is indeed given by the following result.

Proposition 1.3. *Let us assume the existence of a solution Y for (1.4). Let $u(t, \cdot)$ be the law density of Y_t , $t > 0$, that we suppose to exist.*

Then $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}_+$ provides a solution in the sense of distributions of (1.1) with $u_0 = u(0, \cdot)$.

The proof is well-known, but we recall here the basic argument for illustration purposes.

Proof. Let $\varphi \in C_0^\infty(\mathbb{R})$, Y be a solution of the problem (1.4). We apply Itô's formula to $\varphi(Y)$ to obtain :

$$\varphi(Y_t) = \varphi(Y_0) + \int_0^t \varphi'(Y_s) \Phi(u(s, Y_s)) dW_s + \frac{1}{2} \int_0^t \varphi''(Y_s) \Phi^2(u(s, Y_s)) ds.$$

Taking the expectation we get :

$$\int_{\mathbb{R}} \varphi(y) u(t, y) dy = \int_{\mathbb{R}} \varphi(y) u_0(y) dy + \frac{1}{2} \int_0^t ds \int_{\mathbb{R}} \varphi''(y) \Phi^2(u(s, y)) u(s, y) dy.$$

Using then integration by parts and the expression of β , the expected result follows. \square

In the literature there are several contributions about approximation of non-linear PDE's of parabolic type using a stochastic particles system, with study of the chaos propagation. We recall that the chaos propagation takes place if the components of a vector describing the interacting particle system become asymptotically independent, when the number of particles goes to infinity. Note that, physically motivated applications can be found, for instance in numerical studies in hydro- or plasma-physics; [20] and [24] are contributions expressing a heuristic or formal point of view.

When the non-linearity is of the first order, a significant contribution was given by [48]; [10, 11] study a McKean-Vlasov model containing as a particular case, the Burgers equation where, as in [28], the initial condition is a probability (cumulated) distribution function. They performed the rate of convergence of an empirical measure associated with a stochastic particle algorithm. [33] enlarged the model of [32] and [48] and they established the related chaos propagation result. We also quote [17], where authors obtained Burgers equation as a mean field limit of suitable diffusion processes.

In the case of porous media type equation (1.1) with β Lipschitz, [29] investigated the probabilistic representation for (1.1) and a mollified related equation. There, the authors provided a rigorous proof of the propagation of chaos in the case of Lipschitz coefficients, see Proposition 2.3, Proposition 2.5 and Theorem 2.7 of [29].

Outside the Lipschitz case, an alternative method for studying convergence was investigated by [34, 35, 36], whose limiting PDEs concerned a class of equations including the case $\beta(u) = u + u^2$, $u \geq 0$. In fact [36] computed the numerical solution of a viscous porous medium equation through a particle algorithm and studied the L^2 -convergence rate to the analytical solution. More recent papers concerning the chaos propagation when $\beta(u) = u^2$ first and $\beta(u) = |u|^{m-1}u$, $m > 1$ was proposed in [39] and [21].

As far as the coefficient β is discontinuous, at our knowledge, up to now, there are no such results. As we announced, we are particularly interested in an empirical investigation of the stochastic particle algorithm approaching the solution u of (1.1) at some instant t , in several situations with regular or irregular coefficient. We recall that $u(t, \cdot)$ is a probability density. That algorithm involves Euler schemes of stochastic differential equations, Monte-Carlo simulations expressing the empirical law and non-parametric density estimation of $u(t, \cdot)$ using Gaussian kernels, see [46] for an introduction to the kernel method. This technique crucially depends on the window width ε of the smoothing kernel. Classical statistical tools for choosing that parameter are described for instance in [46], where the following formula for choosing the optimal bandwidth ε , in the sense of minimizing the asymptotic *mean integrated squared*

error (MISE), is given by

$$\varepsilon_t = (2n\sqrt{\pi}\|\partial_{xx}^2 u(t, \cdot)\|^2)^{-\frac{1}{5}}, \quad (1.5)$$

where, n is the sample size and $\|\cdot\|$ denotes the classical $L^2(\mathbb{R})$ norm.

Of course, the above expression does not yield an immediately practicable method for choosing the optimal ε since (1.5) depends on the second derivative of the density u , which we are trying indeed to estimate. Therefore, several techniques were proposed to get through this problem. First, a natural and easy approach, often called the *rule of thumb*, replaced the target density u at time t in the functional $\|\partial_{xx}^2 u\|$, by a reference distribution function. For instance, [46] assumed that the unknown density is a standard normal function and obtained the following practically used formula

$$\varepsilon = \left(\frac{4}{3n}\right)^{\frac{1}{5}} \hat{\sigma}, \quad (1.6)$$

$\hat{\sigma}$ being the empirical standard deviation. A version which is more robust to outliers in the sample, consists in replacing $\hat{\sigma}$ by a measure of spread of the variance involving the interquartile range. For instance, see [46] for detailed computations.

The *oversmoothing* methods rely on the fact that there is a simple upper bound for the MISE-optimal bandwidth. In fact, [49], gave a lower bound for the functional $\|\partial_{xx}^2 u\|$ and thus an upper bound for ε in (1.5); it proposed to use this upper bound as an optimal window width, see also [50] for histograms.

The two methods above seem to work well for unimodal densities. However, they lead to arbitrarily bad estimates of the bandwidth ε , when for instance, the true density is far from being Gaussian, especially when it is a multimodal law.

The *least squares cross validation* (LSCV) method aimed to estimate the bandwidth that minimizes the integrated squared error (ISE), based on a "leave-one-out" kernel density estimator, see [41, 14]. The problem is that, for the same target distribution, the estimated bandwidth through different samples has a big variance, which produces instability.

The *biased cross-validation* (BCV) approach, introduced in [42] minimizes the score function obtained by replacing the functional $\|\partial_{xx}^2 u\|$ in the formula of the MISE by an estimator $\|\partial_{xx}^2 \hat{u}\|$, where \hat{u} is the kernel estimator of u . In fact, [42] proposed the use of the minimizer of that score function as optimal bandwidth. This method seemed to be more stable than the LSCV but still has large bias. The slow rate of convergence of both the LSCV and BCV approaches encouraged significant research on faster converging methods.

A popular approach, commonly called *plug-in* method, makes use of an indirect estimator of the density functional $\|\partial_{xx}^2 u\|$ in formula (1.5). This technique comes back to the early paper [52]; in this framework the estimator of $\|\partial_{xx}^2 u\|$ requires the computation of a *pilot bandwidth* h , which is quite different from the window width

ε used for the kernel density estimate. Indeed, this optimal bandwidth h depends on unknown density functionals involving partial derivatives greater than 2. Following an idea of [51], one could express h iteratively through higher order derivatives. In this spirit, the natural associated problem consists in estimating for some positive integer s , the quantity $\|\partial_{x^s}^s u\|$, in terms of $\|\partial_{x^{s+\ell}}^{s+\ell} u\|$ for some positive integer ℓ ; an ℓ -stage *direct plug-in* approach may consist in replacing the norm $\|\partial_{x^{s+\ell}}^{s+\ell} u\|$ by the norm of the $s + \ell$ derivative of a Gaussian density. In the present paper we implement this idea with $s = \ell = 2$. Important contributions to that topic were [43] and [27] who improved the method via the so-called "solve-the-equation" plug-in method. By this technique, the pilot bandwidth h used to estimate $\|\partial_{xx}^2 u\|$, is written as a function of the kernel bandwidth ε . We shall describe in Section 4 in details this bandwidth selection procedure applied in the case of our probabilistic algorithm.

We point out, that a more recent tool was developed in [12] which improved the idea in [43, 27] in the sense that [12] did not postulate any normal reference rule. However, the numerical experiments that we have performed using the Matlab routine developed by the first author of [12] have not produced better results in the case when β is defined by (1.3).

In the paper we examine empirically the stochastic particle algorithm for approaching the solution to the PDE in the case $\beta(u) = u^3$ and in the case β given by (1.3). For this more peculiar case, we compare the approximation with the one obtained by one recent analytic deterministic numerical method.

Problems of the same type as (1.1), in the case when β is Lipschitz but possibly degenerate, were extensively studied from both the theoretical and numerical deterministic points of view. In general, the numerical analysis of (1.1) is difficult for at least one reason: the appearance of singularities for compactly supported solutions in the case of an irregular initial condition. An usual technique to approximate (1.1) involves implicit discretization in time: it requires, at each time step, the discretization of a nonlinear elliptic problem. However, when dealing with nonlinear problems, one generally tries to linearize them in order to take advantage of efficient linear solvers. Linear approximation schemes based on the so-called non linear Chernoff's formula with a suitable relaxation parameter and which arises in the theory of nonlinear semi groups, were studied for instance in [8]. We also cite [30], where the authors approximated degenerate parabolic problems including those of porous media type. In fact, they used nonstandard semi-discretization in time and applied a Newton-like iterations to solve the corresponding elliptic problems. More recently, different approaches based on kinetic schemes for degenerate parabolic systems have been investigated in [1]. Finally a new scheme based on the maximum principle and on the perturbation and regularization approach was proposed in [40].

At the best of our knowledge, up to now, there are no analytical methods dealing with the case when β is given by (1.3). However, we are interested in a sophisticated approach developed in [18] and which appears to be best suited to describe the evolu-

tion of singularities and efficient for computing discontinuous solutions. In fact, [18] focuses onto diffusive *relaxation schemes* for the numerical approximation of nonlinear parabolic equations, see [26, 25], and references therein. Those relaxed schemes are based on a suitable semi-linear hyperbolic system with relaxation terms. Indeed, this reduction is carried out in order to obtain schemes that are easy to implement. Moreover, with this approach it is possible to improve such schemes by using different numerical approaches i.e. either finite volumes, finite differences or high order accuracy methods.

In particular, the authors in [18] coupled ENO (Essentially Non Oscillatory) interpolating algorithms for space discretization, see [45], in order to deal with discontinuous solutions and prevent the onset of spurious oscillation, with IMEX (implicit explicit) Runge-Kutta schemes for time advancement, see [37], to obtain a high order method. We point out that [18] studied convergence and stability of the scheme only in the case when β is Lipschitz but possibly degenerate and $u_0 \in L^1(\mathbb{R})$.

As a byproduct of numerical experiments we can forecast the longtime behavior of $u(t, \cdot)$ where $(t, x) \mapsto u(t, x)$ is the solution of the considered PDE. We can reasonably postulate that the closure of $\{u \in L^1(\mathbb{R}), u \geq 0, \int_{\mathbb{R}} u(x)dx = 1 \mid \beta(u) = 0\}$ is a limiting set, provided it is not empty as in the case $\beta(u) = u^3$.

In Section 6.3, we discuss empirically the dependence of the error on the two parameters: the number of particles n and the time step, for the Heaviside and the porous media cases. If we simulate a fixed static probability density (the initial condition) with a large n , a Monte Carlo error of order 10^{-2} is produced. This is a sort of threshold that we cannot reasonably improve for the solution at any positive time. We have observed, that decreasing the time step Δt one would have rapidly approached that threshold. This suggests that to keep the algorithm performing it is not useful to use a time step smaller than some value $(\Delta t)_0$.

The paper is organized as follows. Section 2 is devoted to the statements of existence and uniqueness results for both the deterministic problem (1.1) and the non-linear SDE (1.4) rendering the probabilistic representation of (1.1). We in particular, recall the results given by authors of [9, 3] and we establish some additional theoretical results in the case when the initial condition of (1.1) belongs to $L^1(\mathbb{R})$ but it is not necessarily bounded.

In Section 3, we settle the theoretical basis for the implementation of our probabilistic algorithm. We first approximate the NLSDE (1.4) by a mollified version replacing $u(t, \cdot)$, the law density of Y_t , by a given smooth function. We then construct an interacting particle system for which we supposed that propagation of chaos result is verified. We drive the attention on Theorem 3.2 which links the mollified PDE (3.3) with its probabilistic representation.

Section 4 is devoted to the numerical procedure implementing the probabilistic algorithm. We first introduce an Euler scheme to obtain a discretized version of the interacting particles system defined in Section 3. We then discuss the optimal choice

of the window width ε .

In Section 5, we describe the numerical deterministic approach we use to simulate solutions of (1.1). In fact, following [18], we first use finite differences and ENO schemes for the space discretization, then we perform an explicit Runge-Kutta scheme for time integration.

In Section 6, we proceed to the validation of the algorithms. The first numerical experiments discussed in that section concern the classical porous media equation whose exact solution, in the case when the initial condition is a delta Dirac function, is the so-called *Barrenblatt-Pattle* density, see [4]. Then, we concentrate on the Heaviside case, i.e. with β of the form (1.3). In fact, we perform several test cases according to the critical threshold u_c and to the initial condition u_0 . Finally, we conclude this section by some considerations about the long time behavior of solutions of (1.1) and the performance of the algorithm.

2. Existence and uniqueness results

We start with some basic analytical framework. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a bounded function we will denote $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$. By $S(\mathbb{R})$ we denote the space of rapidly decreasing infinitely differentiable functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$. We denote by $\mathcal{M}(\mathbb{R})$ and $\mathcal{M}_+(\mathbb{R})$ the set of finite measures and positive finite measures respectively.

2.1. The deterministic PDE

Based on some clarifications of some classical papers [6, 15, 7], [9] states the following theorem about existence and uniqueness in the sense of distributions (in a proper way).

Proposition 2.1. *Let $u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$, $u_0 \geq 0$. We suppose the validity of Assumptions (A) and (B). Then there is a unique solution in the sense of distributions $u \in (L^1 \cap L^\infty)([0, T] \times \mathbb{R})$ of*

$$\begin{cases} \partial_t u & \in \frac{1}{2} \partial_{xx}^2 \beta(u), \\ u(0, x) & = u_0(x), \end{cases} \quad (2.1)$$

in the sense that, there exists a unique couple $(u, \eta_u) \in ((L^1 \cap L^\infty)([0, T] \times \mathbb{R}))^2$ such that

$$\int u(t, x) \varphi(x) dx = \int u_0(x) \varphi(x) dx + \frac{1}{2} \int_0^t ds \int \eta_u(s, x) \varphi''(x) dx, \forall \varphi \in S(\mathbb{R})$$

and

$$\eta_u(t, x) \in \beta(u(t, x)) \text{ for } dt \otimes dx\text{-a.e. } (t, x) \in [0, t] \times \mathbb{R}$$

Furthermore, $\|u(t, \cdot)\|_\infty \leq \|u_0\|_\infty$ for every $t \in [0, T]$ and there is a unique version of u such that $u \in C([0, T]; L^1(\mathbb{R})) \subset L^1([0, T] \times \mathbb{R})$.

One significant difficulty of previous framework is that the coefficient β is discontinuous; this forces us to consider β as a multivalued function even though u is single-valued. Being β , in general, discontinuous it is difficult to imagine the level of space regularity of the solution $u(t, \cdot)$ at time t . In fact, Proposition 4.5 of [3] says that almost surely $\eta_u(t, \cdot)$ belongs dt -a.e in $H^1(\mathbb{R})$ if $u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$. This helps in some cases to visualize the behavior of $u(t, \cdot)$. The proposition below makes some assertions when β is of the type of (1.3), which constitutes our pattern situation.

Proposition 2.2. *Let us suppose $u_0 \in (L^1 \cap L^\infty)(\mathbb{R})$ and β defined by (1.3). For $t \geq 0$, we denote by*

$$E_t^0 = \{x \mid u(t, x) = u_c\}.$$

For almost all $t > 0$,

- (i) E_t^0 has a non empty interior;
- (ii) every point of E_t^0 is either a local minimum or a local maximum.

Remark 2.3. The first point of the previous proposition means that at almost each time $t > 0$, the function $u(t, \cdot)$ remains constant on some interval.

The second point means that if the function $u(t, \cdot)$ crosses the barrier u_c , it has first to stay constant for some time.

Proof of Proposition 2.2. For the sake of simplicity we fix $t > 0$ such that $\eta_u(t, \cdot) \in H^1(\mathbb{R})$ and we write $u = u(t, \cdot)$, $\eta_u = \eta_u(t, \cdot)$.

- (i) Since $\eta_u \in H^1(\mathbb{R})$ it is continuous, then the set $\mathcal{D}_0 = \{x \in \mathbb{R} \mid \eta_u(x) \in]0, u_c[\}$ is open. If $\eta_u(x) \in]0, u_c[$ necessarily we have $u(x) = u_c$; in fact, if $u(x) < u_c$ then $\eta_u(x) = 0$ and if $u(x) > u_c$ then $\eta_u(x) = u(x) > u_c$. Since \mathcal{D}_0 is open and it is included in E_t^0 the result is established.
- (ii) Suppose the existence of sequences (x_n) and (y_n) such that $x_n \rightarrow x$ with $u(x_n) < u_c$ and $y_n \rightarrow y$ with $u(y_n) > u_c$. By continuity of η_u we have

$$\eta_u(x_n) = 0 \xrightarrow{n \rightarrow \infty} 0 = \eta_u(x)$$

$$u(y_n) = \eta_u(y_n) \xrightarrow{n \rightarrow \infty} \eta_u(x) = 0,$$

this is not possible because $u(y_n) > u_c$ for every n .

□

If $u_0 \in \mathcal{M}(\mathbb{R})$, we do not know any existence or uniqueness theorem for (1.1). Our first target consisted in providing some generalization to Proposition 2.1 in the case when u_0 is a finite measure. A solution in that case would be, $u :]0, T] \times \mathbb{R} \rightarrow L^1(\mathbb{R})$ continuous and such that

$$\lim_{t \rightarrow 0} u(t, dx) = u_0(dx),$$

weakly and where $u(t, dx)$ denotes $u(t, x)dx$. This is still an object of further technical investigations. For the moment, we are only able to consider the case u_0 having a $L^1(\mathbb{R})$ density still denoted by u_0 , not necessarily bounded as in Proposition 2.1, at least when Φ characterized by (1.2) is continuous. In particular β is also continuous, but possibly degenerate. In that case, we can prove existence of a distributional solution to (1.1). Even though this is not a very deep observation, this will settle the basis of the corresponding probabilistic representation, completely unknown in the literature. In fact, we provide the following result.

Proposition 2.4. *Let $u_0 \in L^1(\mathbb{R})$. Furthermore, we suppose that Assumption(A) and Assumption(B) are fulfilled. We assume that Φ is continuous on \mathbb{R}_+ .*

(1) *There is a solution u , in the sense of distributions, to the problem*

$$\begin{cases} \partial_t u(t, x) &= \frac{1}{2} \partial_{xx}^2 \beta(u(t, x)), \quad t \in [0, \infty[, \\ u(0, x) &= u_0(dx), \quad x \in \mathbb{R}, \end{cases} \quad (2.2)$$

in the sense that for every $\alpha \in S(\mathbb{R})$

$$\int_{\mathbb{R}} u(t, x) \alpha(x) dx = \int_{\mathbb{R}} u_0(x) \alpha(x) dx + \frac{1}{2} \int_0^t ds \int_{\mathbb{R}} \alpha''(x) \beta(u(s, x)) dx. \quad (2.3)$$

(2) *If u_0 is locally of bounded variation except eventually on a discrete number of points D_0 , then $\Phi(u(t, \cdot))$ has at most countable discontinuities for every $t \in [0, T]$.*

Proof. (1) Let $u_0 \in L^1(\mathbb{R})$, $u_0^N = u_0 * \phi_{\frac{1}{N}}$, $N \in \mathbb{N}^*$, where ϕ is a kernel with compact support and $\phi_{\frac{1}{N}}(x) = N\phi(Nx)$, $x \in \mathbb{R}$. So u_0^N is of class C^1 , therefore locally with bounded variation. Since $\|u_0^N\|_{\infty} \leq \|\phi_{\frac{1}{N}}\|_{\infty} \|u_0\|_{L^1}$, then $u_0^N \in (L^1 \cap L^{\infty})(\mathbb{R})$. Moreover, we have

$$\int_{\mathbb{R}} |u_0^N(x) - u_0(x)| dx \rightarrow 0, \quad \text{as } N \rightarrow +\infty.$$

On one hand, according to Proposition 2.1, there is a unique solution u^N of (2.3), i.e. for every $\alpha \in S(\mathbb{R})$

$$\int_{\mathbb{R}} u^N(t, x) \alpha(x) dx = \int_{\mathbb{R}} u_0^N(x) \alpha(x) dx + \frac{1}{2} \int_0^t ds \int_{\mathbb{R}} \alpha''(x) \beta(u^N(s, x)) dx. \quad (2.4)$$

On the other hand, according to Corollary 8.2 in Chap IV of [44], we have

$$\sup_{t \leq T} \int_{\mathbb{R}} |u^N(t, x) - u(t, x)| dx \rightarrow 0, \quad \text{as } N \rightarrow +\infty. \quad (2.5)$$

Therefore, there is a subsequence $(N_k)_{k \in \mathbb{N}}$ such that

$$u^{N_k}(t, x) \rightarrow u(t, x) \quad dt \otimes dx\text{-a.e.}, \text{ as } k \rightarrow +\infty.$$

Since β is continuous, it follows that

$$\beta(u^{N_k}(t, x)) \rightarrow \beta(u(t, x)) \quad dt \otimes dx\text{-a.e.}, \text{ as } k \rightarrow +\infty.$$

Consequently, (2.4) implies

$$\int_{\mathbb{R}} u(t, x) \alpha(x) dx = \int_{\mathbb{R}} u_0(x) \alpha(x) dx + \lim_{k \rightarrow +\infty} \frac{1}{2} \int_0^t ds \int_{\mathbb{R}} \alpha''(x) \beta(u^{N_k}(s, x)) dx. \quad (2.6)$$

In order to show that u solves (2.3), we verify

$$\lim_{N \rightarrow \infty} \int_0^t ds \int_{\mathbb{R}} \alpha''(x) \beta(u^N(s, x)) dx = \int_0^t ds \int_{\mathbb{R}} \alpha''(x) \beta(u(s, x)) dx, \quad (2.7)$$

where for notational simplicity we have replaced N_k with N . So, we can suppose that

$$u^N \rightarrow u, \quad \beta(u^N) \rightarrow \beta(u), \quad dt \otimes dx\text{-a.e.} \quad \text{as } N \rightarrow +\infty. \quad (2.8)$$

Since $|\beta(u^N)| \leq c|u^N|$ and $u^N \rightarrow u$ in $L^1([0, T] \times \mathbb{R})$, it follows that $\beta(u^N)$ are equi-integrable. Consequently, by (2.8), $\beta(u^N) \rightarrow \beta(u)$ in $L^1([0, T] \times \mathbb{R})$, and therefore (2.7) follows. Finally, u solves equation (2.3).

(2) For this purpose we state a lemma concerning an elliptic equation whose first statement item constitutes the kernel of the proof of Proposition 2.1.

Given $f : \mathbb{R} \rightarrow \mathbb{R}$, for $h \in \mathbb{R}$, we denote

$$f^h(x) = f(x + h) - f(x).$$

Lemma 2.5. *Let $f \in L^1$, $\lambda > 0$.*

(i) *There is a unique solution in the sense of distributions of*

$$u - \lambda(\beta(u))'' = f.$$

(ii) *Let χ be a smooth function with compact support. Then for each h*

$$\int_{\mathbb{R}} \chi(x) |u^h(x)| dx \leq \int_{\mathbb{R}} \chi(x) |f^h(x)| dx + C\lambda|h| \|u\|_{L^1}, \quad (2.9)$$

where C is a constant depending on β and χ .

Proof of Lemma 2.5. (i) is stated in Theorem 4.1 of [6] and Theorem 1 of [7].

(ii) The statement appears in Lemma 3.6 of [3] in the case when $f \in L^1 \cap L^\infty$ but the proof remains the same for $f \in L^1$. \square

We go on with the proof of Proposition 2.4, point (2). Let χ be a smooth nonnegative function with compact support on $\mathbb{R} \setminus D_0$. We prove in fact

$$\limsup_{h \rightarrow 0} \frac{1}{h} \int_{\mathbb{R}} \chi(x) |u^h(t, x)| dx \leq \|u_0 \chi\|_{\text{var}} + C \int_{[0, T] \times \mathbb{R}} |u(s, x)| ds dx, \quad (2.10)$$

where $\|\cdot\|_{\text{var}}$ denotes the total variation and C is a generic universal constant. For this purpose, we proceed exactly as in the proof of Proposition 4.20 of [3] making use of Lemma 2.5. Inequality (2.10) allows, similarly as in [3] to show that $u(t, \cdot)$ restricted to any compact interval of $\mathbb{R} \setminus D_0$ has bounded variation. Therefore it has at most countable discontinuities. Consequently, $\Phi(u(t, \cdot))$ has the same property since Φ is supposed to be continuous. \square

2.2. The non-linear stochastic differential equation (NLSDE)

Definition 2.6. We say that a process Y is a solution to the NLSDE associated to problem (1.1), if there exists χ belonging to $L^\infty([0, T] \times \mathbb{R})$ such that;

$$\begin{cases} Y_t &= Y_0 + \int_0^t \chi(s, Y_s) dW_s, \\ \chi(t, x) &\in \Phi(u(t, x)), \text{ for } dt \otimes dx - \text{a.e. } (t, x) \in [0, T] \times \mathbb{R}, \\ u(t, x) &= \text{Law density of } Y_t, \quad t > 0, \\ u(0, \cdot) &= u_0, \end{cases} \quad (2.11)$$

where W is a Brownian motion on some suitable filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. In particular, the first identity of (2.11) holds in law. We introduce a notion appearing in [3].

Definition 2.7. We say that β is *strictly increasing after some zero* if there is a constant $c > 0$, such that

- i) $\beta|_{[0, c]} = 0$.
- ii) β is strictly increasing on $[c, +\infty[$.

Up to now, two results are available concerning existence and uniqueness of solutions to (2.11). In fact, the first one is stated in the case where β is not degenerate and the second one in the case when β is degenerate, see respectively [9, 3]. We summarize these two results in the following theorem for easy reference later on.

Theorem 2.8. *Let $u_0 \in L^1 \cap L^\infty$ such that $u_0 \geq 0$ and $\int_{\mathbb{R}} u_0(x) dx = 1$. Furthermore, we suppose that Assumptions (A) and (B) are fulfilled.*

- (i) *If β is non-degenerate then it exists a solution Y to (2.11), unique in law.*
- (ii) *Suppose β is degenerate and either β is strictly increasing after some zero or u_0 has locally bounded variation. Then there is a solution Y not necessarily unique to (2.11).*

A step forward is constituted by the proposition below. This provides an existence result for the NLSDE, when u_0 is not necessarily bounded at least whenever Φ is continuous. This does not require a non-degenerate hypothesis on β .

Theorem 2.9. *Let $u_0 \in L^1(\mathbb{R})$ having locally bounded variation except on a discrete set of points D_0 . Furthermore we suppose that Assumption(A) and Assumption(B) are fulfilled. We assume that Φ is continuous on \mathbb{R}_+ .*

The probabilistic representation related to (1.1) holds, i.e. there is a process Y solving (1.4) in law.

Proof. Let u_0^N be the function considered at the beginning of the proof of Proposition 2.4. According to Theorem 2.8, let Y_0^N be the solution to

$$\begin{cases} Y_t^N &= Y_0^N + \int_0^t \Phi(u^N(s, Y_s^N)) dW_s, \\ u^N(t, \cdot) &= \text{Law density of } Y_t^N, \quad \forall t \geq 0, \\ u^N(0, \cdot) &= u_0^N. \end{cases} \quad (2.12)$$

Since Φ is bounded, using Burkholder-Davies-Gundy inequality one obtains

$$E (Y_t^N - Y_s^N)^4 \leq \text{const}(t - s)^2.$$

This implies (see for instance Problem 4.11, Section 2.4 of [31]) that the laws of Y^N , $N \geq 1$ are tight. Consequently, there is a subsequence $Y^k := Y^{N_k}$ converging in law (as $C([0, T])$ -valued random elements) to some process Y . We set $u^k = u^{N_k}$, where we recall that $u^k(t, \cdot)$ is the law of Y_t^k . We also set $X_t^k = Y_t^k - Y_0^k$. Since $[X^k]_t = \int_0^t \Phi^2(u^k(s, Y_s^k)) ds$ and Φ is bounded, the continuous local martingales X^k are indeed martingales.

By Skorokhod's theorem there is a new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and processes \tilde{Y}^k , with the same distribution as Y^k so that \tilde{Y}^k converges \tilde{P} -a.s. to some process \tilde{Y} , of course distributed as Y , as $C([0, T])$ -valued random element. In particular, the processes $\tilde{X}_t^k = \tilde{Y}_t^k - \tilde{Y}_0^k$ remain martingales with respect to the filtration generated by \tilde{Y}^k . We denote the sequence \tilde{Y}^k (resp. \tilde{Y}), again by Y^k (resp. Y).

We now aim to prove that

$$Y_t = Y_0 + \int_0^t \Phi(u(s, Y_s)) dW_s, \quad (2.13)$$

for some standard Brownian motion W with respect with some filtration (\mathcal{F}_t) .

We consider the stochastic process X (vanishing at zero) defined by $X_t = Y_t - Y_0$. We also set again $X_t^k = Y_t^k - Y_0^k$. Taking into account Theorem 4.2 in Chap 3 of [31], to establish (2.13), it will be enough to prove that X is an \mathcal{Y} -martingale with quadratic variation $[X]_t = \int_0^t \Phi^2(u(s, Y_s)) ds$, where \mathcal{Y} is the canonical filtration associated with Y .

Let $s, t \in [0, T]$, with $t > s$ and ψ a bounded continuous function from $C([0, s])$ to \mathbb{R} . In order to prove the martingale property for X , we need to show that

$$E [(X_t - X_s)\psi(Y_r, r \leq s)] = 0. \quad (2.14)$$

Since Y^k are martingales, we have

$$E [(X_t^k - X_s^k)\psi(Y_r^k, r \leq s)] = 0. \quad (2.15)$$

Consequently (2.14) follows from (2.15) and the fact that $Y^k \rightarrow Y$ a.s. ($X^k \rightarrow X$ a.s.) as $C([0, T])$ -valued random process. In fact for each $t \geq 0$, $X_t^k \rightarrow X_t$ in $L^1(\Omega)$ since $(X_t^k, k \in \mathbb{N})$ is bounded in $L^2(\Omega)$.

It remains to show that $X_t^2 - \int_0^t \Phi^2(u(s, Y_s))ds$, $t \in [0, T]$, defines an \mathcal{Y} -martingale, that is, we need to verify

$$E \left[\left(X_t^2 - X_s^2 - \int_s^t \Phi^2(u(r, Y_r))dr \right) \psi(Y_r, r \leq s) \right] = 0.$$

We proceed similarly as in the proof of Theorem 4.3 in [9] but even with some simplification. For the comfort of the reader we give a complete proof.

The left-hand side decomposes into $I^1(k) + I^2(k) + I^3(k)$, where

$$\begin{aligned} I^1(k) &= E \left[\left(X_t^2 - X_s^2 - \int_s^t \Phi^2(u(r, Y_r))dr \right) \psi(Y_r, r \leq s) \right] \\ &\quad - E \left[\left((X_t^k)^2 - (X_s^k)^2 - \int_s^t \Phi^2(u(r, Y_r^k))dr \right) \psi(Y_r^k, r \leq s) \right], \\ I^2(k) &= E \left[\left((X_t^k)^2 - (X_s^k)^2 - \int_s^t \Phi^2(u^k(r, Y_r^k))dr \right) \psi(Y_r^k, r \leq s) \right], \\ I^3(k) &= E \left[\left(\int_s^t (\Phi^2(u^k(r, Y_r^k)) - \Phi^2(u(r, Y_r^k)))dr \right) \psi(Y_r^k, r \leq s) \right]. \end{aligned}$$

We start by showing the convergence of $I^3(k)$. Now, $\psi(Y_r^k, r \leq s)$ is dominated by a constant C . Clearly we have

$$I^3(k) \leq C \int_s^t dr \int_{\mathbb{R}} |\Phi^2(u^k(r, y)) - \Phi^2(u(r, y))| u^k(r, y) dy.$$

The right hand side of this inequality is equal to $C[J^1(k) + J^2(k)]$, where

$$\begin{aligned} J^1(k) &= \int_s^t dr \int_{\mathbb{R}} |\Phi^2(u^k(r, y)) - \Phi^2(u(r, y))| (u^k(r, y) - u(r, y)) dy, \\ J^2(k) &= \int_s^t dr \int_{\mathbb{R}} |\Phi^2(u^k(r, y)) - \Phi^2(u(r, y))| u(r, y) dy. \end{aligned}$$

Since $u^k \rightarrow u$ in $C([0, T]; L^1)$ and Φ^2 is bounded then $\lim_{k \rightarrow +\infty} J^1(k) = 0$.

Furthermore, there is a subsequence $(k_n)_{n \in \mathbb{N}}$ such that

$$u^{k_n}(r, y) \rightarrow u(r, y) \quad dr \otimes dy - \text{a.e. as } n \rightarrow +\infty.$$

Since Φ^2 is continuous, it follows that

$$\Phi^2(u^{k_n}(r, y)) \rightarrow \Phi^2(u(r, y)) \quad dr \otimes dy - \text{a.e., } N \rightarrow +\infty$$

On the other hand, since

$$|\Phi^2(u^{k_n}(r, y)) - \Phi^2(u(r, y))| \leq 2 \sup_{u \in \mathbb{R}} \Phi^2(u) |u(r, y)|,$$

Lebesgue's dominated convergence Theorem implies that $\lim_{k \rightarrow +\infty} J^2(k) = 0$.

Now we go on with the analysis of $I^2(k)$ and $I^1(k)$. $I^2(k)$ equals zero since X^k is a martingale with quadratic variation given by $[X]_t = \int_0^t \Phi^2(u^k(r, Y_r^k)) dr$.

Finally, we treat $I^1(k)$. We recall that $X^k \rightarrow X$ a.s. as a random element in $C([0, T])$ and that the sequence $E((X_t^k)^4)$ is bounded, so $(X_t^k)^2$ are uniformly integrable.

Therefore, we have

$$E [((X_t^k)^2 - (X_s^k)^2) \psi(Y_r, r \leq s)] - E \left[((X_t^k)^2 - (X_s^k)^2) \psi(Y_r^k, r \leq s) \right] \rightarrow 0$$

when k goes to infinity. It remains to prove that

$$\int_s^t E \left[\left(\Phi^2(u(r, Y_r)) - \Phi^2(u(r, Y_r^k)) \right) \psi(Y_r, r \leq s) dr \right] \rightarrow 0. \quad (2.16)$$

Now, for fixed dr -a.e., $r \in [0, T]$, the set $\mathcal{S}(r)$ of discontinuities of $\Phi(u(r, \cdot))$ is countable because of Proposition 2.4, point (2). The law of Y_r has a density and it is therefore non-atomic. Let $N(r)$ be the event of all $\omega \in \Omega$ such that $Y_r(\omega)$ belongs to $\mathcal{S}(r)$. The probability of $N(r)$ equals $E(1_{\mathcal{S}(r)}(Y_r)) = \int_{\mathbb{R}} 1_{\mathcal{S}(r)}(y) dv(y) = 0$, where v is the law of Y_r . Consequently $N(r)$ is a negligible set.

For $\omega \notin N(r)$, we have $\lim_{k \rightarrow +\infty} \Phi^2(u(r, Y_r^k(\omega))) = \Phi^2(u(r, Y_r(\omega)))$. Since Φ is bounded, Lebesgue's dominated convergence theorem implies (2.16).

Concerning the question whether $u(t, \cdot)$ is the law of Y_t , we recall that for all t , Y_t^k converges (even in probability) to Y_t and $u^k(t, \cdot)$, which is the law density of Y_t^k , goes to $u(t, \cdot)$ in $L^1(\mathbb{R})$. By the uniqueness of the limit in (2.3), this obviously implies that $u(t, \cdot)$ is the law density of Y_t . \square

3. Some complements related to the NLSDE

3.1. A mollified version

We suppose here that $u_0(dx)$ is a probability measure. Let Y_0 be a random variable distributed according to $u_0(dx)$ and independent of the Brownian motion W .

In preparation to numerical probability simulations, we define K_ε for every $\varepsilon > 0$, as a smooth regularization kernel obtained from a fixed probability density function K by the scaling :

$$K_\varepsilon(x) = \frac{1}{\varepsilon} K\left(\frac{x}{\varepsilon}\right), \quad x \in \mathbb{R}. \quad (3.1)$$

We suppose in this section that Φ is single valued, therefore continuous. This hypothesis will not be in force in Sections 4 and 6.

In this subsection we wish to comment about the mollified version of the NLSDE (1.4), given by

$$\begin{cases} Y_t^\varepsilon &= Y_0 + \int_0^t \Phi((K_\varepsilon * v^\varepsilon)(s, Y_s^\varepsilon)) dW_s, \\ v^\varepsilon(t, \cdot) &= \text{Law of } Y_t^\varepsilon, \quad \forall t > 0, \\ v^\varepsilon(0, \cdot) &= u_0 \end{cases} \quad (3.2)$$

and its relation to the nonlinear integro-differential PDE

$$\begin{cases} \partial_t v^\varepsilon(t, x) &= \frac{1}{2} \partial_{xx}^2 (\Phi^2(K_\varepsilon * v^\varepsilon(t, x)) v^\varepsilon(t, x)), \quad (t, x) \in]0, +\infty[\times \mathbb{R}, \\ v^\varepsilon(0, \cdot) &= u_0. \end{cases} \quad (3.3)$$

where, $t \mapsto v^\varepsilon(t, \cdot)$ may be measure-valued.

Remark 3.1. (i) When Φ is Lipschitz, the authors of [29] proved in Proposition 2.2, that the problem (3.2) is well-posed. Their proof is based on a fixed point theorem with respect to the Kantorovitch-Rubinstein metric.

(ii) At our knowledge, there are no existence and uniqueness results for (3.3) at least when Φ is not smooth.

(iii) By Itô's formula, similarly to the proof of Proposition 1.3, it is easy to see that a solution Y^ε of (3.2) provides a solution v^ε of (3.3), in the sense of distributions.

When β is non-degenerate it is possible to show that formulations (3.3) and (3.2) are equivalent. In particular we have the following result.

Theorem 3.2. *We suppose that β is non-degenerate and $\varepsilon > 0$ is fixed.*

(1) *If Y^ε is a solution of (3.2) then $v^\varepsilon : [0, T] \rightarrow \mathcal{M}(\mathbb{R})$, where $v^\varepsilon(t, \cdot)$ is the law of Y_t^ε , is a solution of (3.3) and fulfills the following property*

(P) *v^ε has a density, still denoted v^ε such that: $(t, x) \mapsto v^\varepsilon(t, x) \in L^2([0, T] \times \mathbb{R})$.*

(2) If v^ε is a solution to (3.3) fulfilling (P) then there is a process $Y = Y^\varepsilon$ solving (3.2).

Proof. (1) If Y^ε is a solution to (3.2) by Remark 3.1.(iii) it follows that v^ε fulfills (3.3).

On the other hand, since $K_\varepsilon * v^\varepsilon$ is bounded and Φ is lower bounded by a constant C_ε on $[-\inf K_\varepsilon * v^\varepsilon, \sup K_\varepsilon * v^\varepsilon]$ it follows that $a(t, x) = \Phi^2(K_\varepsilon * v^\varepsilon(t, x))$ is lower bounded by C^ε .

Using then Exercise 7.3.3 of [47], i.e., Krylov estimates, it follows that for every smooth function $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ with compact support, we have

$$E \left(\int_0^T f(Y_s^\varepsilon) ds \right) \leq \text{const} \|f\|_{L^2([0, T] \times \mathbb{R})}.$$

Then, developing the left hand side to obtain

$$\int_0^T ds \int_{\mathbb{R}} f(y) v^\varepsilon(s, y) dy \leq \text{const} \|f\|_{L^2([0, T] \times \mathbb{R})}.$$

we deduce that (P) is verified.

(2) We retrieve here some arguments used in the proof of Proposition 4.2 of [9].

Given $v = v^\varepsilon$, by Remark 4.3 of [9], see also Exercise 7.3.2-7.3.4 of [47], we can construct a unique solution $Y = Y^\varepsilon$ in law to the SDE constituted by

$$Y_t = Y_0 + \int_0^t \sqrt{a(s, Y_s)} dW_s, \quad (3.4)$$

where here $a(t, x) = \Phi^2(K_\varepsilon * v(t, x))$. Indeed, this is possible again because a is Borel bounded and lower bounded by a strictly positive constant.

A further use of Itô's formula says that the law $z(t, dx)$ of Y_t solves

$$\begin{cases} \partial_t z(t, \cdot) &= \frac{1}{2} \partial_{xx}^2 (a(t, \cdot) z(t, \cdot)), \\ z(0, \cdot) &= u_0, \end{cases} \quad (3.5)$$

in the sense of distributions.

Using again Krylov estimates as in the second part of the proof of point (1), it follows that z admits a density $(t, y) \mapsto p_t(y)$ which verifies $p \in L^2([0, T] \times \mathbb{R})$. This shows that Hypothesis 3.4 in Theorem 3.3 below is fulfilled, which implies that $v \equiv z$. \square

Theorem 3.3 was stated and proved in [9], see Theorem 3.8.

Theorem 3.3. *Let a be a Borel nonnegative bounded function on $[0, T] \times \mathbb{R}$.*

Let $z_i : [0, T] \rightarrow \mathcal{M}_+(\mathbb{R})$, $i = 1, 2$, be continuous with respect to the weak topology on finite measures on $\mathcal{M}(\mathbb{R})$.

Let z^0 be an element of $\mathcal{M}_+(\mathbb{R})$. Suppose that both z_1 and z_2 solve the problem $\partial_t z = \partial_{xx}^2(az)$ in the sense of distributions with initial condition $z(0, \cdot) = z^0$.

More precisely

$$\int_{\mathbb{R}} \phi(x) z(t, dx) = \int_{\mathbb{R}} \phi(x) z^0(dx) + \int_0^t ds \int_{\mathbb{R}} \phi''(x) a(s, x) z(s, dx)$$

for every $t \in [0, T]$ and any $\phi \in C_0^\infty(\mathbb{R})$.

Then $(z_1 - z_2)(t, \cdot)$ is identically zero for every t , if $z := z_1 - z_2$, satisfies the following:

Hypothesis 3.4. *There is $\rho : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ belonging to $L^2([\kappa, T] \times \mathbb{R})$ for every $\kappa > 0$ such that $\rho(t, \cdot)$ is the density of $z(t, \cdot)$ for almost all $t \in]0, T]$.*

3.2. The interacting particles system

We recall that in this paper, we want to approximate solutions of problem (1.1). For this purpose we will concentrate on a probabilistic particles system of the same nature as in [29] when the coefficients are Lipschitz.

In general, the particles probabilistic algorithms for non linear PDEs are based on the simulation of particles trajectories animated by a random motion. The solution of the PDE is approximated through the smoothing of the empirical measure of the particles, which is a linear combination of Dirac masses centered on particles positions. This procedure is heuristically justified by the chaos propagation phenomenon which will be explained in the sequel.

The dynamics of the particles is described by the following stochastic differential system:

$$Y_t^{i,\varepsilon,n} = Y_0^i + \int_0^t \Phi \left(\frac{1}{n} \sum_{j=1}^n K_\varepsilon(Y_s^{i,\varepsilon,n} - Y_s^{j,\varepsilon,n}) \right) dW_s^i, \quad i = 1, \dots, n \quad (3.6)$$

where $W = (W^1, \dots, W^n)$ is an n -dimensional Brownian motion, $(Y_0^i)_{1 \leq i \leq n}$ is a sequence of independent random variables with law density u_0 and independent of the Brownian motion W and K_ε is the same kernel as in Subsection 3.1.

Remark 3.5. If Φ in the system of ordinary SDEs (3.6) were not continuous but only measurable, that problem would not have necessarily a solution, even if β were non-degenerate. In fact, contrarily to (3.4), here $n \geq 2$. Since Φ is continuous, then (3.6) has at least a solution; if Φ is non-degenerate even uniqueness holds, see Chapter 6 and 7 of [47].

Now, owing to the interacting kernel K_ε , the particles motions are a priori dependent. For a given integer n , we consider $(Y_t^{1,\varepsilon,n}, \dots, Y_t^{n,\varepsilon,n})$ as the solution of the interacting particle system (3.6). Propagation of chaos for the mollified equation happens if for any integer m , the vector $(Y_t^{1,\varepsilon,n}, \dots, Y_t^{m,\varepsilon,n})_{n \geq m}$ converges in law to $\mu_t^{\otimes m}$ where μ_t is the law of Y_t^ε the solution of (3.2).

A consequence of chaos propagation is that one expects that the empirical measure of the particles, i.e. the linear combination of Dirac masses denoted $\mu_t^n = \frac{1}{n} \sum_{j=1}^n \delta_{Y_t^{j,\varepsilon,n}}$ converges in law as a random measure to the deterministic solution $v_\varepsilon(t, \cdot)$ of the regularized PDE (3.3) which in fact depends on ε . This fact was established for instance when β is Lipschitz, in Proposition 2.2 of [29]. On the other hand when ε goes to zero, the same authors show that v_ε converge to the solution u of (1.1). They prove the existence of a sequence $(\varepsilon(n))$ slowly converging to zero when n goes to infinity such that the empirical measure $\frac{1}{n} \sum_{j=1}^n \delta_{Y_t^{j,\varepsilon(n),n}}$, converges in law to u , see Theorem 2.7 of [29]. One consequence of the slow convergence is that the regularized empirical measure

$$\frac{1}{n} \sum_{j=1}^n K_{\varepsilon(n)}(\cdot - Y_t^{j,\varepsilon(n),n})$$

also converges to u . Consequently, this probabilistic interpretation provides an algorithm allowing to solve numerically (1.1).

We recall however that one of the significant objects of this paper is the numerical implementation related to the case when, β is possibly discontinuous; for the moment we do not have convergence results but we implement the same type of algorithm and we compare with some existing deterministic schemes.

4. About probabilistic numerical implementations

In this section we will try to construct an approximation method for solutions u of (1.1), based upon the time discretization of the system (3.6). For now on, the number n of particles is fixed.

In fact, to get a simulation procedure for a trajectory of each $(Y_t^{i,\varepsilon,n})$, $i = 1, \dots, n$, we discretize in time: for fixed $T > 0$, we choose a time step $\Delta t > 0$ and $N \in \mathbb{N}$, such that $T = N\Delta t$. We denote by $t_k = k\Delta t$, the discretization times for $k = 0, \dots, N$.

The Euler explicit scheme of order one, leads then to the following discrete time system, i.e., for every $i = 1, \dots, n$

$$X_{t_{k+1}}^i = X_{t_k}^i + \Phi \left(\frac{1}{n} \sum_{j=1}^n K_\varepsilon(X_{t_k}^i - X_{t_k}^j) \right) (W_{t_{k+1}}^i - W_{t_k}^i), \quad (4.1)$$

where at each time step t_k , we approximate $u(t_k, \cdot)$ by the smoothed empirical measure of the particles :

$$u^{\varepsilon,n}(t_k, x) = \frac{1}{n} \sum_{j=1}^n K_\varepsilon(x - X_{t_k}^j), \quad k = 1, \dots, N, \quad x \in \mathbb{R}, \quad (4.2)$$

at each time step and for every $i = 1, \dots, n$, the Brownian increment $(W_{t_{k+1}}^i - W_{t_k}^i)$ is given by the simulation of the realization of a Gaussian random variable of law $\mathcal{N}(0, \Delta t)$.

One difficult issue concerns the smoothing parameter ε related to the kernel K_ε . It will be chosen according to the *kernel density estimation*.

From now on, we will assume that K , as defined in (3.1), is a Gaussian probability density function with mean 0 and unit standard deviation. In this case, in (4.2), the function $u^{\varepsilon,n}(t_k, \cdot)$ becomes the so-called Gaussian kernel density estimator of $u(t_k, \cdot)$ for every time step t_k with $k = 1, \dots, N$.

Finally, the only unknown parameter in (4.2), is ε ; most of the authors refer to it as the *bandwidth* or the *window width*.

The optimal choice of ε was the object of an enormous amount of research, because its value strongly determines the performance of $u^{\varepsilon,n}$ as an estimator of u , see, e.g. [46] and references therein. The most widely used criterion of performance for the estimator (4.2) is the *Mean Integrated Squared Error* (MISE), defined by

$$\begin{aligned} \text{MISE}\{u^{\varepsilon,n}(t, x)\} &= \mathbb{E}_u \int [u^{\varepsilon,n}(t, y) - u(t, y)]^2 dy \\ &= \int \left(\underbrace{\mathbb{E}_u [u^{\varepsilon,n}(t, y)] - u(t, y)}_{\text{point-wise bias}} \right)^2 dy + \underbrace{\int \mathbb{V}_u [u^{\varepsilon,n}(t, y)] dy}_{\text{integrated point-wise variance}}, \end{aligned}$$

where, \mathbb{E}_u and \mathbb{V}_u are respectively the expectation and the variance of X_t^j , $j = 1, \dots, n$ under the assumption that they are independent and distributed as $u(t, \cdot)$.

We emphasize that the MISE expression is the sum of two components: the integrated bias and variance.

The asymptotic properties of (4.2) under the MISE criterion are well-known (see [46],[12]), but we summarize them below for convenience of the reader.

Theorem 4.1. (Properties of the Gaussian kernel estimator)

Under the assumption that ε depends on n such that $\lim_{n \rightarrow +\infty} \varepsilon = 0$, $\lim_{n \rightarrow +\infty} n\varepsilon = +\infty$ and $\partial_{xx}^2 u$ is a continuous square integrable function, the estimator (4.2) has integrated

squared bias and integrated variance given by

$$\|E_u [u^{\varepsilon,n}(t, \cdot) - u(t, \cdot)]\|^2 = \frac{1}{4}\varepsilon^4 \|\partial_{xx}^2 u\|^2 + o(\varepsilon^2), \quad n \rightarrow +\infty, \quad (4.3)$$

$$\int V_u [u^{\varepsilon,n}(t, y)] dy = \frac{1}{2\varepsilon n \sqrt{\pi}} + o((n\varepsilon)^{-1}), \quad n \rightarrow +\infty. \quad (4.4)$$

Remark 4.2. (i) Here $\|\cdot\|$ denotes the standard L^2 norm. The first order asymptotic approximation of MISE, denoted AMISE, is thus given by

$$\text{AMISE}\{u^{\varepsilon,n}(t, x)\} = \frac{1}{4}\varepsilon^4 \|\partial_{xx}^2 u(t, x)\|^2 + (2\varepsilon n \sqrt{\pi})^{-1}. \quad (4.5)$$

(ii) The asymptotically optimal value of ε is the minimizer of AMISE and by simple calculus it can be shown (see [38], Lemma 4A) to be equal to $\varepsilon_t^{\text{opt}}$ defined in formula (1.5).

As argued in the introduction, we have chosen to use the "solve-the-equation" bandwidth selection plug-in procedure developed in [43, 27], to perform the optimal window width of the Gaussian kernel density estimator $u^{\varepsilon,n}$ of u , defined in (4.2).

Remark 4.3. According to [43], for every positive integer s , the identity

$$\|\partial_{x^s}^s u(t, \cdot)\|^2 = (-1)^s \int_{\mathbb{R}} \partial_{x^{2s}}^{2s} u(t, x) \cdot u(t, x) dx,$$

suggests the following estimator for that density functional:

$$\|\partial_{x^s}^s u^{\varepsilon,n}(t, x)\|^2 = \frac{(-1)^s}{n^2 \varepsilon^{2s+1}} \sum_{i=1}^n \sum_{j=1}^n K^{(2s)} \left(\frac{X_t^i - X_t^j}{\varepsilon} \right) \quad (4.6)$$

where, $K^{(r)}$ is the r^{th} derivative of the Gaussian kernel K and $\partial_{x^r}^r u$ is the r^{th} partial spacial derivative of u .

Inspired by (1.5), the authors of [43, 27] look for an approached optimal bandwidth for the AMISE as the solution of the equation

$$\varepsilon := \varepsilon_t = \left(2n\sqrt{\pi} \|\partial_{xx}^2 u^{\gamma(\varepsilon_t),n}(t, x)\|^2 \right)^{-1/5} \quad (4.7)$$

where, $\|\partial_{xx}^2 u^{\gamma(\varepsilon),n}\|^2$ is an estimate of $\|\partial_{xx}^2 u\|^2$ using (4.6), for $s = 2$, and the pilot bandwidth $\gamma(\varepsilon)$, which depends on the kernel bandwidth ε . The pilot bandwidth $\gamma(\varepsilon)$ is then chosen through an intermediate step which consists in obtaining a quantity h_t minimizing the asymptotic mean squared error (AMSE) for the estimation of $\|\partial_{xx}^2 u\|^2$.

AMSE is in fact some approximation via Taylor expansion of the MSE which is defined as follows:

$$\text{MSE}\{\|\partial_{xx}^2 u^{h,n}\|^2\} = \mathbb{E}_u \left[\|\partial_{xx}^2 u^{h,n}\|^2 - \|\partial_{xx}^2 u\|^2 \right]^2. \quad (4.8)$$

Similarly one can define an analogous quantity for the third derivative

$$\text{MSE}\{\|\partial_{x^3}^3 u^{h,n}\|^2\} = \mathbb{E}_u \left[\|\partial_{x^3}^3 u^{h,n}\|^2 - \|\partial_{x^3}^3 u\|^2 \right]^2. \quad (4.9)$$

and related AMSE. Exhaustive details concerning those computations are given in [51]. In fact, the authors in [51] computed those minimizers and provided the following explicit formulae

$$h_t = \left[\frac{2K^{(4)}(0)}{n\|\partial_{x^3}^3 u(t, x)\|^2} \right]^{1/7}, \quad h_t^* = \left[\frac{-2K^{(6)}(0)}{n\|\partial_{x^4}^4 u(t, x)\|^2} \right]^{1/9}, \quad (4.10)$$

where h_t and h_t^* minimize the AMSE corresponding respectively to (4.8) and (4.9).

Solving (1.5), with respect to n and replacing n in the first equality of (4.10), gives the following expression of h_t in term of ε_t

$$h_t = \left[\frac{4\sqrt{\pi}K^{(4)}(0)\|\partial_{xx}^2 u(t, x)\|^2}{\|\partial_{x^3}^3 u(t, x)\|^2} \right]^{1/7} \varepsilon_t^{5/7}.$$

This suggests to define

$$\gamma(\varepsilon_t) = \left[\frac{4\sqrt{\pi}K^{(4)}(0)\|\partial_{xx}^2 u^{h_t^1, n}(t, x)\|^2}{\|\partial_{x^3}^3 u^{h_t^2, n}(t, x)\|^2} \right]^{1/7} \varepsilon_t^{5/7}, \quad (4.11)$$

where, $\|\partial_{xx}^2 u^{h_t^1, n}\|^2$ and $\|\partial_{x^3}^3 u^{h_t^2, n}\|^2$ are estimators of $\|\partial_{xx}^2 u\|^2$ and $\|\partial_{x^3}^3 u\|^2$ using formula (4.6) and pilot bandwidths h_t^1 and h_t^2 given by

$$h_t^1 = \left[\frac{2K^{(4)}(0)}{n\widehat{\|\partial_{x^3}^3 u(t, x)\|^2}} \right]^{1/7}, \quad h_t^2 = \left[\frac{-2K^{(6)}(0)}{n\widehat{\|\partial_{x^4}^4 u(t, x)\|^2}} \right]^{1/9};$$

$\widehat{\|\partial_{x^3}^3 u(t, x)\|^2}$ and $\widehat{\|\partial_{x^4}^4 u(t, x)\|^2}$ will be suitably defined below. Indeed, h_t^1 and h_t^2 estimate h_t and h_t^* defined in (4.10).

According to the strategy in [43, 51], we will first suppose that $\partial_{x^3}^3 u(t, x)$ and $\partial_{x^4}^4 u(t, x)$ are the third and fourth partial space derivatives of a Gaussian density with standard deviation σ_t of X_t . In a second step we replace σ_t with the empirical standard deviation $\hat{\sigma}_t$ of the sample X_t^1, \dots, X_t^n . This leads naturally to

$$\widehat{\|\partial_{x^3}^3 u(t, x)\|^2} = \frac{15}{16\sqrt{\pi}} \hat{\sigma}_t^{-7}, \quad \widehat{\|\partial_{x^4}^4 u(t, x)\|^2} = \frac{105}{32\sqrt{\pi}} \hat{\sigma}_t^{-9}.$$

Coming back to (4.7), where $\gamma(\varepsilon_t)$ is defined through (4.11), it suffices then to perform a root-finding algorithm for it at each discrete time step t_k , in order to obtain the approached optimal bandwidth ε_{t_k} .

5. Deterministic numerical approach

We recall that the final aim of our work is to approximate solutions of a nonlinear problem given by

$$\begin{cases} \partial_t u(t, x) \in \frac{1}{2} \partial_{xx}^2 \beta(u(t, x)), & t \in [0, +\infty[, \\ u(0, x) = u_0(x), & x \in \mathbb{R}, \end{cases} \quad (5.1)$$

in the case where β is given by (1.3). Despite the fact that, up to now at our knowledge, there are no analytical approaches dealing such issues, we got interested into a recent method, proposed in [18]. Actually, we are heavily inspired by [18] to implement a deterministic procedure simulating solutions of (5.1) which will be compared to the probabilistic one. [18] handles with the propagation of a discontinuous solution, even though coefficient β is Lipschitz. It seems to us that in the numerical analysis literature, [18] is the closest one to our spirit. We describe now the fully discrete scheme we will use for this purpose.

5.1. Relaxation approximation

The schemes proposed in [18] follow the same idea as the well-known relaxation schemes for hyperbolic conservation laws, see [26] for a review of the subject. For the convenience of the reader, we retrieve here some arguments of [18], where we recall that the coefficient β is Lipschitz. In that case \in , of course, becomes $=$.

The equation (5.1) can be formally expressed by the first order system on $\mathbb{R}_+ \times \mathbb{R}$:

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ v + \frac{1}{2} \partial_x \beta(u) = 0. \end{cases} \quad (5.2)$$

(5.2), is relaxed with the help of a parameter $\varepsilon > 0$, in order to obtain the following scheme

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + \frac{1}{2\varepsilon} \partial_x \beta(u) = -\frac{1}{\varepsilon} v. \end{cases} \quad (5.3)$$

Then, another function $w : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ is introduced in order to remove the non-linear term in the second line of system (5.3). So, we obtain

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + \frac{1}{2\varepsilon} \partial_x w = -\frac{1}{\varepsilon} v, \\ \partial_t w + \partial_x v = -\frac{1}{\varepsilon} (w - \beta(u)). \end{cases} \quad (5.4)$$

Note that (5.4) is a particular case of the BGK system previously studied in [13]. In fact, authors of [13] proved that w (resp. v) converges to $\beta(u)$ (resp. $-\frac{1}{2}\partial_x\beta(u)$), as $\varepsilon \rightarrow 0^+$. Furthermore, they showed the convergence of solutions of (5.4) to those of PDE (5.1), in $L^1(\mathbb{R})$, as ε goes to zero.

Finally, we introduce a supplementary parameter $\varphi > 0$, according to usual numerical analysis techniques; while preserving the hyperbolic character of the system. Therefore, we get

$$\begin{cases} \partial_t u + \partial_x v = 0, \\ \partial_t v + \varphi^2 \partial_x w = -\frac{1}{\varepsilon}v + (\varphi^2 - \frac{1}{2\varepsilon})\partial_x w, \\ \partial_t w + \partial_x v = -\frac{1}{\varepsilon}(w - \beta(u)). \end{cases} \quad (5.5)$$

Now, setting

$$z = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \quad \mathcal{F}(z) = \mathbb{A}z, \quad \mathbb{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \varphi^2 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad g(z) = \begin{pmatrix} 0 \\ -v + (\varphi^2\varepsilon - \frac{1}{2})\partial_x w \\ \beta(u) - w \end{pmatrix},$$

the system (5.5) is rewritten in matrix form as follows

$$\partial_t z + \partial_x \mathcal{F}(z) = \frac{1}{\varepsilon}g(z). \quad (5.6)$$

Using the change of variable $Z = \mathbb{P}^{-1}z$, where

$$\mathbb{P}^{-1} = \begin{pmatrix} 0 & \frac{1}{2\varphi} & \frac{1}{2} \\ 0 & \frac{-1}{2\varphi} & \frac{1}{2} \\ 1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbb{P}^{-1}\mathbb{A}\mathbb{P} = \mathbb{D} = \begin{pmatrix} \varphi & 0 & 0 \\ 0 & -\varphi & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we obtain

$$Z = \begin{pmatrix} \mathcal{U} \\ \mathcal{V} \\ \mathcal{W} \end{pmatrix}, \quad \text{with} \quad \mathcal{U} = \frac{v + \varphi w}{2\varphi}, \quad \mathcal{V} = \frac{-v + \varphi w}{2\varphi}, \quad \mathcal{W} = u - w, \quad (5.7)$$

where, \mathcal{U} , \mathcal{V} , \mathcal{W} are called characteristic variables.

Since $z = \mathbb{P}Z$, equation (5.6) leads to

$$\partial_t Z + \mathbb{D}\partial_x Z = \frac{1}{\varepsilon}\mathbb{P}^{-1}g(\mathbb{P}Z). \quad (5.8)$$

By rewriting the system (5.8) in terms of the characteristic variables, we obtain

$$\begin{pmatrix} \partial_t \mathcal{U} + \varphi \partial_x \mathcal{U} \\ \partial_t \mathcal{V} - \varphi \partial_x \mathcal{V} \\ \partial_t \mathcal{W} \end{pmatrix} = \frac{1}{\varepsilon} \mathbb{P}^{-1} g(\mathbb{P}Z). \quad (5.9)$$

Finally, solving (5.5) is equivalent to the resolution of a three advection equations system, (5.9), with respectively a positive, a negative and a zero advection velocity.

Remark 5.1. Note that we can deduce from (5.7) the following relation

$$u = \mathcal{U} + \mathcal{V} + \mathcal{W}. \quad (5.10)$$

5.2. Space discretization

In the sequel of this chapter and in Annex 7, given two integers $i < j$, $[[i, j]]$, will denote the integer interval $\{i, i+1, \dots, j\}$. We will now provide a space discretization scheme for system (5.9). Let us introduce a uniform grid on $[a, b] \subset \mathbb{R}$.

We denote $x_i = a - \frac{\Delta x}{2} + i\Delta x$, $i \in [[1, N_x]]$ and $x_{i+1/2} = a + i\Delta x$, $i \in [[0, N_x]]$, where $\Delta x = \frac{b-a}{N_x}$ is the grid spacing and N_x the number of cells. Note that x_i is the center of the interval $[x_{i-1/2}, x_{i+1/2}]$. Moreover, we denote the boundary conditions by $u(t, a) = u_a(t)$ and $u(t, b) = u_b(t)$, for every $t > 0$.

Then, we evaluate (5.9) on the grid of discrete points (x_i) getting,

$$\begin{cases} \frac{d\mathcal{U}}{dt}(t, x_i) + \varphi \frac{d\mathcal{U}}{dx}(t, x_i) = G_1(t, x_i), \quad \forall t > 0, \forall i \in [[1, N_x]], \\ \frac{d\mathcal{V}}{dt}(t, x_i) - \varphi \frac{d\mathcal{V}}{dx}(t, x_i) = G_2(t, x_i), \quad \forall t > 0, \forall i \in [[1, N_x]], \\ \frac{d\mathcal{W}}{dt}(t, x_i) = G_3(t, x_i), \quad \forall t > 0, \forall i \in [[1, N_x]], \end{cases} \quad (5.11)$$

where,

$$(G_1, G_2, G_3)^t = \frac{1}{\varepsilon} \mathbb{P}^{-1} g(\mathbb{P}Z). \quad (5.12)$$

Remark 5.2. We can easily deduce from (5.12), that for every $t \in]0, +\infty[$ and every $i \in [[1, N_x]]$, we have : $\sum_{j=1}^3 G_j(t, x_i) = 0$.

In order to ensure the convergence of the semi-discrete scheme (5.11) it is necessary to write it in a conservative form. To this aim, following [18], we suppose the existence

of functions $\widehat{\mathcal{U}}$ and $\widehat{\mathcal{V}}$ such that

$$\begin{cases} \mathcal{U}(t, x) = \frac{1}{\Delta x} \int_{x-\Delta x/2}^{x+\Delta x/2} \widehat{\mathcal{U}}(t, y) dy, & \forall x \in]a, b[, \forall t > 0, \\ \mathcal{V}(t, x) = \frac{1}{\Delta x} \int_{x-\Delta x/2}^{x+\Delta x/2} \widehat{\mathcal{V}}(t, y) dy, & \forall x \in]a, b[, \forall t > 0. \end{cases}$$

Substituting in (5.11), we obtain for every $t > 0$ and every $i \in \llbracket 1, N_x \rrbracket$,

$$\begin{cases} \frac{d\mathcal{U}}{dt}(t, x_i) + \frac{\varphi}{\Delta x} \left(\widehat{\mathcal{U}}(t, x_{i+1/2}) - \widehat{\mathcal{U}}(t, x_{i-1/2}) \right) = G_1(t, x_i), \\ \frac{d\mathcal{V}}{dt}(t, x_i) - \frac{\varphi}{\Delta x} \left(\widehat{\mathcal{V}}(t, x_{i+1/2}) - \widehat{\mathcal{V}}(t, x_{i-1/2}) \right) = G_2(t, x_i), \\ \frac{d\mathcal{W}}{dt}(t, x_i) = G_3(t, x_i). \end{cases} \quad (5.13)$$

Let us now denote by $\widetilde{\mathcal{U}}_{i+1/2}(t)$ and $\widetilde{\mathcal{V}}_{i+1/2}(t)$ the so-called *semi-discrete numerical fluxes* that approximate respectively $\widehat{\mathcal{U}}(t, x_{i+1/2})$ and $\widehat{\mathcal{V}}(t, x_{i+1/2})$. For the sake of simplicity, we chose to expose only the calculations necessary to obtain the first semi-discrete flux $\widetilde{\mathcal{U}}_{i+1/2}(t)$, the same procedure being applied for the other one.

In order to compute the numerical flux $\widetilde{\mathcal{U}}_{i+1/2}(t)$, we reconstruct boundary extrapolated data $\mathcal{U}_{i+1/2}^\pm(t)$, from the point values $\mathcal{U}_i(t) = \mathcal{U}(t, x_i)$ of the variables at the center of the cells, with an essentially non oscillatory interpolation (ENO) method. The ENO technique allows to better localize discontinuities and fronts that may appear when β is possibly degenerate; see [23, 45] for an extensive presentation of the subject. In fact, $\mathcal{U}_{i+1/2}^+(t)$ (resp. $\mathcal{U}_{i+1/2}^-(t)$) is calculated from an interpolating polynomial of degree d , on the interval $[x_{i+1/2}, x_{i+3/2}]$ (resp. $[x_{i-1/2}, x_{i+1/2}]$) using a so-called *ENO stencil*, see [45] and formula (7.5) in Annex 7.1.

Next, we shall apply a numerical flux to these boundary extrapolated data. In order to minimize the numerical viscosity and according to authors of [18], we choose the so-called *Godunov flux*, \mathfrak{F}_G , associated to the advection equation

$$\partial_t \mathcal{U} + \partial_x f(\mathcal{U}) = 0,$$

and defined as follows

$$\mathfrak{F}_G[\alpha, \gamma] = \begin{cases} \min_{\alpha \leq \xi \leq \gamma} f(\xi), & \text{if } \alpha \leq \gamma, \\ \max_{\gamma \leq \xi \leq \alpha} f(\xi), & \text{if } \gamma \leq \alpha. \end{cases}$$

where $f(\xi) = \varphi \xi$, with $\varphi > 0$. So we have, $\mathfrak{F}_G[\alpha, \gamma] = \varphi \alpha$.

In fact, we set

$$\forall t > 0, \quad \tilde{\mathcal{U}}_{i+1/2}(t) = \mathfrak{F}_G[\mathcal{U}_{i+1/2}^-(t), \mathcal{U}_{i+1/2}^+(t)]. \quad (5.14)$$

Therefore, we obtain the following semi-discrete flux

$$\forall t > 0, \quad \tilde{\mathcal{U}}_{i+1/2}(t) = \varphi \mathcal{U}_{i+1/2}^-(t). \quad (5.15)$$

Applying the previous procedure to compute $\tilde{\mathcal{V}}_{i+1/2}(t)$ and replacing in (5.13), we get for every $t > 0$ and every $i \in \llbracket 1, N_x \rrbracket$,

$$\left\{ \begin{array}{l} \frac{d\mathcal{U}}{dt}(t, x_i) + \frac{\varphi}{\Delta x} \left(\mathcal{U}_{i+1/2}^-(t) - \mathcal{U}_{i-1/2}^-(t) \right) = G_1(t, x_i), \\ \frac{d\mathcal{V}}{dt}(t, x_i) - \frac{\varphi}{\Delta x} \left(\mathcal{V}_{i+1/2}^+(t) - \mathcal{V}_{i-1/2}^+(t) \right) = G_2(t, x_i), \\ \frac{d\mathcal{W}}{dt}(t, x_i) = G_3(t, x_i). \end{array} \right. \quad (5.16)$$

Consequently summing up the three equation lines in (5.16) and using Remarks 5.1 and 5.2, we obtain

$$\frac{du}{dt}(t, x_i) + \frac{\varphi}{\Delta x} \left(\mathcal{U}_{i+1/2}^-(t) - \mathcal{U}_{i-1/2}^-(t) - \left(\mathcal{V}_{i+1/2}^+(t) - \mathcal{V}_{i-1/2}^+(t) \right) \right) = 0.$$

Now, coming back to the conservative variables, we obtain for every $i \in \llbracket 1, N_x \rrbracket$ and every $t > 0$,

$$\left\{ \begin{array}{l} \frac{du}{dt}(t, x_i) = -\frac{1}{2\Delta x} \left(v_{i+1/2}^-(t) - v_{i-1/2}^-(t) + \varphi(w_{i+1/2}^-(t) - w_{i-1/2}^-(t)) \right) \\ \quad + \frac{1}{2\Delta x} \left(v_{i-1/2}^+(t) - v_{i+1/2}^+(t) + \varphi(w_{i+1/2}^+(t) - w_{i-1/2}^+(t)) \right), \\ u(0, x_i) = u_0(x_i), \\ u(t, a) = u_a(t), \\ u(t, b) = u_b(t). \end{array} \right. \quad (5.17)$$

We recall that by formally setting $\varepsilon = 0$ in the scheme (5.5), we have $v = -\frac{1}{2}\partial_x w$ and $w = \beta(u)$. Therefore we can compute

$$v_{i+1/2}^\pm = -\frac{1}{2}(\partial_x w)_{i+1/2}^\pm \quad \text{and} \quad w_{i+1/2}^\pm = \beta(u_{i+1/2}^\pm),$$

where, $w_{i+1/2}^\pm$, are performed using again an ENO reconstruction, see formulae (7.4)-(7.5) in Annex 7.1; while the derivatives of $w_{i+1/2}^\pm$ are approximated using a reconstruction polynomial with a centered stencil, see formula (7.10)-(7.12) in Annex 7.2.

We wish to emphasize that the scheme of system (5.9) reduces to the time advancement of the single variable u solution of (5.1).

5.3. Time discretization

In order to have a fully discrete scheme, we still need to specify the time discretization. According to [18], we use a discretization based on an explicit Runge-Kutta scheme, see [37], for instance.

We start discretizing the system (5.17) using, for simplicity, a uniform time step Δt . For every $i \in \llbracket 1, N_x \rrbracket$, we denote by u_i^m the numerical approximation of $u(t^m, x_i)$ with $t^m = m\Delta t$, $m = 0, \dots, N_t$, where N_t is the number of time steps.

The ν -stage explicit Runge-Kutta scheme with $\nu \geq 1$, associated to (5.17) can be written for every $i \in \llbracket 1, N_x \rrbracket$ as follows,

$$u_i^{m+1} = u_i^m - \frac{\lambda}{2} \sum_{k=1}^{\nu} \tilde{b}_k F_i^{(k)}, \quad (5.18)$$

where, $\lambda = \frac{\Delta t}{\Delta x}$ and the stage values are computed at each time step t^m and for every $k \in \llbracket 1, \nu \rrbracket$, as

$$\begin{cases} F_i^{(k)} = v_{i+1/2}^{(k)-} - v_{i-1/2}^{(k)-} + \varphi(w_{i+1/2}^{(k)-} - w_{i-1/2}^{(k)-}) - v_{i-1/2}^{(k)+} + v_{i+1/2}^{(k)+} - \varphi(w_{i+1/2}^{(k)+} - w_{i-1/2}^{(k)+}), \\ u_i^{(k)} = u_i^m - \frac{\lambda}{2} \sum_{l=1}^{k-1} \tilde{a}_{kl} F_i^{(l)}, \quad v_{i+1/2}^{(l)\pm} = -\frac{1}{2}(\partial_x w^{(l)})_{i+1/2}^{\pm}, \quad w_{i+1/2}^{(l)\pm} = \beta(u_{i+1/2}^{(l)\pm}). \end{cases} \quad (5.19)$$

Here $(\tilde{a}_{kl}, \tilde{b}_k)$ is a pair of Butcher's tableaux [22], of diagonally explicit Runge-Kutta schemes. This finally completes the description of the deterministic numerical method.

Remark 5.3. In the case when β is Lipschitz but possibly degenerate, the authors of [18], showed the L^1 -convergence of a semi-discrete in time relaxed scheme, see Theorem 1, Section 3 in [18]. In fact, they extended the proof of [8] to the case of a ν -stages Runge-Kutta scheme. Moreover, [18] provided the following stability condition of parabolic type,

$$\Delta t \leq C\Delta x^2, \quad (5.20)$$

where, C is a constant depending on β . At the best of our knowledge, no such results are available in the case where β is not Lipschitz.

6. Numerical experiments

We use a Matlab implementation to simulate both the deterministic and probabilistic solutions. Concerning the plug-in bandwidth selection procedure described in Section 4 and based on [43], we have improved the code produced by J. S. Marron and available on http://www.stat.unc.edu/faculty/marron/marron_software.html, by speeding up the root-finding algorithm used to solve (4.7). Furthermore, the

deterministic numerical solutions are performed using the ENO spatial reconstruction of order 3 and a third order explicit Runge-Kutta scheme for time stepping. We point out that the deterministic time step, denoted from now on by Δt_{det} , is chosen with respect to the stability condition (5.20).

6.1. The Classical porous media equation

We recall that when $\beta(u) = u|u|^{m-1}$, $m > 1$, the PDE in (1.1) is nothing else but the classical porous media equation (PME). The first numerical experiments discussed here, will be for the mentioned β . Indeed, in the case when the initial condition u_0 is a delta Dirac function at zero, we have an exact solution provided in [4], known as the *density of Barenblatt-Pattle* and given by the following explicit formula,

$$U(t, x) = t^{-\beta} (C - \kappa x^2 t^{-2\beta})_+^{\frac{1}{m-1}}, \quad x \in \mathbb{R}, t > 0, \quad (6.1)$$

where

$$\beta = \frac{1}{m+1}, \quad \kappa = \frac{m-1}{2(m+1)m}, \quad C = \left(\frac{\sqrt{\kappa}}{\gamma_m} \right)^{\frac{2(m-1)}{m+1}}, \quad \gamma_m = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [\cos(x)]^{\frac{m+1}{m-1}}.$$

We would now compare the exact solution (6.1) to an approximated probabilistic solution. However, up to now, we are not able to perform an efficient bandwidth selection procedure in the case when the initial condition is the law of a deterministic random variable. Since we are nevertheless interested in exploiting (6.1), we considered a time translation of the exact solution U defined as follows

$$v(t, x) = U(t+1, x), \quad \forall x \in \mathbb{R}, \forall t \geq 0. \quad (6.2)$$

Note that one can immediately deduce from (6.2), that v still solves the PME but now with a smooth initial condition given by

$$v_0(x) = U(1, x), \quad \forall x \in \mathbb{R}. \quad (6.3)$$

In fact, in the case when the exponent m is equal to 3, the exact solution v of the PME with initial condition $v_0(x) = U(1, x)$ is given by the following explicit formula,

$$v(t, x) = \begin{cases} (t+1)^{-\frac{1}{4}} \sqrt{\frac{1}{\pi\sqrt{3}} - \frac{x^2}{12\sqrt{t+1}}} & \text{if } |x| \leq (t+1)^{\frac{1}{4}} \sqrt{\frac{2}{\pi}}, \\ 0 & \text{otherwise.} \end{cases} \quad (6.4)$$

Simulation experiments: we first compute both the deterministic and probabilistic numerical solutions over the time-space grid $[0, 1.5] \times [-2.5, 2.5]$, with space step

$\Delta x = 0.02$. We set $\Delta t_{det} = 4 \times 10^{-6}$, while, we use $n = 50000$ particles and a time step $\Delta t = 2 \times 10^{-4}$, for the probabilistic simulation. Figures 1.(a)-(b)-(c)-(d), display the exact and the numerical (deterministic and probabilistic) solutions at times $t = 0$, $t = 0.5$, $t = 1$ and $t = T = 1.5$ respectively. The exact solution of the PME, defined in (6.4), is depicted by solid lines.

Besides, Figure 1.(e) describes the time evolution of both the discrete L^2 deterministic and probabilistic errors on the time interval $[0, 1.5]$.

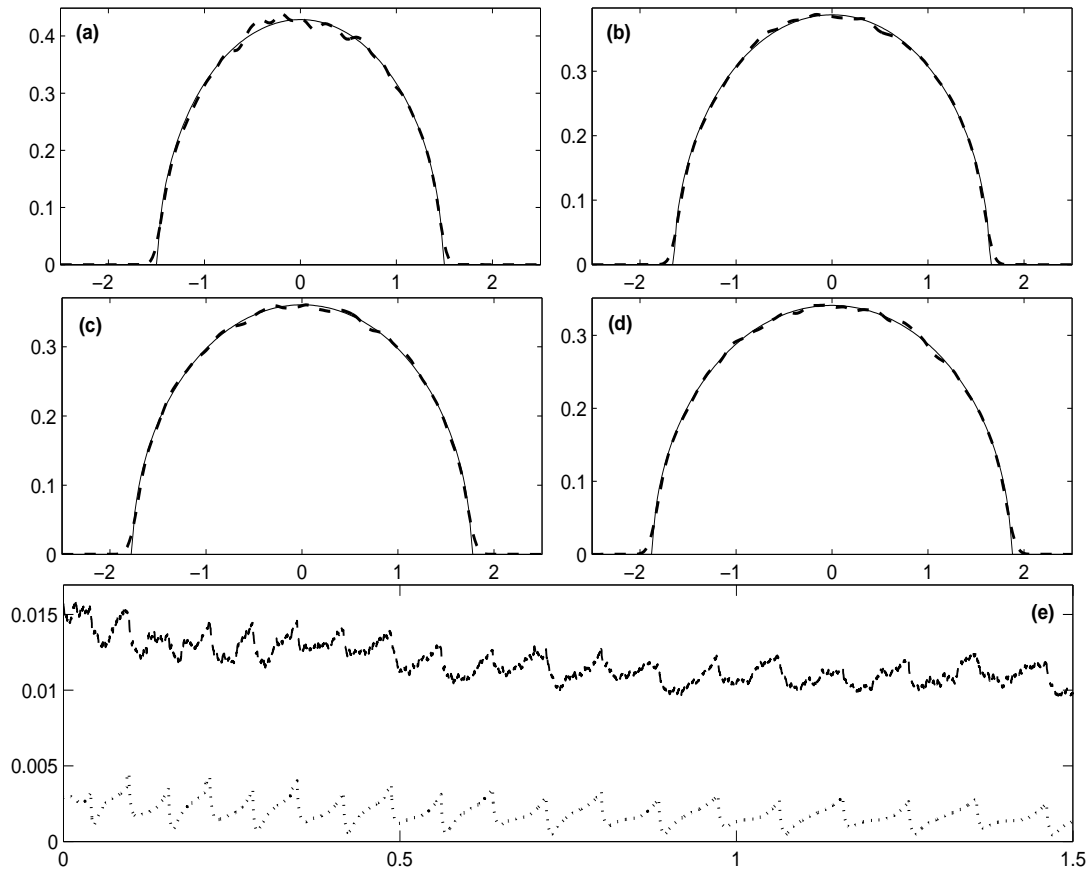


Figure 1. - Deterministic (dotted line), probabilistic (dashed line) and exact solutions (solid line) values at $t=0$ (a), $t=0.5$ (b), $t=1$ (c) and $t=1.5$ (d). The evolution of the L^2 deterministic (dotted line) and probabilistic (dashed line) errors over the time interval $[0, 1.5]$ (e).

The L^1 errors behave very similarly as well in the present case as in the Heaviside case, treated in subsection 6.2.

6.2. The Heaviside case

The second family of numerical experiments discussed here, concerns β defined by (1.3). Since we do not have an exact solution of the diffusion problem (1.1), for the mentioned β , we decided to compare the probabilistic solution to the approximation obtained via the deterministic algorithm described in Section 5. Indeed, we shall simulate both solutions according to several types of initial data u_0 and with different values of the critical threshold u_c .

Empirically, after various experiments, it appears that for a fixed threshold u_c , the numerical solution approaches some limit function which seems to belong to the "attracting" set

$$\mathcal{J} = \{f \in L^1(\mathbb{R}) \mid \int f(x)dx = 1, |f| \leq u_c\}; \quad (6.5)$$

in fact \mathcal{J} is the closure in L^1 of $\mathcal{J}_0 = \{f : \mathbb{R} \rightarrow \mathbb{R}_+ \mid \beta(f) = 0\}$. At this point, the following theoretical questions arise.

- (1) Does indeed $u(t, \cdot)$ have a limit u_∞ when $t \rightarrow \infty$?
- (2) If yes does u_∞ belong to \mathcal{J} ?
- (3) If (2) holds, do we have $u(t, \cdot) = u_\infty$ for t larger than a finite time τ ?

A similar behavior was observed for different β which are strictly increasing after some zero.

6.2.1 Trimodal initial condition

For the β given by (1.3), we consider an initial condition being a mixture of three Gaussian densities with three modes at some distance from each other, i.e.

$$u_0(x) = \frac{1}{3} (p(x, \mu_1, \sigma_1) + p(x, \mu_2, \sigma_2) + p(x, \mu_3, \sigma_3)), \quad (6.6)$$

where,

$$p(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right). \quad (6.7)$$

Simulation experiments: for this specific type of initial condition u_0 , we consider two test cases depending on the value taken by the critical threshold u_c . We set, for instance, $\mu_1 = -\mu_3 = -4$, $\mu_2 = 0$ and $\sigma_1 = 0.1$, $\sigma_2 = 0.2$, $\sigma_3 = 0.3$.

Test case 1 : we start with $u_c = 0.15$, and a time-space grid $[0, 0.6] \times [-7, 7]$, with a space step $\Delta x = 0.02$. For the deterministic approximation, we set $\Delta t_{det} = 4 \times 10^{-6}$. The probabilistic simulation uses $n = 50000$ particles and a time step $\Delta t = 2 \times 10^{-4}$. Figures 2.(a)-(b)-(c), display both the deterministic and probabilistic numerical solutions at times $t = 0$, $t = 0.3$ and $t = T = 0.6$, respectively. On the other hand, the time evolution of the L^2 -norm of the difference between the two numerical solutions

is depicted in Figure 2.(d).

Test case 2 : we choose now as critical value $u_c = 0.08$ and a time-space grid $[0, 4] \times [-8.5, 8.5]$, with a space step $\Delta x = 0.02$. We set $\Delta t_{det} = 4 \times 10^{-6}$ and the probabilistic approximation is performed using $n = 50000$ particles and a time step $\Delta t = 5 \times 10^{-4}$. Figures 3.(a)-(b)-(c) and 3.(d), show respectively the numerical (probabilistic and deterministic) solutions and the L^2 -norm of the difference between the two.

6.2.2 Uniform and Normal densities mixture initial condition

We proceed with β given by (1.3). We are now interested in an initial condition u_0 , being a mixture of a Normal and an Uniform density, i.e.,

$$u_0(x) = \frac{1}{2} (p(x, -1, 0.2) + \mathbb{1}_{[0,1]}(x)), \quad (6.8)$$

where, p is defined in (6.7).

Simulation experiments:

Test case 3 : we perform both the approximated deterministic and probabilistic solutions in the case where $u_c = 0.3$, on the time-space grid $[0, 0.5] \times [-2.5, 2]$, with a space step $\Delta x = 0.02$. We use $n = 50000$ particles and a time step $\Delta t = 2 \times 10^{-4}$, for the probabilistic simulation. Moreover, we set $\Delta t_{det} = 4 \times 10^{-6}$. Figures 4.(a)-(b)-(c) illustrate those approximated solutions at times $t = 0$, $t = 0.1$ and $t = T = 0.5$. Furthermore, we compute the L^2 -norm of the difference between the numerical deterministic solution and the probabilistic one. Values of this error, are displayed in Figure 4.(d), at each probabilistic time step.

6.2.3 Uniform densities mixture initial condition

Now, with β given by (1.3), we consider an initial condition u_0 being a mixture of Uniform densities, i.e.,

$$u_0(x) = \frac{1}{5} \mathbb{1}_{[0,1]}(x) + \frac{3}{4} \mathbb{1}_{[-\frac{1}{5}, \frac{1}{5}]}(x) + \frac{5}{8} \mathbb{1}_{[\frac{6}{5}, 2]}(x), \quad (6.9)$$

Simulation experiments:

Test case 4 : we approximate the deterministic and probabilistic solutions in the case where $u_c = 0.3$, on the time-space grid $[0, 0.6] \times [-1.5, 3.5]$, with a space step $\Delta x = 0.02$. The deterministic time step $\Delta t_{det} = 4 \times 10^{-6}$, while the probabilistic solution is computed using $n = 50000$ particles and a time step $\Delta t = 2 \times 10^{-4}$.

We illustrate in Figure 5.(a)-(b)-(c), both the deterministic and probabilistic numerical solutions at times $t = 0$, $t = 0.1$ and $t = T = 0.6$, respectively; while the time evolution of the L^2 -norm of the difference between them, is shown in Figure 5.(d).

6.2.4 Square root initial condition

Finally, the last test case concerns an initial condition u_0 defined as follows :

$$u_0(x) = \frac{3}{4} \sqrt{|x|} \mathbb{1}_{[-1,1]}(x). \quad (6.10)$$

Simulation experiments:

Test case 5 : we simulate the probabilistic and deterministic solutions over the time-space grid $[0, 0.45] \times [-2, 2]$, using a space step $\Delta x = 0.02$ and setting the critical threshold $u_c = 0.35$. Moreover, the deterministic time step $\Delta t_{det} = 4 \times 10^{-6}$. On the other hand, we use $n = 50000$ particles and a time step $\Delta t = 2 \times 10^{-4}$ to compute the probabilistic approximation.

Figures 6.(a)-(b)-(c), show both the deterministic and probabilistic numerical solutions at times $t = 0$, $t = 0.04$ and $t = T = 0.45$, respectively.

The evolution of the L^2 -norm of the difference between these two solutions, over the time interval $[0, 0.45]$, is depicted in Figure 6.(d).

6.2.5 Long-time behavior of the solutions

As it was mentioned previously, we are interested in the empirical behavior of solutions to (1.1), in the Heaviside case.

For this, we first provide Figure 8, which displays the time evolution of the L^1 -norm of the difference between two successive (in time) numerical solutions. That quantity was computed for both deterministic and probabilistic numerical solutions and in the different test cases 1 to 5. In fact, Figure 8, shows that the numerical solutions approach some limit function \hat{u}_∞ . Indeed, they seem to reach \hat{u}_∞ , after a finite time τ .

In addition, according to Figure 9, this suggests the existence of a limit function u_∞ , which of course depends on the initial condition u_0 , such that $u(t, \cdot) = u_\infty$, for $t \geq \tau$. Moreover, u_∞ , is expected to belong to the "attracting set" \mathcal{J} , since $\|\beta(u(t, \cdot))\|_{L^1(\mathbb{R})}$ equals zero when t is larger than τ , for both numerical solutions.

Besides, Figure 7, displays a single trajectory of the corresponding stochastic process, for each one of the test cases described above. In fact, we observe that, in all cases, the process trajectory stops not later than the instant of stabilization of the macroscopic distribution.

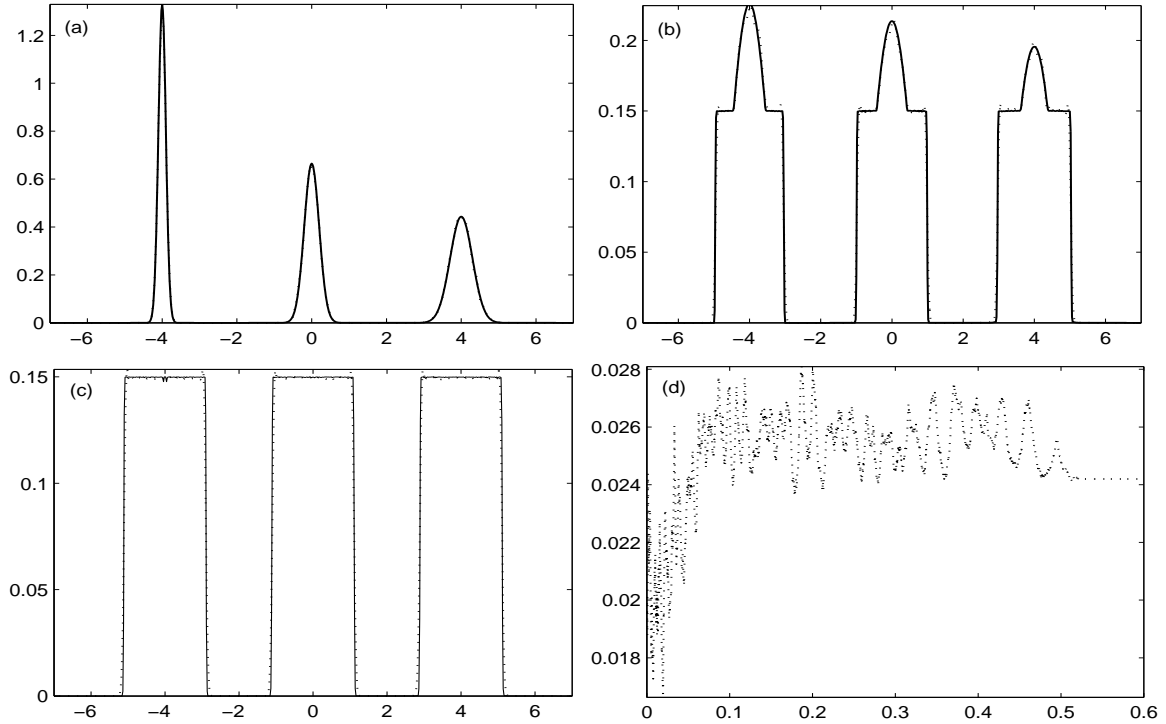


Figure 2. - Test case 1: Deterministic (solid line) and probabilistic (dotted line) solution values at $t=0$ (a), $t=0.3$ (b), $t=0.6$ (c). The evolution of the L^2 -norm of the difference over the time interval $[0, 0.6]$ (d).

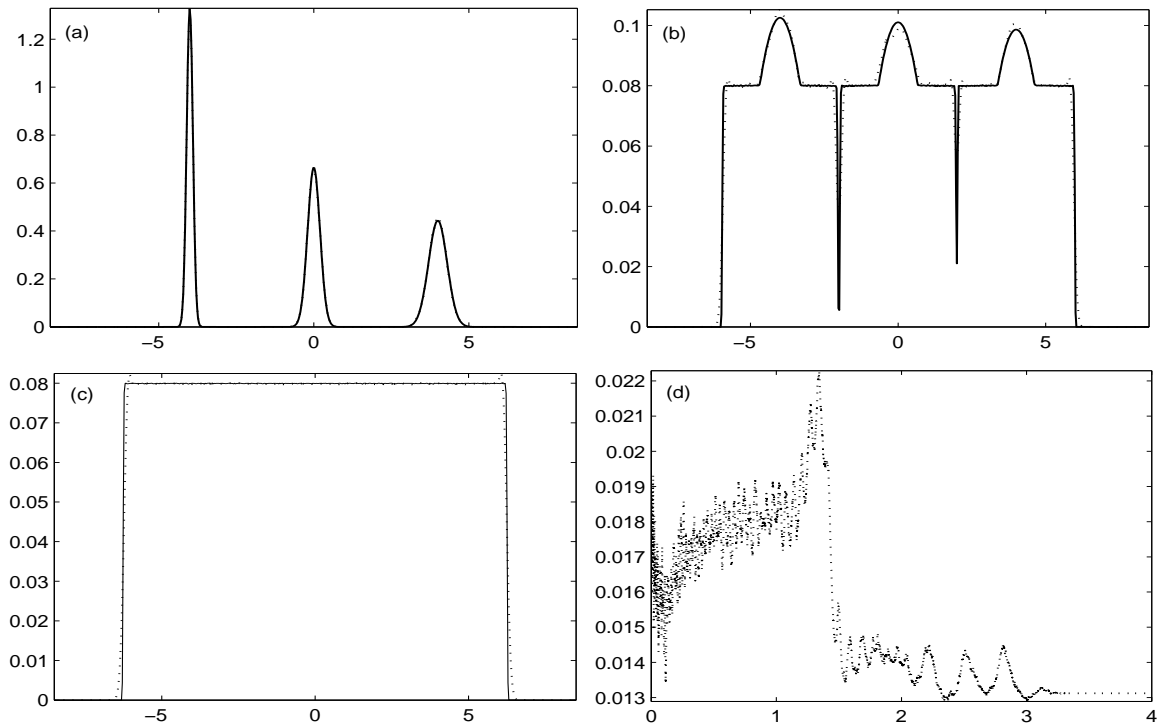


Figure 3. - Test case 2: Deterministic (solid line) and probabilistic (dotted line) solution values at $t=0$ (a), $t=2$ (b), $t=4$ (c). The evolution of the L^2 -norm of the difference over the time interval $[0, 4]$ (d).

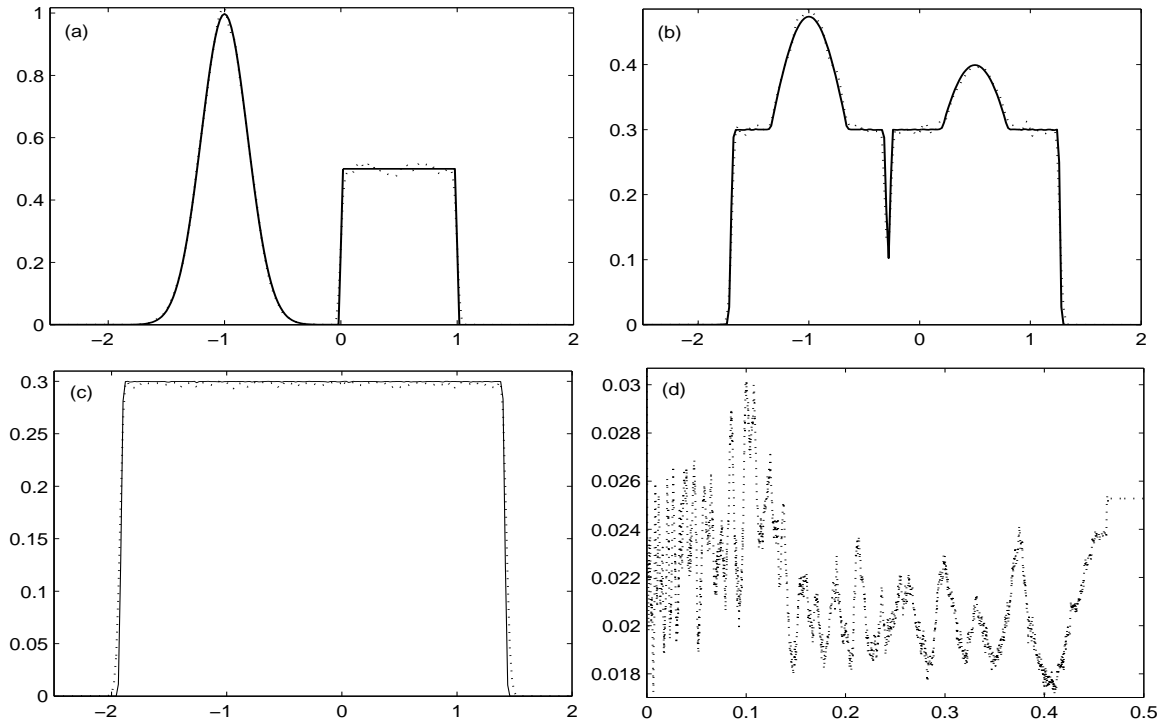


Figure 4. - Test case 3: Deterministic (solid line) and probabilistic (dotted line) solutions values at $t=0$ (a), $t=0.1$ (b), $t=0.5$ (c). The evolution of the L^2 -norm of the difference over the time interval $[0, 0.5]$ (d).

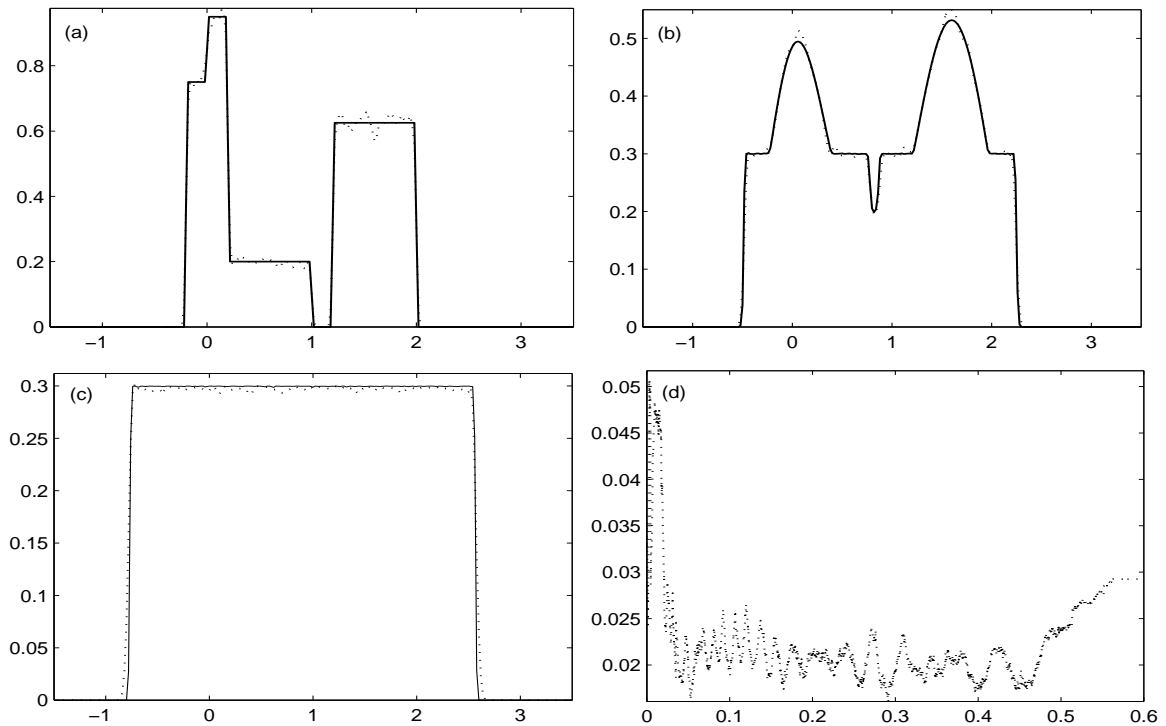


Figure 5. - Test case 4: Deterministic (solid line) and probabilistic solution (dotted line) values at $t=0$ (a), $t=0.1$ (b), $t=0.6$ (c). The evolution of the L^2 -norm of the difference over the time interval $[0, 0.6]$ (d).

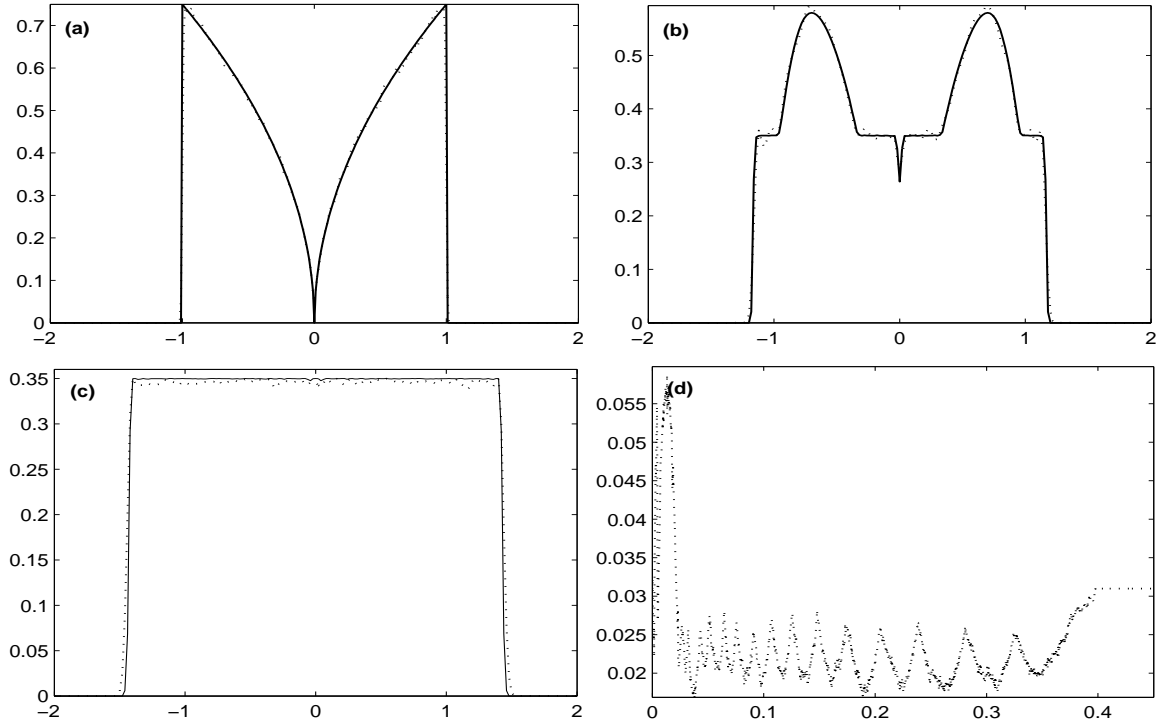


Figure 6. - Test case 5: Deterministic (solid line) and probabilistic (dotted line) solution values at $t=0$ (a), $t=0.04$ (b), $t=0.45$ (c). The evolution of the L^2 -norm of the difference over the time interval $[0, 0.45]$ (d).

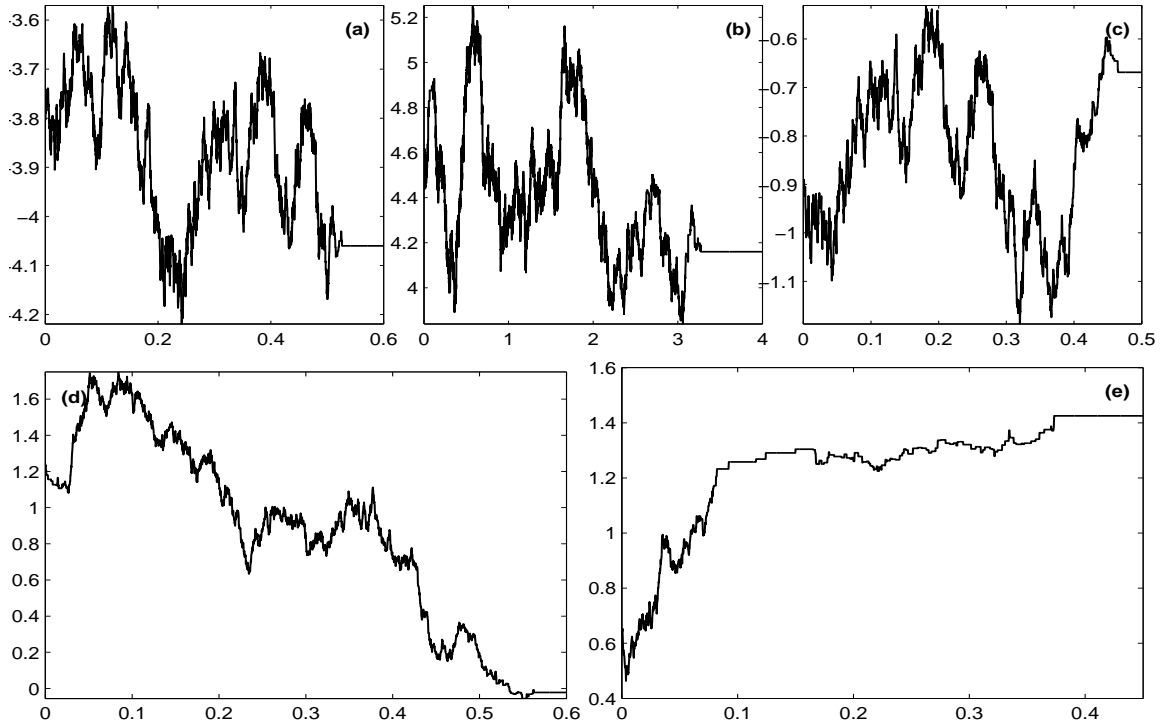


Figure 7. - Representation of a process trajectory for the Test case 1 (a), Test case 2 (b), Test case 3 (c), Test case 4 (d) and Test case 5 (e), respectively.

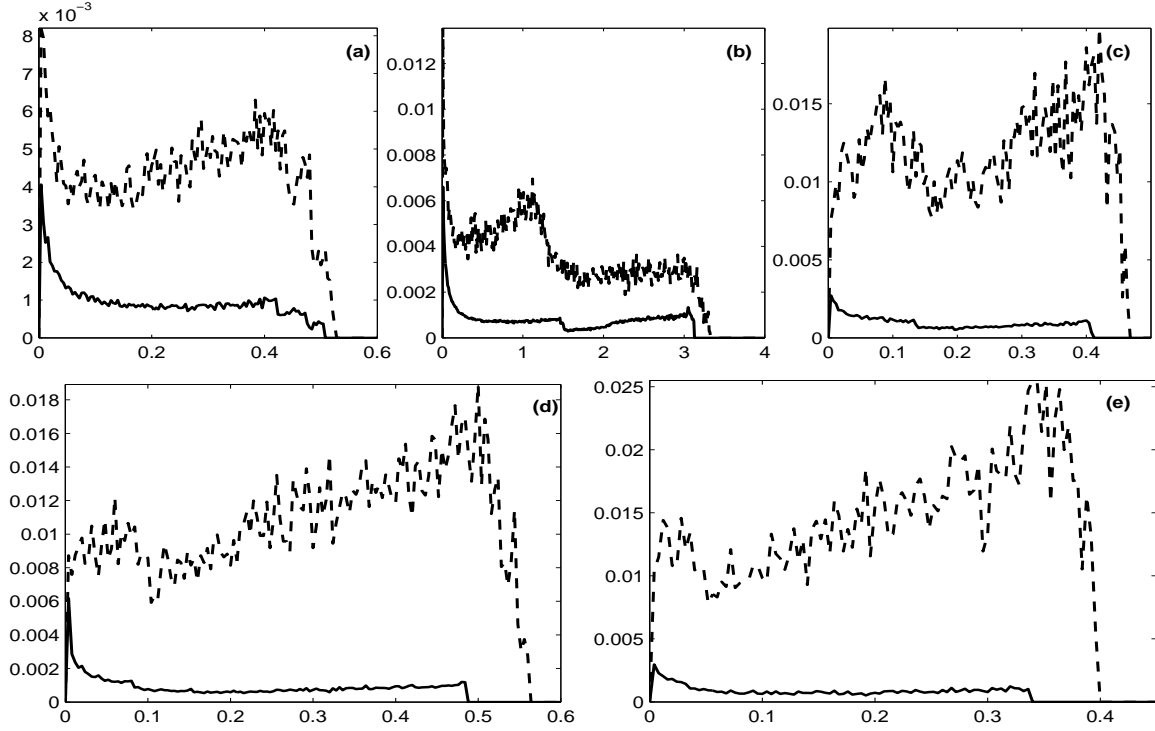


Figure 8. - Time evolution of $\|u^{\varepsilon,n}(t_{i+1}, \cdot) - u^{\varepsilon,n}(t_i, \cdot)\|_{L^1(\mathbb{R})}$ (dashed lines) and $\|\hat{u}_{det}(t_{i+1}, \cdot) - \hat{u}_{det}(t_i, \cdot)\|_{L^1(\mathbb{R})}$ (solid lines) for the Test case 1 (a), Test case 2 (b), Test case 3 (c), Test case 4 (d) and Test case 5 (e), respectively .

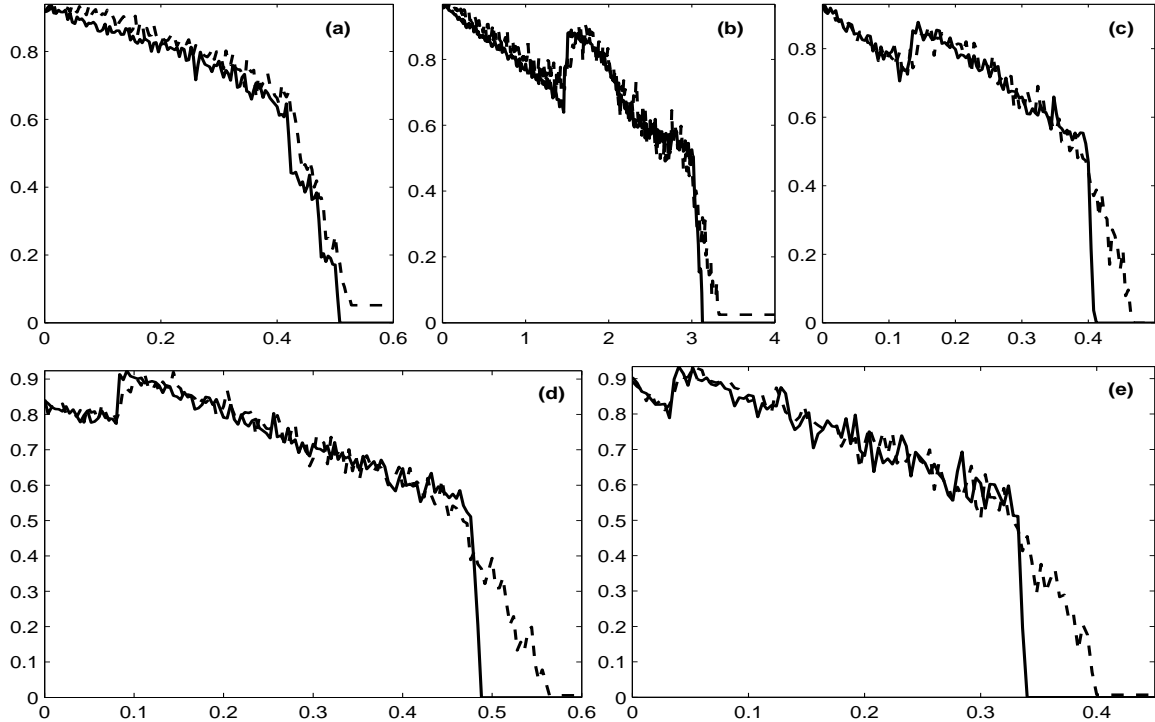


Figure 9. - Time evolution of $\|\beta(u^{\varepsilon,n}(t, \cdot))\|_{L^1(\mathbb{R})}$ (dashed lines) and $\|\beta(\hat{u}_{det}(t, \cdot))\|_{L^1(\mathbb{R})}$ (solid lines) for the Test case 1 (a), Test case 2 (b), Test case 3 (c), Test case 4 (d) and Test case 5 (e), respectively.

6.3. Performance of the probabilistic algorithm

If u is the solution to (1.1), we are now interested in the dependence of the error $\|u(T, \cdot) - u^{\varepsilon, n}(T, \cdot)\|_{L^2(\mathbb{R})}$, on the two parameters of the algorithm: the number of particles n and the time step Δt . For this, we perform numerical experiments for two test cases: the classical porous media (PME) and the Heaviside case with trimodal initial condition (see Section 6.2.1). For the PME case, we have $u(T, \cdot) = v(T, \cdot)$, where v is given by (6.4). Since we do not know an exact solution in the Heaviside case, we set $u(T, \cdot) = \hat{u}_{det}(T, \cdot)$, where \hat{u}_{det} denotes the numerical deterministic solution.

6.3.1 Dependence on n

First, we provide in Table 1 the numerical initial error, $E_0 = \|u(0, \cdot) - u^{\varepsilon, n}(0, \cdot)\|_{L^2(\mathbb{R})}$, in both cases and for the different values of n .

Table 2 shows the estimates of the L^2 norm of the error. In order to study the dependence on n , we will proceed as follows: we fix $\Delta t = 2 \cdot 10^{-4}$, expecting that this value will be small enough to only observe the effect of n on the error and we provide the results corresponding to $n = 3125, 6250, 12500, 25000, 50000$. Figure 10 represents the evolution of the error in terms of n , in a logarithmic scale.

In the Heaviside case and in the PME case, we remark that the error is not of order $\mathcal{O}(n^{-\frac{1}{2}})$. Actually, it seems to be of order $\mathcal{O}(n^{-\frac{1}{3}})$, at least for regular initial conditions; we emphasize that this happens for two radically different β .

Table 1. Dependence on n of the initial error E_0

Number of particles n	3125	6250	12500	25000	50000
Heaviside case	0,062301	0,049737	0,028496	0,021775	0,016778
PME case	0,033808	0,030369	0,022378	0,019164	0,015543

Table 2. Dependence on n of the error; $T = 1$

Number of particles n	3125	6250	12500	25000	50000
Heaviside case	0,057249	0,047548	0,038643	0,033445	0,026366
PME case	0,026296	0,020959	0,016516	0,014389	0,011471

Table 3. Dependence on n of CPU time (sec)

Number of particles n	3125	6250	12500	25000	50000
Heaviside case	1171,3	2955,1	8381,1	26555	75344
PME case	1722,4	3452,4	9917,9	23506	80201

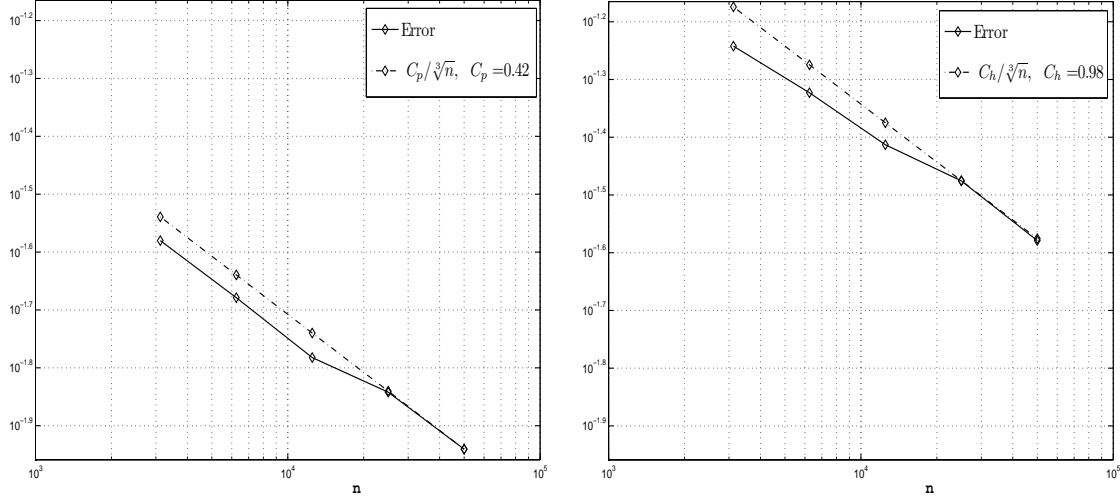


Figure 10. Evolution of the error in terms of n , in the PME (left) and the Heaviside case (right), in logarithmic scale.

6.3.2 Dependence on Δt

We apply the same procedure as previously, but this time we fix $n = 50000$ and we estimate the error for $\Delta t = 2^{-i}$, $i = 4, \dots, 14$. The obtained results are summarized, for both cases, in Table 4. Figure 11 displays, in a logarithmic scale, the error versus Δt .

When β is given by (1.3), we observe that the logarithm of the error seems to be affine from $\Delta t = 1/16$ to $\Delta t = 1/1024$ and then more or less constant, for $\Delta t < 1/1024$. In the PME case, the error seems to be more or less constant for all the values of Δt that we have considered.

Our impression is that there could be a value $(\Delta t)_0$ for which the error remains more or less constant (close to E_0), for $\Delta t < (\Delta t)_0$. We cannot exclude that the error increases for very small Δt .

Moreover, Table 4, seems to indicate that for the PME case one reaches more rapidly the error E_0 than for the Heaviside case, being the solutions more regular in space. This shows that in the PME case, where the coefficients are regular, Δt may be chosen quite big; this would allow us to approximate efficiently the solutions for a large time $T > 0$.

In the Heaviside case, the method becomes performing when Δt is quite smaller; in that case however, the solution reaches the "absorbing set" relatively fast.

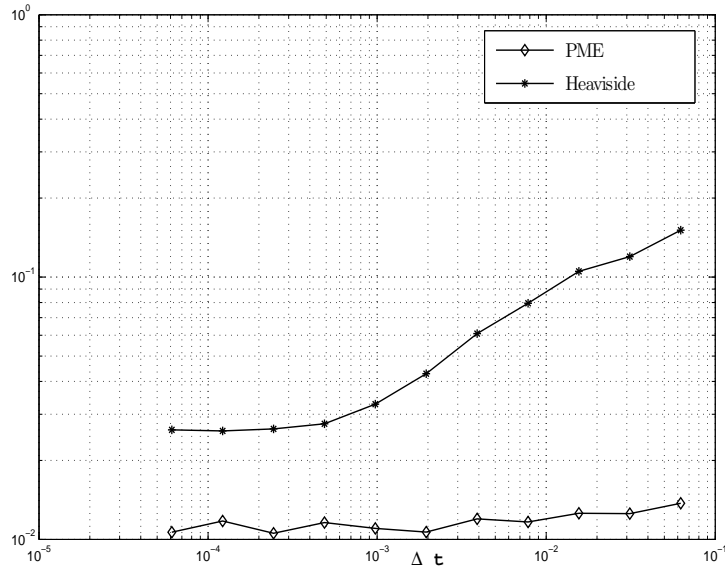


Figure 11. Evolution of the error in terms of Δt , in the PME and the Heaviside cases, in logarithmic scale.

Table 4. Dependence on Δt of the error; $T = 1$

Δt	Error (PME)	Error (Heav)	CPU time(sec) (PME)	CPU time(sec) (Heav)
1/16	0,013729	0,15085	380,75	285,57
1/32	0,012537	0,11963	783,55	609,61
1/64	0,012578	0,10513	1437	1301,9
1/128	0,011653	0,079444	2748,4	2655,1
1/256	0,011978	0,060906	5449,8	8028,8
1/512	0,010659	0,042826	10930	13285
2 ⁻¹⁰	0,011009	0,032785	21839	32394
2 ⁻¹¹	0,011592	0,027595	44070	54152
2 ⁻¹²	0,010549	0,026366	89346	75344
2 ⁻¹³	0,011753	0,0259	1,7355e+005	1,6456e+005
2 ⁻¹⁴	0,010651	0,026166	3,6805e+005	3,7103e+005

6.4. Long-time behavior of the probabilistic algorithm

Now, we inquire about the long time behavior of the probabilistic particle algorithm. In fact, we are interested in the dependence of the error over the time. For this, we have simulated the solution of the PME with $T = 50$, $\Delta t = 0.02$ and $m = 3$. Figure 12 displays the time evolution of the L^2 -norm of the error.

In particular, Figure 12 shows (in the PME case) that the probabilistic algorithm seems to remain stable for a large time T .

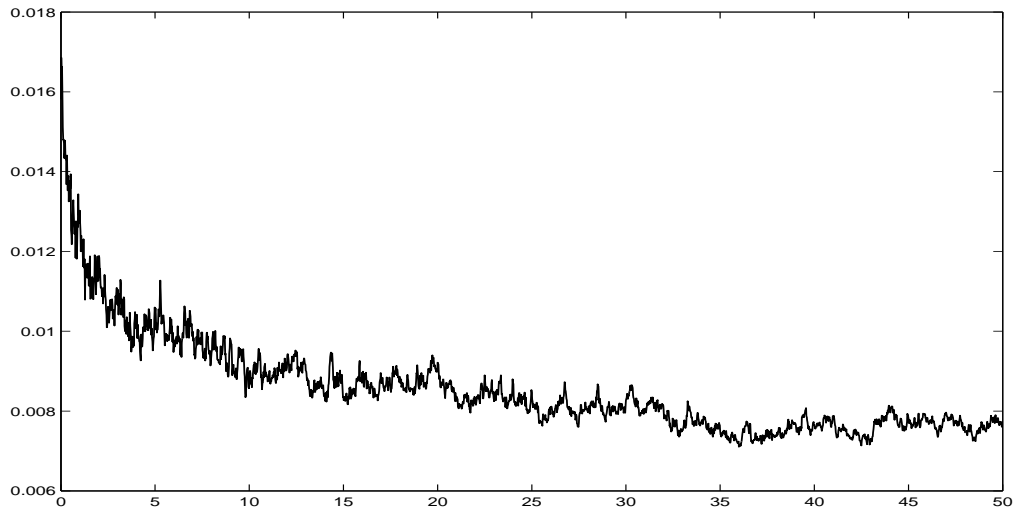


Figure 12. Evolution of the L^2 -norm of the error over the time interval $[0, 50]$, in the PME case .

Remark 6.1.

- (i) It appears through Table 4 that the CPU time and Δt are inversely proportional. On the other hand, according to Table 3, the CPU time seems to be proportional to the number of particles n .
- (ii) The probabilistic algorithm can be parallelized on a Graphical Processor Unit (GPU), such that we can speed-up its time machine execution; on the other hand, for the deterministic algorithm, this transformation is far from being obvious, see [19].

7. Annexes

Let $\mathbf{V} \in \mathbb{R}^{N_x}$ such that $V_i = v(x_i)$, $\forall i \in \llbracket 1, N_x \rrbracket$, where v is a function defined on $[a, b]$. Note that the points (x_i) are still defined as in Section 5. Moreover, $\mathcal{M}_{m,n}(\mathbb{R})$ denotes the linear space of real matrices with m rows and n columns.

7.1. Interpolating polynomial of a function

We aim to approximate $v(x_{i+1/2})$ and $v(x_{i-1/2})$ for every $i \in \llbracket 1, N_x \rrbracket$. In order to do this, we use properly chosen Lagrange interpolation polynomials of degree $k - 1$.

On every interval (or cell) $I_i = [x_{i-1/2}, x_{i+1/2}]$, with $i \in \llbracket 1, N_x \rrbracket$, we construct an interpolation polynomial $\mathbb{P}_{k,i}^{[r]}$ by selecting k consecutive points containing x_i : the so-called *stencil* denoted by

$$S(i) \equiv \{I_{i-r}, \dots, I_{i+s}\}$$

and defined by $\{x_{i-r}, x_{i-r+1}, \dots, x_{i+s-1}, x_{i+s}\}$, where r, s are positive integers and $r + s + 1 = k$. We denote by $R(i)$, the value taken by r for the interval I_i with an ENO stencil, see [45].

The Lagrange interpolation polynomial of degree $k - 1$, on the interval I_i , associated to the stencil $S(i)$ is then given by :

$$\mathbb{P}_{k,i}^{[r]}(x) = \sum_{j=0}^{k-1} V_{i-r+j} L_j^{[r]}(x), \quad \forall x \in I_i, \quad (7.1)$$

where,

$$L_j^{[r]}(x) = \prod_{\substack{l=0 \\ l \neq j}}^{k-1} \frac{x - x_{i-r+l}}{x_{i-r+j} - x_{i-r+l}}. \quad (7.2)$$

Now, we need to compute the polynomial defined in (7.2) at the points $x_{i-1/2}$ and $x_{i+1/2}$. In fact, since the points are equidistant, we have for every $(r, j) \in \llbracket 0, k - 1 \rrbracket^2$,

$$\left\{ \begin{array}{l} L_j^{[r]}(x_{i-1/2}) = \prod_{\substack{l=0 \\ l \neq j}}^{k-1} \frac{r - l - 1/2}{j - l}, \\ L_j^{[r]}(x_{i+1/2}) = \prod_{\substack{l=0 \\ l \neq j}}^{k-1} \frac{r - l + 1/2}{j - l}. \end{array} \right.$$

Then, we define $\mathbb{C} \in \mathcal{M}_{k+1,k}(\mathbb{R})$, as follows

$$\mathbb{C}_{r+1,j+1} = \prod_{\substack{l=0 \\ l \neq j}}^{k-1} \frac{r - l - 1/2}{j - l}, \quad \forall (r, j) \in \llbracket 0, k \rrbracket \times \llbracket 0, k - 1 \rrbracket. \quad (7.3)$$

Substituting (7.3) in (7.1) and using the ENO stencil, we get $\forall i \in \llbracket 1, N_x \rrbracket$

$$v(x_{i-1/2}) \approx v_{i-1/2}^+ = \mathbb{P}_{k,i}^{[R(i)]}(x_{i-1/2}) = \sum_{j=0}^{k-1} V_{i-R(i)+j} \mathbb{C}_{R(i)+1,j+1}, \quad (7.4)$$

$$v(x_{i+1/2}) \approx v_{i+1/2}^- = \mathbb{P}_{k,i}^{[R(i)]}(x_{i+1/2}) = \sum_{j=0}^{k-1} V_{i-R(i)+j} \mathbb{C}_{R(i)+2,j+1}. \quad (7.5)$$

7.2. Interpolation polynomial for the derivative of a function

Now, we would like to approximate $\frac{dv}{dx}(x_i)$, $\frac{dv}{dx}(x_{i-1/2})$ and $\frac{dv}{dx}(x_{i+1/2})$, for every $i \in \llbracket 1, N_x \rrbracket$. In fact, deriving equation (7.1), implies

$$\frac{d\mathbb{P}_{k,i}^{[r]}}{dx}(x) = \sum_{j=0}^{k-1} V_{i-r+j} \frac{dL_j^{[r]}}{dx}(x), \quad \forall x \in I_i. \quad (7.6)$$

On the other hand, for every $j \in \llbracket 0, l-1 \rrbracket$, we have

$$\frac{dL_j^{[r]}}{dx}(x) = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (x - x_{i-r+l})}{\prod_{\substack{l=0 \\ l \neq j}}^{k-1} (x_{i-r+j} - x_{i-r+l})}, \quad \forall x \in I_i. \quad (7.7)$$

Since the points are equidistant, we get

$$\frac{dL_j^{[r]}}{dx}(x_i) = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (r-l)}{\Delta x \prod_{\substack{l=0 \\ l \neq j}}^{k-1} (j-l)}$$

and

$$\frac{dL_j^{[r]}}{dx}(x_{i-1/2}) = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (r-l-1/2)}{\Delta x \prod_{\substack{l=0 \\ l \neq j}}^{k-1} (j-l)}, \quad \frac{dL_j^{[r]}}{dx}(x_{i+1/2}) = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (r-l+1/2)}{\Delta x \prod_{\substack{l=0 \\ l \neq j}}^{k-1} (j-l)}.$$

Then, we define $\mathbb{D} \in \mathcal{M}_{k,k}(\mathbb{R})$ by

$$\mathbb{D}_{r+1,j+1} = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (r-l)}{\Delta x \prod_{\substack{l=0 \\ l \neq j}}^{k-1} (j-l)}, \quad \forall (r, j) \in [[0, k-1]]^2, \quad (7.8)$$

and $\bar{\mathbb{D}} \in \mathcal{M}_{k+1,k}(\mathbb{R})$ as follows

$$\bar{\mathbb{D}}_{r+1,j+1} = \frac{\sum_{\substack{m=0 \\ m \neq j}}^{k-1} \prod_{\substack{l=0 \\ l \neq j, m}}^{k-1} (r-l-1/2)}{\Delta x \prod_{\substack{l=0 \\ l \neq j}}^{k-1} (j-l)}, \quad \forall (r, j) \in [[0, k]] \times [[0, k-1]]. \quad (7.9)$$

Therefore, replacing (7.8) and (7.9) in (7.6), for every $i \in [[1, N_x]]$, we obtain :

$$\frac{dv}{dx}(x_i) \approx dv_i = \frac{d\mathbb{P}_{k,i}^{[r]}(x_i)}{dx} = \sum_{j=0}^{k-1} V_{i-r+j} \mathbb{D}_{r+1,j+1}, \quad (7.10)$$

$$\frac{dv}{dx}(x_{i-1/2}) \approx dv_{i-1/2}^+ = \frac{d\mathbb{P}_{k,i}^{[r]}(x_{i-1/2})}{dx} = \sum_{j=0}^{k-1} V_{i-r+j} \bar{\mathbb{D}}_{r+1,j+1}, \quad (7.11)$$

$$\frac{dv}{dx}(x_{i+1/2}) \approx dv_{i+1/2}^- = \frac{d\mathbb{P}_{k,i}^{[r]}(x_{i+1/2})}{dx} = \sum_{j=0}^{k-1} V_{i-r+j} \bar{\mathbb{D}}_{r+2,j+1}. \quad (7.12)$$

Acknowledgments. Part of the work was done during the stay of the first and third named authors at the Bielefeld University (SFB 701 and BiBoS). They are grateful for the invitation.

References

1. D. Aregba-Driollet, R. Natalini and S. Tang, *Explicit diffusive kinetic schemes for nonlinear degenerate parabolic systems*, Math. Comp. 73 (2004), pp. 63–94 (electronic).
2. P. Bak, *How Nature Works: The science of Self-Organized Criticality*. Springer-Verlag New York, Inc, 1986.

3. V. Barbu, M. Röckner and F. Russo, *Probabilistic representation for solutions of an irregular porous media type equation: the irregular degenerate case*, Probab. Theory Related Fields 151 (2011), pp. 1–43.
4. G. I. Barenblatt, *On some unsteady motions of a liquid and gas in a porous medium*, Akad. Nauk SSSR. Prikl. Mat. Meh. 16 (1952), pp. 67–78.
5. S. Benachour, P. Chassaing, B. Roynette and P. Vallois, *Processus associés à l'équation des milieux poreux*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 23 (1996), pp. 793–832 (1997).
6. P. Benilan, H. Brezis and M. G. Crandall, *A semilinear equation in $L^1(\mathbb{R}^N)$* , Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 2 (1975), pp. 523–555.
7. P. Benilan and M. G. Crandall, *The continuous dependence on φ of solutions of $u_t - \Delta\varphi(u) = 0$* , Indiana Univ. Math. J. 30 (1981), pp. 161–177.
8. A. E. Berger, H. Brézis and J. C. W. Rogers, *A numerical method for solving the problem $u_t - \Delta f(u) = 0$* , RAIRO Anal. Numér. 13 (1979), pp. 297–312.
9. P. Blanchard, M. Röckner and F. Russo, *Probabilistic representation for solutions of an irregular porous media type equation*, Ann. Probab. 38 (2010), pp. 1870–1900.
10. M. Bossy and D. Talay, *A stochastic particle method for some one-dimensional nonlinear p.d.e.*, Math. Comput. Simulation 38 (1995), pp. 43–50. Probabilités numériques (Paris, 1992).
11. ———, *A stochastic particle method for the McKean-Vlasov and the Burgers equation*, Math. Comp. 66 (1997), pp. 157–192.
12. Z. I. Botev, J. F. Grotowski and D. P. Kroese, *Kernel density estimation via diffusion*, Ann. Statist. 38 (2010), pp. 2916–2957.
13. F. Bouchut, F. R. Guarguaglini and R. Natalini, *Diffusive BGK approximations for nonlinear multidimensional parabolic equations*, Indiana Univ. Math. J. 49 (2000), pp. 723–749.
14. A. W. Bowman, *An alternative method of cross-validation for the smoothing of density estimates*, Biometrika 71 (1984), pp. 353–360.
15. H. Brezis and M. G. Crandall, *Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$* , J. Math. Pures Appl. (9) 58 (1979), pp. 153–163.
16. R. Cafiero, V. Loreto, L. Pietronero, A. Vespignani and S. Zapperi, *Local rigidity and self-organized criticality for avalanches*, Europhysics Letters 29 (1995), pp. 111–116.
17. P. Calderoni and M. Pulvirenti, *Propagation of chaos for Burgers' equation*, Ann. Inst. H. Poincaré Sect. A (N.S.) 39 (1983), pp. 85–97.
18. F. Cavalli, G. Naldi, G. Puppo and M. Semplice, *High-order relaxation schemes for nonlinear degenerate diffusion problems*, SIAM J. Numer. Anal. 45 (2007), pp. 2098–2119 (electronic).
19. F. Cuvelier, *Implementing Kernel Density Estimation on GPU: application to a probabilistic algorithm for PDEs of porous media type*, Technical report. In preparation.
20. J. M. Dawson, *Particle simulation of plasmas*, Rev. Modern Phys. 55 (1983), pp. 403–447.
21. A. Figalli and R. Philipowski, *Convergence to the viscous porous medium equation and propagation of chaos*, ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008), pp. 185–203.
22. E. Hairer, S. P. Nørsett and G. Wanner, *Solving ordinary differential equations. I*, second ed, Springer Series in Computational Mathematics 8. Springer-Verlag, Berlin, 1993, Nonstiff problems.

23. A. Harten and S. Osher, *Uniformly high-order accurate nonoscillatory schemes. I*, SIAM J. Numer. Anal. 24 (1987), pp. 279–309.
24. R. W. Hockney and J. W. Eastwood, *Computer simulation using particles*. McGraw-Hill, New York, 1981.
25. S. Jin and C. D. Levermore, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, J. Comput. Phys. 126 (1996), pp. 449–467.
26. S. Jin and Z. P. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Comm. Pure Appl. Math. 48 (1995), pp. 235–276.
27. M. C. Jones, J. S. Marron and S. J. Sheather, *A brief survey of bandwidth selection for density estimation*, J. Amer. Statist. Assoc. 91 (1996), pp. 401–407.
28. B. Jourdain, *Probabilistic approximation for a porous medium equation*, Stochastic Process. Appl. 89 (2000), pp. 81–99.
29. B. Jourdain and S. Méléard, *Propagation of chaos and fluctuations for a moderate model with smooth initial data*, Ann. Inst. H. Poincaré Probab. Statist. 34 (1998), pp. 727–766.
30. J. Kačur, A. Handlovičová and M. Kačurová, *Solution of nonlinear diffusion problems by linear approximation schemes*, SIAM J. Numer. Anal. 30 (1993), pp. 1703–1722.
31. I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed, Graduate Texts in Mathematics 113. Springer-Verlag, New York, 1991.
32. H. P. Jr. McKean, *Propagation of chaos for a class of non-linear parabolic equations.*, Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), Air Force Office Sci. Res., Arlington, Va., 1967, pp. 41–57.
33. S. Méléard and S. Roelly-Coppoletta, *A propagation of chaos result for a system of particles with moderate interaction*, Stochastic Process. Appl. 26 (1987), pp. 317–332.
34. K. Oelschläger, *A law of large numbers for moderately interacting diffusion processes*, Z. Wahrsch. Verw. Gebiete 69 (1985), pp. 279–322.
35. ———, *A fluctuation theorem for moderately interacting diffusion processes*, Probab. Theory Related Fields 74 (1987), pp. 591–616.
36. ———, *Simulation of the solution of a viscous porous medium equation by a particle method*, SIAM J. Numer. Anal. 40 (2002), pp. 1716–1762 (electronic).
37. L. Pareschi and G. Russo, *Implicit-Explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation*, J. Sci. Comput. 25 (2005), pp. 129–155.
38. E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist. 33 (1962), pp. 1065–1076.
39. R. Philipowski, *Interacting diffusions approximating the porous medium equation and propagation of chaos*, Stochastic Process. Appl. 117 (2007), pp. 526–538.
40. I. S. Pop and W. Yong, *A numerical approach to degenerate parabolic equations*, Numer. Math. 92 (2002), pp. 357–381.
41. M. Rudemo, *Empirical choice of histograms and kernel density estimators*, Scand. J. Statist. 9 (1982), pp. 65–78.
42. D. W. Scott and G. R. Terrell, *Biased and unbiased cross-validation in density estimation*, J. Amer. Statist. Assoc. 82 (1987), pp. 1131–1146.
43. S. J. Sheather and M. C. Jones, *A reliable data-based bandwidth selection method for kernel density estimation*, J. Roy. Statist. Soc. Ser. B 53 (1991), pp. 683–690.

44. R. E. Showalter, *Monotone operators in Banach space and nonlinear partial differential equations*, Mathematical Surveys and Monographs 49. American Mathematical Society, Providence, RI, 1997.
45. C. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997), Lecture Notes in Math. 1697, Springer, Berlin, 1998, pp. 325–432.
46. B. W. Silverman, *Density estimation for statistics and data analysis*, Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1986.
47. D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics. Springer-Verlag, Berlin, 2006, Reprint of the 1997 edition.
48. A. S. Sznitman, *Topics in propagation of chaos*, École d'Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math. 1464, Springer, Berlin, 1991, pp. 165–251.
49. G. R. Terrell, *The maximal smoothing principle in density estimation*, J. Amer. Statist. Assoc. 85 (1990), pp. 470–477.
50. G. R. Terrell and D. W. Scott, *Oversmoothed nonparametric density estimates*, J. Amer. Statist. Assoc. 80 (1985), pp. 209–214.
51. M. P. Wand and M. C. Jones, *Kernel smoothing*, Monographs on Statistics and Applied Probability 60. Chapman and Hall Ltd., London, 1995.
52. M. Woodroffe, *On choosing a delta-sequence*, Ann. Math. Statist. 41 (1970), pp. 1665–1671.

Author information

Nadia Belaribi, Laboratoire d'Analyse, Géométrie et Applications (LAGA), Université Paris 13, 99, avenue Jean-Baptiste Clément, F-93430 Villetaneuse and ENSTA ParisTech, Unité de Mathématiques appliquées, 32, Boulevard Victor, F-75739 Paris Cedex 15, France.
Email: belaribi@math.univ-paris13.fr

François Cuvelier, Laboratoire d'Analyse, Géométrie et Applications (LAGA), Université Paris 13, 99, avenue Jean-Baptiste Clément F-93430 Villetaneuse, France.
Email: cuvelier@math.univ-paris13.fr

Francesco Russo, ENSTA ParisTech, Unité de Mathématiques appliquées, 32, Boulevard Victor, F-75739 Paris Cedex 15, INRIA Rocquencourt and Cermics Ecole des Ponts et Chaussées, Projet MATHFI Domaine de Voluceau, BP 105 F-78153 Le Chesnay Cedex, France.
Email: francesco.russo@ensta-paristech.fr