Probabilistic Numerical Methods for Partial Differential Equations: Elements of Analysis*

Denis Talay

The objective of these notes is to present recent results on the convergence rate of Monte–Carlo methods for linear Partial Differential Equations and integrodifferential Equations, and for stochastic particles methods for some nonlinear evolution problems (McKean–Vlasov equations, Burgers equation, convection-reaction-diffusion equations). The given bounds for the numerical errors are non asymptotic: one wants to estimate the global errors of the methods for different possible values taken by their parameters (discretization step, number of particles or of simulations, etc).

Only a selection of existing results is presented. Most of the proofs are only sketched but the methodologies are described carefully. Deeper information should be available in Talay and Tubaro [49]. A companion review paper of these notes, with an emphasis on applications in Random Mechanics, is Talay [48].

PART I - Monte Carlo Methods for Parabolic PDE's

1 Notation

We fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, and a r-dimensional Brownian motion (W_t) on this space.

*THESE NOTES HAVE BEEN PUBLISHED IN 'PROBABILISTIC MODELS FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS', D. TALAY AND L. TUBARO (Eds.), LECTURE NOTES IN MATHEMATICS 1627, 1996.

Usually a time interval $[0,T]$ will be fixed.

The notation $(X_t(x))$ stands for a process (X_t) such that $X_0 = x$ $a.s.$

Given a smooth function φ and a multiindex α of the form

$$\alpha = (\alpha_1, \ldots, \alpha_k)\ ,\ \ \alpha_i \in \{1, \ldots, d\}$$

the notation $\partial^x_\alpha \varphi(t,x,y)$ means that the multiindex α concerns the differentiation with respect to the coordinates of x, the variables t and y being fixed.

When $\gamma = (\gamma^i_j)$ is a matrix, $\hat{\gamma}$ denotes the determinant of γ, and γ_j denotes the $j-th$ column of γ.

When V is a vector, ∂V denotes the matrix $(\partial_i V_j)^i_j$.

Finally, the same notation $K(\cdot)$, q, Q, μ, etc is used for different functions and positive real numbers, having the common property of being independent of T and of the approximation parameters (discretization step, number of simulations or number of particles, etc).

2 The Euler and Milshtein schemes for SDE's

Let (X_t) be the process taking values in $\mathbb{R}^d$ solution to

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s\ , \tag{1}$$

where (W_t) is a r-dimensional Brownian motion.

Our objective is to approximate the unknown process (X_t) by an approximate process whose trajectories can easily be simulated on a computer. Typically we must simulate a large number of independent trajectories of this process. Therefore, the cost of the simulation of one trajectory must be so low as possible.

An efficient procedure consists in choosing a discretisation step $\frac{T}{n}$ of the time interval $[0,T]$ and in simulating the Euler scheme defined by

$$\left\{ \begin{array}{rcl} X^n_0 & = & X_0\ , \\ X^n_{(p+1)T/n} & = & X^n_{pT/n} + b(X^n_{pT/n})\frac{T}{n} \\ & & \quad +\sigma(X^n_{pT/n})(W_{(p+1)T/n} - W_{pT/n}). \end{array} \right. \tag{2}$$

To simulate one trajectory of $(X^n_t, 0 \le t \le T)$, one simply has to simulate the family

$$(W_{T/n}, W_{2T/n} - W_{T/n}, \ldots, W_T - W_{T-T/n})$$

of independent Gaussian random variables. For $\frac{kT}{n} \leq t < \frac{(k+1)T}{n}$, X_t^n is defined by

$$X_t^n = X_{kT/n}^n + b(X_{kT/n}^n)\left(t - \frac{kT}{n}\right) + \sigma(X_{kT/n}^n)(W_t - W_{kT/n}). \tag{3}$$

The convergence rate of this scheme has been studied according to various convergence criterions. In the sequel we will present estimates on the discretization error according to several different criterions, all of them being related to probabilistic numerical procedures for Partial Differential Equations. The proofs of most of these estimates use an elementary result concerning the convergence in $L^p(\Omega)$.

Proposition 2.1 *Suppose that the functions $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz.*

Let $p \geq 1$ be an integer such that $\mathbb{E}|X_0|^{2p} < \infty$.

Then there exists an increasing function $K(\cdot)$ such that, for any integer $n \geq 1$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - X_t^n|^{2p}\right] \leq \frac{K(T)}{n^p}. \tag{4}$$

The function $K(\cdot)$ depends on the Lipschitz constants of the functions $b(\cdot)$ and $\sigma(\cdot)$, on the dimension d, on p and on $\mathbb{E}|X_0|^{2p}$.

Sketch of the proof. Let L_b be the Lipschitz constant of $b(\cdot)$:

$$|b(x)| \leq L_b|x| + |b(0)| \ ;$$

a similar inequality holds for $\sigma(\cdot)$. Thus, from (3) and Itô's formula, an induction on k shows: there exists an increasing function $K(\cdot)$ such that for any $n \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t^n|^{2p}\right] \leq K(T)(1 + \mathbb{E}|X_0|^{2p}) \exp(K(T)). \tag{5}$$

Here, the function K depends on L_b, L_σ, p and the dimension d.

Consider, for $t \in [kT/n, (k+1)T/n]$, the process

$$\begin{aligned} \varepsilon_t \ &:= \ X_{kT/n} - X_{kT/n}^n + \int_{kT/n}^t (b(X_s) - b(X_{kT/n}^n))ds \\ &\quad + \int_{kT/n}^t (\sigma(X_s) - \sigma(X_{kT/n}^n))dW_s. \end{aligned} \tag{6}$$

Apply the Itô formula to $|\varepsilon_t|^{2p}$ between $t = \frac{kT}{n}$ and $t = \frac{(k+1)T}{n}$; standard computations and (5) then show that, for a new increasing function $K(\cdot)$,

$$\mathbb{E}|\varepsilon_{(k+1)T/n}|^{2p} \leq \left(1 + \frac{K(T)}{n}\right) \mathbb{E}|\varepsilon_{kT/n}|^{2p} + \frac{K(T)}{n^{p+1}}.$$

Noting that $\varepsilon_0 = 0$, an induction on k provides the estimate

$$\sup_{0 \leq k \leq n} \mathbb{E}|\varepsilon_{kT/n}|^{2p} \leq \frac{C_1 \exp(C_2 T)}{n^p}.$$

To conclude, it remains to use (6) again. ■

Applying the Borel-Cantelli lemma, one readily deduces the

Proposition 2.2 *Suppose that the functions $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz. Suppose that $\mathbb{E}|X_0|^{2p} < \infty$ for any integer p. Then*

$$\forall 0 \leq \alpha < \frac{1}{2}\ ,\ n^\alpha \sup_{t \in [0,T]} |X_t - X_t^n| \overset{n \to \infty}{\longrightarrow} 0 \ a.s. \tag{7}$$

The details of the easy proofs of the two preceding propositions can be found in Faure's thesis [13] or in Kanagawa [23] e.g.

Concerning the path by path convergence of the Euler scheme, one has an even better information, which we briefly now present; we refer to Roynette [43] for a complete exposition.

Let $g(\cdot)$ be a function from $[0,T]$ to $\mathbb{R}$ and let p be a strictly positive integer. Set

$$\omega_p(g,t) := \sup_{|h| \leq t} \left(\int_{I_h} |g(x+h) - g(x)|^p dx \right)^{1/p}$$

with $I_h := \{x \in [0,T]; x+h \in [0,T]\}$. For $0 < \alpha < 1$ and $1 \leq q \leq +\infty$, set

$$\|g\|_{\alpha,p,q} :=\| g \|_{L^p(\mathbb{R})} + \left(\int_0^T \left(\frac{\omega_p(g,t)}{t^\alpha} \right)^q \frac{dt}{t} \right)^{1/q}.$$

The Besov space $\mathcal{B}^\alpha_{p,q}$ is the Banach space of the functions $g(\cdot)$ such that $\|g\|_{\alpha,p,q} < \infty$, endowed with the norm $\| \cdot \|_{\alpha,p,q}$. The Besov space $\mathcal{B}^\alpha_{\infty,\infty}$ is the usual space of Hölder functions of order α.

Theorem 2.3 (Roynette [43]) *Suppose that the functions $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz. Suppose that $I\!\!E|X_0|^{2p} < \infty$ for any integer p. Let $\tilde{X}^n := X^{2^n}$. For any integer $p > 1$, for any n large enough, there exists a constant $C_T(p)$ uniform with respect to n such that, for any $\gamma < \frac{1}{2}$,*

$$\| X_\cdot - \tilde{X}^n_\cdot \|_{1/2,p,\infty} \leq C_T(p) 2^{-n\gamma} \ a.s. \tag{8}$$

Sketch of the proof. Consider the process $\varepsilon(\cdot)$ defined in (6). Define

$$b^n_s := b(X_s) - b(\tilde{X}^n_s) \quad , \quad \sigma^n_s := \sigma(X_s) - \sigma(\tilde{X}^n_s).$$

Thus,

$$\varepsilon_t := \int_0^t b^n_s ds + \int_0^t \sigma^n_s dW_s.$$

From the estimates (4) and (7), one can easily show that, for any $\gamma < \frac{1}{2}$, for any n large enough,

$$\sup_{0 \leq s \leq T} (|b^n_s| + |\sigma^n_s|) \leq C 2^{-n\gamma} \ a.s.,$$

and for any integer $p \geq 1$,

$$\sup_{0 \leq s \leq T} I\!\!E|\sigma^n_s|^{2p} \leq C_T(p) 2^{-np}.$$

The technical proposition 1 in Roynette [43] then implies: for any integer $p > 1$, there exists a (new) constant $C_T(p)$ uniform w.r.t. n such that, for any $\gamma < \frac{1}{2}$,

$$\| \varepsilon_\cdot \|_{1/2,p,\infty} \leq C_T(p) 2^{-n\gamma} \ a.s.$$

■

The asymptotic distribution of the normalized Euler scheme error

$$U^n := \sqrt{n}(X_\cdot - X^n_\cdot)$$

is analysed in Kurtz and Protter [25] (see also their contribution to this volume): (X, U^n) converges in law to the process (X, U) where U is the solution to

$$U_t := \int_0^t \partial b(X_s) U_s ds + \sum_{j=1}^r \int_0^d \partial\sigma_j(X_s) U_s dW^j_s + \frac{1}{\sqrt{2}} \sum_{i,j=1}^r \partial\sigma_i(X_s)\sigma_j(X_s) dB^{ij}_s,$$

where $(B^{ij}, 1 \leq i,j \leq r)$ is a r^2-dimensional standard Brownian motion independent of X.

In the sequel, when $d = r = 1$, we also use the Milshtein scheme

$$\begin{cases} X_0^n & = X_0 \ , \\ X^n_{\frac{(p+1)T}{n}} & = X^n_{pT/n} + b(X^n_{pT/n})\frac{T}{n} \\ & \quad + \sigma(X^n_{pT/n})(W_{(p+1)T/n} - W_{pT/n}) \\ & \quad + \frac{1}{2}\sigma\left(X^n_{pT/n}\right)\sigma'\left(X^n_{pT/n}\right)\left((W_{(p+1)T/n} - W_{pT/n})^2 - \frac{T}{n}\right) . \end{cases} \tag{9}$$

Our reason for considering that scheme here comes from the

Proposition 2.4 *Suppose that the functions $b(\cdot)$ and $\sigma(\cdot)$ are twice continuously differentiable with bounded derivatives.*

Let $p \geq 1$ be an integer such that $\mathbb{E}|X_0|^{4p} < \infty$. Then there exists an increasing function $K(\cdot)$ such that

$$\sup_{t\in[0,T]} \mathbb{E}|X_t - X_t^n|^{2p} \leq \frac{K(T)}{n^{2p}}. \tag{10}$$

If $\mathbb{E}|X_0|^{4p} < \infty$ holds for any integer p, then for any $0 \leq \alpha < 1$,

$$n^\alpha \sup_{t\in[0,T]} \mathbb{E}|X_t - X_t^n|^{2p} \leq \frac{K(T)}{n^{2p}}. \tag{11}$$

The proof follows the same guidelines as the proof of Propositions 2.1 and 2.2. See Faure[13]. Note that the Milshtein scheme has better convergence rates than the Euler scheme for the convergence in $L^p(\Omega)$ and the almost sure convergence. A similar remark is true for the convergence in Besov spaces, see Roynette [43].

Here we consider the Milshtein scheme only when $d = r = 1$. In the multidimensional case, generally it requires double stochastic integrals which are not simple to simulate: see Talay [48] for a discussion and Gaines and Lyons [16] for a method of resolution.

3 Monte Carlo methods for parabolic PDE's

3.1 Principle of the method

Define the $d \times d$ matrix valued function $(a_j^i(\cdot))$ by

$$a(\cdot) := \sigma(\cdot)\sigma^*(\cdot).$$

Define the second-order differential operator $\mathcal{L}$ by

$$\mathcal{L} := \sum_{i=1}^{d} b_i(\cdot)\partial_i + \frac{1}{2}\sum_{i,j=1}^{d} a_j^i(\cdot)\partial_{ij}. \tag{12}$$

Consider the problem

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + \mathcal{L}u(t,x) & = \quad 0 \ in \ [0,T) \times \mathbb{R}^d \ , \\ u(T,x) & = \quad f(x) \ , \ x \in \mathbb{R}^d \ . \end{cases} \tag{13}$$

In the two different sets of hypotheses that we will consider for $b(\cdot)$, $\sigma(\cdot)$ and $f(\cdot)$, the following holds: the problem (13) has a unique solution which belongs to the set $\mathcal{C}^{1,2}([0,T) \times \mathbb{R}^d)$ and is continuous on $[0,T] \times \mathbb{R}^d$. This unique solution is given by

$$u(t,x) = \mathbb{E}_x f(X_{T-t}) = P_{T-t}f(x) \ , \ \forall (t,x) \in [0,T] \times \mathbb{R}^d$$

where P_θ denotes the transition operator of the Markov process (X_t).

Let $\{X_\cdot^{(i)}, i \in \mathbb{N}\}$ be a sequence of independent trajectories of the process X. If the Strong Law of Large Numbers applies for the sequence $f(X_t^{(i)}(x))$, then

$$u(t,x) = \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f(X_{T-t}^{(i)}(x)) \ , a.s.$$

In practice one must approximate the $X_t^{(i)}(x)$'s. We consider the simplest approximation method: the Euler scheme. As we will see, this simple method has very interesting properties in the present context, even from the point of view of the convergence rate and of the numerical efficiency. The "Monte Carlo+Euler" approximation of $u(t,x)$ is

$$u^{n,N}(t,x) := \frac{1}{N}\sum_{i=1}^{N} f(X_{T-t}^{n,(i)}(x)) \ , \tag{14}$$

where $\{X_\cdot^{n,(i)} \ , \ 1 \leq i \leq N\}$ denotes a set of independent trajectories of the process X^n.

Why are we interested in this method? Is it competitive with the usual deterministic algorithms of resolution of (13)? Of course there is no general answer to such a question. The answer depends on the dimension d and on the functions $b(\cdot)$, $\sigma(\cdot)$. Roughly speaking, the Monte Carlo method seems unuseful when a finite difference method, a finite element method, a finite volume method or a suitable deterministic algorithm is numerically stable and does not require too a long computation time.

Nevertheless we can give examples of situations where a Monte Carlo method is efficient.

First, the computational cost of the deterministic algorithms growths exponentially with the dimension d of the state space: these algorithms use grids whose number of points growths exponentially with d. Thus, when d is large ($d \geq 4$, say), the numerical resolution of (13) may even be impossible without a Monte Carlo procedure, whose computational cost growths only linearly with the dimension of the process X^n to simulate.

A Monte Carlo algorithm may also be interesting when one wants to compute $u(t, \cdot)$ at only a few points. This situation occurs in financial problems (evaluation of an option price in terms of the spot prices of the stocks) or in Physics (computation of the probability that a random process reaches given thresholds). One can also think to use a Monte Carlo method to compute $u(t, \cdot)$ on artificial boundaries in view of a decomposition of domains procedure: one divides the whole space in a set of subdomains; then the objective is to solve the problem (13) in each subdomain with Dirichlet boundary conditions by deterministic methods; these Dirichlet boundary conditions, i.e. the values of $u(t, \cdot)$ along the boundaries, can be approximated by a Monte Carlo algorithm. This combination of numerical methods may have several advantages. The resolution in the subdomains can be distributed to different processors. The convergence rate results for the Monte Carlo+Euler method suppose much weaker assumptions than the strong ellipticity condition of the operator $\mathcal{L}$; moreover if $f(\cdot)$ is a smooth function, no assumption on $\mathcal{L}$ is required; therefore, if the matrix $a(\cdot)$ degenerates locally, the domain of the numerical integration by a deterministic method can be reduced to the nondegeneracy region by an approximation of $u(t, \cdot)$ along its boundaries, which may considerably improve the efficiency of the deterministic method.

3.2 Introduction to the error analysis

Our objective is to give estimates for

$$|u(T, x) - u^{n,N}(T, x)|.$$

A natural decomposition of this error is as follows:

$$\begin{aligned} |u(T, x) - u^{n,N}(T, x)| &\leq |u(T, x) - \mathbb{E}f(X_T^n(x))| \\ &\quad + |\mathbb{E}f(X_T^n(x)) - u^{n,N}(T, x)| \\ &=: \alpha^n + \beta^{n,N}. \end{aligned}$$

The analysis of $\beta^{n,N}$ is related to usual considerations on the Strong Law of Large Numbers: Central Limit Theorems, Berry-Esseen inequalities, etc. The difficulty here is to obtain estimates uniform w.r.t. to n. This can be solved by the convergence in $L^p(\Omega)$ of X^n to X which holds under the hypotheses we make below.

Consequently we concentrate our attention to α^n.

When $f(\cdot)$ is a Lipschitz function, one can bound α^n from above by using the estimate (4) for $p=2$. This gives the estimate:

$$|\alpha^n| \leq \frac{C}{\sqrt{n}}.$$

We now show that one can be much more clever.

For the rest of the section we suppose

(H) The functions b and σ are $\mathcal{C}^\infty$ functions, whose derivatives of any order are bounded (but b and σ are not supposed bounded themselves).

Define $\Psi(t,\cdot)$ by

$$\begin{aligned}\Psi(t,\cdot) \;=\; & \frac{1}{2}\sum_{i,j=1}^{d} b^i(\cdot)b^j(\cdot)\partial_{ij}u(t,\cdot) + \frac{1}{2}\sum_{i,j,k=1}^{d} b^i(\cdot)a^j_k(\cdot)\partial_{ijk}u(t,\cdot) \\ & + \frac{1}{8}\sum_{i,j,k,l=1}^{d} a^i_j(\cdot)a^k_l(\cdot)\partial_{ijkl}u(t,\cdot) + \frac{1}{2}\frac{\partial^2}{\partial t^2}u(t,\cdot) \\ & + \sum_{i=1}^{d} b^i(\cdot)\frac{\partial}{\partial t}\partial_i u(t,\cdot) + \frac{1}{2}\sum_{i,j=1}^{d} a^i_j(\cdot)\frac{\partial}{\partial t}\partial_{ij}u(t,\cdot). \qquad (15)\end{aligned}$$

Lemma 3.1 *It holds that*

$$\mathbb{E}f(X^n_T(x)) - \mathbb{E}f(X_T(x)) = \frac{T^2}{n^2}\sum_{k=0}^{n-2}\mathbb{E}\Psi\left(\frac{kT}{n}, X^n_{kT/n}(x)\right) + \sum_{k=0}^{n-1}R^n_k(x)\ , \quad (16)$$

where

$$R^n_{n-1}(x) := \mathbb{E}f(X^n_T(x)) - \mathbb{E}(P_{T/n}f)(X^n_{T-T/n}(x))\ ,$$

and for $k<n-1$, $R^n_k(x)$ can be explicited under a sum of terms, each of them being of the form

$$\begin{aligned}\mathbb{E}\Big[\varphi^\natural_\alpha(X^n_{kT/n}(x)) & \int_{kT/n}^{(k+1)T/n}\int_{kT/n}^{s_1}\int_{kT/n}^{s_2}(\varphi^\sharp_\alpha(X^n_{s_3}(x))\partial_\alpha u(s_3,X^n_{s_3}(x)) \\ & +\varphi^\flat_\alpha(X_{s_3}(x))\partial_\alpha u(s_3,X_{s_3}(x)))ds_3ds_2ds_1\Big]\ , \qquad (17)\end{aligned}$$

where $|\alpha| \leq 6$, and the $\varphi^{\natural}_{\alpha}$'s, $\varphi^{\sharp}_{\alpha}$'s, $\varphi^{\flat}_{\alpha}$'s are products of functions which are partial derivatives up to the order 3 of the a^{ij}'s and b^i's.

Proof. For $z \in \mathbb{R}^d$ define the differential operator $\mathcal{L}_z$ by

$$\mathcal{L}_z g(\cdot) := \sum_{i=1}^{d} b^i(z)\partial_i g(\cdot) + \frac{1}{2}\sum_{i,j=1}^{d} a^i_j(z)\partial_{ij}.$$

As $u(t,\cdot) = P_{T-t}f(\cdot) = \mathbb{E}f(X_{T-t}(\cdot))$, one has

$$\mathbb{E}_x f(X^n_T) - \mathbb{E}_x f(X_T) = \mathbb{E}_x u(T, X^n_T) - u(0,x) = \sum_{k \leq n-1} \delta^n_k$$

with

$$\delta^n_k := \mathbb{E}_x\left[u\left(\frac{(k+1)T}{n}, X^n_{(k+1)T/n}\right) - u\left(\frac{kT}{n}, X^n_{kT/n}\right)\right]. \tag{18}$$

The Itô formula implies

$$\delta^n_k = \mathbb{E}_x \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \left(\partial_t u(t, X^n_t) + \mathcal{L}_z u(t, X^n_t)\mathcal{B}_{z=X^n_{kT/n}}\right) dt\ ,$$

from which, using (13), one gets

$$\delta^n_k = \mathbb{E}_x \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \left(-\mathcal{L}u(t, X^n_t) + \mathcal{L}_z u(t, X^n_t)\mathcal{B}_{z=X^n_{kT/n}}\right) dt.$$

Denote

$$I^n_k(t) := \mathcal{L}_z u(t, X^n_t)\mathcal{B}_{z=X^n_{kT/n}} - \mathcal{L}_z u\left(\frac{kT}{n}, X^n_{kT/n}\right)\mathcal{B}_{z=X^n_{kT/n}}$$

and

$$\begin{aligned} J^n_k(t) &:= \mathcal{L}_z u\left(\frac{kT}{n}, X^n_{kT/n}\right)\mathcal{B}_{z=X^n_{kT/n}} - \mathcal{L}u(t, X^n_t) \\ &= \mathcal{L}u\left(\frac{kT}{n}, X^n_{kT/n}\right) - \mathcal{L}u(t, X^n_t). \end{aligned}$$

We have:

$$\delta^n_k = \mathbb{E}_x \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} (I^n_k(t) + J^n_k(t))dt.$$

We now consider $I_k^n(t)$ and $J_k^n(t)$ as smooth functions of the process (X_t^n) and recursively apply the Itô formula, using the fact that the function u solves (13), so that $\mathcal{L}u$ solves a similar PDE. ■

The expansion (16) can be rewritten as follows:

$$\begin{aligned}
&\mathbb{E}_x f(X_T^n) - \mathbb{E}_x f(X_T) \\
&\quad = \frac{T}{n}\int_0^T \mathbb{E}_x \Psi(s, X_s) ds \\
&\qquad + \frac{T^2}{n^2}\sum_{k=0}^{n-2} \mathbb{E}_x \Psi\left(\frac{kT}{n}, X_{kT/n}\right) - \frac{T}{n}\int_0^T \mathbb{E}_x \Psi(s, X_s) ds \\
&\qquad + \frac{T^2}{n^2}\sum_{k=0}^{n-2} \mathbb{E}_x \left(\Psi\left(\frac{kT}{n}, X_{kT/n}^n\right) - \Psi\left(\frac{kT}{n}, X_{kT/n}\right)\right) \\
&\qquad + \sum_{k=0}^{n-2} r_k^n(x) + \mathbb{E}_x f(X_T^n) - \mathbb{E}_x (P_{T/n} f)(X_{T-T/n}^n) \\
&\quad = \frac{T}{n}\int_0^T \mathbb{E}_x \Psi(s, X_s) ds \\
&\qquad + A^n + B^n + \sum_{k=0}^{n-2} r_k^n(x) + C^n.
\end{aligned} \tag{19}$$

From this expansion, it is reasonable to expect that the error

$$\mathbb{E}_x f(X_T^n) - \mathbb{E}_x f(X_T)$$

is *equal* to

$$\frac{T}{n}\int_0^T \mathbb{E}_x \Psi(s, X_s) ds$$

plus a remainder of order n^{-2}, because

$$\sum_{k=0}^{n-2} \left(\mathbb{E}_x \Psi\left(\frac{kT}{n}, X_{kT/n}^n\right) - \mathbb{E}_x \Psi\left(\frac{kT}{n}, X_{kT/n}\right)\right)$$

should be uniformly bounded w.r.t. n since each term of the sum should be of order $\frac{1}{n}$.

More precisely, for $1 \leq k \leq n-2$, one applies the expansion (19), substituting the function

$$f_{n,k}(\cdot) := \Psi\left(\frac{kT}{n}, \cdot\right)$$

to $f(\cdot)$. Set $u_{n,k}(t,x) := P_{kT/n-t} f_{n,k}(\cdot)$ and denote by $\Psi_{n,k}(t,\cdot)$ the function defined in (15) with $u_{n,k}$ instead of u and kT/n instead of T; thus, for some functions $g_\lambda(\cdot) \in \mathcal{C}_b^\infty(\mathbb{R}^d)$ one has that, for $t \leq \frac{kT}{n}$,

$$\Psi_{n,k}(t,\cdot) = \sum_\lambda g_\lambda(\cdot)\partial_\lambda \left[P_{kT/n-t} \Psi\left(\frac{kT}{n},\cdot\right)\right].$$

There holds

$$\mathbb{E}_x \Psi\left(\frac{kT}{n}, X^n_{kT/n}\right) - \mathbb{E}_x \Psi\left(\frac{kT}{n}, X_{kT/n}\right) = \frac{T^2}{n^2}\sum_{j=0}^{k-2} \mathbb{E}_x \Psi_{n,k}\left(\frac{jT}{n}, X^n_{jT/n}\right) + \sum_{j=0}^{k-1} r_j^{n,k}(x) \ , \tag{20}$$

where the $r_j^{n,k}(x)$'s are sums of terms of type (17) with $u_{n,k}$ instead of u.

It is now clear that one key problem is as follows. Let γ and λ be multiindices, let $g_\gamma(\cdot)$ and $g_\lambda(\cdot)$ be smooth functions with polynomial growth. Set

$$\varphi(\theta,\cdot) := g_\gamma(\cdot)\partial_\gamma P_{T-\theta} f(\cdot).$$

We want to prove that quantities of the type

$$\left| \mathbb{E}_x \left[g_\lambda(X^n_t)\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z)\mathcal{B}_{z=X^n_t}\right]\right| \tag{21}$$

can be bounded uniformly w.r.t. $n \in \mathbb{N}^*$, $\theta \in \left[0, T-\frac{T}{n}\right]$, $t \in \left[0, \theta - \frac{T}{n}\right]$.

We distinguish two different situations. When $f(\cdot)$ is a smooth function, we make no assumption on the operator $\mathcal{L}$. When $f(\cdot)$ is only measurable and bounded, we suppose that $\mathcal{L}$ satisfies an assumption of the Hörmander type.

3.3 Smooth functions $f(\cdot)$

Let $\mathcal{H}_T$ be the class of functions $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with the following properties: ϕ is of class $\mathcal{C}^\infty$ and for any multiindex α there exist a positive integer s and an increasing function $K(\cdot)$ such that

$$\forall \theta \in [0,T] \quad , \quad \forall x \in \mathbb{R}^d \quad , \quad |\ \partial_\alpha \phi(\theta,x)\ | \leq K(T)(1+|\ x\ |^s). \tag{22}$$

A function ϕ of $\mathcal{H}_T$ is said *homogeneous* if it does not depend on the time variable: $\phi(\theta,x) = \phi(x)$.

In this subsection we suppose

(H1) The function $f(\cdot)$ is a *homogeneous* function of $\mathcal{H}_T$.

It is well known that the condition (H) implies that there exists a smooth version of the stochastic flow $x \longrightarrow X_t(x)$. For the sake of simplicity we denote this smooth version $X_t(\cdot)$. Besides, for any integer $k > 0$ the family of the processes equal to the partial derivatives of the flow up to the order k solves a system of stochastic differential equations with Lipschitz coefficients: see, e.g., Kunita [24] and Protter [37]. Thus, for any $0 \leq t \leq T$,

$$\partial_i u(t,x) = \partial_i \mathbb{E} f(X_{T-t}(x)) = \mathbb{E} \sum_{j=1}^{d} \partial_j f(X_{T-t}(x)) \partial_i X_{T-t}(x). \tag{23}$$

From (H) and (H1) one easily deduces that, for some increasing function $K(\cdot)$ and some integer m,

$$|\partial_i u(t,x)| \leq K(T)(1 + |x|^m).$$

Differentiations of (23) provide a probabilistic interpretation of $\partial_\alpha u(t,x)$ for any multiindex α. It is easy to prove by induction that, for any multiindex α, there exist an increasing function $K_\alpha(\cdot)$ and an integer m_α such that

$$|\partial_\alpha u(t,x)| \leq K_\alpha(T)(1 + |x|^{m_\alpha}). \tag{24}$$

For $\phi \in \mathcal{H}_T$ and θ fixed in $[0,T]$ the function $u(\theta;t,x)$ defined by

$$u(\theta;t,x) := \mathbb{E}\phi(\theta, X_{T-t}(x)) = \mathbb{E}_x \phi(\theta, X_{T-t})$$

belongs to $\mathcal{H}_T$ and satisfies

$$\begin{cases} \dfrac{\partial u}{\partial t} + \mathcal{L}u &= \; 0 \;, 0 \leq t < T, \\ u(\theta;T,x) &= \; \phi(\theta,x). \end{cases} \tag{25}$$

Similarly to (24) one has: for any multiindex α, there exists an increasing function $K_\alpha(\cdot)$ and an integer m_α such that

$$\forall \theta \in [0,T] \;, \quad |\; \partial_\alpha u(\theta;t,x) \;| \leq K_\alpha(T)(1 + |\; x \;|^{k_\alpha}).$$

This result can be used to prove:

Lemma 3.2 *Suppose (H) and (H1).*

Let γ et λ be multiindices, let $g(\cdot)$ and $g_\gamma(\cdot)$ be smooth functions with polynomial growth. Set

$$\varphi(\theta,\cdot) := g_\gamma(\cdot)\partial_\gamma P_{T-\theta} f(\cdot).$$

There exist an increasing function $K(\cdot)$ and an integer m such that

$$\left| \mathbb{E}_x \left[g(X_t^n) \partial_\lambda P_{\theta-t} \varphi(\theta,\cdot)(z) \mathcal{B}_{z=X_t^n} \right] \right| \le K(T)(1+|x|^m). \tag{26}$$

Coming back to (19) and (20) one deduces the

Theorem 3.3 (Talay and Tubaro [50]) *Suppose (H) and (H1). The Euler scheme error satisfies*

$$u(T,x) - \mathbb{E}_x f(X_T^n) = -\frac{T}{n} \int_0^T \mathbb{E}_x \Psi(s, X_s) ds + \frac{Q_T^n(x)}{n^2} \tag{27}$$

and there exist an increasing function $K(\cdot)$ and an integer m such that

$$|Q_T^n(x)| \le K(T)(1+|x|^m). \tag{28}$$

Here $\Psi(\cdot,\cdot)$ is defined by (15).

Observe that in the preceding statement the differential operator $\mathcal{L}$ may be degenerate.

3.4 Non smooth functions $f(\cdot)$

Theorem 3.3 supposes that $f(\cdot)$ is a smooth function. From an applied point of view this is a stringent condition: often one wants to compute quantities of the type

$$\mathbb{P}\left[|X_T(x)| > y\right]$$

for a given threshold $y > 0$. Our objective now is to show that an expansion of the type (27) still holds even when $f(\cdot)$ is only supposed measurable and bounded. In the proof that we give, the boundedness could be relaxed: as in the preceding section we could suppose that $f(\cdot)$ belongs to the set $\mathcal{H}_T$. To realize this programme a nondegeneracy condition is supposed. As we now see, this condition is less restrictive than the uniform strong ellipticity of the operator $\mathcal{L}$.

We need some basic elements of the Malliavin calculus. For a complete exposition of this theory we refer to Nualart [34] (we use the notation of this book) and Ikeda-Watanabe [22]; the applications to the existence of a density for the law of a diffusion process can also be found in Pardoux [36].

For $h(\cdot) \in L^2(I\!\!R_+, I\!\!R^r)$, $W(h)$ denotes the quantity $\int_0^T < h(t), dW_t >$. $\mathcal{S}$ is the space of "simple " functionals of the Wiener process W, i.e. the sub-space of $L^2(\Omega, \mathcal{F}, I\!\!P)$ of random variables F which can be written under the form

$$F = f(W(h_1), \dots, W(h_n))$$

for some n, some polynomial function $f(\cdot)$, some $h_i(\cdot) \in L^2(I\!\!R_+, I\!\!R^r)$.

For $F \in \mathcal{S}$, $(D_t F)$ denotes the $I\!\!R^r$-dimensional process defined by

$$D_t F = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), \dots, W(h_n)) h_i(t).$$

The operator D is closable as an operator from $L^p(\Omega)$ to $L^p(\Omega; L^2(0,T))$, for any $p \geq 1$. Its domain is denoted by $I\!\!D^{1,p}$. Define the norm

$$\|F\|_{1,p} := \left[I\!\!E|F|^p + \|DF\|^p_{L^p(\Omega;L^2(0;T))} \right]^{1/p} \quad ,$$

The j-th component of $D_t F$ is denoted by $D_t^j F$. The k-th order derivative is the the random vector on $[0,T]^k \times \Omega$ whose coordinates are

$$D^{j_1,\dots,j_k}_{t_1,\dots,t_k} F := D^{j_k}_{t_k} \dots D^{j_1}_{t_1} F \ ,$$

and $I\!\!D^{N,p}$ denotes the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{N,p} := \left[I\!\!E|F|^p + \sum_{k=1}^N I\!\!E\|D^k F\|^p_{L^2((0;T)^k)} \right]^{1/p} \quad .$$

$I\!\!D^\infty$ denotes the space $\bigcap_{p\geq 1} \bigcap_{j \geq 1} I\!\!D^{j,p}$.

For $F := (F^1, \dots, F^m) \in (I\!\!D^\infty)^m$, γ_F denotes the Malliavin covariance matrix associated to F, i.e. the $m \times m$-matrix defined by

$$(\gamma_F)^i_j := < DF^i, DF^j >_{L^2(0,T)} \ .$$

Definition 3.4 *We say that the random vector F satisfies the* nondegeneracy assumption *if the matrix γ_F is a.s. invertible, and the inverse matrix $\Gamma_F := \gamma_F^{-1}$ satisfies*

$$|\det(\Gamma_F)| \in \bigcap_{p \geq 1} L^p(\Omega).$$

Remark 3.5 The above condition can also be written as follows:

$$\frac{1}{\det(\gamma_F)} \in \bigcap_{p\geq 1} L^p(\Omega).$$

The main ingredient of our analysis is the following integration by parts formula (cf. the section V-9 in Ikeda-Wanabe [22]):

Proposition 3.6 *Let $F \in (\mathbb{D}^\infty)^m$ satisfy the nondegeneracy condition 3.4, let g be a smooth function with polynomial growth, and let G in $\mathbb{D}^\infty$. Let $\{H_\beta\}$ be the family of random variables depending on multiindices β of length strictly larger than 1 and with coordinates $\beta_j \in \{1,\dots,m\}$, recursively defined in the following way:*

$$\begin{aligned} H_i(F,G) &= H_{(i)}(F,G) \\ &:= -\sum_{j=1}^m \Big\{ G < D\Gamma_F^{ij}, DF^j >_{L^2(0,T)} \\ &\qquad\qquad + \Gamma_F^{ij} < DG, DF^j >_{L^2(0,T)} \qquad (29) \\ &\qquad\qquad + \Gamma_F^{ij} \cdot G \cdot \hat{L} F^j \Big\} \ , \\ H_\beta(F,G) &= H_{(\beta_1,\dots,\beta_k)}(F,G) \\ &:= H_{\beta_k}(F, H_{(\beta_1,\dots,\beta_{k-1})}(F,G)) \ , \qquad (30) \end{aligned}$$

where $\hat{L}$ is the so called Ornstein-Uhlenbeck operator whose domain includes $\mathbb{D}^\infty$. Then, for any multiindex α,

$$\mathbb{E}[(\partial_\alpha g)(F) G] = \mathbb{E}[g(F) H_\alpha(F,G)]. \qquad (31)$$

One has the following estimate:

Proposition 3.7 *For any $p > 1$ and any multiindex β, there exist a constant $C(p,\beta) > 0$ and integers $k(p,\beta)$, $m(p,\beta)$, $m'(p,\beta)$, $N(p,\beta)$, $N'(p,\beta)$, such that, for any measurable set $A \subset \Omega$ and any F, G as above, one has*

$$\begin{aligned} \mathbb{E}[|H_\beta(F,G)|^p \ \mathbb{1}_{\mathrm{A}}]^{\frac{1}{p}} &\leq C(p,\beta) \ \|\Gamma_F \ \mathbb{1}_{\mathrm{A}}\|_{k(p,\beta)} \ \|G\|_{N(p,\beta),m(p,\beta)} \\ &\qquad \|F\|_{N'(p,\beta),m'(p,\beta)}. \qquad (32) \end{aligned}$$

We now state another classical result, which concerns the solutions of stochastic differential equations considered as functionals of the driving Wiener process. $[A, A']$ denotes the Lie brackett of two vector fields A and A'.

Definition 3.8 *Denote by $A_0, A_1, \dots, A_r$ the vector fields defined by*

$$A_0(x) := \sum_{i=1}^{d} b^i(x)\partial_i \ ,$$
$$A_j(x) := \sum_{i=1}^{d} \sigma_j^i(x)\partial_i \quad , \quad j = 1, \dots, r.$$

For a multiindex $\alpha = (\alpha_1, \dots, \alpha_k) \in \{0, 1, \dots r\}^k$, define the vector fields A_i^α ($1 \le i \le r$) by induction: $A_i^\emptyset := A_i$ and for $0 \le j \le r$, $A_i^{(\alpha,j)} := [A^j, A_i^\alpha]$.

Finally set

$$V_L(x, \eta) := \sum_{i=1}^{r} \sum_{|\alpha| \le L-1} < A_i^\alpha(x), \eta >^2 .$$

Set

$$V_L(x) = 1 \wedge \inf_{\|\eta\|=1} V_L(x, \eta). \tag{33}$$

Under the hypothesis (H), $X_t(x) \in \mathbb{D}^\infty$ for any $x \in \mathbb{R}^d$. Let $\gamma_t(x)$ denote the Malliavin covariance matrix of $X_t(x)$ and let $\Gamma_t(x)$ denote its inverse.

We replicate Corollary 3.25 in Kusuoka and Stroock [26] in a weakened form.

Proposition 3.9 *Suppose (H) and*

(UH) $C_L := \inf_{x \in \mathbb{R}^d} V_L(x) > 0$ *for some integer L.*

Let L be an integer such that (UH) holds. Then

$$\|\Gamma_t(x)\| \in \bigcap_{p \ge 1} L^p(\Omega) \ , \forall x \in \mathbb{R}^d \ ,$$

and for any $p \ge 1$, for some constant μ and some increasing function $K(\cdot)$,

$$\|\Gamma_t(x)\|_p \le K(T) \frac{1 + |x|^\mu}{t^{dL}} \ , \ \forall x \in \mathbb{R}^d \ , \ \forall 0 < t \le T. \tag{34}$$

Thus, for any $t > 0$ and any $x \in \mathbb{R}^d$ the law of $X_t(x)$ has a smooth density $p_t(x, \cdot)$. Besides, for any integers m, k and any multiindices α and β such that $2m + |\alpha| + |\beta| \le k$, there exist an integer $M(k, L)$, a non decreasing function

$K(\cdot)$ and real numbers C, q, Q depending on $L, T, m, k, \alpha, \beta$ and on the bounds associated to the coefficients of the stochastic differential equation and their derivatives up to the order $M(k, L)$, such that the following inequality holds[1]*:*

$$|\partial_t^m \partial_x^\alpha \partial_y^\beta p_t(x,y)| \leq \frac{K(T)(1+|x|^Q)}{t^q(1+|y-x|^2)^k} \exp\left(-C\frac{(|x-y| \wedge 1)^2}{t(1+|x|)^2}\right) \quad , \quad \forall 0 < t \leq T. \tag{35}$$

Equipped with this result we can prove the

Theorem 3.10 (Bally and Talay [1]) *Let $f(\cdot)$ be a measurable and bounded function. Under the hypotheses (UH) and (H), the Euler scheme error satisfies*

$$\mathbb{E}f(X_T(x)) - \mathbb{E}f(X_T^n(x)) = -\frac{C_f(T,x)}{n} + \frac{Q_n(f,T,x)}{n^2}. \tag{36}$$

The terms $C_f(T,x) := \int_0^T \mathbb{E}\Psi(s, X_s(x))ds$ and $Q_n(f,T,x)$ have the following property: there exists an integer m, a non decreasing function $K(\cdot)$ depending on the coordinates of a and b and on their derivatives up to the order m, and positive real numbers q, Q such that

$$|C_f(T,x)| + sup_n|Q_n(f,T,x)| \leq K(T)\|f\|_\infty \frac{1+\|x\|^Q}{T^q}. \tag{37}$$

Sketch of the proof. As for Theorem 3.3, the main part of the proof consists in bounding terms of the type (21) from above. In the present context there is a serious difficulty: when $f(\cdot)$ is not smooth, the spatial derivatives of $u(t,\cdot)$ explode when t goes to T. Indeed,

$$u(t,x) = \int_{\mathbb{R}^d} p_{T-t}(x,y)f(y)dy$$

and the estimate (35) shows that for any $|\gamma| \geq 1$,

$$|\partial_\gamma^x p_{T-t}(x,y)| \leq \frac{K(T-t)}{(T-t)^q}(1+\|x\|^Q)\frac{1}{(1+\|y-x\|^2)^{|\gamma|}} \ ,$$

from which

$$|\partial_\alpha^x u(t,x)| \leq K(T)\frac{\|f\|_\infty}{(T-t)^q}(1+\|x\|^Q). \tag{38}$$

It can be shown that there is no hope to improve the explosion rate in power of $T-t$.

[1] The constant γ_0 of the statement of Kusuoka and Stroock is equal to 1 under (H).

But a miracle occurs: in (21) the derivatives of the function

$$P_{\theta-t}\varphi(\theta,\cdot)$$

are integrated w.r.t. the law of $X_t^n(x)$. Let us give an intuition of what happens. Consider the case $\theta = t$ and replace $X_t^n(x)$ by $X_t(x)$. Then, in view of (21) the problem becomes to bound from above an expression of the type

$$|\mathbb{E}_x\left[g_\gamma(X_t)\partial_\gamma u(t,X_t)\right]| = |\mathbb{E}_x\left[g_\gamma(X_t)\partial_\gamma(P_{T-t}f)(X_t)\right]|$$

uniformly w.r.t. $t\in[0,T)$. When t is "small" i.e $t\leq\frac{T}{2}$ the transition operator P_{T-t} has smoothened enough the initial condition $f(\cdot)$: the inequality (38) implies

$$|\partial_\alpha^x u(t,x)| \leq K(T)\frac{\|f\|_\infty}{T^q}(1+\|x\|^Q).$$

When t is "large" i.e $t\geq\frac{T}{2}$ the estimate (38) cannot be used. Instead, we observe that the matrix $\Gamma_t(x)$ has L^p-norms which satisfy (see (34))

$$\|\Gamma_t(x)\| \leq K(T)\frac{1+|x|^\mu}{T^{dL}}.$$

Thus, one can apply the integration by part formula (31) with

$$g(\cdot) = (P_{T-t}f)(\cdot) = u(t,\cdot)$$

and

$$F = X_t(x).$$

Using (32) one deduces that for $T\geq t\geq\frac{T}{2}$,

$$|\mathbb{E}_x\left[g_\gamma(X_t)\partial_\gamma u(t,X_t)\right]| \leq K(T)\|f\|_\infty\frac{1+|x|^\mu}{T^{dL}}.$$

This would be perfect if we would not have to consider $F = X_t^n(x)$ rather than $F = X_t(x)$: we must take care that $X_t^n(x)$ does not satisfy the nondegeneracy condition (3.4). We now explain the reason.

On one hand, one can easily prove the following: for any $p>1$ and $j\geq 1$, there exist an integer Q and a non decreasing function $K(\cdot)$ such that

$$\sup_{n\geq 1}\|X_t^n(x)\|_{j,p} < K(t)(1+\|x\|^Q) \tag{39}$$

and

$$\sup_{n\geq 1}\|X_t(x)-X_t^n(x)\|_{j,p} < \frac{K(t)}{\sqrt{n}}(1+\|x\|^Q). \tag{40}$$

On the other hand this result is far from satisfactory in view of the condition (3.4): indeed, if (Z^n) is a sequence of random variables, the convergence to a random variable Z in $L^p(\Omega)$ does not imply that $\frac{1}{Z^n}$ is in $L^p(\Omega)$.

At this step of the proof a localization argument seems necessary. Let γ_t^n denote the Malliavin covariance matrix of X_t^n and let Γ_t^n denotes its inverse (where it is defined). We recall that we are considering the case $T \geq t \geq \frac{T}{2}$. Let Ω_0 be the set of events where $|\hat{\gamma}_t^n - \hat{\gamma}_t|$ is larger than $\frac{\hat{\gamma}_t}{4}$. Using (34) and (40) one proves that $\mathbb{P}(\Omega_0)$ is small. On the complementary set of Ω_0, $|\hat{\gamma}_t^n - \hat{\gamma}_t|$ is small, which (roughly speaking) means that the Malliavin covariance matrix of $X_t^n(x)$ behaves like that of $X_t(x)$ (see (34)), which allows integrations by parts of the type (31) with a good control of the L^p-norms of the variables H_α. ■

3.5 Extensions

In the preceding proof, we have integrated by parts in order to make appear $f(\cdot)$ instead of derivatives of $u(t,\cdot)$. One can refine the method to get an expansion for

$$p_T(x,y) - \tilde{p}_T^n(x,y)$$

where $p_T(x,y)$ denotes the density of $X_T(x)$ and $\tilde{p}_t^n(x,\cdot)$ denotes the density of the law of a suitable small perturbation of $X_t^n(x)$ (the law of $X_t^n(x)$ may have no density, see our remark above on $\Gamma_t^n(x)$). To treat this problem, it is natural to fix y, choose $f_\delta(\xi) = \rho_\delta(y-\xi)$ where the $\rho_\delta(\cdot)$'s are such that the sequence of measures $(\rho_\delta(\xi)d\xi)$ converges weakly to the Dirac measure at 0, and make δ tend to 0. Theorem 3.10 is not sufficient since, when δ tends to 0, $(\| f_\delta \|_\infty)$ tends to infinity. Nevertheless, if F_δ is the distribution function of the measure $f_\delta(\xi)d\xi$, the sequence $(\| F_\delta \|_\infty)$ is constant: this gives the idea of proving inequalities of the type (37) with $\| F \|_\infty$ instead of $\| f \|_\infty$ when $f(\cdot)$ has a compact support, $F(\cdot)$ being the distribution function of the measure $f(\xi)d\xi$. A supplementary difficulty is to prove that, instead of $(1+\|x\|^Q)$ appears a function which satisfies an exponential upper bound and that the function $C_{f_\delta}(T,x)$ itself satisfies an exponential upper bound: such estimates permit to conclude that, when the differential operator $\mathcal{L}$ in (12) is strongly uniformly elliptic, that the density $p_T^n(x,y)$ of $X_t^n(x)$ (which does exist in this case) satisfies:

$$\forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^d,\ p_T(x,y) - p_T^n(x,y) = -\frac{1}{n}\pi_T(x,y) + \frac{1}{n^2}R_T^n(x,y) \quad (41)$$

and there exists a strictly positive constant c, an integer q and an increasing function $K(\cdot)$ such that

$$|\pi_T(x,y)| + |R_T^n(x,y)| \leq \frac{K(T)}{T^q} \exp\left(-c\frac{\|x-y\|^2}{T}\right) \,. \tag{42}$$

For a complete exposition and a precise result, see Bally and Talay [2].

Observe also that the expansion of the error (36) justifies the Romberg extrapolation procedure. Indeed, for some function $e(\cdot)$ one has

$$\mathbb{E}f(X_T) - \mathbb{E}f\left(X_T^n\right) = \frac{T}{n}e(T) + \mathcal{O}\left(\frac{1}{n^2}\right) \,,$$

and

$$\mathbb{E}f(X_T) - \mathbb{E}f\left(X_T^{2n}\right) = \frac{T}{2n}e(T) + \mathcal{O}\left(\frac{1}{n^2}\right).$$

Consider the new approximate value

$$Z_T^n := 2\mathbb{E}f(X_T^{2n}) - \mathbb{E}f(X_T^n) \,, \tag{43}$$

then

$$\mathbb{E}f(X_T) - Z_T^n = \mathcal{O}(n^{-2}).$$

Thus, a precision of order n^{-2} is achieved by a linear combination of the results produced by the Euler scheme with 2 different step sizes. For numerical examples and comments, see Talay and Tubaro [50].

In the context of the present subsection, the Milshtein scheme (9) (for $d = r = 1$) has the same convergence rate as the Euler scheme (contrarily to the approximation in $L^p(\Omega)$). The expansion of the Milshtein scheme error makes appear a different function $\Psi(\cdot)$.

The results given above only concern SDE's driven by a Wiener process. One can extend both the convergence rate analysis and the simulation technique to SDE's driven by Lévy processes, which corresponds to the analysis of Monte-Carlo methods for integro-differential equations of the type

$$\frac{\partial u}{\partial t}(t,x) = \tag{44}$$
$$Lu(t,x) + \int_{\mathbb{R}^d} \{u(t,x+z) - u(t,x) - < z, \nabla u(t,x) > \ \mathbb{1}_{[\|z\|\leq 1]}\} M(x,dz)$$

where L is the elliptic operator (12) and the measure $M(x,\cdot)$ is defined as follows: let ν be a measure on $\mathbb{R}^d - \{0\}$ such that

$$\int_{\mathbb{R}^d} (\| x \|^2 \wedge 1)\nu(dx) < \infty$$

and let $g(\cdot)$ be a $d \times r$-matrix valued function defined in $\mathbb{R}^d$; then, for any Borel set $B \subset \mathbb{R}^d$ whose closure does not contain 0, set

$$M(x, B) := \nu\{z \; ; \; < g(x), z > \in B\}.$$

Consider a Lévy process (Z_t) and (X_t) solution to

$$X_t = X_0 + \int_0^t g(X_{s-}) dZ_s. \tag{45}$$

For $K > 0$, $m > 0$ and $p \in \mathbb{N} - \{0\}$, set

$$\begin{aligned} \rho_p(m) \;\; := \;\; & 1 + \|\beta\|^2 + \|\sigma\|^2 + \int_{-m}^{m} \|z\|^2 \nu(dz) + \|\beta\|^p + \|\sigma\|^p \\ & + \left(\int_{-m}^{m} \|z\|^2 \nu(dz) \right)^{p/2} + \int_{-m}^{m} \|z\|^p \nu(dz) \end{aligned} \tag{46}$$

where ν is the Lévy measure of (Z_t), and

$$\eta_{K,p}(m) := \exp\left(K \rho_p(m) \right). \tag{47}$$

For $m > 0$ we define

$$h(m) := \nu(\{x; \|x\| \geq m\}). \tag{48}$$

Theorem 3.11 (Protter and Talay [38]) *Suppose:*

(H1) *the function $f(\cdot)$ is of class $\mathcal{C}^4$; $f(\cdot)$ and all derivatives up to order 4 are bounded;*

(H2) *the function $g(\cdot)$ is of class $\mathcal{C}^4$; $g(\cdot)$ and all derivatives up to order 4 are bounded;*

(H3) $X_0 \in L^4(\Omega)$.

Then there exists a strictly increasing function $K(\cdot)$ depending only on d, r and the L^∞-norm of the partial derivatives of $f(\cdot)$ and $g(\cdot)$ up to order 4 such that, for any discretization step of type $\frac{T}{n}$, for any integer m,

$$|\mathbb{E}g(X_T) - \mathbb{E}g(\bar{X}_T^n)| \leq 4\|g\|_{L^\infty(\mathbb{R}^d)}(1 - \exp(-h(m)T)) + \frac{\eta_{K(T),8}(m)}{n}. \tag{49}$$

Thus, the convergence rate is governed by the rate of increase to infinity of the functions $h(\cdot)$ and $\eta_{K(T),8}(\cdot)$.

Stronger hypotheses permit to get much more precise results:

Theorem 3.12 *Suppose:*

(H1') *the function $f(\cdot)$ is of class $\mathcal{C}^4$; all derivatives up to order 4 of $f(\cdot)$ are bounded;*

(H2') *the function $g(\cdot)$ is of class $\mathcal{C}^4$ and moreover $|\partial_I g(x)| = \mathcal{O}(\|x\|^{M'})$ for $|I| = 4$ and some $M' \geq 2$;*

(H3') $\int_{\|x\|\geq 1} \|x\|^\gamma \nu(dx) < \infty$ *for* $2 \leq \gamma \leq M'^* := \max(2M', 8)$ *and* $X_0 \in L^{M'^*}(\Omega)$.

Then there exists an increasing function $K(\cdot)$ such that, for all $n \in \mathbb{N} - \{0\}$,

$$|\mathbb{E}g(X_T) - \mathbb{E}g(\bar{X}_T^n)| \leq \frac{\eta_{K(T),M'^*}(\infty)}{n}. \tag{50}$$

Suppose now:

(H1") *the function $f(\cdot)$ is of class $\mathcal{C}^8$; all derivatives up to order 8 of $f(\cdot)$ are bounded;*

(H2") *the function $g(\cdot)$ is of class $\mathcal{C}^8$ and moreover $|\partial_I g(x)| = \mathcal{O}(\|x\|^{M''})$ for $|I| = 8$ and some $M'' \geq 2$;*

(H3") $\int_{\|x\|\geq 1} \|x\|^\gamma \nu(dx) < \infty$ *for* $2 \leq \gamma \leq M''^* := 2\max(2M'', 16)$ *and* $X_0 \in L^{M''^*}(\Omega)$.

Then there exists a function $C(\cdot)$ and an increasing function $K(\cdot)$ such that, for any discretization step of type $\frac{T}{n}$, one has

$$\mathbb{E}g(X_T) - \mathbb{E}g(\bar{X}_T^n) = \frac{C(T)}{n} + R_T^n \tag{51}$$

and $\sup_n n^2 |R_T^n| \leq \eta_{K(T),M''^*}(\infty)$.

3.6 Newton's variance reduction technique

In Newton [33] are presented variance reduction techniques for the Monte Carlo computation of quantities of the type

$$\mathbb{E}\Phi(X_\cdot)$$

where $\Phi(\cdot)$ is a real valued functional defined on $\mathcal{C}([0,T];\mathbb{R}^d)$. In the preceding subsection we have considered a much less general situation:

$$\Phi(\omega_\cdot) = f(\omega_T). \tag{52}$$

Newton proposes a general methodology to reduce the variance of the Monte Carlo procedure. His rather complex approach is based upon Haussmann's integral representation theorem applied to $\Phi(X_\cdot)$. The analysis is considerably simplified in the context (52) to which we limit ourselves here. In this context, the principle of Newton's method is as follows. Write

$$f(X_T(x)) = \mathbb{E}f(X_T(x)) + \int_0^T (\partial u)(t, X_t(x))\sigma(X_t(x))dW_t \ , \ a.s.$$

and set

$$Z := f(X_T(x)) - \int_0^T (\partial u)(t, X_t(x))\sigma(X_t(x))dW_t.$$

Of course Z is an unbiased estimator of $\mathbb{E}f(X_T(x))$ and the variance of the error is 0. Now suppose that one knows an approximation $\bar{v}$ of ∂u. Then it is natural to consider

$$\bar{Z} := f(X_T(x)) - \int_0^T \bar{v}(t, X_t(x))\sigma(X_t(x))dW_t.$$

$\bar{Z}$ is an unbiased estimator of $\mathbb{E}f(X_T(x))$; the error of the variance is

$$\begin{aligned} &\mathbb{E}|\bar{Z} - \mathbb{E}f(X_T(x))|^2 \\ &\quad = \mathbb{E}\left| f(X_T(x)) - \int_0^T \bar{v}(t, X_t(x))\sigma(X_t(x))dW_t - \mathbb{E}f(X_T(x))\right|^2 \\ &\quad = \mathbb{E}\int_0^T |((\partial u)(t, X_t(x)) - \bar{v}(t, X_t(x)))\sigma(X_t(x))|^2 \, dt. \end{aligned}$$

Thus, the variance may be small is $\bar{v}(\cdot,\cdot)$ is a good approximation of $u(\cdot,\cdot)$ in the sense that the right hand side of the preceding inequality is small ($\bar{v}(\cdot,\cdot)$

can be seen as an approximation of $\partial u(\cdot,\cdot)$ in a suitable Hilbert space). In such a case, one approximates

$$\int_0^T (\partial u)(t, X_t(x))\sigma(X_t(x))dW_t$$

by the sum

$$\sum_{p=1}^{n} \bar{v}(pT/n, X^n_{pT/n}(x))\sigma(X^n_{pT/n}(x))(W_{(p+1)T/n} - W_{pT/n}).$$

Such a variance reduction technique is called a “control variate” technique. Newton also proposes a methodology to construct “importance sampling” methods. See [33].

3.7 Lépingle's reflected Euler scheme

Elliptic and parabolic PDE's with a Dirichlet condition at the boundary lead to probabilistic interpretations in terms of diffusion processes stopped at the boundary. If the boundary condition is of the Neumann type then the probabilistic interpretations involve reflected diffusion proceses. See Bensoussan and Lions [3] or Freidlin [14] e.g.

We do not discuss here the approximation of stopped diffusions. Only a few convincing results are available, see Milshtein [32].

For reflected diffusions on the boundary of the half-space, Lépingle [27] has constructed and analysed a version of the Euler scheme which mimics the reflection and is numerically efficient in the sense that the random variables involved in the scheme are easy to simulate.

We first define a diffusion process obliquely reflected at the boundary of the half-space.

For $d > 1$ consider the domain $D := \mathbb{R}^{d-1} \times \mathbb{R}^*_+$. Suppose that $X_0 \in \bar{D}$ *a.s.*. Fix a vector

$$\gamma := (\gamma_1, \ldots, \gamma_{d-1}, 1)$$

in $\mathbb{R}^d$.

Suppose that the functions $b(\cdot)$ and $\sigma(\cdot)$ are globally Lipschitz. Then there exists a unique adapted continuous process X with values in $\bar{D}$ and a unique adapted continuous nondecreasing process L such that for any $t \in [0,T]$,

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s + \gamma L_t \ , \\ L_t &= \int_0^t \mathbb{1}_{\{X^d_s = 0\}} dL_s. \end{aligned}$$

The process L is given by

$$L_t = \sup_{0\le s\le t} (X_s^d - L_s)^- .$$

The reflected Euler scheme is as follows:

$$\left\{\begin{array}{rcl} X_0^n & = & X_0 \ , \\ X_{(p+1)T/n}^n & = & X_{pT/n}^n + b(X_{pT/n}^n)\frac{T}{n} \\ & & \quad +\sigma(X_{pT/n}^n)(W_{(p+1)T/n} - W_{pT/n}) \\ & & \quad +\gamma\max(0, A_{T/n}^n(p) - X_{pT/n}^{n,d}) \end{array}\right. \tag{53}$$

where

$$A_\theta^n(p) := \sup_{pT/n\le s\le \theta+pT/n} \{-b^d(X_{pT/n}^n)(s-pT/n) - \sum_{j=1}^r \sigma_j^d(X_{pT/n}^n)(W_s^j - W_{pT/n}^j)\}.$$

The simulation of the reflected Euler scheme requires the simulation of the pair

$$\left(W_{(p+1)T/n} - W_{pT/n}, A_{T/n}^n(p)\right)$$

at each step. This can be efficiently done, as proven in Lépingle [27]:

Proposition 3.13 *Let* $\alpha = (\alpha_1, \ldots, \alpha_r)$ *be a vector of* $\mathbb{R}^r$ *and let* c *be a real number. Set*

$$S_t := \sup_{s\le t} (< \alpha, W_s > + cs) \ .$$

Let $U = (U_1, \ldots, U_r)$ *be a Gaussian vector of zero mean with covariance matrix* $t\ Id$*, and let* V *be an exponential random variable with parameter* $(2t)^{-1}$ *independent of* U*. Set*

$$Y := \frac{1}{2}(< \alpha, U > + ct + (|\alpha|^2 V + (< \alpha, U > + ct)^2)^{1/2}).$$

Then the vectors (W_t, S_t) *and* (U, Y) *have the same law.*

Now define a continuous-time version of the preceding scheme, coinciding with $X_{pT/n}^n$ at each time pT/n: for $\frac{kT}{n} \le t < \frac{(k+1)T}{n}$,

$$\left\{\begin{array}{rcl} X_0^n & = & X_0 \ , \\ X_t^n & = & X_{kT/n}^n + b(X_{kT/n}^n)\left(t - \frac{kT}{n}\right) + \sigma(X_{kT/n}^n)(W_t - W_{kT/n}) \\ & & \quad +\gamma \sup_{pT/n\le s\le t}(A_{s-pT/n}^n(p) - X_{pT/n}^{n,d}). \end{array}\right. \tag{54}$$

One has the following convergence result, similar to (4):

Theorem 3.14 (Lépingle [27]) *Suppose $b(\cdot)$ and $\sigma(\cdot)$ are Lipschitz functions and that $\mathbb{E}|X_0|^2 < \infty$. Then, for some constants C_1 and C_2 uniform w.r.t. n,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - X_t^n|^2\right] \le \frac{C_1 \exp(C_2 T)}{n}. \tag{55}$$

In [28] Lépingle extends his analysis to the case of hypercubes with normal reflections.

An estimate for the approximation of $\mathbb{E}f(X_T)$ would be useful: this work is in progress. A result of this nature has just appeared in a manuscript by Costantini, Pacchiarotti and Sartoretto [12].

An original numerical procedure is proposed by Liu [29]. This procedure is based upon a penalization technique.

For Monte Carlo methods coupled with the simulation of obliquely reflecting Brownian motions, Calzolari, Costantini and Marchetti [8] give confidence intervals.

Other approximation problems are investigated by Slominski [44] and [45] for much wider classes of semimartingales and much larger types of reflections. As expected the approximating processes are less easy to simulate than Lépingle's scheme and the convergence rates are lower. Other references can be found in [27].

3.8 The stationary case

In this subsection we assume

(H3) the functions b, σ are of class $\mathcal{C}^\infty$ with bounded derivatives of any order; the function σ is bounded;

(H4) the operator L is uniformly elliptic: there exists a strictly positive constant α such that

$$\forall x,\ \xi \in \mathbb{R}^d\ ,\ \sum_{i,j} a_j^i(\xi) x^i x^j \ge \alpha |x|^2\ ;$$

(H5) there exists a strictly positive constant β and a compact set K such that:

$$\forall x \in \mathbb{R}^d - K\ ,\ x \cdot b(x) \le -\beta |x|^2.$$

It is well known that (H3)-(H5) are (even too strong) sufficient conditions for the ergodicity of (X_t): see for instance Hasminskii [21]. Thus, (X_t) has a unique invariant probability measure μ. The hypothesis (H4) implies the existence of a smooth density $p(\cdot)$ for μ. This density solves the stationary parabolic PDE

$$L^* p(\cdot) = 0. \tag{56}$$

Our objective is to approximate

$$\int_{\mathbb{R}^d} f(y)p(y)dy$$

for a given function $f(\cdot)$ in $L^1(\mu)$.

Theorem 3.15 (Talay and Tubaro [50]) *Assume (H3)-(H5).*

The Euler scheme defines an ergodic Markov chain.

Let $f(\cdot)$ be a real function of class $\mathcal{C}^\infty(\mathbb{R}^d)$. Assume that $f(\cdot)$ and any of its partial derivatives have a polynomial growth at infinity.

Let Ψ be defined as in (15). Set

$$\lambda := \int_0^{+\infty} \int_{\mathbb{R}^d} \Psi(t,y)\mu(dy)dt.$$

Then the Euler scheme with step size $\frac{1}{n}$ satisfies: for any deterministic initial condition $\xi = \overline{X}^h_0$,

$$\int f(y)\mu(dy) - a.s. \lim_{N\to+\infty} \frac{1}{N} \sum_{p=1}^{N} f(X^n_{p/n}(\xi)) = -\frac{\lambda}{n} + \mathcal{O}\left(\frac{1}{n^2}\right). \tag{57}$$

Sketch of the proof. The ergodicity of the Euler scheme can be proven by using a sufficient criterion due to Tweedie [51]: first, one can check that there exits a compact set which is reached in finite time by the chain $(X^n_{p/n})$ with a strictly positive probability; second, it is easy to check that for any n large enough, there exists $\epsilon > 0$ such that for all deterministic starting point $X^n_0 = x$ outside this compact set,

$$\mathbb{E}|X^n_{1/n}|^2 \le |x|^2 - \epsilon.$$

Next, we oberve that the measure μ has finite moments of any order. Similarly,

$$\begin{aligned} &\forall p \in \mathbb{N}\ ,\ \exists C_p > 0\ ,\ \exists \gamma_p > 0\ ,\ \exists n_0 > 0\ ,\ \ \forall n \ge n_0\ , \\ &\mathbb{E}|X^n_t(x)|^p \le C_p(1 + |x|^p e^{-\gamma_p t})\ ,\ \forall t > 0\ ,\ \forall x \in \mathbb{R}. \end{aligned} \tag{58}$$

Note that (58) imply that $\mathbb{E}_x f(X_t^n)$ is well defined.

Equipped with these preliminary results, our main ingredient to prove (57) is the following. Set

$$u(t,x) := \mathbb{E}_x f(X_t) - \int f(y)d\mu(y).$$

Then, for any multiindex α there exist an integer s_α, there exist strictly positive constants Γ_α and γ_α such that

$$|\partial_\alpha u(t,x)| \leq \Gamma_\alpha(1+|x|^{s_\alpha})e^{-\gamma_\alpha t} \ , \ \forall t>0 \ , \ \forall x \in \mathbb{R}^d. \tag{59}$$

The proof of this estimate is technical (see Talay [47]). One step is to show that for any multiindex I, if M_I is defined by

$$|I| = \mathit{integer\ part\ of}\ (M_I - d/2) \ ,$$

and if

$$\pi_s(x) := \frac{1}{(1+|x|^2)^s} \ ,$$

there holds, for $s \in \mathbb{N}$ large enough:

$$\begin{aligned} &\exists C_I > 0 \ , \ \exists \lambda_I > 0 \ , \ \forall |\alpha| \leq M_I \ , \ \forall t > 0 \ , \\ &\int |\partial_\alpha u(t,x)|^2 \pi_s(x) dx \leq C_I \exp(-\lambda_I t). \end{aligned} \tag{60}$$

An easy computation shows that the preceding inequality implies that

$$\exists C_I \ , \ \lambda_I \quad : \quad \forall |\alpha| \leq M_I \ , \ \forall t>0 \ , \ \int |\partial_\alpha (u(t,x)\pi_s(x))|^2 dx \leq C_I \exp(-\lambda_I t).$$

We then can deduce (59) by using the Sobolev imbedding Theorem.

Next, one observes that

$$\begin{aligned} \frac{1}{N}\sum_{p=1}^{N} \mathbb{E}_x f(X_{p/n}^n) \ &= \ \frac{1}{N}\sum_{p=1}^{N} u(p/n,x) + \frac{1}{Nn^2}\sum_{j=1}^{N}\sum_{p=0}^{N-j} \mathbb{E}_x \Psi(j/n, X_{p/n}^n) \\ &\quad + \frac{1}{Nn^3}\sum_{p=1}^{N} \mathcal{R}_p^n \end{aligned}$$

where $\mathcal{R}_p^n$ is a sum of terms, each term being a product of derivatives of $b(\cdot)$, $\sigma(\cdot)$ and $u(p/n,\cdot)$. Then one makes N tend to infinity. The exponential decay in (59) permits to control the sum of the remainders $\mathcal{R}_p^n$ and to prove that

$$\lim_{N\to\infty} \frac{1}{Nn} \sum_{j=1}^{N} \sum_{p=0}^{N-j} \mathbb{E}_x \Psi(j/n, X_{p/n}^n) = \int_0^{+\infty} \int_{\mathbb{R}^d} \Psi(t,y)\mu(dy)dt.$$

■

The Milshtein scheme (9) (for $d = r = 1$) has the same convergence rate as the Euler scheme. The expansion of the Milshtein scheme error makes appear a different function $\Psi(\cdot)$.

As in the non stationary case, the expansion of the error in terms of $\frac{1}{n}$ justifies a Romberg extrapolation which permits to accelate the convergence rate. See [50] for numerical experiments.

PART II - Stochastic Particle Methods

In this part, we analyse stochastic particle methods for nonlinear PDE's in a few special cases. Our objective is to establish the convergence rates which can be observed in numerical experiments for PDE's such that an explicit solution is known, especially the rate $N^{-1/2}$ where N is the number of particles of the algorithm.

Works in progress at Inria, based on the results presented below, have for objective the analysis of the random vortex methods for the incompressible 2D Navier-Stokes equation developed by Chorin, Hald, etc (Chorin [11], Chorin and Marsden [10], Goodman [17], Hald ([19] and [20]), Long [30], Puckett ([40], [39]) e.g.; see also the bibliography in [11] and in the different contributions of [18]).

From now on, we suppose

$$d = r = 1.$$

We also suppose

(H6) $b(\cdot)$ and $\sigma(\cdot)$ are bounded functions of class $\mathcal{C}^\infty(I\!\!R)$; any derivative of any order is assumed bounded;

(H7) $\sigma(x) \geq \sigma_0 > 0$, $\forall x \in I\!\!R$.

We continue to set $a(\cdot) := \sigma^2(\cdot)$.

4 Introduction to the stochastic particle methods

Let $V_0(\cdot)$ be the distribution function of a probability law. Consider the PDE in $(0,T] \times I\!\!R$

$$\left\{ \begin{array}{lcl} \frac{\partial V}{\partial t}(t,x) & = & \frac{1}{2} a(x) \frac{\partial^2}{\partial x^2} V(t,x) \\ & & + \left(\frac{1}{2} a'(x) - b(x)\right) \frac{\partial}{\partial x} V(t,x) \textit{ in } (0,T] \times I\!\!R \ , \\ \lim_{t \to 0} V(t,x) & = & V_0(x) \textit{ at all continuity points of } V_0(\cdot). \end{array} \right. \qquad (61)$$

It is well known (see for instance [15]) that under (H6)-(H7) the law of $X_t(x)$ has a smooth density $p_t(x,\cdot)$ for all $x \in I\!\!R$ and all $t > 0$; this density satifies

$$\begin{cases} \frac{\partial p}{\partial t}(t,x) = L^* p(t,x) \ , \ \forall t>0 \ , \ \forall x \in I\!R \ , \\ \\ p_t(x,\xi)d\xi \xrightarrow{w} \delta_x. \end{cases} \tag{62}$$

Besides, there exists an increasing function $K(\cdot)$ and a constant $\lambda > 0$ such that, for all $(x,y) \in I\!R^2$,

$$p_t(x,y) \le \frac{K(t)}{\sqrt{t}} \exp\left(-\frac{(x-y)^2}{2\lambda t}\right). \tag{63}$$

From (62) it is easy to see that $V(t,\cdot)$ is the distribution function of the law of X_t when the law of X_0 is $\mu_0(d\xi) := dV_0(\xi)$.

Let $H(\cdot)$ denote the Heaviside function ($H(z) = 0$ if $z < 0$, $H(z) = 1$ if $z \ge 1$). Set

$$\omega_0^i = \frac{1}{N} \ , \ \text{for } i = 1, \ldots, N \ ; \quad V_0^N(x) = \sum_{i=1}^N \omega_0^i \ H(x - x_0^i) \ ,$$

where

$$\forall i > 1 \ , \ x_0^i := V_0^{-1}\left(\frac{i}{N}\right) \ , \ x_0^1 := V_0^{-1}\left(\frac{1}{2N}\right). \tag{64}$$

Thus, $V_0^N(\cdot)$ is a piecewise constant approximation to $V_0(\cdot)$.

Now consider N independent copies of the process (W_t) and the corresponding N copies (X_t^i) of (X_t) $(1 \le i \le N)$, with $X_0^i = x_0^i$. Define

$$\mu_t^N := \frac{1}{N} \sum_{i=1}^N H(x - X_t^i). \tag{65}$$

The distribution function of the measure μ_t^N is

$$V^N(t,x) = \frac{1}{N} \sum_{i=1}^N H(x - X_t^i).$$

Proposition 4.1 *Suppose (H6)-(H7) and*

(H8) *There exist strictly positive constants C_1, C_2 such that, for any x in $I\!R$, $|V_0(x)| \le C_1 e^{-C_2 x^2}$.*

Then there exists an increasing function $K(\cdot)$ such that, for all $N \in \mathbb{N}^*$,

$$\mathbb{E} \parallel V(t,\cdot) - V^N(t,\cdot) \parallel_{L^1(\mathbb{R})} \leq \frac{K(t)}{\sqrt{N}}. \tag{66}$$

Sketch of the proof. We have:

$$\begin{aligned} V(t,x) - V^N(t,x) &= V(t,x) - \mathbb{E}V^N(t,x) + \mathbb{E}V^N(t,x) - V^N(t,x) \\ &= \mathbb{P}_{\mu_0}(X_t \leq x) - \frac{1}{N}\sum_{i=1}^N \mathbb{P}(X_t^i \leq x) \\ &\quad + \frac{1}{N}\sum_{i=1}^N \left\{ \mathbb{E}H(x - X_t^i) - H(x - X_t^i) \right\} \\ &=: A(x) + B(x). \end{aligned}$$

An integration by parts for Stieljes integrals leads to

$$\begin{aligned} A(x) &= \int \mathbb{P}_y(X_t \leq x) dV_0(y) - \int \mathbb{P}_y(X_t \leq x) dV_0^N(y) \\ &= \int (V_0(y) - V_0^N(y)) \frac{d}{dy} \mathbb{P}_y(X_t \leq x) dy. \end{aligned}$$

Therefore

$$\parallel A(\cdot) \parallel_{L^1(\mathbb{R})} \leq \int \int \left| \frac{d}{dy} \mathbb{P}_y(X_t \leq x) \right| dx |V_0(y) - V_0^N(y)| dy.$$

The function $y \to X_t(y)$ is a.s. increasing since its derivative is an exponential (see Kunita [24, Ch.2] e.g., for the diffeomorphism property of stochastic flows associated with stochastic differential equations). Thus, again denoting by $X_t(\cdot)$ the flow defined by (1),

$$P(X_t(y) \leq x) = P(y \leq X_t^{-1}(x)).$$

X_t^{-1} is the solution to a stochastic differential equation. The coefficients of this SDE are such that the above mentioned result of Friedman [15] applies: for some new function $K(\cdot)$,

$$\left| \frac{d}{dy} \mathbb{P}_x(X_t^{-1} \leq y) \right| \leq \frac{K(t)}{\sqrt{t}} \exp\left(-\frac{(x-y)^2}{2\lambda t} \right).$$

It comes:

$$\parallel A(\cdot) \parallel_{L^1(\mathbb{R})} \leq K(t) \parallel V_0 - V_0^N \parallel_{L^1(\mathbb{R})} \leq \frac{K(t)\sqrt{\log(N)}}{N}.$$

Now consider $B(x)$.

$$\mathbb{E} \parallel B(\cdot) \parallel_{L^1(\mathbb{R})} = \frac{1}{N} \int \mathbb{E} \left| \sum_{i=1}^{N} (H(x - X_t^i) - \mathbb{E}H(x - X_t^i)) \right| dx.$$

The random variables $(H(x - X_t^i) - \mathbb{E}H(x - X_t^i))_{1 \leq i \leq N}$ have mean 0 and are independent. Thus, by the Cauchy-Schwarz inequality,

$$\begin{aligned} \mathbb{E} \parallel B(\cdot) \parallel_{L^1(\mathbb{R})} &\leq \frac{1}{N} \int \sqrt{\sum_{i=1}^{N} \mathbb{E}(H(x - X_t^i) - \mathbb{E}H(x - X_t^i))^2 dx} \\ &= \frac{1}{N} \int \sqrt{\sum_{i=1}^{N} \mathbb{P}(X_t^i \leq x)\mathbb{P}(X_t^i \geq x) dx} \\ &\leq \frac{K(t)}{N} \int \sqrt{\sum_{i=1}^{N} \int_{\frac{x - x_0^i}{\sqrt{\lambda t}}}^{+\infty} \exp\left\{-\frac{y^2}{2}\right\} dy} \quad dx. \end{aligned}$$

For fixed x the function

$$\xi \overset{\psi_x}{\longmapsto} \frac{1}{\sqrt{2\pi}} \int_{\frac{x - V_0^{-1}(\xi)}{\sqrt{\lambda t}}}^{+\infty} \exp\left\{-\frac{y^2}{2}\right\} dy$$

is decreasing from $(0,1)$ to $(0,1)$; therefore, the definition of the x_0^i implies

$$\frac{1}{2N} \sum_{i=1}^{N} \int_{\frac{x - x_0^i}{\sqrt{\lambda t}}}^{+\infty} \exp\left\{-\frac{y^2}{2}\right\} dy \leq \int_0^1 \int_{\frac{x - V_0^{-1}(s)}{\sqrt{\lambda t}}}^{+\infty} \exp\left\{-\frac{y^2}{2}\right\} dy\, ds.$$

Easy computations wich use (H8) (see [4]) then lead to the following estimate:

$$\mathbb{E} \parallel B(\cdot) \parallel_{L^1(\mathbb{R})} \leq \frac{K(t)}{\sqrt{N}}.$$

■

In practice, one cannot use exact values of X_t^i. Thus, we consider N independent processes defined by the Euler or Milshtein scheme $(\bar{X}_t^i)$ with $X_0^i = x_0^i$ and the new approximate measure

$$\bar{\mu}_t^N := \frac{1}{N} \sum_{i=1}^{N} H(x - \bar{X}_t^i). \tag{67}$$

The distribution function of the measure $\bar{\mu}_t^N$ is

$$\bar{V}^N(t,x) := \frac{1}{N}\sum_{i=1}^N H(x - \bar{X}_t^i).$$

Proposition 4.2 *Suppose (H6)-(H7) and*

(H8) *There exist strictly positive constants C_1, C_2 such that, for any x in $\mathbb{R}$, $|V_0(x)| \le C_1 e^{-C_2 x^2}$.*

Then there exists an increasing function $K(\cdot)$ such that, for all $N \in \mathbb{N}^*$,

$$\mathbb{E} \parallel V(T,\cdot) - \bar{V}^N(t,\cdot) \parallel_{L^1(\mathbb{R})} \le K(T)\left(\frac{1}{\sqrt{N}} + \frac{1}{n^\alpha}\right) \tag{68}$$

with $\alpha = \frac{1}{2}$ for the Euler scheme and $\alpha = 1$ for the Milshtein scheme.

Proof. The conclusion readily follows from Section 2 of Part I, the preceding theorem and the inequality

$$\mathbb{E}\|H(x - X_T^i) - H(x - \bar{X}_T^i)\|_{L^1(\mathbb{R})} \le \mathbb{E}|X_T^i - \bar{X}_T^i| \le \frac{K(T)}{n^\alpha}.$$

■

The convergence rate $\frac{1}{\sqrt{N}}$ is optimal. Indeed, in the following example, the error estimate is equal to $\frac{1}{\sqrt{N}}$ plus a negligible term. Let X be a random variable taking the values 0 and 1 with probability $\frac{1}{2}$. Let μ^N be the empirical distribution of N independent copies of X and let V^N be the distribution function of μ^N. It is easy to see that

$$\|V - V^N\|_{L^1(\mathbb{R})} = \frac{1}{2^N}\sum_{k=0}^N \left|\frac{1}{2} - \frac{k}{N}\right| \frac{N!}{k!(N-k)!}.$$

For example, suppose that $N = 2n$. Then,

$$\|V - V^N\|_{L^1(\mathbb{R})} = 2^{-2n}\sum_{k=0}^n \frac{(2n)!}{k!(2n-k)!} - 2^{-2n}\sum_{k=0}^n \frac{k}{n}\frac{(2n)!}{k!(2n-k)!}.$$

Now, an easy induction shows that, for all $n > 0$,

$$\sum_{k=0}^n k\frac{(2n)!}{k!(2n-k)!} = n2^{2n-1}.$$

Besides,

$$2\sum_{k=0}^{n-1}\frac{(2n)!}{k!(2n-k)!}+\frac{(2n)!}{n!n!}=2^{2n}.$$

Thus,

$$\|V-V^N\|_{L^1(\mathbb{R})}=\frac{1}{2^{2n+1}}\frac{(2n)!}{n!n!}.$$

Applying Stirling's formula, one deduces

$$\|V-V^N\|_{L^1(\mathbb{R})}=\frac{1}{\sqrt{N}}(1+o(1)).$$

5 The Chorin-Puckett method for convection-reaction-diffusion equations

Let $f(\cdot)$ be a real function such that

(H9) f is a C^2 function on $[0,1]$ such that $f(0)=f(1)=0$, $f(u)\geq 0$ for $u\in[0,1]$ (therefore, $\frac{f(u)}{u}$ is bounded in $(0,1]$ and continuous in 0).

Let $V_0(\cdot)$ be as in the preceding subsection. Consider the convection-reaction-diffusion PDE

$$\begin{cases} \dfrac{\partial u}{\partial t}=L\,u+f(u)\ , \\ \\ u(0,\cdot)=u_0(\cdot)=1-V_0(\cdot). \end{cases} \tag{69}$$

In [4], Bernard, Talay and Tubaro have analysed a stochastic particle method introduced by Puckett [40]. They have proven Puckett's conjecture, based on numerical obervations, on the convegence rate of the method. The analysis is based on an original probabilistic interpretation of the solution.

Theorem 5.1 *Under (H7)-(H9), if u_0 is of class $\mathcal{C}_b^\infty(\mathbb{R})$, we have the following representation:*

$$u(t,x)=\mathbb{E}\left[H\big(X_t-x\big)\,\exp\left(\int_0^t f'\circ u\big(s,X_s\big)\,ds\right)\right], \tag{70}$$

where (X_t) is the solution to

$$dX_t=\sigma(X_t)\,dB_t-\{b(X_t)-\sigma(X_t)\,\sigma'(X_t)\}\,dt. \tag{71}$$

Here, the law of X_0 has a density equal to $-u_0'$, and (B_t) is a standard Brownian motion.

Sketch of the proof. The function $v(t,x) := \frac{\partial u}{\partial x}(t,x)$ satifies the following equation:

$$\begin{cases} \dfrac{\partial v}{\partial t}(t,x) &= \dfrac{1}{2}\sigma^2(x)\dfrac{\partial^2 v}{\partial x^2}(t,x) + (b(x)+\sigma(x)\,\sigma'(x))\dfrac{\partial v}{\partial x}(t,x) \\ & \quad +(b'(x) + f'\circ u(t,x))\, v(t,x)\ , \\ v(0,x) &= u_0'(x). \end{cases}$$

By applying the Feynman-Kac formula, we obtain

$$v(t,x) = \mathbb{E}\left[u_0'(Y_t(x))\,\exp\left\{ \int_0^t [b'(Y_s(x)) + f'\circ u(t-s, Y_s(x))]\, ds \right\}\right], \quad (72)$$

where (Y_t) is the solution to

$$dY_t = (b(Y_t) + \sigma(Y_t)\,\sigma'(Y_t))\; dt + \sigma(Y_t)\; dB_t. \quad (73)$$

One can easily check that $u(t,x) \to 1$ as $x \to -\infty$. Thus,

$$u(t,x) = -\mathbb{E}\int_x^{+\infty} u_0'(Y_t(y))\exp\left\{ \int_0^t [b'(Y_s(y)) + f'\circ u(t-s, Y_s(y))]\, ds \right\} dy.$$

Let $\xi_{0,t}(\cdot)$ be the flow associated with the stochastic differential equation (73). Hence, we set $y = \xi_{0,t}^{-1}(z)$.

Using results of the second chapter of Kunita [24], we have, for $\theta < t$,

$$\xi_{\theta,t}^{-1}(z) = z - \int_\theta^t \sigma(\xi_{s,t}^{-1}(z))\,\hat{d}B_s - \int_\theta^t b(\xi_{s,t}^{-1}(z))\; ds,$$

where $\hat{d}B_\theta$ denotes the "backward" stochastic integral [2]. One infers that

$$\frac{\partial}{\partial z}\xi_{0,t}^{-1}(z)$$

$$= \exp\left(\int_0^t \left\{ -b'(\xi_{\theta,t}^{-1}(z)) - \frac{1}{2}\sigma'^2(\xi_{\theta,t}^{-1}(z)) \right\} d\theta - \int_0^t \sigma'(\xi_{\theta,t}^{-1}(z))\; \hat{d}B_\theta \right)$$

from which

$$-u(t,x)$$

[2]For a definition, cf. Kunita [24, end of Ch. I]

$$= \mathbb{E}\Bigg[\int_{\xi_{0,t}(x)}^{+\infty} u_0'(z) \exp\Bigg\{ \int_0^t \Big(b'(\xi_{0,s}(\alpha)) + f' \circ u(t-s, \xi_{0,s}(\alpha)) \Big) ds \Bigg|_{\alpha = \xi_{0,t}^{-1}(z)} \Bigg\}$$

$$\exp\left\{ \int_0^t \Big[- b'(\xi_{s,t}^{-1}(z)) - \frac{1}{2} \sigma'^2(\xi_{s,t}^{-1}(z)) \Big] ds - \int_0^t \sigma'(\xi_{s,t}^{-1}(z)) \hat{d}B_s \right\} dz \Bigg].$$

One now uses Kunita [24, Lemma 6.2, Ch. II]: for any continuous function $g(s,x)$ we have

$$\int_0^t g(s, \xi_{0,s}(\alpha)) \Bigg|_{\alpha = \xi_{0,t}^{-1}(z)} ds = \int_0^t g(s, \xi_{s,t}^{-1}(z)) \, ds.$$

Thus,

$$-u(t,x) = \mathbb{E}\Bigg[\int_{-\infty}^{+\infty} H(-\xi_{0,t}(x) + z) \exp\left\{ \int_0^t f' \circ u(t-s, \xi_{s,t}^{-1}(z)) ds \right\}$$

$$M_0^t(z) u_0'(z) dz \Bigg]$$

where $(M_\theta^t(z))_{\theta \le t}$ is the exponential (backward) $(\mathcal{F}_\theta^t)_{\theta \le t}$-martingale defined by

$$M_\theta^t(z) = \exp\left\{ -\frac{1}{2} \int_\theta^t \sigma'^2(\xi_{s,t}^{-1}(z)) \, ds - \int_\theta^t \sigma'(\xi_{s,t}^{-1}(z)) \hat{d}B_s \right\}.$$

The application $x \to \xi_{0,t}(x)$ is a.s. increasing (its derivative is an exponential), thus $H(-\xi_{0,t}(x) + z) = H(\xi_{0,t}^{-1}(z) - x)$.

Hence,

$-u(t,x)$

$$= \mathbb{E}\Bigg[\int_{\mathbb{R}} H(\xi_{0,t}^{-1}(z) - x) \exp\left\{ \int_0^t f' \circ u(s, \xi_{t-s,t}^{-1}(z)) ds \right\} M_0^t(z) u_0'(z) dz \Bigg].$$

We observe that the law of the process $(\xi_{t-\theta,t}^{-1})_{0 \le \theta \le t}$, on $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_0^t)$, is identical to the law of the process $(X_\theta)_{0 \le \theta \le t}$ solution to

$$dX_\theta = \sigma(X_\theta) \, dB_\theta - b(X_\theta) \, d\theta.$$

Hence, $\mathbb{E}_0$ denoting the expectation under the law $\mathbb{P}_0$ for which the initial law of the process (X_θ) has a density equal to $-u_0'(z)$, and (M_t) denoting the exponential martingale defined by

$$M_t = \exp\left\{ -\frac{1}{2} \int_0^t \sigma'^2(X_s)\, ds + \int_0^t \sigma'(X_s) dB_s \right\},$$

we have

$$u(t,x) = \mathbb{E}_0 \left[H(X_t - x) \exp\left\{ \int_0^t f' \circ u(t-s, X_s)\, ds \right\} M_t \right].$$

On $(\Omega, \mathcal{F}, \mathbb{P}_0, \mathcal{F}_0^T)$, one performs the Girsanov transformation defined by

$$\widetilde{\mathbb{P}}(A) := \mathbb{E}_0 \left[1_A\, M_T \right] \quad , \quad A \in \mathcal{F}_0^T \ ;$$

then, for $t \leq T$,

$$u(t,x) = \widetilde{\mathbb{E}} \left[H(X_t - x) \exp\left\{ \int_0^t f' \circ u(s, X_s)\, ds \right\} \right].$$

Under $\widetilde{\mathbb{P}}$, (X_t) solves

$$dX_t = \sigma(X_t)\, d\tilde{B}_t - \{ b(X_t) - \sigma(X_t)\, \sigma'(X_t) \}\, dt.$$

Here, $(\tilde{B}_\theta)$ defined by

$$\tilde{B}_\theta = B_\theta - \int_0^\theta \sigma'(X_s)\, ds,$$

is a Brownian motion under $\widetilde{\mathbb{P}}$. Obviously, the above representation of u is identical to (70). ∎

Define the initial weights and the initial approximation by

$$\omega_0^i = \frac{1}{N}, \text{for } i = 1, \ldots, N \ ; \qquad \bar{u}_0(x) = \sum_{i=1}^{N} \omega_0^i\ H(x_0^i - x)\ ,$$

where

$$\forall i < N \quad : \quad x_0^i = u_0^{-1}\left(1 - \frac{i}{N}\right) \quad , \quad x_0^N = u_0^{-1}\left(\frac{1}{2N}\right). \tag{74}$$

Let X^n be defined by the Milshtein scheme (9). From now on, we write $\bar{X}$ instead of X^n. We set

$$\begin{aligned}\bar{X}^i_{(p+1)T/n} &= \bar{X}^i_{pT/n} - \big(b(\bar{X}^i_{pT/n}) - \sigma(\bar{X}^i_{pT/n})\sigma'(\bar{X}^i_{pT/n})\big)\frac{T}{n} \\ &\quad + \sigma(\bar{X}^i_{pT/n})(B^i_{(p+1)T/n} - B^i_{pT/n}) \\ &\quad + \frac{1}{2}\sigma(\bar{X}^i_{pT/n})\sigma'(\bar{X}^i_{pT/n})\left((B^i_{(p+1)T/n} - B^i_{pT/n})^2 - \frac{T}{n}\right).\end{aligned} \tag{75}$$

Let $\pi_k(i)$ denotes the label number of the particle located immediately at the right side of the particle of label i at the time kT/n. We define

$$\begin{aligned}&\omega^i_{pT/n} = \\ &\quad \omega^i_{(p-1)T/n} \\ &\left(1 + \frac{T}{n}\frac{f\circ\bar{u}((p-1)T/n, \bar{X}^i_{(p-1)T/n}) - f\circ\bar{u}((p-1)T/n, \bar{X}^{\pi_{p-1}(i)}_{(p-1)T/n})}{\omega^i_{(p-1)T/n}}\right)\end{aligned} \tag{76}$$

and

$$\bar{u}(pT/n, x) = \sum_{i=1}^{N} \omega^i_{pT/n}\, H(\bar{X}^i_{pT/n} - x) \tag{77}$$

for $p = 1, \ldots, n$.

Theorem 5.2 (Bernard, Talay and Tubaro [4]) (i) *Under (H7)-(H9), there exists an increasing function $K(\cdot)$ and an integer n_0 such that, for any $n > n_0$ and any $N \geq 1$,*

$$\|u(T,\cdot) - \bar{u}(T,\cdot)\|_{L^1(\mathbb{R}\times\Omega)} \leq K(T)\left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{n}}\right).$$

(ii) *When the functions $b(\cdot)$ and $\sigma(\cdot)$ are constant, then the rate of convergence is given by*

$$\|u(T,\cdot) - \bar{u}(T,\cdot)\|_{L^1(\mathbb{R}\times\Omega)} \leq K(T)\left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{n}}\right).$$

The same estimates hold for the standard deviation of $\|u(T,\cdot) - \bar{u}(T,\cdot)\|_{L^1(\mathbb{R})}$.

Sketch of the proof. The lengthy proof consists in observing that the algorithm is a discretization of the representation (70).

Indeed, the approximation of $-u_0'(z)\,dz$ by

$$\sum_{i=1}^{N} \omega_0^i \, \delta_{x_0^i}$$

leads to

$$u(T,x) \simeq \sum_{i=1}^{N} \omega_0^i \; I\!E \left[H(X_T(x_0^i) - x) \; \exp\left\{ \int_0^T f' \circ u(s, X_s(x_0^i)) \; ds \right\} \right].$$

Let $\{(B_\theta^i), i = 1, \dots, N\}$ be N independent Brownian motions and let (X_θ^i) be the (independent) solutions to the following SDE's (in forward time):

$$\begin{cases} dX_\theta^i = \sigma(X_\theta^i)\, dB_\theta^i - \{b(X_\theta^i) - \sigma(X_\theta^i)\, \sigma'(X_\theta^i)\}\; d\theta, \\ \\ X_0^i = x_0^i. \end{cases}$$

One has

$$u(T,x) \simeq \sum_{i=1}^{N} \omega_0^i \; I\!E \left[H(X_T^i - x) \; \exp\left\{ \int_0^T f' \circ u(s, X_s^i) \; ds \right\} \right].$$

The particle algorithm replaces the expectation by a point estimation:

$$u(T,x) \simeq \sum_{i=1}^{N} \omega_0^i \; H(X_T^i - x) \; \exp\left\{ \int_0^T f' \circ u(s, X_s^i) \; ds \right\}.$$

Then, one approximates $\exp\left\{ \int_0^T f' \circ u(s, X_s^i) \; ds \right\}$. The integral is discretized with a step T/n and the Milshtein scheme is used to approximate the $X_{pT/n}^i$'s. Besides, the unknown function $u(pT/n, \cdot)$ is replaced by its approximation $\bar{u}(pT/n, \cdot)$. It is this substitution which introduces a dependency in the algorithm, because the computation of $\bar{u}(pT/n, \cdot)$ requires to sort the positions of the particles at each step of the algorithm (see the role of the functions $\pi_k(\cdot)$ in (76)). Without this substitution, the weights would be recursively defined by

$$\bar{\rho}_{(p+1)T/n}^i = \bar{\rho}_{pT/n}^i + \frac{T}{n} f \circ u(pT/n, \bar{X}_{pT/n}^i).$$

The following key estimate shows that the true weights are not far from being independent, which explains that the global error of the algorithm is of order $N^{-1/2}$ as if the weights were independent. Set $\alpha_p^i := \mathbb{E}|\omega_p^i - \rho_p^i|^2$, and $\alpha_p := \sup_i \alpha_p^i$. One can show (the proof is very technical) :

$$\forall p = 1, \ldots, n \ , \ \alpha_p \leq \frac{C}{nN^2} + \frac{C}{N^3}. \tag{78}$$

Then, one must carefully estimate the error produced by each one of the successive approximations that have just been described. In particular, the difficulty is to avoid the summation over p of the "statistical error" involved in the algorithm which identifies $\bar{u}(pT/n, \cdot)$ and $\mathbb{E}\bar{u}(pT/n, \cdot)$, because such a summation would lead to an estimate on the global error of order $\frac{n}{\sqrt{N}}$. In fact, a more clever analysis shows that the algorithm propagates the error

$$u(pT/n, \cdot) - \mathbb{E}\bar{u}(pT/n, \cdot)$$

in a rather complex way whereas the "statistical error" can be taken into account only at time T; this latter error can be controlled owing to the estimate (78). ∎

6 One-dimensional Mc-Kean Vlasov equations

Consider two Lipschitz kernels $b(x,y)$, $s(x,y)$ from $\mathbb{R}^2$ to $\mathbb{R}$, a probability distribution function V_0 and the nonlinear problem

$$\begin{cases} \frac{\partial V}{\partial t}(t,x) = \frac{1}{2}\frac{\partial}{\partial x}\left[\left(\int_{\mathbb{R}} s(x,y) \frac{\partial V}{\partial x}(t,y)dy\right)^2 \frac{\partial V}{\partial x}(t,x)\right] \\ \qquad\qquad - \left[\int_{\mathbb{R}} b(\cdot,y) \frac{\partial V}{\partial x}(t,y)dy \right] \frac{\partial V}{\partial x}(t,x) \ , \\ V(0,x) = V_0(x) \ , \end{cases} \tag{79}$$

Later on, we will see that Burgers equation can be interpreted as a special case of this family of problems.

Our objective is to develop an algorithm of simulation of a discrete time particle system $\{Y_{kT/n}^i, i = 1, \ldots, N\}$ such that the empirical distribution

$$\bar{V}_{kT/n}(x) := \frac{1}{N}\sum_{i=1}^{N} H(x - Y_{kT/n}^i)$$

approximates the solution $V(t,x)$ of (79). Contrarily to the stochastic particle method of the previous subsection, the weights are constant but, in counterpart, the positions of the particles are given by dependent stochastic processes.

Consider the system of weakly interacting particles described by

$$\left\{ \begin{array}{l} dX_t^{i,N} = \int_{I\!R} b(X_t^{i,N}, y)\mu_t^N(dy)\, dt \; + \; \int_{I\!R} s(X_t^{i,N}, y)\mu_t^N(dy) dW_t^i \;, \\ \\ X_0^{i,N} = X_0^i \;, i = 1, \ldots, N \;, \end{array} \right. \tag{80}$$

where $(W_t^1), \ldots, (W_t^N)$ are independent one-dimensional Brownian motions and μ_t^N is the random empirical measure

$$\mu_t^N \; = \; \frac{1}{N} \sum_{i=1}^{N} \; \delta_{X_t^{i,N}} .$$

The functions b and s are the "interaction kernels". When the initial distribution of the particles is symmetric and when the kernels are Lipschitz, one has the propagation of chaos property: the sequence of random probability measures (μ^N) on the space of trajectories defined by

$$\mu^N \; = \; \frac{1}{N} \sum_{i=1}^{N} \; \delta_{X_\cdot^{i,N}}$$

converges in law as N goes to infinity to a deterministic probability measure μ. Besides, if for each t we denote by μ_t the one-dimensional distribution of μ (μ_t is the limit in law of μ_t^N), then there exists a unique strong solution (X_t) to the nonlinear stochastic differential equation

$$\left\{ \begin{array}{l} X_t \; = \; X_0 \; + \; \int_0^t \int_{I\!R} b(X_\theta, y)\mu_\theta(dy) dt \; + \; \int_0^t \int_{I\!R} s(X_\theta, y)\mu_\theta(dy) dW_\theta \;, \\ \\ \mu_t \; \text{ is the law of the random variable } X_t, \text{ for all } t \geq 0 \end{array} \right. \tag{81}$$

(see S. Méléard's contribution to this volume or Sznitman [46] e.g.). One consequence of the propagation of chaos is that the law of one particle, for example the law of $(X_t^{1,N})$, tends to the law of the process (X_t) when N goes to infinity.

Defining the differential operator $L(\mu)$ by

$$L(\mu) f(x) = \frac{1}{2} \left(\int_{I\!R} s(x,y) d\mu(y) \right)^2 f''(x) \; + \; \left(\int_{I\!R} b(x,y) d\mu(y) \right) f'(x) \;,$$

Itô's formula shows that μ_t is the solution to the McKean-Vlasov equation

$$\begin{cases} \frac{d}{dt} < \mu_t, f > = < \mu_t, L(\mu_t) f > \ , \ \forall f \in C_K^\infty(\mathbb{R}) \ , \\ \mu_{t=0} = \mu_0. \end{cases} \tag{82}$$

Consequently, the distribution function $V(t,x)$ of μ_t solves (79) where $V_0(\cdot)$ is the distribution function of μ_0.

We suppose that the following assumptions hold:

(H10) There exists a strictly positive constant s_* such that

$$s(x,y) \geq s_* > 0 \ , \ \forall (x,y).$$

(H11) The kernels $b(\cdot,\cdot)$ and $s(\cdot,\cdot)$ are uniformly bounded functions of $\mathbb{R}^2$; $b(\cdot,\cdot)$ is globally Lipschitz and $s(\cdot,\cdot)$ has uniformly bounded first partial derivatives.

(H12) The initial law μ_0 has a continuous density $u_0(\cdot)$ satisfying: there exist constants $M > 0$, $\eta \geq 0$ and $\alpha > 0$ such that

$$u_0(x) \ \leq \ \eta \exp(-\alpha \frac{x^2}{2}) \text{ for } \ | x | > \ M.$$

The initial distribution function $V(0,\cdot) = V_0(\cdot)$ is approximated as in the preceding subsection. We set

$$y_0^i := V_0^{-1}(i/N).$$

Consider the system (80) with the initial condition $X_0^{i,N} = y_0^i$, and denote its solution by $(X_t^i \ , \ 1 \leq i \leq N)$. There holds

$$\begin{cases} dX_t^i = \frac{1}{N} \sum_{j=1}^N b\left(X_t^i, X_t^j\right) \ dt \ + \ \frac{1}{N} \sum_{j=1}^N s\left(X_t^i, X_t^j\right) \ dw_t^i \ , \ t \in [0,T] \ , \\ \\ X_0^i = y_0^i \ , \ i = 1, \ldots, N. \end{cases}$$

To get a simulation procedure of a trajectory of each (X_t^i), we discretize in time and we approximate μ_t by the empirical measure of the simulated particles. The Euler scheme then leads to

$$\begin{cases} Y^i_{(k+1)T/n} & = & Y^i_{kT/n} + \frac{1}{N} \sum_{j=1}^N b(Y^i_{kT/n} \ , Y^j_{kT/n}) \frac{T}{n} \\ & & + \frac{1}{N} \sum_{j=1}^N s(Y^i_{kT/n}, Y^j_{kT/n}) \left(W^i_{(k+1)T/n} - W^i_{kT/n}\right) \ , \\ \\ Y_0^i & = & y_0^i \ , \ i = 1, \ldots, N. \end{cases} \tag{83}$$

In the same way, we approximate $V(t,\cdot)$ solution to (79) by the cumulative distribution function of μ_t:

$$\overline{V}_{kT/n}(x) \;:=\; \frac{1}{N}\sum_{i=1}^{N} H(x - Y^i_{kT/n})\;,\;\forall\, x\;\in\; I\!\!R. \tag{84}$$

Theorem 6.1 (Bossy and Talay [7], Bossy [5]) *Suppose (H10)-(H12). Let $V(t,x)$ be the solution of the PDE (79).*

There exists an increasing function $K(\cdot)$ such that, $\forall\, k \in \{1,\ldots,n\}$:

$$\begin{aligned} &I\!\!E\;\left\|V(kT/n,.) - \overline{V}_{kT/n}(.)\right\|_{L^1(I\!\!R)} \\ &\quad\le\; K(T)\;\left(\|V_0\;-\;\overline{V}_0\|_{L^1(I\!\!R)}\;+\;\frac{1}{\sqrt{N}}\;+\;\frac{1}{\sqrt{n}}\right) \end{aligned} \tag{85}$$

and

$$\begin{aligned} &Var\left(\left\|V(kT/n,.) - \overline{V}_{kT/n}(.)\right\|_{L^1(I\!\!R)}\right) \\ &\quad\le\; K(T)\;\left(\|V_0\;-\;\overline{V}_0\|^2_{L^1(I\!\!R)}\;+\;\frac{1}{N}\;+\;\frac{1}{n}\right). \end{aligned} \tag{86}$$

Besides,

$$\left\|V_0\;-\;\overline{V}_0\right\|_{L^1(I\!\!R)}\;\le\;\frac{C\sqrt{\log(N)}}{N}.$$

Sketch of the proof. Define $\beta\;:\;[0,T]\times I\!\!R\longrightarrow I\!\!R$ by

$$\beta(t,x) := \int_{I\!\!R} b(x,y)\;\mu_t(dy)\;,$$

and $\sigma\;:\;[0,T]\times I\!\!R\longrightarrow I\!\!R$ by

$$\sigma(t,x) := \int_{I\!\!R} s(x,y)\;\mu_t(dy).$$

Under our hypotheses, there exists a unique strong solution to

$$\begin{cases} dz_t \;=\; \beta(t,z_t)dt\;+\;\sigma(t,z_t)\;dw_t\;, \\[2ex] z_{t=0}\;=\;z_0\;, \end{cases} \tag{87}$$

where z_0 is a square integrable random variable. When the law of z_0 is μ_0, the two processes (z_t) and (X_t) solution to (81) have the same law and

$$V(t,x)\;=\;I\!\!EH(x-X_t)\;=\;I\!\!E_{\mu_0}H(x-z_t)\;=\;\int_{I\!\!R}\;I\!\!EH(x-z_t(y))\;\mu_0(dy).$$

Consider the independent processes $(z_t^i)_{(i=1,..,N)}$ solutions to

$$\begin{cases} dz_t^i \;=\; \beta(t, z_t^i)\; dt \;+\; \sigma(t, z_t^i)\; dW_t^i\;, \\ \\ z_0^i \;=\; y_0^i. \end{cases} \tag{88}$$

Applying the Euler scheme to (88), one defines the independent discrete-time processes $(\overline{z}_{kT/n}^i)$:

$$\begin{cases} \overline{z}_{(k+1)T/n}^i = \overline{z}_{kT/n}^i + \beta(kT/n, \overline{z}_{kT/n}^i)\,\frac{T}{n} \\ \qquad\qquad\qquad +\sigma(kT/n, \overline{z}_{kT/n}^i)\,\left(W_{(k+1)/n}^i - W_{kT/n}^i\right)\;, \\ \\ \overline{z}_0^i = y_0^i. \end{cases} \tag{89}$$

The global error is decomposed as follows:

$$\begin{aligned} &\mathbb{E}\left\|V(kT/n,x)\;-\;\overline{V}_{kT/n}(x)\right\|_{L^1(\mathbb{R})} \\ &\quad\leq \left\|\mathbb{E}_{\mu_0}H(x-z_{kT/n})\;-\;\mathbb{E}_{\overline{\mu}_0}H(x-z_{kT/n})\right\|_{L^1(\mathbb{R})} \\ &\quad+\mathbb{E}\left\|\mathbb{E}_{\overline{\mu}_0}H(x-z_{kT/n})\;-\;\frac{1}{N}\sum_{i=1}^{N}H(x-z_{kT/n}^i)\right\|_{L^1(\mathbb{R})} \\ &\quad+\mathbb{E}\left\|\frac{1}{N}\sum_{i=1}^{N}H(x-z_{kT/n}^i)\;-\;\frac{1}{N}\sum_{i=1}^{N}H(x-\overline{z}_{kT/n}^i)\right\|_{L^1(\mathbb{R})} \\ &\quad+\mathbb{E}\left\|\frac{1}{N}\sum_{i=1}^{N}H(x-\overline{z}_{kT/n}^i)-\;\frac{1}{N}\;\sum_{i=1}^{N}H(x-Y_{kT/n}^i)\right\|_{L^1(\mathbb{R})}. \end{aligned} \tag{90}$$

The first term of the right handside corresponds to the approximation error of the measure μ_0; the second term essentially is a statistical error related to the Strong Law of Large Numbers; the third term is the discretization error induced by the Euler scheme; the last term corresponds to the approximation of the coefficients $\beta(kT/n,\cdot)$ and $\sigma(kT/n,\cdot)$ by means of the empirical measure $\overline{\mu}_{kT/n}$, which introduces the family of dependent processes $(Y_{kT/n}^i)$.

One successively proves:

$$\left\| \mathbb{E}_{\mu_0} H(x-z_t) \; - \; \mathbb{E}_{\overline{\mu}_0} H(x-z_t) \right\|_{L^1(\mathbb{R})} \; \leq \; C \, \| \, V_0 \; - \; \overline{V}_0 \, \|_{L^1(\mathbb{R})} \; ,$$

$$\mathbb{E} \left\| \mathbb{E}_{\overline{\mu}_0} H(x-z_t) \; - \; \frac{1}{N} \sum_{i=1}^{N} \; H(x-z_t^i) \right\|_{L^1(\mathbb{R})} \; \leq \; \frac{C}{\sqrt{N}} \; ,$$

$$\mathbb{E} \left\| \frac{1}{N} \sum_{i=1}^{N} H(x-z_{kT/n}^i) \; - \; \frac{1}{N} \sum_{i=1}^{N} H(x-\bar{z}_{kT/n}^i) \right\|_{L^1(\mathbb{R})} \; \leq \; \frac{C}{\sqrt{n}} \; ,$$

$$\mathbb{E} \left\| \frac{1}{N} \sum_{i=1}^{N} H(x-\bar{z}_{kT/n}^i) - \frac{1}{N} \sum_{i=1}^{N} H(x-Y_{kT/n}^i) \right\|_{L^1(\mathbb{R})} \leq C \left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{N}} + \|V_0 - \bar{V}_0\|_{L^1(\mathbb{R})} \right).$$

The three first inequalities are obtained following the guidelines presented in Section 4 and observing that an inequality of the type (63) holds for the density of the law of $z_t(x)$. The proof of the last inequality is based upon an induction formula with respect to k which mimics the propagation of the global error. More precisely, set

$$E_k := \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} |\bar{z}_{t_k}^i - Y_{t_k}^i|^2$$

and

$$\delta := \|V_0 - \bar{V}_0\|_{L^1(\mathbb{R})}^2 \; + \; \frac{1}{N} \; + \; \frac{T}{n}.$$

A tedious computation, where the Lipschitz condition on the kernels and the estimate (4) play a role, shows that

$$\begin{cases} E_k \leq (1 + \frac{CT}{n}) E_{k-1} + \frac{CT}{n} \left(\delta \; + \; \frac{T}{n}\right) \; + \; \frac{CT}{n} \frac{\sqrt{E_{k-1}}}{\sqrt{t_{k-1}}} \sqrt{\delta} \quad \text{for } k > 1, \\ E_1 \leq \frac{CT}{n}, \end{cases}$$

from which one can deduce

$$\mathbb{E} \left\| \frac{1}{N} \sum_{i=1}^{N} H(x-\bar{z}_{kT/n}^i) - \frac{1}{N} \sum_{i=1}^{N} H(x-Y_{kT/n}^i) \right\|_{L^1(\mathbb{R})} \leq C \left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{N}} + \|V_0 - \bar{V}_0\|_{L^1(\mathbb{R})} \right).$$

■

Suppose now that the objective is to approximate the solution of the equation (82) rather than (79).

The above hypotheses imply that for all $t > 0$ the measure μ_t has a density $u(t,\cdot)$ w.r.t. Lebesgue's measure. To obtain an approximation of $u(kT/n,\cdot)$, we construct a regularization by convolution of the discrete measure $\overline{\mu}_{kT/n}$.

Let $\Phi_\varepsilon(\cdot)$ be the density of the Gaussian law $N(0,\varepsilon^2)$ and set

$$\overline{u}^{\varepsilon}_{kT/n}(x) := \left(\Phi_\varepsilon * \overline{\mu}_{kT/n}\right)(x) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\varepsilon} \exp\left(-\frac{(x - Y^i_{kT/n})^2}{2\varepsilon^2}\right).$$

We strengthen our hypotheses:

(H11') The kernel $b(\cdot,\cdot)$ is in $C_b^2(\mathbb{R}^2)$ and $s(\cdot,\cdot)$ is in $C_b^3(\mathbb{R}^2)$.

(H12') The initial law μ_0 has a strictly positive density $u_0(\cdot)$ in $C^2(\mathbb{R})$ satisfying: there exist strictly positive constants M, η and α such that

$$u_0(x) + |u_0'(x)| + |u_0''(x)| \le \eta \exp\left(-\alpha\frac{x^2}{2}\right), \text{ for } |x| > M.$$

One then have the

Theorem 6.2 (Bossy and Talay [7], Bossy [5]) *Suppose (H10), (H11') and (H12'). Let $u(t,\cdot)$ be the classical solution to the PDE*

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[u(t,x)\ \left(\int_{\mathbb{R}} s(x,y)u(t,y)dy\right)^2\right] \\ \qquad\qquad - \frac{\partial}{\partial x}\left[u(t,x)\ \int_{\mathbb{R}} b(x,y)u(t,y)dy\right], \\ u(0,x) = u_0(x), \end{cases} \tag{91}$$

where $u_0(\cdot)$ is the density of μ_0.

Then there exists an increasing function $K(\cdot)$ such that, $\forall k \in \{1,\dots,n\}$,

$$\mathbb{E}\ \left\|u(kT/n,.) - \overline{u}^{\varepsilon}_{kT/n}(.)\right\|_{L^1(\mathbb{R})} \le K(T)\left[\varepsilon^2 + \frac{1}{\varepsilon}\left(\|V_0 - \overline{V}_0\|_{L^1(\mathbb{R})} + \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{n}}\right)\right] \tag{92}$$

and

$$Var\left(\left\|u(kT/n,.) - \overline{u}^{\varepsilon}_{kT/n}(.)\right\|_{L^1(\mathbb{R})}\right) \le K(T)\left[\varepsilon^4 + \frac{1}{\varepsilon^2}\left(\|V_0 - \overline{V}_0\|^2_{L^1(\mathbb{R})} + \frac{1}{N} + \frac{1}{n}\right)\right]. \tag{93}$$

Sketch of the proof. The first step consists in proving that the density $u(t,\cdot)$ belongs to the Sobolev space $W^{2,1}(I\!\!R)$ and that the norm of $u(t,\cdot)$ in $W^{2,1}(I\!\!R)$ is bounded uniformly in $t \in [0,T]$. This is done by using a criterion due to Cannarsa and Vespri [9] to check that the function $(1+x^2)u(t,x)$ belongs to $\mathcal{C}^1([0,T];L^2(I\!\!R)) \bigcap \mathcal{C}([0,T];W^{2,2}(I\!\!R))$. Equipped with this result, one can then use the well-known estimate (cf. Raviart [41])

$$\|u(t_k,\cdot) - (u(t_k,\cdot) * \Phi_\varepsilon)\|_{L^1(I\!\!R)} \leq C\,\varepsilon^2 \,\|u(t_k,\cdot)\|_{W^{2,1}(I\!\!R)}. \tag{94}$$

The second step is easy. It consists in checking that

$$I\!\!E\|\,(u(t_k,\cdot) - \overline{u}_{t_k}(\cdot)) * \Phi_\varepsilon\|_{L^1(I\!\!R)} \leq \frac{C}{\varepsilon}\, I\!\!E\,\|V(t_k,.) - \overline{V}_{t_k}(.)\|_{L^1(I\!\!R)}.$$

Therefore, one can conclude by applying Theorem 6.1. ■

Thus, the rate of convergence depends on relations between ε, N and n. This is not estonishing: roughly speaking, if ε is too large, the smoothing by $\Phi_\varepsilon(\cdot)$ is too crude whereas, when ε is too small w.r.t. N, there may be too few particles in the windows of size ε.

7 The Burgers equation

For all the results of this section we refer to Bossy and Talay [6] and Bossy [5].

Consider the Burgers equation:

$$\begin{cases} \dfrac{\partial V}{\partial t} = \dfrac{1}{2}\sigma^2 \dfrac{\partial^2 V}{\partial x^2} - V\,\dfrac{\partial V}{\partial x} \;, \text{ in } [0,T]\times I\!\!R\,, \\[2ex] V(0,x) = V_0(x)\;. \end{cases} \tag{95}$$

This PDE can be seen as the Fokker-Planck equation for the limit law of particle systems corresponding to a kernel $b(\cdot,\cdot)$, roughly speaking, equal to a Dirac measure (see Sznitman [46]). The corresponding algorithm must involve a smoothing of this kernel. The analysis of its convergence rate is still in progress. Another stochastic particle method for the Burgers equation has been proposed by Roberts [42].

In order to construct a stochastic particle algorithm involving a kernel $b(\cdot,\cdot)$ less irregular than a Dirac measure (therefore more interesting from a numerical point of view), we interpret the solution of the Burgers equation as the distribution function of the probability measure U_t solution to the following McKean-Vlasov PDE:

$$\begin{cases} \dfrac{\partial U}{\partial t} \;=\; \dfrac{1}{2}\sigma^2 \dfrac{\partial^2 U}{\partial x^2} \;-\; \dfrac{\partial}{\partial x}\left(\left(\displaystyle\int_{I\!R} H(x-y)U_t(dy)\right)U_t\right) \;, \\ \\ U_{t=0} \;=\; U_0 \;. \end{cases} \tag{96}$$

The above PDE is understood in the distribution sense. Its nonlinear part makes appear the discontinuous interaction kernel $b(x, y) = H(x - y)$.

To this McKean-Vlasov equation, is associated the nonlinear stochastic differential equation

$$\begin{cases} dX_t = \sigma dW_t + \displaystyle\int_{I\!R} H(X_t - y)Q_t(dy)\, dt \quad , \; Q_t(dy) \text{ is the law of } X_t \;, \\ \\ X_{t=0} = X_0 \; \text{ of law } Q_0 \;. \end{cases} \tag{97}$$

As the kernel $H(x - y)$ is discontinuous, the “classical” results of the propagation of chaos for weakly interacting particles do not apply. Thus, one first must prove that there exists a solution to (97) and that the propagation of chaos holds for the corresponding particles system.

Let $\mathcal{M}(I\!R)$ denote the set of probability measures on $I\!R$. For any measure $\mu \in \mathcal{M}(I\!R)$ the differential operator $\mathcal{L}_{(\mu)}$ is defined by

$$\mathcal{L}_{(\mu)} f(x) \;=\; \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2}(x) \;+\; \left(\int_{I\!R} H(x-y)\mu(dy)\right)\frac{\partial f}{\partial x}(x) \;.$$

One can prove the existence and the uniqueness of a solution to the following nonlinear martingale problem (98) associated to the operator $\mathcal{L}_{(.)}$: for any initial distribution $Q_0 \in \mathcal{M}(I\!R)$, there exists a unique $\mathcal{Q}$ in $\mathcal{M}(C([0, T]; I\!R))$ (we denote by $\mathcal{Q}_t$, $t \in [0, T]$, its onedimensional distributions) such that

$$\left.\begin{array}{ll} (i) & \mathcal{Q}_0 \;=\; U_0 \;, \\ (ii) & \forall f \in C^2_K(I\!R), f(x(t)) - f(x(0)) - \displaystyle\int_0^t \mathcal{L}_{(\mathcal{Q}_s)} f(x(s))ds, \; t \in [0, T] \\ & \text{is a } \mathcal{Q} \text{ martingale} \;, \end{array}\right\} \tag{98}$$

where $x(\cdot)$ denotes the canonical process on the space of continuous functions from $[0, T]$ to $I\!R$ (this is done by showing the convergence of the solutions of the martingale problems corresponding to an appropriate sequence of smoothened Heaviside functions). Equipped with this result, one can prove that there is a unique solution $\mathbb{Q}$ in the sense of probability law to (97). (Besides, one can show that the distribution function of $\mathbb{Q}_t$ is the *classical* solution to the Burgers equation.)

One can also prove the following

Proposition 7.1 *The propagation of chaos holds for the sequence of measures* (μ^N) *defined by*

$$\mu^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^{i,N}_\cdot}$$

where

$$dX_t^{i,N} = \sigma dW_t^i + \frac{1}{N} \sum_{j=1}^{N} H(X_t^{i,N} - X_t^{j,N})\, dt.$$

Sketch of the proof. First, one easily shows that the sequence of the laws of the μ^N's is tight. Then, let Π_1^∞ be a limit point of a convergent subsequence of $\{\mathcal{L}aw(\mu^N)\}$. Similarly to what is done in Section 4.2 of S. Méléard's contribution in this volume, set

$$F(m) :=< m, \left(f(x(t)) - f(x(s)) - \int_s^t \mathcal{L}_{(m_\theta)} f(x(\theta)) d\theta \right) g(x(s_1), \ldots, x(s_k)) >$$

where $f \in C_b^2(\mathbb{R})$, $g \in C_b(\mathbb{R}^k)$, $0 < s_1 < \ldots < s_k \leq s \leq T$ and m is a probability on $C([0,T];\mathbb{R})$. Then use the two following arguments.

(a) First, $\lim_{N\to+\infty} \mathbb{E}[F(\mu^N)]^2 = 0$ since

$$\begin{aligned}\lim_{N\to+\infty} \mathbb{E}[F(\mu^N)]^2 &\leq \lim_{N\to+\infty} \frac{C}{N^2} \sum_{i=1}^{N} \mathbb{E} \left(\int_s^t \sigma dW_\theta^i \right)^2 \\ &= 0.\end{aligned}$$

(b) Second, one can show that the support of Π_1^∞ is the set of solutions to the nonlinear martingale problem (98). As the uniqueness of such a solution holds, one gets that $\Pi_1^\infty = \delta_{\mathbb{Q}}$. Here, one cannot use the continuity of $F(\cdot)$ in $\mathcal{P}(\mathcal{C}([0,T];\mathbb{R}))$ endowed with the Vaserstein metric because the Heaviside function is discontinuous, but one can take advantage of the explicit form of F. The key argument is as follows. Let ν^N be defined by

$$\nu^N := \frac{1}{N^4} \sum_{i,j,k,l=1}^{N} \delta_{(X^{i,N}_\cdot, X^{j,N}_\cdot, X^{k,N}_\cdot, X^{l,N}_\cdot)} \ .$$

Let $\Pi^\infty \in \mathcal{P}(\mathcal{P}(\mathcal{C}([0,T];\mathbb{R})^4))$ be the limit of a convergent subsequence of the tight family $\{Law(\nu^N)\}$. Denote by ν^1 the first marginal of a measure $\nu \in \mathcal{P}(\mathcal{C}([0,T];\mathbb{R})^4)$ (for all Borel sets A in $\mathcal{C}([0,T];\mathbb{R})$, $\nu^1(A) = \nu(A \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{C}([0,T];\mathbb{R})))$. Then, one observes that

$$\Pi^\infty - \text{a.e},\ \nu = \nu^1 \otimes \nu^1 \otimes \nu^1 \otimes \nu^1.$$

Besides, one can prove that

$$\begin{aligned}
&\lim_{N\to+\infty} \mathbb{E}[F(\mu^N)]^2 = \\
&\int_{\mathcal{P}(\mathcal{C}([0,T];\mathbb{R})^4)} \left\{ \int_{\mathcal{C}([0,T];\mathbb{R})^4} \left[f(x_t^1) - f(x_s^1) - \frac{\sigma^2}{2}\int_s^t f''(x_\theta^1)d\theta \right.\right. \\
&\qquad\qquad \left. - \int_s^t H(x_\theta^1 - x_\theta^2) f'(x_\theta^1) d\theta \right] \qquad (99) \\
&\qquad\qquad \left. g(x_{s_1}^1, \ldots, x_{s_p}^1) d\nu(x^1, x^2, x^3, x^4) \right\}^2 d\Pi^\infty(\nu)\ .
\end{aligned}$$

One then proves that Π^∞-a.e.,

$$\begin{aligned}
&\int_{\mathcal{C}([0,T];\mathbb{R})^2)} \left[f(x_t^1) - f(x_s^1) - \frac{\sigma^2}{2}\int_s^t f''(x_\theta^1)d\theta - \int_s^t H(x_\theta^1 - x_\theta^2) f'(x_\theta^1) d\theta \right] \\
&\qquad\qquad g(x_{s_1}^1, \ldots, x_{s_p}^1) d\nu^1(x^1) \otimes d\nu^1(x^2) = 0. \qquad (100)
\end{aligned}$$

Then, (100) and the uniqueness to the nonlinear martingale problem (98) imply that $\nu^1 = \mathbb{Q}$ which is equivalent to

$$\lim_{N\to\infty} (Law(\mu^N)) = \delta_{\mathbb{Q}}\ .$$

See [6] for details.

■

We now turn our attention to the numerical approximation of the preceding particle system. We set:

$$\begin{aligned}
Y^i_{(k+1)T/n} &:= Y^i_{kT/n} + \frac{1}{N}\sum_{j=1}^N H\left(Y^i_{kT/n} - Y^j_{kT/n}\right)\frac{T}{n} \\
&\quad + \frac{1}{N}\left(W^i_{(k+1)T/n} - W^i_{kT/n}\right)\ , \qquad (101) \\
\bar{V}_{kT/n}(\cdot) &:= \frac{1}{N}\sum_{i=1}^N H\left(x - Y^i_{kT/n}\right). \qquad (102)
\end{aligned}$$

A much more complex and technical analysis than for the McKean-Vlasov equations with Lipschitz kernels (the study of the propagation from $\frac{kT}{n}$ to $\frac{(k+1)T}{n}$ of the error is very intricate when the kernels are not globally Lipschitz) leads to

Theorem 7.2 (Bossy and Talay [6], Bossy [5]) *Let $V(t,x)$ be the classical solution of the Burgers equation (95) with the initial condition V_0. Suppose (H12).*

Let $\overline{V}_{kT/n}(x)$ be defined as above, N being the number of particles.

There exists an increasing function $K(\cdot)$ such that for all $k \in \{1,\ldots,n\}$:

$$\begin{aligned} &\mathbb{E}\,\|V(kT/n,.) - \overline{V}_{kT/n}(.)\|_{L^1(\mathbb{R})} \\ &\quad \leq\; K(T)\left(\|V_0 \;-\; \overline{V}_0\|_{L^1(\mathbb{R})} \;+\; \frac{1}{\sqrt{N}} \;+\; \frac{1}{\sqrt{n}}\right). \end{aligned} \tag{103}$$

The monotonicity of the function $V_0(\cdot)$ can be relaxed: see [7] for the modification of the algorithm when $V_0(\cdot)$ is non monotonic and for the corresponding error analysis.

In a forthcoming paper, Bossy and Talay extend this analysis to Chorin's random vortex method for the 2-D incompressible Navier-Stokes equation. The interpretation of the Navier-Stokes equation in terms of limit law of weakly interacting particles has been given by Marchioro and Pulvirenti [31] and Osada [35]. The interaction kernel is still less smooth than the Heaviside function since it is the Biot and Savart kernel, which is singular at 0. This makes the error analysis delicate.

For numerical experiments on the above stochastic particle methods related to McKean-Vlasov equations, see M. Bossy's thesis [5].

References

[1] Bally, V., Talay, D., "The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function", *Probability Theory and Related Fields*, 104, 43-60 (1996).

[2] Bally, V., Talay, D., "The law of the Euler scheme for stochastic differential equations (II) : convergence rate of the density", *Monte Carlo Methods and Applications*, 2, 93-128 (1996).

[3] Bensoussan A., Lions, J.L., *Applications des Inéquations Variationnelles en Contrôle Stochastique*, Dunod (1978).

[4] Bernard, P., Talay, D., Tubaro, L., "Rate of convergence of a stochastic particle method for the Kolmogorov equation with variable coefficients", *Math. Comp.*, 63(208), 555-587 (1994).

[5] Bossy, M., *Vitesse de Convergence d'Algorithmes Particulaires Stochastiques et Application à l'Equation de Burgers*, PhD thesis, Université de Provence (1995).

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

[7] Bossy, M., Talay, D., "A stochastic particle method for the McKean-Vlasov and the Burgers equation" (1995). *Math. Comp.*, 66(217), 157-192 (1997).

[8] Calzolari, A., Costantini, C., Marchetti F., "A confidence interval for Monte Carlo methods with an application to simulation of obliquely reflecting Brownian motion", *Stochastic Processes and their Applications*, 29, 209-222 (1988).

[9] Cannarsa, P., Vespri, C., "Generation of analytic semigroups by elliptic operators with unbounded coefficients", *SIAM J. Math. Anal.*, 18(3), 857-872 (1987).

[10] Chorin, A.J., Marsden, J.E., *A Mathematical Introduction to Fluid Mechanics*, Springer Verlag (1993).

[11] Chorin, A.J., "Vortex methods and Vortex Statistics – Lectures for Les Houches Summer School of Theoretical Physics", *Lawrence Berkeley Laboratory Prepublications* (1993).

[12] Costantini, C., Pacchiarotti, B., SARTORETTO, F., "Numerical approximation of functionals of diffusion processes", (1995).

[13] Faure, O., *Simulation du Mouvement Brownien et des Diffusions*, PhD thesis, Ecole Nationale des Ponts et Chaussées (1992).

[14] Freidlin, M., *Functional Integration and Partial Differential Equations.* Annals of Mathematics Studies, Princeton University (1985).

[15] Friedman, A., *Stochastic Differential Equations and Applications*, volume 1, Academic Press, New-York (1975).

[16] Gaines, J.G., Lyons, T.J., "Random generation of stochastic area integrals", *SIAM Journal of Applied Mathematics*, 54(4), 1132-1146 (1994).

[17] Goodman, J., "Convergence of the random vortex method", *Comm. Pure Appl. Math.*, 40, 189-220 (1987).

[18] Gustafson, K.E., Sethian, J.A., (eds.), *Vortex Methods and Vortex Motions*, SIAM (1991).

[19] Hald, O.H., "Convergence of random methods for a reaction diffusion equation", *SIAM J. Sci. Stat. Comput.*, 2, 85-94 (1981).

[20] Hald, O.H., "Convergence of a random method with creation of vorticity", *SIAM J. Sci. Stat. Comput.*, 7, 1373-1386 (1986).

[21] Has'minskii, R.Z., *Stochastic Stability of Differential Equations*, Sijthoff & Noordhoff (1980).

[22] Ikeda, N, Watanabe, S., *Stochastic Differential Equations and Diffusion Processes*, North Holland (1981).

[23] Kanagawa, S., "On the rate of convergence for Maruyama's approximate solutions of stochastic differential equations", *Yokohama Math. J.*, 36, 79-85 (1988).

[24] Kunita, H., "Stochastic differential equations and stochastic flows of diffeomorphisms", *Ecole d'Été de Saint-Flour XII*, LNM 1097, Springer (1984).

[25] Kurtz, T.G., Protter, P., "Wong–Zakai corrections, random evolutions and numerical schemes for S.D.E.'s", in *Stochastic Analysis: Liber Amicorum for Moshe Zakai*, 331-346 (1991).

[26] Kusuoka S., Stroock, D., "Applications of the Malliavin Calculus, part II", *J. Fac. Sci. Univ. Tokyo*, 32, 1-76 (1985).

[27] Lépingle, D., "Un schéma d'Euler pour équations différentielles réfléchies", *Note aux Comptes-Rendus de l'Académie des Sciences*, 316(I), 601-605 (1993).

[28] Lépingle, D., "Euler scheme for reflected stochastic differential equations", *Mathematics and Computers in Simulation*, 38 (1995).

[29] Liu, Y., "Numerical approaches to reflected diffusion processes", (submitted for publication).

[30] Long, D.G., Convergence of the random vortex method in two dimensions, *J. Amer. Math. Soc.*, 1(4) (1988).

[31] Marchioro, C., Pulvirenti, M., "Hydrodynamics in two dimensions and vortex theory", *Comm. Math. Phys.*, 84, 483-503 (1982).

[32] Milshtein, G.N., "The solving of the boundary value problem for parabolic equation by the numerical integration of stochastic equations", (to appear).

[33] Newton, N.J., "Variance reduction for simulated diffusions", *SIAM J. Appl. Math.*, 54(6), 1780-1805 (1994).

[34] Nualart, D., *Malliavin Calculus and Related Topics*. Probability and its Applications, Springer-Verlag (1995).

[35] Osada, H., "Propagation of chaos for the two dimensional Navier-Stokes equation", in K.Itô and N. Ikeda, editors, *Probabilistic Methods in Mathematical Physics*, pages 303–334, Academic Press (1987).

[36] Pardoux, E., "Filtrage non linéaire et équations aux dérivées partielles stochastiques associées", *Cours à l'Ecole d'Eté de Probabilités de Saint-Flour XIX*, LNM 1464, Springer-Verlag (1991).

[37] Protter, P., *Stochastic Integration and Differential Equations*, Springer–Verlag, Berlin (1990).

[38] Protter, P., Talay, D., "The Euler scheme for Lévy driven stochastic differential equations", Annals Prob., 25(1), 393-423 (1997).

[39] Puckett, E.G., A study of the vortex sheet method and its rate of convergence. *SIAM J. Sci. Stat. Comput.*, 10(2), 298-327 (1989).

[40] Puckett, E.G., "Convergence of a random particle method to solutions of the Kolmogorov equation", *Math. Comp.*, 52(186), 615-645 (1989).

[41] Raviart, P.A., "An analysis of particle methods", in *Numerical Methods in Fluid Dynamics*, F. Brezzi (ed.), vol. 1127 of *Lecture Notes in Math.*, 243-324, Springer-Verlag (1985).

[42] Roberts, S., "Convergence of a random walk method for the Burgers equation", *Math. Comp.*, 52(186), 647-673 (1989).

[43] Roynette, B., “Approximation en norme Besov de la solution d’une équation différentielle stochastique”, *Stochastics and Stochastic Reports*, 49, 191-209 (1994).

[44] Slominski, L., “On existence, uniqueness and stability of solutions of multidimensional SDE’s with reflecting boundary conditions”, *Ann. Inst. H. Poincaré*, 29 (1993).

[45] Slominski, L., “On approximation of solutions of multidimensional S.D.E.’s with reflecting boundary conditions”, *Stochastic Processes and their Applications*, 50(2), 197-219 (1994).

[46] Sznitman, A.S., “Topics in propagation of chaos”, *Ecole d’Eté de Probabilités de Saint Flour XIX*, (P.L. Hennequin, ed.), LNM 1464, Springer, Berlin, Heidelberg, New York (1989).

[47] Talay, D., “Second order discretization schemes of stochastic differential systems for the computation of the invariant law”, *Stochastics and Stochastic Reports*, 29(1),13-36 (1990).

[48] Talay, D., “Simulation and numerical analysis of stochastic differential systems: a review”, *Probabilistic Methods in Applied Physics*, (P. Krée and W. Wedig, eds.), *Lecture Notes in Physics* 451, 54-96, Springer-Verlag (1995).

[49] Talay, D., Tubaro, L., *Probabilistic Numerical Methods for Partial Differential Equations*, (book in preparation).

[50] Talay, D., Tubaro, L., “Expansion of the global error for numerical schemes solving stochastic differential equations”, *Stochastic Analysis and Applications*, 8(4), 94-120 (1990).

[51] Tweedie, R.L., “Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space”, *Stochastic Processes and Applications*, 3 (1975).