THE LAW OF THE EULER SCHEME FOR STOCHASTIC DIFFERENTIAL EQUATIONS: I. CONVERGENCE RATE OF THE DISTRIBUTION FUNCTION *

Vlad BALLY
Université du Maine
and
Laboratoire de Probabilités
Université Paris 6
4 Place Jussieu
75252 Paris Cédex 05
France
e-mail : bally@ccr.jussieu.fr

Denis TALAY
INRIA
2004 Route des Lucioles
Sophia-Antipolis
B.P. 109
06561 Valbonne Cedex
France
e-mail : talay@sophia.inria.fr

Abstract

We study the approximation problem of $\mathbb{E}f(X_T)$ by $\mathbb{E}f(X_T^n)$, where (X_t) is the solution of a stochastic differential equation, (X_t^n) is defined by the Euler discretization scheme with step $\frac{T}{n}$, and f is a given function. For smooth f's, Talay and Tubaro have shown that the error $\mathbb{E}f(X_T) - f(X_T^n)$ can be expanded in powers of $\frac{1}{n}$, which permits to construct Romberg extrapolation procedures to accelerate the convergence rate. Here, we prove that the expansion exists also when f is only supposed measurable and bounded, under an additional nondegeneracy condition of Hörmander type for the infinitesimal generator of (X_t): to obtain this result, we use the stochastic variations calculus.

In the second part of this work, we will consider the density of the law of X_T^n and compare it to the density of the law of X_T.

AMS(MOS) classification: 60H07, 60H10, 60J60, 65C05, 65C20, 65B05. KEY WORDS: Stochastic Differential Equations, Malliavin Calculus, Approximation of diffusion processes, Euler scheme for SDE's.

*Published in PROB. THEORY and RELATED FIELDS 104, 43-60 (1996).

1. Introduction

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

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \,, \tag{1.1}$$

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

The problem of computing the expectation $\mathbb{E}f(X_t)$ on a time interval $[0,T]$ by a Monte–Carlo algorithm, (X_t) being a diffusion process, arises from various motivations. For example, in Random Mechanics, a random dynamical system with a white noise being given, one wants to get the two first moments of the response of the system, or the probability that the response reaches a certain level. In numerical analysis, this permits to solve parabolic or elliptic Partial Differential Equations in situations where deterministic algorithms become difficult to use or unefficient, especially when the dimension of the state space is large, when the underlying differential operator is degenerate, or when the objective is to compute the solution only at a few points. In economy, this permits to compute option prices based upon a large panel of assets.

The algorithm consists in approximating the unknown process (X_t) by an approximate process (depending on a parameter denoted by n) X_t^n which can be simulated on a computer, and in simulating a large number M of independent trajectories of X_t^n, so that $\mathbb{E}f(X_t)$ is approximated by:

$$\frac{1}{M}\sum_{i=1}^{M} f(X_t^n(\omega_i)) \,.$$

The resulting error of the algorithm depends on the choice of the approximate process and the two parameters M and n.

The effects of M can be described by the Central–Limit Theorem or large deviation results; in practice, one estimates the maximum value of the variance of $f(X_t)$ for t in $[0,T]$, and then chooses M according to the desired accuracy and the power of the available computer. A sophisticated variance reduction technique has been developed and analysed by Nigel Newton in [8]. Generally M must be large,and, as just mentioned, one chooses a probabilistic technique because the problem is degenerate or high dimensional: therefore, one takes advantages of simple procedures to approximate (X_t).

A natural mean is to use a time discretization scheme of the stochastic differential equation whose (X_t) is the solution: T/n represents the discretization step. For example, the Euler scheme is defined by

$$X^n_{(p+1)T/n} = X^n_{pT/n} + b(X^n_{pT/n})\frac{T}{n} + \sigma(X^n_{pT/n})(W_{(p+1)T/n} - W_{pT/n}) \,. \tag{1.2}$$

For $\frac{pT}{n} \le t < \frac{(p+1)T}{n}$, X_t^n is defined by

$$X_t^n = X^n_{pT/n} + b(X^n_{pT/n})\left(t - \frac{pT}{n}\right) + \sigma(X^n_{pT/n})(W_t - W_{pT/n}) \,.$$

The effects of n can be measured by the quantity:

$$|\mathbb{E}f(X_T) - \mathbb{E}f(X_T^n)| \, . \tag{1.3}$$

Milshtein [6] was the first to show that the schemes built for the quadratic mean convergence, and L^2 estimates of the corresponding errors, are not relevant in that context, since the objective is to approximate the law of (X_t).

Talay ([12] and [13]) and, independently, Milshtein [7], have introduced the appropriate methodology to analyse the error (1.3): it consists in writing this difference as a sum of terms involving the solution of a parabolic PDE (this technique will be used also below). These references provide schemes such that, under smoothness conditions on b, σ, f:

$$|\mathbb{E}f(X_T) - \mathbb{E}f(X_T^n)| \leq \frac{C}{n^\alpha} \, , \; \alpha = 1, 2 \, .$$

Several other schemes have then been proposed by Kloeden and Platen [4].

In Talay and Tubaro [14], a more precise result is proven : under the same conditions, the errors corresponding to these schemes can be expanded in terms of powers of $\frac{1}{n}$, and formulae for the coefficients of the expansion can be derived. In Protter and Talay [11], the same result is shown for the Euler scheme applied to stochastic differential equations driven by general (discontinuous) Levy processes.

Here, we will focus our analysis on the simplest scheme, the Euler scheme: as a consequence of the existence of the expansion, linear combinations of results obtained with this scheme and different step-sizes permit to reach any desired convergence rate (Romberg extrapolation technique: see Talay and Tubaro [14]).

Our objective is to show the existence of the expansion under a much weaker hypothesis on f than in [14]: we will suppose it measurable and bounded (the boundedness could be relaxed); for example, f can be the indicatrix function of a domain: our result applies when one wants to compute probabilities of the type $\mathbb{P}[|X_T| > K]$. In counterpart, we suppose a nondegeneracy condition which in particular ensures that, for any $t > 0$ and any $x \in \mathbb{R}^d$, the law of the random variable $X_t(x)$ has a smooth density with respect to the Lebesgue measure (essentially, this condition is the Hörmander condition for the infinitesimal generator of the process): that condition is less restrictive than the uniform ellipticity of the generator, and therefore our result applies for dynamical systems whose solution, representing a pair (position, velocity), cannot have a uniformly elliptic generator.

The organization of the paper is the following : in the section 2, we recall some results of the Malliavin calculus that we will use in the sequel, in particular an estimate due to Kusuoka and Stroock on the derivatives of the density of $X_t(x)$; in the section 3, we state and comment our main result; the section 4 is devoted to the proof, except two technical lemmas which are proved in the section 5; in the section 6, we give some extensions of the result.

In the second part of this work [2], we will consider the density of the law of X_T^n and compare it to the density of the law of X_T.

Notation. In all the paper, φ being a smooth function, the notation $\partial_\alpha^x \varphi(t,x,y)$ means that the multiindex α concerns the derivation with respect to the coordinates of x, the variables t and y being fixed.

When $\gamma = (\gamma^{ij})$ 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 $(\partial_i V_j)^{ij}$.

Finally, we will use the same notation $K(\cdot)$, q, Q, μ, etc, for different functions and postive real numbers, having the common property to be independent of T and of the approximation parameter n: typically, they will only depend on L^∞-norms of a finite number of partial derivatives of the coordinates of b and σ and on an integer L to be defined below (see the hypothesis (HU)).

2. Some basic results of the Malliavin calculus

One can now find several expositions of the Malliavin calculus: see, for example, Nualart [9] (we use the notation of this book) and Ikeda-Watanabe [3]; a short presentation on the applications to the existence of a density for the law of a diffusion process can be found in Pardoux [10].

We only introduce the material necessary to our computations.

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

For $h(\cdot) \in L^2(\mathbb{R}_+, \mathbb{R}^r)$, we denote by $W(h)$ the quantity $\int_0^T < h(t), dW_t >$.

Let $\mathcal{S}$ be the space of "simple " functionals of the Wiener process W, i.e. the sub-space of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ of random variables F which can be written under the form

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

for some n, polynomial function $f(\cdot)$, $h_i(\cdot) \in L^2(\mathbb{R}_+, \mathbb{R}^r)$.

For $F \in \mathcal{S}$, we denote by $(D_t F)$ the $\mathbb{R}^r$-dimensional process defined by

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

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 $\mathbb{D}^{1,p}$, and we define the norm

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

The j-th component of $D_t F$ will be denoted by $D_t^j F$.

One also defines the k-th order derivative as the the random vector on $[0,T]^k \times \Omega$ whose coordinates are defined by

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

and we denote by $\mathbb{D}^{N,p}$ the completion of $\mathcal{S}$ with respect to the norm

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

$\mathbb{D}^\infty$ will denote the space $\bigcap_{p\geq 1}\bigcap_{j\geq 1}\mathbb{D}^{j,p}$.

For $F \in \mathcal{S}$, one also defines the Ornstein-Uhlenbeck operator L by

$$LF = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1),\ldots,W(h_n))W(h_i) - \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(W(h_1),\ldots,W(h_n)) < h_i, h_j > ,$$

which is a closable operator. The domain of L includes $\mathbb{D}^\infty$.

For $F := (F^1,\ldots,F^m) \in (\mathbb{D}^\infty)^m$, we denote by γ_F the Malliavin covariance matrix associated to F, i.e. the $m \times m$-matrix defined by

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

Definition 2.1. *We will 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*

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

Remark 2.2. The above condition can also be written (we recall that $\hat{\gamma}_F$ denotes the determinant of γ_F):

$$\frac{1}{\hat{\gamma}_F} \in \bigcap_{p\geq 1} L^p(\Omega) \,. \;\blacksquare$$

Our main ingredient is the following integration by parts formula (cf. the section V-9 in Ikeda-Wanabe [3]):

Proposition 2.3. *Let $F \in (\mathbb{D}^\infty)^m$ satisfy the nondegeneracy condition 2.1, 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 $\beta_j \in \{1,\ldots,m\}$, recursively defined in the following way:*

$$\begin{aligned} H_i(F,G) &= H_{(i)}(F,G) \\ &:= -\sum_{j=1}^{m} \left\{ G < D\Gamma_F^{ij}, DF^j >_{L^2(0,T)} + \Gamma_F^{ij} < DG, DF^j >_{L^2(0,T)} + \Gamma_F^{ij} \cdot G \cdot LF^j \right\} , \\ H_\beta(F,G) &= H_{(\beta_1,\ldots,\beta_k)}(F,G) \\ &:= H_{\beta_k}(F, H_{(\beta_1,\ldots,\beta_{k-1})}(F,G)) \,. \end{aligned} \tag{2.1}$$

Then, for any multiindex α,

$$\mathbb{E}[(\partial_\alpha g)(F)G] = \mathbb{E}[g(F)H_\alpha(F,G)]\,. \tag{2.2}$$

We can get the following estimate:

Proposition 2.4. *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*

$$\mathbb{E}[|H_\beta(F,G)|^p \;\; 1\!\!1_{\mathrm{A}}]^{\frac{1}{p}} \le C(p,\beta)\; \|\Gamma_F \; 1\!\!1_{\mathrm{A}}\|_{k(p,\beta)} \; \|G\|_{N(p,\beta),m(p,\beta)} \; \|F\|_{N'(p,\beta),m'(p,\beta)}\,. \tag{2.3}$$

Proof. We apply the Meyer inequality on $\|LF\|_p$ (see the theorem 8.4 of the chapter 5 in Ikeda-Watanabe [3], with $k = 2$, taking into account the definition 8.2 in the same chapter):

$$\|LF\|_p \le C\|F\|_{2,p}\,,$$

and the equality

$$D\Gamma^{ij} = -\sum_{k,l} \Gamma^{ik}\Gamma^{jl} D\gamma^{kl}\,;$$

the result readily follows from the definition (2.1). ∎

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

Definition 2.5. *Let us denote by* $A_0, A_1, \ldots, A_r$ *the vector fields defined by*

$$\begin{aligned} A_0(x) &= \sum_{i=1}^{d} b^i(x)\partial_i\,, \\ A_j(x) &= \sum_{i=1}^{d} \sigma^{ij}(x)\partial_i\,, \quad j = 1, \ldots, r\,. \end{aligned}$$

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

The Hörmander condition is said to hold at the point x *if the vector space spanned by all the vector fields* A_i^α, $1 \le i \le r$ *and* α *multiindex, at the point* x*, is* $\mathbb{R}^d$.

Theorem 2.6. *Assume that the coefficients* b *and* σ *are infinitely differentiable, with bounded derivatives of order strictly larger than 1. Then, for all* x*, all* $t > 0$ *and* $i = 1, \ldots, d$, $X_t^i(x)$ *belongs to* $\mathbb{D}^\infty$.

Besides, suppose that the Hörmander condition holds at some point x. Let $\gamma_t(x)$ denote the Malliavin covariance matrix corresponding to $X_t(x)$, and $\Gamma_t(x)$ its inverse.

Then, for any $t > 0$, one has

$$\|\Gamma_t(x)\| \in \bigcap_{p\geq 1} L^p(\Omega)\ ,$$

and the random vector $X_t(x)$ has an infinitely differentiable density $p_t(x,\cdot)$.

Actually, Kusuoka and Stroock [5] give an exponential bound for $p_t(x,\cdot)$ in terms of the following quadratic forms:

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

Set

$$V_L(x) = 1 \wedge \inf_{\|\eta\|=1} V_L(x,\eta)\ . \tag{2.4}$$

The exponential bounds require some smoothness conditions on b and σ, and are valid for x in the set $\{x \in \mathbb{R}^d : V_L(x) > 0\}$; as we will apply this estimate for $x = X_t^n$, we assume

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

(C) 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).

Under the conditions (C) and (UH), and a corresponding L being fixed (the smallest one, for example), the corollary 3.25 in Kusuoka and Stroock asserts: for any integers m, k and any multiindices α and β such that $2m+|\alpha|+|\beta| \leq k$, there exist an integer $M(k,L)$, a non decreasing function $K(T)$ 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]:

$$|\partial_t^m \partial_x^\alpha \partial_y^\beta p_t(x,y)| \leq \frac{K(T)(1+\|x\|^Q)}{t^q(1+\|y-x\|^2)^k} e^{-C\frac{(\|x-y\|\wedge 1)^2}{t(1+\|x\|)^2}}\quad,\ \forall 0 < t \leq T\ . \tag{2.5}$$

The same theorem provides the following estimate for $\Gamma_t(x)$: for any $p \geq 1$, for some constants C, μ one has

$$\|\Gamma_t(x)\|_p \leq K(t)\frac{1+\|x\|^\mu}{t^{dL}}\ . \tag{2.6}$$

Remark 2.7. The rate of degeneracy of $p_t(x,y)$ as $t \to 0$ is controlled by that of $\Gamma_t(x)$; (2.6) gives an upper bound of order $1/t^{dL}$; the lower bound $1/t^{L-1}$ is proved in the theorem 5.1 in Bally [1]. ∎

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

3. Our main result

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

$$\mathcal{L} := \sum_{i=1}^{d} b^i(x)\partial_i + \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^*)^{ij}(x)\partial_{ij} \,. \tag{3.1}$$

Consider a measurable bounded function f and $u(t,x) := \mathbb{E}f(X_{T-t}(x))$ which solves:

$$\begin{cases} \dfrac{\partial u}{\partial t} + \mathcal{L}u = 0 \ ,0 \le t < T \,, \\ u(T,\cdot) = f(\cdot) \,. \end{cases} \tag{3.2}$$

Denote by a the matrix $\sigma\sigma^*$. Let $\Psi(t,x)$ be defined by

$$\begin{aligned} \Psi(t,x) &= \frac{1}{2}\sum_{i,j=1}^{d} b^i(x)b^j(x)\partial_{ij}u(t,x) + \frac{1}{2}\sum_{i,j,k=1}^{d} b^i(x)a^{jk}(x)\partial_{ijk}u(t,x) \\ &\quad + \frac{1}{8}\sum_{i,j,k,l=1}^{d} a^{ij}(x)a^{kl}(x)\partial_{ijkl}u(t,x) + \frac{1}{2}\frac{\partial^2}{\partial t^2}u(t,x) \\ &\quad + \sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial t}\partial_i u(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial}{\partial t}\partial_{ij}u(t,x) \,. \end{aligned} \tag{3.3}$$

Theorem 3.1. *Let f be a measurable and bounded function. Under the hypotheses (UH) and (C), the Euler scheme error satisfies*

$$\mathbb{E}f(X_T(x)) - \mathbb{E}f(X_T^n(x)) = -\frac{C_f(T,x)}{n} + \frac{Q_n(f,T,x)}{n^2} \,; \tag{3.4}$$

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 $K(T)$ 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

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

The expansion (3.4) was first obtained by Talay and Tubaro in [14]. No nondegeneracy assumption of Hörmander type was necessary, but in counterpart the function $f(\cdot)$ was supposed smooth enough; in that context, one obtains a bound of type

$$|C_f(T,x)| + sup_n|Q_n(f,T,x)| \le K(T)\sum_{|\alpha|\le 6}\|\partial_\alpha^x f\|_\infty \,.$$

Besides, when f is smooth, the analysis shows that the simulation of Brownian paths is unnecessary to get the existence of the expansion: the algorithm may involve appropriate discrete lawed random variables as well (see [14]). In our context, this property does not remain true. This has no practical incidence: the simulation of the increments of the Wiener process can efficiently be performed.

4. Proof of Theorem 3.1

The proof of the preceding theorem is based upon the two following technical lemmas, which are interesting by themselves. Their statements suppose that the hypotheses of the theorem 3.1 hold.

Lemma 4.1. *Let the function u be defined by (3.2). Then, for any multiindex α and for any smooth function with polynomial growth g, there exist a non decreasing function $K(T)$ and positive constants q, Q, uniform with respect to n and T, such that*

$$\forall t \in [0,T] \quad , \quad |\mathbb{E}[g(X_t(x))\partial_\alpha u(t, X_t(x))]| \leq K(T)\|f\|_\infty \frac{1+\|x\|^Q}{T^q} \,, \tag{4.1}$$

and

$$\forall t \in \left[0, T-\frac{T}{n}\right] \quad , \quad |\mathbb{E}[g(X_t^n(x))\partial_\alpha u(t, X_t^n(x))]| \leq K(T)\|f\|_\infty \frac{1+\|x\|^Q}{T^q} \,. \tag{4.2}$$

Lemma 4.2. *Let γ et λ be multiindices, let g and g_γ be smooth functions with polynomial growth. Set*

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

There exist a non decreasing function $K(T)$ and positive constants q, Q, uniform with respect to n and T, such that

$$\forall \theta \in \left[0, T-\frac{T}{n}\right] \quad , \quad \forall t \in \left[0, \theta-\frac{T}{n}\right] \quad ,$$
$$\left|\mathbb{E}\left[g(X_t^n(x))\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z)\mathcal{B}_{z=X_t^n(x)}\right]\right| \leq K(T)\|f\|_\infty \frac{1+\|x\|^Q}{T^q} \,. \tag{4.3}$$

Lemma 4.3. *For some integer q and some non decreasing function $K(T)$, one has that*

$$|\mathbb{E}f(X_T^n(x)) - \mathbb{E}(P_{T/n}f)(X_{T-T/n}^n(x))| \leq \frac{K(T)}{n^2}\|f\|_\infty(1+\|x\|^Q) \,. \tag{4.4}$$

For a while, we admit the lemmas 4.1,4.2,4.3 which will be proven in the section 5.

We start with an easy other lemma.

Lemma 4.4. *It holds that*

$$\mathbb{E}f(X_T^n(x)) - \mathbb{E}f(X_T(x)) = \frac{T^2}{n^2}\sum_{k=0}^{n-2}\mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}^n(x)\right) + \sum_{k=0}^{n-1} R_k^n \,, \tag{4.5}$$

where

$$R_{n-1}^n := \mathbb{E}f(X_T^n(x)) - \mathbb{E}(P_{T/n}f)(X_{T-T/n}^n(x)) \,,$$

and for $k < n-1$, R_k^n can be explicited under a sum of terms, each of them being of the form

$$\mathbb{E}\left[\varphi_\alpha^\natural(X^n_{kT/n}(x))\int_{kT/n}^{(k+1)T/n}\int_{kT/n}^{s_1}\int_{kT/n}^{s_2}(\varphi_\alpha^\sharp(X^n_{s_3}(x))\partial_\alpha u(s_3,X^n_{s_3}(x))\right.$$
$$\left.\varphi_\alpha^\dagger(X_{kT/n})\int_{kT/n}^{(k+1)T/n}\int_{kT/n}^{s_1}\int_{kT/n}^{s_2}\varphi_\alpha^\flat(X_{s_3})\partial_\alpha u(s_3,X_{s_3}))ds_3ds_2ds_1\right], \quad (4.6)$$

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

Proof. We follow [14], just changing the presentation.

For a fixed z in $\mathbb{R}^d$, we define the differential operator $\mathcal{L}_z$ by

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

We note that, for $z = X^n_{kT/n}(x)$, $\mathcal{L}_z$ is the infinitesimal generator of the diffusion process $\left(X_t^n(x), \frac{kT}{n} \leq t < \frac{(k+1)T}{n}\right)$.

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

$$\mathbb{E}f(X_T^n(x)) - \mathbb{E}f(X_T(x)) = \mathbb{E}u(T, X_T^n(x)) - u(0,x) = \sum_{k\leq n-1}\delta_k^n$$

with

$$\delta_k^n := \mathbb{E}\left[u\left(\frac{(k+1)T}{n}, X^n_{(k+1)T/n}(x)\right) - u\left(\frac{kT}{n}, X^n_{kT/n}(x)\right)\right]. \quad (4.7)$$

The Itô formula implies

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

from which, using (3.2), one gets

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

Denote

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

and

$$\begin{aligned} J_k^n(t) &:= \mathcal{L}_z u\left(\frac{kT}{n}, X_{kT/n}^n(x)\right)\mathcal{B}_{z=X_{kT/n}^n(x)} - Lu(t, X_t^n(x)) \\ &= Lu\left(\frac{kT}{n}, X_{kT/n}^n(x)\right) - Lu(t, X_t^n(x)) \,. \end{aligned}$$

We have:

$$\delta_k^n = \mathbb{E}\int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} (I_k^n(t) + J_k^n(t))dt \,.$$

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 (3.2), so that Lu solves a similar PDE. ∎

We can deduce the following corollary (note that this upper bound in terms of $\|f\|_\infty$ was stated nowhere else in the literature before, even for smooth functions f (to our knowledge); but we do not focus on it, since our objective is a stronger statement):

Corollary 4.5. *There exist a non decreasing function $K(T)$ and constants q, Q such that:*

$$|\mathbb{E}f(X_T^n(x)) - \mathbb{E}f(X_T(x))| \leq \frac{K(T)\|f\|_\infty}{T^q}(1 + \|x\|^Q)\frac{1}{n} \,. \tag{4.8}$$

Proof. We apply the lemma 4.3, and for $k < n-1$, we apply the estimates of the lemma 4.1 to $\Psi(\frac{kT}{n}, X_{kT/n}^n(x))$ and R_k^n. ∎

Before ending our proof, we make an easy remark.

We recall that $\partial_\alpha^x \varphi(t, x, y)$ means that the multiindex α concerns the derivation with respect to the coordinates of x, the variables t and y being fixed. As $u(t, x) = P_{T-t}f(x)$, one has

$$\partial_\alpha^x u(t, x) = \int \partial_\alpha^x p_{T-t}(x, y) f(y) dy \,.$$

Now we use the estimate (2.5) and get, for some $k \geq |\alpha| \geq 1$:

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

so that

$$|\partial_\alpha^x u(t, x)| \leq K(T)\frac{\|f\|_\infty}{(T-t)^q}(1 + \|x\|^Q) \,. \tag{4.9}$$

Now we are in position to prove (3.4).

The expansion (4.5) can be rewritten under the following form:

$$\begin{aligned}
\mathbb{E}f(X_T^n(x)) - \mathbb{E}f(X_T(x)) \quad = \quad & \frac{T^2}{n^2}\sum_{k=0}^{n-2} \mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}(x)\right) \\
& + \frac{T^2}{n^2}\sum_{k=0}^{n-2} \mathbb{E}\left[\Psi\left(\frac{kT}{n}, X_{kT/n}^n(x)\right) - \Psi\left(\frac{kT}{n}, X_{kT/n}(x)\right)\right] \\
& + \sum_{k=0}^{n-1} R_k^n \, . \qquad (4.10)
\end{aligned}$$

For any k such that $\frac{k}{n} \leq \frac{1}{2}$, we apply the inequalities (4.8) (with $\Psi(kT/n, \cdot)$ instead of $f(\cdot)$) and (4.9) (to upper bound $\|\Psi(kT/n, \cdot)\|_\infty$):

$$\left|\mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}^n(x)\right) - \mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}(x)\right)\right| \leq \frac{K(T)}{nT^q}\|f\|_\infty(1 + \|x\|^Q) \, . \qquad (4.11)$$

For $\frac{k}{n} \geq \frac{1}{2}$, one applies the expansion (4.5), substituting the function $f_{n,k}(x) := \Psi(\frac{kT}{n}, \cdot)$ to $f(\cdot)$. Set $u_{n,k}(t,x) := P_{kT/n-t}f_{n,k}(x)$ and denote by $\Psi_{n,k}(t, \cdot)$ the function defined in (3.3) with $u_{n,k}$ instead of u; then one has that

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

where the $R_j^{n,k}$'s are sums of terms of type (4.6) with $u_{n,k}$ instead of u. We apply the lemma 4.2 to upper bound the right hand-side.

Combining this estimate with (4.11), we get (for a new function $K(\cdot)$ and new constants):

$$\frac{T^2}{n^2}\sum_{k=0}^{n-2}\left|\mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}^n(x)\right) - \mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}(x)\right)\right| \leq \frac{K(T)}{n^2T^q}\|f\|_\infty(1 + \|x\|^Q) \, .$$

We proceed similarly to upper bound $\left|\sum_{k=0}^{n-2} R_k^n\right|$, and we apply the lemma 4.3 to upper bound $\left|R_{n-1}^n\right|$.

Finally, using Itô's formula and the estimate (4.1), we get

$$\left|\frac{T}{n}\sum_{k\leq n-1} \mathbb{E}\Psi\left(\frac{kT}{n}, X_{kT/n}(x)\right) - \int_0^T \mathbb{E}\Psi\left(s, X_s(x)\right) ds\right| \leq \frac{K(T)}{n}\|f\|_\infty(1 + \|x\|^Q) \, .$$

We note that (4.1) also ensures that $\int_0^T \mathbb{E}|\Psi\left(s, X_s(x)\right)| ds$ is finite.

That ends the proof of the theorem 3.1. ∎

5. Proof of Lemmas 4.1, 4.2 and 4.3

We first state a technical lemma.

Lemma 5.1. *Under the above hypotheses, for any $p > 1$ and $j \geq 1$, there exist an integer Q and a non decreasing function $K(t)$ such that*

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

and

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

Proof. We just have to mimic the classical computations giving estimates for $\|X_t(x)\|_{j,p}$. For example, let us show how we can proceed for $j = 1$. Let $\eta^n(t)$ denote $\frac{pT}{n}$, where the integer p is such that $\frac{pT}{n} \leq t < \frac{(p+1)T}{n}$. We remark that, for $t - \theta > \frac{T}{n}$, $D_\theta^k X_t^n(x)$ satisfies:

$$\begin{aligned} D_\theta^k X_t^n(x) &= \sigma(X_{\eta^n(\theta)}^n(x)) + \sum_{l=1}^r \int_{\eta^n(\theta)+T/n}^t \partial\sigma_l(X_{\eta^n(s)}^n(x)) D_\theta^k X_{\eta^n(s)}^n(x) dW_s \\ &\quad + \int_{\eta^n(\theta)+T/n}^t \partial b(X_{\eta^n(s)}^n(x)) D_\theta^k X_{\eta^n(s)}^n(x) ds \,. \end{aligned}$$

Under (C), a classical use of Gronwall's lemma permits to get

$$\sup_{n\geq 1} \sup_{0\leq t\leq T} \|X_t^n(x)\|_{1,p} < K(T)(1 + \|x\|^Q).$$

For other values of j and for the difference $X_t(x) - X_t^n(x)$, we proceed in the same way. ∎

5.1. Proof of Lemma 4.1.

We only prove the part (4.2), the part (4.1) being treated with the same arguments.

5.1.1. Small t.

We first consider the case $0 \leq t \leq \frac{T}{2}$. As $T - t \geq T/2$, the inequality (4.9) yields (4.2).

5.1.2. Large t.

Now, let us treat the situation $\frac{T}{2} \leq t \leq T - \frac{T}{n}$. The preceding argument cannot be used, since for $\theta \to T$ the measure $p_{T-\theta}(x, y)dy$ converges weakly to the Dirac measure at point x.

The principle of the rest of the proof is the following: in order to get rid of the derivatives of $u(t, x)$, we will use Malliavin's integration by parts formula with respect to the functional $X_t^n(x)$, which is expected non degenerated (with a high probability) for

$t \geq T/2$ because $X_t^n(x)$ approximates $X_t(x)$; estimates on the L^p-norm of the inverse of the Malliavin covariance matrix of $X_t^n(x)$, Γ_t^n [2], can be directly obtained under a uniform ellipticity condition but, under (UH) only, we are led to compare $\|\Gamma_t^n\|_p$ and $\|\Gamma_t\|_p$ (where Γ_t denotes the inverse of the Malliavin covariance matrix of $X_t(x)$) and we will use a localization argument:

- let Ω_0 be the set of events where $|\hat{\gamma}_t^n - \hat{\gamma}_t|$ is larger than $C\hat{\gamma}_t$; to prove that $\mathbb{P}(\Omega_0)$ is small, we will use two facts: first, $(X_t^n(x))$ is a "good" approximation of $(X_t(x))$; second, the $\|\hat{\gamma}_t^{-1}\|_p$'s are finite;
- 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 (2.6)), which allows integrations by parts of type 2.2 with a good control of the L^p-norms of the variables H_α.

Let $\phi \in \mathcal{C}_b^\infty(\mathbb{R})$ such that $\phi(x) = 1$ for $|x| \leq \frac{1}{4}$, $\phi(x) = 0$ for $|x| \geq \frac{1}{2}$ and $0 < \phi(x) < 1$ for $|x| \in (\frac{1}{4}, \frac{1}{2})$.

Set

$$r_t^n := \frac{(\hat{\gamma}_t^n - \hat{\gamma}_t)}{\hat{\gamma}_t}\,.$$

One has

$$\begin{aligned} \mathbb{E}[g(X_t^n(x))\partial_\alpha^x u(t, X_t^n(x))] &= \mathbb{E}[g(X_t^n(x))\partial_\alpha^x u(t, X_t^n(x))(1-\phi(r_t^n))] \\ &\quad + \mathbb{E}[g(X_t^n(x))\partial_\alpha^x u(t, X_t^n(x))\phi(r_t^n)] \\ &=: A + B\,. \end{aligned}$$

To upper bound $|A|$, we use (4.9):

$$|A| \leq K(T)\frac{\|f\|_\infty}{(T-t)^q}\mathbb{E}|1-\phi(r_t^n)|\,. \tag{5.1}$$

Since $1 - \phi(r_t^n) = 0$ for $|r_t^n| \leq \frac{1}{4}$, one has

$$\begin{aligned} \mathbb{E}|1-\phi(r_t^n)| &\leq \mathbb{P}\left[|r_t^n| \geq \frac{1}{4}\right] = \mathbb{P}\left[|\hat{\gamma}_t^n - \hat{\gamma}_t| \geq \frac{\hat{\gamma}_t}{4}\right] \\ &\leq \mathbb{P}\left[\hat{\gamma}_t \leq \frac{1}{n^{1/4}}\right] + \mathbb{P}\left[|\hat{\gamma}_t^n - \hat{\gamma}_t| \geq \frac{1}{4n^{1/4}}\right] \\ &= \mathbb{P}\left[\hat{\gamma}_t^{-1} \geq n^{1/4}\right] + \mathbb{P}\left[|\hat{\gamma}_t^n - \hat{\gamma}_t| \geq \frac{1}{4n^{1/4}}\right]\,. \end{aligned}$$

[2] In the sequel, γ_t^n (resp. γ_t) will denote the Malliavin covariance matrix of $X_t^n(x)$ (resp. $X_t(x)$); note that, in order to simplify the notation, we systematically drop the dependancy on x except for X_t^n and X_t theirselves.

Thus, for any $p \geq 1$, one has

$$\mathbb{E}|1-\phi(r_t^n)| \leq n^{-p/4}\mathbb{E}|\hat{\gamma}_t|^{-p} + (4n^{1/4})^p \mathbb{E}|\hat{\gamma}_t^n - \hat{\gamma}_t|^p \,.$$

But (see the lemma 5.1): $\|X_t(x) - X_t^n(x)\|_{1,p} \leq K(t)(1+\|x\|^Q)n^{-1/2}$, so that

$$\|\hat{\gamma}_t^n - \hat{\gamma}_t\|_p \leq (1+\|x\|^Q)K(t)n^{-1/2} \,.$$

On the other hand, under our nondegeneracy assumption, one has (see (2.6))

$$\sup_{\frac{T}{2}\leq t\leq T} \|\hat{\gamma}_t^{-1}\|_p < K(T)\frac{1+\|x\|^\mu}{T^{dL}} \,.$$

As a consequence, for any $p > 0$ there exists an increasing function $K(\cdot)$ and an integer Q such that

$$\mathbb{E}|1-\phi(r_t^n)| \leq K(T)(1+\|x\|^Q)n^{-p/4}\frac{1}{T^{pdL}} \,, \tag{5.2}$$

from which and (5.1), remembering that for $t \leq T - \frac{T}{n}$, $(T-t)^{-q} \leq Cn^q$, one gets

$$|A| \leq \|f\|_\infty n^{q-p/4}K(T)(1+\|x\|^Q)\frac{1}{T^{pdL}} \,. \tag{5.3}$$

To obtain the desired result, it remains to choose $p = 4q$.

We now treat B. We want to apply the proposition (2.3). As X_t^n may not satisfy the nondegeneracy condition, we make a slight modification: we change the time interval for $[0, T+\epsilon]$ with $\epsilon > 0$, and on $[0, T+\epsilon]$ we set

$$X_t^{n,\epsilon} := X_t^n + \epsilon W_{T+\epsilon} \,.$$

Then $X_t^{n,\epsilon}$ satisfies the nondegeneracy condition 2.1 for all $\epsilon > 0$ and all $t \in [0, T+\epsilon]$.

In order to simplify the notation, we continue to write X^n instead of $X^{n,\epsilon}$. In the computations which follow, the Sobolev norms below are computed for the time interval $[0, T+\epsilon]$. It must be understood that, at the end, we make ϵ tend to 0: the constants which appear in these computations can be chosen uniform w.r.t. ϵ.

The proposition (2.3) implies:

$$B = \mathbb{E}[u(t, X_t^n(x))H_\alpha^n]$$

where $H_\alpha^n := H_\alpha(X_t^n(x), g(X_t^n(x))\phi(r_t^n))$. First, we observe that H_α^n is a sum of terms, each one being a product which includes a partial derivative of ϕ evaluated at point r_t^n. From the definition of ϕ it follows that $H_\alpha^n = H_\alpha^n \ \mathbb{1}_{[0,1/2]}(|r_t^n|)$. On $|r_t^n| \leq \frac{1}{2}$ one has $\frac{3}{2}\hat{\gamma}_t \geq \hat{\gamma}_t^n \geq \frac{1}{2}\hat{\gamma}_t$ and therefore

$$H_\alpha^n = H_\alpha^n \ \mathbb{1}_{[\frac{3}{2}\hat{\gamma}_t \geq \hat{\gamma}_t^n \geq \frac{1}{2}\hat{\gamma}_t]} \,.$$

Consequently,

$$|B| \leq C\|f\|_\infty \, E|H^n_\alpha \;\; 1\!\!1_{[\hat\gamma^n_t \geq \frac12 \hat\gamma_t]}| \,.$$

We apply (2.3) and obtain, for some integers k, N, N':

$$\begin{aligned} |B| &\leq C\|f\|_\infty \left\|\Gamma^n_t \;\; 1\!\!1_{[\hat\gamma^n_t \geq \frac12 \hat\gamma_t]}\right\|_k \|X^n_t(x)\|_{N,m} \, \|g(X^n_t(x))\phi(r^n_t)\|_{N',m'} \\ &\leq K(T)\|f\|_\infty \, \|\Gamma_t\|_k \frac{1+\|x\|^\mu}{t^{dL}} \;\leq\; K(T)\|f\|_\infty \frac{1+\|x\|^\mu}{t^{dL}} \,. \end{aligned}$$

As $\frac{T}{2} \leq t < T$ we obtain

$$B \leq K(T)\|f\|_\infty \frac{1+\|x\|^\mu}{T^{dL}} \,.$$

∎

5.2. Proof of Lemma 4.2.

For $\theta \leq \frac{T}{2}$, the derivatives of the function $\varphi(\theta,\cdot)$ can be uniformly bounded in θ (remember (4.9)); thus, to get (4.3) one can simply use the lemma 4.1 with $\varphi(\theta,\cdot)$ instead of $f(\cdot)$ and $P_{\theta-t}\varphi(\theta,\cdot)$ instead of $u(t,\cdot) = P_{T-t}f(\cdot)$.

For $\theta \geq \frac{T}{2}$ and $\theta - t \geq \frac{T}{4}$, one can note that

$$\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z) = \int g_\gamma(y) \int f(y')\partial^y_\gamma p_{T-\theta}(y,y')dy' \; \partial^z_\lambda p_{\theta-t}(z,y)dy \quad ; \tag{5.4}$$

an integration by parts w.r.t. y and the inequality (4.9) give the result.

Consider now the case where $\theta \geq \frac{T}{2}$ and $\theta - t \leq \frac{T}{4}$; in that case, $t \geq \frac{T}{4}$ and, in order to get rid of the explosion of the upper bounds for the derivatives of $\varphi(\theta,\cdot)$, we are going to use the law of $X^n_t(x)$: this argument is similar to what we have done for the lemma 4.1, but here it is unsufficient because we must deal with $\partial_\gamma P_{T-\theta}f$ instead of f. In any case, we start as in the subsection 5.1.2:

$$\begin{aligned} E\left[g(X^n_t(x))\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z)\mathcal{B}_{z=X^n_t(x)}\right] &= E\left[(1-\phi(X^n_t(x)))g(X^n_t(x))\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z)\mathcal{B}_{z=X^n_t(x)}\right] \\ &\quad + E\left[\phi(r^n_t)g(X^n_t(x))\partial_\lambda P_{\theta-t}\varphi(\theta,\cdot)(z)\mathcal{B}_{z=X^n_t(x)}\right] \\ &=: \tilde A + \tilde B \,. \end{aligned} \tag{5.5}$$

We upper bound $|\tilde A|$ by using (5.4), (4.9) and (5.2): we obtain an estimate similar to the right hand-side of (5.3).

Now consider $\tilde B$. First, we apply the proposition 2.3 again [3], and get

$$\tilde B = E[P_{\theta-t}\varphi(\theta,\cdot)(X^n_t(x))H^n_\lambda(t)] \,.$$

[3] As in the preceding proof, we should introduce $X^{n,\epsilon}$. We omit this detail.

We now use a probabilistic representation of $P_{\theta-t}\varphi(\theta,\cdot)$, based upon a process $(\tilde{X}_t)$ which is a weak solution of (1.1) independent of (X_t); we denote by $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}})$ the probability space on which $(\tilde{X}_t)$ is defined, and $\tilde{\mathbb{E}}$ the expectation under $\tilde{\mathbb{P}}$. Applying the proposition 5.2 below, one gets:

$$\begin{aligned} P_{\theta-t}\varphi(\theta,\cdot) &= \tilde{\mathbb{E}}\varphi(\theta,\tilde{X}_{\theta-t}(\cdot)) \\ &= \tilde{\mathbb{E}}[g_\gamma(\tilde{X}_{\theta-t}(\cdot))(\partial_\gamma u)(\theta,\tilde{X}_{\theta-t}(\cdot))] \\ &= \sum_\rho \tilde{\mathbb{E}}[g_\gamma(\tilde{X}_{\theta-t}(\cdot))\partial_\gamma\{u(\theta,\tilde{X}_{\theta-t}(\cdot))\}\tilde{Q}^\rho_t(\cdot)]\,. \end{aligned}$$

Thus,

$$\begin{aligned} \tilde{B} &= \sum_\rho \mathbb{E}\left[H^n_\lambda(t)\tilde{\mathbb{E}}\left[g_\gamma(\tilde{X}_{\theta-t}(z))\partial^z_\gamma\left\{u(\theta,\tilde{X}_{\theta-t}(z))\tilde{Q}^\rho_t(z)\right\}\mathcal{B}_{z=X^n_t(x)}\right]\right] \\ &= \sum_\rho \tilde{\mathbb{E}}\mathbb{E}\left[H^n_\lambda(t)\left[g_\gamma(\tilde{X}_{\theta-t}(z))\partial^z_\gamma\left\{u(\theta,\tilde{X}_{\theta-t}(z))\tilde{Q}^\rho_t(z)\right\}\mathcal{B}_{z=X^n_t(x)}\right]\right]\ ; \end{aligned}$$

we fix $\tilde{\omega}$ and use the integration by parts formula (2.2) with $F(\omega)=X^n_t(x,\omega)$: for some random variable $\tilde{H}^n_{\lambda,\gamma,\rho}(\theta,t,x)$, one has:

$$\tilde{B}=\sum_\rho \tilde{\mathbb{E}}\mathbb{E}\left[u(\theta,\tilde{X}_{\theta-t}(z))\mathcal{B}_{z=X^n_t(x)}\tilde{H}^n_{\lambda,\gamma,\rho}(\theta,t,x)\right]\,.$$

We now conclude as at the end of the subsection 5.1.2. ■

In the above proof, we have used the

Proposition 5.2. *Let $\tilde{X}_t(\cdot,\tilde{\omega})$ denote a version of class $\mathcal{C}^\infty$ of the stochastic flow $y\to\tilde{X}_t(y,\tilde{\omega})$; let $\tilde{Y}_t(\cdot,\tilde{\omega})$ denote its Jacobian matrix and $\tilde{Z}_t(\cdot,\tilde{\omega})$ the inverse matrix of $\tilde{Y}_t(\cdot,\tilde{\omega})$.*

For any multiindex γ, there exists processes $(\tilde{Q}^\rho_t)$ such that: for any smooth real function F, for any $y\in\mathbb{R}^d$,

$$(\partial_\gamma F)(\tilde{X}_t(y))=\sum_{|\rho|\le|\gamma|}\tilde{Q}^\rho_t(y)\partial_\rho\{F\circ\tilde{X}_t(\cdot,\tilde{\omega})\}(y)\ a.s.\,, \tag{5.6}$$

and $\tilde{Q}^\rho_t(y)$ is a polynomial function of the coordinates of $\tilde{Z}_t(y,\tilde{\omega})$.

Proof. We proceed by induction. For $|\gamma|=1$, we observe that

$$\nabla F\circ\tilde{X}_t(\cdot,\tilde{\omega})=\tilde{Y}_t(\cdot,\tilde{\omega})\times(\nabla F)(\tilde{X}_t(\cdot))\ ;$$

it now remains to multiply the two sides of the equality by $\tilde{Z}_t(\cdot,\tilde{\omega})$.

Suppose that the relation (5.6) holds for $|\gamma|\le k$; as

$$\nabla(\partial_\gamma F\circ\tilde{X}_t(\cdot,\tilde{\omega}))=\tilde{Y}_t(\cdot,\tilde{\omega})\times(\nabla\partial_\gamma F)(\tilde{X}_t(\cdot))\,,$$

a multiplication of the two sides of that equality by $\tilde{Z}_t(y,\tilde{\omega})$ and (5.6) (for $|\gamma|\le k$) imply that (5.6) also holds for $|\gamma|=k+1$. ■

5.3. Proof of Lemma 4.3. The proof of the lemma 4.1 cannot apply to treat δ^n_{n-1} (defined in (4.7)) because the last argument before (5.3) cannot be used. We still localize by introducing ϕ, but we do it at once: set

$$\begin{aligned} A^* &:= |\mathbb{E}[(f(X^n_T(x)) - P_{T/n}f(X^n_{T-T/n}(x)))\,(1-\phi(r^n_{T/2}))]|\,, \\ B^* &:= |\mathbb{E}[(f(X^n_T(x)) - P_{T/n}f(X^n_{T-T/n}(x)))\,\phi(r^n_{T/2})]| \\ &= |\mathbb{E}[(u(T,X^n_T(x)) - u(T-T/n, X^n_{T-T/n}(x)))\phi(r^n_{T/2})]|\,. \end{aligned}$$

Clearly (remember (5.2)):

$$A^* \leq 2\|f\|_\infty\, \mathbb{E}|1-\phi(r^n_{T/2})| \;\leq\; \frac{K(T)}{T^q n^2}\|f\|_\infty(1+\|x\|^Q)\,.$$

To treat the term B^*, we proceed as in the proof of the lemma 4.4 to express δ^n_k as an integration of $I^n_k(t)$ and $J^n_k(t)$, and we apply the arguments used in the subsection 5.1.2, especially the integration by parts formula (2.2) with $F = X^n_t(x)$. ∎

6. Extensions

In the theorem 3.1, the boundedness hypothesis on f can be relaxed: the preceding technique can be improved to treat the case of functions f which are measurable and have a polynomial growth, i.e such that

$$\forall x \quad , \quad \|f(x)\| \leq C_f(1+\|x\|^{q_f})$$

for some C_f and q_f; then in all the estimates of the proof, $\|f\|_\infty$ must be replaced by a constant C depending on C_f and q_f; indeed, instead of upper bounding quantities of type $\|f(X^n_t(x))\|_{L^p(\Omega)}$ by $\|f\|_\infty$, one can use that $\mathbb{E}\|X^n_t(x)\|^p$ can be upper bounded by $C(1+\|x\|^q)$, where the constants C and q are uniform in n.

One can also show the existence of an expansion up to any order: as in [14], instead of bounding the I^n_k's of the expansion (4.5), one can apply the technique used at the end of the proof of the theorem 3.1 and make appear an integral of type $\frac{1}{n^2}\int_0^T \mathbb{E}\Psi_1\left(s, X_s(x)\right) ds$ (for some appropriate function Ψ_1); this operation makes appear a new remaining term for which one applies the lemmas 4.1 and 4.3.

Finally, the result can be extended to other schemes. The numerical benefit is less clear, since they involve derivatives of the coefficients and therefore require a larger computational effort than the Euler scheme (they also may lead to larger coefficients of $\frac{T}{n^i}$ in the expansion of the error).

7. Conclusion

We have proved that the error corresponding to the approximation of $\mathbb{E}f(X_T)$ by $\mathbb{E}f(X_T^n)$, X_T^n being given by the Euler scheme, can be expanded in terms of $\frac{1}{n}$ when f is a bounded and measurable function, under an hypothesis of uniform hypoellipticity.

It now remains to give estimates on the convergence rate of the density of X_T^n to the density of X_T (when they do exist). This will done in the second part of this work, in preparation [2].

References

[1] V. BALLY. On the connection between the Malliavin covariance matrix and Hörmander's condition. *J. of Functional Analysis*, 96(2):219–255, 1991.

[2] V. BALLY and D. TALAY. The law of the Euler scheme for stochastic differential equations (II) : convergence rate of the density. Submitted for publication, 1995.

[3] N. IKEDA and S. WATANABE. *Stochastic Differential Equations and Diffusion Processes*. North Holland, 1981.

[4] P.E. KLOEDEN and E.PLATEN. *Numerical Solution of Stochastic Differential Equations*. Springer–Verlag, 1992.

[5] S. KUSUOKA and D. STROOCK. Applications of the Malliavin Calculus, part II. *J. Fac. Sci. Univ. Tokyo*, 32:1–76, 1985.

[6] G.N. MILSHTEIN. A method of second-order accuracy integration of stochastic differential equations. *Theory of Probability and Applications*, 23, 1976.

[7] G.N. MILSHTEIN. Weak approximation of solutions of systems of stochastic differential equations. *Theory of Probability and Applications*, 30:750–766, 1985.

[8] N.J. NEWTON. Variance reduction for simulated diffusions. *SIAM Journal of Applied Mathematics*, to appear.

[9] D. NUALART. *Malliavin Calculus and Related Topics*. To appear.

[10] E. PARDOUX. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. In *Cours à l'Ecole d'Eté de Probabilités de Saint-Flour XIX*, number 1464 in Lecture Notes in Mathematics. Springer-Verlag, 1991.

[11] P. PROTTER and D. TALAY. The Euler scheme for Lévy driven stochastic differential equations. Submitted for publication, 1995.

[12] D. TALAY. Efficient numerical schemes for the approximation of expectations of functionals of S.D.E. In J. Szpirglas H. Korezlioglu, G. Mazziotto, editor, *Filtering and Control of Random Processes*, volume 61 of *Lecture Notes in Control and Information Sciences*. Proceedings of the ENST-CNET Colloquium, Paris, 1983, Springer-Verlag, 1984.

[13] D. TALAY. Discrétisation d'une E.D.S. et calcul approché d' espérances de fonctionnelles de la solution. *Mathematical Modelling and Numerical Analysis*, 20(1), 1986.

[14] D. TALAY and L. TUBARO. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):94–120, 1990.