3.Simulation of Stochastic Differential Systems

Denis Talay

Abstract[1]

We present approximation methods for quantities related to solutions of stochastic differential systems, based on the simulation of time-discrete Markov chains. The motivations come from Random Mechanics and the numerical integration of certain deterministic P.D.E.'s by probabilistic algorithms.

We state theoretical results concerning the rates of convergence of these methods. We give results of numerical tests, and we describe an application of this approach to an engineering problem (the study of stability of the motion of a helicopter blade).

1 Introduction

Let us consider a differential system in $\mathbb{R}^d$ excited by a r-dimensional multiplicative random noise $(\xi(t,\omega))$:

$$\frac{d}{dt}X(t,\omega) = b_0(X(t,\omega)) + \sigma(X(t,\omega))\xi(t,\omega) \tag{1}$$

where $b_0(\cdot)$ is an application from $\mathbb{R}^d$ to $\mathbb{R}^d$, $\sigma(\cdot)$ is an application from $\mathbb{R}^d$ to the space of $d \times r$-matrices, and ω denotes the random parameter (in the sequel, we will omit it).

The characteristics of the random noise (bandwidth, energy, law, ...) depend on the modelled physical problem.

Here we are essentially interested in the white-noise case, which is a limit case of systems with "physically realizable" random perturbations (cf. Kushner [32] e.g.); in the section devoted to the computation of Lyapunov exponents, we will examine also systems with coloured noises.

The aim of this paper is to present efficient numerical methods to compute certain quantities depending on the unknown process $(X(t))$, with algorithms based on simulations on a computer of other processes.

We will try to justify this approach (at least, to explain why it may be interesting), we also will underline its limitations; we will estimate the theoretical errors of our approximations, and we will give the results of some illustrative numerical experiments; we will describe an application to an engineering problem (a study of stability for the motion of a helicopter blade) and, finally, we will describe PRESTO, a system of automatic generation of Fortran programs corresponding to the different problems and methods presented here.

[1] Reprint of Chapter 3 in "PROBABILISTIC METHODS IN APPLIED PHYSICS", P. Kree and W. Wedig (Eds), Lecture Notes in Physics 451, Springer-Verlag, 1995.

Before going on, it must be emphasized that the numerical analysis of stochastic differential systems is at its very beginning: at our knowledge, at the present time only a few algorithms have been proposed (some of them irrealistic ...), and only a few systematic numerical investigations have been pursued. Besides, the theoretical results are not very numerous. Nevertheless, it already appears that this field is not at all a direct continuation of what has been done for the numerical solving of ordinary differential equations. For example, we will underline that **it is often unuseful and even clumsy to try to approximate the diffusion process on the space of trajectories, when one wants to compute a quantity which depends on its law**: approximate processes efficient for simulations may not converge almost surely to the considered diffusion process.

In any case, this paper treats a very few approximation problems, and we have chosen to present only algorithms which have been studied from a theoretical **and** numerical point of view. Therefore, this paper must be read as a subjective description of the present state of a new art, and also as a hope that numerical problems which can be efficiently solved by probabilistic algorithms will justify and cause new developments; in particular, recent results related to Random Mechanics (Arnold & Kloeden [4], Schenk [49] e.g.) or related to the numerical integration of certain deterministic nonlinear P.D.E.'s by stochastic particles methods (see Bossy & Talay ([10], [11]), Bossy [9], Bernard & Talay & Tubaro [7] e.g.), show that the mathematical or numerical techniques developed to establish some of the results stated below are useful in various contexts.

For complements and variations on the themes of this paper, one can also read the contributions to the volume [14], and consult the extended list of references in Kloeden & Platen's book [30].

The book by Bouleau & Lepingle [13] presents the various mathematical tools necessary to construct and analyse the numerical methods of approximation of a wide class of stochastic processes.

2 Examples of applications and objectives

2.1 Preliminaries

For the theoretical prerequisites, we refer to the basic book of Arnold [1], or the books of Ikeda & Watanabe [28] and Karatzas & Shreve [29] for example. Here we will just briefly introduce some very elementary concepts, which explain the construction of the discretization schemes.

From a mathematical point of view, one must first give a sense to the limit system of systems of type (1) when $(\xi(t))$ tends to a white noise. The answer is provided by the stochastic calculus; the limit system (in a sense we do not precise here) is a stochastic differential system in the Stratonovich sense (cf. Kushner [32]).

Let us consider r independent Wiener processes, $(W^i(t))$, i.e of Gaussian processes with almost surely continuous trajectories, such that

$$E(W^i(t)) = 0 \ ,$$

$$E(W^i(s)W^j(t)) = \delta_{ij} \inf(s,t) \quad .$$

The limit system is written in the Stratonovich sense as follows:

$$dX(t) = b_0(X(t))dt + \sigma(X(t)) \circ dW(t) \quad , \tag{2}$$

equivalent to the "integral" formulation:

$$X(t) = X(0) + \int_0^t b_0(X(s))ds + \int_0^t \sigma(X(s)) \circ dW(s) \quad .$$

Almost surely, the trajectories of the Wiener process have unbounded variations on each finite time interval (this implies that they are nowhere differentiable), therefore the integral $\int_0^t \sigma(X(s)) \circ dW(s)$ cannot be defined as a Stieljes integral. Let us give indications on its construction.

A process $(Y(t))$ is said adapted to the filtration generated by the Wiener process $(W(t))$ (we will also simply say "adapted") if, for each t, $Y(t)$ is measurable w.r.t. the σ-field generated by $(W(s), s \leq t)$; in particular, $Y(t)$ is independent of all the

$$(W(t_1) - W(t), W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})) \quad ,$$

for any n and $t < t_1 < t_2 < \ldots < t_n$.

Let us consider a one-dimensional Wiener process $(B(t))$. For the class $\mathcal{Q}$ of real continuous adapted (to the filtration generated by $(B(t))$) processes $(Y(t))$ which can be represented as $Y(t) = Y(0) + M(t) + A(t)$, where $Y(0)$ is a r.v., $(M(t))$ is a continuous locally square integrable martingale relative to the previous filtration, and $(A(t))$ is a continuous adapted process of bounded variation on every finite time interval, one can show that for any $T > 0$, the following limit in probability exists:

$$\lim_{|\Delta| \to 0} \sum_{i=1}^{n} \frac{Y(t_i) + Y(t_{i-1})}{2} (B(t_i) - B(t_{i-1})) \quad , \tag{3}$$

where Δ denotes a partition $0 = t_0 < t_1 < \ldots < t_n = T$, and $|\Delta|$ denotes $\max_{1 \leq i \leq n}(t_i - t_{i-1})$.

This limit is called the Stratonovich integral of $(Y(t))$ w.r.t. $(B(t))$ on $[0,T]$ and denoted $\int_0^T Y(s) \circ dB(s)$. If $(Y(t))$ is a matrix-valued process, and the Wiener process multi-dimensional, the integral is defined coordinate by coordinate.

If b_0 and σ are continuous functions such that each component is twice continuously differentiable with bounded derivatives of first and second orders, then for each Borel probability measure μ on $\mathbb{R}^d$, there exists a process $(X(t))$ in $\mathcal{Q}$, satisfying (2) and such that the law of $X(0)$ coincides with μ, unique in the sense that, if $(Y(t))$ is another solution with $Y(0) = X(0)$ *a.s.*, then for each t, $X(t) = Y(t)$ *a.s.*

We will see that the discretization of a Stratonovich integral leads to some difficulties, which do not exist for another stochastic integral, the Itô stochastic integral.

To simplify, we restrict ourselves to consider those of the previous processes $(Y(t))$ which also satisfy:

$$\forall t > 0 \ , \ E\left[\int_0^t |Y(s)|^2 ds\right] < \infty \ .$$

One can show that the following limit exists in the space of the square integrable random variables:

$$\lim_{|\Delta|\to 0} \sum_{i=1}^{n} Y(t_i)(B(t_i) - B(t_{i-1})) \ ,$$

where Δ denotes a partition $0 = t_0 < t_1 < \ldots < t_n = t$, and $|\Delta|$ denotes $\max_{1\leq i\leq n}(t_i - t_{i-1})$.

This limit is called the Itô integral of $(Y(t))$ w.r.t. $(B(t))$, and is denoted by $\int_0^t Y(s)dB(s)$. Under our hypothesis on $(Y(t))$, the process $(\int_0^t Y(s)dB(s))$ is a square integrable martingale, which is often used in the proofs of most approximation theorems stated below.

Moreover, under the above assumptions on $b_0(\cdot)$ and $\sigma(\cdot)$, one can also show that the solution of the Stratonovich system (2) is the unique solution of:

$$X(t) = X(0) + \int_0^t b(X(s))ds + \int_0^t \sigma(X(s))dW(s)$$

where, if σ_j denotes the j^{th} column of σ and $\partial\sigma_j$ the matrix whose element of the i^{th} row and k^{th} column is $\partial_k \sigma_j^i$:

$$b(\cdot) = b_0(\cdot) + \frac{1}{2}\sum_{j=1}^{r} \partial\sigma_j(\cdot)\sigma_j(\cdot) \ .$$

In differential notations, the previous equation is written under the form of an Itô stochastic differential system:

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t) \ . \tag{4}$$

As a consequence of the definition of the Itô integral, it appears that the differential chain rule is different from the deterministic case: for any real function of class $\mathcal{C}^2$, we have the following formula (the Itô formula):

$$df(X(t)) = Lf(X(t))dt + \sigma(X(t))\nabla f(X(t)) \cdot dW(t) \ , \tag{5}$$

where, if the matrix a is $a = \sigma\sigma^*$, L is the following differential operator:

$$L = \sum_{i=1}^{d} b^i(x)\partial_i + \frac{1}{2}\sum_{i,j=1}^{d} a_j^i(x)\partial_{ij} \ . \tag{6}$$

Let us now briefly present the situations that we will treat.

2.2 Simulation of trajectories

We are supposed to have at our disposal exact or approximate trajectories of the Wiener process $(W(t))$, and we want to "see" the corresponding approximate trajectories of $(X(t))$. As we will show later on, when the dimension of the noise is larger than 1 and only approximate trajectories of $(W(t))$ are available, this problem has a signification under a stringent condition which must be fulfilled by $\sigma(\cdot)$ and will be called in the sequel the "commutativity condition" (its precise formulation will be given below).

Let us give 2 examples of situations where one may wish to get trajectories of the solution of a S.D.E.

First, let us suppose that the process $(X(t))$ depends on a parameter θ, that one wants to estimate from a unique observation of $(X(t))$ during a time interval $[0,T]$.

In order to test the quality of different estimators, one may choose a particular value for θ, simulate a few trajectories of the corresponding process $(X(t))$, and then apply the estimators on these simulated trajectories. For applications to financial models, see Fournié and Talay [23] and Fournié [22], e.g.

A less elementary example is a filtering situation, where $(X(t))$ is a non observed process solution of (2), whereas one observes realizations of a process $(Y(t))$ satisfying:

$$dY(t) = g(X(t))dt + \alpha dW(t) + \beta dV(t) \quad ,$$

where $(V(t))$ is a Wiener process independent of $(W(t))$; one wants to get the conditional law of $(X(t))$, given the observations $(Y(s), 0 \leq s \leq t)$. In the non-linear case, under some regularity assumptions on the functions b_0, σ, g, the answer is given by the so-called Zakai equation satisfied by $p(t, x_1, \ldots, x_d)$, the unnormalized density of this conditional law:

$$dq(t) = Aq(t)dt + B_0 q(t)dt + B_1 q(t)dt + C_0 q(t) \circ dY(t) + C_1 q(t) \circ dY(t) \quad ,$$

where A is a second-order operator, B_1 and C_1 are first-order operators, B_0 and C_0 are zero order operators.

The previous equation is a stochastic partial differential equation. To solve it numerically, Florchinger and Le Gland [19] propose the following algorithm.

Let (t_p) be a dicretization of the time interval $[0,t]$, and $\overline{q}$ be the approximate density. On each time interval, $[t_p, t_{p+1}]$, one first numerically solves the deterministic P.D.E.

$$\begin{aligned} \frac{d}{dt}u(t) &= Au(t) \quad , \\ u(t_p) &= \overline{q}(t_p) \quad , \end{aligned}$$

and then one considers the stochastic P.D.E.

$$\begin{aligned} dv(t) &= B_0 v(t)dt + B_1 v(t)dt + C_0 v(t) \circ dY(t) + C_1 v(t) \circ dY(t) \quad , \\ v(t_p) &= u(t_{p+1}) \quad . \end{aligned}$$

Let us write the operators B_1 and C_1 under the form: $B_1 = b_1(x)\nabla$ and $C_1 = c_1(x)\nabla$.

Let $(Z(t;s,x))$ the flow associated to the stochastic differential equation:

$$dZ(t) = -b_1(Z(t))dt - c_1(Z(t)) \circ dY(t) \quad . \tag{7}$$

Then one computes the value of $v(t,x)$ at points $Z(t;t_p,z)$ according to the formula (d_1 and d_2 being appropriate functions):

$$v(t,Z(t;t_p,z)) = v(t_p,z)\exp\left\{\int_{t_p}^{t} d_1(Z(s;t_p,z))ds + \int_{t_p}^{t} d_2(Z(s;t_p,z)) \circ dY(s)\right\} \quad ,$$

and $\overline{q}(t_{p+1})$ is given by $\overline{q}(t_{p+1}) = v(t_{p+1})$.

This procedure requires to solve (7) in a pathwise sense: one wants to get the path of $(Z(t))$ corresponding to the particular observed path of $(Y(t))$.

2.3 Computation of statistics of $(X(t))$ on a finite time interval

For example, one wants to compute the first moments of the response of the dynamical system $(X(t))$, or, more generally, $Ef(X(t))$, $f(\cdot)$ being an explicitly given function.

Another motivation is to construct Monte Carlo methods to solve parabolic P.D.E.'s in $\mathbb{R}^d$

$$\begin{cases} \frac{d}{dt}u(t,x) = Lu(t,x) \quad , \\ u(0,x) \quad = f(x) \quad , \end{cases}$$

in some situations where deterministic methods are not efficient: the theoretical accuracy and the numerical behaviour of the probabilistic algorithms are not affected by the possible non coercivity of the second-order elliptic differential operator L, and the computational cost growths only linearly w.r.t. the dimension d of the state space.

Therefore these methods and the stochastic particles methods (random vortex methods for the integration of certain non-linear P.D.E.'s in Fluid Mechanics e.g.) which also require the simulation of stochastic processes (see the references given at the end of the Introduction) can be useful in degenerate situations or when the state space has a large dimension; in Random Mechanics, often $(X(t))$ is a vector (position, speed), and therefore both degeneracy and a high dimensional state space occur. Other examples are the situations where $u(t,x)$ needs to be computed only at a small number of points, for example in order to separate the integration space in subdomains where deterministic methods become efficient.

To compute $Ef(X(t))$, if we could simulate the process $(X(t))$ itself, we would simulate several independent paths of $(X(t))$, denoted by $(X(t,\omega_1),\ldots,X(t,\omega_N))$ and then we would compute the average

$$\frac{1}{N}\sum_{i=1}^{N} f(X(t,\omega_i)) \quad .$$

Instead of $(X(t))$, we propose to simulate another process $(\overline{X}(t))$; as we are interested in the approximation of the law of $(X(t))$, it is unnecessary that $(\overline{X}(t))$ is a trajectorial approximation of $(X(t))$, and the "commutativity condition" will not be required (better, some efficient processes $(\overline{X}(t))$ in that context are not at all approximations of $(X(t))$ in the pathwise sense, and even do not converge almost surely to $(X(t))$.

2.4 Asymptotic behaviour of $(X(t))$, Lyapunov exponents

In the section 6.1, we present an industrial problem leading to the study of a bilinear system for which it can be shown that, $(X(t,x))$ denoting the solution of (2) with initial condition x, the almost-sure limit

$$\lambda = \lim_{t \longrightarrow +\infty} \frac{1}{t} \log |X(t,x)|$$

exists and is independent of x (it is the upper Lyapunov exponent of the system); the problem is to determine the sign of that limit: if it is strictly negative, almost surely $(X(t))$ tends to 0 exponentially fast for any initial condition x, the system (2) is then said to be stable.

The proposed algorithm consists in simulating one particular path of a process $(\overline{X}(t))$ over a long time $[0,T]$ and in computing

$$\overline{\lambda}_T = \frac{1}{T} \log |\overline{X}(T,x)| \quad .$$

We will classify different processes $(\overline{X}(t))$ according to the following criterium: how large is

$$|\lambda - \overline{\lambda}|$$

where $\overline{\lambda}$ is the Lyapunov exponent of the process $(\overline{X}(t))$, defined by

$$\begin{aligned} \overline{\lambda} &= \lim_{T \to \infty} \overline{\lambda}_T \\ &= \lim_{T \to \infty} \frac{1}{T} \log |\overline{X}(T,x)| \quad . \end{aligned}$$

Remark: this criterium does not take into account the error due to the necessary approximation of $\overline{\lambda}$ by $\overline{\lambda}_T$, which only depends on the choice of the integration time T. We will see that, from a practical point of view, this choice may be very difficult.

An extension of the method has been developed and analysed for nonlinear systems.

2.5 Computation of the stationary law

Under some conditions on the coefficients $b_0(\cdot)$ and $\sigma(\cdot)$, one a priori knows that the process $(X(t))$ is ergodic. Let us denote by μ its unique invariant probability law.

One may be interested in computing $\int f(x)d\mu(x)$ for a given function f (for example, in order to get the asymptotic value of $Ef(X(t))$ when t goes to infinity, i.e to describe the stationary distribution of the response of the system, which often is of prime importance in Random Mechanics). For reasons already underlined, the numerical solving of the stationary Fokker-Planck equation may be extremely difficult.

Here, we propose to choose T "large enough" and to compute

$$\frac{1}{T}\int_0^T f(\overline{X}(s,x))ds \quad .$$

2.6 Remark

As it has been mentioned above, it will appear that the choice of the convenient process $(\overline{X}(t))$ must be related to the final purpose of the simulation.

Our basic tool to construct this process is the time discretization of the system (4).

3 Discretization methods

3.1 Introduction to the Milshtein scheme

Let us consider the expression

$$X(t) = X(0) + \int_0^t b_0(X(s))ds + \int_0^t \sigma(X(s)) \circ dW(s) \quad .$$

From the definition (3) of the Stratonovich integral, for small t the integral $\int_0^t \sigma(X(s)) \circ dW(s)$ can be approximated by

$$\frac{1}{2}(\sigma(X(t)) + \sigma(X(0)))(W(t) - W(0)) \quad ,$$

and therefore this procedure would lead to an implicit discretization scheme.

According to the definition of an Itô integral, a rough approximation of $(X(t))$ (for a small t) can be:

$$X(t) \simeq X(0) + b(X(0))t + \sigma(X(0))(W(t) - W(0)) \quad .$$

Let h be a discretization step.

The above remark justifies the Euler scheme for (4):

$$\overline{X}^h_{p+1} = \overline{X}^h_p + b(\overline{X}^h_p)h + \sigma(\overline{X}^h_p)(W((p+1)h) - W(ph)) \quad . \tag{8}$$

This scheme can easily be simulated on a computer: at each step p, one has just to simulate the vector $W((p+1)h) - W(ph)$, whose law is Gaussian.

As we will see, nevertheless this scheme may be unsatisfying: for example, it is divergent for the pathwise approximation of $(X(t))$.

Let us introduce a new scheme, and first consider the case $d = r = 1$.

If we perform a Taylor expansion of $\sigma(X(t))$, we easily get:

$$X(t) \simeq X(0) + b(X(0))t + \sigma(X(0))(W(t) - W(0)) + \sigma(X(0))\sigma'(X(0)) \int_0^t (W(s) - W(0))dW(s) \quad .$$

At a first glance, the situation is more complex than previously, because of the presence of the stochastic integral $\int_0^t (W(s) - W(0))dW(s)$. But the Itô formula shows:

$$\int_0^t (W(s) - W(0))dW(s) = \frac{1}{2}(W(t)^2 - t) \quad .$$

Then, again only Gaussian laws are involved in the previous scheme, due to Milshtein [37] who introcuded it in 1974 for the mean-square approximation of $(X(t))$.

3.2 The multi-dimensional Milshtein scheme

Let us now examine the general case.

Let us introduce the notation

$$\Delta_{p+1}^h W := W((p+1)h) - W(ph) \quad .$$

The same procedure as before leads to the multi-dimensional Milshtein scheme:

$$\overline{X}_{p+1}^h = \overline{X}_p^h + \sum_{j=1}^r \sigma_j(\overline{X}_p^h)\Delta_{p+1}^h W^j + b(\overline{X}_p^h)h + \sum_{j,k=1}^r \partial\sigma_j(\overline{X}_p^h)\sigma_k(\overline{X}_p^h) \int_{ph}^{(p+1)h} (W^k(s) - W^k(ph))dW^j(s) \quad . \quad (9)$$

Now, the situation is really complex, because of the presence of the multiple stochastic integrals $\int_{ph}^{(p+1)h} (W^k(s) - W^k(ph))dW^j(s)$: these integrals do not depend continuously on the trajectories of $(W(t))$ (therefore are annoying for the trajectorial approximation), and the joint law of these integrals and the increments $\Delta_{p+1}^h W$ seems difficult to simulate: in particular, it cannot be seen as the law of a simple transformation of a Gaussian vector (see the work by Gaines & Lyons [25]).

How to get rid of this difficulty is one of the main features of the numerical analysis of stochastic differential systems.

3.3 Mean-square approximation and Taylor formula

Milshtein [37] proved the following result:

Theorem 1. *Let us suppose that the functions b and σ are of class $\mathcal{C}^2$, with bounded derivatives of first and second orders.*

Then the Euler scheme satisfies: for any integration time T, there exists a positive constant $C(T)$ such that, for any step-size h of type $\frac{T}{n}$, $n \in \mathbb{N}$:

$$\left[E|X(T) - \overline{X}_n^h)|^2\right]^{\frac{1}{2}} \leq C(T)\sqrt{h} \quad .$$

For the Milshtein scheme, we can substitute the following bound for the error:

$$\left[E|X(T) - \overline{X}_n^h)|^2\right]^{\frac{1}{2}} \leq C(T)h \quad .$$

It can be shown that the Milshtein scheme is not "asymptotically efficient" in the sense that the leading coefficient in the expansion of the mean square error as power series in h is not the smallest possible. Clark [17] and Newton ([40], [41] and [42]) have introduced new schemes which are asymptotically efficient; these schemes may be seen as versions of the Milshtein scheme with additional terms of order $h\Delta_{p+1}^h W$ and $(\Delta_{p+1}^h W)^3$ (when the Wiener process is scalar). In the two last references, efficient schemes based upon first passage times of the Wiener samples through given points, and efficient Runge-Kutta schemes are presented. Another very interesting approach can be found in Castell & Gaines [16], based upon the representation of diffusion processes in terms of the solutions of ordinary differential equations.

Let us now examine the question of the order of convergence.

Let us say that a random variable is of order k if its variance is upper bounded by $Constant \times h^{2k}$.

The Milshtein scheme involves only random variables of order less than 1.

To get a better rate of convergence in the mean-square sense than this scheme, one must involve multiple stochastic integrals of order strictly larger than 1, for example:

$$\int_{ph}^{(p+1)h} (W^k(s) - W^k(ph))(W^l(s) - W^l(ph))dW^j(s) \quad ,$$

$$\int_{ph}^{(p+1)h} (s - ph)dW^j(s) \quad , \quad \int_{ph}^{(p+1)h} (W^k(s) - W^k(ph))ds \quad ,$$

in order to get an error of order $h^{\frac{3}{2}}$.

The coefficients of these integrals in the schemes are given by a Taylor formula (see Platen & Wagner [47]).

Of course, most of these integrals, as those involved in the multi-dimensional Milshtein scheme, have probability laws difficult to simulate (see Gaines [24]). Therefore, in the general case, the Euler scheme is the only efficient scheme for the mean-square approximation.

Nevertheless, there exists a situation where the multi-dimensional Milshtein scheme involves only the increments of the Wiener process $\Delta_{p+1}^h W$.

3.4 The commutativity condition

Suppose that the column vectors of the matrix σ satisfy the following condition:

$$\forall j, \forall k \quad : \quad \partial\sigma_j(\cdot)\sigma_k(\cdot) = \partial\sigma_k(\cdot)\sigma_j(\cdot) \quad . \tag{10}$$

That condition means that the vector fields defined by the column vectors of σ commute. It is obviously satisfied when the noise is one-dimensional, or when the function σ is constant.

The Itô formula and this hypothesis imply:

$$\begin{aligned}
&\partial\sigma_k(\cdot)\sigma_j(\cdot)\int_{ph}^{(p+1)h}(W^k(s)-W^k(ph))dW^j(s)\\
&+\partial\sigma_j(\cdot)\sigma_k(\cdot)\int_{ph}^{(p+1)h}(W^j(s)-W^j(ph))dW^k(s)\\
&=\partial\sigma_k(\cdot)\sigma_j(\cdot)(W^k((p+1)h)-W^k(ph))(W^j((p+1)h)-W^j(ph)) \quad , \quad (11)
\end{aligned}$$

and therefore the Milshtein scheme can be rewritten:

$$\begin{aligned}
\overline{X}^h_{p+1} &= \overline{X}^h_p + \sum_{j=1}^r \sigma_j(\overline{X}^h_p)\Delta^h_{p+1}W^j + b(\overline{X}^h_p)h\\
&\quad + \sum_{k=2}^r \sum_{j<k} \partial\sigma_j(\overline{X}^h_p)\sigma_k(\overline{X}^h_p)\Delta^h_{p+1}W^j\Delta^h_{p+1}W^k\\
&\quad + \frac{1}{2}\sum_{j=1}^r \partial\sigma_j(\overline{X}^h_p)\sigma_j(\overline{X}^h_p)\left[\left(\Delta^h_{p+1}W^j\right)^2 - h\right] \quad . \qquad (12)
\end{aligned}$$

A very nice result due to Clark & Cameron [18] shows that, under the commutativity condition, the Milshtein scheme leads to the best possible rate of convergence for the mean-square approximation (i.e h) among all the discretization schemes involving only values of the process $(W(t))$ at times $(ph, 0 \leq p \leq n = \frac{T}{h})$.

4 Almost sure and pathwise approximation

4.1 Statement of the problems

First, let us suppose that we observe or simulate increments of the Wiener process during time intervals of length h; then we construct a continous time process $(\overline{X}(t))$ by using the Euler scheme and by interpolating linearly between the times ph. Does this process converge almost surely to $(X(t))$ on finite time intervals when h goes to 0 ?

Second, let us now suppose that we dispose of a deterministic function $t \to u(t)$ which approximates a given trajectory of $(W(t))$ in the sense of the topology of uniform convergence on the space of continuous functions on $[0, T]$.

We hope to get an approximation of the trajectory of $(X(t))$ on $[0,T]$ corresponding to this particular trajectory of $(W(t))$.

To give a sense to this new problem, a natural condition is that there exists a continuous mapping F from $\mathbb{R}^d \times \mathcal{C}_0(\mathbb{R}_+;\mathbb{R}^r)$ [2] to $\mathcal{C}(\mathbb{R}_+;\mathbb{R}^r)$ such that $X(t) = F(X(0),W)(t)$, a.s.

A result due to McShane [36], Doss [20] and Sussman [51] shows that this mapping exists if the above "commutativity condition" is satisfied.

Now, the problem is to build a scheme defined by functionals ϕ_p^h on the space $\mathbb{R}^d \times \mathcal{C}([0,ph])$:

$$\overline{X}_0^{h,u} = X(0) \quad , \quad \overline{X}_{p+1}^{h,u} = \phi_p^h(\overline{X}_p^{h,u}, (u(t), 0 \le t \le ph))$$

such that: for any entry $(u(t))$ belonging to a large set of functions (including the trajectories of $(W(t))$ and their reasonable approximations), if $x_u(t)$ denotes $F(X(0),u)(t)$, and if $t \to \overline{X}_u^h(t)$ is the function defined by

$$\overline{X}_u^h(t) = \overline{X}_p^{h,u} \quad , \quad ph \le t < (p+1)h \quad ,$$

then

$$\lim_{h \to 0} \sup_{0 \le t \le T} |x_u(t) - \overline{X}_u^h(t)| = 0 \quad .$$

When this property is fulfilled, the scheme is robust w.r.t. small pertubations of the trajectory $u(\cdot)$; we say that the scheme converges in the pathwise (or trajectorial) sense.

4.2 Example

Let $(B(t))$ be a one-dimensional Wiener process. From the Itô formula (5), the process

$$X(t) = \exp(t + B(t))$$

solves the one-dimensional stochastic differential system:

$$dX(t) = \frac{3}{2} X(t)dt + X(t)dB(t) \quad .$$

It is easy to see that the Euler scheme (8) converges almost surely, and converges in the above pathwise sense only if the function $(u(t))$ has the same quadratic variation as the trajectories of the Wiener process.

The situation is different with the Milshtein scheme.

[2] $\mathcal{C}_0(\mathbb{R}_+;\mathbb{R}^r)$ denotes the set of continuous functions f from $\mathbb{R}_+$ to $\mathbb{R}^d$ such that $f(0) = 0$.

4.3 Main results

For the almost sure convergence problem of the Euler problem, a first result appears in Newton ([40] and [42]). More precise statements appear in Faure ([21]), for example:

Theorem 2. *Let us suppose that the coefficients $b(\cdot)$ and $\sigma(\cdot)$ are Lipschitz.*

(i) If for some integer $K > 1$ the initial condition X_0 satisfies $E|X_0|^{2K} < \infty$, then the interpolated Euler scheme with step-size $\frac{T}{n}$, $(\overline{X}^h(t))$, converges almost surely to $(X(t))$ on $[0,T]$ when n goes to infinity.

(ii) If the initial condition has moments of any order, then the order of convergence is given by

$$\forall \alpha < \frac{1}{2} \quad , \quad n^\alpha \sup_{t\in[0,T]} |X(t) - \overline{X}(t)| \overset{n\to+\infty}{\longrightarrow} 0 \quad , \quad a.s.$$

Let us now turn to the trajectorial problem.

In the multidimensional case, the remarkable point is that the commutativity condition, which is necessary to have a well-posed problem, is also sufficient to make the Milshtein scheme (9) depend only on the values of $(W(t))$ at the discretization points (formula (11)). This leads us to introduce the trajectorial Milshtein scheme defined by

$$\begin{aligned}\overline{X}^{h,u}_{p+1} &= \overline{X}^{h,u}_p + \sum_{j=1}^r \sigma_j(\overline{X}^{h,u}_p)\Delta^h_{p+1}u^j + b(\overline{X}^{h,u}_p)h \\ &+ \sum_{k=2}^r \sum_{j<k} \partial\sigma_j(\overline{X}^{h,u}_p)\sigma_k(\overline{X}^{h,u}_p)\Delta^h_{p+1}u^k\Delta^h_{p+1}u^j \\ &+ \frac{1}{2}\sum_{j=1}^r \partial\sigma_j(\overline{X}^{h,u}_p)\sigma_k(\overline{X}^{h,u}_p)\left[\left(\Delta^h_{p+1}u^j\right)^2 - h\right] \quad . \end{aligned} \tag{13}$$

In Talay [52] the following result is proven:

Theorem 3. *Let us suppose that b and σ are bounded, of class $\mathcal{C}^3$ with bounded derivatives up to the order 3, and that the function $(u(t))$ satisfies:*

$$\lim_{|\Delta|\to 0} \sum_i |u(t_i) - u(t_{i-1})|^3 = 0 \quad , \tag{14}$$

where Δ denotes a partition $0 = t_0 < t_1 < \ldots < t_m = t$, and $|\Delta|$ denotes $\max_{1\le i\le m}(t_i - t_{i-1})$.

Then,if $(\overline{X}^h_u(t))$ is defined as in the previous section, under the commutativity condition:

$$\lim_{h\to 0} \sup_{0\le t\le T} |x_u(t) - \overline{X}^h_u(t)| = 0 \quad . \tag{15}$$

Remarks

- The condition (14) is satisfied by the paths of the Wiener process, but also by a much larger class of functions (for example, the differentiable functions).
- The proof of the theorem (3) is based on an analytical expression of the mapping F.
- It is also shown that the Milshtein scheme has the best possible rate of convergence for the criterium (15).
- The commutativity condition is a strong limitation to the trajectorial approximation of the solution of an Itô differential system. But, in some sense, this problem forgets the fact that $(X(t))$ is a stochastic process, whose statistics may be more interesting than some particular paths. We will not need this condition to approximate quantities depending on the law of $(X(t))$.
- The asymptotic distribution of the normalized Euler scheme error is analysed in Kurtz & Protter [31].

4.4 Numerical example

The following numerical test illustrates the divergent behaviour of the Euler scheme for the pathwise approximation.

Let $(X(t))$ the 2-dimensional process defined by

$$X(t) = (sin(W(t)), cos(W(t))) \quad ,$$

where $(W(t))$ is a one-dimensional Wiener process.

This process solves a system with the function $b(x_1, x_2)$ defined by

$$b^1 = -\frac{x^1}{2} \quad ,$$
$$b^2 = -\frac{x^2}{2} \quad ,$$

and the matrix σ is defined by

$$\sigma_{11} = x^2 \quad ,$$
$$\sigma_{21} = -x^1 \quad .$$

We have simulated a trajectory of $(W(t))$, and a perturbation of it: we have simulated a second Brownian trajectory (denoted by $t \to V(t)$), and, for each t, we have added $\varepsilon V(t)$ to $W(t)$.

The figure 1 shows the "exact" path of $(X(t))$ corresponding to the simulated path of $(W(t))$. The figure 2 compares the evolution in time of the errors (in the trajectorial sense) due to the Milshtein scheme (thick line) and the Euler scheme (thin line), corresponding to $\varepsilon = 0.001$ and $h = 0.01$.

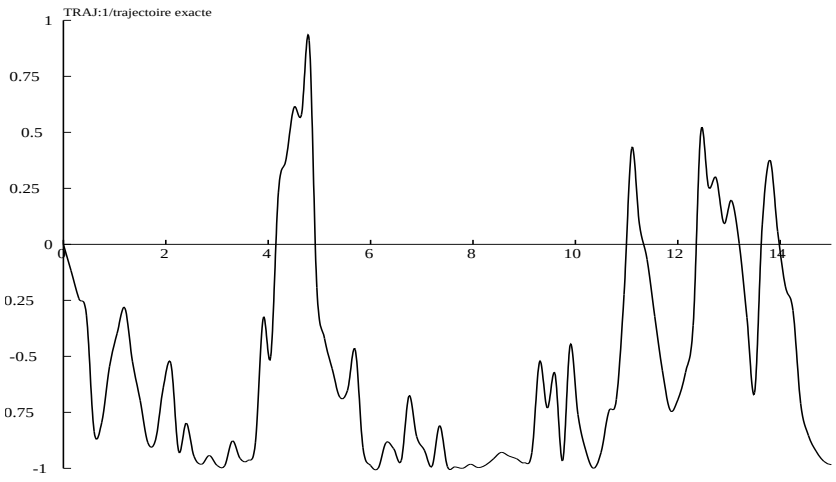

Fig. 1. Exact path.

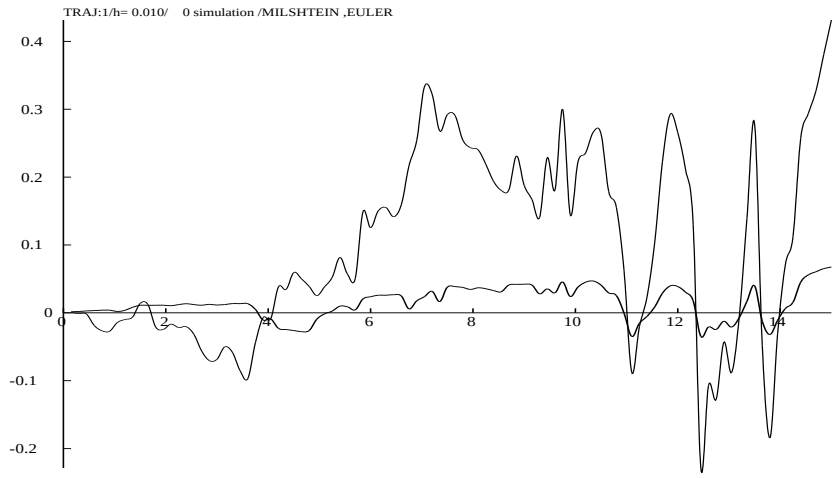

Fig. 2. Milshtein and Euler schemes.

5 Computation of $Ef(X(t))$

5.1 Methodology

Now, we are interested in the approximation of $Ef(X(t))$ on a fixed finite time interval $[0,T]$.

Suppose that the coefficients b and σ are smooth enough, and that the operator L defined in (6) is hypoelliptic; if the law of the initial condition $X(0)$ has a density $p_0(\cdot)$, then for any $t>0$ the law of $X(t)$ has a density $p(t,\cdot)$ solution of the Fokker-Planck equation:

$$\left\{ \begin{array}{ll} \frac{d}{dt}p(t,x) & = L^*p(t,x) \quad , \\ p(0,x) & = p_0(x) \quad , \end{array} \right.$$

where L^* is the adjoint of the differential operator L (see [28] e.g.).

Thus a first method to compute $Ef(X(t))$ consists in integrating the previous P.D.E. But the numerical solving of this P.D.E. can be difficult, for example when the dimension d of the process $(X(t))$ is very large, or when the differential operator L is degenerate (*it very often is the case in problems coming from Mechanics, in particular each time* $(X(t))$ *is a vector (position,velocity)*).

A second method consists in using a Monte Carlo method. Let us begin by choosing the Euler scheme (8), and let us simulate (if possible, in parallel) a large number N of independent realizations of the Gaussian sequence $(\Delta_{p+1}^h W, p \in \mathbb{N})$. Then, for each discretization step, we get N independent realizations of $\overline{X}_p^h$, denoted by $\overline{X}_p^h(\omega_i)$, and we can compute

$$\frac{1}{N}\sum_{i=1}^{N} f(\overline{X}_p^h(\omega_i)) \quad . \tag{16}$$

By the strong law of large numbers, this gives us an approximate value of $Ef(\overline{X}_p^h)$. The quality of this approximation depends only on the choice of N.

It remains to estimate the error $|Ef(X(ph)) - Ef(\overline{X}_p^h)|$. It can be shown that, under some smoothness assumptions on b, σ, if the law of $X(0)$ has moments of any order, then, for any time T, there exists a positive constant $C(T)$ such that, for any discretization step h of type $h=\frac{T}{n}$, $n \in \mathbb{N}$:

$$|Ef(X(T)) - Ef(\overline{X}_n^h)| \leq C(T)h \quad .$$

It can be also shown that, even under the commutativity condition, the Milshtein scheme has the same rate of convergence.

This is illustrated by the following example: choose $d=r=1$, $b(x)=\frac{1}{2}x$, $\sigma(x)=x$ and $f(x)=x^4$. Then, for the Euler or Milshtein scheme, there exist constants C_1, C_2 such that:

$$|Ef(X(T)) - Ef(\overline{X}_n^h)| = C_1 T \exp(C_2 T)h + O(h^2) \quad .$$

Therefore, the Milshtein scheme which is "optimal" in the mean-square sense and in the trajectorial sense is so poor as the Euler scheme for an approximation of the law of $(X(t))$.

The technique introduced in Talay [53] or [54], Milshtein [38] permits to analyse the error on $Ef(X(T))$ without using estimates in L^p of $X(T) - \overline{X}_n^h$. It also permits to construct second-order schemes. A refinement of the analysis leads to a very efficient procedure (see the section 5.6 below).

5.2 Second-order schemes

Let $\mathcal{P}$ be the set of numerical functions of $\mathbb{R}^d$, of class $\mathcal{C}^6$, such that f and its partial derivatives up to order 6 have a growth at most polynomial at infinity.

A scheme is said of second-order if it satisfies for any system whose coefficients b, σ are smooth and have bounded derivatives of any order: for any function f in $\mathcal{P}$, for any time T, there exists a positive constant $C(T)$ such that, for any discretization step h of type $h = \frac{T}{n}$, $n \in \mathbb{N}$:

$$|Ef(X(T)) - Ef(\overline{X}_n^h)| \leq C(T)h^2 \quad . \tag{17}$$

Let $\mathcal{F}_p$ be the σ-algebra generated by $(\overline{X}_0^h, \ldots, \overline{X}_p^h)$.

In Talay [54], it is shown that a sufficient condition for a scheme to satisfy (17) is the set of hypotheses (C1), (C2), (C3):

(C1) $\overline{X}_0^h = X(0)$;

(C2) $\forall n \in \mathbb{N}$, $\forall N \in \mathbb{N}$, $\exists C > 0$, $\forall p \leq N$, $E|\overline{X}_p^h|^n \leq C$;

(C3) the following properties are satisfied for all $p \in \mathbb{N}$, where all the right-side terms of the equalities must be understood evaluated at $\overline{X}_p^h$:

$$E\left(\Delta_{p+1}^h \overline{X} | \mathcal{F}_p\right) = bh + \frac{1}{2}(Lb)h^2 + \xi_{p+1} \quad , \quad E|\xi_{p+1}| \leq Ch^3 \quad ,$$

$$\begin{aligned} E\left((\Delta_{p+1}^h \overline{X})^{i_1}(\Delta_{p+1}^h \overline{X})^{i_2} | \mathcal{F}_p\right) = {} & \sigma_j^{i_1}\sigma_j^{i_2}h + (b^{i_1}b^{i_2} + \frac{1}{2}\partial_{k_1}\sigma_j^{i_1}\partial_{k_2}\sigma_j^{i_2}\sigma_l^{k_1}\sigma_l^{k_2} \\ & + \frac{1}{2}\partial_k b^{i_2}\sigma_j^{i_1}\sigma_j^k + \frac{1}{2}\partial_k b^{i_1}\sigma_j^{i_2}\sigma_j^k \\ & + \frac{1}{2}\sigma_j^{i_1}\partial_k\sigma_j^{i_2}b^k + \frac{1}{2}\sigma_j^{i_2}\partial_k\sigma_j^{i_1}b^k \\ & + \frac{1}{4}\sigma_j^{i_1}\partial_{kl}\sigma_j^{i_2}\sigma_n^k\sigma_n^l + \frac{1}{4}\sigma_j^{i_2}\partial_{kl}\sigma_j^{i_1}\sigma_n^k\sigma_n^l)h^2 \\ & + \xi_{p+1}^{i_1 i_2} \quad , \quad E|\xi_{p+1}^{i_1 i_2}| \leq Ch^3 \quad , \end{aligned}$$

$$\begin{aligned} E\left((\Delta_{p+1}^h \overline{X})^{i_1} \ldots (\Delta_{p+1}^h \overline{X})^{i_3} | \mathcal{F}_p\right) = {} & (b^{i_1}\sigma_j^{i_2}\sigma_j^{i_3} + b^{i_2}\sigma_j^{i_3}\sigma_j^{i_1} + b^{i_3}\sigma_j^{i_1}\sigma_j^{i_2} \\ & + \frac{1}{2}\sigma_l^{i_2}\partial_k\sigma_l^{i_3}\sigma_j^{i_1}\sigma_j^k + \frac{1}{2}\sigma_l^{i_3}\partial_k\sigma_l^{i_2}\sigma_j^{i_1}\sigma_j^k \\ & + \frac{1}{2}\sigma_l^{i_3}\partial_k\sigma_l^{i_1}\sigma_j^{i_2}\sigma_j^k + \frac{1}{2}\sigma_l^{i_1}\partial_k\sigma_l^{i_3}\sigma_j^{i_2}\sigma_j^k \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sigma_l^{i_1}\partial_k\sigma_l^{i_2}\sigma_j^{i_3}\sigma_j^k+\frac{1}{2}\sigma_l^{i_2}\partial_k\sigma_l^{i_1}\sigma_j^{i_3}\sigma_j^k)h^2\\
&+\xi_{p+1}^{i_1i_2i_3}\quad,\quad E|\xi_{p+1}^{i_1i_2i_3}|\le Ch^3\ ,\\
E\left((\Delta_{p+1}^h\overline{X})^{i_1}\dots(\Delta_{p+1}^h\overline{X})^{i_4}|\mathcal{F}_p\right)&=(\sigma_j^{i_1}\sigma_j^{i_2}\sigma_l^{i_3}\sigma_l^{i_4}+\sigma_j^{i_1}\sigma_j^{i_3}\sigma_l^{i_2}\sigma_l^{i_4}\\
&+\sigma_j^{i_1}\sigma_j^{i_4}\sigma_l^{i_2}\sigma_l^{i_3})h^2+\xi_{p+1}^{i_1\dots i_4}\quad,\\
&E|\xi_{p+1}^{i_1\dots i_4}|\le Ch^3\ ,\\
E\left((\Delta_{p+1}^h\overline{X})^{i_1}\dots(\Delta_{p+1}^h\overline{X})^{i_5}|\mathcal{F}_p\right)&=\xi_{p+1}^{i_1\dots i_5}\quad,\quad E|\xi_{p+1}^{i_1\dots i_5}|\le Ch^3\ ,\\
E\left((\Delta_{p+1}^h\overline{X})^{i_1}\dots(\Delta_{p+1}^h\overline{X})^{i_6}|\mathcal{F}_p\right)&=\xi_{p+1}^{i_1\dots i_6}\quad,\quad E|\xi_{p+1}^{i_1\dots i_6}|\le Ch^3\ .
\end{aligned}
$$

An interesting fact happens.

To satisfy the condition $(C3)$, a scheme does not need to involve any stochastic integral (even $\Delta_{p+1}^h W=\int_{ph}^{(p+1)h}dW(s)$!). Very simple random variables may be used, the only requirement is that the law of these random variables must have the same small number of statistics (for example, the same expectation vector, the same correlation matrix, and so on) as a certain finite family of stochastic integrals. Let us see 2 examples.

5.3 Two examples of efficient second-order schemes

Let us introduce the random variables that will be involved in our schemes.

– The sequence

$$(U_{p+1}^j,\tilde{U}_{p+1}^{kj},j,k=1,\dots,r;p\in\mathbb{N})$$

is a family of independent random variables; the (U_{p+1}^j) are i.i.d. and satisfy the following conditions:

$$E[U_{p+1}^j]=E[U_{p+1}^j]^3=E[U_{p+1}^j]^5=0\ , \tag{18}$$

$$E[U_{p+1}^j]^2=1\ , \tag{19}$$

$$E[U_{p+1}^j]^4=3\ , \tag{20}$$

$$E[U_{p+1}^j]^6<+\infty\ ; \tag{21}$$

the $(\tilde{U}_{p+1}^{kj})$ are i.i.d., their common law being defined by

$$P(\tilde{U}_p^{kj}=\frac{1}{2})=P(\tilde{U}_p^{kj}=-\frac{1}{2})=\frac{1}{2}\ .$$

For example, one can choose

$$U_{p+1}^j=\frac{1}{\sqrt{h}}\Delta_{p+1}^h W^j$$

and as well one can choose for U_{p+1}^j the discrete law of mass $\frac{2}{3}$ at 0 and of mass $\frac{1}{6}$ at the points $+\sqrt{3}$ and $-\sqrt{3}$;

– the family (Z_p^{kj}) is defined by

$$Z_{p+1}^{kj} = \frac{1}{2} U_{p+1}^k U_{p+1}^j + \tilde{U}_{p+1}^{kj} \quad , \quad k < j \quad ,$$
$$Z_{p+1}^{kj} = \frac{1}{2} U_{p+1}^k U_{p+1}^j - \tilde{U}_{p+1}^{jk} \quad , \quad k > j \quad ,$$
$$Z_{p+1}^{jj} = \frac{1}{2} \left((U_{p+1}^j)^2 - 1 \right) \quad .$$

Now, a being $a = \sigma\sigma^*$, we define the vectors A_j by

$$A_j = \frac{1}{2} \sum_{k,l=1}^{d} a_l^k \partial_{kl} \sigma_j \quad .$$

We recall that we denote by L the infinitesimal generator of the process $(X(t))$:

$$L = \sum_{i=1}^{d} b^i(x) \partial_i + \frac{1}{2} \sum_{i,j=1}^{d} a_j^i(x) \partial_{ij} \quad .$$

We consider the scheme defined by

$$\begin{aligned} \overline{X}_{p+1}^h = \overline{X}_p^h &+ \sum_{j=1}^{r} \sigma_j(\overline{X}_p^h) U_{p+1}^j \sqrt{h} + b(\overline{X}_p^h) h + \sum_{j,k=1}^{r} \partial\sigma_j(\overline{X}_p^h) \sigma_k(\overline{X}_p^h) Z_{p+1}^{kj} h \\ &+ \frac{1}{2} \sum_{j=1}^{r} \left\{ \partial b(\overline{X}_p^h) \sigma_j(\overline{X}_p^h) + \partial\sigma_j(\overline{X}_p^h) b(\overline{X}_p^h) + A_j(\overline{X}_p^h) \right\} U_{p+1}^j h^{\frac{3}{2}} \\ &+ \frac{1}{2} Lb(\overline{X}_p^h) h^2 \quad . \end{aligned} \tag{22}$$

It has been shown that this scheme is of second order (see Talay [53] or [54], Milshtein [38]). The Taylor formula given in Platen & Wagner [47] helps to understand how it has been constructed.

Another example of second-order scheme is the following "MCRK"scheme of Talay [54], which is of Runge-Kutta type, and therefore may be more interesting than the previous scheme from a numerical point of view, since most derivatives of the coefficients are avoided.

We define the new family

$$(V_{p+1}^j, \tilde{V}_{p+1}^{kj}, j, k = 1, \ldots, r; p \in \mathbb{N})$$

so that the family

$$(U_{p+1}^j, V_{p+1}^j, \tilde{U}_{p+1}^{kl}, \tilde{V}_{p+1}^{mn})_{j,k,l,m,n,p}$$

is a sequence of independent variables, the V_{p+1}^j's having the same distribution as the U_{p+1}^j's, and the $\tilde{V}_{p+1}^{mn}$'s having the same distribution as the $\tilde{U}_{p+1}^{kl}$'s.

Now we define (S_p^{kj}), (T_p^{kj}) and (Z_p^{kj}) by

$$
\begin{aligned}
S_{p+1}^{kj} &= \frac{1}{4}(U_{p+1}^k U_{p+1}^j + \tilde{U}_{p+1}^{kj}) \quad , \quad k < j \quad , \\
S_{p+1}^{kj} &= \frac{1}{4}(U_{p+1}^k U_{p+1}^j - \tilde{U}_{p+1}^{jk}) \quad , \quad k > j \quad , \\
S_{p+1}^{jj} &= \frac{1}{4}\left((U_{p+1}^j)^2 - 1\right) \quad , \\
T_{p+1}^{kj} &= \frac{1}{4}(V_{p+1}^k V_{p+1}^j + \tilde{V}_{p+1}^{kj}) \quad , \quad k < j \quad , \\
T_{p+1}^{kj} &= \frac{1}{4}(V_{p+1}^k V_{p+1}^j - \tilde{V}_{p+1}^{jk}) \quad , \quad k > j \quad , \\
T_{p+1}^{jj} &= \frac{1}{4}\left((V_{p+1}^j)^2 - 1\right) \quad , \\
Z_{p+1}^{kj} &= S_{p+1}^{kj} + T_{p+1}^{kj} + \frac{1}{2}U_{p+1}^k V_{p+1}^j \quad , \quad k \neq j \quad , \\
Z_{p+1}^{jj} &= \frac{1}{4}\left((U_{p+1}^j + V_{p+1}^j)^2 - 2\right) \quad .
\end{aligned}
$$

The MCRK scheme proceeds in 2 steps.
From $\overline{X}_p^h$, one first computes:

$$
\begin{aligned}
\overline{X}_{p+\frac{1}{2}}^h = \overline{X}_p^h &+ \frac{\sqrt{2}}{2}\sigma(\overline{X}_p^h)U_{p+1}^j\sqrt{h} + \frac{1}{2}b(\overline{X}_p^h)h \\
&+ \sum_{j,k=1}^{r} \partial\sigma_j(\overline{X}_p^h)\sigma_k(\overline{X}_p^h)S_{p+1}^{kj}h \quad .
\end{aligned}
$$

Then the new value $\overline{X}_{p+1}^h$ is obtained according to the formula:

$$
\begin{aligned}
\overline{X}_{p+1}^h = \overline{X}_p^h &+ \left[\sigma(\overline{X}_p^h)U_{p+1} + \sigma(\overline{X}_{p+\frac{1}{2}}^h)V_{p+1} - \frac{1}{2}\sigma(\overline{X}_p^h)(U_{p+1} + V_{p+1})\right]\sqrt{2h} \\
&+ b(\overline{X}_{p+\frac{1}{2}}^h)h \\
&+ \sum_{j,k=1}^{r}\Big[2\partial\sigma_j(\overline{X}_p^h)\sigma_k(\overline{X}_p^h)S_{p+1}^{kj}h + 2\partial\sigma_j(\overline{X}_{p+\frac{1}{2}}^h)\sigma_k(\overline{X}_{p+\frac{1}{2}}^h)T_{p+1}^{kj}h \\
&- \partial\sigma_j(\overline{X}_p^h)\sigma_k(\overline{X}_p^h)Z_{p+1}^{kj}h\Big] \quad . \qquad (23)
\end{aligned}
$$

Other examples of Runge-Kutta schemes are studied for the quadratic mean error, in Rumelin [48] and in Newton [42].

5.4 Remarks

– The law of the family $(\sqrt{h}U_{p+1}^j, hS_p^{kj}, hT_p^{kj}, hZ_p^{kl})$ can be chosen in various ways. A precise formulation of families such that the above schemes are of second order is given by the notion of "Monte Carlo equivalence" in Talay [54].

The idea is that the law of the family must only have a small number of properties; in particular, these properties imply that the expectation vector and the correlation matrix of

$$(\sqrt{h}U_{p+1}^j, hS_{p+1}^{kj})$$

and

$$\left(\Delta_{p+1}^h W^j, \int_{ph}^{(p+1)h} (W^k(s) - W^k(ph))dW^j(s)\right)$$

differ only by terms of order h^3.

- As already mentioned, the above schemes may diverge for the almost sure and pathwise approximations (even under the commutativity condition), in particular if one chooses for the U_{p+1}^j's a discrete law.
- The choice of the number of the independent realizations to simulate in order to perform the Monte Carlo computation depends on the wished accuracy. In practice, one may first roughly estimate the maximal value of the variance of $f(\overline{X}_p^h)$ on the considered time interval, by simulating a small number N_0 of sample paths of $(\overline{X}_p^h)$; then, one chooses N according to the central-limit theorem, and finally one simulates $N - N_0$ other samples to approximate $E[f(\overline{X}_p^h)]$ with a better accuracy.
  Of course, this is the critical point of the procedure: the Monte Carlo algorithms converge slowly, and in practice N must be large, especially when the variance of $f(X(t))$ increases with t. If a particular discretization method permits to obtain a variance reduction is an extremely difficult question, deeply examined in a recent work by N. Newton [43], based on Haussmann's integral representation for functional of Itô processes (see also, for a different approach, Wagner [59]).

5.5 Numerical experiments

We now show that the numerical performances of the above different schemes may be extremely surprizing: the Euler scheme may be more efficient than second-order schemes (as for the ordinary differential equations, this will be explained by an expansion of the discretization error as a power series with respect to the discretization step).

First, let us consider an example where the second-order schemes are much better than the Euler and Milshtein schemes.

The processes $(X(t))$ and $(W(t))$ are one-dimensional, and $X(t) = atan(Z(t))$, where $(Z(t))$ is a stationary $\mathcal{N}(0,1)$ Ornstein-Uhlenbeck process solution of

$$dZ(t) = -Z(t)dt + \sqrt{2}dW(t)$$

so that $(X(t))$ solves a stochastic differential equation whose coefficients are:

$$b(x) = -\frac{1}{4}\sin(4x) - \sin(2x) \quad , \quad \sigma(x) = \sqrt{2}\cos^2(x) \quad .$$

We compute $E(\cosh(1.3X(t)+2)) \sim 5.36168895$.

The figure 3 compares the evolution in time of the errors due to the Euler scheme (thin line) and the second-order scheme (22) (thick line). The figure 4 compares the evolution in time of the errors due to the Milshtein scheme (thin line) and the second-order scheme (22) (thick line).

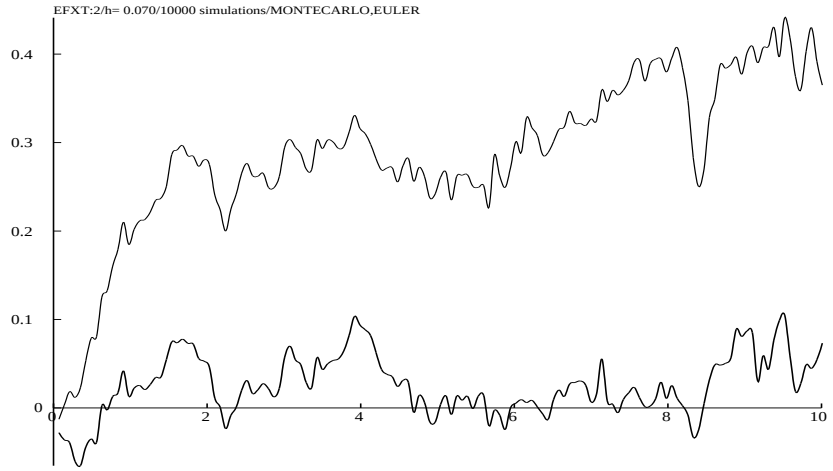

Fig. 3. Euler and second-order schemes.

But, in the next example, a strange fact occurs: the Euler scheme gives as good results as the second-order schemes, whereas the Milshtein scheme may give very bad results.

The function $b(\cdot)$ is defined by

$$b^1 = (3\sqrt{2}x^1 + 6\sqrt{2}x^2 - 2\sin(\Omega t)\Omega x^1 - 12x^1 - 6x^2)/(4(\cos(\Omega t)+2)) \quad ,$$
$$b^2 = (3\sqrt{2}x^2 + 6\sqrt{2}x^1 - 2\sin(\Omega t)\Omega x^2 - 12x^2 - 6x^1)/(4(\cos(\Omega t)+2)) \quad ,$$

and the matrix σ is defined by

$$\sigma_1^1 = \sin\left(\nu\left(x^1+x^2\right)\right) \quad ,$$
$$\sigma_1^2 = \cos\left(\nu\left(x^1+x^2\right)\right) \quad ,$$
$$\sigma_2^1 = \sin\left(\frac{\pi + 3\nu x^1 + 3\nu x^2}{3}\right) \quad ,$$
$$\sigma_2^2 = \cos\left(\frac{\pi + 3\nu x^1 + 3\nu x^2}{3}\right) \quad .$$

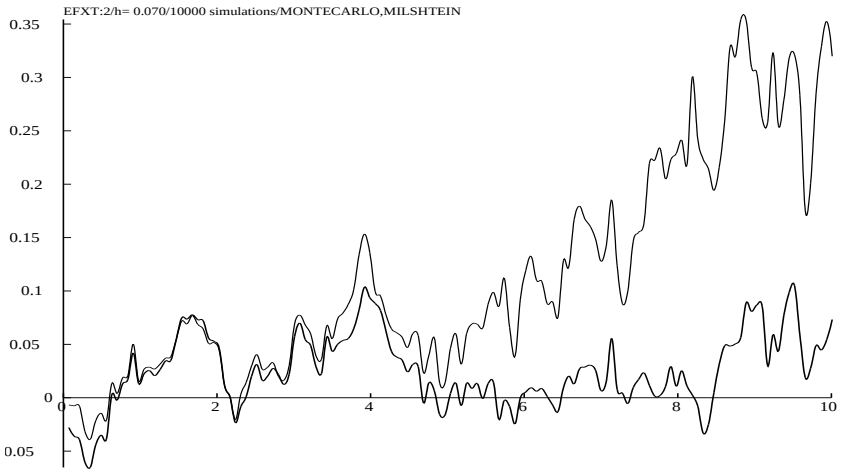

Fig. 4. Milshtein and second-order schemes.

One can check that, if the initial law is Gaussian with zero mean and a covariance matrix equal to

$$C = \begin{bmatrix} 1 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{bmatrix}$$

then the law of $X(t)$ is also Gaussian with zero mean and a covariance matrix equal to $\frac{2+cos(\Omega t)}{3} C$.

Let us fix $\nu = 2$, $\Omega = 5$, and the number of simulations $N = 10,000$.

The figure 5 shows the time evolution of the true value (thick line) of $E|X^1(t)|^2$ and of the approximate value corresponding to the Euler scheme (thin line): the approximation error is weak.

The figure 6 compares the time evolution of the errors due to the Euler scheme (thin line) and the second-order scheme (22) (thick line): these errors are similar.

The figure 7 compares the time evolution of the errors due to the Euler scheme (thin line) and the Milshtein scheme (thick line). These two schemes have identical theoretical convergence rates, but, in that particular situation, the Milshtein scheme leads to absurd results.

Numerical experiments have shown that this bad behaviour of the Milshtein scheme is not avoided by an (even large) increase of the number of simulations. The only remedy is a choice of a smaller discretization step. One also observes that the Milshtein scheme has a better behaviour for small ν. Let us explain why, and present a new algorithm.

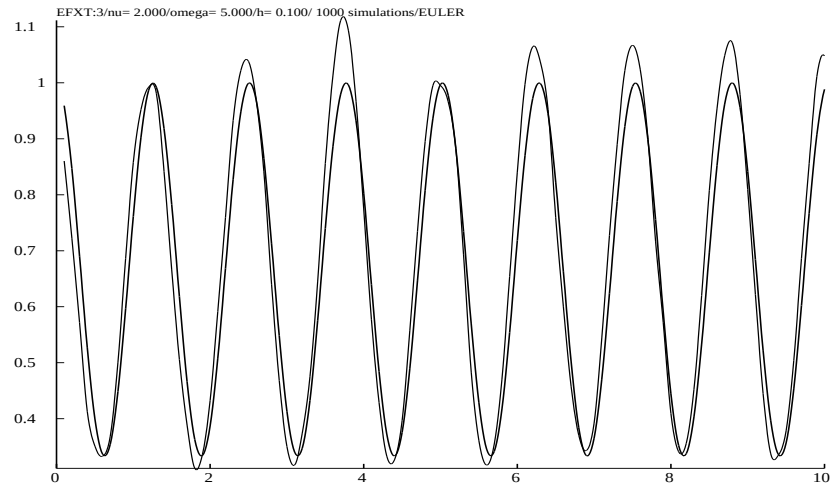

Fig. 5. Exact value and Euler scheme.

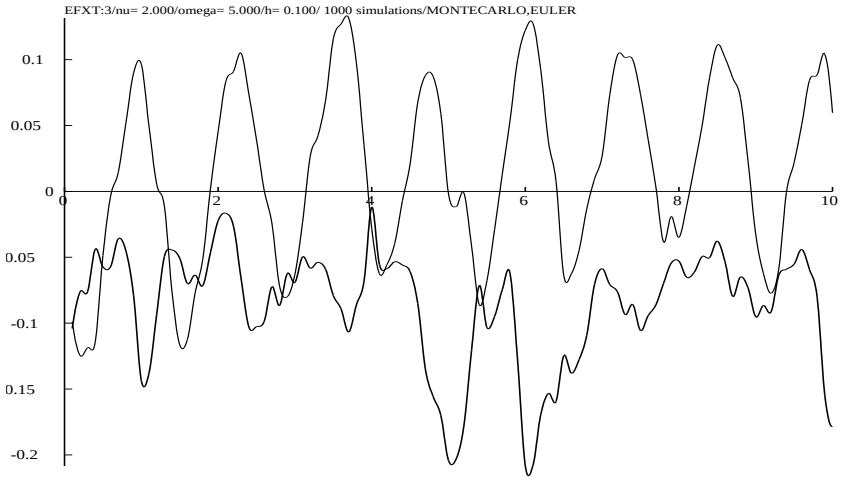

Fig. 6. Second-order and Euler schemes.

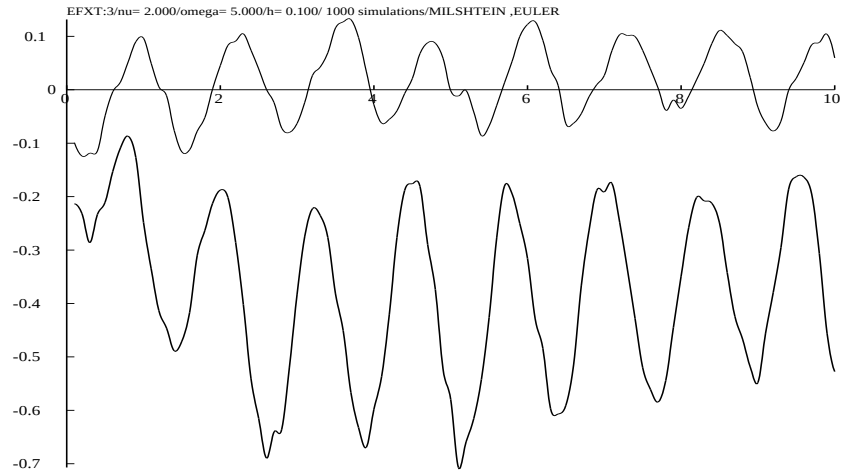

Fig. 7. Milshtein and Euler schemes.

5.6 Expansion of the error

The following theorem (Talay & Tubaro [58]) explains the numerical results of the previous section.

Theorem 4. *Let us suppose that the functions $b(\cdot)$, $\sigma(\cdot)$ and $f(\cdot)$ are C^∞; the derivatives of all orders of b and σ are supposed bounded, those of f are supposed to have a growth at most polynomial at infinity*[3]*. Then, for any step-size h of the form $\frac{T}{n}$:*

(i) For the Euler scheme, the error at time T is given by

$$Err_e(T,h) = Ef(X(T)) - Ef\left(\overline{X}_n^h\right) = -h\int_0^T E\psi_e(s,X_s)ds + \mathcal{O}(h^2) \quad , \tag{24}$$

where, if $u(t,x) := Ef(X(t,x))$:

$$\psi_e(t,x) = \frac{1}{2}\sum_{i,j=1}^{d} b^i(t,x)b^j(t,x)\partial_{ij}u(t,x) + \frac{1}{2}\sum_{i,j,k=1}^{d} b^i(t,x)a_k^j(t,x)\partial_{ijk}u(t,x)$$

[3] See the remark thereafter for a generalization of the result when f is not a smooth function.

$$+\frac{1}{8}\sum_{i,j,k,l=1}^{d} a_j^i(t,x)a_l^k(t,x)\partial_{ijkl}u(t,x) + \frac{1}{2}\frac{\partial^2}{\partial t^2}u(t,x)$$
$$+\sum_{i=1}^{d} b^i(t,x)\frac{\partial}{\partial t}\partial_i u(t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_j^i(t,x)\frac{\partial}{\partial t}\partial_{ij}u(t,x) \quad . \quad (25)$$

(ii) The same result extends to the Milshtein scheme:

$$Err_m(T,h) = -h\int_0^T E\psi_m(s,X_s)ds + \mathcal{O}(h^2)$$

where $\psi_m(\cdot)$ is defined by

$$\psi_m(t,x) = \psi_e(t,x) + \frac{1}{4}\sum_{i_1,i_2,j,k,l} a_k^l(t,x)\partial_l\sigma_j^{i_1}(t,x)\partial_k\sigma_j^{i_2}(t,x)\partial_{i_1 i_2}u(t,x)$$
$$+\frac{1}{2}\sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,k}} \sigma_{j_1}^{i_1}(t,x)\sigma_{j_2}^{i_2}(t,x)\sigma_{j_1}^{k}(t,x)\partial_k\sigma_{j_2}^{i_3}(t,x)\partial_{i_1 i_2 i_3}u(t,x). \quad (26)$$

(iii) The same result also extends to the schemes (22) and MCRK; besides, for these 2 schemes, as well as for the Euler and Milshtein schemes, an expansion of the error as power series in h exists: for any integer n, there exist constants $C_1,\ldots,C_n$ independent of h (but depending on the scheme) such that:

$$Ef(X(T)) - Ef\left(\overline{X}_n^h\right) = C_1 h + C_2 h^2 + \ldots + C_n h^n + \mathcal{O}(h^{n+1}) \quad .$$

Remarks. In the preceding statement, f is supposed smooth. In Bally & Talay [5], this hypothesis is relaxed (f is only supposed measurable with a polynomial growth at infinity) under a condition on L which is slightly more than hypoellipticity; the technique of the proof uses the Malliavin calculus.

For the example illustrated by the figure 7, tedious computations show (cf [58]) that the difference between the Euler and Milshtein schemes errors behaves (for large T) like

$$\frac{\nu^2}{2}(2+\cos(\Omega T))h \quad .$$

5.7 Romberg extrapolations

An interesting consequence of the previous theorem is the justification of a Romberg extrapolation between values corresponding to two different step-sizes. More precisely, let us consider a scheme such that ($h = \frac{T}{n}$):

$$Err(T,h) = Ef(X(T)) - Ef\left(\overline{X}_n^h\right) = e_1(T)h + \mathcal{O}(h^2) \quad ,$$

and consider the following new approximation (the Romberg extrapolation):

$$Z_T^h = 2Ef(\bar{X}_{2n}^{h/2}) - Ef(\bar{X}_n^h) \ , \tag{27}$$

then:

$$Ef(X_T) - Z_T^h = \mathcal{O}(h^2) \ .$$

That is, it is possible to get a result of precision of order h^2 from results given by a first-order scheme.

This procedure seems to be very robust w.r.t the choice of the discretization step. Let us give an illustration of this remark: let us compare the time evolution of the errors due to the extrapolation based on the Milshtein scheme with the 2 step-sizes $h = 0.05$ and $h = 0.1$ (thick line), and to the Milshtein scheme itself with the step-size $h = 0.05$ (thin line): the extrapolation has extremely improved the accuracy (see figure 8).

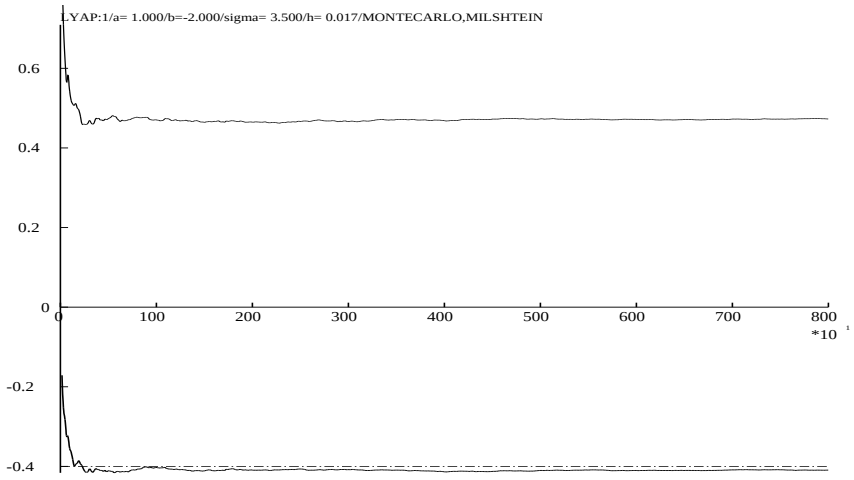

Fig. 8. Romberg extrapolation.

5.8 High order schemes

Of course, from a theoretical point of view it is possible to construct schemes with a convergence rate of an arbitrary order, or to get a precision of an arbitrary order by linearly combining results corresponding to a given scheme (Euler, MCRK, etc) and an appropriate number of different step-sizes.

We are not sure that high order procedures can be useful. Generally, it is impossible to choose a step-size too large without losing too much information

on the law of the increments of $(X(t))$. For small values of h (0.05 for example), it is difficult to see the gain in accuracy due to a 3rd order method, compared to a 2nd order one: the error due to the approximation of $Ef(\overline{X}_p^h)$ by the average (16) cannot be reduced enough (one cannot choose N so large as we would like!). Besides, the coefficients C_i in the expansion of the error depend on the successive derivatives of $b(\cdot)$, $\sigma(\cdot)$ and $f(\cdot)$. Often, C_i is rapidly increasing w.r.t. i.

6 Computation of Lyapunov exponents, study of stability

A summary of the theory of Lyapunov exponents of dynamical stochastic systems can be found in Arnold [2]; for bilinear systems, complements can be found in Pardoux & Talay [45].

To compute the upper Lyapunov exponent of 2-dimensional bilinear systems, W. Wedig [60] proposes an ingenious deterministic algorithm.

This section is a summary of results concerning the approximation of Lyapunov exponents of (non necessarily 2-dimensional) bilinear stochastic differential systems and the application to a helicopter blade problem, based on simulations and presented in Talay [56].

To be complete, we mention that an algorithm of computation of the Lyapunov spectrum for nonlinear systems in $\mathbb{R}^d$ or on compact manifolds has been developed, and its convergence rate given (see Grorud & Talay [27]).

6.1 An engineering stability problem

Let us study the stability of the motion of the movement of a rotor blade with 2 freedom degrees in terms of various physical parameters: velocity of the helicopter, geometric characteristics of the blade, statistical characteristics of the process modelizing the turbulency around the blade.

In first approximation, the stability of the movement of the blade is equivalent to the stability of the solution of a linearized ordinary differential equation in $\mathbb{R}^4$

$$\frac{dX(t)}{dt} = A(t)X(t) + F(t) \quad ,$$

where the matrix-valued function $A(t)$ and the vector-valued function $F(t)$ are periodic of same period (the period of rotation of the blade).

When one takes into account the turbulent flow around the blade, one may consider the following linearized model

$$\frac{dX(t)}{dt} = A(t)X(t) + F(t) + [B(t)X(t) + G(t)]\sigma(t)\xi^\varepsilon(t) \quad , \tag{28}$$

where $b(t)$ (resp. $G(t)$) has the same property as $A(t)$ (resp. $F(t)$), and $(\xi^\varepsilon(t))$ is a one-dimensional noise. The intensity of the noise, $\sigma(t)$ is also a periodic function of the azimuth angle Ωt, where Ω is the angular velocity of the blade.

All the coefficients $A(t)$, $b(t)$, $F(t)$, $G(t)$ are explicitly known in terms of different physical parameters of the blade.

Here the "stability" we are interested in, is the following: the system is stable when it admits a unique periodic in law solution $(Y(t))$ and when, for each initial deterministic condition, the corresponding process $(X(t))$ satisfies:

$$\lim_{t\to+\infty} |Y(t) - X(t)| = 0 \quad a.s. \tag{29}$$

First we will consider the white-noise case. The system (28) becomes:

$$dX(t) = [A(t)X(t) + F(t)]dt + [B(t)X(t) + G(t)]\sigma(t) \circ dW(t) \quad , \tag{30}$$

where $(W(t))$ is a standard one-dimensional Wiener process.

One can show (cf Pardoux[44]):

Theorem 5. *Let us suppose there exists $\lambda_0 < 0$ such that the solution of the system*

$$dX(t) = A(t)X(t)dt + B(t)X(t)\sigma(t) \circ dW(t) \tag{31}$$

satisfies, for any deterministic initial condition x :

$$\limsup_{t\to+\infty} \frac{1}{t} \log|X(t)| \leq \lambda_0 \quad a.s. \tag{32}$$

Then the system (30) is stable in the sense (29).

In her thesis, M. Pignol ([46]) has proven the existence of the Lyapunov exponent $\lim_{t\to+\infty} \frac{1}{t} \log|X(t)|$ for the blade system. Let us see how we can compute it (in order to know its sign!).

6.2 Numerical tests

Let us consider an example of Baxendale for which there exists an explicit formula giving the Lyapunov exponent. More precisely, let us consider a one-dimensional Wiener process $(W(t)$ and the system:

$$dX(t) = AX(t)dt + \sigma BX(t) \circ dW(t) \quad , \tag{33}$$

with

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \quad , \quad B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad .$$

Then:

$$\lambda = \frac{1}{2}(a+b) + \frac{1}{2}(a-b)\frac{\int_0^{2\pi} \cos(2\theta) \exp(\frac{a-b}{2\sigma^2}\cos(2\theta))d\theta}{\int_0^{2\pi} \exp(\frac{a-b}{2\sigma^2}\cos(2\theta))d\theta} \quad .$$

We discretize (33). For each Markov chain defined by one of our schemes, and for any h small enough, one can show that there exists $\overline{\lambda}^h$ such that, for any deterministic initial condition:

$$\overline{\lambda}^h = a.s \lim_{p\to+\infty} \frac{1}{ph} \log|\overline{X}_p^h| \quad .$$

We want to compare $\overline{\lambda}^h$ and λ.

It is important to note that we cannot use the formula $\overline{\lambda}^h \sim \frac{1}{ph} \log |\overline{X}_p^h|$ in practice, because it leads to numerical instabilities, the process $(|\overline{X}_p^h|)$ decreasing to 0 or increasing to infinity exponentially fast. This is avoided by a projection at each step technique, described in the next section.

We have tested the Milshtein method and the second-order method (22).

By example, let us choose $a = 1, b = -2, \sigma = 3.5$. Then an accurate numerical computation gives $\lambda = -0.4$. In this example, the second-order schemes are extremely accurate, whereas the Euler scheme and the Milshtein scheme lead to completely wrong results, as illustrated by the 2 figures below, which represent the evolution (in terms of time ph) of our estimator of $\overline{\lambda}^h$.

First, we compare the scheme (22) (thick line) and the Euler scheme (thin line): figure 9.

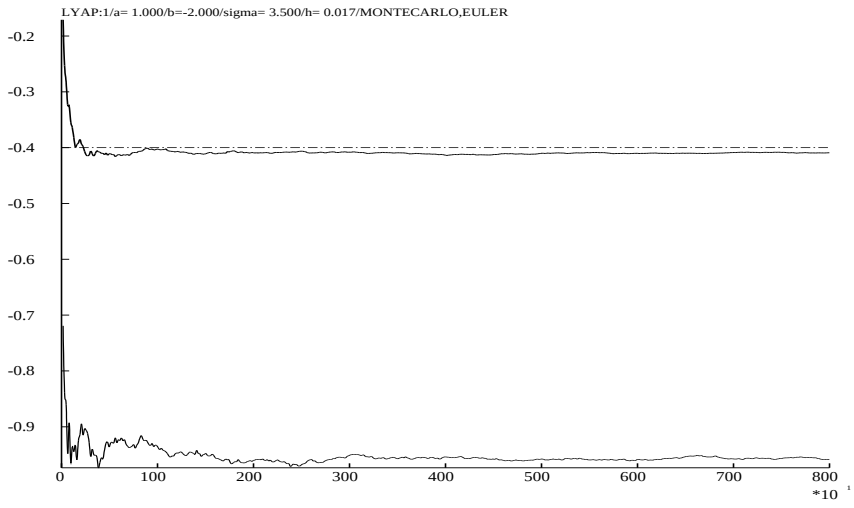

Fig. 9. Euler and second-order schemes.

Next, we compare the scheme (22) (thick line) and the Milshtein scheme (thin line): figure 10.

6.3 Algorithm for the helicopter problem

In the deterministic case (corresponding to $\sigma(t) \equiv 0$), only the Runge-Kutta methods of order larger than 4 have given good results (because of the numerical instability of the system, due to the large coefficients of the matrix $A(t)$ and their very short period).

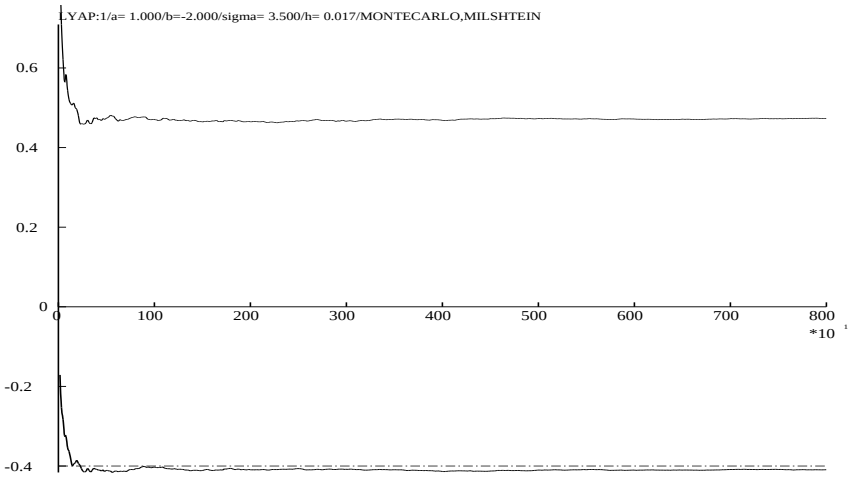

Fig. 10. Milshtein and second-order schemes.

Moreover, the discretization step had to be choosen smaller than 10^{-4}.

For the stochastic case, the use of a second-order scheme has also been necessary. We have observed the need of a large final integration time, therefore too small steps would have led to too large CPU times. Morevover, the second-order schemes have been less sensitive than the first-order ones to the above underlined strong instability of the system.

Besides, the above study of the deterministic case shows the necessity to improve the scheme (22), in order that it reduces to the Runge-Kutta scheme of order 4 when the intensity of the noise is nought.

Finally, our algorithm has been the following:

1. One chooses an initial condition on the unit sphere $\mathcal{S}^3$;
2. At step $(p+1)$, one proceeds in two stages:
   - One applies to a single step the Runge-Kutta method of order 4 in order to integrate the system

$$\begin{aligned} y(0) &= \overline{X}_p^h \\ \dot{y}(t) &= \tilde{A}(ph+t)y(t) \end{aligned}$$

   (the presence of $\tilde{A} = A + \frac{1}{2}B^2$ is due to the discretization of the system written in the Itô sense);
   - Then one computes:

$$\overline{X}_{p+1}^h = y(h) + \left\{ \sigma(ph)B\Delta_{p+1}^h W + \frac{1}{2}\sigma(ph)^2 B^2((\Delta_{p+1}^h W)^2 - h) \right.$$

$$+\frac{1}{2}(\sigma(ph)(\tilde{A}B + B\tilde{A} + B') + \sigma'(ph)B)h\Delta_{p+1}^h W\}\overline{X}_p^h \quad , \tag{34}$$

where we have used the following notation:
$(\Delta_{p+1}^h W)$:= sequence of mutually independent Gaussian random variables $\mathcal{N}(0, \sqrt{h})$;
B' : derivative of the matrix $B(\cdot)$;
all the matrices are computed at time $t = ph$;

3. One computes the new approximate value λ_{p+1} from the previous one, λ_p by

$$\lambda_p\left(1 - \frac{1}{p+1}\right) + \frac{\log(|\overline{X}_{p+1}^h|)}{(p+1)h} \quad ; \tag{35}$$

4. One projects $\overline{X}_{p+1}^h$ on the unit sphere.

The theorem 8 below shows that this algorithm is of second order, in the sense that, for any deterministic initial condition x:

$$\left|\lambda - \lim_{p \longrightarrow +\infty} \frac{1}{ph} \log|\overline{X}_p^h(x)|\right| = O(h^2) \quad .$$

For the given models of blades, the deterministic system was extremely stable for admissible velocities.

Let us suppose that $\sigma(t)$ is a constant function $\sigma(t) \equiv \sigma_0$ (i.e the effects of the turbulency are independent of the azimuth angle).

For a velocity equal to $100m/s$, we obtain the figure 11.

The destabilization of the system could occur only for intensities of the noise larger than 0.3; such intensities are not realistic for the turbulent winds around the blade.

Moreover, let us consider a more precise modelization of the noise. We suppose that its intensity is a periodic function of the azimuth angle $\Psi = \Omega t$, reaching its maximum for $\Psi = \frac{3\pi}{4}$, and defined by the function:

$$\sigma_0 \exp(\delta(cos(0.75\pi - 0.5\Omega t)^2 - 1)) \quad .$$

We have observed a strong dependency of the Lyapunov exponent upon δ but in that case also the instability could not appear for realistic intensities of the noise.

6.4 Choice of the integration time

This choice is much more complex that the choice of the number of simulations to perform a Monte Carlo approximation: the error due to the fact that one integrates during a finite time asymptotically has a Gaussian law, but the variance of this law is given by the solution of a P.D.E., and moreover this P.D.E. depends on the Lyapunov exponent that one wants to compute!

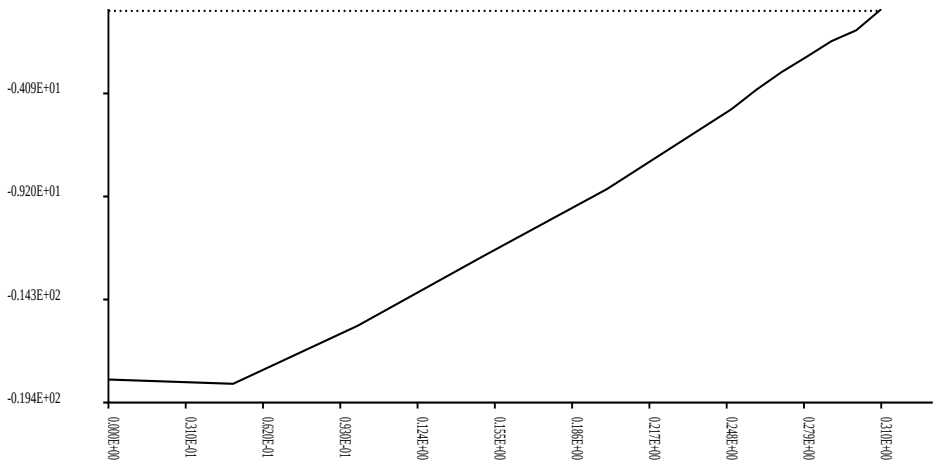

Fig. 11. Variations of the Lyapunov exponent in terms of σ_0.

More precisely, considering again the case of matrices A and B independent of t for the sake of simplicity, let $\mathcal{L}$ define the infinitesimal generator of the process $(s(t))$ defined by $s(t) = \frac{X(t)}{|X(t)|}$; under technical assumptions on the matrices A and B (satisfied by the blade system and described below), this process has a unique invariant probability measure μ on the projective space P^{d-1} (i.e the set obtained by identifying s and $-s$ on the sphere).

We have the following central-limit theorem (Bhattacharya [8]):

Theorem 6. *Let the function Q be defined by*

$$Q(s) := (As, s) + \frac{1}{2}\left[(B^2 s, s) + |Bs|^2 - 2(Bs, s)^2\right] \quad .$$

For t tending to $+\infty$:

$$\frac{1}{\sqrt{t}} \int_0^t (Q(s(\theta)) - \lambda) d\theta \to \mathcal{N}(0, \nu^2) \ \ in \ \ distribution \ \ ,$$

where the constant ν^2 depends only on the coefficients of the system and is given by ($< \cdot, \cdot >$ denoting the inner product in $\mathcal{L}^2(P^{d-1}, \mu)$):

$$\nu^2 = -2 < Q - \lambda, \mathcal{L}^{-1}(Q - \lambda) > \quad .$$

An estimation of the integration time can be got by a first estimation of λ, and by a numerical resolution of the Poisson P.D.E. on the projective space

$$\mathcal{L}u = Q - \lambda \quad .$$

Of course, this phase of the algorithm is very critical. Further researches in that direction are necessary.

6.5 Algorithm for the wideband noise case

Let us consider the system

$$\frac{dX(t)}{dt} = A(t)X(t) + B(t)X(t)\sigma(t)\xi^{\varepsilon}(t) \quad , \tag{36}$$

with a wide-band noise of the form

$$\xi^{\varepsilon}(t) = \frac{1}{\sqrt{\varepsilon}} Z(\frac{t}{\varepsilon}) \quad ,$$

where $(Z(t))$ is a stationary$\mathcal{N}(0,1)$ Ornstein-Uhlenbeck process.

The above system is an ordinary differential system, so that the pathwise simulation of the solution can be achieved using, by example, the Euler scheme:

$$\overline{X}^h_{p+1} = \overline{X}^h_p + (A(ph)\overline{X}^h_p + \sigma(ph)B(ph)\xi^{\varepsilon}(ph))\overline{X}^h_p h \quad .$$

Pardoux [44] and Kushner [33] have shown that the Lyapunov exponent of (36) converges to the Lyapunov exponent of (31).

But it appears that, even for small h, the Lyapunov exponent of that dicrete-time process defined by the Euler scheme does not converge, when ε goes to 0, towards the Lyapunov exponent of the system (31).

One reason is that the process $\{\xi^{\varepsilon}(ph)\}$ does not converge in law, so that the scheme must rather involve the sequence

$$\Delta^h_{p+1}\xi = \int_{ph}^{(p+1)h} \xi^{\varepsilon}(s)ds$$

which converges in law to the sequence $(\Delta^h_{p+1}W)$.

But one has to be careful: the new scheme

$$\overline{X}^h_{p+1} = \overline{X}^h_p + (A(ph)h + \sigma(ph)B(ph)\Delta^h_{p+1}\xi)\overline{X}^h_p$$

does not converge to a discretization scheme of (31) ($A \neq \tilde{A}$).

Finally, we introduce a convenient second-order scheme, similar to (34).

First, one applies to a single step the Runge-Kutta method of order 4 in order to integrate the system

$$\begin{aligned} y(0) &= \overline{X}^h_p \\ \dot{y}(t) &= A(ph+t)y(t) \end{aligned}$$

and then

$$\begin{aligned} \overline{X}^h_{p+1} = y(h) + \Big\{ &\sigma(ph)B\Delta^h_{p+1}\xi + \frac{1}{2}\sigma(ph)^2B^2(\Delta^h_{p+1}\xi)^2 \\ &+ \frac{1}{2}(\sigma(ph)(AB+BA+B') + \sigma'(ph)B)h\Delta^h_{p+1}\xi \Big\} \overline{X}^h_p \quad . \end{aligned} \tag{37}$$

It is interesting to note that the limit of the above scheme when ε goes to 0 is not the scheme (34), the difference including only terms of order $h\Delta^h_{p+1}W$

and h^2. We have not succeeded to build a second-order scheme of the wideband system converging to a second-order scheme of the white noise system.

Let us describe our simulation of the integrals $\Delta^h_{p+1}\xi$.

Let $(V(t))$ is a Wiener process independent of $(W(t))$ such that:

$$dZ(t) = -Z(t)dt + \sqrt{2}dV(t) \quad .$$

Then we have the formula:

$$\int_{ph}^{(p+1)h} \xi^\varepsilon(s)ds = \sqrt{\varepsilon}\left[\frac{1-e^{-\frac{2h}{\varepsilon}}}{2}\xi^\varepsilon(ph) + \Delta^{\frac{1}{\varepsilon}}_{p+1}V - e^{-\frac{2(p+1)h}{\varepsilon}}\int_{ph}^{(p+1)h} e^{2s}dV_s\right] \quad .$$

Therefore, it is possible to simulate the vector

$$(\xi^\varepsilon(ph), \Delta^h_{p+1}\xi)$$

by the simulation of the Gaussian vector

$$(\Delta^{\frac{1}{\varepsilon}}_{p+1}V, \int_{ph}^{(p+1)h} e^s dV_s, \int_{ph}^{(p+1)h} e^{2s}dV_s) \quad .$$

For our models, we have observed that the coloration of the noise tended to stabilize the system (the limit case being the case of very large ε, equivalent to the deterministic case).

6.6 Remarks

The models for the blade and the noise were simplified; in particular, only physical experiments during a flight could permit to improve the modelling of the noise, and overall a more realistic model should be nonlinear.

In this simplified context, the conclusion is that the turbulency around the blade has small effects on the stability of the blade.

6.7 Convergence rate

Let us consider a bilinear system

$$dX(t) = AX(t)dt + \sum_{i=1}^{r} B_i X(t) \circ dW_i(t) \quad . \tag{38}$$

Let $(s(t))$ be the process on the projective space of $\mathbb{R}^d$, P^{d-1}, defined as the equivalence class of $\frac{X(t)}{|X(t)|}$ with respect to the equivalence relation: $x \sim y$ iff $x = -y$ or $x = y$.

The process $(s(t))$ is the solution of the following Stratonovich stochastic differential equation, describing a diffusion process in P^{d-1}:

$$ds(t) = g(A, s(t))dt + \sum_{i=1}^{r} g(B_i, s(t)) \circ dW_i(t) \quad , \tag{39}$$

where

$$g(C,s) := Cs - (Cs,s)s \quad .$$

Now, let us introduce the Lie algebra $\Lambda = L.A.\{g(A,\cdot), g(B_1,\cdot), \ldots, g(B_k,\cdot)\}$, i.e. the smallest vector space of differential operators containing the operators

$$\sum_i g^i(A,\cdot)\partial_i \quad , \quad \sum_i g^i(B_j,\cdot)\partial_i \quad (j = 1, \ldots r)$$

and closed under the brackett operation $[P_1, P_2] = P_1 \circ P_2 - P_2 \circ P_1$.

For s in the projective space P^{d-1}, $\Lambda(s)$ denotes the space obtained by considering all the elements of Λ with all the coefficients of the operators frozen at their value in s.

In Arnold, Oeljeklaus and Pardoux [3] is proven the following theorem:

Theorem 7. *Let us suppose:*

(H) $dim\Lambda(s) = d-1 \quad , \quad \forall s \in P^{d-1} \quad .$

Then the process $(s(t))$ on P^{d-1} has a unique invariant probability measure μ, and there exists a real number λ such that, for any x in $\mathbb{R}^d - \{0\}$:

$$\lambda = \lim_{t \longrightarrow +\infty} \frac{1}{t} \log |X(t,x)| \quad , \quad a.s.$$

In Talay [56], is proven the

Theorem 8. *Let us suppose that the system (39) satisfies the hypothesis:*

(H0) *The infinitesimal generator $\mathcal{L}$ of the process $(s(t))$ on S^{d-1} is uniformly elliptic, i.e there exists a strictly positive constant α such that, for any x in S^{d-1} and any vector ξ in the tangent space $T_{S^{d-1}}(x)$:*

$$\sum_{i=1}^{r} (h(B_i,x),\xi)^2 \geq \alpha |\xi|^2 \quad .$$

(HU) **(i)** *The (U^j_{p+1})'s are i.i.d., and the following conditions on the moments are fulfilled:*

$$\begin{aligned} E[U^j_{p+1}] = E[U^j_{p+1}]^3 = E[U^j_{p+1}]^5 &= 0 \quad , \\ E[U^j_{p+1}]^2 &= 1 \quad , \\ E[U^j_{p+1}]^4 &= 3 \quad , \\ E[U^j_{p+1}]^n &< +\infty \quad , \quad \forall n > 5 \quad . \end{aligned} \tag{40}$$

(ii) *The common law of the (U^j_{p+1})'s has a continuous density w.r.t. the Lebesgue measure; the support of this density contains an open interval including 0 and is compact.*

Let $(\overline{X}_p^h, p \in \mathbb{N})$ *be defined by the Euler scheme, the Milshtein scheme or the scheme (22).*

Then, if λ *is the upper Lyapunov exponent of (38):*

(i) *for* $(\overline{X}_p^h)$ *defined by the Euler or Milshtein scheme,* $|\lambda - \overline{\lambda}^h| = O(h)$;

(ii) *for* $(\overline{X}_p^h)$ *defined by the scheme (22),* $|\lambda - \overline{\lambda}^h| = O(h^2)$.

Remark: the hypothesis (HU) is not limitative from a practical point of view, but (ii) was unnecessary to obtain the results concerning the Monte Carlo type approximation.

7 Computation of the invariant law

7.1 Position of the problem

We again consider the general system (4).

Under the hypotheses below, the system has a unique invariant measure μ, which has a smooth density, p. One way to compute $\int f(x)d\mu(x)$ for a given function f is to solve the stationary Fokker-Plank equation $L^*p = 0$, where L^* is the adjoint of the infinitesimal generator of the process $(X(t))$.

This stationary Fokker-Plank equation is a P.D.E., and its numerical resolution could be extremely difficult or impossible, especially when the dimension of the state-space, d, is large, or when L is degenerate (remember the Remark of the section 5.1).

In [26], Gerardi, Marchetti & Rosa propose to approximate $(X(t))$ by a sequence of ergodic pure jump processes which converges in law.

Since for any μ-integrable function f we have:

$$\int f(x)d\mu(x) = \lim_{t\to+\infty} \frac{1}{t}\int_0^t f(X(s))ds \quad ,$$

we propose to simulate one long trajectory of a process $(\overline{X}_p^h)$, and to approximate $\int f(x)d\mu(x)$ by

$$\frac{1}{N}\sum_{p=1}^{N} f(\overline{X}_p^h) \quad .$$

As in the preceding section, the critical point is the choice of N: again the random variable

$$\frac{1}{\sqrt{t}}\int_0^t [f(X(s)) - \int f(\theta)d\mu(\theta)]ds$$

is asymptotically Gaussian, but the variance of the limit law depends on the solution of a P.D.E. which itself depends on the unknown $\int f(x)d\mu(x)$.

At the present time, we do not know what could be a good procedure to estimate this variance.

7.2 Second-order schemes

As for the approximation of Lyapunov exponents, these schemes seem to have a better long-time behaviour than simpler ones.

For example, one can show (cf Talay [55]):

Theorem 9. *Suppose that the hypotheses (H1), (H2), (H3) hold:*

(H1) *the functions b, σ are of class $\mathcal{C}^\infty$ with bounded derivatives of any order; the function σ is bounded;*

(H2) *the operator L is uniformly elliptic: there exists a positive constant α such that*

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

(H3) *there exists a strictly positive constant β and a compact set K such that:*

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

Consider the scheme (22) and the MCRK scheme (23), with the law of the involved random variables defined as in the section (5.3), with the additional hypothesis: the law of the U_{p+1}^j's and of the U_{p+1}^j's has a continuous density w.r.t. the Lebesgue measure.

The schemes (22) and MCRK define ergodic Markov chains and for any function f of $\mathbb{R}^d$ of class $\mathcal{C}^\infty$, having the property that f, as well as all its derivatives, have an at most polynomial growth at infinity:

$$\forall x \in \mathbb{R}^d \quad : \quad \lim_{N \to \infty} \frac{1}{N} \sum_{p=1}^{N} f(\overline{X}_p^h(x)) = \int f(\theta) d\mu(\theta) + O(h^2) \quad , \quad a.s.$$

For the Euler and the Milshtein schemes, the convergence rate is of order h.

We underline that now we do not require anymore that the law of the U_{p+1}^j's has a compact support, therefore the Gaussian law of the increments of the Wiener process is allowed.

Again, another good (and probably usually better) algorithm is to perform a Romberg extrapolation (see the subsection 5.7); indeed, as for the problem of computation of $Ef(X(t))$ on a finite time interval, the errors due to the different schemes introduced above can be expanded as power series in the discretization step h (see Talay & Tubaro [58]):

Theorem 10. *Suppose that the hypotheses of the preceding theorem hold. Let ψ_e and ψ_m be defined by (25) and (26) respectively and set*

$$\lambda_e := \int_0^{+\infty} \int_{\mathbb{R}^d} \phi_e(t, y) \mu(dy) dt \quad ,$$

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

Then the Euler scheme error satisfies: for any deterministic initial condition $\xi = \overline{X}_0^h$,

$$\int f(y)\mu(dy) - a.s. \lim_{N\to+\infty} \frac{1}{N}\sum_{p=1}^{N} f(\overline{X}_p^h(\xi)) = -\lambda_e h + O(h^2) \quad . \tag{41}$$

For the Milshtein scheme, an analogous result can be written, substituting λ_m *to* λ_e.

7.3 Numerical experiments

The example is the same as in the section (5.5), with $\Omega = 0$, so that the invariant law is Gaussian of zero mean and of covariance matrix C.

We have observed a strong numerical instability of all the schemes (the results are very different when $h = 0.01$ and $h = 0.015$), but the second-order schemes lead to good results.

Below (figure 12), we compare $\frac{1}{N}\sum_{p=1}^{N} f(\overline{X}_p^h(x))$ for the Milshtein scheme (thin line) and the scheme (22) (thick line), for

$$f(x^1, x^2) = |x^1|^4$$

(the correct value is 3.0; in x-axis : Nh).

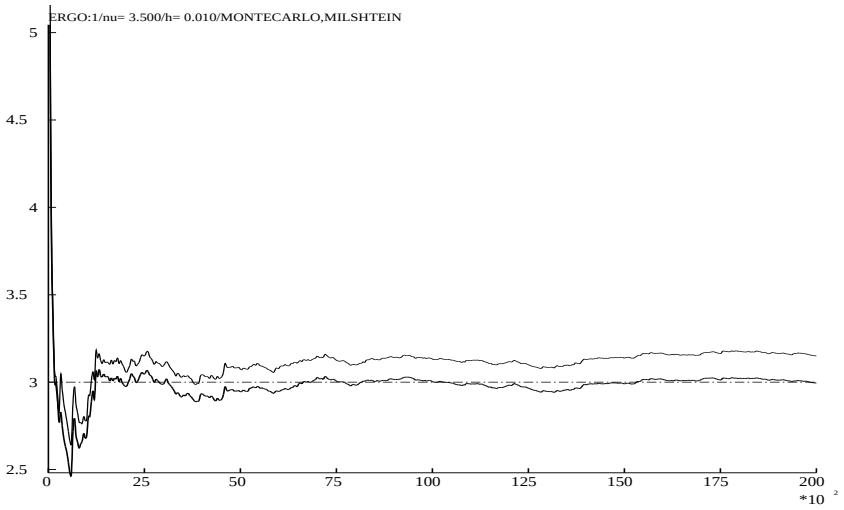

Fig. 12. Milshtein and second-order schemes.

Remark : for the two schemes, $\frac{1}{N}\sum_{p=1}^{N} f(\overline{X}_p^h(x))$ converges to $\int f(\theta)d\mu^h(\theta)$, where μ^h is the invariant law of the Markov chain $(\overline{X}_p^h)$.

8 PRESTO : a generator of Fortran programs

PRESTO is a system which generates Fortran programs solving Stochastic Differential Systems.

The user describes his problem using a bitmap environment; then PRESTO treats the data, performs the transformation Stratonovich/Itô of the system if necessary, uses its knowledge base in order to decide what particular scheme can be used in the context described by the user, what random variables must be involved and how they must be simulated, and finally writes a commented complete Fortran program ready to be run.

Internally, the analytical expressions of the coefficients of the Itô system and of the chosen scheme are computed by procedures written in a Computer Algebra Programming System Language (Reduce in the first version, Maple in the current one).

A complete description can be found in Talay [57].

9 Conclusion

We have proposed some discretization methods of stochastic differential systems, which seem efficient when one wants to simulate trajectories of diffusion processes, or when one wants to compute certain quantities depending on the law of a diffusion process, by techniques involving simulations.

A lot of open problems still remain, some of them are being studied: as examples, we could quote the discretization of reflected diffusions processes (which are studied in extremely recent interesting papers, see e.g Calzolari & Costantini & Marchetti [15] on the simulation of obliquely reflecting Brownian motions, Liu [35] who uses a penalization technique, Lépingle [34] for reflections at the boundary of a half–space or an orthant, and in a more abstract way Slominski [50]), the approximation of stopped diffusions and the numerical approximation of elliptic P.D.E.'s in bounded domains (see Milshtein [39]), the estimation of the necessary simulation time corresponding to a given accuracy for ergodic computations (computation of the stationary law, Lyapunov exponents, ...), etc.

From the numerical implementation point of view, Bouleau [12] and Ben Alaya [6] have just opened new perspectives by the use of the shift method to generate Brownian paths with few calls to a random number generator, and by their mathematical analysis of their algorithm (which is a nice application of the ergodic theory) in particular when the objective is to compute expectations of functionals of diffusion processes.

All these works show that the numerical analysis of diffusion processes is a field which is developing so fast that a new review paper will be necessary in a next future.

References

1. L. ARNOLD. *Stochastic Differential Equations.* Wiley, New-York, 1974.
2. L. ARNOLD. Lyapunov exponents of nonlinear stochastic systems. In *Nonlinear Stochastic Dynamic Engineering Systems (G.I. Schueller & F. Ziegler (Eds.).* Springer-Verlag, 1988.
3. L. ARNOLD, E. OELJEKLAUS, E. PARDOUX. Almost sure and moment stability for linear Itô equations. In *Lyapunov Exponents, L. Arnold & V. Wihstutz (Eds.)*, volume 1186 of *Lecture Notes in Mathematics.* Springer, 1986.
4. L. ARNOLD, P. KLOEDEN. Discretization of a random system near a hyperbolic point. Report 302, Institut für Dynamische Systeme, Universität Bremen, 1994.
5. V. BALLY, D. TALAY. The law of the Euler scheme for stochastic differential equations (I) : convergence rate of the distribution function. *Inria Research Report*, 2244, 1994. Submitted for publication.
6. M. BEN ALAYA. *Les Théorèmes Ergodiques en Simulation.* PhD thesis, Ecole Nationale des Ponts et Chaussées, 1993.
7. P. BERNARD, D. TALAY, L.TUBARO. Rate of convergence of a stochastic particle method for the Kolmogorov equation with variable coefficients. *Mathematics of Computation*, 63, 1994.
8. R.N. BHATTACHARYA. On the functional central limit theorem and the law of the iterared logarithm for Markov processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 60:185–201, 1982.
9. M. BOSSY. *Vitesse de Convergence d'Algorithmes Particulaires Stochastiques et Application à l'Equation de Burgers.* PhD thesis, Université de Provence, 1995.
10. M. BOSSY, D. TALAY. A stochastic particles method for nonlinear PDE's of Burgers type. Submitted for publication (a part of the paper has appeared as a Rapport de Recherche INRIA (number 2180), under the title "Convergence rate for the approximation of the limit law of weakly interacting particles (I) : smooth interacting kernels"), 1994.
11. M. BOSSY, D. TALAY. Convergence rate for the approximation of the limit law of weakly interacting particles : application to the Burgers equation. Rapport de Recherche 2410, Inria, novembre 1994. Submitted for publication.
12. N. BOULEAU. On numerical integration by the shift and application to Wiener space. *Acta Applicandae Mathematicae*, 3(25):201–220, 1991.
13. N. BOULEAU, D. LEPINGLE. *Numerical Methods for Stochastic Processes.* J. Wiley, 1993.
14. N. BOULEAU, D. TALAY (Eds.). *Probabilités Numériques.* Collection Didactique. INRIA, 1992.
15. A. CALZOLARI, C. COSTANTINI, F. MARCHETTI. A confidence interval for Monte Carlo methods with an application to simulation of obliquely reflecting Brownian motion. *Stochastic Processes and their Applications*, 29(2), 1988.
16. F. CASTELL, J.G. GAINES. The ordinary differential equation approach to asymptotically efficient schemes for solutions of stochastic differential equations. To appear in "Annales de l'Institut H. Poincaré".
17. J.M.C. CLARK. An efficient approximation for a class of stochastic differential equations. In W. Fleming et L. Gorostiza, editor, *Advances in Filtering and Optimal Stochastic Control*, volume 42 of *Lecture Notes in Control and Information Sciences.* Proceedings of the IFIP Working Conference, Cocoyoc, Mexico, 1982, Springer-Verlag, 1982.

18. J.M.C. CLARK, R.J. CAMERON. The maximum rate of convergence of discrete approximations for stochastic differential equations. In B. Grigelionis, editor, *Stochastic Differential Systems – Filtering and Control*, volume 25 of *Lecture Notes in Control and Information Sciences*. Proceedings of the IFIP Working Conference, Vilnius, Lithuanie, 1978, Springer-Verlag, 1980.
19. D. FLORENS-ZMIROU. Estimation de la variance d'une diffusion à partir d'une observation discrétisée. *Note au Compte-Rendu de l'Académie des Sciences*, 309(I):195–200, 1989.
20. H. DOSS. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré*, XIII(2):99–125, 1977.
21. O. FAURE. Simulation du Mouvement Brownien et des Diffusions. *Thèse de Doctorat, Ecole des Ponts et Chaussées*, 1992.
22. E. FOURNIE. *Statistiques des Diffusions Ergodiques avec Applications en Finance*. PhD thesis, Université de Nice-Sophia-Antipolis, 1993.
23. E. FOURNIE, D. TALAY. Application de la statistique des diffusions à un modèle de taux d'intérêt. *Finance*, 12(2), 1991.
24. J.G. GAINES. The algebra of iterated stochastic integrals. *Stochastics and Stochastic Reports*, 49:169–179, 1994.
25. J.G. GAINES, T.J. LYONS. Random generation of stochastic area integrals. *SIAM Journal of Applied Mathematics*, 54(4):1132–1146, 1994.
26. A. GERARDI, F. MARCHETTI, A.M. ROSA. Simulation of diffusions with boundary conditions. *Systems & Control Letters*, 4, 1984.
27. A. GRORUD, D. TALAY. Approximation of Lyapunov exponents of nonlinear stochastic differential systems. *SIAM J. Applied Math.*, to appear.
28. N. IKEDA, S. WATANABE. *Stochastic Differential Equations and Diffusion Processes*. North Holland, 1981.
29. I. KARATZAS, S.E. SHREVE. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, 1988.
30. P.E. KLOEDEN, E.PLATEN. *Numerical Solution of Stochastic Differential Equations*. Springer–Verlag, 1992.
31. T. KURTZ, P. PROTTER. Wong–Zakai corrections, random evolutions and numerical schemes for S.D.E.'s. In *Stochastic Analysis: Liber Amicorum for Moshe Zakai*, pages 331–346, 1991.
32. H.J. KUSHNER. *Approximation and Weak Convergence Methods for Random Processes*. M.I.T. Press, 1984.
33. H.J. KUSHNER. Approximations and optimal control for the pathwise average cost per unit time and discounted problems for wideband noise driven systems. *SIAM J. on Control and Optimization*, 27(3):546–562, 1987.
34. D. LEPINGLE. Un schéma d'Euler pour équations différentielles réfléchies. *Note aux Comptes-Rendus de l'Académie des Sciences*, 316(I):601–605, 1993.
35. Y. LIU. Numerical approaches to reflected diffusion processes. Submitted for publication, 1992.
36. E.J. Mc SHANE. *Stochastic Calculus and Stochastic Models*. Acad. Press, 1974.
37. G.N. MILSHTEIN. Approximate integration of stochastic differential equations. *Theory of Probability and Applications*, 19:557–562, 1974.
38. G.N. MILSHTEIN. Weak approximation of solutions of systems of stochastic differential equations. *Theory of Probability and Applications*, 30:750–766, 1985.
39. G.N. MILSHTEIN. The solving of the boundary value problem for parabolic equation by the numerical integration of stochastic equations. To appear, 1992.

40. N.J. NEWTON. An asymptotically efficient difference formula for solving stochastic differential equations. *Stochastics*, 19:175–206, 1986.
41. N.J. NEWTON. An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times. *Stochastics and Stochastic Reports*, 29:227–258, 1990.
42. N.J. NEWTON. Asymptotically efficient Runge-Kutta methods for a class of Ito and Stratonovich equations. *SIAM Journal of Applied Mathematics*, 51(2):542–567, 1991.
43. N.J. NEWTON. Variance reduction for simulated diffusions. *SIAM Journal of Applied Mathematics*, to appear.
44. E. PARDOUX. Stabilité du mouvement des pales d'hélicoptères dans le cas d'un écoulement de l'air turbulent. In *Actes du Colloque "L'Automatique pour l'Aéronautique"*, Paris, 1986.
45. E. PARDOUX, D. TALAY. Stability of linear differential systems with parametric excitation. In *Nonlinear Stochastic Dynamic Engineering Systems (G.I.Schueller & F.Ziegler (Eds.)*. Springer-Verlag, 1988.
46. M. PIGNOL. *Stabilité Stochastique des Pales d'Hélicoptère*. PhD thesis, Université de Provence, 1985.
47. E. PLATEN, W. WAGNER. On a Taylor formula for a class of Ito processes. *Probability and Mathematical Statistics*, 3:37–51, 1982.
48. W. RUMELIN. Numerical treatment of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 19:604–613, 1982.
49. K.R. SCHENK-HOPPE. Bifurcation scenarios of the noisy Duffing-van der Pol oscillator. Report 295, Institut für Dynamische Systeme, Universität Bremen, 1993.
50. L. SLOMINSKI. On approximation of solutions of multidimensional s.d.e.'s with reflecting boundary conditions. *Stochastic Processes and their Applications*, 50(2):197–219, 1994.
51. H.J. SUSSMANN. On the gap between deterministic and stochastic differential equations. *Ann. Prob.*, 6:19–44, 1978.
52. D. TALAY. Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. *Stochastics*, 9:275–306, 1983.
53. 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.
54. 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.
55. D. TALAY. Second order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics and Stochastic Reports*, 29(1):13–36, 1990.
56. D. TALAY. Approximation of upper Lyapunov exponents of bilinear stochastic differential systems. *SIAM Journal on Numerical Analysis*, 28(4):1141–1164, 1991.
57. D. TALAY. Presto: a software package for the simulation of diffusion processes. *Statistics and Computing Journal*, 4(4), 1994.
58. D. TALAY, L. TUBARO. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):94–120, 1990.

59. W. WAGNER. MonteCarlo evaluation of functionals of stochastic differential equations—variance reduction and numerical examples. *Stochastic Analysis and Applications*, 6:447–468, 1988.
60. W. WEDIG. Contribution in this volume. 1995.

This article was processed using the LaTeX macro package with ICM style