Convergence in total variation distance for a third order scheme for one dimensional diffusion process

Clément Rey

Convergence in total variation distance for a third order scheme for one dimensional diffusion process

Clément Rey [1]

Abstract

In this paper, we study a third weak order scheme for diffusion processes which has been introduced by Alfonsi [1]. This scheme is built using cubature methods and is well defined under an abstract commutativity condition on the coefficients of the underlying diffusion process. Moreover, it has been proved in [1], that the third weak order convergence takes place for smooth test functions. First, we provide a necessary and sufficient explicit condition for the scheme to be well defined when we consider the one dimensional case. In a second step, we use a result from [3] and prove that, under an ellipticity condition, this convergence also takes place for the total variation distance with order 3. We also give an estimate of the density function of the diffusion process and its derivatives.

1 Introduction

In this paper, we study the total variation distance between a one dimensional diffusion process and a third weak order scheme based on a cubature method and introduced by Alfonsi [1]. In his work, Alfonsi proved that it converges with weak order three for smooth test functions with polynomial growth. We will show that the convergence also takes place with order three if we consider measurable and bounded test functions. In this case, we say that the total variation distance between the diffusion process and the scheme converges towards zero with order three. In order to do it, we will use a result from [3] based on en abstract Malliavin calculus introduced by Bally and Clément [2]. A main interest of this approach is that the random variables used to build the scheme are not necessarily Gaussian but belong to a class of random variables with no specific law. Consequently our result can be seen as an invariance principle.

Let us be more specific. We consider the $\mathbb{R}$-valued one dimensional Markov diffusion process

$$dX_t = V_0(X_t)dt + V_1(X_t) \circ dW_t, \quad (1)$$

with $V_i : \mathcal{C}_b^\infty(\mathbb{R},\mathbb{R})$, $i = 0, 1$, $(W_t)_{t\geqslant 0}$ a one dimensional standard Brownian motion and $\circ dW_t$ the Stratonovich integral with respect to W_t. In this paper, we will study an approximation

[1] reyc@cermics.enpc.fr, CERMICS-École des Ponts Paris Tech, 6 et 8 avenue Blaise Pascal, Cité Descartes - Champs sur Marne, 77455 Marne la Vallée Cedex 2 - France.
Projet MathRisk ENPC-INRIA-UMLV, This research benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

scheme for (1) which is defined on an homogeneous time grid. It is relevant to notice that the results we will obtain remain true for non homogeneous time grids, but we do not treat that case for sake of clarity. We fix $T>0$ and we denote $n\in\mathbb{N}^*$, the number of time step between 0 and T. Then, for $k\in\mathbb{N}$ we define $t^n_k=kT/n$ and we introduce the homogeneous time grid $\pi_{T,n}=\{t^n_k=kT/n,k\in\mathbb{N}\}$ and its bounded version $\pi^{\tilde{T}}_{T,n}=\{t\in\pi_{T,n},t\leqslant\tilde{T}\}$ for $\tilde{T}\geqslant 0$. Finally, for $S\in[0,\tilde{T})$ we will denote $\pi^{S,\tilde{T}}_{T,n}=\{t\in\pi^{\tilde{T}}_{T,n},t>S\}$. Now, for $t^n_k=kT/n$, we introduce the abstract $\mathbb{R}$-valued Markov chain

$$X^n_{t^n_{k+1}}=\psi_k(X^n_{t^n_k},\frac{Z_{k+1}}{\sqrt{n}},\delta^n_{k+1}),\qquad k\in\mathbb{N}, \tag{2}$$

where $\psi_k:\mathbb{R}\times\mathbb{R}\times\mathbb{R}_+\to\mathbb{R}$ is a smooth function such that $\psi_k(x,0,0)=x$, $Z_{k+1}\in\mathbb{R}^N,\ k\in\mathbb{N}$ is a sequence of independent and centered random variables and $\sup_{k\in\mathbb{N}^*}\delta^n_k\leqslant C/n$.

Before estimating the distance between X and X^n, we introduce some notations. For $f\in\mathcal{C}^\infty(\mathbb{R}^d)$ and for a multi-index $\alpha=(\alpha_1,\cdots,\alpha_d)\in\mathbb{N}^d$ we denote $|\alpha|=\alpha_1+...+\alpha_d$ and $\partial_\alpha f=\partial^\alpha_x f=\partial^{\alpha_1}_{x_1}...\partial^{\alpha_d}_{x_d}f(x)$. We include the multi-index $\alpha=(0,...,0)$ and in this case $\partial_\alpha f=f$. We will use the norms

$$\|f\|_{q,\infty}=\sup_{x\in\mathbb{R}^d}\sum_{0\leqslant|\alpha|\leqslant q}|\partial_\alpha f(x)|,\quad q\in\mathbb{N}. \tag{3}$$

In particular $\|f\|_{0,\infty}=\|f\|_\infty$ is the usual supremum norm and we will denote $\mathcal{C}^q_b(\mathbb{R}^d)=\{f\in\mathcal{C}^q(\mathbb{R}^d),\|f\|_{q,\infty}<\infty\}$.

A first standard result is the following: Let us assume that there exists $h>0,\ q\in\mathbb{N}$ such that for every test function $f\in\mathcal{C}^q_b(\mathbb{R})$, $k\in\mathbb{N}$ and $x\in\mathbb{R}$,

$$|\mathbb{E}[f(X_{t^n_{k+1}})-f(X^n_{t^n_{k+1}})|X_{t^n_k}=X^n_{t^n_k}=x]|\leqslant C\|f\|_{q,\infty}/n^{h+1}. \tag{4}$$

Then, we have

$$\sup_{t\in\pi^T_{T,n}}|\mathbb{E}[f(X_t)-f(X^n_t)]|\leqslant C\|f\|_{q,\infty}/n^h. \tag{5}$$

It means that $(X^n_{t_k})_{k\in\mathbb{N}}$ is an approximation scheme of weak order h for the Markov process $(X_t)_{t\geqslant 0}$ for the test functions $f\in\mathcal{C}^q_b(\mathbb{R};\mathbb{R})$. The value h thus measures the efficiency of the scheme whereas q stands for the required regularity on the test functions in order to obtain convergence with order h. This subject has already been widely studied in the literature and we point out some famous examples. However, the reader may notice that in all those works, the required order of regularity q is greater than one. Concerning the Euler scheme for diffusion processes, the result (5), with $h=1$, has initially been proved in the seminal papers of Milstein [19] and of Talay and Tubaro [22] (see also [12]). Since then, various situations have been studied: Diffusion processes with jumps (see [21], [10]) or diffusion processes with boundary conditions (see [7], [6], [8]). An overview of the subject is proposed in [11]. More recently, discretization schemes of higher orders (*e.g.*, $h=2$), based on cubature methods, have been introduced and studied by Kusuoka [16], Lyons [18], Ninomiya, Victoir [20] or Alfonsi

[1]. The reader may also refer to the work Kohatsu-Higa and Tankov [13] for a higher weak order for jump processes. Finally, in [1], a third weak order scheme (with $h = 3$) has been introduced following similar cubature ideas. This is the one we will study in this paper.

As we already precised, all those schemes converge for some $q \geqslant 1$ in (5). Another point of interest relies thus on the study of the set of test functions which enable the converge with weak order h. The purpose is to extend this set beyond $\mathcal{C}_b^q(\mathbb{R};\mathbb{R})$ and to obtain (5) with $\|f\|_{q,\infty}$ replaced by $\|f\|_\infty$ when f is a measurable and bounded function. In this case, we say that the scheme converges for the total variation distance. A first result of this type has been obtained by Bally and Talay [4], [5]. They treat the case of the Euler scheme using the Malliavin calculus (see also Guyon [9] when f is a tempered distribution). Afterwards Konakov, Menozzi and Molchanov [14], [15] established some local limit theorems using a parametrix method. Recently Kusuoka [17], also using Malliavin Calculus, obtained estimates of the error in total variation distance for the Ninomiya Victoir scheme (which corresponds to the case $h = 2$) under a Hörmander type condition.

Under an ellipticity condition, we will obtain a similar result for the case $h = 3$, using a scheme introduced in [1]. This scheme is well defined if the Lie bracket between V_1^2 and V_0 is equal to $2\tilde{V}^2$, with $\tilde{V}$ a first order differential operator. Since we consider one dimensional processes with form (1), we will be able to give an explicit necessary and sufficient condition in order to obtain this property on the Lie bracket.

Moreover, we will not work in a Gaussian framework and then we will have to use a variant of the Malliaivin calculus introduced by Bally and Clement [2] for which we can apply the results from [3]. A main interest of this approach is that the random variables involved in the scheme do not have a specific law but simply belong to a class of random variables which are Lebesgue lower bounded and satisfy some moment conditions. In this way, our final result can be seen as an invariance principle. The ambit of this scheme thus goes well beyond the Gaussian case.

We will begin presenting the framework of this paper in Section 2. In Section 3, we will give some third weak order convergence results for smooth test functions and for bounded measurable test functions. The latter is presented in Theorem 3.2 and constitutes the main result of this paper. It gives the convergence for the total variation distance with order three of the scheme from [1], toward the Markov process (1). We will also obtain an estimate of the density function of the diffusion and its derivatives. We will follow with a short numerical illustration in order to check the order of convergence for a suited example. This paper will end with the proof of our main theorems in Section 5.

2 The third weak order scheme

We consider the one dimensional $\mathbb{R}$-valued diffusion process

$$dX_t = V_0(X_t)dt + V_1(X_t) \circ dW_t, \tag{6}$$

with $V_0, V_1 \in \mathcal{C}_b^\infty(\mathbb{R};\mathbb{R})$, $(W_t)_{t\geqslant 0}$ a standard Brownian motion. Moreover, $\circ dW_t$ denotes the Stratonovich integral with respect to W. The infinitesimal operator of this Markov process is

$$A = V_0 + \frac{1}{2}V_1^2, \tag{7}$$

with the notation $Vf(x) = V(x)\partial f(x)$. Let us define $\exp(V)(x) := \Phi_V(x,1)$ where Φ_V solves the deterministic equation

$$\Phi_V(x,t) = x + \int_0^t V(\Phi_V(x,s))ds. \tag{8}$$

By a change of variables one obtains $\Phi_{\varepsilon V}(x,t) = \Phi_V(x,\varepsilon t)$, so we have

$$\exp(\varepsilon V)(x) := \Phi_{\varepsilon V}(x,1) = \Phi_V(x,\varepsilon).$$

We also notice that the semigroup of the above Markov process, which is given by $P_t^V f(x) = f(\Phi_V(x,t))$, has the infinitesimal operator $A_V f(x) = Vf(x)$. In particular the relation $P_t^V A_V = A_V P_t^V$ reads

$$Vf(\Phi_V(x,t)) = A_V P_t^V f = P_t^V A_V f = V(x)\partial_x\left(f\circ\Phi_V\right)(x,t).$$

Using m times Dynkin's formula $P_t^V f(x) = f(x) + \int_0^t P_s^V A_V f(x)ds$ we obtain

$$f(\Phi_V(x,t))) = f(x) + \sum_{r=1}^{m}\frac{t^r}{r!}V^r f(x) + \frac{1}{m!}\int_0^t (t-s)^m V^{m+1}P_s^V f(x)ds. \tag{9}$$

We present now the third weak order scheme introduced in [1]. In order to do it, we introduce the following commutation property:

$$V_1^2V_0 - V_0V_1^2 = 2\tilde{V}^2, \tag{10}$$

where $\tilde{V}$ is a first order operator. We consider some sequences $\epsilon_k, \rho_k,\ k\in\mathbb{N}$ of independent uniform random variables with values in $\{-1,1\}$ and $\{1,2,3\}$, and we define $\psi : \{-1,1\}\times\{1,2,3\}\times\mathbb{R}^3\to\mathbb{R}$ using the following splitting procedure:

$$\psi(\epsilon_k,\rho_k,x,w_{k+1}^1,w_{k+1}^0) = \begin{cases} \exp(\epsilon_k w_{k+1}^0\tilde{V})\circ\exp(w_{k+1}^0V_0)\circ\exp(w_{k+1}^1V_1)(x), & \text{if } \rho_k = 1,\\ \exp(w_{k+1}^0V_0)\circ\exp(\epsilon_k w_{k+1}^0\tilde{V})\circ\exp(w_{k+1}^1V_1)(x), & \text{if } \rho_k = 2,\\ \exp(w_{k+1}^0V_0)\circ\exp(w_{k+1}^1V_1)\circ\exp(\epsilon_k w_{k+1}^0\tilde{V})(x), & \text{if } \rho_k = 3,\end{cases} \tag{11}$$

with $w_k^0 = T/n$, $w_k^1 = \sqrt{T}Z_k/\sqrt{n}$. We notice that $\psi(\epsilon_k,\rho_k,x,0,0) = x$, which is relevant with the definition of a scheme. Moreover $Z_k,\ k\in\mathbb{N}^*$ are independent random variables which are lower bounded by the Lebesgue measure: There exists $z_{*,k}\in\mathbb{R}$ and $\varepsilon_*, r_* > 0$ such that for every Borel set $A\subset\mathbb{R}$ and every $k\in\mathbb{N}^*$

$$L_{z_*}(\varepsilon_*,r_*) \qquad \mathbb{P}(Z_k\in A)\geqslant \varepsilon_*\lambda(A\cap B_{r_*}(z_{*,k})). \tag{12}$$

Moreover, we assume that the sequence Z_k satisfies the following moment conditions:

$$\mathbb{E}[Z_k]=\mathbb{E}[Z_k^3]=\mathbb{E}[Z_k^5]=\mathbb{E}[Z_k^7]=0,\quad \mathbb{E}[Z_k^2]=1,\quad \mathbb{E}[Z_k^4]=3,\quad \mathbb{E}[Z_k^6]=15,$$
$$\forall p\geqslant 1,\qquad \mathbb{E}[|Z_k|^p]<\infty. \tag{13}$$

One step of our scheme (between times t_k^n and t_{k+1}^n) is given by

$$X^n_{t^n_{k+1}}=\psi(\epsilon_k,\rho_k,X^n_{t^n_k},w^1_{k+1},w^0_{k+1}). \tag{14}$$

Using the notation from (2), we also have

$$X^n_{t^n_{k+1}}=\psi_k(X^n_{t^n_k},w^1_{k+1},w^0_{k+1}). \tag{15}$$

with $\psi_k(x,z,t)=\psi(\epsilon_k,\rho_k,x,z,t)$. In the sequel, we will study the third order convergence of this scheme towards the Markov process given in (1) for smooth test functions and for bounded measurable test functions.

3 Convergence Results

We begin introducing some notations. Let $r\in\mathbb{N}^*$. For a sequence of functions $\psi_k\in\mathcal{C}^r(\mathbb{R}\times\mathbb{R}\times\mathbb{R}_+;\mathbb{R})$, $k\in\mathbb{N}$, we denote

$$\|\psi\|_{1,r,\infty}=1\vee\sup_{k\in\mathbb{N}}\sum_{|\alpha|=0}^{r}\sum_{|\beta|+|\gamma|=1}^{r-|\alpha|}\|\partial_x^\alpha\partial_z^\beta\partial_t^\gamma\psi_k\|_\infty, \tag{16}$$

and for $r\in\mathbb{N}^*$,

$$\mathfrak{K}_r(\psi)=(1+\|\psi\|_{1,r,\infty})\exp(\|\psi\|^2_{1,3,\infty}). \tag{17}$$

3.1 Smooth test functions

In this Scetion, we study the convergence of the scheme given in (15) for smooth test functions. We state a first result, which is the starting point in order to prove the convergence in total variation distance.

Theorem 3.1. *Suppose that $V_0,V_1,\tilde{V}\in\mathcal{C}_b^\infty(\mathbb{R};\mathbb{R})$. We also assume that (10) and (13) hold. Then, there exists some universal constant $l\in\mathbb{N}^*$, $C\geqslant 1$ such that for every $f\in\mathcal{C}_b^8(\mathbb{R})$, we have*

$$\sup_{t\in\pi^T_{T,n}}|\mathbb{E}[f(X_t))-\mathbb{E}[f(X^n_t)]|\leqslant CC_8(V)^l\|f\|_{8,\infty}/n^3, \tag{18}$$

with $C_q(V):=\sup_{i=0,1}\|V_i\|_{q,\infty}+\|\tilde{V}\|_{q,\infty}$.

Remark 3.1. *This result has already been obtained in [1] in the case of test functions with polynomial growth. The proof is similar and since we intend to obtain this result with the supremum norm of f we do not treat that case.*

We give a proof of this result in Section 5. Once we have used the Lindeberg decomposition, it relies on short time estimates using the Dynkin's formula. Now, we are going to take a step further and consider simply bounded and measurable test functions. Notice that, it means the convergence for total variation distance.

3.2 Bounded mesurable test functions

We see that the estimate (18) involves the derivatives of order eight of the test function. We will see that it is possible to obtain similar estimates with $\|f\|_{8,\infty}$ replaced by $\|f\|_{\infty}$. This is a consequence of a result from [3] in which the authors provide some sufficient conditions for the scheme in order to obtain the convergence for the total variation distance. The scheme (14) satisfy those conditions and, under an ellipticity assumption on the diffusion coefficient V_1, we are going to obtain an estimate of its total variation distance with the diffusion process (6).

Before doing it, we introduce a necessary and sufficient explicit condition in order to obtain (10) as son as for all $x \in \mathbb{R}, V_1(x) \neq 0$. Notice that, since we assume that V_1 is continuous, it has a constant sign. Moreover, this hypothesis will not be restrictive in this application. Indeed, the ellipticity condition required to use the result from [3] implies that $\inf_x V_1(x)^2 \geqslant \lambda_* > 0$ for a constant λ_*. We will suppose without loss of generality that V_1 is positive. The necessary and sufficient condition for (10) is the following: We assume that the function

$$\begin{aligned} g :&\mathbb{R} \to \mathbb{R} \\ &x \mapsto V_0(x)/V_1(x) \end{aligned} \tag{19}$$

is increasing. Notice that if V_1 is negative, g has to be decreasing.

Moreover, we propose an alternative scheme in order to approximate the density function of X and its derivatives. We consider a standard normal random variable G which is independent from $Z_k, k \in \mathbb{N}$, and for $\theta > 0$, we introduce $(X^{n,\theta}_t)_{t\in\pi_{T,n}}$ as follows

$$X^{n,\theta}_t(x) = \frac{1}{n^\theta} G + X^n_t(x). \tag{20}$$

where $X^n(x)$ is the process which starts from x that is $X^n_0 = x$. We denote by $p^{\theta,n}_t(x, y)$ the density of the law of $X^{\theta,n}_t(x)$ and for $t \in \pi_{T,n}$, we define

$$Q^{n,\theta}_t f(x) := \mathbb{E}[f(\frac{1}{n^\theta} G + X^n_t(x))]. \tag{21}$$

Now, we can state our main result.

Theorem 3.2. *Suppose that $V_0, V_1, \tilde{V} \in \mathcal{C}^\infty_b(\mathbb{R};\mathbb{R})$. We fix $T > 0$ and we also assume that (19), (12) and (13) hold and that*

$$V_1(x)^2 \geqslant \lambda_* > 0 \qquad \forall x \in \mathbb{R}. \tag{22}$$

Let $S \in (0, T/2)$. Then there exists $n_0 \in \mathbb{N}^*$ such that for every $n \geqslant n_0$, we have the following properties.

A. *There exists $l \in \mathbb{N}^*$ and $C \geqslant 1$ which depends on m_*, r_* and the moments of Z such that, for every bounded and measurable function $f : \mathbb{R} \to \mathbb{R}$,*

$$\sup_{t \in \pi_{T,n}^{2S,T}} |\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]| \leqslant C \frac{C_8(V)^l \mathfrak{K}_{11}(\psi)^l}{(\lambda_* S)^{\eta(8)}} \|f\|_\infty / n^3. \tag{23}$$

with $\mathfrak{K}_r(\psi)$ and $C_q(V)$ given in (17) and (18) and $\eta(r) = r(r+1)$.

B. *Moreover, for every $t > 0, P_t(x, dy) = p_t(x, y)dy$ with $(x, y) \mapsto p_t(x, y)$ belonging to $\mathcal{C}^\infty(\mathbb{R}^d \times \mathbb{R}^d)$.*

C. *Let $\theta \geqslant h+1$. We recall the $Q^{n,\theta}$ is defined in (21) and verifies $Q_t^{n,\theta}(x, dy) = p_t^{n,\theta}(x, y)dy$. Then, there exists $l \in \mathbb{N}^*$ such that for every $R > 0, \varepsilon \in (0, 1)$, $x_0, y_0 \in \mathbb{R}^d$, and every multi-index α, β with $|\alpha| + |\beta| = u$, we also have*

$$\sup_{t \in \pi_{T,n}^{2S,T}} \sup_{(x,y) \in \overline{B}_R(x_0,y_0)} |\partial_x^\alpha \partial_y^\beta p_t(x, y) - \partial_x^\alpha \partial_y^\beta p_t^{n,\theta}(x, y)| \leqslant C \frac{C_8(V)^l \mathfrak{K}_{11}(\psi)^l}{(\lambda_* S)^{\eta(p_{u,\varepsilon} \vee 8)}} / n^{3(1-\varepsilon)} \tag{24}$$

with a constant C which depends on R, x_0, y_0, T and on $|\alpha| + |\beta|$ and $p_{u,\varepsilon} = (u + 2d + 1 + 2\lceil(1 - \varepsilon)(u + d)/(2\varepsilon)\rceil)$.

Remark 3.2. *It is relevant to notice that we have the same result if we assume that the function defined in (19) is decreasing (resp. increasing) for V_1 positive (resp. V_1 negative). In this case $V_0 V_1^2 - V_1^2 V_0 = 2\tilde{V}^2$ and we have to define the scheme differently. In the construction (11), we invert the terms containing V_1 with the ones containing V_0.*

Remark 3.3. *The property (12) is crucial here, since we will use a result from [3] which employs abstract integration by parts formulae based on the noise Z_k. However it is not restrictive for concrete applications.*

The result (23) signifies the convergence in total variation with order 3. The proof of this theorem is given in Section 5. Since we have already obtained some short time estimates of the form (4) in the proof of Thereom 3.1 and (19) holds, the key point of this proof does not rely on the weak order of the scheme. This is the fact that, the splitting procedure (11) in order to build the scheme, always includes a diffusion part through $\exp(Z_k/\sqrt{n/T} V_1)$, with Z_k satisfying (12) and the ellipticity condition (22) for V_1. The proof is then a consequence from Theorem 3.3 in [3] which employs an abstract Malliavin calculus based on such noise Z_k and initially presented by Bally and Clément [2]. A similar approach can be used in order to prove the convergence for the total variation distance for even higher order scheme built as in (11). The main difficulty will then rely on the proof of the short time estimate (4).

A main interest of this result is that it can be seen as an invariance principle as well. Indeed, it does not require that Z_k follows a particular law but only the properties (12) and (13). In particular, we do not restrict ourselves to the Gaussian framework which is necessary to use the Malliavin Calculus in order to proove the convergence for the total variation distance as in [4], [5], or [17]. In this way, the condition (12) might be a hint to find a necessary condition on the random variables $(Z_k)_{k\in\mathbb{N}^*}$ in order to obtain the total variation convergence with order $h=3$.

Moreover, using Remark 3.2, we can define a third order scheme as soon as the function defined in (19) is monotonic. If it is increasing (recall that $V_1(x)\geqslant 0$), the Lie bracket between V_1^2 and V_0 is given by $[V_1^2,V_0]f=V_1^2V_0f-V_0V_1^2f=2\tilde{V}^2f$ with

$$\tilde{V}(x)=\sqrt{|V_1(x)(V_1(x)\partial_x V_0(x)-\partial_x V_1(x)V_0(x))|},\quad x\in\mathbb{R}.$$

If it is decreasing, we have $[V_0,V_1^2]f=2\tilde{V}^2f$ as well. This explicit representation for $\tilde{V}$ is crucial for concrete applications since the scheme is defined using the solution of (8) with $V=\tilde{V}$. Moreover, looking at (18) and (23), we have to control its derivatives.

4 Numerical illustration

In this section, we study the numerical approximation of a one dimensional SDE with schemes defined on homogeneous time grids with form $\pi_{T,n}=\{kT/n,k\in\mathbb{N}\}$. We will fix T and we will analyze the behavior of the total variation distance between the diffusion process $(X_t)_{t\geqslant 0}$ and miscellaneous discretization schemes $(X_t^n)_{t\in\pi_{T,n}}$ with respect to the number of time step n. More particularly, we will study the weak error $|\mathbb{E}[f(X_t)]-\mathbb{E}[f(X_t^n)]|$ for bounded measurable functions f and various n.

In concrete applications, once we have selected a scheme X^n, $\mathbb{E}[f(X_t^n)]$ will be used to estimate $\mathbb{E}[f(X_t)]$. The next step is thus to approximate $\mathbb{E}[f(X_t^n)]$. A standard way to do it, is to use a Monte Carlo method. Given an independent sampling of size M, and using the Central Limit Theorem, we can easily show that those algorithms converge toward the real expectancy with rate $\sqrt{M}$. Moreover, discretization schemes provide an estimation of $\mathbb{E}[f(X_t)]$ with any desired precision since we can choose any value for n. However, the cost of calculation will also increase with n since we have n iterations of the scheme function (2). At this point, it is important to notice that there is a trade off to make between the precision we want to obtain and the time of calculation we can afford. Indeed, if our scheme converges with order h, we have to choose $M=\mathcal{O}(n^{2h})$ and then choose n large enough in order to obtain the desired precision. We will see that even if the time of calculation of one step of the scheme we study in this paper is much longer than the time of a lower order scheme (*e.g.* the Euler scheme), the third weak order scheme is better in time of calculation and precision as soon as the precision is high enough. In order to illustrate the reason why we point out such properties, we now present our example.

We consider the Markov diffusion process $(X_t)_{t\geqslant 0}$ given by the following SDE,

$$dX_t=adt+\frac{\sigma}{\arctan(X_t)+\pi}\circ dW_t, \tag{25}$$

with $\sigma>0$ and $a\in\mathbb{R}$. Notice that the coefficients of the SDE (25) belong to $\mathcal{C}_b^\infty(\mathbb{R})$ and moreover $V_1:x\mapsto\sigma/(\arctan(x)+\pi)$ satisfies $\inf_x V_1(x)>2\sigma/\pi$ and the function $V_0/V_1=a/V_1$ is increasing. Therefore, the scheme (11) is well-defined and we have the required hypothesis in order to obtain the results from Theorem 3.2. Moreover, we have an explicit representation for the first order operator $\tilde{V}$, that is : $\tilde{V}(x) = \sigma\sqrt{a}/(\sqrt{1+x^2}(\arctan(x)+\pi)^{3/2})$.

The next step consists then in solving the ODE (8) for $V = V_0, V_1, \tilde{V}$. Looking closer to (11), we will use each of these solutions once for each step of the discretization algorithm. In this example, it is easy to find an analytic solution to (8) when $V = V_0$. However for $V = V_1, \tilde{V}$, it is much more cumbersome and we will use some numerical algorithms. A naive algorithm consists in using the Riemann approximation of $\int_0^t V(\Phi_V(x,s))ds$ on a time grid of $[0,t]$ in the following way : For a number N of time steps, we put $\Phi_V^N(x,0) = x$ and for $i \in \{0,\dots,N-1\}$, $\Phi_V^N(x,(i+1)t/N) = \Phi_V^N(x,it/N) + TN^{-1}V(\Phi_V^N(x,it/N))$. This is the method we will use to approximate $\Phi_{\tilde{V}}^N(x,t)$. Finally, in this case, we can use an alternative way to approximate $\Phi_{V_1}^N(x,t)$. Indeed, we can show that $g(\Phi_{V_1}^N(x,t)) = g(x)+t$ where g is the bijective function in $\mathcal{C}^1(\mathbb{R})$ defined by

$$g(x) = (x\arctan(x) - 0.5\log(1+x^2) + x\pi)/\sigma.$$

Then, we can find $\Phi_{V_1}(x,t)$ using a Newton algorithm in order to invert g. Likewise the naive Riemann approximation, this method provides an approximation given a parameter of precision (which is N for Riemann sums). Obviously, the more this parameter is tight, the more the cost of the algorithm is high. Compared to one step the Euler scheme,

$$X_{t_{k+1}^n}^{n,Eul} = X_{t_k^n}^{n,Eul} + \left(a - \frac{\sigma^2}{2(1+(X_{t_k^n}^{n,Eul})^2)(\arctan(X_{t_k^n}^{n,Eul})+\pi)^3}\right)T/n \qquad (26)$$
$$+ \frac{\sigma}{\arctan(X_{t_k^n}^{n,Eul})+\pi}\sqrt{T/N}Z_{k+1}, \quad (Z_k)_{k\in\mathbb{N}^*}\ i.i.d \sim \mathcal{N}(0,1),$$

the cost of one step of (11) can thus be very important. However, despite that cost, the third order scheme become more effective as soon as we want to compute $\mathbb{E}[f(X_t)]$ with a sufficiently high precision.
Heuristically, let $\epsilon > 0$, the precision of the weak error that is $|\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]| \leqslant \epsilon$. In order to reach that precision, we will have to run $M = \epsilon^{-2}$ Monte Carlo iterations. Now let $n \in \mathbb{N}$ such that $n^3 = \epsilon^{-1}$. Then, if we want to reach this precision, we will have to simulate $M = \epsilon^{-2}$ realizations of the third order scheme with time step t/n, or of the Euler scheme with time step t/n^3. Now, we assume that the cost in time of calculation of one step of the third order scheme is given by τ_{NV3} and by τ_{Eul} for the Euler scheme. Then the total cost to reach the precision ϵ will be $\tau_{NV3}nM = \tau_{NV3}\epsilon^{-2-1/3}$ for the third order scheme and $\tau_{Eul}n^3M = \tau_{Eul}\epsilon^{-3}$ for the Euler scheme. Then, as soon as $\tau_{NV3}/\tau_{Eul} \leqslant \epsilon^{-2/3}$, the cost of the third order scheme will be lower than the cost of the Euler scheme. Controversially, if τ_{NV3} and τ_{Eul} are fixed we can find a precision ϵ_0 such that the cost of the three order scheme is lower than the one of the Euler scheme for all $\epsilon \leqslant \epsilon_0$.

In Figure 1, we represent the error $|\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]|$ [2], with respect to the number of time steps n, in Log Log scale, for the third order scheme we study in this paper and when f

[2]We do not estimate $\mathbb{E}[f(X_t)]$ using Monte Carlo methods with exact simulation but with the third order scheme for $n = 50$.

is a Heavyside function. We observe that the scheme converges with the expected rate, that is $h \approx 2.91$. This numerical experiment thus confirms the total variation convergence result from Theorem 3.2. Notice that we have also implemented the Euler scheme and the Ninomiya Victoir scheme of order 2 [20] in order to compare the cost of the different approaches. With the precision parameters we have selected in the algorithms solving (8) in order to obtain Figure 1, we have $\tau_{NV3} \approx 7.8\tau_{NV2} \approx 51.9\tau_{Eul}$ which is quite reasonable given the gain which is made with respect to the number of time steps. In this case, the third order scheme thus become more effective than the Euler scheme as soon as the precision ϵ of the weak error satisfies $|\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]| \leqslant \epsilon \leqslant (51.9)^{-3/2}$.

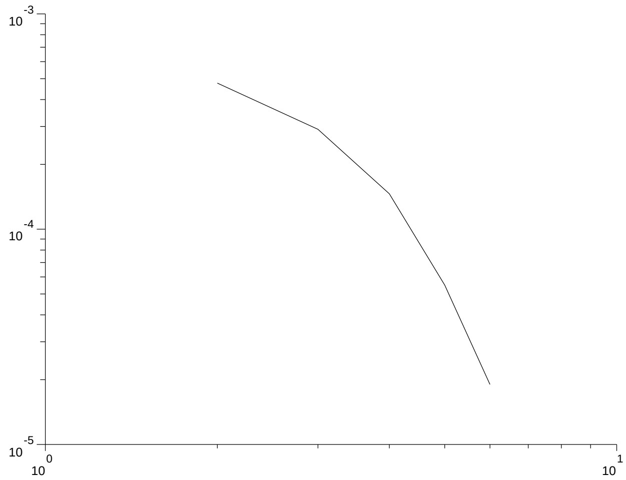

Figure 1: Log-Log representation of $|\mathbb{E}[f(X_t)] - \mathbb{E}[f(X_t^n)]|$ for $x = 0.8$, $T = 1$, $a = 0.2$, $\sigma = 2$, with respect to n for $f(x) = \mathbb{1}_{x>1.1}$.

5 Proof of the main theorems

Proof of Theorem 3.1. **Step 1.** We define $(P_{t,s})_{t,s\in\pi_{T,n};t\leqslant s}$ by

$$P^n_{t,t}f(x) = f(x), \qquad \forall k \leqslant r \in \mathbb{N},\ P^n_{t^n_k,t^r_k}f = \mathbb{E}[f(X_{t^n_r})|X_{t^n_k} = x],$$
$$Q^n_{t,t}f(x) = f(x), \qquad \forall k \leqslant r \in \mathbb{N},\ Q^n_{t^n_k,t^r_k}f = \mathbb{E}[f(X^n_{t^n_r})|X^n_{t^n_k} = x],$$

and we notice that for $t, s, u \in \pi_{T,n}$ with $t \leqslant s \leqslant u$, then $P^n_{t,u}f = P^n_{t,s}P^n_{s,u}f$. It follows that

$$\begin{aligned}\left|\mathbb{E}[f(X_{t^n_m})] - \mathbb{E}[f(X^n_{t^n_m})]\right| &\leqslant \|P^n_{0,t^n_m}f - Q^n_{0,t^n_m}f\|_\infty \qquad (27)\\ &\leqslant \sum_{k=0}^{m-1} \|P^n_{t^n_{k+1},t^n_m}P^n_{t^n_k,t^n_{k+1}}Q_{t^n_{k+1}}f - P^n_{t^n_{k+1},t^n_m}Q_{t^n_k,t^n_{k+1}}Q^n_{t^n_k}f\|_\infty\\ &= \sum_{k=0}^{m-1} \|P_{t^n_{k+1},t^n_m}(P_{t^n_k,t^n_{k+1}} - Q_{t^n_k,t^n_{k+1}})Q_{t^n_k}f\|_\infty.\end{aligned}$$

$$\|P^n_{t^n_m} f - Q^n_{t^n_m} f\|_\infty \leqslant \sum_{k=0}^{m-1} \|P^n_{t^n_k} P^n_{t^n_k,t^n_{k+1}} Q^n_{t^n_{k+1},t^n_m} f - P^n_{t^n_k} Q^n_{t^n_k,t^n_{k+1}} Q^n_{t^n_{k+1},t^n_m} f\|_\infty \quad (28)$$
$$= \sum_{k=0}^{m-1} \|P^n_{t^n_k}(P_{t^n_k,t^n_{k+1}} - Q_{t^n_k,t^n_{k+1}}) Q^n_{t^n_{k+1},t^n_m} f\|_\infty .$$

We notice that it easy to prove that, for $t, s \in \pi_{T,n}$, $t \leqslant s$, $\|P_{t,s}f\|_{p,\infty} \leqslant C\|f\|_{p,\infty}$ and $\|Q_{t,s}f\|_{p,\infty} \leqslant C\|f\|_{p,\infty}$.

Step 2. It remains to show $\|P_{t^n_k,t^n_{k+1}} f - Q_{t^n_k,t^n_{k+1}} f\|_\infty \leqslant C\|f\|_{8,\infty}/n^4$ and using (28) the proof will be completed. In order to simplify the notation, we fix $T = 1$ without loss of generality. For $\epsilon = -1, 1$, we denote

$$\mathcal{T}_0 f(x) = f(\exp(\frac{1}{n}V_0)(x)), \quad \mathcal{T}_1 f(x) = f(\exp(\frac{Z}{\sqrt{n}}V_1)(x)), \quad \tilde{\mathcal{T}}_\epsilon f(x) = f(\exp(\frac{\epsilon}{n}\tilde{V})(x)).$$

Notice that, using the notation introduced in the beginning of this section with $V = n^{-1/2}ZV_1$, we have $\mathcal{T}_1 f(x) = P_1^{n^{-1/2}ZV_1} f(x)$. Using (9) with $t = 1$ and $V = n^{-1/2}ZV_1$ we obtain

$$\mathcal{T}_1 f(x) = f(x) + \sum_{r=1}^{m} \frac{Z^r}{n^{r/2}} \frac{1}{r!} V_1^r f(x) + \frac{Z^{m+1}}{n^{(m+1)/2}} R_{m+1,1} f(x) \quad (29)$$

with

$$R_{m+1,1} f(x) = \frac{1}{m!} \int_0^1 (1-\lambda)^m V_1^{m+1} P_\lambda^{n^{-1/2}ZV_1} f(x) d\lambda \quad (30)$$

and we recall that $P_\lambda^{n^{-1/2}ZV_1} f(x) = f(\exp(\lambda Z V_1/\sqrt{n}))$. We have a similar development if we put $V = V_0/n$ or $V = \epsilon\tilde{V}/n$ in (9). Our aim is to give a development of order 4 (with respect to n) for $\mathbb{E}[f(\psi_k(x, w^1_{k+1}, w^0_{k+1}))]$ (see (31) below). We replace each $\mathcal{T} \in \{\mathcal{T}_0, \mathcal{T}_1, \tilde{\mathcal{T}}_\epsilon\}$, with an expansion of order $m \leqslant 7$ given above with $Z = Z_{k+1}$ for $\mathcal{T}_1$ and $m \leqslant 3$ for $\mathcal{T} = \mathcal{T}_0, \tilde{\mathcal{T}}$. Then, we calculate the products of the miscellaneous expansions, each with a well chosen order such that there is no term with factor n^{-r}, $r > 4$, appearing in those productss. Moreover, all the terms containing n^{-4} go in the remainder. The last step consists in computing the expectancy. We notice that $\mathbb{E}[P_t^{n^{-1/2}ZV_1}] = P_t^{V_1^2/(2n)}$ and $\mathbb{E}[Z^r_{k+1}] = 0$ for odd $r \leqslant 7$. Finally, since $\mathbb{E}[Z^2_{k+1}] = 1$, $\mathbb{E}[Z^4_{k+1}] = 6$ and $\mathbb{E}[Z^6_{k+1}] = 15$, the calculus is completed and we obtain:

$$\mathbb{E}[f(\psi_k(x, w^1_{k+1}, w^0_{k+1}))] = \frac{1}{6} \sum_{\epsilon=-1,1} \mathbb{E}[(\tilde{\mathcal{T}}_\epsilon \mathcal{T}_0 \mathcal{T}_1 + \mathcal{T}_0 \tilde{\mathcal{T}}_\epsilon \mathcal{T}_1 + \mathcal{T}_0 \mathcal{T}_1 \tilde{\mathcal{T}}_\epsilon) f(x)] \quad (31)$$
$$= f(x) + \frac{1}{n}\big(V_0 + \frac{1}{2}V_1^2\big) f(x) + \frac{1}{2n^2}\big(V_0^2 + \frac{1}{4}V_1^4 + 2V_0\frac{1}{2}V_1^2 + \tilde{V}^2\big) f(x)$$
$$+ \frac{1}{6n^3}\big(\frac{1}{8}V_1^6 + V_0^3 + 3V_0\frac{1}{4}V_1^4 + 3V_0^2\frac{1}{2}V_1^2 + 2\tilde{V}^2\frac{1}{2}V_1 + 2V_0\tilde{V}^2 + \frac{1}{2}V_1^2\tilde{V}^2 + \tilde{V}^2 V_0\big) f(x)$$
$$+ \frac{1}{n^4} R f(x)$$
$$= f(x) + \frac{1}{n}\big(V_0 + \frac{1}{2}V_1^2\big) f(x) + \frac{1}{2n^2}\big(V_0 + \frac{1}{2}V_1^2\big)^2 f(x) + \frac{1}{6n^3}\big(V_0 + \frac{1}{2}V_1^2\big)^3 f(x) + \frac{1}{n^4} R f(x)$$

The remainder R is a sum of terms of the following form:

$$C(\tilde{\mathcal{T}}_{\epsilon,\alpha_\epsilon}\mathcal{T}_{0,\alpha_0}\mathcal{T}_{1,\alpha_1}+\mathcal{T}_{0,\alpha_0}\tilde{\mathcal{T}}_{\epsilon,\alpha_\epsilon}\mathcal{T}_{1,\alpha_1}+\mathcal{T}_{0,\alpha_0}\mathcal{T}_1\tilde{\mathcal{T}}_{\epsilon,\alpha_\epsilon})f(x) \tag{32}$$

with $\alpha=(\alpha_0,\alpha_1,\alpha_2)\in\{0,\dots,4\}^3$, $|\alpha|=\alpha_0+\alpha_1+\alpha_2=4$, and, using the notation given in (30),

$$\begin{aligned}
\mathcal{T}_{0,k}\in\{V_0^k,R_{k,0}\},\quad \tilde{\mathcal{T}}_{\epsilon,k}\in\{\tilde{V}^k,\tilde{R}_{k,\epsilon}\},\qquad &\mathcal{T}_{1,k}\in\{V_1^{2k},R_{2k,1}\},\quad k=0,\dots,3,\\
\mathcal{T}_{0,4}=R_{4,0},\quad \tilde{\mathcal{T}}_{\epsilon,4}=\tilde{R}_{4,\epsilon},\qquad &\mathcal{T}_{1,4}=\mathbf{R}_{8,1},
\end{aligned}$$

with

$$\mathbf{R}_{8,1}=\mathbb{E}[Z^8R_{8,1}]=\int_0^1(1-\lambda)^7\mathbb{E}[Z^8V_1^8P_\lambda^{U_1}f(x)]d\lambda.$$

It is easy to check that for every $g\in\mathcal{C}^{k+p}(\mathbb{R})$, we have the following property

$$\|\mathcal{T}_{0,k}g\|_{p,\infty}+\|\mathcal{T}_{1,k}g\|_{p,\infty}+\|\tilde{\mathcal{T}}_{\epsilon,k}g\|_{p,\infty}\leqslant CC_{2k+p}(V)^l\|g\|_{2k+p,\infty}$$

for some constants $l\in\mathbb{N}^*$, $C\geqslant 1$. So

$$\|Rf\|_\infty\leqslant CC_8(V)^l\|f\|_{8,\infty}. \tag{33}$$

We turn now to the diffusion process X_t. We have the development

$$\mathbb{E}[f(X_t(x))]=P_t^Af(x)=f(x)+tAf(x)+\frac{t^2}{2}A^2f(x)+\frac{t^3}{6}A^3(x)+\frac{t^4}{4!}R'_tf(x).$$

with

$$R'_tf(x)=t^{-1}\int_0^tP_\lambda^AA^4f(x)(1-\lambda/t)^3d\lambda. \tag{34}$$

We take $t=n^{-1}$ and make the difference between (34) and (31). All the terms cancel except for the remainders so we obtain

$$\begin{aligned}
&\forall k\in\{0,\dots,n-1\},\\
&\quad\mathbb{E}[f(X_{t^n_{k+1}})]-\mathbb{E}[f(X^n_{t^n_{k+1}})\mid X_{t^n_k}=X^n_{t^n_k}=x]=(R'_{1/n}f(x)/4!-Rf(x))/n^4.
\end{aligned} \tag{35}$$

We clearly have $\|R'_{1/n}f\|_\infty\leqslant CC_8(V)^l\|f\|_{8,\infty}$. This, together with (33) and (28), completes the proof. □

Proof of Theorem 3.2. **Step 1.** Let us prove that (10) is satisfied. We have

$$\begin{aligned}
\frac{1}{2}(V_1^2V_0-V_0V_1^2)f(x)=&\big(\partial_xV_1(x)(V_1(x)\partial_xV_0(x)-\partial_xV_1(x)V_0(x))\\
&+V_1(x)(\partial_x^2V_0(x)V_1(x)-\partial_x^2V_1(x)V_0(x))\big)\partial_xf(x)\\
&+V_1(x)(V_1(x)\partial_xV_0(x)-\partial_xV_1(x)V_0(x))\partial_x^2f(x).
\end{aligned}$$

Since $V_1(x) \neq 0$, if we take

$$\tilde{V}(x) = \sqrt{V_1(x)(V_1(x)\partial_x V_0(x) - \partial_x V_1(x) V_0(x))}, \tag{36}$$

then, using (19), $\tilde{V}$ is well defined and satisfies (10).

Step 2. Now we are going to show the convergence in total variation distance. In order to do it we will use a result from [3]. First, applying the same reasoning as in the proof of Theorem 3.1 we can show that there exists some universal constants $C, l \geqslant 1$ such that

$$|\langle g, P^n_{t^n_k, t^n_{k+1}} f - Q_{t^n_k, t^n_{k+1}} f \rangle| \leqslant n^{-4} C C_8(V)^l \|g\|_{1,8} \|f\|_\infty, \tag{37}$$

with $\langle \cdot, \cdot \rangle$ the scalar product in $\mathrm{L}^2(\mathbb{R})$. Now we have (35) and (37), the result will be a consequence of Theorem 3.3. in [3], as soon as we check that the following ellipticity assumption holds:

$$\exists \lambda_* > 0, \qquad \inf_{k \leqslant n} \inf_{x \in \mathbb{R}} (\partial_{w^1} \psi_k(x, w^1, w^0)_{|w^1 = w^0 = 0})^2 \geqslant \lambda_*. \tag{38}$$

We fix k and we look at $\psi_k(x, w^1, w^0)$ defined in (15). We suppose that $\rho_k = 3, \epsilon_k = 1$ (the proof for $\rho_k = 1, 2$ or $\epsilon_k = -1$ is similar). We consider the process $x_t(\tilde{w}), 0 \leqslant t \leqslant T_3$, with $T_i = i$, $\tilde{w} = (w^1, w^0)$, solution of the following equation:

$$\begin{aligned}
x_t(\tilde{w}) &= x + w^0 \int_0^t \tilde{V}(x_s(\tilde{w})) ds, \qquad T_0 \leqslant t \leqslant T_1, \\
x_t(\tilde{w}) &= x_{T_1}(\tilde{w}) + w^1 \int_{T_1}^t V_1(x_s(\tilde{w})) ds, \qquad T_1 \leqslant t \leqslant T_2, \\
x_t(\tilde{w}) &= x_{T_2}(\tilde{w}) + w^0 \int_{T_2}^t V_0(x_s(\tilde{w})) ds, \qquad T_2 \leqslant t \leqslant T_3.
\end{aligned}$$

We notice that $\psi_k(x, w^1_{k+1}, w^0_{k+1}) = x_{T_3}(\tilde{w}_{k+1})$ and consequently, we have $\partial_z \psi_k(x, w^1_{k+1}, w^0_{k+1}) = \partial_{w^1} x_{T_3}(\tilde{w}_{k+1})$. Moreover, $\partial_{w^1} x_t(w) = 0$ for $t \leqslant T_1$. Now, let $T_1 \leqslant t \leqslant T_2$. Then $\partial_{w^1} x_t(\tilde{w})$ solves the equation

$$\partial_{w^1} x_t(\tilde{w}) = \int_{T_1}^t V_1(x_s(\tilde{w})) ds + w^1 \int_{T_1}^t \partial V_1(x_s(\tilde{w})) \partial_{w^1} x_s(\tilde{w}) ds.$$

It follows that

$$\partial_{w^1} x_t(\tilde{w}) \mid_{\tilde{w}=0} = \int_{T_1}^t V_1(x_s(0)) ds = V_1(x)(t - T_1).$$

Notice that $T_2 - T_1 = 1$. Then, we have

$$\partial_{w^1} x_{T_3}(\tilde{w}) \mid_{w^1=0} = \partial_{w^1} x_{T_2}(\tilde{w}) \mid_{\tilde{w}=0} = V_1(x).$$

and then, by (22),

$$(\partial_{w^1} x_{T_3}(0))^2 \geqslant \lambda_*.$$

□

References

[1] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.

[2] V. Bally and E. Clément. Integration by parts formula and applications to equations with jumps. *Probab. Theory Related Fields*, 151(3-4):613–657, 2011.

[3] V. Bally and C. Rey. Approximation of Markov semigroup in total variation disctance. January 2015.

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

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

[6] M. Bossy, E. Gobet, and D. Talay. A symmetrized Euler scheme for an efficient approximation of reflected diffusions. *J. Appl. Probab.*, 41(3):877–889, 2004.

[7] E. Gobet. Weak approximation of killed diffusion using Euler schemes. *Stochastic Process. Appl.*, 87(2):167–197, 2000.

[8] Emmanuel Gobet and Stéphane Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Process. Appl.*, 120(2):130–162, 2010.

[9] J. Guyon. Euler scheme and tempered distributions. *Stochastic Process. Appl.*, 116(6):877–904, 2006.

[10] J. Jacod, T. G. Kurtz, S. Méléard, and P Protter. The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(3):523–558, 2005.

[11] B. Jourdain and A. Kohatsu-Higa. *A Review of Recent Results on Approximation of Solutions of Stochastic Differential Equations*, volume 65 of *Progress in Probability*. Springer Basel, 2011.

[12] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[13] A. Kohatsu-Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Process. Appl.*, 120(11):2258–2285, 2010.

[14] V. Konakov and S. Menozzi. Weak error for stable driven stochastic differential equations: expansion of the densities. *J. Theoret. Probab.*, 24(2):454–478, 2011.

[15] V. Konakov, S. Menozzi, and S. Molchanov. Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(4):908—923, 2010.

[16] S. Kusuoka. Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in mathematical economics. Vol. 6*, volume 6 of *Adv. Math. Econ.*, pages 69–83. Springer, Tokyo, 2004.

[17] S. Kusuoka. Gaussian K-scheme: justification for KLNV method. In *Advances in mathematical economics. Vol. 17*, volume 17 of *Adv. Math. Econ.*, pages 71–120. Springer, Tokyo, 2013.

[18] T. Lyons and N. Victoir. Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):169–198, 2004. Stochastic analysis with applications to mathematical finance.

[19] G.N. Milstein. Weak approximation of solutions of systems of stochastic differential equations. In *Numerical Integration of Stochastic Differential Equations*, volume 313 of *Mathematics and Its Applications*, pages 101–134. Springer Netherlands, 1995.

[20] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15(1-2):107–121, 2008.

[21] P. Protter and D. Talay. The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.*, 25(1):393–423, 1997.

[22] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.