High order discretization schemes for the CIR process: application to Affine Term Structure and Heston models

Aurélien Alfonsi

High order discretization schemes for the CIR process: application to Affine Term Structure and Heston models

Aurélien Alfonsi

CERMICS, MATHFI project, Ecole des Ponts, 6-8 avenue Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne-la-vallée, France.
e-mail : alfonsi@cermics.enpc.fr
June 18, 2008

Abstract

This paper presents weak second and third order schemes for the Cox-Ingersoll-Ross (CIR) process, without any restriction on its parameters. At the same time, it gives a general recursive construction method to get weak second-order schemes that extends the one introduced by Ninomiya and Victoir [17]. Combining these both results, this allows to propose a second-order scheme for more general affine diffusions. Simulation examples are given to illustrate the convergence of these schemes on CIR and Heston models. Algorithms are stated in a pseudocode language.

Keywords: simulation, discretization scheme, squared Bessel process, Cox-Ingersoll-Ross model, Heston model, Affine Term Structure Models (ATSM).

Acknowledgments. Most part of this work has been done when I was at the TU Berlin, thanks to the support of MATHEON. I would like to thank Vlad Bally (Univ. Marne-la-Vallée) and Benjamin Jourdain (Ecole des Ponts) for fruitful comments, and Victor Reutenauer (CALyon) for stimulating discussions on ATSM.

Introduction

In this paper, we are interested in discretization schemes for the Cox-Ingersoll-Ross process (CIR for short), and more generally for multidimensional diffusion processes that contain a square-root diffusion coefficient, like in Affine Term Structure and Heston models [9, 13]. Initially introduced in 1985 to model the short interest rate [8], the CIR process is now widely used in finance because it presents interesting qualitative features such as positivity

and mean-reversion. Moreover, it belongs to the class of affine models for which some standard expectations are analytically or semi-analytically known. We will use in this paper the following parametrization of the CIR process

$$\begin{cases} X_t^x = x + \int_0^t (a - kX_s^x)ds + \sigma \int_0^t \sqrt{X_s^x} dW_s, t \in [0,T] \\ x \geq 0 \end{cases} \tag{1}$$

with parameters $(a, k, \sigma) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+$. It is a nonnegative process. Moreover, if $x > 0$ and $2a \geq \sigma^2$ the process $(X_t, t \geq 0)$ is always positive. We will exclude the trivial case $\sigma = 0$ and assume $\sigma > 0$ in the whole paper.

First, let us say that exact simulation methods exist for the CIR process (see Glasserman [12]) and also for the Heston model (Broadie and Kaya [7]). With respect to discretization schemes, the drawback of these exact simulation methods is the computation time that they require. This is analysed in Alfonsi [1], Broadie and Kaya [7], and Lord, Koekkoek and van Dijk [16]. What comes out is that exact methods are competitive when one has to simulate the process just at one time (or few times), for example to compute European options prices with a Monte-Carlo algorithm. On the contrary, they are drastically too slow if one has to simulate the process along a time-grid, which occurs when computing pathwise options prices. At least for that reason, it is worth studying discretization schemes for square-root SDEs.

The main difficulty when discretizing the CIR process is located in 0, where the square-root is not Lipschitzian. Usual schemes such as the Euler scheme or the Milstein scheme are in general not well defined. They can indeed lead to negative values for which the square root is not defined. One has therefore to modify them or to create ad-hoc schemes. Discretization schemes dedicated to square-root diffusions have thus been studied in the recent years by Deelstra and Delbaen [10], Bossy, Diop and Berkaoui ([11, 5]), Alfonsi [1], Kahl and Schurz [15], Lord, Koekoek and van Dijk [16] and recently Andersen [2]. A possible criteria to chose the scheme may be its capacity to support large values of σ (we mean here $\sigma^2 \gg 4a$). In finance, such large values do not occur when the CIR diffusion is used to represent the short interest rate. They are instead often observed when the CIR stands for the default intensity in credit risk or the stock volatility like in the Heston model (see [6] and [2] for numerical examples in these three cases). Heuristically, the larger is σ, the more the CIR process spends time in the neighbourhood of 0 where the square-root is very sensitive. This is intuitively why most of the schemes fail to be accurate for large σ. The QE scheme proposed by Andersen is in fact the only one among those cited that is really well suited for these large values, but no theoretical convergence result is given for this scheme. In another direction, Ninomiya and Victoir [17] have proposed recently a general method to get weak second-order discretization schemes for a broad class of multidimensional SDEs. We will present their method in detail in the first part. They apply it to the Heston model and get encouraging results in that case, but once again it is restricted to small values of σ, because their scheme may not be defined for $\sigma^2 > 4a$.

The main contribution of this paper is to present very efficient schemes for general affine diffusions, without any restriction on the parameters. More precisely, we introduce second

and third order discretization schemes for the CIR process that support even large values of σ. We give theoretical results of convergence for these schemes, analyzing the weak error. Moreover, we also present a simple recursive method to construct second order schemes in a general framework that encompasses affine diffusions. This method allows to get a second order scheme for Affine Term Structure Models (ATSM) and also a very efficient scheme for the Heston model. Again, these schemes support large volatility coefficient values. The paper is structured as follows.

The first part introduces notations and assumptions. It presents the analysis of the weak error made by Talay and Tubaro [20] and then gives a recursive construction of second order schemes for multidimensional SDEs that extends the results of Ninomiya and Victoir [17]. This method relies on the idea of scheme composition, that dates back to Strang [18] in the field of ODEs. Let us emphasize here that most of result presented in this part are already known, but usually with $\mathcal{C}^\infty$ SDE coefficients with uniformly bounded derivatives, which is not satisfied by CIR and more general affine diffusions. The scope of this part is thus to give a rigorous framework for the weak error analysis that embeds affine diffusions. The second and third parts are respectively devoted to the construction of a weak second and third order discretization scheme for the CIR. In both cases, our solution consists in switching scheme near the origin in order to keep nonnegativity. The fourth part presents schemes for ATSM and Heston models, putting into practice the general results of the first part. Algorithms coming from these schemes are described in a pseudocode language. Simulations results are gathered in the last part for the CIR process and for the Heston model. European, Asian and exotic options prices are in particular computed. The numerical behaviour of these schemes is really satisfactory.

1 Second order discretization schemes for SDEs.

1.1 Assumptions on the SDE and notations

We consider a d_W-dimensional standard Brownian motion $(W_t, t \geq 0)$ and will denote in the sequel $(\mathcal{F}_t)_{t\geq 0}$ its augmented associated filtration that satisfies the usual conditions. Let $d \in \mathbb{N}^*$, and $\mathbb{D} \subset \mathbb{R}^d$ a domain that we assume for sake of simplicity to be a product of d intervals. Typically, we will consider $\mathbb{D} = \mathbb{R}_+^{d_1} \times \mathbb{R}^{d_2}$ with $d_1 + d_2 = d$ in this paper. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, we define $\partial_\alpha = \partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d}$ and $|\alpha| = \sum_{l=1}^d \alpha_l$. We introduce the following functional space:

$$\mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D}) \;=\; \{f \in \mathcal{C}^\infty(\mathbb{D}, \mathbb{R}), \forall \alpha \in \mathbb{N}^d, \exists C_\alpha > 0, e_\alpha \in \mathbb{N}^*, \forall x \in \mathbb{D},\ |\partial_\alpha f(x)| \leq C_\alpha(1 + \|x\|^{e_\alpha})\}$$

where $\|.\|$ is a norm on $\mathbb{R}^d$. We will say that $(C_\alpha, e_\alpha)_{\alpha\in\mathbb{N}^d}$ is a good sequence for $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ if one has $\forall x \in \mathbb{D}, |\partial_\alpha f(x)| \leq C_\alpha(1 + \|x\|^{e_\alpha})$.

Assumptions. *We assume that $b : \mathbb{D} \to \mathbb{R}^d$ and $\sigma : \mathbb{D} \to \mathcal{M}_{d\times d_W}(\mathbb{R})$ are such that for $1 \leq i, j \leq d$, the functions $x \in \mathbb{D} \mapsto b_i(x)$ and $x \in \mathbb{D} \mapsto (\sigma\sigma^*)_{i,j}(x)$ are in $\mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$. For*

$x \in \mathbb{D}$, we introduce the general $\mathbb{R}^d$-valued SDE:

$$t \geq 0,\ X_t^x = x + \int_0^t b(X_s^x)ds + \int_0^t \sigma(X_s^x)dW_s. \tag{2}$$

We assume that for any $x \in \mathbb{D}$, there is a unique weak solution defined for $t \geq 0$, and therefore

$$\mathbb{P}(\forall t \geq 0, X_t^x \in \mathbb{D}) = 1.$$

It satisfies then the strong Markov property (Theorem 4.20, p. 322 in [14]). The differential operator associated to the SDE is given by

$$f \in \mathcal{C}^2(\mathbb{D}, \mathbb{R}),\ Lf(x) = \sum_{i=1}^{d} b_i(x)\partial_i f(x) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{d_W} \sigma_{i,k}(x)\sigma_{j,k}(x)\partial_i\partial_j f(x). \tag{3}$$

If $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$, thanks to the regularity assumptions made on b and σ, all the iterated functions $L^k f(x)$ are well defined on $\mathbb{D}$ and belong to $\mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ for any $k \in \mathbb{N}$.

Definition 1.1. *We will say (for short) that the operator L satisfies the required assumptions on $\mathbb{D}$ if it is defined by (3) for some functions $b(x)$ and $\sigma(x)$ and satisfies all the assumptions above.*

Now, let us turn to discretization schemes for the SDE (2). Let us fix a time horizon $T > 0$. We will consider in the whole paper the time interval $[0, T]$ and the regular time discretization $t_i^n = iT/n$ for $i = 0, 1, \ldots, n$.

Definition 1.2. *A family of transition probabilities $(\hat{p}_x(t)(dz), t > 0, x \in \mathbb{D})$ on $\mathbb{D}$ is such that $\hat{p}_x(t)$ is a probability law on $\mathbb{D}$ for $t > 0$ and $x \in \mathbb{D}$.*

A discretization scheme with transition probabilities $(\hat{p}_x(t)(dz), t > 0, x \in \mathbb{D})$ is a sequence $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ of $\mathbb{D}$-valued random variables such that:

- *for $0 \leq i \leq n$, $\hat{X}^n_{t_i^n}$ is a $\mathcal{F}_{t_i^n}$-measurable random variable on $\mathbb{D}$,*
- *the law of $\hat{X}^n_{t_{i+1}^n}$ is given by $\mathbb{E}[f(\hat{X}^n_{t_{i+1}^n})|\mathcal{F}_{t_i^n}] = \int_{\mathbb{D}} f(z)\hat{p}_{\hat{X}^n_{t_i^n}}(T/n)(dz)$ and thus only depends on $\hat{X}^n_{t_i^n}$ and T/n.*

For convenience, we will denote, for $t > 0$ and $x \in \mathbb{D}$, $\hat{X}_t^x$ a random variable distributed according to the probability law $\hat{p}_x(t)(dz)$. The law of a discretization scheme $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ is thus entirely determined by its initial value and its transition probabilities. Since the initial value is quite always taken equal to the initial value of the SDE, we will identify with a slight abuse of language the scheme $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ with its transition probabilities $(\hat{p}_x(t)(dz)$ or $\hat{X}_t^x)$.

Definition 1.3. *Let us denote $\mathcal{C}^\infty_K(\mathbb{D}, \mathbb{R})$ the set of the $\mathcal{C}^\infty$ real valued functions with a compact support in $\mathbb{D}$. Let $x \in \mathbb{D}$. A discretization scheme $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ is a weak νth-order scheme for the SDE $(X_t^x, t \in [0, T])$ if:*

$$\forall f \in \mathcal{C}^\infty_K(\mathbb{D}, \mathbb{R}), \exists K > 0,\ |\mathbb{E}(f(X_T^x)) - \mathbb{E}(f(\hat{X}^n_{t_n^n}))| \leq K/n^\nu.$$

The quantity $\mathbb{E}(f(X_T^x)) - \mathbb{E}(f(\hat{X}^n_{t_n^n}))$ is called the weak error associated to f.

1.2 Analysis of the weak error

In this section, we develop in our setting the weak error analysis of Talay and Tubaro [20]. For that purpose, we introduce the following definitions.

Definition 1.4. *A discretization scheme $(\hat{X}^n_{t^n_i}, 0 \leq i \leq n)$ has uniformly bounded moments if one has*

$$\exists n_0 \in \mathbb{N}^*, \forall q \in \mathbb{N}^*, \sup_{n \geq n_0, 0 \leq i \leq n} \mathbb{E}[\|\hat{X}^n_{t^n_i}\|^q] < \infty.$$

Proposition 1.5. *Let us suppose that there is $\eta > 0$ such that for $t \in (0, \eta)$,*

$$\forall q \in \mathbb{N}^*, \exists C_q > 0, \forall x \in \mathbb{D},\ \mathbb{E}[\|\hat{X}^x_t\|^q] \leq \|x\|^q(1 + C_q t) + C_q t. \tag{4}$$

Then, the discretization scheme has uniformly bounded moments.

Proof. If $n > T/\eta$, we have clearly $\mathbb{E}[\|\hat{X}^n_{t^n_{i+1}}\|^q] \leq (1 + C_q T/n)\mathbb{E}[\|\hat{X}^n_{t^n_i}\|^q] + C_q T/n$ and thus $\mathbb{E}[\|\hat{X}^n_{t^n_i}\|^q] \leq u_i$ where $u_0 = \|\hat{X}^n_{t^n_0}\|^q$ and $u_{i+1} = (1 + C_q T/n)u_i + C_q T/n$. Since $u_i = (1 + C_q T/n)^i u_0 - 1 \leq \|\hat{X}^n_{t^n_0}\|^q e^{C_q T}$, we get the desired result. □

Definition 1.6. *Let us consider a mapping $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D}) \mapsto Rf$ such that $Rf : \mathbb{R}^*_+ \times \mathbb{D} \to \mathbb{R}$. It is a remainder of order $\nu \in \mathbb{N}$ if for any function $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ with a good sequence $(C_\alpha, e_\alpha)_{\alpha \in \mathbb{N}^d}$, there exist positive constants C, E, and η depending only on $(C_\alpha, e_\alpha)_{\alpha \in \mathbb{N}^d}$ such that*

$$\forall t \in (0, \eta), \forall x \in \mathbb{D}, |Rf(t, x)| \leq C t^\nu (1 + \|x\|^E).$$

The upper bound of a remainder is thus assumed to be the same for two functions that have the same good sequence. To get upper bounds, we will say in the following with a slight abuse of language that a constant depends on a good sequence of f when this constant can be chosen only with a good sequence of f, independently from f itself. From the definition, we get the following straightforward properties.

Proposition 1.7. *Let $\nu \in \mathbb{N}$, and R_1 and R_2 be remainders of order ν. Then, $R_1 + R_2$ and μR_1 (with $\mu \in \mathbb{R}$) are remainders of order ν. If $\nu' \leq \nu$, R_1 is also a remainder of order ν'.*

Definition 1.8. *For any scheme $(\hat{p}_x(t)(dz), t > 0, x \in \mathbb{D})$ we define*

$$\forall f \in \mathcal{C}^\infty,\ R^{\hat{p}(t)}_{\nu+1} f(x) = \mathbb{E}[f(\hat{X}^x_t)] - \left[f(x) + \sum_{k=1}^{\nu} \frac{1}{k!} t^k L^k f(x) \right].$$

as soon as $\mathbb{E}[|f(\hat{X}^x_t)|] < \infty$.

We will say that $\hat{p}_x(t)(dz)$ is a potential weak νth-order scheme for the operator L if $R^{\hat{p}(t)}_{\nu+1} f(x)$ is defined for $f \in \mathcal{C}^\infty_{pol}(\mathbb{D})$ and $t > 0$, and is a remainder of order $\nu + 1$.

Thanks to the previous proposition, a potential weak νth-order scheme $\hat{X}_t^x$ for the operator L is also a potential weak ν'th-order scheme for the operator L when $\nu' \leq \nu$. In particular taking $\nu' = 0$, there are constants C , E , $\eta > 0$ that depend only on a good sequence of $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ such that

$$\forall t \in (0, \eta), |\mathbb{E}[f(\hat{X}_t^x)]| \leq C(1 + \|x\|^E). \tag{5}$$

Now, we state the following key result which is a direct consequence of the weak error analysis proposed by Talay and Tubaro [20]. Its proof is left in Appendix A.

Theorem 1.9. *Let us consider an operator L that satisfies the required assumptions on $\mathbb{D}$ and a discretization scheme $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ with transition probabilities $\hat{p}_x(t)(dz)$ on $\mathbb{D}$ that starts from $\hat{X}^n_{t_0^n} = x \in \mathbb{D}$. We assume that*

1. *the scheme has uniformly bounded moments and is a potential weak νth-order discretization scheme for the operator L.*
2. *$f : \mathbb{D} \to \mathbb{R}$ is a function such that $u(t, x) = \mathbb{E}[f(X^x_{T-t})]$ is defined on $[0, T] \times \mathbb{D}$, $\mathcal{C}^\infty$, solves $\forall t \in [0, T], \forall x \in \mathbb{D}, \partial_t u(t, x) = -Lu(t, x)$, and satisfies:*

$$\forall l \in \mathbb{N}, \alpha \in \mathbb{N}^d, \exists C_{l,\alpha}, e_{l,\alpha} > 0, \forall x \in \mathbb{D}, t \in [0, T], \ |\partial_t^l \partial_\alpha u(t, x)| \leq C_{l,\alpha}(1+\|x\|^{e_{l,\alpha}}). \tag{6}$$

Then, there is $K > 0$, $n_0 \in \mathbb{N}$, such that $|\mathbb{E}[f(\hat{X}^n_{t_n^n})] - \mathbb{E}[f(X^x_T)]| \leq K/n^\nu$ for $n \geq n_0$.

In this statement, the first assumption *1* concerns the discretization scheme and the second one mainly relies on the test function f and the diffusion coefficients b and σ. When $\mathbb{D} = \mathbb{R}^d$, $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$, b and σ are $\mathcal{C}^\infty$ with bounded derivatives, Talay [19] has shown that the second point is automatically satisfied. In that case, a potential weak νth-order scheme leads indeed to a weak error of order ν.

Let us give now on the one hand two Propositions that allow to extend easily potential weak νth-order scheme, when a coordinate is simply a function of the time and of the other coordinates. On the other hand, we check that exact schemes are indeed potential weak νth-order schemes for any ν. Their proof is left in Appendix A.

Proposition 1.10. *If $\hat{X}_t^x$ is a potential weak νth-order scheme for the operator L on $\mathbb{D}$, then $(\hat{X}_t^x, t)$ is a potential weak νth-order scheme for the operator $L + \partial_t$ on $\mathbb{D} \times \mathbb{R}_+$.*

Proposition 1.11. *Let $h \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$. We define the operator L^h for $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D} \times \mathbb{R})$ by $L^h f(x) = L\tilde{f}(x)$ where $\tilde{f}(x) = f(x, h(x))$. If $\hat{X}_t^x$ is a potential weak νth-order scheme for the operator L, then $(\hat{X}_t^x, h(\hat{X}_t^x))$ is a potential weak νth-order scheme for the operator L^h.*

Proposition 1.12. *Let $b : \mathbb{D} \to \mathbb{R}^d$ and $\sigma : \mathbb{D} \to \mathcal{M}_{d \times d_W}(\mathbb{R})$ such that $\|b(x)\| + \|\sigma(x)\| \leq C(1+\|x\|)$ for some $C > 0$, and assume that the associated operator L satisfies the required assumption on $\mathbb{D}$. Then, for any $\nu \in \mathbb{N}$, the exact scheme is a potential weak νth-order scheme for L.*

1.3 Composition of discretization schemes

In this section, we will introduce the notion of composition of discretization schemes via their transition probabilities.

Definition 1.13. *Let us consider two transition probabilities $\hat{p}^1_x(t)(dz)$ and $\hat{p}^2_x(t)(dz)$ on $\mathbb{D}$. Then, we define the composition $\hat{p}^2(t_2) \circ \hat{p}^1_x(t_1)(dz)$ simply as*

$$\hat{p}^2(t_2) \circ \hat{p}^1_x(t_1)(dz) = \int_{\mathbb{D}} \hat{p}^1_y(t_2)(dz)\hat{p}^1_x(t_1)(dy).$$

This amounts to first use the scheme 1 with a time step t_1 and then the scheme 2 with a time step t_2 with independent samples. We name $\hat{X}^{2\circ 1,x}_{t_2,t_1} = \hat{X}^{2,\hat{X}^{1,x}_{t_1}}_{t_2}$ a random variable with the law $\hat{p}^2(t_2) \circ \hat{p}^1_x(t_1)(dz)$.

More generally, if one has m transition probabilities $\hat{p}^1_x, \dots, \hat{p}^m_x$ on $\mathbb{D}$, we define $\hat{p}^m(t_m) \circ \cdots \circ \hat{p}^1_x(t_1)(dz)$ as the composition of $\hat{p}^{m-1}(t_{m-1}) \circ \cdots \circ \hat{p}^1_x(t_1)(dz)$ and then $\hat{p}^m_x(t_m)$.

Remark 1.14. *The criterion (4) that ensures the uniform boundedness of the moments is easy to use with the scheme composition. Indeed, let us fix $\lambda_1, \lambda_2 > 0$. One checks easily that if $\hat{p}^1_x(t)(dz)$ and $\hat{p}^2_x(t)(dz)$ satisfy (4), then $\hat{p}^2(\lambda_2 t) \circ \hat{p}^1_x(\lambda_1 t)(dz)$ satisfies also (4) and has thus uniformly bounded moments.*

Proposition 1.15. *Let L_1 and L_2 be two operators that satisfy the required assumptions on $\mathbb{D}$, and assume that $\hat{p}^1_x(t)(dz)$ and $\hat{p}^2_x(t)(dz)$ are respectively potential weak νth-order discretization schemes on $\mathbb{D}$ for these operators. Then, for $\lambda_1, \lambda_2 > 0$, $\hat{p}^2(\lambda_2 t) \circ \hat{p}^1_x(\lambda_1 t)(dz)$ is such that for $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$:*

$$\mathbb{E}[f(\hat{X}^{2\circ 1,x}_{\lambda_2 t,\lambda_1 t})] = \sum_{l_1+l_2 \le \nu} \frac{\lambda_1^{l_1}\lambda_2^{l_2}}{l_1! l_2!} t^{l_1+l_2} L_1^{l_1} L_2^{l_2} f(x) + R^{\hat{p}^2(\lambda_2 t)\circ\hat{p}^1(\lambda_1 t)} f(x)$$

where $R^{\hat{p}^2(\lambda_2 t)\circ\hat{p}^1(\lambda_1 t)} f(x)$ is a remainder of order $\nu + 1$.

The proof is left in Appendix A. Thanks to that result, one can think a potential scheme of order ν with a time step t as an operator $I + tL + \cdots + \frac{t^\nu}{\nu!} L^\nu + rem$ on f where rem is a remainder of order $\nu+1$. The composition of two schemes is thus simply the composition of their operators (in the reverse order) because $\sum_{l_1+l_2\le\nu} \frac{\lambda_1^{l_1}\lambda_2^{l_2}}{l_1! l_2!} t^{l_1+l_2} L_1^{l_1} L_2^{l_2} f(x) = [I + \lambda_1 t L_1 + \cdots + \frac{(\lambda_1 t)^\nu}{\nu!} L_1^\nu][I + \lambda_2 t L_2 + \cdots + \frac{(\lambda_2 t)^\nu}{\nu!} L_2^\nu] f(x)$. We deduce also the following result.

Corollary 1.16. *Let us assume that $\hat{p}^1_x(t)(dz)$ and $\hat{p}^2_x(t)(dz)$ are potential weak νth-order discretization scheme on $\mathbb{D}$ for the operators L_1 and L_2. If $L_1 L_2 = L_2 L_1$, then $\hat{p}^2(t) \circ \hat{p}^1_x(t)(dz)$ is a potential weak νth-order discretization scheme for $L_1 + L_2$.*

1.4 The Ninomiya-Victoir discretization scheme revisited

In this section, we extend in our framework the idea of the Ninomiya-Victoir scheme.

Theorem 1.17. *Let $L_1, \dots, L_m$ be m operators that satisfy the required assumption on $\mathbb{D}$. Let us consider $\hat{p}_x^1, \dots, \hat{p}_x^m$ m potential second order discretization schemes on $\mathbb{D}$ for the operators $L_1, \dots, L_m$. Then, the transition probabilities*

$$\hat{p}^m(t/2) \circ \cdots \circ \hat{p}^2(t/2) \circ \hat{p}^1(t) \circ \hat{p}^2(t/2) \circ \cdots \circ \hat{p}_x^m(t/2) \tag{7}$$

$$\frac{1}{2}\left(\hat{p}^m(t) \circ \cdots \circ \hat{p}^2(t) \circ \hat{p}_x^1(t) + \hat{p}^1(t) \circ \hat{p}^2(t) \circ \cdots \circ \hat{p}_x^m(t)\right) \tag{8}$$

are potential second order discretization schemes for the operator $\Sigma L = L_1 + L_2 + \cdots + L_m$.

Proof. Thanks to Proposition 1.15, the following expansions are justified. The first scheme gives: $(I + \frac{t}{2}L_m + \frac{t^2}{8}L_m^2 + rem) \times \cdots \times (I + \frac{t}{2}L_2 + \frac{t^2}{8}L_2^2 + rem)(I + tL_1 + \frac{t^2}{2}L_1^2 + rem)(I + \frac{t}{2}L_2 + \frac{t^2}{8}L_2^2 + rem) \times \cdots \times (I + \frac{t}{2}L_2 + \frac{t^2}{8}L_m^2 + rem) = I + t\Sigma L + \frac{t^2}{2}\Sigma L^2 + rem$ where rem denotes a remainder of order 3. In the same manner, $(I + tL_1 + \frac{t^2}{2}L_1^2 + rem) \times \cdots \times (I + tL_m + \frac{t^2}{2}L_m^2 + rem) = I + tL + \frac{t^2}{2}(\sum_{j=1}^m L_j^2 + 2\sum_{j<k} L_jL_k) + rem$ and therefore the second scheme is also a potential second order discretization scheme for ΣL. □

Let us discuss now which of the two schemes is the more efficient for computational purposes. If we suppose that each transition requires one sample, the first one requires a priori $2m-1$ samples for each step while the second one only $m+1$ (m for the schemes themselves and 1 to draw an independent Bernoulli random variable of parameter $1/2$). Since $2m-1 \geq m+1$ for $m \geq 2$, the second one is therefore a priori more efficient. There is however an exception when one of the scheme is deterministic. For example, let us assume that $\hat{p}_x^2(t)$ is a Dirac mass measure. Then, $\hat{p}^2(t/2) \circ \hat{p}^1(t) \circ \hat{p}_x^2(t/2)$ requires only one sample while the scheme $\frac{1}{2}\left(\hat{p}^2(t) \circ \hat{p}_x^1(t) + \hat{p}^1(t) \circ \hat{p}_x^2(t)\right)$ needs two samples.

Theorem 1.18. *(Ninomiya-Victoir) Let us consider the operator L defined by (3) that satisfies the required assumptions on $\mathbb{D}$. Let us assume that $\sigma(x)$ is such that the operators*

$$\begin{aligned} V_0 f(x) &= \sum_{i=1}^d b_i(x)\partial_i f(x) - \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^{d_W} \partial_j \sigma_{i,k}\sigma_{j,k}\partial_i f(x) \\ V_k f(x) &= \sum_{i=1}^d \sigma_{i,k}(x)\partial_i f \text{ for } k = 1, \dots, d_W. \end{aligned}$$

are well defined on $\mathbb{D}$ and we assume that V_0 and $\frac{1}{2}V_k^2$ (for $k = 1, \dots, d_W$) satisfy the required assumptions on the same domain $\mathbb{D}$. Then, we have

$$L = V_0 + \frac{1}{2}\sum_{k=1}^{d_W} V_k^2.$$

Defining $v_k : \mathbb{D} \to \mathbb{R}^d$ *as* $V_k f(x) =: v_k(x).\nabla f$ *for* $k = 1, \dots, d_W$, *we assume that* $\exists K > 0, \|v_k(x)\| \leq K(1 + \|x\|)$ *and that* $X_0(t,x)$ *(resp.* $X_k(t,x)$, $k = 1, \dots, d_W$*) is a* $\mathbb{D}$*-valued solution to the ODE*

$$\frac{dX_0(t,x)}{dt} = v_0(X_0(t,x)),\ t \geq 0 \quad \textit{(resp. } \frac{dX_k(t,x)}{dt} = v_k(X_k(t,x)),\ t \in \mathbb{R}\textit{)}$$

that starts from $x \in \mathbb{D}$ *at* $t = 0$. *Let denote* $\hat{p}_x^0(t)(dz)$ *(resp.* $\hat{p}_x^k(t)(dz)$*) the law of* $X_0(t,x)$ *(resp.* $X_k(\sqrt{t}N, x)$ *where* $N \sim \mathcal{N}(0,1)$*, for* $k = 1, \dots, d_W$*).*

Then, for any $\nu \in \mathbb{N}^*$, $\hat{p}_x^0(t)(dz)$ *(resp.* $\hat{p}_x^k(t)(dz)$*) is a potential* ν*th-order scheme on* $\mathbb{D}$ *for* V_0 *(resp.* $\frac{1}{2}V_k^2$*). Moreover*

$$\frac{1}{2}\left(\hat{p}^0(t/2) \circ \hat{p}^m(t) \circ \cdots \circ \hat{p}^1(t) \circ \hat{p}_x^0(t/2) + \hat{p}^0(t/2) \circ \hat{p}^1(t) \circ \cdots \circ \hat{p}^m(t) \circ \hat{p}_x^0(t/2)\right) \tag{9}$$

is well-defined and is a potential second order scheme on $\mathbb{D}$ *for* L.

This result is proven in Appendix A. A close look at the proof and especially (32) shows that $X_k(\sqrt{t}Y, x)$ is a potential νth-order scheme for $\frac{1}{2}V_k^2$ as soon as Y has uniformly bounded moments and $\mathbb{E}[Y^q] = \mathbb{E}[N^q]$ for $q \leq 2\nu + 1$. This gives the following corollary.

Corollary 1.19. *Let* Y *be a random variable with finite moments of any order such that* $\mathbb{E}[Y^q] = \mathbb{E}[N^q]$ *for* $q \leq 5$ *with* $N \sim \mathcal{N}(0,1)$. *Let us consider the framework of Theorem 1.18 but denote* $\hat{p}_x^k(t)(dz)$ *the law of* $X_k(\sqrt{t}Y, x)$ *for* $k = 1, \dots, d_W$. *Then,* $\hat{p}_x^k(t)(dz)$ *is a potential second order scheme on* $\mathbb{D}$ *for* $\frac{1}{2}V_k^2$. *Moreover,*

$$\frac{1}{2}\left(\hat{p}^0(t/2) \circ \hat{p}^m(t) \circ \cdots \circ \hat{p}^1(t) \circ \hat{p}_x^0(t/2) + \hat{p}^0(t/2) \circ \hat{p}^1(t) \circ \cdots \circ \hat{p}^m(t) \circ \hat{p}_x^0(t/2)\right) \tag{10}$$

is well-defined and is a potential second order scheme on $\mathbb{D}$ *for* L.

As an aside, we notice that when Y is chosen to be a discrete r.v. (such as in Example 2.3 later), the simulation of the scheme amounts to sample a discrete variable (a Bernoulli variable and d_W independent samples of Y) and can easily be done with only one sample of a uniform random variable on $[0,1]$.

Now, we would like to give a rather general way to split in two the operator L. Of course, a recursive application of this method allow to split L as the sum of many operators. Let us consider $I \subset \{1, \dots, d_W\}$ and denote W_t^I the $\mathbb{R}^{d_W}$-valued process such that $(W_t^I)_i = (W_t)_i$ if $i \in I$, and $(W_t^I)_i = 0$ if $i \notin I$. Let us assume that $b^I(x)$ and $b^{I^c}(x)$ are such that $b^I(x) + b^{I^c}(x) = b(x)$. Then, it is easy to see that $L = L^I + L^{I^c}$ where L^I (resp. L^{I^c}) is the operator associated to the SDE:

$$dX_t^I = b^I(X_t^I)dt + \sigma(X_t^I)dW_t^I \ (\textit{resp. } dX_t^{I^c} = b^I(X_t^{I^c})dt + \sigma(X_t^{I^c})dW_t^{I^c}).$$

The splitting of L proposed by Ninomiya and Victoir is easily obtained if one writes the SDE of $(X_t, t \geq 0)$ with the Stratonovitch integral. The operator V_0 is associated to the ODE $dX_t^\emptyset = v_0(X_t^\emptyset)dt$ and for $k = 1, \dots, d_W$, $\frac{1}{2}V_k^2$ is associated to $dX_t^{\{k\}} = \sigma(X_t^{\{k\}}) \star$

$dW_t^{\{k\}} = v_k(X_t^{\{k\}}) \star d(W_t)_k$ where $\star$ denotes the Stratonovitch integral. This splitting has the main advantage to reduce the problem to the resolution of ODEs instead of SDEs. The laws of $X_0(t,x)$ and $X_k(\sqrt{t}N,x)$ give exact schemes for their associated SDEs. If one has exact or very accurate methods to integrate the ordinary differential equations (such as Runge-Kutta method), one can get easily a weak second order scheme. Typically, the numerical integration should be accurate up to t^3 for $X_0(t,x)$ and up to t^6 for $X_k(t,x)$ to get a remainder of order 3 and thus a potential second order scheme.

2 A second order scheme for the CIR process

In this section, we focus on the discretization scheme for the CIR process (1) and have thus $d_W = 1$ and $\mathbb{D} = \mathbb{R}_+$. We introduce its operator

$$f \in \mathcal{C}^2(\mathbb{R}_+, \mathbb{R}),\ L^{\mathbf{CIR}} f(x) = (a - kx)\partial_x f(x) + \frac{1}{2}\sigma^2 x \partial_x^2 f(x)$$

that satisfies the required assumptions on $\mathbb{D}$. The main result of this section is the construction of a second order scheme for the CIR process without any restrictions on the CIR parameters $(a,k,\sigma) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+^*$. When $\sigma^2 \leq 4a$, the scheme of Ninomiya and Victoir is well defined and gives a second order scheme. For $\sigma^2 > 4a$, this scheme is no longer defined when the scheme comes near 0. Our solution consists in keeping the nonnegativity of the discretization scheme, taking different schemes whether the discretization is in a neighbourhood of 0 or not.

2.1 Ninomiya-Victoir's scheme for the CIR.

We split the operator $L^{\mathbf{CIR}}$ according to Theorem 1.18, and get $L^{\mathbf{CIR}} = V_0^{\mathbf{CIR}} + \frac{1}{2}(V_1^{\mathbf{CIR}})^2$ with

$$V_0^{\mathbf{CIR}} f(x) = (a - kx - \frac{\sigma^2}{4})f'(x) \text{ and } V_1^{\mathbf{CIR}} f(x) = \sigma\sqrt{x} f'(x).$$

In that case, we can solve explicitly the ODEs associated to $V_0^{\mathbf{CIR}}$ and $V_1^{\mathbf{CIR}}$. Defining

$$\psi_k(t) = \frac{1 - e^{-kt}}{k} = t, k \neq 0 \text{ and } \psi_0(t) = t,$$

we get for $x \geq 0$:

$$X_0^{\mathbf{CIR}}(t,x) = xe^{-kt} + (a - \sigma^2/4)\psi_k(t),\ X_1^{\mathbf{CIR}}(t,x) = ((\sqrt{x} + \frac{\sigma}{2}t)^+)^2.$$

We have $\frac{1}{2}(V_1^{\mathbf{CIR}})^2 f(x) = \frac{\sigma^2}{4} f'(x) + \frac{1}{2}\sigma^2 x f''(x)$. It is easy to see that the assumptions of Theorem 1.18 are satisfied for $\frac{1}{2}(V_1^{\mathbf{CIR}})^2$ and for $V_0^{\mathbf{CIR}}$ when $\sigma^2 \leq 4a$. When $\sigma^2 > 4a$, V_0 does no longer satisfy the required assumptions on $\mathbb{R}_+$: we do not have $\forall x \in \mathbb{R}_+, \forall t \geq 0, X_0(t,x) \in \mathbb{R}_+$. Thus, the Ninomiya-Victoir scheme is not well defined for in that case for small values of x. Last, let us remark here that Ninomiya and Victoir consider $\tilde{X}_1^{\mathbf{CIR}}(t,x) =$

$(\sqrt{x}+\frac{\sigma}{2}t)^2$ instead of $X_1^{\mathbf{CIR}}(t,x)$, which does not satisfy $\frac{d\tilde{X}_1^{\mathbf{CIR}}(t,x)}{dt}=\sigma\sqrt{\tilde{X}_1^{\mathbf{CIR}}(t,x)}$ when $\sqrt{x}+\frac{\sigma}{2}t<0$. However, a close look at the proof of Theorem 1.18 convinces that a similar expansion as (32) holds for $\mathbb{E}[f(\tilde{X}_1^{\mathbf{CIR}}(\sqrt{t}N,x))]$ ($N\sim\mathcal{N}(0,1)$), and therefore $\tilde{X}_1^{\mathbf{CIR}}(\sqrt{t}N,x)$ defines as $X_1^{\mathbf{CIR}}(\sqrt{t}N,x)$ a potential ν-th order scheme for $\frac{1}{2}(V_1^{\mathbf{CIR}})^2$. We thus get the following result.

Proposition 2.1. *When $\sigma^2\le 4a$, the Ninomiya-Victoir scheme writes for the CIR process $\hat{X}_t^x=\varphi(x,t,\sqrt{t}N)$ where $N\sim\mathcal{N}(0,1)$ and*

$$\varphi(x,t,w) \;=\; e^{-\frac{kt}{2}}\left(\sqrt{(a-\sigma^2/4)\psi_k(t/2)+e^{-\frac{kt}{2}}x}+\frac{\sigma}{2}w\right)^2+(a-\sigma^2/4)\psi_k(t/2). \quad (11)$$

It is well defined and is a potential second order scheme for $L^{\mathbf{CIR}}$.

When $\sigma^2>4a$, the idea that we use here is to consider different schemes whether we are or not in a neighborhood of 0, similarly as the QE scheme presented by Andersen [2]. Away from 0, it is natural to take the Ninomiya-Victoir scheme, provided that it is well defined and keeps nonnegativity. Unfortunately, since the standard Gaussian variable has a positive density over $\mathbb{R}$, there is still a positive probability for which the scheme leads to negative values. However, thanks to Corollary 1.19, we can replace the standard Gaussian by any random variable Y with bounded moments that matches the five first moments: this does not degrade the order of convergence of the scheme. Thus, choosing a bounded variable, we are then able to control the sign of the discretization scheme thanks to the following proposition.

Proposition 2.2. *Assume $\sigma^2>4a$ and let $A>0$. Then, $X_0^{\mathbf{CIR}}(t/2,X_1^{\mathbf{CIR}}(\sqrt{t}y,X_0^{\mathbf{CIR}}(t/2,x)))$ is well defined and nonnegative for $y\in[-A,A]$ if, and only if*

$$x\ge e^{\frac{kt}{2}}\left((\frac{\sigma^2}{4}-a)\psi_k(t/2)+\left[\sqrt{e^{\frac{kt}{2}}[(\frac{\sigma^2}{4}-a)\psi_k(t/2)]}+\frac{\sigma}{2}A\sqrt{t}\right]^2\right).$$

In that case, $X_0^{\mathbf{CIR}}(t/2,X_1^{\mathbf{CIR}}(\sqrt{t}y,X_0^{\mathbf{CIR}}(t/2,x)))=\varphi(x,t,\sqrt{t}y)$.

The proof is easy to check if one observes that $X_0^{\mathbf{CIR}}$ is increasing w.r.t x and $X_1^{\mathbf{CIR}}(\sqrt{t}y,x)$ is increasing w.r.t x and y on $\{X_1^{\mathbf{CIR}}(\sqrt{t}y,x)>0\}$. One has just then to compute the reciprocal image on 0 when $y=-A$. Let us turn to a practical example.

Example 2.3. *A suitable bounded variable that fits the five first moments of a standard Gaussian variable is Y such that $\mathbb{P}(Y=\sqrt{3})=\frac{1}{6}$, $\mathbb{P}(Y=-\sqrt{3})=\frac{1}{6}$, and $\mathbb{P}(Y=0)=2/3$. If we set*

$$\mathbf{K}_2(t)=\mathbf{1}_{\{\sigma^2>4a\}}e^{\frac{kt}{2}}\left((\frac{\sigma^2}{4}-a)\psi_k(t/2)+\left[\sqrt{e^{\frac{kt}{2}}[(\frac{\sigma^2}{4}-a)\psi_k(t/2)]}+\frac{\sigma}{2}\sqrt{3t}\right]^2\right),$$

we have, for $x\ge\mathbf{K}_2(t)$, $X_0^{\mathbf{CIR}}(t/2,X_1^{\mathbf{CIR}}(\sqrt{t}Y,X_0^{\mathbf{CIR}}(t/2,x))\ge 0$ and $\mathbf{K}_2(t)\underset{t\to 0}{\sim}[\frac{1}{2}(\frac{\sigma^2}{4}-a)+(\sqrt{\frac{1}{2}(\frac{\sigma^2}{4}-a)}+\frac{\sigma}{2}\sqrt{3})^2]t$ for $\sigma^2>4a$.

From now on, we have a threshold above which the scheme composition is well-defined and positive. Similarly as in Theorem 1.18, we are able to prove that it defines a potential second order scheme above this threshold. This is stated in the following proposition.

Proposition 2.4. *Let Y and $\mathbf{K}_2(t)$ be defined as in Example 2.3. Then, for any $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$, there are positive constants η, C and E that depend on a good sequence of f such that $\forall t \in (0,\eta), \forall x \geq \mathbf{K}_2(t)$,*

$$\left|\mathbb{E}[f(\varphi(x,t,\sqrt{t}Y))] - \left(f(x) + tL^{\mathbf{CIR}}f(x) + \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x)\right)\right| \leq Ct^3(1+|x|^E). \qquad (12)$$

Remark 2.5. *Looking for a scheme that writes $\hat{X}^x_t = \phi(x,t,\sqrt{t}N)$ with $N \sim \mathcal{N}(0,1)$ and $\phi(x,t,w) = \sum_{l+2l'\leq 4} \frac{\phi_{l,l'}(x)}{l!l'!} w^l t^{l'}$, we can get by Taylor expansions necessary conditions to obtain a second order scheme for the CIR process. In particular, a necessary condition is*

$$\frac{\sigma(a - 3kx - \sigma^2/4)}{2\sqrt{x}} = 2\phi_{1,1}(x) + \phi_{3,0}(x),$$

which implies that $\phi_{1,1}$ or $\phi_{3,0}$ explodes in the neighborhood of 0 when $\sigma^2 \neq 4a$. Due to this explosion, it is rather hard to control and get an upper bound on the remainder $|\mathbb{E}[f(\phi(x,t,\sqrt{t}N))] - \left(f(x) + tL^{\mathbf{CIR}}f(x) + \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x)\right)|$. The splitting introduced by Ninomiya and Victoir amounts to integrate the CIR when $\sigma^2 = 4a$ and is thus the right splitting to get round that difficulty.

2.2 A potential second order scheme in a neighbourhood of 0.

From now on, we turn to the simulation of the CIR near 0, namely on $[0, \mathbf{K}_2(t)]$. Near the origin, as soon as $\sigma^2 > 4a$, it does not seem possible to find even a first-order scheme that writes $\hat{X}^x_t = \varphi(x,t,\sqrt{t}Y)$ with Y matching the two first moments of a standard Gaussian variable, and that ensures nonnegativity. We therefore have to consider a different kind of scheme when the discretization approaches 0 to keep nonnegativity, as it is also done in Andersen [2]. Let us give at least two reasons for which it is really important to preserve nonnegativity. First, the CIR diffusion itself is nonnegative and it seems natural and preferable that its approximation have the same property. Second, keeping the nonnegativity ensures that the scheme is well-defined, or, more exactly, it avoids to define a discretization scheme in the negative values. Beyond the fact that it sounds strange to discretize a process where it is not defined, it is not clear how to do so and then control the weak error to get a second order scheme (see remark below).

Remark 2.6. *Defining a discretization scheme for negative values roughly amounts to extend the CIR process $(X^x_t, t \geq 0)$ for $x < 0$, and find a scheme on the domain $\mathbb{D} = \mathbb{R}$ for this extension. This approach has already been considered in the literature. For example, Deelstra and Delbaen [10] (resp. Lord and al. [16]) have chosen $dX^x_t = (a - kX^x_t)dt + \sigma\sqrt{(X^x_t)^+}dW_t$ (resp. $dX^x_t = (a - k(X^x_t)^+)dt + \sigma\sqrt{(X^x_t)^+}dW_t$) which boils down to extend*

$L^{\mathbf{CIR}}$ *by* $L^{\mathbf{CIR}}f(x) = (a-kx)f'(x)$ *(resp.* $L^{\mathbf{CIR}}f(x) = af'(x)$*) on* $x < 0$*. Keeping in mind the Talay and Tubaro method to control the weak error, it is required to have regularity assumptions on the function* $u : (t,x) \in \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{E}[f(X_t^x)]$*. Since* $\partial_t^k u = (-L^{\mathbf{CIR}})^k u$*, one should at least extend the CIR on* $\mathbb{R}_-$ *to get spatially continuous iterated operators* $(L^{\mathbf{CIR}})^k$*, which seems not obvious. For example, for the extension taken by Deelstra and Delbaen, we have* $(L^{\mathbf{CIR}})^2 f(0+) = -akf'(0) + a(a+\sigma^2/2)f''(0)$ *and* $(L^{\mathbf{CIR}})^2 f(0-) = -akf'(0) + a^2 f''(0)$*.*

To approximate the CIR near 0 and keep nonnegativity, we decide here to take a discrete random variable that matches the two first moments. Namely, we are looking for $\hat{X}_t^x$ that takes two possible values $0 \le x_-(t,x) < x_+(t,x)$ with respective probabilities $1-\pi(t,x)$ and $\pi(t,x)$ such that

$$\begin{cases} \pi(t,x)x_+(t,x) + (1-\pi(t,x))x_-(t,x) = \tilde{u}_1(t,x) \\ \pi(t,x)x_+(t,x)^2 + (1-\pi(t,x))x_-(t,x)^2 = \tilde{u}_2(t,x) \end{cases} \quad \text{where } \tilde{u}_q(t,x) = E((X_t^x)^q) \text{ for } q \in \mathbb{N}. \tag{13}$$

Some calculations give:

$$\tilde{u}_1(t,x) = xe^{-kt} + a\psi_k(t) \text{ and } \tilde{u}_2(t,x) = \tilde{u}_1(t,x)^2 + \sigma^2\psi_k(t)[a\psi_k(t)/2 + xe^{-kt}]. \tag{14}$$

Let us define $\gamma_\pm(t,x) = \frac{x_\pm(t,x)}{\tilde{u}_1(t,x)}$. The equations to solve write

$$\begin{cases} \pi(t,x)\gamma_+(t,x) + (1-\pi(t,x))\gamma_-(t,x) = 1 \\ \pi(t,x)\gamma_+(t,x)^2 + (1-\pi(t,x))\gamma_-(t,x)^2 = \frac{\tilde{u}_2(t,x)}{\tilde{u}_1(t,x)^2}. \end{cases} \tag{15}$$

We arbitrarily take $\gamma_+(t,x) = 1/(2\pi(t,x))$ and $\gamma_-(t,x) = 1/(2(1-\pi(t,x)))$ which ensures the first equation and the positivity of the random variable when $\pi(t,x) \in (0,1)$. One has thus from the last equation

$$\pi^2(t,x) - \pi(t,x) + \tilde{u}_1(t,x)^2/(4\tilde{u}_2(t,x)) = 0.$$

The discriminant is $\Delta(t,x) = 1 - \tilde{u}_1(t,x)^2/\tilde{u}_2(t,x) \in [0,1]$, and since we want $\gamma_+ > \gamma_-$, we take

$$\pi(t,x) = \frac{1-\sqrt{\Delta(t,x)}}{2}. \tag{16}$$

We have thus $0 \le \pi(t,x) \le 1/2$. Besides, we have $\tilde{u}_2(t,x)/\tilde{u}_1(t,x)^2 \le 1 + \sigma^2/a$ because $\tilde{u}_1(t,x)^2 \ge \max(a^2(\frac{1-e^{-kt}}{k})^2, 2a\frac{1-e^{-kt}}{k}xe^{-kt})$. Therefore, $\Delta(t,x) \ge 1 - 1/(1+\sigma^2/a)$ and we get $0 < \pi_{\min} = (1-\sqrt{1-1/(1+\sigma^2/a)})/2 \le \pi(t,x) \le 1/2$. Since $\mathbf{K}_2(t) \underset{t\to 0}{\sim} [\frac{1}{2}(\frac{\sigma^2}{4} - a) + (\sqrt{\frac{1}{2}(\frac{\sigma^2}{4}-a)} + \frac{\sigma}{2}\sqrt{3})^2]t$, there is a constant $C > 0$ that depends on the CIR parameters such that $\tilde{u}_1(t,x) \le Ct$ for $x \in [0, \mathbf{K}_2(t)]$ and $t \le 1$. Therefore $0 \le \hat{X}_t^x \le \frac{C}{2\pi_{\min}}t$ and

$$\forall t \in (0,1), \forall x \in [0, \mathbf{K}_2(t)], \forall q \in \mathbb{N},\ \mathbb{E}[(\hat{X}_t^x)^q] \le \left(\frac{C}{2\pi_{\min}}\right)^q t^q. \tag{17}$$

Proposition 2.7. *Let $U \sim \mathcal{U}([0,1])$. The scheme $\hat{X}_t^x = \mathbf{1}_{\{U \le \pi(t,x)\}} \frac{\tilde{u}_1(t,x)}{2\pi(t,x)} + \mathbf{1}_{\{U > \pi(t,x)\}} \frac{\tilde{u}_1(t,x)}{2(1-\pi(t,x))}$ is a potential second order scheme on $x \in [0, \mathbf{K}_2(t)]$: for any $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$, there are positive constants C and η that depend on a good sequence of f s.t.*

$$\forall t \in (0,\eta), \forall x \in [0, \mathbf{K}_2(t)], |\mathbb{E}[f(\hat{X}_t^x)] - f(x) - tL^{\mathbf{CIR}} f(x) - \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x)| \le Ct^3.$$

Proof. Let us consider a function $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$. From Proposition 1.12, the exact scheme is a potential second order scheme, i.e. it exists positive constants $C,\ E, \eta$ depending on a good sequence of $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$ s.t.

$$\forall x \ge 0, \forall t \in (0,\eta),\ |\mathbb{E}[f(X_t^x)] - f(x) - tL^{\mathbf{CIR}} f(x) - \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x)| \le Ct^3(1+x^E).$$

It is therefore sufficient to check that one has $\forall x \in [0, \mathbf{K}t], |\mathbb{E}(f(\hat{X}_t^x)) - \mathbb{E}(f(X_t^x))| \le Ct^3$ for a constant C that depends on a good sequence of f. We make a Taylor expansion of f up to order 3:

$$x \ge 0,\ f(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \int_0^x \frac{(x-y)^2}{2} f^{(3)}(y)dy.$$

Since $\hat{X}_t^x$ matches the two first moments and $|f^{(3)}(y)| \le C_3(1+|y|^q)$, we get $|\mathbb{E}(f(\hat{X}_t^x)) - \mathbb{E}(f(X_t^x))| \le C_3\mathbb{E}[(\hat{X}_t^x)^3 + (\hat{X}_t^x)^{q+3} + (X_t^x)^3 + (X_t^x)^{q+3}]$. We have shown in (17) that $\mathbb{E}[(\hat{X}_t^x)^q] \le \left(\frac{C}{2\pi_{\min}}\right)^q t^q$ for $q \in \mathbb{N}$ and $t \in (0,1)$. We have $\frac{d\tilde{u}_q(t,x)}{dt} = [aq + \frac{1}{2}q(q-1)\sigma^2]\tilde{u}_{q-1}(t,x) - kq\tilde{u}_q(t,x)$, we can prove by induction using Gronwall lemma that $\exists K_q > 0, \forall x \in [0, \mathbf{K}_2(t)], \mathbb{E}[(X_t^x)^q] \le K_q t^q$. Therefore, there is a constant $K > 0$, such that $\forall t \le 1, \mathbb{E}[(\hat{X}_t^x)^3 + (\hat{X}_t^x)^{q+3} + (X_t^x)^3 + (X_t^x)^{q+3}] \le Kt^3$. We finally get $\forall t \in (0,1), |\mathbb{E}[f(\hat{X}_t^x)] - f(x) - tL^{\mathbf{CIR}} f(x) - \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x)| \le C_3 K t^3$. Last, we observe that $C_3 K$ depends on f only through C_3 and q and thus just depends on a good sequence of f. □

2.3 The second order scheme

Theorem 2.8. *Let Y and $\mathbf{K}_2(t)$ be defined as in Example 2.3, φ the function defined in (11), $\tilde{u}_1(t,x)$ and $\pi(t,x)$ the quantities defined in (14) and (16). Let us define for $t > 0$, $\hat{p}_x(t)(dz)$ the law of $\varphi(x,t,\sqrt{t}Y)$ for $x \ge \mathbf{K}_2(t)$ and $\hat{p}_x(t)(dz) = \pi(t,x)\delta_{\frac{\tilde{u}_1(t,x)}{2\pi(t,x)}}(dz) + (1 - \pi(t,x))\delta_{\frac{\tilde{u}_1(t,x)}{2(1-\pi(t,x))}}(dz)$ for $0 \le x < \mathbf{K}_2(t)$. The scheme $(\hat{X}^n_{t_i^n}, 0 \le i \le n)$ associated to the transition probabilities $(\hat{p}_x(t)(dz), t > 0)$ and starting from $\hat{X}^n_{t_0^n} = x \in \mathbb{R}_+$ is well defined and nonnegative. It is a second order scheme:*

$$\forall f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+), \exists K > 0, \forall n \in \mathbb{N}^*,\ |\mathbb{E}[f(\hat{X}^n_{t_n^n})] - \mathbb{E}[f(X_T^x)]| \le K/n^2.$$

The main advance made here with respect to the scheme proposed by Ninomiya and Victoir is that we have a second order scheme well defined, without restriction on the parameters. To get this result, we need before the following technical result stated in [1].

Proposition 2.9. *Let us assume that* $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$. *Then,* $u(t,x) = \mathbb{E}[f(X^x_{T-t})]$ *is* $\mathcal{C}^\infty$, *solves* $\partial_t u(t,x) = -L^{\mathbf{CIR}} u(t,x)$ *on* $(t,x) \in [0,T] \times \mathbb{R}_+$ *and its derivatives satisfy*

$$\forall l, \alpha \in \mathbb{N}, \exists C_{l,\alpha}, e_{l,\alpha} > 0, \forall x \in \mathbb{R}_+, t \in [0,T],\ |\partial_t^l \partial_\alpha u(t,x)| \leq C_{l,\alpha}(1 + x^{e_{l,\alpha}}). \tag{18}$$

Proof of Theorem 2.8. The fact that the scheme is well defined is clear since the domain $\mathbb{R}_+$ is preserved by the schemes. The uniform boundedness of the moments is ensured by (17), Proposition 1.5 and Remark 1.14 since for each q, there is $C_q > 0$ s.t. $\forall x \geq 0, \max(\mathbb{E}[X_1^{\mathbf{CIR}}(\sqrt{t}Y, x)^q], X_0^{\mathbf{CIR}}(t,x)^q) \leq x^q(1 + C_q t) + C_q t$. Last, $\hat{p}_x(t)$ defines a potential scheme of order 2 thanks to Propositions 2.4 and 2.7. Point *1* of Theorem 1.9 is thus satisfied, and the second point is given by Proposition 2.9, which concludes the proof. □

`function CIR_O2` (x):
`if` $(x \geq \mathbf{K}_2(t))$ $x \leftarrow \varphi(x,t,\sqrt{t}Y)$
`else` $\pi \leftarrow \frac{1-\sqrt{1-\tilde{u}_1(t,x)^2/\tilde{u}_2(t,x)}}{2}$ `if` $(U < \pi)$ $x \leftarrow \frac{\tilde{u}_1(t,x)}{2\pi}$ `else` $x \leftarrow \frac{\tilde{u}_1(t,x)}{2(1-\pi)}$

Table 1: Algorithm computing the value at the next time-step of the 2nd-order scheme of the CIR, U (resp. Y) being sampled uniformly on $[0,1]$ (resp. as Example 2.3).

3 A third-order scheme for the CIR process

In this section, we present a third-order scheme for the CIR diffusion using the same technique as for the second order scheme. This enlightens the key ingredients to get a ν-th order scheme for the CIR. Roughly speaking, it is sufficient to have on the one hand a potential ν-order scheme for $x \geq \mathbf{K}_\nu(t)$ that preserves nonnegativity with a threshold satisfying $\mathbf{K}_\nu(t) \underset{t\to 0}{=} O(t)$ and, on the other hand, to sample under that threshold a nonnegative random variable that matches the ν first moments of the CIR process.

3.1 A third-order scheme away from a neighbourhood of 0.

Our construction will rely on the following remark. Let us assume that L_1 and L_2 are operators such that $L_1 L_2 = L_2 L_1 + L_3^2$ for some operator L_3. Let $S_i(t)$ denote the formal series $S_i(t) = I + tL_i + \frac{t^2}{2}L_i^2 + \frac{t^3}{6}L_i^3 + \dots$ where the dots represent the terms of order 4 and more. Then, we have

$$\begin{aligned}&\frac{1}{6}\sum_{\varepsilon\in\{-1,1\}} [S_2(t)S_1(t)S_3(\varepsilon t) + S_2(t)S_3(\varepsilon t)S_1(t) + S_3(\varepsilon t)S_2(t)S_1(t)] \\ &= I + t(L_1+L_2) + \frac{t^2}{2}(L_1+L_2)^2 + \frac{t^3}{6}(L_1+L_2)^3 + \dots\ .\end{aligned} \tag{19}$$

Indeed, it is easy to check the first order term. The second (resp. third) order term is $\frac{t^2}{2}(L_1^2+L_2^2+2L_2L_1+L_3^2)$ (resp. $\frac{t^3}{6}(L_1^3+L_2^3+3L_2L_1^2+3L_2^2L_1+2L_3^2L_1+2L_2L_3^2+L_1L_3^2+L_3^2L_2)$),

and since $L_3^2 = L_1L_2 - L_2L_1$, it is equal to $\frac{t^2}{2}(L_1+L_2)^2$ (resp. $\frac{t^3}{6}(L_1+L_2)^3$). Thanks to the results stated in Section 1 and especially Proposition 1.15, if one has potential third-order schemes $\hat{p}_x^i(t)$ for L_i, $i \in \{1,2,3\}$,

$$\frac{1}{6}\left(\sum_{\varepsilon\in\{-1,1\}} \hat{p}^3(\varepsilon t)\circ\hat{p}^1(t)\circ\hat{p}_x^2(t) + \hat{p}^1(t)\circ\hat{p}^3(\varepsilon t)\circ\hat{p}_x^2(t) + \hat{p}^1(t)\circ\hat{p}^2(t)\circ\hat{p}_x^3(\varepsilon t)\right) \tag{20}$$

is a potential third-order scheme for L_1+L_2. This construction requires also L_3 to be a first-order operator, because it has to be approximated for negative times when $\varepsilon=-1$.

We are going to illustrate this method on the CIR. We know from Theorem 1.18 that $X_0^{\mathbf{CIR}}(t,x)$ and $X_1^{\mathbf{CIR}}(\sqrt{t}N,x)$ are potential third-order schemes for $V_0^{\mathbf{CIR}}$ and $\frac{1}{2}(V_1^{\mathbf{CIR}})^2$. Looking at its proof, we get easily that $X_1^{\mathbf{CIR}}(\sqrt{t}Y,x)$ is a potential third-order scheme for $\frac{1}{2}(V_1^{\mathbf{CIR}})^2$ for any random variable Y with bounded moments that matches the seven first moments of $\mathcal{N}(0,1)$. Like in the second order case, we consider a random variable Y that is bounded in order to control the sign of the discretization scheme.

Example 3.1. *A suitable bounded variable that fits the seven first moments of a standard Gaussian variable is Y such that $\mathbb{P}(Y=\sqrt{3+\sqrt{6}}) = \mathbb{P}(Y=-\sqrt{3+\sqrt{6}}) = \frac{\sqrt{6}-2}{4\sqrt{6}}$, and $\mathbb{P}(Y=\sqrt{3-\sqrt{6}}) = \mathbb{P}(Y=-\sqrt{3-\sqrt{6}}) = \frac{1}{2}-\frac{\sqrt{6}-2}{4\sqrt{6}}$. This can be easily obtained thanks to Lemma 3.5 matching the moments of N^2 where $N\sim\mathcal{N}(0,1)$.*

We first focus on the particular case $k=0$, where we simply have:

$$\frac{1}{2}\left(V_0^{\mathbf{CIR}}(V_1^{\mathbf{CIR}})^2 - (V_1^{\mathbf{CIR}})^2V_0^{\mathbf{CIR}}\right) = \frac{\sigma^2}{2}\left(a-\frac{\sigma^2}{4}\right)\partial_x^2 \tag{21}$$

Let us define

$$\tilde{X}(t,x) = x + t\frac{\sigma}{\sqrt{2}}\sqrt{\left|a-\frac{\sigma^2}{4}\right|}, \tag{22}$$

the solution to the ODE associated to the operator $\tilde{L} = \frac{\sigma}{\sqrt{2}}\sqrt{|a-\frac{\sigma^2}{4}|}\partial_x$. We are then exactly in the framework described above with $L_1 = V_0^{\mathbf{CIR}}$, $L_2 = \frac{1}{2}(V_1^{\mathbf{CIR}})^2$ (resp. $L_1 = \frac{1}{2}(V_1^{\mathbf{CIR}})^2$, $L_2 = V_0^{\mathbf{CIR}}$) and $L_3 = \tilde{L}$ when $\sigma^2 \le 4a$ (resp. $\sigma^2 > 4a$). We just have to find conditions similar as those given in Proposition 2.2 that ensure that all the compositions in (20) are well defined.

Proposition 3.2. *Assume $k=0$ and let $A>0$.*

- *If $\sigma^2 \le 4a$, the compositions $\tilde{X}(\varepsilon t, X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(\sqrt{t}y,x)))$, $X_0^{\mathbf{CIR}}(t,\tilde{X}(\varepsilon t, X_1^{\mathbf{CIR}}(\sqrt{t}y,x)))$ and $X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(\sqrt{t}y,\tilde{X}(\varepsilon t,x)))$ are well defined and nonnegative for any $y\in[-A,A]$, $\varepsilon\in\{-1,1\}$ if, and only if,*

$$\begin{cases} x \ge t\frac{\sigma}{\sqrt{2}}\sqrt{a-\sigma^2/4} \text{ when } \sigma^2 \le \frac{4}{3}a \\ x \ge t\max\left(\frac{\sigma}{\sqrt{2}}\sqrt{a-\sigma^2/4}, \left(\sqrt{\frac{\sigma^2}{4}-a+\frac{\sigma}{\sqrt{2}}\sqrt{a-\frac{\sigma^2}{4}}}+\frac{\sigma}{2}A\right)^2\right) \text{ when } \frac{4}{3}a<\sigma^2<4a, \end{cases}$$

and without any restriction on $x \geq 0$ when $\sigma^2 = 4a$.

- *If $\sigma^2 > 4a$, the compositions $\tilde{X}(\varepsilon t, X_1^{\mathbf{CIR}}(\sqrt{t}y, X_0^{\mathbf{CIR}}(t,x)))$, $X_1^{\mathbf{CIR}}(\sqrt{t}y, \tilde{X}(\varepsilon t, X_0^{\mathbf{CIR}}(t,x)))$ and $X_1^{\mathbf{CIR}}(\sqrt{t}y, X_0^{\mathbf{CIR}}(t, \tilde{X}(\varepsilon t, x)))$ are well defined and nonnegative for any $y \in [-A, A]$, $\varepsilon \in \{-1, 1\}$ if, and only if,*

$$x \geq t\left[\frac{\sigma^2}{4} - a + \left(\sqrt{\frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - a}} + \frac{\sigma}{2}A\right)^2\right].$$

The proof is given in Appendix A. Let us observe that when $A \geq \sqrt{2}$ and $4a/3 < \sigma^2 < 4a$, $\frac{\sigma}{\sqrt{2}}\sqrt{a - \sigma^2/4} \leq \frac{\sigma^2}{2} \leq \left(\sqrt{\frac{\sigma^2}{4} - a + \frac{\sigma}{\sqrt{2}}\sqrt{a - \frac{\sigma^2}{4}}} + \frac{\sigma}{2}A\right)^2$.

Proposition 3.3. *Let ε and ζ be respectively independent uniform r.v. on $\{-1, 1\}$ and $\{1, 2, 3\}$, and Y be sampled independently according to Example 3.1. Let*

$$\mathbf{K}_3(t) = \psi_{-k}(t)\left[\mathbf{1}_{\{4a/3<\sigma^2<4a\}}\left(\sqrt{\frac{\sigma^2}{4} - a + \frac{\sigma}{\sqrt{2}}\sqrt{a - \frac{\sigma^2}{4}}} + \frac{\sigma}{2}\sqrt{3 + \sqrt{6}}\right)^2\right. \tag{23}$$
$$\left. + \mathbf{1}_{\{\sigma^2 \leq 4a/3\}}\frac{\sigma}{\sqrt{2}}\sqrt{a - \sigma^2/4} + \mathbf{1}_{\{4a<\sigma^2\}}\left[\frac{\sigma^2}{4} - a + \left(\sqrt{\frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - a}} + \frac{\sigma}{2}\sqrt{3 + \sqrt{6}}\right)^2\right]\right].$$

For $\sigma^2 \leq 4a$ (resp. $\sigma^2 > 4a$), the following scheme

$$\hat{X}_t^{x,k=0} = \begin{cases} \tilde{X}(\varepsilon t, X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(\sqrt{t}Y, x))) \text{ (resp. } \tilde{X}(\varepsilon t, X_1^{\mathbf{CIR}}(\sqrt{t}Y, X_0^{\mathbf{CIR}}(t,x)))) \text{ if } \zeta = 1, \\ X_0^{\mathbf{CIR}}(t, \tilde{X}(\varepsilon t, X_1^{\mathbf{CIR}}(\sqrt{t}Y, x))) \text{ (resp. } X_1^{\mathbf{CIR}}(\sqrt{t}Y, \tilde{X}(\varepsilon t, X_0^{\mathbf{CIR}}(t,x)))) \text{ if } \zeta = 2, \\ X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(\sqrt{t}Y, \tilde{X}(\varepsilon t, x))) \text{ (resp. } X_1^{\mathbf{CIR}}(\sqrt{t}Y, X_0^{\mathbf{CIR}}(t, \tilde{X}(\varepsilon t, x)))) \text{ if } \zeta = 3, \end{cases} \tag{24}$$

is well defined and nonnegative for $t \geq 0$ and $x \geq \mathbf{K}_3(t)t/\psi_{-k}(t)$. Then, for $x \geq \mathbf{K}_3(t)$, the scheme

$$\hat{X}_t^x = e^{-kt}\hat{X}_{\psi_{-k}(t)}^{x,k=0} \tag{25}$$

is a potential third-order scheme, i.e. for any $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$, there are positive constants η, C and E that depend on a good sequence of f such that $\forall t \in (0, \eta), \forall x \geq \mathbf{K}_3(t)$,

$$\left|\mathbb{E}[f(\hat{X}_t^x)] - \left(f(x) + tL^{\mathbf{CIR}}f(x) + \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x) + \frac{t^3}{6}(L^{\mathbf{CIR}})^3 f(x)\right)\right| \leq Ct^4(1 + |x|^E). \tag{26}$$

Here, for sake of clearness, we have written the scheme using three random variables ε, ζ and Y. Since these variables are discrete and independent, the scheme just requires to sample only one random variable (ε, ζ, Y) that takes 24 values.

Proof. The fact that $\hat{X}_t^{x,k=0}$ is well defined is a direct consequence of Proposition 3.2. When $k = 0$, (26) comes from (19) and Proposition 1.15. When $k \neq 0$, since $\hat{X}_t^{x,k=0}$ is a potential third-order scheme for $L_{k=0}^{\mathbf{CIR}}$ and since the multiplication by e^{-kt} is the exact scheme associated to $\Lambda = -kx\partial_x$, it is sufficient by Proposition 1.15 to check that $(I + \psi_{-k}(t)L_{k=0}^{\mathbf{CIR}} + \frac{\psi_{-k}(t)^2}{2}(L_{k=0}^{\mathbf{CIR}})^2 + \frac{\psi_{-k}(t)^3}{6}(L_{k=0}^{\mathbf{CIR}})^3 + \dots)(I + t\Lambda + \frac{t^2}{2}\Lambda^2 + \frac{t^3}{6}\Lambda^3 + \dots) = (I + tL^{\mathbf{CIR}} + \frac{t^2}{2}(L^{\mathbf{CIR}})^2 + \frac{t^3}{6}(L^{\mathbf{CIR}})^3 + \dots)$ where $L_{k=0}^{\mathbf{CIR}}$ denotes the operator associated to the CIR diffusion when $k = 0$. This can be done by some calculations using that $\psi_{-k}(t) = t + \frac{k}{2}t^2 + \frac{k^2}{6}t^3 + \dots$, $L^{\mathbf{CIR}} = L_{k=0}^{\mathbf{CIR}} + \Lambda$ and $L_{k=0}^{\mathbf{CIR}}\Lambda - \Lambda L_{k=0}^{\mathbf{CIR}} = -kL_{k=0}^{\mathbf{CIR}}$. □

Remark 3.4. *When $k \in \mathbb{R}$, one has $\frac{1}{2}(V_0^{\mathbf{CIR}}(V_1^{\mathbf{CIR}})^2 - (V_1^{\mathbf{CIR}})^2 V_0^{\mathbf{CIR}}) = \frac{\sigma^2}{2}(a - \frac{\sigma^2}{4} + kx)\partial_x^2 + \frac{\sigma^2}{4}k\partial_x = sign(a - \frac{\sigma^2}{4} + kx) \times (\frac{\sigma}{\sqrt{2}}\sqrt{|a - \frac{\sigma^2}{4} + kx|}\partial_x)^2$. Unless $a - \frac{\sigma^2}{4} + kx$ does not change sign on $x \geq 0$, it is not clear how to apply directly the method (20), mainly because the sets $\{a - \frac{\sigma^2}{4} + kx \geq 0\}$ and $\{a - \frac{\sigma^2}{4} + kx \leq 0\}$ are then no longer stable for the schemes $X_0^{\mathbf{CIR}}(t, x)$ and $X_1^{\mathbf{CIR}}(\sqrt{t}Y, x)$. To avoid that difficulty and to extend the third-order scheme when $k \neq 0$, we have used here instead the identity $(X_t^x, t \geq 0) \overset{law}{=} (e^{-kt}X_{\psi_{-k}(t)}^{x,k=0}, t \geq 0)$ between the CIR process and the CIR process with the same parameters but $k = 0$.*

3.2 A potential third-order scheme in a neighbourhood of 0.

On $x \in [0, \mathbf{K}_3(t)]$, we will approximate the CIR with a discrete random variable that matches the three first moments of the CIR. We will use the following lemma.

Lemma 3.5. *Let us consider a (non constant) random variable X such that for $i \in \{1, 2, 3\}$, $\mathbb{E}[|X|^i] < \infty$, and set $m_i = \mathbb{E}[X^i]$. Let $s = \frac{m_3 - m_1 m_2}{m_2 - m_1^2}$ and $p = \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2}$. Then, $\Delta = s^2 - 4p > 0$ and defining $x_\pm = \frac{s \pm \sqrt{\Delta}}{2}$ and $\pi = \frac{m_1 - x_-}{x_+ - x_-}$, the random variable defined by:*

$$x_+ \mathbf{1}_{\{U \leq \pi\}} + x_- \mathbf{1}_{\{U > \pi\}} \text{ with } U \sim \mathcal{U}([0, 1])$$

matches the three first moments of X. Moreover, it is nonnegative if $X \geq 0$.

Proof. We look for a random variable taking two values $x_- < x_+$ such that $\pi x_+^i + (1 - \pi)x_-^i = m_i$ for $i \in \{1, 2, 3\}$. Some calculations show that this is equivalent to the following system:

$$\pi = \frac{m_1 - x_-}{x_+ - x_-}, \; s = \frac{m_3 - m_1 m_2}{m_2 - m_1^2}, \; p = \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2},$$

where $s = x_1 + x_+$ and $p = x_- x_+$. We thus consider the polynomial function $P(x) = x^2 - sx + p$. Introducing the cumulants $\kappa_i = \mathbb{E}((X - m_1)^i)$, we check that its discriminant writes $\Delta = (4\kappa_2^3 + \kappa_3^2)/\kappa_2 > 0$. Since $P(m_1) = -\kappa_2 < 0$, we get that $m_1 \in (x_-, x_+)$ and thus $\pi \in (0, 1)$. Last, when X is nonnegative, Cauchy-Schwarz inequality gives that s and p are nonnegative and therefore $x_+ \geq x_- \geq 0$. □

Some calculations give the following formula for the third moment of the CIR:

$$\tilde{u}_3(t,x) = \tilde{u}_1(t,x)\tilde{u}_2(t,x) + \sigma^2\psi_k(t)[2x^2e^{-2kt} + \psi_k(t)(a+\frac{\sigma^2}{2})(3xe^{-kt} + a\psi_k(t))] \quad (27)$$

Let us denote from now and **till the end of Section 3** $\pi(t,x)$, $x_+(t,x)$ and $x_-(t,x)$ the parameter of the discrete random variable matching the three moments $\tilde{u}_1(t,x)$, $\tilde{u}_2(t,x)$ and $\tilde{u}_3(t,x)$ given by Lemma 3.5. Let $\hat{X}_t^x = x_+(t,x)\mathbf{1}_{\{U\leq\pi(t,x)\}} + x_-\mathbf{1}_{\{U>\pi(t,x)\}}$ with $U \sim \mathcal{U}([0,1])$. By Lemma 3.5, we obtain from (14) and (27) that for $t \geq 0$ and $0 \leq x \leq \mathbf{K}_3(t)$, $x_+(t,x)+x_-(t,x) = \frac{2x^2e^{-2kt}+\psi_k(t)(a+\frac{\sigma^2}{2})(3xe^{-kt}+a\psi_k(t))}{a\psi_k(t)/2+xe^{-kt}} \leq \frac{4e^{-2kt}}{a}\frac{\mathbf{K}_3(t)^2}{\psi_k(t)} + (2+\sigma^2/a)(3\mathbf{K}_3(t)e^{-kt} + a\psi_k(t)) \underset{t\to 0}{=} O(t)$. Thus, we get a result analogous to (17), i.e. there is a constant $C > 0$ such that

$$\forall t \in (0,1), \forall x \in [0, \mathbf{K}_3(t)], \forall q \in \mathbb{N},\ \mathbb{E}[(\hat{X}_t^x)^q] \leq C^q t^q,$$

and we can show the following result exactly like in Proposition 2.7, just doing a Taylor expansion one order further.

Proposition 3.6. *The scheme* $\hat{X}_t^x = \mathbf{1}_{\{U\leq\pi(t,x)\}}x_+(t,x) + \mathbf{1}_{\{U>\pi(t,x)\}}x_-(t,x)$ *is a potential third-order scheme on* $x \in [0, \mathbf{K}_3(t)]$*: for any* $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+)$*, there are positive constants* C *and* η *that depend on a good sequence of* f *s.t. for* $t \in (0,\eta)$ *and* $x \in [0, \mathbf{K}_3(t)]$*,*

$$|\mathbb{E}[f(\hat{X}_t^x)] - f(x) - tL^{\mathbf{CIR}}f(x) - \frac{t^2}{2}(L^{\mathbf{CIR}})^2 f(x) - \frac{t^3}{6}(L^{\mathbf{CIR}})^3 f(x)| \leq Ct^4.$$

3.3 The third-order scheme

Like for the second-order scheme, Propositions 3.3, 3.6 and 2.9 give easily, thanks to Theorem 1.9, the following result. The corresponding algorithm is given in Table 2.

Theorem 3.7. *Let* $\mathbf{K}_3(t)$ *be defined as in* (23)*,* $\hat{X}_t^x$ *the scheme defined in Proposition 3.3 (resp. Proposition 3.6) for* $x \geq \mathbf{K}_3(t)$ *(resp.* $x < \mathbf{K}_3(t)$*) and* $\hat{p}_x(t)(dz)$ *the law of* $\hat{X}_t^x$*. Then,* $\hat{p}_x(t)(dz)$ *is a potential third-order scheme for* $L^{\mathbf{CIR}}$ *on* $\mathbb{R}_+$*. Moreover, the scheme* $(\hat{X}^n_{t_i^n}, 0 \leq i \leq n)$ *associated to the transition probabilities* $(\hat{p}_x(t)(dz), t > 0)$ *and starting from* $\hat{X}^n_{t_0^n} = x \in \mathbb{R}_+$ *is a third-order scheme:*

$$\forall f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{R}_+), \exists K > 0, \forall n \in \mathbb{N}^*,\ |\mathbb{E}[f(\hat{X}^n_{t_n^n})] - \mathbb{E}[f(X_T^x)]| \leq K/n^3.$$

4 Application to Affine Term Structure Models

4.1 A second order scheme for general affine diffusions

In this section, we deal with the discretization of general affine diffusions $(X_t, t \geq 0)$. These diffusions write in their general form,

$$dX_t = (A - KX_t)dt + \Sigma\sqrt{D_t}dW_t, \quad (28)$$

`function X0`(x): $x \leftarrow x + (a - \sigma^2/4)\psi_{-k}(t)$
`function X1`(x): $x \leftarrow ((\sqrt{x} + \sigma\sqrt{\psi_{-k}(t)}Y/2)^+)^2$
`function Xt`(x): $x \leftarrow x + \frac{\sigma}{\sqrt{2}}\sqrt{|a - \sigma^2/4|}\varepsilon\psi_{-k}(t)$
`function CIR_O3`(x):
`if` $(x \geq \mathbf{K}_3(t))$ `{`
`if` $(\zeta = 1)$ `{ if` $(\sigma^2 \leq 4a)$ `{ X1`(x) `X0`(x) `Xt`(x) `} else { X0`(x) `X1`(x) `Xt`(x) `} }`
`if` $(\zeta = 2)$ `{ if` $(\sigma^2 \leq 4a)$ `{ X1`(x) `Xt`(x) `X0`(x) `} else { X0`(x) `Xt`(x) `X1`(x) `} }`
`if` $(\zeta = 3)$ `{ if` $(\sigma^2 \leq 4a)$ `{ Xt`(x) `X1`(x) `X0`(x) `} else { Xt`(x) `X0`(x) `X1`(x) `} }`
$x \leftarrow xe^{-kt}$ `}`
`else {` $s \leftarrow \frac{\tilde{u}_3(t,x)-\tilde{u}_1(t,x)\tilde{u}_2(t,x)}{\tilde{u}_2(t,x)-\tilde{u}_1(t,x)^2}$, $p \leftarrow \frac{\tilde{u}_1(t,x)\tilde{u}_3(t,x)-\tilde{u}_2(t,x)^2}{\tilde{u}_2(t,x)-\tilde{u}_1(t,x)^2}$, $\delta = \sqrt{s^2 - 4p}$, $\pi \leftarrow \frac{\tilde{u}_1-(s-\delta)/2}{\delta}$
`if` $(U < \pi)$ $x \leftarrow (s + \delta)/2$ `else` $x \leftarrow (s - \delta)/2$ `}`

Table 2: Algorithm computing the 3rd-order scheme next value, starting from x with a time-step t. Here, U is sampled uniformly on $[0, 1]$ and ε, ζ and Y as stated in Proposition 3.3.

where $A \in \mathbb{R}^d$, $K, \Sigma \in \mathbb{R}^{d\times d}$, D_t is a diagonal matrix such that $(D_t)_{ii} = \gamma_{i0} + \sum_{j=1}^d \gamma_{ij}(X_t)_j$, and $(W_t, t \geq 0)$ is a standard d-dimensional Brownian motion. We consider here the following canonical parametrization that ensures that the process is well-defined on the domain $\mathbb{D} = \mathbb{R}_+^{d'} \times \mathbb{R}^{d-d'}$:

1. $A, X_0 \in \mathbb{D}$, $\Sigma = I_d$,
2. $(K_{ij})_{1\leq i\leq d', d'+1\leq j\leq d} = 0$ and $K_{ij} \leq 0$ for $1 \leq i, j \leq d'$, $i \neq j$,
3. for $1 \leq i \leq d'$, $\gamma_{ii} \geq 0$ and $\gamma_{ij} = 0$ for $j \neq i$,
4. for $m + 1 \leq i \leq d$, $\gamma_{ij} \geq 0$ for $0 \leq j \leq d'$ and $\gamma_{ij} = 0$ for $d' + 1 \leq j \leq d$.

Then, more general admissible affine diffusions can be obtained from these canonical affine processes by affine transformations, diffusion rescaling and Brownian rotation. We refer to [9] for further details. However, for simulation purposes, it is therefore sufficient to be able to generate paths of affine processes that satisfy the four properties above. In that case, the associated operator is given by

$$
\begin{aligned}
&f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D}), Lf = L_A f + L_B f + L_C f, \text{ with} \qquad (29)\\
&L_A f = \sum_{i=1}^{d'} \left((A_i - K_{ii}x_i)\partial_i + \frac{\gamma_{ii}}{2} x_i \partial_i^2 \right), L_B f = -\sum_{i=1}^{d}\sum_{j=1}^{d} \tilde{K}_{ij} x_j \partial_i f,\\
&L_C f = \sum_{i=d'+1}^{d} \left(A_i \partial_i f + \frac{1}{2}(\gamma_{i0} + \sum_{j=1}^{d'} \gamma_{ij} x_j)\partial_i^2 f \right),
\end{aligned}
$$

where $\tilde{K}_{ij} = 0$ if $1 \leq i = j \leq d'$, and $\tilde{K}_{ij} = K_{ij}$ otherwise. We have already written the splitting that we use here to get a potential second-order scheme. First, L_A is the operator associated with d' independent CIR processes and one gets from Theorem 2.8 (or even Theorem 3.7) a second-order scheme for L_A, taking d' independent samples. We denote

$p_x^A(t)$ such a scheme. Let $p_x^B(t)$ be the Dirac mass in $\exp(-\tilde{K}t)x$: this solves exactly the ODE associated to L_B. Last, the SDE associated to L_C can be solved also exactly and for $x = (x_1, \ldots, x_d)'$, we denote $p_x^C(t)$ the law of $(x_1(t), \ldots, x_d(t))'$ with $x_i(t) = x_i$ for $i \leq d'$ and $x_i(t) = x_i + A_i t + \sqrt{\gamma_{i0} + \sum_{j=1}^{d'} \gamma_{ij} x_j} \times (W_t)_i$ for $i > d'$. We draw the attention on the fact that the domain $\mathbb{D}$ is stable for the schemes $p_x^A(t)$, $p_x^B(t)$ and $p_x^C(t)$ for any $t > 0$. We can thus compose them and from Proposition 1.12 and Theorem 1.17, we get the following result.

Proposition 4.1. *The scheme* $\frac{1}{2}(p^B(t/2) \circ p^A(t) \circ p^C(t) \circ p_x^B(t/2) + \frac{1}{2}(p^B(t/2) \circ p^C(t) \circ p^A(t) \circ p_x^B(t/2)$ *is a potential second-order scheme for the operator defined in* (29) *on* $\mathbb{D}$.

Let us add that we can prove using Remark 1.14 that this scheme has uniformly bounded moments. Therefore, it just lacks controls like (6) to get from Theorem 1.9 a second-order scheme. Since (6) holds for CIR and Vasicek processes, we may hope that it holds also for more general affine processes, but we do not tackle this technical point in this paper.

`function Affine` $(x_1, \ldots, x_d)$:
$x \leftarrow \exp(-\tilde{K}t/2)x$
`if` $(B = 1)$ `{` `for` $i = 1$ `to` d', `CIR_O2`(x_i) `//` or `CIR_O3` with parameters $(A_i, K_{ii}, \sqrt{\gamma_{ii}})$
`for` $i = d' + 1$ `to` d, $x_i \leftarrow x_i + A_i t + \sqrt{\gamma_{i0} + \sum_{j=1}^{d'} \gamma_{ij} x_j}\sqrt{t} N_i$`}`
`else {` `for` $i = d' + 1$ `to` d, $x_i \leftarrow x_i + A_i t + \sqrt{\gamma_{i0} + \sum_{j=1}^{d'} \gamma_{ij} x_j}\sqrt{t} N_i$
`for` $i = 1$ `to` d', `CIR_O2`(x_i) `} //` or `CIR_O3` with parameters $(A_i, K_{ii}, \sqrt{\gamma_{ii}})$
$x \leftarrow \exp(-\tilde{K}t/2)x$

Table 3: Algorithm for affine diffusions of Proposition 4.1, where B is a Bernoulli r.v. with parameter $1/2$ and $(N_i, d' + 1 \leq i \leq d)$ are independent standard Gaussian variables.

4.2 An efficient scheme for the Heston model

In this part, we are going to use the ideas developed in the Section 1 to the Heston model [13]. This approach has already been used by Ninomiya and Victoir [17], but the difference here is that we have at our disposal a second-order scheme for the CIR, without restriction on its parameters. Thus, we will use a different splitting of the Heston SDE that allows to use directly our CIR discretization. Before going into the details, let us mention that the Heston model (considered with a log-transformation of the stock price) belongs to the Affine Term Structure Models. We could therefore easily get a scheme from the general one given in Proposition 4.1 for affine diffusions. Due to the importance of the Heston model in finance, we prefer however to give directly the scheme that we consider in that case.

Let W and Z be two independent Brownian motions. We would like to discretize the following SDE:

$$\begin{cases} X_t^1 = X_0^1 + \int_0^t (a - kX_s^1)ds + \sigma \int_0^t \sqrt{X_s^1} dW_s \\ X_t^2 = \int_0^t X_s^1 ds \\ X_t^3 = X_0^3 + \int_0^t rX_s^3 ds + \int_0^t \sqrt{X_s^1} X_s^3 (\rho dW_s + \sqrt{1-\rho^2} dZ_s) \\ X_t^4 = \int_0^t X_s^3 ds \end{cases} \tag{30}$$

with $X_0^1 \geq 0$, $X_0^3 > 0$, $r \in \mathbb{R}$, $\rho \in [-1, 1]$ and $(a, k, \sigma) \in \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{R}_+^*$. The processes X^1 and X^3 are respectively the volatility process and the stock process, and X^2 and X^4 their respective integrals. From a financial point of view, it is common to assume moreover $r > 0$, $k > 0$ and $\rho \leq 0$, but these assumptions are not required for what follows.

First, we have to say that there is no hope that the theory developed in Section 1 works for the Heston model. Indeed, all that theory is thought to work when the discretization scheme has uniformly bounded moments. Since the discretization scheme is supposed to stick rather closely to the SDE, this roughly amounts to assume that the SDE has uniformly bounded moments, which holds when the drift $b(x)$ and the volatility function $\sigma(x)$ have a sublinear growth. In the Heston model the diffusion coefficient $\sigma(x)$ has not a sublinear growth, and it is proved indeed that the moments explode in a finite time (see Andersen and Piterbarg [3] for details). Therefore, the framework developed in this paper is not well suited to get a rigorous estimate of the weak error within the Heston model. However, it is not meaningless to apply the results stated in the Section 1 to the Heston model. The recursive construction of second-order scheme is a way to cancel many biased terms of order 1, and improve really the convergence as it will be observed in the simulation part.

We will then apply the results of Section 1 in a non rigorous manner. To do so, we split the operator of the SDE (30) $L = L^W + L^Z$, where the two operators L^W and L^Z are associated to the following respective SDEs:

$$\begin{cases} dX_t^1 = (a - kX_t^1)dt + \sigma\sqrt{X_t^1} dW_t \\ dX_t^2 = X_t^1 dt \\ dX_t^3 = (r - \frac{1}{2}(1-\rho^2)X_t^1)X_t^3 dt + \rho\sqrt{X_t^1} X_t^3 dW_t \\ dX_t^4 = X_t^3 dt \end{cases} \quad \text{and} \quad \begin{cases} dX_t^1 = 0 \\ dX_t^2 = 0 \\ dX_t^3 = \sqrt{(1-\rho^2)X_t^1} X_t^3 \star dZ_t \\ dX_t^4 = 0. \end{cases}$$

Here, $\star$ denotes the Stratonovitch integral. The second SDE is easy to integrate exactly. Concerning the first SDE, we use the second or the third order scheme described in this paper for the CIR. To discretize X_t^2, we then use the construction (7) of Theorem 1.17 with the exact scheme for $x_2\partial_1$, which amounts to use the trapezoidal rule. Then, we observe that X^3 can be integrated exactly in function of the increments of X^1 and X^2:

$$X_t^3 = X_0^3 \exp\left[(r - \frac{\rho}{\sigma}a)t + [\frac{\rho}{\sigma}k - \frac{1}{2}](X_t^2 - X_0^2) + \frac{\rho}{\sigma}(X_t^1 - X_0^1)\right],$$

and we use this formula with the increments of the discretization. Last, we discretize X^4 like X^2 using the trapezoidal scheme. Instead of writing the cumbersome formula of our

scheme, we prefer to write here directly the algorithm that computes the discretization at the next time-step (see Table 4). The function `HW` (resp. `HZ`) calculates the discretization of the SDE associated to L^W (resp. L^Z).

`function HW` (x_1, x_2, x_3, x_4):
$\Delta x_1 \leftarrow -x_1$, `CIR_O2` (x_1), $\Delta x_1 \leftarrow \Delta x_1 + x_1$ `// CIR_O3` can be used instead of `CIR_O2`
$x_2 \leftarrow x_2 + (x_1 + 0.5\Delta x_1)t$
$x_4 \leftarrow x_4 + 0.5x_3 t$
$x_3 \leftarrow x_3 \exp\left[(r - \rho a/\sigma)t + \rho \Delta x_1/\sigma + (\rho k/\sigma - 0.5)(x_1 + 0.5\Delta x_1)t\right]$
$x_4 \leftarrow x_4 + 0.5x_3 t$
$x_1 \leftarrow x_1 + \Delta x_1$
`function HZ` (x_1, x_2, x_3, x_4): $x_3 \leftarrow x_3 \exp(\sqrt{(1-\rho^2)x_1 t}N)$
`function Heston` (x_1, x_2, x_3, x_4):
`if` $(B = 1)$ `HZ`(x_1, x_2, x_3, x_4) `HW`(x_1, x_2, x_3, x_4) `else` `HW`(x_1, x_2, x_3, x_4) `HZ`(x_1, x_2, x_3, x_4)

Table 4: Algorithm for the Heston model, B being a Bernoulli sample or parameter $1/2$ and N an independent standard Gaussian variable.

5 Simulation results

5.1 Simulations for the CIR process

In this section, we want to illustrate the convergence of our second and third order schemes for the CIR presented in Sections 2 and 3. In particular, we will consider an example with parameters such that $\sigma^2 \gg 4a$, for which few existing discretization schemes are accurate as it has been mentioned in the introduction. We will consider different schemes. Schemes 1 and 2 are respectively the second and the third order schemes that we recommend. Their simulations are plotted in solid line in Figure 1. We consider also three distortions of the second-order scheme that illustrate the importance of the choice of $\mathbf{K}_2(t)$, the threshold around which we switch between the schemes given by Propositions 2.4 and 2.7. First, a look at the proof of Theorem 2.8 shows that any other threshold $\tilde{\mathbf{K}}(t)$ greater than $\mathbf{K}_2(t)$ s.t. $\tilde{\mathbf{K}}(t) \underset{t\to 0}{=} O(t)$ would have lead to another second-order scheme. Instead, if one takes a threshold smaller than $\mathbf{K}_2(t)$ forcing nonnegativity by taking positive parts, it is not clear mathematically that we get a second order scheme. We can however wonder if this is just a mathematical restriction or if it leads indeed to a worse scheme. We thus consider the following schemes:

3. second order scheme of Theorem 2.8, with switching threshold $3\mathbf{K}_2(t)/2$,
4. second order scheme of Theorem 2.8, with switching threshold $\mathbf{K}_2(t)/2$, forcing nonnegativity with positive parts.

Last, the way to obtain $\mathbf{K}_2(t)$ is closely linked with the support of Y, the moment-matching random variable that we have chosen for N. Taking a bounded random variable was

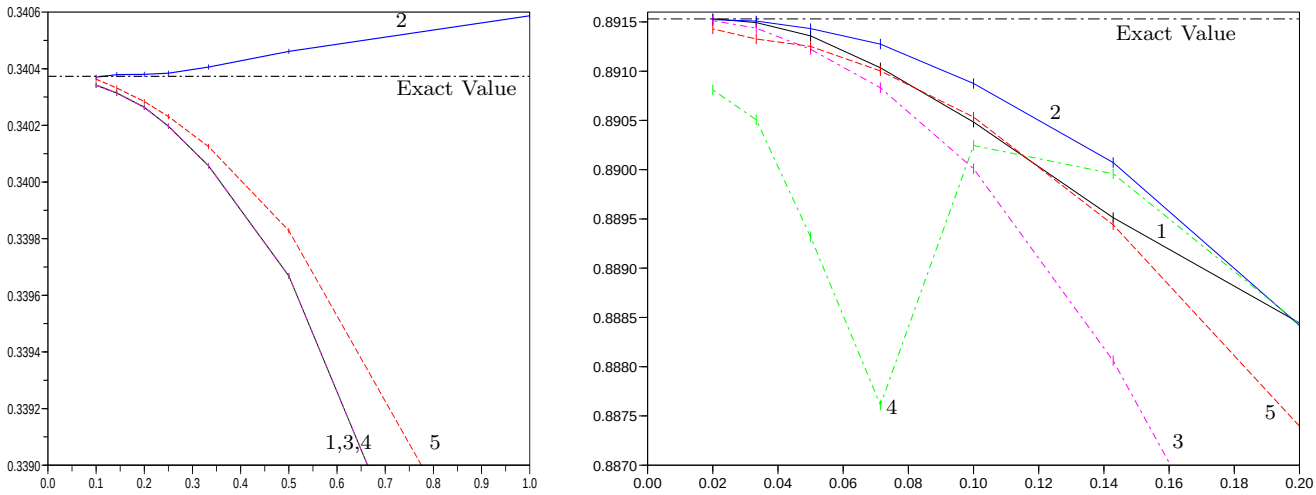

Figure 1: $\mathbb{E}(\exp(-\hat{X}^n_{t^n_n}))$ in function of $1/n$ with $x_0 = 3/2$, $k = 1/2$, $a = 1/2$ and $\sigma = 0.8$ (left) and $x_0 = 0.3$, $k = 0.1$, $a = 0.04$ and $\sigma = 2$ (right). The width of each point gives the precision up to two standard deviations.

important to prove the convergence of our scheme, but once again, we can wonder if it is of numerical importance and we consider the following scheme:

5. second order scheme of Theorem 2.8, with $N \sim \mathcal{N}(0,1)$ instead of Y, forcing nonnegativity with positive parts.

In Figure 1, we have set $T = 1$ and plotted the values of $\mathbb{E}(\exp(-\hat{X}^n_{t^n_n}))$ in function of the time step $1/n$ for two choices of parameters: $\sigma^2 < 4a$ (left) and $\sigma^2 \gg 4a$ (right). The first set of parameters is such that $\sigma^2 < 4a$, and the schemes are most of the time largely above the switching threshold, which explain that we observe no differences between the schemes 1, 3 and 4. For the same reason, the scheme 5 has also a qualitatively quadratic convergence and is even slightly better than scheme 1. Last, the third order scheme 2 converges here much better than the other schemes, giving in that case a five digit precision from $n = 5$.

The second set of parameters such that $\sigma^2 \gg 4a$ is more interesting to discuss the choice of the threshold, because the schemes are often around its value. First, we observe that the convergence of the schemes 1 and 2 is compatible with the theoretical results, and the third order scheme 2 converges more quickly to the right value than the second order scheme 1. Then, the scheme 3 converges as expected with a quadratic speed. Nonetheless with respect to scheme 1, the convergence has been slightly downgraded with the increasing of the threshold. Thus, even if theoretically any switching threshold $\tilde{\mathbf{K}}(t)$ greater than $\mathbf{K}_2(t)$ s.t. $\tilde{\mathbf{K}}(t) \underset{t\to 0}{=} O(t)$ gives a second order scheme, it seems better to take the smaller one possible as in scheme 1. The erratic behaviour of scheme 4 is sufficient to convince that

our choice of $\mathbf{K}_2(t)$ is not just a convenient choice for the proofs, but has a real impact on the convergence. Last, the convergence of scheme 5 is also worse when the time-step gets smaller than the scheme 1 and 3 for the following reason. The threshold $\mathbf{K}_2(t)$ has been calculated for a random variable Y that takes value in $[-\sqrt{3}, \sqrt{3}]$, which is of course not satisfied by a standard Gaussian variable.

To illustrate that most of the usual schemes are not accurate for large values of σ, we have also calculated the same expectations with the Full Truncation scheme proposed by Lord and al. [16]. This scheme is defined by $\hat{X}_t^x = x + (a - kx^+)t + \sigma\sqrt{x^+}W_t$. We give the values obtained apart in the following table, because they are outside the Figure 1. It

$x_0 = 3/2$, $k = 1/2$, $a = 1/2$ and $\sigma = 0.8$:

n	1	2	3	4	5	7	10
$\mathbb{E}(\exp(-\hat{X}_{t_n^n}^n))$	0.3864	0.36836	0.35924	0.35442	0.35151	0.34822	0.3458

$x_0 = 0.3$, $k = 0.1$, $a = 0.04$, and $\sigma = 2$:

n	5	7	10	14	20	30	50
$\mathbb{E}(\exp(-\hat{X}_{t_n^n}^n))$	0.80636	0.82799	0.84635	0.85974	0.8704	0.87883	0.88522

is important to notice here that for the second set of parameters, the number of samples for the Monte-Carlo method to get a precision up to four digits is about 10^8. Therefore, when $\sigma^2 \gg 4a$, the choice of the scheme is really crucial to make calculations within limited time or computational means. Of course, this holds also for the Heston model.

5.2 Simulations for the Heston model

In this section, we want to test the scheme described in Table 4 to price claims under the Heston model. More precisely, we will denote scheme 1 (resp. 2) the scheme that uses the second (resp. third) order scheme for the nested CIR. As explained in Section 4.2, we may hope at the best that these both schemes have a second order of convergence since they are constructed from the result of Theorem 1.17. Nonetheless, we would like to see numerically if there is some interest to use the third-order scheme for the CIR instead of the second-order one. Last, for comparison, we introduce the following scheme which coincides for the first and the third coordinates to the one suggested by Lord and al. [16]:

$$\hat{X}_t^x = \begin{pmatrix} x_1 + (a - kx_1^+)t + \sigma\sqrt{x_1^+}W_t \\ x_2 + x_1 t \\ x_3 \exp\left((r - x_1^+/2)t + \sqrt{x_1^+}(\rho W_t + \sqrt{1-\rho^2}Z_t)\right) \\ x_4 + x_3 t \end{pmatrix}.$$

This is the scheme 3.

In all the simulations, we have fixed $T = 1$. To test the schemes, we have calculated European put prices for different strikes with rather high values of σ in Figure 2 and Figure 3. It is hard to say qualitatively from the curves that the convergence is indeed quadratic for the schemes 1 and 2. Nonetheless in the European put case we can compare

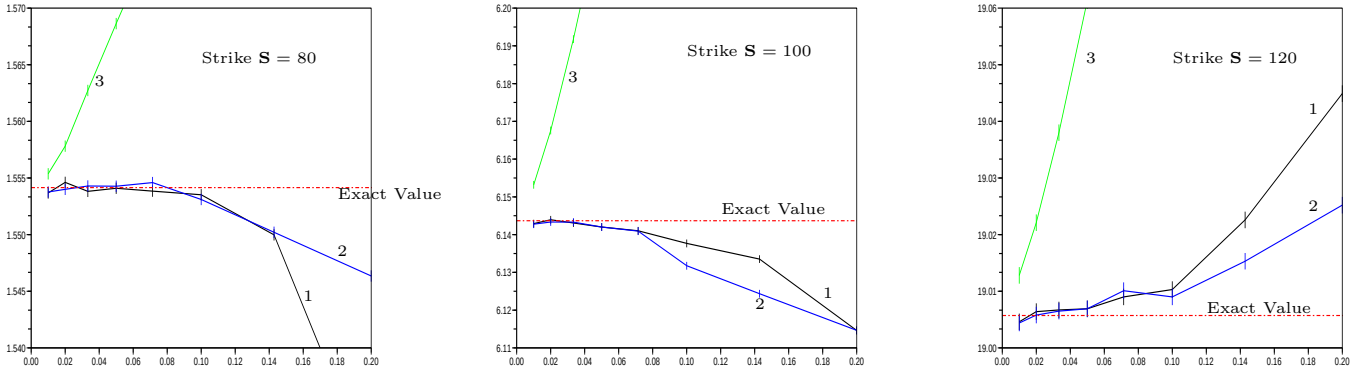

Figure 2: $\mathbb{E}[e^{-r}(\mathbf{S} - (\hat{X}^n_{t^n_n})_3)^+]$ in function of $1/n$ with $X^1_0 = 0.04$, $k = 0.5$, $a = 0.02$, $\sigma = 0.4$, $r = 0.02$, $X^3_0 = 100$ and $\rho = -0.5$. Point width gives 95% confidence interval.

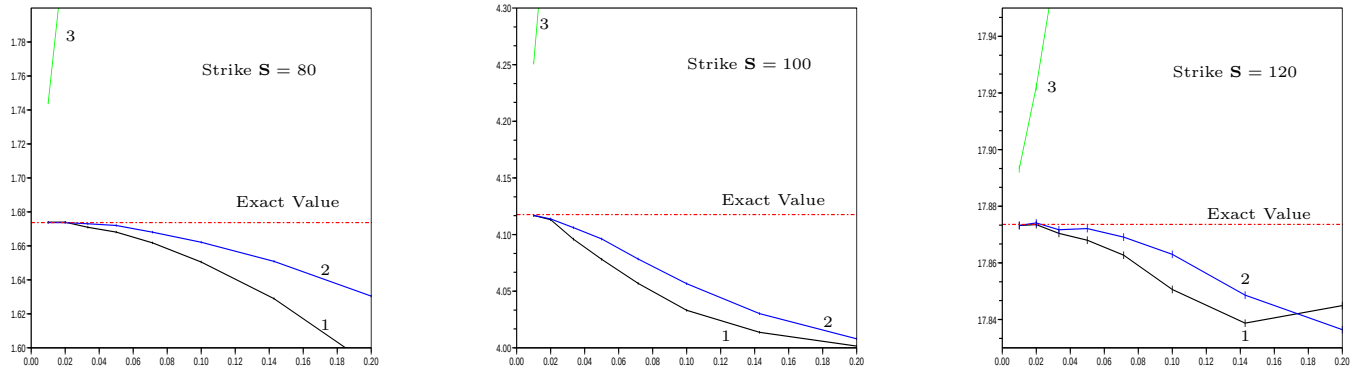

Figure 3: $\mathbb{E}[e^{-r}(\mathbf{S} - (\hat{X}^n_{t^n_n})_3)^+]$ in function of $1/n$ with $X^1_0 = 0.04$, $k = 0.5$, $a = 0.02$, $\sigma = 1$, $r = 0.02$, $X^3_0 = 100$ and $\rho = -0.8$. Point width gives 95% confidence interval.

the value obtained with the exact value. For example in Figure 2, for a time step 1/50 and for each strike, the exact value is in the two standard deviations window of which width is between 0.5×10^{-3} and 1.5×10^{-3} according to the strike value. Therefore, the bias is not much big as $(1/50)^2 = 0.4 \times 10^{-3}$ and the convergence quality is not far from being the one of a true second-order scheme. In comparison, the scheme 3 has in that case a rather linear convergence and is still far from the exact value for $n = 50$. Last, we observe that schemes 1 and 2 give similar convergence orders. In Figure 2 where σ is not that big, the difference between the schemes is not really significant. Instead, in Figure 3, when the volatility of the volatility is really high ($\sigma^2 \gg 4a$), the use of the third-order scheme for the CIR in scheme 2 allows to reduce the bias with respect to the scheme 1.

We have also plotted in Figure 4 the prices of an Asian put and of an exotic option

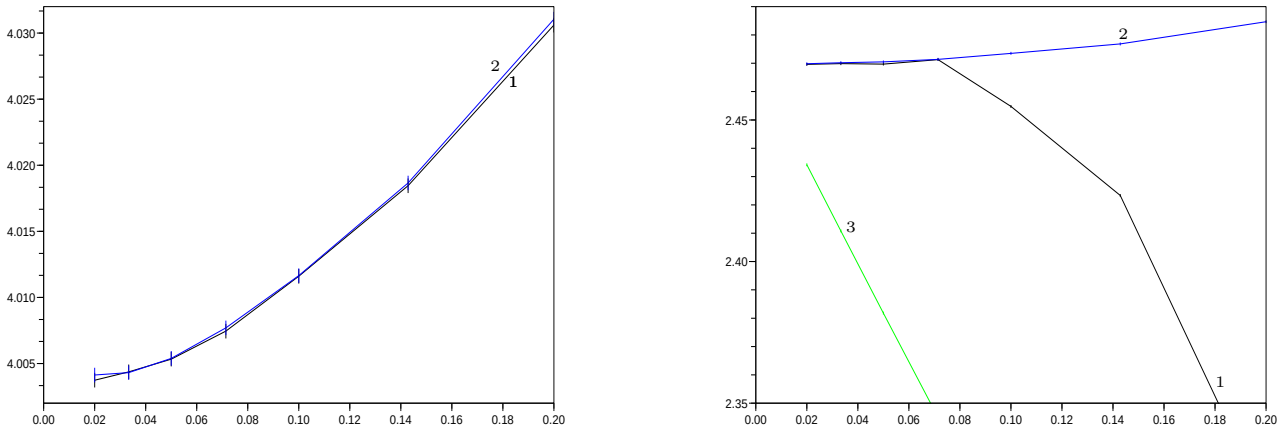

Figure 4: For the scheme 1: $\mathbb{E}[e^{-r}(100-(\hat{X}^n_{t^n_n})_4)^+]$ (left) and $\mathbb{E}[e^{-r}\mathbf{1}_{(\hat{X}^n_{t^n_n})_2>a/k}((\hat{X}^n_{t^n_n})_4-(\hat{X}^n_{t^n_n})_3)^+]$ (right) in function of $1/n$ with $X^1_0=0.04$, $k=0.5$, $a=0.02$, $\sigma=0.2$, $r=0.02$, $X^3_0=100$ and $\rho=-0.3$. Point width gives the two standard deviation precision.

that gives the right to earn the difference between the average stock and the stock when the realized variance is above a certain level. We have chosen here a rather low value of σ ($\sigma^2<4a$). Thus, the CIR process X^1 does not spend much time near 0 and the convergence observed for the schemes 1 and 2 is qualitatively parabolic in function of the time-step. For the exotic option considered here, we also notice that the scheme 2 gives minor bias than scheme 1 for large time-steps. In comparison and to underline the importance of the method chosen, we have put in Table 5 the values obtained with the scheme 3 for the Asian option, because they could not have been plotted on the same scale. For that scheme, the convergence is in that case quasi-linear.

n	5	7	10	14	20	30	50
$\mathbb{E}[e^{-r}(100-(\hat{X}^n_{t^n_n})_4)^+]$	4.6189	4.4427	4.3108	4.2235	4.1570	4.1062	4.0646

Table 5: Results for the scheme 3. Parameters as in Figure 4. Precision up to two standard deviations: 5×10^{-4}.

Conclusion and prospects

To sum up, the contribution of this paper is twofold. On the one hand, we have proposed second and third order schemes for the CIR process that work without any restriction on the parameters chosen. On the other hand, we have presented the scheme composition technique in a framework that encompasses affine diffusions. Hence, we have been able to

propose a second order scheme for these processes. We have also illustrated on examples the accuracy and the good convergence of these schemes.

Let us now hint at some possible continuations of this work. First, as it has been mentioned, the framework that we have presented here is well suited when the discretization has uniformly bounded moments, which is roughly the case when the diffusion itself has bounded moments. This holds for ATSM. Instead, we no longer have this property in the Heston model and a rigorous analysis of the weak error in that case seems to be a challenging topic. Second, the weak error has been studied in this paper for smooth functions f. It would be interesting to get the convergence for a wider set of test functions, like in the work of Bally and Talay [4] for the Euler scheme. Last, we have been able here to construct a third-order scheme for the CIR using scheme compositions. To the best of our knowledge, there is no simple recursive construction of ν-th order scheme like Theorem 1.17 for $\nu > 2$. However, we can wonder if in the special case of affine diffusions, it is possible to construct automatically schemes of order greater than 2.

A Proofs of Sections 1 and 3.

Proof of Theorem 1.9. Following Talay and Tubaro [20], we write the weak error $\mathbb{E}[f(\hat{X}^n_{t^n_n})] - \mathbb{E}[f(X^x_T)]$ as $\mathbb{E}[f(\hat{X}^n_{t^n_n})] - \mathbb{E}[f(X^x_T)] = \mathbb{E}[u(T, \hat{X}^n_{t^n_n}) - u(0, \hat{X}^n_{t^n_0})] = \sum_{i=0}^{n-1} \mathbb{E}[u(t^n_{i+1}, \hat{X}^n_{t^n_{i+1}}) - u(t_i, \hat{X}^n_{t^n_i})]$. From the Taylor expansion of u at the point $(t^n_{i+1}, \hat{X}^n_{t^n_i})$ and $\partial_t u = -Lu$ (assumption *2*), we obtain

$$\left| u(t^n_i, \hat{X}^n_{t^n_i}) - [u(t^n_{i+1}, \hat{X}^n_{t^n_i}) + \sum_{k=1}^{\nu} \frac{1}{k!} \left(\frac{T}{n}\right)^k L^k u(t^n_{i+1}, \hat{X}^n_{t^n_i})] \right| \leq \frac{(T/n)^{\nu+1}}{(\nu+1)!} C_{\nu+1,0}(1 + \|\hat{X}^n_{t^n_i}\|^{e_{\nu+1,0}}).$$

On the other hand, we deduce from (6) and assumption *1* that there are positive constants $C,\ E,\ n_0$ that depend on ν and $(C_{0,\alpha}, e_{0,\alpha})_\alpha$ such that for $n \geq n_0$,

$$u(t^n_{i+1}, \hat{X}^n_{t^n_{i+1}}) = u(t^n_{i+1}, \hat{X}^n_{t^n_i}) + \sum_{k=1}^{\nu} \frac{1}{k!} \left(\frac{T}{n}\right)^k L^k u(t^n_{i+1}, \hat{X}^n_{t^n_i}) + R^{\hat{p}(T/n)}_{\nu+1} u(t^n_{i+1}, .)(\hat{X}^n_{t^n_i})$$

with

$$|R^{\hat{p}(T/n)}_{\nu+1} u(t^n_{i+1}, .)(x)| \leq C(T/n)^{\nu+1}(1 + \|x\|^E).$$

Since the scheme has uniformly bounded moments, there is n_0 s.t. for any $q > 0$, $\kappa(q) = \sup_{n \geq n_0, 0 \leq i \leq n} \mathbb{E}[\|\hat{X}^n_{t^n_i}\|^q] < \infty$. Gathering the both previous expansions, we get $|\mathbb{E}[u(t_{i+1}, \hat{X}^n_{t^n_{i+1}}) - u(t^n_i, \hat{X}^n_{t^n_i})]| \leq \frac{K}{n^{\nu+1}}$ for $n \geq n_0$, with $K = T^{\nu+1}\left(\frac{C_{\nu+1,0}}{(\nu+1)!}(1 + \kappa(e_{\nu+1,0})) + C(1 + \kappa(E))\right)$. Thus, we deduce $|\mathbb{E}(f(X^x_T)) - \mathbb{E}(f(\hat{X}^n_{t^n_n}))| \leq K/n^\nu$. □

Proof of Proposition 1.10. Let $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D} \times \mathbb{R}_+)$. Then, there is a family $(C_\alpha, e_\alpha)_{\alpha \in \mathbb{N}^d}$ such that

$$\forall x \in \mathbb{D}, \forall t \in [0, 1), |\partial_\alpha f(x, t)| \leq C_\alpha(1 + \|x\|^{e_\alpha}),$$

and therefore there are constants $C,\ E,\ \eta > 0$, depending on $(C_\alpha, e_\alpha)_{\alpha\in\mathbb{N}^d}$ such that

$$\forall t \in (0,\eta),\ \left|\mathbb{E}[f(\hat{X}_t^x,t)] - \sum_{k=0}^{\nu}\frac{1}{k!}t^k L^k f(x,t)\right| \leq Ct^{\nu+1}(1+\|x\|^E).$$

The quantity $\mathbb{E}[f(\hat{X}_t^x,t)] - \sum_{k=0}^{\nu}\frac{1}{k!}t^k L^k f(x,t)$ is thus a remainder of order $\nu+1$. The Taylor's formula applied to $L^k f(x,t)$ up to order $\nu-k+1$ gives:

$$L^k f(x,t) = L^k f(x,0) + \cdots + \frac{t^{\nu-k}}{(\nu-k)!}\partial_t^{\nu-k}L^k f(x,0) + \int_0^t \frac{(t-s)^{\nu-k}}{(\nu-k)!}\partial_t^{\nu-k+1}L^k f(x,s)ds.$$

It is easy then to check that the integral is a remainder of order $\nu-k+1$, and therefore $\mathbb{E}[f(\hat{X}_t^x,t)] - \sum_{k=0}^{\nu}\sum_{l=0}^{\nu-k}\frac{1}{k!l!}t^{k+l}L^k\partial_t^l f(x,0) = \mathbb{E}[f(\hat{X}_t^x,t)] - \sum_{k=0}^{\nu}\frac{1}{k!}t^k(L+\partial_t)^k f(x,0)$ is a remainder of order $\nu+1$. □

Proof of Proposition 1.11. Let $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D}\times\mathbb{R})$. Then $\tilde{f}(x) \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$, and therefore we get

$$\forall t\in(0,\eta), \left|\mathbb{E}[\tilde{f}(\hat{X}_t^x)] - \left[\tilde{f}(x) + \sum_{k=1}^{\nu}\frac{1}{k!}t^k L^k\tilde{f}(x)\right]\right| \leq Ct^{\nu+1}(1+\|x\|^E),$$

for constants $C,\ E,\ \eta$ that only depend on a good sequence of $\tilde{f}$. The function $h \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ being fixed, these constants only depend also on a good sequence of f. □

Proof of Proposition 1.12. Let $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$. Thanks to the sublinear growth condition, we have bounds on the moments of X_t^x: $\forall q\in\mathbb{N}^*, \exists C_q > 0, \forall t\in[0,1], \mathbb{E}[\|X_t^x\|^q] \leq C_q(1+x^q)$. Using iterations of Itô's Formula, we get then easily for $t\in[0,1]$,

$$\mathbb{E}[f(X_t^x)] = \sum_{k=0}^{\nu}\frac{t^k}{k!}L^k f(x) + \int_0^t \frac{(t-s)^\nu}{\nu!}\mathbb{E}[L^{\nu+1}f(X_s^x)]ds.$$

Since $f\in\mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ and L satisfies the required assumptions, there are constants $C>0$ and $q\in\mathbb{N}^*$ depending only on f such that $|L^{\nu+1}f|(x) \leq C(1+\|x\|^q)$. Thus, we deduce that $|\mathbb{E}[f(X_t^x)] - \sum_{k=0}^{\nu}\frac{t^k}{k!}L^k f(x)| \leq \frac{t^{\nu+1}}{(\nu+1)!}C(1+C_q(1+\|x\|^q))$. □

Proof of Proposition 1.15. One has $\mathbb{E}[f(\hat{X}^{2\circ1,x}_{\lambda_2 t,\lambda_1 t})|\hat{X}^{1,x}_{\lambda_1 t}] = f(\hat{X}^{1,x}_{\lambda_1 t}) + \sum_{k=1}^{\nu}\frac{1}{k!}\lambda_2^k t^k L_2^k f(\hat{X}^{1,x}_{\lambda_1 t}) + R^{\hat{p}^2(\lambda_2 t)}_{\nu+1}f(\hat{X}^{1,x}_{\lambda_1 t})$ and then

$$\mathbb{E}[f(\hat{X}^{2\circ1,x}_{\lambda_2 t,\lambda_1 t})] = \sum_{l_1+l_2\leq\nu}\frac{\lambda_1^{l_1}\lambda_2^{l_2}}{l_1!l_2!}t^{l_1+l_2}L_1^{l_1}L_2^{l_2}f(x) + R^{\hat{p}^2(\lambda_2 t)\circ\hat{p}^1(\lambda_1 t)}f(x)$$

where $R^{\hat{p}^2(\lambda_2 t)\circ\hat{p}^1(\lambda_1 t)}f(x) = \mathbb{E}[R^{\hat{p}^2(\lambda_2 t)}_{\nu+1}f(\hat{X}^{1,x}_{\lambda_1 t})] + \sum_{k=0}^{\nu}\frac{1}{k!}\lambda_2^k t^k R^{\hat{p}^1(\lambda_1 t)}_{\nu+1-k}L_2^k f(x)$.
Since $R^{\hat{p}^1(\lambda_1 t)}_{\nu+1-k}L_2^k f(x)$ is a remainder of order $\nu+1-k$, it is easy to get that the sum is a remainder of order $\nu+1$ using Proposition 1.7. We have also $t\in(0,\eta_2),\ |R^{\hat{p}^2(\lambda_2 t)}_{\nu+1}f(\hat{X}^{1,x}_{\lambda_1 t})| \leq$

$C_2\lambda_2^{\nu+1}t^{\nu+1}(1+\|\hat{X}^{1,x}_{\lambda_1 t}\|^{E_2})$ for some constants $\eta_2, C_2 > 0$ and $E_2 \in 2\mathbb{N}$ that only depend on a good sequence (C_α, e_α) of f. Defining $\Phi(x) = 1 + x_1^{E_2} + \cdots + x_d^{E_2}$, we have $\Phi \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ and $|R^{\hat{p}^2(\lambda_2 t)}_{\nu+1} f(\hat{X}^{1,x}_{\lambda_1 t})| \le C_2'\lambda_2^{\nu+1}t^{\nu+1}\Phi(\hat{X}^{1,x}_{\lambda_1 t})$ and therefore we get for $\forall t \in (0, \eta_2 \wedge \eta_\Phi)$

$$|\mathbb{E}[R^{\hat{p}^2(\lambda_2 t)}_{\nu+1} f(\hat{X}^{1,x}_{\lambda_1 t})]| \le C_2'\lambda_2^{\nu+1}t^{\nu+1}\mathbb{E}[\Phi(\hat{X}^{1,x}_{\lambda_1 t})] \overset{\text{Ineq. (5)}}{\le} C_2'\lambda_2^{\nu+1}t^{\nu+1}C_\Phi(1+\|x\|^{E_\Phi})$$

for some positive constants η_Φ, C_Φ, E_Φ that only depend on Φ. Since Φ just depends on E_2, these constants depend on a good sequence of f. Therefore, $R^{\hat{p}^2(\lambda_2 t)\circ\hat{p}^1(\lambda_1 t)} f(x)$ is a remainder of order $\nu+1$. □

Proof of Theorem 1.18. We just have to check that $\hat{p}^0_x(t)(dz)$ and $\hat{p}^k_x(t)(dz)$ $(k>0)$ are respectively potential νth-order schemes for V_0 and $\frac{1}{2}V_k^2$ on $\mathbb{D}$. The result is then a straightforward consequence of Theorem 1.17.

Since there is a positive constant K such that $\|v_k(x)\| \le K(1+\|x\|)$ for $k = 0, \ldots, d_W$, the solutions to the ODEs $X_k(t,x)$ are well defined on $\mathbb{R}$, and satisfy thanks to the Gronwall lemma

$$\exists c, c' > 0, \forall t \in \mathbb{R}, k = 0, \ldots, d_W,\ \|X_k(t,x)\| \le ce^{c'|t|}(\|x\|+1).$$

Now let us consider $f \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$. Since $X_k(t,x)$ solves the ODE $dX_k(t,x)/dt = v_k(X_k(t,x))$, we get for $l \in \mathbb{N}$:

$$f(X_k(t,x)) = f(x) + tV_kf(x) + \cdots + \frac{t^l}{l!}V_k^l f(x) + \int_0^t \frac{(t-s)^l}{l!}V_k^{l+1}f(X_k(s,x))ds. \tag{31}$$

Now, let us consider the case $k=0$ and take $l=\nu$ and $t \in (0,1)$. Since V_0 satisfies the required assumption on $\mathbb{D}$, $V_0^{\nu+1}f(x) \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ and there are positive constants $C, E > 0$ that depend on a good sequence of f such that $\|V_0^{\nu+1}f(x)\| \le C(1+\|x\|^E)$. We can bound $|\int_0^t \frac{(t-s)^\nu}{\nu!}V_0^{\nu+1}f(X_0(s,x))ds| \le \frac{t^{\nu+1}}{\nu!}C(1+(ce^{c'}(\|x\|+1)^E) \le C't^{\nu+1}(1+\|x\|^E)$ for a constant $C' > 0$ that depends on a good sequence of f, and therefore $\hat{p}^0_x(t)(dz)$ is a potential νth-order scheme for V_0.

We now consider $k \in \{1, \ldots, d_W\}$ and take $l = 2\nu+1$ in (31). Since $\frac{1}{2}V_k^2$ satisfies the required assumption on $\mathbb{D}$, $V_k^{2\nu+2}f(x) \in \mathcal{C}^\infty_{\mathbf{pol}}(\mathbb{D})$ and there are positive constants $C, E > 0$ that depend on a good sequence of f such that $\|V_0^{2\nu+2}f(x)\| \le C(1+\|x\|^E)$. We get from (31) (recall $\mathbb{E}[N^{2\nu}] = \frac{(2\nu)!}{2^\nu\nu!}$):

$$\begin{aligned}\mathbb{E}[f(X_k(\sqrt{t}N,x))] &= f(x) + \frac{t}{2}V_k^2f(x) + \cdots + \frac{t^\nu}{\nu!}(\frac{1}{2}V_k^2)^\nu f(x) \\ &\quad + \mathbb{E}\left[\int_0^{\sqrt{t}N} \frac{(\sqrt{t}N-s)^{2\nu+1}}{(2\nu+1)!}V_k^{2\nu+2}f(X_k(s,x))ds\right].\end{aligned} \tag{32}$$

We have $|\int_0^{\sqrt{t}N} \frac{(\sqrt{t}N-s)^{2\nu+1}}{(2\nu+1)!}V_k^{2\nu+2}f(X_k(s,x))ds| \le \frac{t^{\nu+1}}{(2\nu+1)!}|N|^{2\nu+2}C(1+ce^{c'\sqrt{t}|N|}(\|x\|+1)^E)$ and remark that for $t \in (0,1)$, $\mathbb{E}[|N|^{2\nu+2}C(1+ce^{c'\sqrt{t}|N|}(\|x\|+1)^E)] \le C''(1+\|x\|^E)$ for a constant C'' that depends on f only through a good sequence. Therefore, $\hat{p}^k_x(t)(dz)$ is a potential νth-order scheme for $\frac{1}{2}V_k^2$. □

Proof of Proposition 3.2. We just give the main arguments here. The functions $\tilde{X}$, $X_0^{\mathbf{CIR}}$ and $X_1^{\mathbf{CIR}}$ are nondecreasing w.r.t. t and x, and it is therefore necessary and sufficient to check that these compositions are well-defined and non negative for the "worst case": $y = -A$ and $\varepsilon = -1$. We remark also that $\tilde{X}(-t, X_0^{\mathbf{CIR}}(t,x)) = X_0^{\mathbf{CIR}}(t, \tilde{X}(-t,x)) = x + \left(a - \frac{\sigma^2}{4} - \frac{\sigma}{\sqrt{2}}\sqrt{\left|a - \frac{\sigma^2}{4}\right|}\right) t$ and the term in bracket is positive when $\sigma^2 < \frac{4}{3}a$ and negative when $\sigma^2 > \frac{4}{3}a$ and $\sigma^2 \neq 4a$. When $\sigma^2 < 4a$, the condition $x \geq t\frac{\sigma}{\sqrt{2}}\sqrt{a - \sigma^2/4}$ ensures that $\tilde{X}(t,x) \geq 0$ and $X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(\sqrt{t}y, \tilde{X}(\varepsilon t, x)))$ is then well defined. When $4a/3 < \sigma^2 < 4a$, $x \geq t\left(\sqrt{\frac{\sigma^2}{4} - a + \frac{\sigma}{\sqrt{2}}\sqrt{a - \frac{\sigma^2}{4}}} + \frac{\sigma}{2}A\right)^2$ guarantees that $\tilde{X}(-t, X_0^{\mathbf{CIR}}(t, X_1^{\mathbf{CIR}}(-\sqrt{t}A, x))) \geq 0$. When $\sigma^2 > 4a$, $\tilde{X}(-t, X_1^{\mathbf{CIR}}(-\sqrt{t}A, X_0^{\mathbf{CIR}}(t,x)))$ is well defined and nonnegative if and only if $x \geq t\left[\frac{\sigma^2}{4} - a + \left(\sqrt{\frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - a}} + \frac{\sigma}{2}A\right)^2\right]$. This condition implies that $x \geq t\left[\frac{\sigma^2}{4} - a + \frac{\sigma}{\sqrt{2}}\sqrt{\frac{\sigma^2}{4} - a}\right]$. Hence, $X_1^{\mathbf{CIR}}(-\sqrt{t}A, X_0^{\mathbf{CIR}}(t, \tilde{X}(-t,x)))$ is well defined too. □

References

[1] Alfonsi, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications, Vol. 11, No. 4, pp. 355-467.*

[2] Andersen, L. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance, Vol. 11, No. 3.*

[3] Andersen, L. and Piterbarg, V. (2007). Moment explosions in stochastic volatility models. *Finance and Stochastics, Vol. 11, No. 1, pp. 29-50.*

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

[5] Berkaoui A., Bossy M. and Diop A. (2005). Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *RR-5637, INRIA, Juillet 2005.*

[6] Brigo, D., and Mercurio, F. (2006), Interest Rate Models - Theory and Practice, *2nd edition with Smile, Inflation and Credit, Springer-Verlag.*

[7] Broadie, M. and Kaya, Ö. (2003). Exact simulation of stochastic volatility and other affine jump diffusion processes, *Working Paper.*

[8] Cox, J.C. Ingersoll, J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53, pp 385-407.*

[9] Dai, Q. and Singleton, K. (2000). Specification Analysis of Affine Term Structure Models, *The Journal of Finance, Vol. LV, No. 5, pp. 1943-1978.*

[10] Deelstra, G. and Delbaen, F (1998). Convergence of Discretized Stochastic (Interest Rate) Processes with Stochastic Drift Term, *Appl. Stochastic Models Data Anal. 14, pp. 77-84.*

[11] Diop, A. (2003). Sur la discrétisation et le comportement à petit bruit d'EDS multidimensionnelles dont les coefficients sont à dérivées singulières, ph.D Thesis, INRIA. *(available at http://www.inria.fr/rrrt/tu-0785.html)*

[12] Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering *Springer, Series : Applications of Mathematics , Vol. 53.*

[13] Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies, Vol. 6, No. 2, pp. 327-343.*

[14] Karatzas, I. and Shreve, S. E. (1991). Brownian Motion and Stochastic Calculus, 2nd edition. *Springer, Series : Graduate Texts in Mathematics, Vol. 113.*

[15] Kahl, C. and Schurz, H. (2006). Balanced Milstein Methods for SDE's. *Monte Carlo Methods and Applications, Vol 12, No. 2, 2006, pp. 143-170.*

[16] Lord, R., Koekkoek, R. and D. van Dijk (2006). A comparison of biased simulation schemes for stochastic volatility models, *working paper, Erasmus University Rotterdam, Rabobank International and Robeco Alternative Investments.*

[17] Ninomiya, S. and Victoir, N. (2008). Weak approximation of stochastic differential equations and application to derivative pricing, *Applied Mathematical Finance, Vol. 15, No. 2, pp. 107-121.*

[18] Strang, G. (1968). On the construction and comparison of difference schemes, *SIAM Journal on Numerical Analysis, Vol. 5, pp. 506-517.*

[19] Talay, D. (1986). Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. *Modélisation Math. Anal. Numér., Vol. 20, No .1, pp. 141-179.*

[20] Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications, Vol. 8 No. 4 pp. 94-120.*