On the discretization schemes for the CIR (and Bessel squared) processes

Aurélien Alfonsi

CERMICS, projet MATHFI, Ecole Nationale des Ponts et Chaussées, 6-8 avenue Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne-la-vallée, France.
e-mail : alfonsi@cermics.enpc.fr
December 27, 2005

Abstract

In this paper, we focus on the simulation of the CIR processes and present several discretization schemes of both the implicit and explicit types. We study their strong and weak convergence. We also examine numerically their behaviour and compare them to the schemes already proposed by Deelstra and Delbaen [5] and Diop [6]. Finally, we gather all the results obtained and recommend, in the standard case, the use of one of our explicit schemes.

1 Introduction

The aim of this paper is to present an overview on the discretization schemes that can be used for the simulation of the square-root diffusions of Cox-Ingersoll-Ross type. These processes, initially introduced to model the short interest rate (Cox, Ingersoll and Ross [4]), are now widely used in modelling because they present interesting features like the nonnegativity and the mean reversion. Moreover, some standard expectations can be analytically calculated which can be useful especially for calibrating the parameters. Thus, they have also been used in finance to model the stochastic volatility of the stock price (Heston [9]) or the credit spread (Brigo and Alfonsi [3]). We will use in this paper the following notation for this diffusion: (X_t) will denote a Cox-Ingersoll-Ross (CIR for short) process of parameter (k, a, σ, x_0) if

$$\begin{cases} X_t = x_0 + \int_0^t (a - kX_s)ds + \sigma \int_0^t \sqrt{X_s} dW_s, t \in [0, T] \\ x_0, \sigma, a \geq 0, k \in \mathbb{R}. \end{cases} \tag{1}$$

Under the above assumption on the parameters that we will suppose valid through all the paper, it is well known that this SDE has a nonnegative solution, and this solution is pathwise unique (see for example Rogers and Williams [13]). Let us recall here that under the assumption (see for example Lamberton and Lapeyre [11])

$$2a > \sigma^2 \text{ and } x_0 > 0 \tag{2}$$

the process is always positive.

When $k > 0$, it is common to define $\theta = a/k$ and rewrite the SDE $dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$. Indeed, θ appears as the asymptotic mean of X_t toward which the process is attracted. In practice, this more intuitive parametrization is preferred.

In the sequel, $(\mathcal{F}_t, t \geq 0)$ will denote the natural filtration of the Brownian motion W, and we will consider the regular grid $t_i^n = \frac{iT}{n}$. Except in cases where it is important to remind the dependency in n, we will write t_i rather than t_i^n. It is well known that the increments of the CIR process are non-central chi-squared random variables that can be simulated exactly. Thus, we can inductively simulate a random vector distributed according to the law of $(X_{t_0}, ..., X_{t_n})$ (see Glasserman [7], pp. 120-134). However, the exact simulation in general requires more time than a simulation with approximation schemes. It may also be restrictive if one wishes to correlate this diffusion with another diffusion via the Brownian motions as in Brigo and Alfonsi [3] where two correlated CIR processes are considered. At least for both these reasons, studying approximation schemes is relevant.

It is important to remark first that the natural way to simulate this process, that is the explicit Euler-Maruyama scheme

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i})$$

with $\hat{X}^n_{t_0} = x_0$ can lead to negative values since the Gaussian increment is not bounded from below. Thus, this scheme is not well defined. To correct this problem, Deelstra and Delbaen [5] have proposed to consider:

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}\mathbf{1}_{\hat{X}^n_{t_i}>0}}(W_{t_{i+1}} - W_{t_i})$$

while Diop proposes in [6]:

$$\hat{X}^n_{t_{i+1}} = |\hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i})|.$$

However, we can as proposed in Brigo and Alfonsi [3] obtain the positivity using an implicit scheme. More precisely, if we rewrite the CIR process with the posticipated stochastic integral, we get, since $d\langle\sqrt{X},W\rangle_s=\frac{\sigma}{2}ds$:

$$\begin{aligned}
X_t &= x_0+\int_0^t(a-kX_s)ds+\sigma\int_0^t\sqrt{X_s}dW_s\\
&= x_0+\lim_{n\to\infty}\left\{\sum_{i;t_i<t}(a-kX_{t_{i+1}})\frac{T}{n}+\sigma\sum_{i;t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i})\right.\\
&\quad\left.-\sigma\sum_{i;t_i<t}(\sqrt{X_{t_{i+1}}}-\sqrt{X_{t_i}})(W_{t_{i+1}}-W_{t_i})\right\}\\
&= x_0+\lim_{n\to\infty}\left\{\sum_{i;t_i<t}(a-\frac{\sigma^2}{2}-kX_{t_{i+1}})\frac{T}{n}+\sigma\sum_{i;t_i<t}\sqrt{X_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i})\right\}.
\end{aligned}$$

It is then natural to consider the following implicit scheme that is well defined under the hypothesis (2) at least when the time step is small enough:

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i}+(a-\frac{\sigma^2}{2}-k\hat{X}^n_{t_{i+1}})\frac{T}{n}+\sigma\sqrt{\hat{X}^n_{t_{i+1}}}(W_{t_{i+1}}-W_{t_i}).$$

More precisely, when $\hat{X}^n_{t_i}\geq 0$ and $\frac{T}{n}\leq 1/k^-$ (where $y^-=\max(-y,0)$), $\sqrt{\hat{X}^n_{t_{i+1}}}$ can then be chosen as the unique positive root (since $2a>\sigma^2$, $P(0)<0$) of the second-degree polynomial $P(x)=(1+k\frac{T}{n})x^2-\sigma(W_{t_{i+1}}-W_{t_i})x-(\hat{X}^n_{t_i}+(a-\frac{\sigma^2}{2})\frac{T}{n})$, and we get

$$\hat{X}^n_{t_{i+1}}=\left(\frac{\sigma(W_{t_{i+1}}-W_{t_i})+\sqrt{\sigma^2(W_{t_{i+1}}-W_{t_i})^2+4(\hat{X}^n_{t_i}+(a-\frac{\sigma^2}{2})\frac{T}{n})(1+k\frac{T}{n})}}{2(1+k\frac{T}{n})}\right)^2. \quad (3)$$

This scheme is well defined and it is also easy to check that it preserves the monotonicity property satisfied by the CIR process: if $x_0<x_0'$ are two initial conditions, the scheme satisfies $\hat{X}^n_{t_i}<\hat{X}^{n\prime}_{t_i}$. This is an interesting example of implicit scheme on the diffusion coefficient whose general form is given by Milstein and al. (2002) since it leads to an analytical formula. In the same spirit, we can look at the SDE that drives the square-root:

$$d\sqrt{X_t}=\frac{a-\sigma^2/4}{2\sqrt{X_t}}dt-\frac{k}{2}\sqrt{X_t}dt+\frac{\sigma}{2}dW_t$$

and consider the scheme obtained by impliciting the drift. This gives also a second-degree equation in $\sqrt{\hat{X}^n_{t_{i+1}}}$:

$$\left(1+\frac{kT}{2n}\right)\hat{X}^n_{t_{i+1}}-\left[\frac{\sigma}{2}(W_{t_{i+1}}-W_{t_i})+\sqrt{\hat{X}^n_{t_i}}\right]\sqrt{\hat{X}^n_{t_{i+1}}}-\frac{a-\sigma^2/4}{2}\frac{T}{n}=0$$

that has also only one positive root when $\sigma^2 < 4a$ and $\frac{T}{n} < 2/k^-$, and it gives:

$$\hat{X}^n_{t_{i+1}} = \left(\frac{\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}} + \sqrt{(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}})^2 + 4(1+\frac{kT}{2n})\frac{a-\sigma^2/4}{2}\frac{T}{n}}}{2(1+\frac{kT}{2n})} \right)^2. \tag{4}$$

In this case $\hat{X}^n_{t_{i+1}}$ is still an increasing function of $\hat{X}^n_{t_i}$ so that the monotonicity property is satisfied. One can wonder whether we can get other schemes looking at the implicit scheme (implicit on the drift and the diffusion coefficients) with the SDE satisfied by X^α. It is not hard to see that the only two values of α that give a second-degree equation are 1 and 1/2. The other powers do not lead to analytical formulas and require a numerical resolution.

It is then interesting to make a rough Taylor expansion of order 1 of these schemes, i.e. we fix $\hat{X}^n_{t_i}$ and only conserve the terms in $\frac{T}{n}$, $(W_{t_{i+1}} - W_{t_i})$ and $(W_{t_{i+1}} - W_{t_i})^2$. We get respectively for the first scheme (3) and the second (4):

$$\hat{X}^n_{t_{i+1}} \approx \hat{X}^n_{t_i}\left(1 - k\frac{T}{n}\right) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \sigma^2/2(W_{t_{i+1}} - W_{t_i})^2 + (a - \sigma^2/2)\frac{T}{n}$$

$$\hat{X}^n_{t_{i+1}} \approx \hat{X}^n_{t_i}\left(1 - k\frac{T}{n}\right) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \sigma^2/4(W_{t_{i+1}} - W_{t_i})^2 + (a - \sigma^2/4)\frac{T}{n}$$

This indicates us a family of explicit schemes $E(\lambda)$ for $0 \leq \lambda \leq a - \sigma^2/4$ that ensure nonnegative values but not the property of monotonicity:

$$\begin{aligned} \hat{X}^n_{t_{i+1}} &= \left(\left(1 - \frac{kT}{2n}\right)\sqrt{\hat{X}^n_{t_i}} + \frac{\sigma(W_{t_{i+1}} - W_{t_i})}{2(1-\frac{kT}{2n})} \right)^2 \\ &\quad + (a - \sigma^2/4)T/n + \lambda[(W_{t_{i+1}} - W_{t_i})^2 - T/n]. \end{aligned} \tag{5}$$

It is well defined for $kT/n \neq 2$. The expansion of the scheme (3) corresponds then to $\lambda = \sigma^2/4$ while the scheme (4) to $\lambda = 0$. It is interesting here to notice that the implicit scheme on the square-root and the explicit scheme $E(0)$ have the same expansion (up to order 1) as the Milstein scheme for (1) (which can lead to negative values like the Euler scheme when $k > 0$) and for $k = 0$, $E(0)$ is exactly the Milstein scheme. Let us mention also that we could have considered as well the schemes obtained by replacing the factor $1 - \frac{kT}{2n}$ by $\sqrt{1 - kT/n}$ in (5).

This paper aims to get results on the weak and strong convergence of these schemes. Let us mention here that Deelstra and Deelbaen have proven in [5] a strong convergence result for their scheme. Diop also gets a strong convergence result in [6] but under some strong assumptions on the coefficients. She also obtains a weak convergence rate that depends on parameters. We introduce a framework in Section 2 that will allow us to study simultaneously several schemes presented above. In Section 3, we will thus establish a result of strong convergence for the schemes that satisfy an hypothesis denoted by $(\mathcal{H}_S)$.

Then we analyze the weak error in Section 4, establishing a convergence result with a $1/n$ rate for schemes satisfying an hypothesis denoted by $(\mathcal{H}_W)$. Moreover, an expansion of the weak error is given for the schemes $E(\lambda)$. Section 5 presents numerical results. We study in particular the strong convergence speed numerically and also calculate the computing time required by the several schemes. All the properties put in evidence by our analysis are listed in the conclusion, and $E(0)$ seems to be the scheme that gathers the most interesting properties.

2 Notations and preliminary lemmas

2.1 Some results on the CIR process

Lemma 2.1. *The moments of $(X_t)_{t\in[0,T]}$ are uniformly bounded by a constant that depends only on the parameters (k, a, σ, x_0), T, and the order of the moment $p \in \mathbb{N}^*$. More precisely, setting $\tilde{u}_p(t, x_0) = \mathbb{E}[X_t^p]$, there exists smooth functions $\tilde{u}_{j,p}(t)$ that depend on (k, a, σ) such that:*

$$\tilde{u}_p(t, x_0) = \sum_{j=0}^{p} \tilde{u}_{j,p}(t) x_0^j.$$

Proof: We have $\tilde{u}_0(t, x_0) = 1$ and in the case $p = 1$, $\tilde{u}_1(t, x_0) = x + \int_0^t (a - k\tilde{u}_1(s, x_0))ds$ than can be solved:

$$\tilde{u}_1(t, x_0) = x_0 e^{-kt} + a\frac{1 - e^{-kt}}{k}$$

with the convention that $\frac{1-e^{-kt}}{k} = t$ for $k = 0$. Let us consider $p \geq 2$ and assume the result true for $1 \leq j \leq p - 1$. One has $\frac{d(\tilde{u}_p(t,x_0))}{dt} = [ap + \frac{1}{2}p(p-1)\sigma^2]\tilde{u}_{p-1}(t, x_0) - kp\tilde{u}_p(t, x_0)$. Hence, we have

$$\tilde{u}_p(t, x_0) = (e^{-kt})^p \left(x_0^p + \int_0^t [ap + \frac{1}{2}p(p-1)\sigma^2](e^{ks})^p \tilde{u}_{p-1}(s, x_0)ds \right)$$

and we get the induction relations

$$\begin{aligned} \forall j \leq p-1,\ \tilde{u}_{j,p}(t) &= (e^{-kt})^p \int_0^t [ap + \frac{1}{2}p(p-1)\sigma^2](e^{ks})^p \tilde{u}_{j,p-1}(s)ds \\ \tilde{u}_{p,p}(t) &= (e^{-kt})^p. \end{aligned}$$

This gives the desired result, and we remark incidentally that $\tilde{u}_{j,p}(t)$ can be written as a polynomial of e^{-kt} or t depending on whether we are in the case $k \neq 0$ or $k = 0$. □

2.2 Introduction of the notations $\mathcal{O}(1/n^\delta)$ and $O(1/n^\delta)$

In this section, we introduce Landau type notations for sequences of random variables that will considerably simplify formulas later. To allow the multiplication of two $\mathcal{O}$, we suppose the existence of moments of any order. The results presented here are elementary, and will largely be used later.

Definition 2.2. *Let us consider a doubly indexed family of random variables* $Z = (Z^n_\gamma)_{n,\gamma}$ *with* $n \in \mathbb{N}$ *and* $\gamma \in \Gamma_n$ *a nonempty set. We will say that* Z *is of order* $\delta \in \mathbb{R}$ *- and use the notation* $Z^n_\gamma = \mathcal{O}(1/n^\delta)$ *- if there exists a family of positive random variables* $(A^n_\gamma)_{\gamma,n}$ *that have moments of any order uniformly bounded (i.e* $\forall p \in \mathbb{N}^*, \exists \kappa(A,p) > 0, \forall n \in \mathbb{N}^*, \sup_{\gamma\in\Gamma_n} \mathbb{E}[(A^n_\gamma)^p] \le \kappa(A,p)$ *) and such that:*

$$|Z^n_\gamma| \le A^n_\gamma / n^\delta$$

This is clearly equivalent to the following property:

$$\forall p \in \mathbb{N}^*, \exists \kappa(p) > 0, \forall n \in \mathbb{N}^*, \sup_{\gamma\in\Gamma_n} \mathbb{E}[(n^\delta |Z^n_\gamma|)^p] \le \kappa(p)$$

When in particular the $(Z^n_\gamma)_{\gamma,n}$ *are deterministic, this is equivalent to the boundedness of* $(n^\delta Z^n_\gamma)_{\gamma,n}$ *and we use the standard notation* $Z^n_\gamma = O(1/n^\delta)$*.*

Remarks 2.3.
1. *It is obvious but important to observe that* $Z^n_\gamma = \mathcal{O}(1/n^\delta)$ *implies that* $\mathbb{E}[Z^n_\gamma] = O(1/n^\delta)$*.*
2. *Typically we will use in the paper this definition for* $\Gamma_n = \{t_0, t_1, .., t_n\}$*.*
3. *A simple but fundamental example is* $W_{t^n_{i+1}} - W_{t^n_i} = \mathcal{O}(1/\sqrt{n})$ *which is clear since* $\sqrt{n}|W_{t^n_{i+1}} - W_{t^n_i}| \overset{law}{=} |\mathcal{N}(0,T)|$ *has moments of any order.*

Proposition 2.4. *If* $(Z^n_\gamma)_{n\in\mathbb{N},\gamma\in\Gamma_n}$ *and* $(Z'^n_{\gamma'})_{n\in\mathbb{N},\gamma'\in\Gamma'_n}$ *are two families such that* $Z^n_\gamma = \mathcal{O}(1/n^\delta)$ *and* $Z'^n_{\gamma'} = \mathcal{O}(1/n^{\delta'})$*, we have:*

$$\begin{array}{ll} 1)\ \forall c \in \mathbb{R}^*, cZ^n_\gamma = \mathcal{O}(1/n^\delta) & 2)\ \forall d \in \mathbb{R}, Z^n_\gamma/n^d = \mathcal{O}(1/n^{\delta+d}) \\ 3)\ Z^n_\gamma + Z'^n_{\gamma'} = \mathcal{O}(1/n^{\inf(\delta,\delta')}) & 4)\ \forall d > 0, (Z^n_\gamma)^d = \mathcal{O}(1/n^{d\delta}) \\ 5)\ Z^n_\gamma Z'^n_{\gamma'} = \mathcal{O}(1/n^{\delta+\delta'}) & \end{array}$$

where the families in 3) and 5) are indexed in $\Gamma_n \times \Gamma'_n$*. In particular, if we have a family of functions* $h_n : \Gamma_n \to \Gamma'_n$*, we have also:*

$$3')\ Z^n_\gamma + Z'^n_{h_n(\gamma)} = \mathcal{O}(1/n^{\inf(\delta,\delta')}) \quad 5')\ Z^n_\gamma Z'^n_{h_n(\gamma)} = \mathcal{O}(1/n^{\delta+\delta'}).$$

Proof : 1) and 2) are obvious. To prove 3), let us assume for example that $\delta \le \delta'$. Then, it is not hard to see that $Z'^n_i = \mathcal{O}(1/n^\delta)$. Since a sum of L^p random variables is L^p, we conclude easily. 4) comes immediately from the definition while 5) requires the use of Cauchy-Schwarz inequality to get the boundedness of the moments.□

By Jensen's inequality, we also easily check the following result.

Lemma 2.5. *Let us consider a family* $(\mathcal{G}_{\gamma'})_{\gamma'\in\Gamma'_n}$ *of* σ*-algebras and* $(Z^n_\gamma)_{n\in\mathbb{N},\gamma\in\Gamma_n}$ *a family of random variables such that* $Z^n_\gamma = \mathcal{O}(1/n^\delta)$*, then* $\mathbb{E}(Z^n_\gamma|\mathcal{G}_{\gamma'}) = \mathcal{O}(1/n^\delta)$*.*

2.3 On the moments of the discretization schemes

First of all, we need the following lemma to control the moments of the schemes presented here.

Lemma 2.6. *Let us suppose that $(\hat{X}^n_{t_i})$ is an nonnegative adapted scheme (i.e. $\hat{X}^n_{t_i}$ is $\mathcal{F}_{t_i}$-measurable) such that for all $n \in \mathbb{N}$,*

$$\begin{aligned} \hat{X}^n_{t_0} &= x_0 \\ \forall i \le n-1,\ \hat{X}^n_{t_{i+1}} &\le (1+b/n)\hat{X}^n_{t_i} + \sigma^n_{t_i}\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n) \end{aligned}$$

where $(\sigma^n_{t_i})$ is also supposed to be adapted with $\sigma^n_{t_i} = \mathcal{O}(1)$ and $b > 0$. Then, $(\hat{X}^n_{t_i})$ has uniformly bounded moments, that is $\hat{X}^n_{t_i} = \mathcal{O}(1)$.

Proof : Let us first remark that it is sufficient to study the case $b = 0$. Indeed, $(1+b/n)^{-i}\hat{X}^n_{t_i}$ satisfies the condition above with $b = 0$: we have for $x \in [0,n]$, $1 \le (1+b/n)^x \le e^b$ and thus on the one hand, $(1+b/n)^{-1-i/2}\sigma^n_{t_i}$ is adapted and thanks to Proposition 2.4 is a $\mathcal{O}(1)$, and on the other hand $(1+b/n)^{-i-1}\mathcal{O}(1/n) = \mathcal{O}(1/n)$. We observe then that $\hat{X}^n_{t_i} = \mathcal{O}(1) \iff (1+b/n)^{-i}\hat{X}^n_{t_i} = \mathcal{O}(1)$.

By Definition 2.2, there is $A^n_i = \mathcal{O}(1)$ such that we can rewrite the inequality (with $b = 0$) as follows :

$$\hat{X}^n_{t_{i+1}} \le \hat{X}^n_{t_i} + \sigma^n_{t_i}\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + A^n_i/n.$$

We denote in this proof $\kappa(A,p) = \sup_{i,n}\mathbb{E}[|A^n_i|^p]$. We are going to check by on p that $\forall p \in \mathbb{N}, \sup_{i,n}\mathbb{E}\left[(\hat{X}^n_{t_i})^p\right] < \infty$. It is easy to check that $\mathbb{E}[\hat{X}^n_{t_i}] \le x_0 + \kappa(A,1)$ since we have $\mathbb{E}[\hat{X}^n_{t_{i+1}}] \le \mathbb{E}[\hat{X}^n_{t_i}] + \kappa(A,1)/n$. Let us assume for any $q \le p-1$, there is a positive constant $\kappa(q)$ such that

$$\mathbb{E}[(\hat{X}^n_{t_i})^q] \le \kappa(q).$$

Since $(\hat{X}^n_{t_{i+1}})^p \le \sum_{l_1+l_2+l_3=p} \frac{p!}{l_1!l_2!l_3!}(\hat{X}^n_{t_i})^{l_1+l_2/2}\left(\sigma^n_{t_i}(W_{t_{i+1}} - W_{t_i})\right)^{l_2}(A^n_i/n)^{l_3}$, it is sufficient to control $E(l_1,l_2,l_3) = \mathbb{E}\left[(\hat{X}^n_{t_i})^{l_1+l_2/2}\left(\sigma^n_{t_i}(W_{t_{i+1}} - W_{t_i})\right)^{l_2}(A^n_i/n)^{l_3}\right]$ for $l_1 + l_2 + l_3 = p$. If $l_1 + l_2/2 \le p - 3/2$, we have necessary $l_3 + l_2/2 \ge 3/2$ and Hölder inequality gives

$$E(l_1,l_2,l_3) \le (\kappa(p-1))^{1/\alpha}\mathbb{E}\left[\left(\left(\sigma^n_{t_i}(W_{t_{i+1}} - W_{t_i})\right)^{l_2}(A^n_i/n)^{l_3}\right)^{\beta}\right]^{1/\beta} \le \frac{C(l_1,l_2,l_3)}{n^{3/2}}$$

where $\alpha = \frac{p-1}{l_1+l_2/2}$ and $1/\alpha + 1/\beta = 1$. Thus, there is a positive constant Cte such that :

$$\begin{aligned} \mathbb{E}[(\hat{X}^n_{t_{i+1}})^p] \le\ & \mathbb{E}[(\hat{X}^n_{t_i})^p] + \frac{p}{n}\mathbb{E}[(\hat{X}^n_{t_i})^{p-1}A^n_i] + \frac{p}{n}\mathbb{E}[(\hat{X}^n_{t_i})^{p-1/2}\sigma^n_{t_i}]\mathbb{E}(W_{t_{i+1}} - W_{t_i}) \\ & + \frac{p(p-1)}{2n}\mathbb{E}[(\hat{X}^n_{t_i})^{p-1}(\sigma^n_{t_i})^2] + Cte/n. \end{aligned}$$

Using once again the Hölder inequality to bound $\mathbb{E}[(\hat{X}^n_{t_i})^{p-1}A^n_i]$ from above, have for a constant $C > 0$

$$\mathbb{E}[(\hat{X}^n_{t_{i+1}})^p] \leq \mathbb{E}[(\hat{X}^n_{t_i})^p] + \frac{C}{n}(\mathbb{E}[(\hat{X}^n_{t_i})^p] + 1) = \mathbb{E}[(\hat{X}^n_{t_i})^p](1 + \frac{C}{n}) + \frac{C}{n}$$

and then we easily conclude that $\mathbb{E}[(\hat{X}^n_{t_i})^p] + 1 \leq (x_0^p + 1)e^C$.□

Now, we present a quite general framework that includes, as we will see, the implicit scheme (3) and the explicit schemes $E(\lambda)$. The hypotheses that are stated below will be useful later to get results of strong and weak convergence.

Hypothesis $(\mathcal{H}_S)$ *We will say that* $(\hat{X}^n_{t_i})$ *satisfies* $(\mathcal{H}_S)$ *if it is a nonnegative adapted scheme such that:*

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m^n_{t_{i+1}} - m^n_{t_i} + \mathcal{O}(1/n^{3/2}) \qquad (6)$$

where $m^n_{t_{i+1}} - m^n_{t_i}$ *is a martingale increment (i.e.* $\mathbb{E}[m^n_{t_{i+1}} - m^n_{t_i}|\mathcal{F}_{t_i}] = 0$*) of order 1:*

$$m^n_{t_{i+1}} - m^n_{t_i} = \mathcal{O}(1/n). \qquad (7)$$

If it is satisfied, we get immediately that $\hat{X}^n_{t_{i+1}} \leq \hat{X}^n_{t_i} + \frac{|k|T}{n}\hat{X}^n_{t_i} + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$ using that $|a - k\hat{X}^n_{t_i}| \leq a + |k|\hat{X}^n_{t_i}$. Therefore, we can apply the Lemma 2.6 and deduce that $\hat{X}^n_{t_i} = \mathcal{O}(1)$. We define in that case the discrete martingale $(M^n_{t_i})$ by

$$\begin{cases} M^n_{t_0} = 0 \\ M^n_{t_{i+1}} - M^n_{t_i} = \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m^n_{t_{i+1}} - m^n_{t_i}. \end{cases} \qquad (8)$$

Thanks to Proposition 2.4 and Remark 2.3, we get

Corollary 2.7. *Under hypothesis* $(\mathcal{H}_S)$*,* $\hat{X}^n_{t_i}$ *has uniformly bounded moments, and we have:*

$$\begin{aligned} (M^n_{t_{i+1}} - M^n_{t_i})^2 &= \sigma^2\hat{X}^n_{t_i}(W_{t_{i+1}} - W_{t_i})^2 + \mathcal{O}(1/n^{3/2}) \\ \hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} &= \mathcal{O}(1/\sqrt{n}). \end{aligned}$$

However, as we will see when studying the weak error, it can be useful to make a stronger assumption to get a faster convergence.

Hypothesis $(\mathcal{H}_W)$ *We say that a scheme* $(\hat{X}^n_{t_i})$ *satisfies* $(\mathcal{H}_W)$ *if it already satisfies* $(\mathcal{H}_S)$ *and moreover*

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m^n_{t_{i+1}} - m^n_{t_i} + \mathcal{O}(1/n^2) \qquad (9)$$

$$\mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^2|\mathcal{F}_{t_i}\right] = \sigma^2\hat{X}^n_{t_i}T/n + \mathcal{O}(1/n^2). \qquad (10)$$

The absence of term of order 3/2 in (10) and the knowledge of the expansion of the scheme (9) up to order 2 play a key role to get a weak error at most proportional to the time step.

Remark 2.8. *Let us suppose that there is a function $\psi^n(x,w)$ which is even with respect to its second argument w such that:*

$$m^n_{t_{i+1}} - m^n_{t_i} = \psi^n(\hat{X}^n_{t_i}, (W_{t_{i+1}} - W_{t_i})) + \mathcal{O}(1/n^{3/2}).$$

Then, $(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^2 = \sigma^2 \hat{X}^n_{t_i}(W_{t_{i+1}} - W_{t_i})^2 + 2\sigma\sqrt{\hat{X}^n_{t_i}}\psi^n(\hat{X}^n_{t_i}, (W_{t_{i+1}} - W_{t_i}))(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n^2)$*, and therefore condition (10) is automatically satified thanks to Lemma 2.5.*

2.4 Study of the expansion of the different schemes

In this section we examine each scheme presented in the introduction and our aim is to discuss whether it satisfies or not Hypotheses $(\mathcal{H}_S)$ and $(\mathcal{H}_W)$ defined before.

2.4.1 Expansion of the implicit scheme (3)

We assume here that $2a > \sigma^2$, and expand the relation that defines the implicit scheme (3):

$$\begin{aligned}\hat{X}^n_{t_{i+1}} &= \frac{1}{4(1+kT/n)^2}\Bigg(2\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}^n_{t_i} + (a - \frac{\sigma^2}{2})T/n)(1+kT/n) \\ &\quad + 2\sigma(W_{t_{i+1}} - W_{t_i})\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}^n_{t_i} + (a - \frac{\sigma^2}{2})T/n)(1+kT/n)}\Bigg) \quad (11)\end{aligned}$$

Let us now observe that

$$\begin{aligned}&\left|\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(\hat{X}^n_{t_i} + (a - \frac{\sigma^2}{2})T/n)(1+kT/n)} - 2\sqrt{\hat{X}^n_{t_i}(1+kT/n)}\right| \\ &\leq \sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4(a - \frac{\sigma^2}{2})(1+kT/n)T/n} = \mathcal{O}(1/\sqrt{n}), \quad (12)\end{aligned}$$

using Proposition 2.4. Thus, we have

$$\hat{X}^n_{t_{i+1}} = \tfrac{1}{1+kT/n}\hat{X}^n_{t_i} + \tfrac{1}{(1+kT/n)^{3/2}}\sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$$

which gives that $\hat{X}^n_{t_i} = \mathcal{O}(1)$ using Lemma 2.6. Once we know this, we can continue the expansion thanks to Proposition 2.4 and it is not hard to get:

$$\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} = \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \frac{\sigma^2}{2}[(W_{t_{i+1}} - W_{t_i})^2 - T/n] + \dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i} + \mathcal{O}(1/n^2) \quad (13)$$

where $\dot{M}^n_{t_i}$ is a discrete $\mathcal{F}_{t_i}$-martingale defined by $\dot{M}^n_{t_0} = 0$ and

$$\dot{M}^n_{t_{i+1}} = \dot{M}^n_{t_i} + \frac{\sigma(W_{t_{i+1}} - W_{t_i})}{2(1+kT/n)^2}\sqrt{\sigma^2(W_{t_{i+1}} - W_{t_i})^2 + 4\left(\hat{X}^n_{t_i} + (a - \frac{\sigma^2}{2})\frac{T}{n}\right)(1+kT/n)}.$$

Indeed, we have $\mathbb{E}(\dot{M}^n_{t_{i+1}}|\mathcal{F}_{t_i}) =$

$$\dot{M}^n_{t_i}+\frac{\sigma\sqrt{T}}{2\sqrt{2\pi n}(1+kT/n)^2}\underbrace{\int_{-\infty}^{\infty} xe^{-\frac{x^2}{2}}\sqrt{\frac{\sigma^2 T}{n}x^2+4(\hat{X}^n_{t_i}+(a-\frac{\sigma^2}{2})T/n)(1+kT/n)}dx}_{0} = \dot{M}^n_{t_i}.$$

Moreover, we have $(\dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i})^2 = \sigma^2(W_{t_{i+1}} - W_{t_i})^2\hat{X}^n_{t_i} + \mathcal{O}(1/n^2)$ and in particular $\dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i} = \mathcal{O}(1/\sqrt{n})$. Now, we can define the martingale $(m^n_{t_i})$ by

$$m^n_{t_{i+1}} - m^n_{t_i} = \frac{\sigma^2}{2}[(W_{t_{i+1}} - W_{t_i})^2 - T/n] + \dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i} - \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i})$$

and it is easy from (13) to see that the properties (6) and (9) are satisfied. Inequality (12) gives us that $\dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i} - \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) = \mathcal{O}(1/n)$ and therefore property (7) is satisfied by m^n since $\frac{\sigma^2}{2}[(W_{t_{i+1}} - W_{t_i})^2 - T/n] = \mathcal{O}(1/n)$. We have first shown thus that $(\mathcal{H}_S)$ is satisfied. Now, using the Proposition 2.4, we get that:

$$\begin{aligned}(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^2 &= \sigma^2(W_{t_{i+1}} - W_{t_i})^2\hat{X}^n_{t_i} \\ &+[\sigma^2((W_{t_{i+1}} - W_{t_i})^2 - T/n) + 2(a-k\hat{X}^n_{t_i})T/n](\dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i}) + \mathcal{O}(1/n^2)\end{aligned}$$

and that the term of order 3/2, $[\sigma^2((W_{t_{i+1}} - W_{t_i})^2 - T/n) + 2(a-k\hat{X}^n_{t_i})T/n](\dot{M}^n_{t_{i+1}} - \dot{M}^n_{t_i})$, has a null conditional expectation respect to $\mathcal{F}_{t_i}$ since it can be written as an odd function respect to the Brownian increment. This shows that we have (10) and $(\mathcal{H}_W)$ is also satisfied by this implicit scheme.

2.4.2 Expansion of the implicit scheme (4)

Let us assume here that $4a > \sigma^2$. Expanding (4), we get:

$$\begin{aligned}\hat{X}^n_{t_{i+1}} &= \frac{1}{4(1+\frac{kT}{2n})^2}\Bigg[2\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}}\right)^2 + 4\left(1+\frac{kT}{2n}\right)\frac{a-\sigma^2/4}{2}T/n \\ &+ 2\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}}\right)\sqrt{\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}}\right)^2 + 4\left(1+\frac{kT}{2n}\right)\frac{a-\sigma^2/4}{2}\frac{T}{n}}\Bigg].\end{aligned}$$

Thus, using the inequality $x^2 + x\sqrt{x^2+y} \le \begin{cases} 2x^2 + y/2 & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}$ for $y \ge 0$, we get that

$$\hat{X}^n_{t_{i+1}} \le \frac{1}{(1+\frac{kT}{2n})^2}\left[\left(\frac{\sigma}{2}(W_{t_{i+1}} - W_{t_i}) + \sqrt{\hat{X}^n_{t_i}}\right)^2 + \left(1+\frac{kT}{2n}\right)\frac{a-\sigma^2/4}{2}T/n\right]$$

and we can therefore apply Proposition 2.6 to deduce that $\hat{X}^n_{t_i}$ has bounded moments. Unfortunately, if we try now to get an expansion of $\hat{X}^n_{t_i}$ up to order 3/2 by expanding the square-root, we get a term in $\frac{1}{\sqrt{\hat{X}^n_{t_i}}}\mathcal{O}(1/n^{3/2})$ which is hard to manage. Despite the good numerical convergence of this scheme, our approach in this paper did not enable us to obtain theoretical results for it.

2.4.3 Expansion of the explicit scheme $E(\lambda)$

Let us assume here that $4a \geq \sigma^2$ and consider $\lambda \in [0, a - \sigma^2/4]$. Expanding (5), we get

$$\begin{aligned}\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} &= (a - k\hat{X}^n_{t_i})\frac{T}{n} + \frac{k^2}{4}\hat{X}^n_{t_i}\left(\frac{T}{n}\right)^2 + \frac{k\sigma^2}{8}\frac{2 - kT/(2n)}{(1 - kT/(2n))^2}\left(\frac{T}{n}\right)^2 \\ &\quad + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \left(\frac{\sigma^2}{4(1 - kT/(2n))^2} + \lambda\right)((W_{t_{i+1}} - W_{t_i})^2 - T/n) \\ &\leq \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \left(k^- + \frac{k^2}{4}\frac{T}{n}\right)\hat{X}^n_{t_i}\frac{T}{n} + \mathcal{O}(1/n).\end{aligned}$$

We can then apply Lemma 2.6 to deduce that $\hat{X}^n_{t_i} = \mathcal{O}(1)$. We have then an expansion analogous to that obtained for the implicit scheme, that is

$$\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} = \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m^n_{t_{i+1}} - m^n_{t_i} + \mathcal{O}(1/n^2) \qquad (14)$$

where $m^n_{t_i}$ is a $\mathcal{F}_{t_i}$-martingale defined by $m^n_{t_0} = 0$ and $m^n_{t_{i+1}} - m^n_{t_i} = (\frac{\sigma^2}{4} + \lambda)[(W_{t_{i+1}} - W_{t_i})^2 - T/n]$. It is in this case straightforward to see that we have the properties (6) and (9) and that the martingale increments satisfy (7) and (10) thanks to Remark 2.8. Hence, explicit scheme $E(\lambda)$ fulfills the conditions of $(\mathcal{H}_S)$ and $(\mathcal{H}_W)$.

3 Strong convergence

In all this section, we consider a scheme $(\hat{X}^n_{t_i})$ that satisfies the hypothesis $(\mathcal{H}_S)$. We will prove the strong convergence for it, following the method proposed by Deelstra and Delbaen [5] that relies on Yamada's functions. Thus, we first need to build a continuous adapted extension of our scheme in order to use then Itô's formula. For that purpose, we need to explicit the $\mathcal{O}$ terms and first define $Z^n_{t_i} = \mathcal{O}(1/n^{3/2})$ as:

$$\hat{X}^n_{t_{i+1}} = \hat{X}^n_{t_i} + \frac{T}{n}(a - k\hat{X}^n_{t_i}) + M^n_{t_{i+1}} - M^n_{t_i} + Z^n_{t_i}$$

We can suppose that $Z^n_{t_i}$ is $\mathcal{F}_{t_i}$-measurable. Indeed, if it were not the case, it would be sufficient then to consider the martingale increment

$$\tilde{M}^n_{t_{i+1}} - \tilde{M}^n_{t_i} = M^n_{t_{i+1}} - M^n_{t_i} + Z^n_{t_i} - \mathbb{E}[Z^n_{t_i}|\mathcal{F}_{t_i}]$$

and $\tilde{Z}^n_{t_i} = \mathbb{E}[Z^n_{t_i}|\mathcal{F}_{t_i}]$ instead of respectively $M^n_{t_{i+1}} - M^n_{t_i}$ and $Z^n_{t_i}$. Thus, we have $\tilde{M}^n_{t_{i+1}} - \tilde{M}^n_{t_i} + \tilde{Z}^n_{t_i} = M^n_{t_{i+1}} - M^n_{t_i} + Z^n_{t_i}$ and, thanks to Lemma 2.5, we get that $\tilde{Z}^n_{t_i} = \mathcal{O}(1/n^{3/2})$, and also $\tilde{M}^n_{t_{i+1}} - \tilde{M}^n_{t_i} = \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + \mathcal{O}(1/n)$.

Now, we apply the martingale representation theorem to the martingales $\{\mathbb{E}(m^n_{t_{i+1}}|\mathcal{F}_t) - m^n_{t_i}, t \in [t_i, t_{i+1}]\}$ to get the existence of an $\mathcal{F}_t$-adapted process $(R^n_t, 0 \le t \le T)$ such that

$$\mathbb{E}(m^n_{t_{i+1}}|\mathcal{F}_t) - m^n_{t_i} = \int_{t_i}^t R^n_s dW_s.$$

In particular, we know that $\int_{t_i}^t R^n_s dW_s = \mathcal{O}(1/n)$ and so $(\int_{t_i}^t R^n_s dW_s)^2 = \mathcal{O}(1/n^2)$ which gives us that, for $t \in [t_i, t_{i+1}]$:

$$\int_{t_i}^t \mathbb{E}[(R^n_s)^2]ds = O(1/n^2). \tag{15}$$

Now, we are able to build a continuous extension $(\hat{X}^n_t, 0 \le t \le T)$ $\mathcal{F}_t$-adapted of our discretization scheme. Indeed, we define for $t \in [t_i, t_{i+1}]$:

$$\hat{X}^n_t = \hat{X}^n_{t_i} + (t - t_i)(a - k\hat{X}^n_{t_i} + \frac{n}{T} Z^n_{t_i}) + \int_{t_i}^t (\sigma\sqrt{\hat{X}^n_{t_i}} + R^n_s)dW_s.$$

Thus, naming $\eta(t)$ the function defined on $[0, T]$ by $\eta(t) = t_i$ for $t \in [t_i, t_{i+1})$, we can rewrite our scheme as follows:

$$\hat{X}^n_t = x_0 + \int_0^t (a - k\hat{X}^n_{\eta(s)} + \frac{n}{T} Z^n_{\eta(s)})ds + \int_0^t (\sigma\sqrt{\hat{X}^n_{\eta(s)}} + R^n_s)dW_s. \tag{16}$$

Let us now introduce a family of Yamada's functions (see Karatzas and Shreve [10]) $\psi_{\epsilon,m}$ parametrized by two positive numbers ϵ and m. Since we have $\int_{\epsilon e^{-\sigma^2 m}}^{\epsilon} \frac{1}{\sigma^2 u} du = m$, there exists a continuous function $\rho_{\epsilon,m}$ with a compact support in $]\epsilon e^{-\sigma^2 m}, \epsilon[$ such that $\rho_{\epsilon,m}(x) \le \frac{2}{\sigma^2 x m}$ for $x > 0$ and $\int_{\epsilon e^{-\sigma^2 m}}^{\epsilon} \rho_{\epsilon,m}(u)du = 1$. We then consider

$$\psi_{\epsilon,m}(x) = \int_0^{|x|}\int_0^y \rho_{\epsilon,m}(u)dudy$$

that can be viewed as a sequence of smooth approximation of $x \longrightarrow |x|$ when m is large and ϵ tends to 0. Indeed functions $\psi_{\epsilon,m}$ thus satisfies:

$$|x| - \epsilon \le \psi_{\epsilon,m}(x) \le |x|, |\psi'_{\epsilon,m}(x)| \le 1, 0 \le \psi''_{\epsilon,m}(x) = \rho_{\epsilon,m}(|x|) \le \frac{2}{\sigma^2 |x| m}.$$

Following the method used by Deelstra and Delbaen [5], we first write

$$|\hat{X}^n_t - X_t| \le \epsilon + \psi_{\epsilon,m}(\hat{X}^n_t - X_t) \tag{17}$$

and then apply Itô's formula :

$$\begin{aligned}
\psi_{\epsilon,m}(\hat{X}^n_t - X_t) &= \int_0^t (kX_s - k\hat{X}^n_{\eta(s)} + \frac{n}{T} Z^n_{\eta(s)})\psi'_{\epsilon,m}(\hat{X}^n_s - X_s)ds \\
&\quad + \int_0^t (\sigma\sqrt{\hat{X}^n_{\eta(s)}} + R^n_s - \sigma\sqrt{X_s})\psi'_{\epsilon,m}(\hat{X}^n_s - X_s)dW_s \\
&\quad + \frac{1}{2}\int_0^t (\sigma\sqrt{\hat{X}^n_{\eta(s)}} + R^n_s - \sigma\sqrt{X_s})^2 \psi''_{\epsilon,m}(\hat{X}^n_s - X_s)ds \\
&=: I_1(t,n) + I_2(t,n) + I_3(t,n).
\end{aligned}$$

The absolute value of the first integral can be bounded using that $\|\psi'_{\epsilon,m}\|_\infty \leq 1$:

$$I_1(t,n) \leq |k| \int_0^t (|X_s - \hat{X}^n_s| + |\hat{X}^n_s - \hat{X}^n_{\eta(s)}|)ds + \int_0^t \frac{n}{T} |Z^n_{\eta(s)}| ds.$$

For the third integral, we have that

$$\begin{aligned}(\sigma\sqrt{\hat{X}^n_{\eta(s)}} + R^n_s - \sigma\sqrt{X_s})^2 &\leq 2(\sigma^2|X_s - \hat{X}^n_{\eta(s)}| + (R^n_s)^2) \\ &\leq 2(\sigma^2|X_s - \hat{X}^n_s| + \sigma^2|\hat{X}^n_s - \hat{X}^n_{\eta(s)}| + (R^n_s)^2). \quad (18)\end{aligned}$$

Therefore, using that $\psi''_{\epsilon,m}(x)|x| \leq \frac{2}{\sigma^2 m}$ and $\|\psi''_{\epsilon,m}\|_\infty \leq \frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m}$ we get:

$$I_3(t,n) \leq \frac{2t}{m} + \frac{2e^{\sigma^2 m}}{\epsilon m} \int_0^t (|\hat{X}^n_s - \hat{X}^n_{\eta(s)}| + \frac{1}{\sigma^2}(R^n_s)^2) ds.$$

Using Lemma 2.1, Corollary 2.7 and (15), we check that $\mathbb{E}[I_2(t,n)] = 0$. Now, taking the expectation in (17), we get

$$\forall t \in [0,T], \mathbb{E}(|\hat{X}^n_t - X_t|) \leq \epsilon + |k| \int_0^t \mathbb{E}(|\hat{X}^n_s - X_s|) ds + \frac{2T}{m} + \left(\frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m} + |k|\right) \frac{Cte}{\sqrt{n}}$$

for some $Cte > 0$, using that $|\hat{X}^n_s - \hat{X}^n_{\eta(s)}| = \mathcal{O}(1/\sqrt{n})$ and $\frac{n}{T} Z^n_{\eta(s)} = \mathcal{O}(1/\sqrt{n})$. Gronwall's lemma leads then to

$$\forall t \in [0,T], \mathbb{E}(|\hat{X}^n_t - X_t|) \leq e^{|k|T} \left[\epsilon + \frac{2T}{m} + \left(\frac{2e^{\sigma^2 m}}{\sigma^2 \epsilon m} + |k|\right) \frac{Cte}{\sqrt{n}}\right]. \quad (19)$$

Now, taking $m = \frac{1}{4\sigma^2} \ln(n)$ and $\epsilon = 1/\ln(n)$, we get that

$$\sup_{0 \leq t \leq T} \mathbb{E}(|\hat{X}^n_t - X_t|) = O(1/\ln(n)). \quad (20)$$

Now, we would like to exchange the supremum and the expectation. Doob's inequality gives $\mathbb{E}[\sup_{0 \leq s \leq t} |I_2(s,n)|] \leq C\sqrt{\mathbb{E}\left[\int_0^t (\sigma\sqrt{\hat{X}^n_{\eta(s)}} + R^n_s - \sigma\sqrt{X_s})^2 (\psi'_{\epsilon,m}(\hat{X}^n_s - X_s))^2 ds\right]}$. We use that $\|\psi'_{\epsilon,m}\|_\infty \leq 1$ and the inequality (18), and then control each terms thanks to relations (20) and (15), and observing that $|\hat{X}^n_s - \hat{X}^n_{\eta(s)}| = \mathcal{O}(1/\sqrt{n})$:

$$\mathbb{E}[\sup_{0 \leq s \leq t} |I_2(s,n)|] = O(1/\sqrt{\ln(n)}).$$

We can then use the same controls as before for I_1 and I_3 to conclude that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |\hat{X}^n_t - X_t|\right) = O(1/\sqrt{\ln(n)}). \quad (21)$$

We sum up our results in the proposition that follows.

Proposition 3.1. *Let us consider a discretization scheme $(\hat{X}^n)$ that satisfies the hypothesis $(\mathcal{H}_S)$. Then, there exists a positive constant C depending on T and on the parameters (k, a, σ, x_0) but not on n such that:*

$$\begin{aligned} \sup_{0\leq i\leq n} \mathbb{E}(|\hat{X}^n_{t_i} - X_{t_i}|) &\leq C/\ln(n) \\ \mathbb{E}\left(\sup_{0\leq i\leq n} |\hat{X}^n_{t_i} - X_{t_i}|\right) &= C/\sqrt{\ln(n)}. \end{aligned}$$

4 Weak convergence

In this section, we will establish a result that gives the convergence rate of $\mathbb{E}[f(\hat{X}^n_T)]$ to $\mathbb{E}[f(X_T)]$. We will use the method introduced by Talay and Tubaro (1990) to study that weak error and get also a convergence rate in $1/n$ provided that f is regular enough. We thus introduce the notation X^x_t to denote the CIR process with initial value x, and we first need to establish the following technical result.

Proposition 4.1. *Let us consider $f : \mathbb{R}_+ \to \mathbb{R}$ a $\mathcal{C}^q$ function with $q \geq 2$, such that there is $A > 0$ and $m \geq q$, $m \in \mathbb{N}$ such that*

$$\forall x \geq 0,\ |f^{(q)}(x)| \leq A(1 + x^m).$$

Then $u : [0, T] \times \mathbb{R}_+ \to \mathbb{R}$ defined by $u(t, x) = \mathbb{E}[f(X^x_{T-t})]$ has successive derivatives $\partial^l_x \partial^{l'}_t u(t, x)$ for $l, l' \in \mathbb{N}$ and $l + 2l' \leq q$, that satisfy the following property:

$$\exists C > 0, \forall (t, x) \in [0, T] \times \mathbb{R}_+,\ \max_{l+2l'\leq q} |\partial^l_x \partial^{l'}_t u(t, x)| \leq C(1 + x^{m+q+l'}) \tag{22}$$

and is a classical solution of the PDE:

$$\begin{cases} \partial_t u(t, x) + (a - kx)\partial_x u(t, x) + \frac{\sigma^2}{2} x \partial^2_x u(t, x) = 0 \\ u(T, x) = f(x). \end{cases} \tag{23}$$

More generally, let us assume that $(f_\theta, \theta \in \Theta)$ is a family of $\mathcal{C}^q$ functions with $q \geq 2$, such that there is $A > 0$ and $m \geq q$, $m \in \mathbb{N}$ such that

$$\forall \theta \in \Theta,\ \forall x \geq 0,\ |f^{(q)}_\theta(x)| \leq A(1 + x^m) \text{ and } \forall l < q, |f^{(l)}_\theta(0)| \leq A. \tag{24}$$

For $0 \leq \tau \leq T$, we consider $u_{\theta,\tau}(t, x) = \mathbb{E}[f_\theta(X^x_{\tau-t})]$ for $0 \leq t \leq \tau$ and $x \geq 0$. Then there is a constant $C > 0$ that does not depend on τ and θ such that

$$\forall \theta \in \Theta, \tau \in [0, T],\ \forall (t, x) \in [0, \tau] \times \mathbb{R}_+,\ \max_{l+2l'\leq q} |\partial^l_x \partial^{l'}_t u_{\theta,\tau}(t, x)| \leq C(1 + x^{m+q+l'}) \tag{25}$$

The proof of this proposition, mainly based on the analytical formula available for the transition density of the CIR process is made in the Appendix A.

We are now able to prove the main results of this section:

Proposition 4.2. *Let $f:\mathbb{R}_+\to\mathbb{R}$ be a C^4 function such that $\exists A,m>0,\forall x\geq 0,|f^{(4)}(x)|\leq A(1+x^m)$. Let us suppose moreover that the scheme $(\hat{X}^n)$ satisfies the hypothesis $(\mathcal{H}_W)$. Then, the weak error is in $1/n$:*

$$\mathbb{E}[f(\hat{X}^n_T)]=\mathbb{E}[f(X_T)]+O(1/n).$$

More generally, if $(f_\theta,\theta\in\Theta)$ is a family of $\mathcal{C}^4$ functions satisfying condition (24) for $q=4$,

$$\mathbb{E}[f_\theta(\hat{X}^n_{t^n_j})]=\mathbb{E}[f_\theta(X_{t^n_j})]+O(1/n)$$

where $O(1/n)$ has to be understood in the sense of Definition 2.2 with $(\theta,t^n_j)\in\Gamma_n=\Theta\times\{t^n_0,...,t^n_n\}$.

Proof : We have $\mathbb{E}[f(\hat{X}^n_T)]=\mathbb{E}[u(T,\hat{X}^n_T)]$ and $\mathbb{E}[f(X_T)]=u(0,x_0)$ so that:

$$\mathbb{E}[f(\hat{X}^n_T)]-\mathbb{E}[f(X_T)]=\mathbb{E}[u(T,\hat{X}^n_T)-u(0,x_0)]=\sum_{i=0}^{n-1}\mathbb{E}[u(t_{i+1},\hat{X}^n_{t_{i+1}})-u(t_i,\hat{X}^n_{t_i})].$$

Let us consider (t,x) and (s,y) in $[0,T]\times\mathbb{R}_+$. We can apply the Taylor formula to $t\mapsto u(t,y)$ up to order 2 and get:

$$u(s,y)=u(t,y)+(s-t)\partial_t u(t,y)+(s-t)^2\int_0^1(1-\tau)\partial_t^2u(t+\tau(s-t),y)d\tau.$$

Now, we apply Taylor formula to $y\mapsto u(t,y)$ and $y\mapsto\partial_t u(t,y)$ and we finally get

$$\begin{aligned}u(s,y)\ =\ &\sum_{0\leq l+2l'<4}\partial_x^l\partial_t^{l'}u(t,x)\frac{(s-t)^{l'}(y-x)^l}{l!l'!}+(s-t)^2\int_0^1(1-\tau)\partial_t^2u(t+\tau(s-t),y)d\tau\\&+(s-t)(y-x)^2\int_0^1(1-\xi)\partial_x^2\partial_t u(t,x+\xi(y-x))d\xi\\&+\frac{(y-x)^4}{3!}\int_0^1(1-\xi)^3\partial_x^4u(t,x+\xi(y-x))d\xi.\end{aligned}$$

Proposition 4.1 allows us then to get:

$$\begin{aligned}&\left|u(s,y)-\sum_{0\leq l+2l'<4}\partial_x^l\partial_t^{l'}u(t,x)\frac{(s-t)^{l'}(y-x)^l}{l!l'!}\right|\\&\qquad\leq C(1+\max(x,y)^{6+m})\left[(s-t)^2+|s-t|(y-x)^2+(y-x)^4\right]\end{aligned}$$

and we apply this bound to $(t_i,\hat{X}^n_{t_i})$ and $(t_{i+1},\hat{X}^n_{t_{i+1}})$. Proposition 2.4 and Corollary 2.7 give immediately that $C(1+\max(\hat{X}^n_{t_i},\hat{X}^n_{t_{i+1}})^{6+m})[(T/n)^2+(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^2T/n+(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^4]=\mathcal{O}(1/n^2)$ and therefore:

$$u(t_{i+1},\hat{X}^n_{t_{i+1}})-u(t_i,\hat{X}^n_{t_i})=\sum_{0<l+2l'<4}\partial_x^l\partial_t^{l'}u(t_i,\hat{X}^n_{t_i})\frac{(T/n)^{l'}(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l}{l!l'!}+\mathcal{O}(1/n^2).\quad(26)$$

Now we expand the powers of $(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})$ up to order 2 using the Hypothesis $(\mathcal{H}_W)$:

$$\begin{aligned}\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} &= \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}} - W_{t_i}) + m^n_{t_{i+1}} - m^n_{t_i} + \mathcal{O}(1/n^2)\\ (\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^3 &= \sigma^3(\hat{X}^n_{t_i})^{3/2}(W_{t_{i+1}} - W_{t_i})^3 + \mathcal{O}(1/n^2)\end{aligned}$$

Therefore, we get that

$$\begin{aligned}\mathbb{E}\left[\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} | \mathcal{F}_{t_i}\right] &= \frac{T}{n}(a - k\hat{X}^n_{t_i}) + \mathcal{O}(1/n^2)\\ \mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^3 | \mathcal{F}_{t_i}\right] &= \mathcal{O}(1/n^2)\\ \text{and according to (10), } \mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^2 | \mathcal{F}_{t_i}\right] &= \sigma^2\hat{X}^n_{t_i}T/n + \mathcal{O}(1/n^2).\end{aligned}$$

The bound (22) and Lemma 2.6 ensure that $\partial_x^l \partial_t^{l'} u(t_i, \hat{X}^n_{t_i}) = \mathcal{O}(1)$ for $l + 2l' < 4$. Thus, using Lemma 2.5, we can deduce from (26) :

$$\begin{aligned}&\mathbb{E}\left[u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i}) | \mathcal{F}_{t_i}\right]\\ &= \sum_{0<l+2l'<4} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}^n_{t_i}) \frac{(T/n)^{l'} \mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^l | \mathcal{F}_{t_i}\right]}{l! l'!} + \mathcal{O}(1/n^2)\\ &= \partial_t u(t_i, \hat{X}^n_{t_i})T/n + \partial_x u(t_i, \hat{X}^n_{t_i})\frac{T}{n}(a - k\hat{X}^n_{t_i}) + \partial_x^2 u(t_i, \hat{X}^n_{t_i})\frac{\sigma^2}{2}\hat{X}^n_{t_i}T/n + \mathcal{O}(1/n^2)\\ &= \mathcal{O}(1/n^2)\end{aligned}$$

since u solves the PDE (23). Therefore, there is a constant $C > 0$ that does not depend on i such that $\left|\mathbb{E}\left[u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i})\right]\right| \leq C/n^2$ and so, we finally get that

$$\left|\mathbb{E}[f(\hat{X}^n_T)] - \mathbb{E}[f(X_T)]\right| \leq C/n$$

which is the desired result.

Now let us explain why this proof can be generalized easily to the case of the family of functions f_θ that satisfy (24) and all times t^n_j. We apply as before Taylor formula to functions u_{θ,t^n_j} and thanks to Proposition 4.1, the bounds we have on its derivatives do not depend on (θ, t^n_j) and we get for $0 \leq i < j \leq n$ as in (26)

$$u_{\theta,t^n_j}(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u_{\theta,t^n_j}(t_i, \hat{X}^n_{t_i}) = \sum_{0<l+2l'<4} \partial_x^l \partial_t^{l'} u_{\theta,t^n_j}(t_i, \hat{X}^n_{t_i}) \frac{(T/n)^{l'}(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^l}{l! l'!} + \mathcal{O}(1/n^2)$$

with the difference that the $\mathcal{O}$ symbol is now meant with $\Gamma_n = \{(t^n_i, t^n_j), 0 \leq i < j \leq n\} \times \Theta$ instead of $\{t^n_i, 0 \leq i \leq n\}$ before. Then, the proof is the same, noticing that we still have $\partial_x^l \partial_t^{l'} u_{\theta,t^n_j}(t_i, \hat{X}^n_{t_i}) = \mathcal{O}(1)$ for $l + 2l' \leq q$. □

Remark 4.3. *We desired to get a weak error in $1/n$ as in the case of the Euler scheme for stochastic differential equations with coefficients regular enough (C^4 and bounded derivatives). Using the argument of Talay and Tubaro, we need then a control on $u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i})$ up to order 2. This is why we assume to know the relation (9) between $\hat{X}^n_{t_{i+1}}$ and $\hat{X}^n_{t_i}$ up to order 2. Expanding $u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i})$, we see that the term of order $1/2$ has a null expectation, the term of order $1/n$ is null since u solves the PDE (23), but we need to require condition (10) so that the term of order $3/2$ has a null expectation. If we had only assumed that the scheme satisfies $(\mathcal{H}_S)$, we would have obtained a weak error in $1/\sqrt{n}$.*

Now, we would like to expand further the weak error, in particular to justify the use the Romberg method that mainly relies on the following remark: if we know that there is $c_1 \in \mathbb{R}$ such that $\mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + c_1/n + O(1/n^2)$, then $2\mathbb{E}[f(\hat{X}^{2n}_T)] - \mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + O(1/n^2)$ converges thus faster toward the desired expectation. If we want to adapt the previous proof, we see that we need to add the following assumptions to get a weak error up to order $\nu \in \mathbb{N}^*$:

- f is regular enough ($\mathcal{C}^{4\nu}$) and its derivatives have a polynomial growth.
- We know the relation between $\hat{X}^n_{t_{i+1}}$ and $\hat{X}^n_{t_i}$ up to order $\nu + 1$.

Moreover, if we wish to have as for the Euler scheme an error that expands only on the integer orders: $\mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + c_1/n + c_2/n^2 + .. + c_{\nu-1}/n^{\nu-1} + O(1/n^\nu)$, we need to make assumptions of the same kind as (10) for any power of $(\hat{X}^n_{t_i+1} - \hat{X}^n_{t_i})$ to get terms of order "integer + one half" with null expectation. However these assumptions would be hardly readable, and practically, they would be clearly satisfied only by the explicit schemes $E(\lambda)$. That's why we prefer to state here directly the result for the explicit schemes $E(\lambda)$.

Proposition 4.4. *Let $\nu \in \mathbb{N}^*$ and $f : \mathbb{R}_+ \to \mathbb{R}$ that we suppose $\mathcal{C}^\infty$ and such that $\forall q, \exists A_q > 0, m_q \in \mathbb{N}$, $|f^{(q)}(x)| \leq A_q(1 + x^{m_q})$. Let $(\hat{X}^n)$ be the explicit scheme $E(\lambda)$ with $0 \leq \lambda \leq a - \sigma^2/4$. Then, the weak error has an expansion up to order ν:*

$$\mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + c_1/n + c_2/n^2 + .. + c_{\nu-1}/n^{\nu-1} + O(1/n^\nu)$$

where $c_1 = T\int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)]dt$ with $\psi_{E(\lambda)}$ defined below in (28).

Proof : With the same argument as in Proposition 4.2, first using the Taylor expansion respect to t and then to x, we get that there is $C(\nu) > 0$ and $M(\nu) \in \mathbb{N}$:

$$\left| u(s,y) - \sum_{0 \leq l+2l' < 2\nu+2} \partial_x^l \partial_t^{l'} u(t,x) \frac{(s-t)^{l'}(y-x)^l}{l!l'!} \right|$$
$$\leq C(\nu)(1 + \max(x,y)^{M(\nu)}) \sum_{j=0}^{\nu+1} |s-t|^{\nu+1-j}(y-x)^{2j}.$$

Similarly, we get that

$$u(t_{i+1},\hat{X}^n_{t_{i+1}})-u(t_i,\hat{X}^n_{t_i})=\sum_{0<l+2l'<2\nu+2}\partial_x^l\partial_t^{l'}u(t_i,\hat{X}^n_{t_i})\frac{(T/n)^{l'}(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l}{l!l'!}+\mathcal{O}(1/n^{\nu+1}).$$

and then

$$\begin{aligned}\mathbb{E}\left[u(t_{i+1},\hat{X}^n_{t_{i+1}})-u(t_i,\hat{X}^n_{t_i})|\mathcal{F}_{t_i}\right] &= \sum_{0<l+2l'<2\nu+2}\partial_x^l\partial_t^{l'}u(t_i,\hat{X}^n_{t_i})\frac{(T/n)^{l'}\mathbb{E}\left[(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l|\mathcal{F}_{t_i}\right]}{l!l'!}\\ &\quad +\mathcal{O}(1/n^{\nu+1}).\end{aligned}$$

Let us first expand (5) to get:

$$\begin{aligned}\hat{X}^n_{t_{i+1}} &= \hat{X}^n_{t_i}+\sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}}-W_{t_i})+(a-k\hat{X}^n_{t_i})\frac{T}{n}+\left(\lambda+\frac{\sigma^2}{4}\right)((W_{t_{i+1}}-W_{t_i})^2-T/n)\\ &\quad +\frac{k^2}{4}\hat{X}^n_{t_i}(T/n)^2+\frac{\sigma^2}{4}(W_{t_{i+1}}-W_{t_i})^2\left(\frac{1}{\left(1-\frac{kT}{2n}\right)^2}-1\right).\end{aligned}$$

Since $\frac{1}{\left(1-\frac{kT}{2n}\right)^2}-1=\sum_{j\geq1}(j+1)(k/2)^j(T/n)^j$, we get that

$$\begin{aligned}\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i} &= \sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}}-W_{t_i})+(a-k\hat{X}^n_{t_i})\frac{T}{n}+\left(\lambda+\frac{\sigma^2}{4}\right)((W_{t_{i+1}}-W_{t_i})^2-T/n)\\ &\quad +\frac{k^2}{4}\hat{X}^n_{t_i}(T/n)^2+\frac{\sigma^2}{4}(W_{t_{i+1}}-W_{t_i})^2\sum_{j=1}^{\nu-1}(j+1)(k/2)^j(T/n)^j+\mathcal{O}(1/n^{\nu+1}).\end{aligned}$$

All the terms here are of integer order but $\sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}}-W_{t_i})$ that is of order 1/2. Now, taking the power l of these expansion, we get using Proposition 2.4 an expansion of $(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l$ up to order $\nu+1$ (even $\nu+1+(l-1)/2$). What is important to remark is that the term of order "integer + one half" comes from an odd power of $\sigma\sqrt{\hat{X}^n_{t_i}}(W_{t_{i+1}}-W_{t_i})$ and a product of the other terms. Since all these other terms are even respect to $(W_{t_{i+1}}-W_{t_i})$, we finally get that all the terms of order "integer + one half" have a null conditional expectation. Thus, we see that we can write for $l\in\mathbb{N}$

$$\mathbb{E}\left[(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l|\mathcal{F}_{t_i}\right]=\sum_{j\geq l/2}^{\nu}\phi^{l,j}(\hat{X}^n_{t_i})(T/n)^j+\mathcal{O}(1/n^{\nu+1})$$

where $\phi^{l,j}$ are polynomial functions that we do not explicit and satisfy $\phi^{l,j}(\hat{X}^n_{t_i})=\mathcal{O}(1)$. Thus, $(T/n)^{l'}\mathbb{E}\left[(\hat{X}^n_{t_{i+1}}-\hat{X}^n_{t_i})^l|\mathcal{F}_{t_i}\right]=\sum_{l/2\leq j\leq\nu}\phi^{l,j}(\hat{X}^n_{t_i})(T/n)^{j+l'}+\mathcal{O}(1/n^{\nu+1})$

$= \sum_{l+2l' \leq 2j < 2\nu+2} \phi^{l,j-l'}(\hat{X}^n_{t_i})(T/n)^j + \mathcal{O}(1/n^{\nu+1})$ and so, $\mathbb{E}\left[u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i})|\mathcal{F}_{t_i}\right] =$

$$\sum_{j=1}^{\nu}(T/n)^j \left(\sum_{0<l+2l' \leq 2j} \partial_x^l \partial_t^{l'} u(t_i, \hat{X}^n_{t_i}) \frac{\phi^{l,j-l'}(\hat{X}^n_{t_i})}{l!l'!} \right) + \mathcal{O}(1/n^{\nu+1}). \tag{27}$$

For $\nu = 2$, one obtains:

$$\begin{aligned}
\mathbb{E}\left[\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i} \,|\mathcal{F}_{t_i}\right] &= (a - k\hat{X}^n_{t_i})T/n + \frac{k^2\hat{X}^n_{t_i} + k\sigma^2}{4}(T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^2 \,|\mathcal{F}_{t_i}\right] &= \sigma^2 \hat{X}^n_{t_i} T/n + \left[(a - k\hat{X}^n_{t_i})^2 + 2(\lambda + \sigma^2/4)^2\right](T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^3 \,|\mathcal{F}_{t_i}\right] &= 3\sigma^2 \hat{X}^n_{t_i}\left[a - k\hat{X}^n_{t_i} + 2(\lambda + \sigma^2/4)\right](T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^4 \,|\mathcal{F}_{t_i}\right] &= 3\sigma^4 (\hat{X}^n_{t_i})^2 (T/n)^2 + \mathcal{O}(1/n^3) \\
\mathbb{E}\left[(\hat{X}^n_{t_{i+1}} - \hat{X}^n_{t_i})^5 \,|\mathcal{F}_{t_i}\right] &= \mathcal{O}(1/n^3)
\end{aligned}$$

and so:

$$\mathbb{E}\left[u(t_{i+1}, \hat{X}^n_{t_{i+1}}) - u(t_i, \hat{X}^n_{t_i}) \,|\mathcal{F}_{t_i}\right] = (T/n)^2 \psi_{E(\lambda)}(t_i, \hat{X}^n_{t_i}) + \mathcal{O}(1/n^3)$$

where

$$\begin{aligned}
\psi_{E(\lambda)}(t,x) &= \frac{1}{2}\partial_t^2 u(t,x) + \frac{k^2 x + k\sigma^2}{4}\partial_x u(t,x) + (a - kx)\partial_x \partial_t u(t,x) \\
&+ \frac{1}{2}\left[(a - kx)^2 + 2(\lambda + \sigma^2/4)^2\right]\partial_x^2 u(t,x) + \frac{\sigma^2}{2} x \partial_x^2 \partial_t u(t,x) \\
&+ \frac{\sigma^2}{2} x\left[a - kx + 2(\lambda + \sigma^2/4)\right]\partial_x^3 u(t,x) + \frac{\sigma^4}{8} x^2 \partial_x^4 u(t,x).
\end{aligned} \tag{28}$$

Therefore, summing and taking the expectation, we get that $\mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + (T/n)^2 \sum_{i=0}^{n-1} \mathbb{E}[\psi_{E(\lambda)}(t_i, \hat{X}^n_{t_i})] + O(1/n^2)$. We then apply Proposition 4.2 to the family of functions $x \mapsto \psi_{E(\lambda)}(t_i, x)$ which satisfies condition (24) thanks to (22) (we incidentally remark that it is sufficient to have f $\mathcal{C}^8$ to get the expansion with $\nu = 2$). It gives that $\mathbb{E}[\psi_{E(\lambda)}(t_i, \hat{X}^n_{t_i})] = \mathbb{E}[\psi_{E(\lambda)}(t_i, X_{t_i})] + \mathcal{O}(1/n)$. Since $\partial_t \mathbb{E}[\psi_{E(\lambda)}(t, X_t)]$ is bounded on $[0,T]$, we have that $(T/n)\sum_{i=0}^{n-1} \mathbb{E}[\psi_{E(\lambda)}(t_i, X_{t_i})] = \int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)]dt + O(1/n)$ and then:

$$\mathbb{E}[f(\hat{X}^n_T)] = \mathbb{E}[f(X_T)] + (T/n)\int_0^T \mathbb{E}[\psi_{E(\lambda)}(t, X_t)]dt + O(1/n^2). \tag{29}$$

To get the expansion for $\nu = 3$ and further, one has to check by induction the desired result for any ν using the same methodology. □

5 Numerical results

In this section, we will analyze numerically the convergence of the discretization schemes. For the theoritical study, an interesting feature of the implicit schemes (3) and (4) and of the explicit schemes $E(\lambda)$, is their "automatic" nonnegativity for the following parameters:

Scheme	Condition on (a, σ)
Implicit (3)	$\sigma^2 \leq 2a$
Implicit (4)	$\sigma^2 \leq 4a$
$E(\lambda)$	$0 \leq \lambda \leq a - \sigma^2/4$

(30)

Indeed, contrary to the schemes using a reflection technique as those proposed by Deelstra-Delbaen or Diop, there is no need to control the reflection. However, we can use the following trick to extend schemes (3), (4) and $E(\lambda)$ to all the values of the parameters (k, a, σ):

- For the implicit schemes which are defined with second-degree polynomials, we will set $\hat{X}^n_{t_{i+1}} = 0$ when the discriminant is negative and else use formulas (3) and (4).
- For the explicit schemes $E(\lambda)$, we simply define $\hat{X}^n_{t_{i+1}}$ as the positive part of the left-hand side of (5)

We will use these extensions when needed for the simulations presented in this section.

5.1 Numerical study of the strong convergence

In this paragraph we present a numerical analysis of the strong convergence of various schemes. It does not seem possible to compute the limit process on the same probability space, and we overcome this difficulty using the following lemma that says that it is sufficient to study the difference between the values obtained with a scheme for a given time step and the ones obtained with the same scheme and a time step twice smaller. Let us recall here that $t_n^i = iT/n = t_{2n}^{2i}$.

Lemma 5.1. *Let us consider a scheme $(\hat{X}^n_{t_i})$ that converges toward a continuous process X_t in the following sense:*

$$\mathbb{E}\left[\sup_{0\leq i\leq n} |\hat{X}^n_{t_i^n} - X_{t_i^n}|\right] \underset{n\to\infty}{\longrightarrow} 0. \tag{31}$$

Then, for any $\alpha > 0$ and $\beta \geq 0$,

$$\mathbb{E}\left[\sup_{0\leq i\leq n} |\hat{X}^n_{t_i^n} - X_{t_i^n}|\right] = O(\frac{(\ln n)^\beta}{n^\alpha}) \iff \mathbb{E}\left[\sup_{0\leq i\leq n} |\hat{X}^n_{t_i^n} - \hat{X}^{2n}_{t_{2i}^{2n}}|\right] = O(\frac{(\ln n)^\beta}{n^\alpha}).$$

The condition (31) has been established in this paper for the explicit schemes $E(\lambda)$ and for the implicit scheme (3) and it has also been proved for the scheme of Deelstra-Delbaen [5]. Under some restrictive conditions of the parameters, the scheme proposed by Diop converges with a $1/\sqrt{n}$ rate [6]. For the other parameters and for the Implicit scheme on the square-root (4), we can check numerically the condition (31) doing the comparison with a scheme on which this comparison has been proved.

Proof of the Lemma. If there is $K>0$ such that $\forall n \in \mathbb{N}^*, \mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - X_{t^n_i}|\right] \le K\frac{(\ln n)^\beta}{n^\alpha}$, then

$$\mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - \hat{X}^{2n}_{t^{2n}_{2i}}|\right] \le K\left(\frac{(\ln n)^\beta}{n^\alpha} + \frac{(\ln n + \ln 2)^\beta}{(2n)^\alpha}\right) \le K'\frac{(\ln n)^\beta}{n^\alpha}.$$

Reciprocally, since $\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - X_{t^n_i}| \le \sum_{k=0}^{l} \sup_{0\le i\le n} |\hat{X}^n_{t^{2^k n}_{2^k i}} - \hat{X}^{2n}_{t^{2^{k+1}n}_{2^{k+1}i}}| + \sup_{0\le i\le n} |\hat{X}^{2n}_{t^{2^{l+1}n}_{2^{l+1}i}} - X_{t^n_i}|$, we get $\mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - X_{t^n_i}|\right] \le K\sum_{k=0}^{l} \frac{(\ln n + k\ln 2)^\beta}{(2^k n)^\alpha} + \mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^{2n}_{t^{2^{l+1}n}_{2^{l+1}i}} - X_{t^n_i}|\right]$ and with $l\to\infty$,

$$\mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - X_{t^n_i}|\right] \le K\sum_{k=0}^{\infty} C_\beta \frac{(\ln n)^\beta + (k\ln 2)^\beta}{(2^k n)^\alpha} \le K'\frac{(\ln n)^\beta}{n^\alpha}$$

for some constant $K'>0$, using that $\sum_{k=0}^{\infty} k^\beta/2^k < \infty$.□

Now for the numerical study, we consider the standard time interval $[0,1]$ $(T=1)$ and set $S_n = \mathbb{E}\left[\sup_{0\le i\le n} |\hat{X}^n_{t^n_i} - \hat{X}^{2n}_{t^{2n}_{2i}}|\right]$. The figures below show the convergence of S_n in function of the time-step $1/n$ for different parameters. Let us first observe that the implicit scheme (4) and the explicit scheme $E(0)$ give errors smaller than the others, for all the values of the parameters tested. Which is also interesting and nontrivial is that the behaviour of the convergence depends on the parameters.

We notice that for the case $2a > \sigma^2$ the schemes (4) and $E(0)$ present an error which looks linear respect to the time-step while the others give a square-root shape (see Fig. 1). This is not totally surprising because we have seen that these schemes correspond to the Milstein expansion, and we also know that under this hypothesis, X_t never reaches 0 so that the non-lipschitzian behaviour of the square-root is less important.

When $2a < \sigma^2 < 4a$, the schemes (3), $E(\sigma^2/8)$, $E(\sigma^2/4)$, Deelstra-Delbaen and Diop, S_n still has a square-root behaviour (see Fig. 2). Finally, let us mention that for the last case $\sigma^2 > 4a$, the schemes (4) and $E(0)$ still give the smaller value of S_n. However, we have to say that when $\sigma^2 >> 4a$, the convergence is really slow.

Lastly, concerning the impact of λ for the explicit schemes $E(\lambda)$, we see (Fig. 1 and 2) that $\lambda = \sigma^2/4$ is the parameter that gives a strong convergence analogous to the schemes of Diop, Deelstra-Delbaen and implicit (3); and the value of S_n for $E(\sigma^2/8)$ is as one can expect between those of $E(0)$ and $E(\sigma^2/4)$.

To get an idea of the speed of convergence in function of the parameters, we postulate that $S_n \sim C/n^\alpha$ with $\alpha > 0\cdot$ Thanks to the lemma, this is equivalent to a strong convergence speed in $1/n^\alpha$. To estimate α, we remark that

$$\log_{10}(S_n) - \log_{10}(S_{10n}) \underset{n\to+\infty}{\longrightarrow} \alpha,$$

and we have reported $\log_{10}(S_n) - \log_{10}(S_{10n})$ for $n=200$ in Figure 3. We have plotted the result in function of the parameter $\sigma^2/(2a)$ since it is the one that plays a key role. This

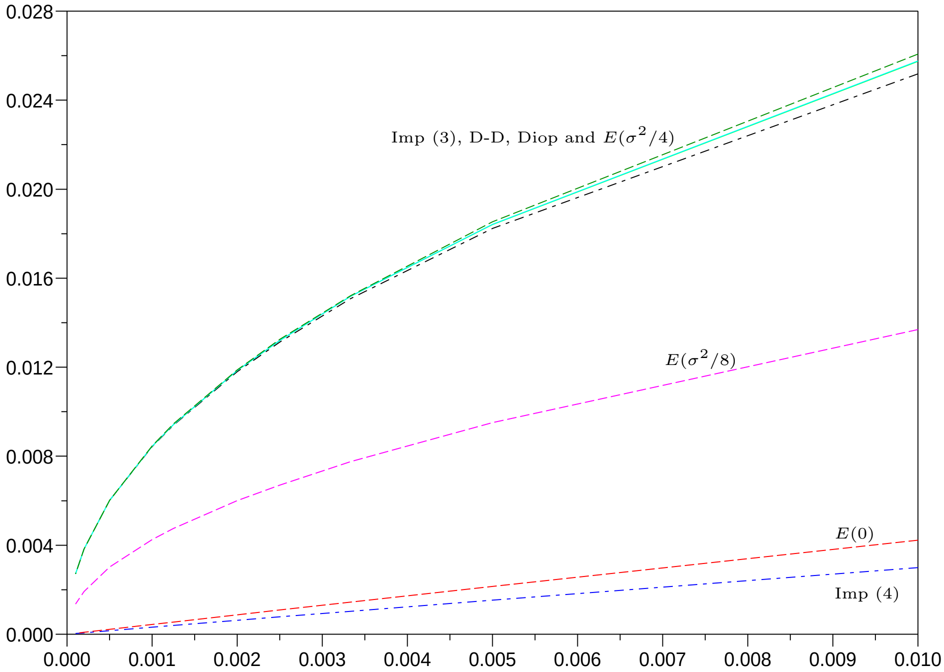

Figure 1: S_n in function of the time-step $1/n$ for $x_0 = 1$, $k = 1$, $a = 1$ and $\sigma = 1$.

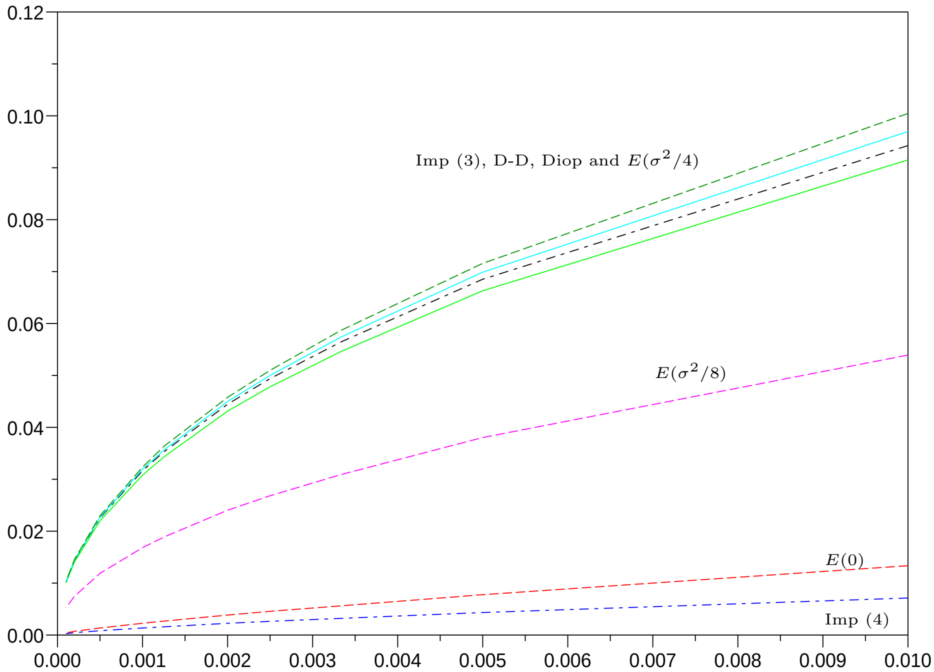

Figure 2: S_n in function of the time-step $1/n$ for $x_0 = 1$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

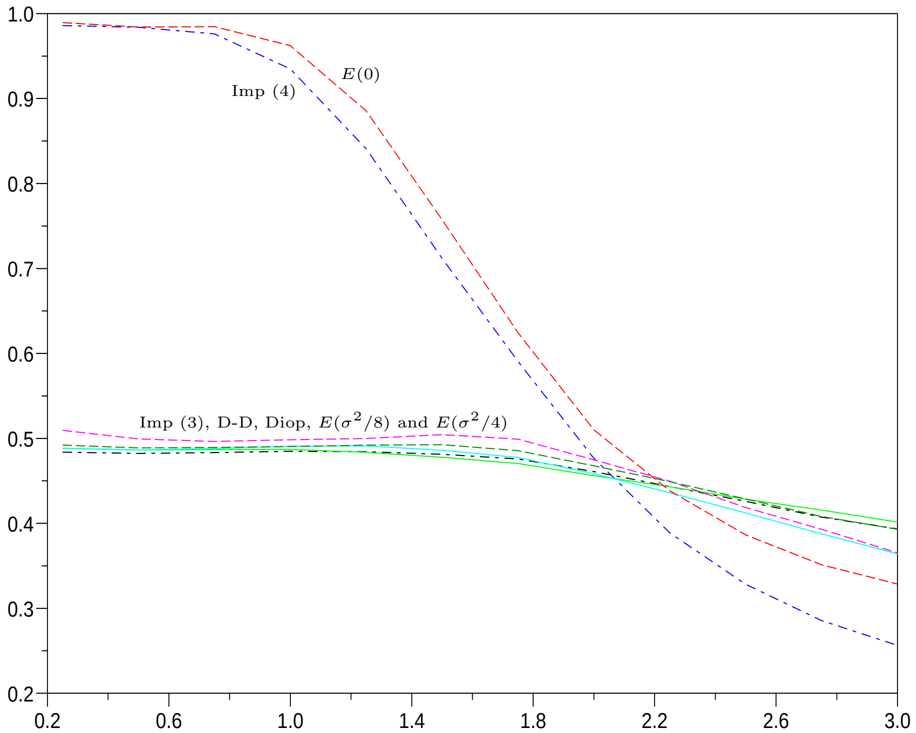

Figure 3: Speed convergence of S_n: estimation of the α parameter in function of $\sigma^2/(2a)$ for $x_0 = 1$, $k = 1$ and $a = 1$.

can be understood easily with a time-scaling. For the schemes (4) and $E(0)$, the estimated α is close to 1 for $\sigma^2 < 2a$ and decreases from 1 to $1/2$ for $2a < \sigma^2 < 4a$ while for the other schemes, the estimated value of α is close to $1/2$ for $\sigma^2 < 4a$. Intuitively, we can understand this decrease because for $\sigma^2 > 2a$, X_t can reach the origin, and a non negligible time is spent in the neighbourhood of 0 where the square root is non Lipschitz. Obviously, the speed of convergence may have a more complicated form than the one postulated, but our method gives nonetheless a good idea of its behaviour.

5.2 Numerical study of the weak convergence

We have plotted in figures 4,5,6 and 7, for fixed parameters of the CIR process, the approximation given by the scheme or a Romberg extrapolation of the expected value $\mathbb{E}[f(X_1)]$ for the function $f(x) = \frac{5+3x^4}{2+5x}$. This function has been chosen to be sensitive to variation for large and small values so that it catches the defaults of the schemes near 0 and ∞. We have taken two sets of parameters that illustrate the cases $\sigma^2 \leq 2a$ and $2a \leq \sigma^2 \leq 4a$.

	Implicit (3)	Implicit (4)	Diop	Deelstra-Delbaen	$E(0)$	$E(\sigma^2/4)$	Exact
$\sigma = 1$	72	64	65	67	67	68	668
$\sigma = \sqrt{2}$	77	64	67	67	66	70	1092

Table 1: Simulation time (in s) for 10^6 paths with a time step equal to 10^{-3} and parameters $k = 1,\ a = 1$ and $x_0 = 1$.

Let us recall here that we have proved here the $O(1/n)$ convergence only for regular functions and for the schemes satisfying $(\mathcal{H}_W)$, that is (3) with $\sigma^2 < 2a$ and $E(\lambda)$ with $0 \leq \lambda \leq a - \sigma^2/4$. What comes out from the computations (see Figures 4 and 5) is that for the small values of σ ($\sigma^2 \leq 2a$, Fig. 4) all the schemes seem to have a behaviour in $O(1/n)$ while for the large values ($\sigma^2 > 2a$, Fig. 5), only the Explicit schemes and the Deelstra-Delbaen scheme give shapes compatible with a behaviour in in $O(1/n)$. On the contrary, the scheme of Diop shows clearly a root shape while the implicit schemes (3) and (4) seem to converge a little bit slower than K/n.

Concerning the Romberg method to calculate $\mathbb{E}(f(\hat{X}_1))$, the figure 6 show that in the both cases $\sigma^2 \leq 2a$ and $\sigma^2 > 2a$, Diop's and implicit schemes (3) and (4) do not show a quadratic convergence. As expected, Explicit schemes have a quadratic shape in all the cases even if, strictly speaking, we have not proved the speed convergence observed for $E(\sigma^2/8)$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$ since $\lambda = \sigma^2/8 > a - \sigma^2/4$. Concerning the Deelstra-Delbaen scheme, let us first say that for large time-steps, negative values may be frequent which explains the strange behaviour observed. However, for time-steps small enough, the convergence seems compatible with a quadratic convergence.

5.3 Computation time required by the schemes

In this paragraph we compare the time required by the schemes and the exact method to simulate 10^6 paths with a time step equal to 10^{-3} on the time interval $[0, 1]$ (see Table 1). Concerning the exact simulation of the increment of the CIR process, we have used the method proposed by Glasserman in [7] (see p. 120-128). As we could expect, this method is more time-consuming (up to a factor 10). Thus, it should be used to compute expectations that depend on the values of the process (X_t) at a few fixed times. On the contrary, for expectations that depends on all the path (such as integrals), discretization schemes should be preferred. As we see in Table 1, the time required by the schemes presented are of the same order. Let us mention here that for the implicit scheme (4), one has to be careful and store at each step the value of $\sqrt{\hat{X}^n_{t_i}}$ so that only one square-root has to be computed at each time step.

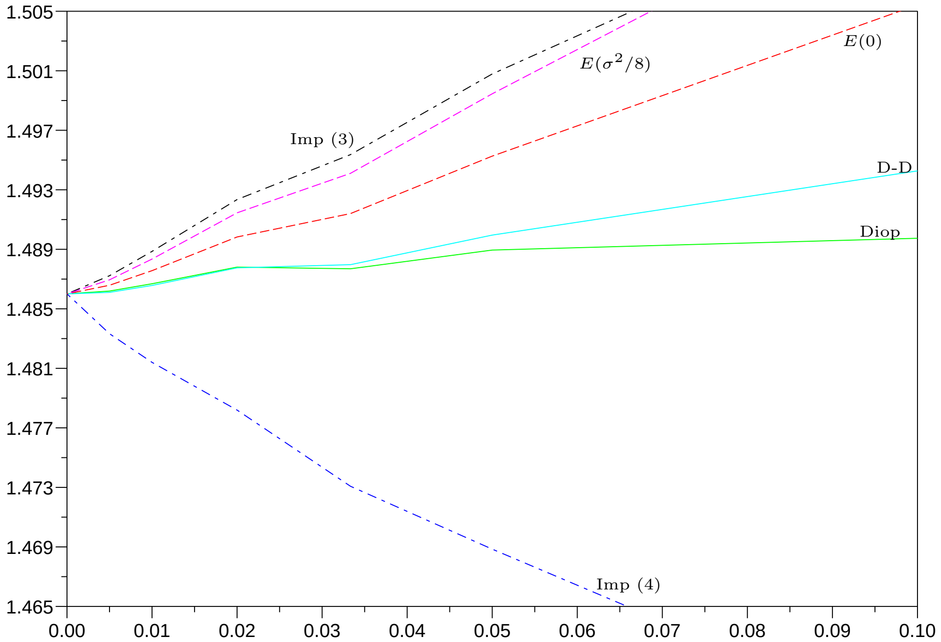

Figure 4: $\mathbb{E}(f(\hat{X}_1^{2n}))$ in function of $1/n$ with $f(x) = \frac{5+3x^4}{2+5x}$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = 1$.

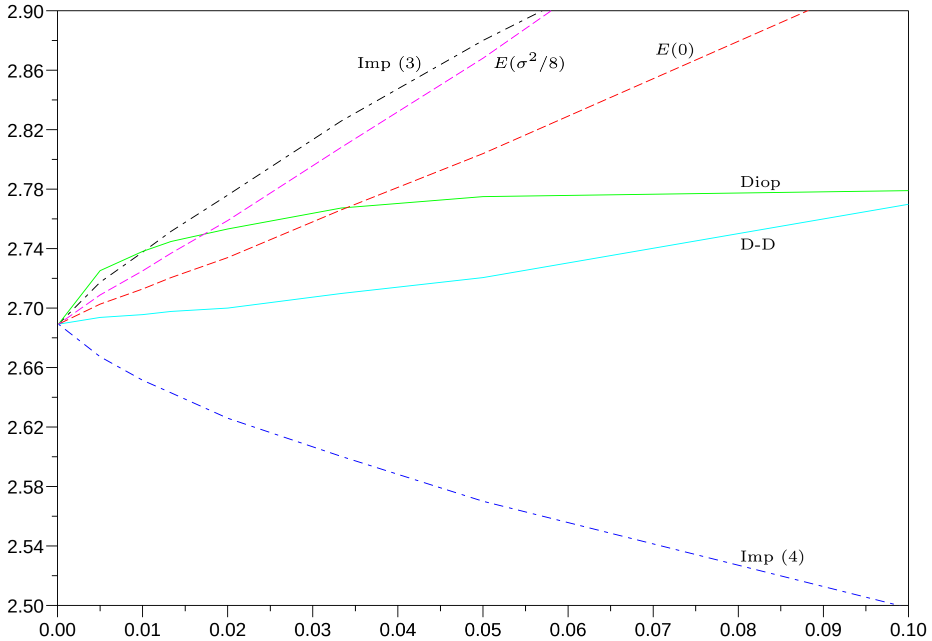

Figure 5: $\mathbb{E}(f(\hat{X}_1^{2n}))$ in function of $1/n$ with $f(x) = \frac{5+3x^4}{2+5x}$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

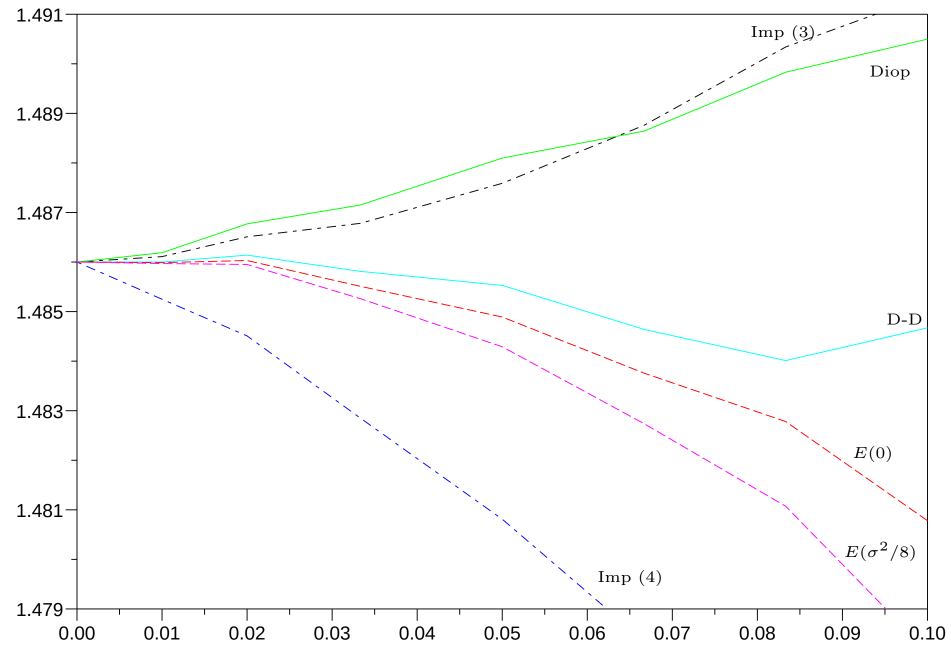

Figure 6: $2\mathbb{E}(f(\hat{X}_1^{2n})) - \mathbb{E}(f(\hat{X}_1^n))$ in function of $1/n$ with $f(x) = (5+3x^4)/(2+5x)$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = 1$.

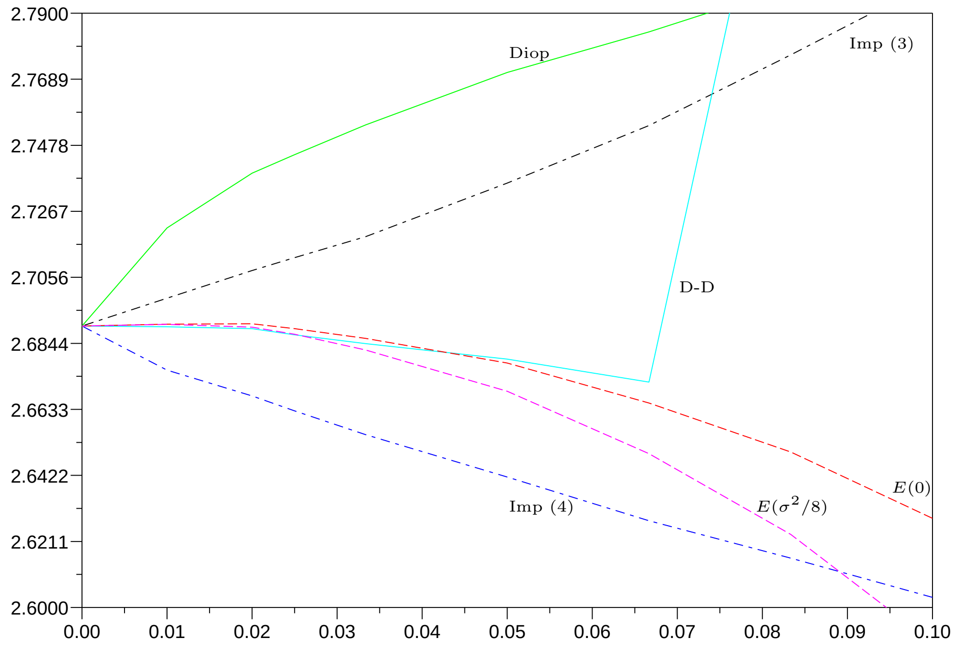

Figure 7: $2\mathbb{E}(f(\hat{X}_1^{2n})) - \mathbb{E}(f(\hat{X}_1^n))$ in function of $1/n$ with $f(x) = (5+3x^4)/(2+5x)$ for $x_0 = 0$, $k = 1$, $a = 1$ and $\sigma = \sqrt{3}$.

	Implicit (3), $\sigma^2 \le 2a$	Implicit (4), $\sigma^2 \le 4a$	Diop	Deelstra Delbaen	$E(0)$, $\sigma^2 \le 4a$	$E(\lambda)$, $0 \le \lambda$, $\lambda \le a - \sigma^2/4$
Nonnegativity	Y	Y	Y	N	Y	Y
Monotonicity	Y	Y	N	N	N	N
Strong CV	Y	?	Y*	Y	Y	Y
Weak CV rate in $1/n$	Y	?	Y*	?	Y	Y
Weak error expansion	?	?	?	?	Y	Y

Table 2: Theoretical results

6 Conclusion

We have sum up in Table 2 the theoretical results obtained in this paper and those of Diop, Deelstra and Delbaen [6, 5]. We first point out which scheme satisfy the algebraic properties of positivity and monotonicity. Then, we examine among the several schemes whether it has been proved

- a result of strong convergence,
- a weak convergence rate in $1/n$,
- an expansion of the weak error along the powers of $1/n$.

The star (Y*) means that the result has been established under some assumption on the parameters while the question mark indicates that no result has been shown yet. Let us mention here that Diop in [6] has also obtained a strong convergence speed in $1/\sqrt{n}$ under some restrictive conditions on parameters. Table 3 presents the results of the numerical tests of Section 5.

All these results tend to show that the explicit scheme $E(0)$ is the one that gathers the most interesting properties. Moreover, it is really easy to implement and is not more time consuming than the other schemes. That is why in the general case, it is recommended to use this scheme, at least for $\sigma^2 \le 4a$.

As a further work, it would be interesting to get an accurate mathematical study on the dependence of the strong convergence of $E(0)$ on $\frac{\sigma^2}{2a}$ (see Fig. 3). It would be also interesting to study the behaviour of the convergence of the various schemes for large values of σ, ($\sigma^2 \ge 4a$). Since none of the scheme studied in this paper seems to be efficient for these large values of σ, designing a relevant scheme appears to be an interesting challenge. Lastly, in a different direction, it would be nice to relax the condition of regularity on f for the weak error and prove estimates on the cumulated distribution function and the density of X_T, as in Bally and Talay [1, 2] or more recently Guyon [8].

Acknowledgement. I am grateful to Benjamin Jourdain (ENPC-CERMICS) for his numerous and helpful comments. I also thank Chalinène Bassinah (Paris 13-Institut Galilée) for having double checked some numerical results.

		Implicit (3)	Implicit (4)	Diop	Deelstra Delbaen	$E(0)$	$E(\lambda)$, $0 < \lambda$, $\lambda \leq a - \sigma^2/4$
$\sigma^2 \in [0, 2a]$	Strong CV order	$\approx 1/2$	≈ 1	$\approx 1/2$	$\approx 1/2$	≈ 1	$\approx 1/2$
	Weak CV rate in $1/n$	Y	Y	Y	Y	Y	Y
	Romberg in $1/n^2$	N	N	N	Y	Y	Y
$\sigma^2 \in [2a, 4a]$	Strong CV order	$\approx 1/2$	$\gtrapprox 1/2$	$\approx 1/2$	$\approx 1/2$	$\gtrapprox 1/2$	$\approx 1/2$
	Weak CV rate in $1/n$	?	?	N	Y	Y	Y
	Romberg in $1/n^2$	N	N	N	Y?	Y	Y

Table 3: Numerical results

A Proof of the Proposition 4.1

We will focus for sake of simplicity on the case of one function and one time T before explaining how to extend the results to the case of a family of functions that satisfy (24).

We will first prove

$$\max_{0\leq l\leq q} |\partial_x^l u(t,x)| \leq C(1+x^{q+m}) \tag{32}$$

for some constant $C > 0$, and then (23), so that (22) will outcome automatically by an induction on l', using that for $l' \geq 1$ such that $l + 2l' \leq q$,

$$\begin{aligned} \partial_x^l \partial_t^{l'} u(t,x) &= -\partial_x^l \left((a-kx)\partial_x \partial_t^{l'-1} u(t,x) + \frac{\sigma^2}{2} x \partial_x^2 \partial_t^{l'-1} u(t,x) \right) \\ &= -\frac{\sigma^2}{2} x \partial_x^{l+2} \partial_t^{l'-1} u(t,x) - (l\frac{\sigma^2}{2} + a - kx)\partial_x^{l+1}\partial_t^{l'-1} u(t,x) + lk\partial_x^l \partial_t^{l'-1} u(t,x). \end{aligned}$$

Let us set $\tilde{u}(t,x) = u(T-t,x) = \mathbb{E}(f(X_t^x))$. By Lemma 2.1, (32) holds for $f(x) = x^p$ ($p \in \mathbb{N}$) and therefore for any polynomial. Now, using the decomposition $f(x) = f(x) - P(x) + P(x)$ with $P(x) = \sum_{l=0}^q f^{(l)}(0)x^l/l!$, we deduce that it is enough to prove (32) for $f \in \mathcal{C}^q$ such that $|f(x)| \leq A(1+x^m)$ and $f^{(l)}(0) = 0$ for $l \leq q$.

Integrating successively, we get easily that $|f^{(l)}(x)| \leq A(1+x^{m+q-l})$ and so, $\forall\, l \leq q$, $|f^{(l)}(x)| \leq A(1+x^{m+q})$. The density of X_t^x is known and is given by:

$$p(t,x,z) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!} \frac{c_t/2}{\Gamma(i+v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2}$$

where $c_t = \frac{4k}{\sigma^2(1-e^{-kt})}$, $v = 4a/\sigma^2$ and $\lambda_t = c_t e^{-kt}$. Let us remark here that

$$c_t \geq c_{\min} := \begin{cases} \frac{4k}{\sigma^2}, & k > 0 \\ \frac{4}{\sigma^2 T}, & k = 0 \\ \frac{4|k|}{\sigma^2(e^{|k|T}-1)}, & k < 0. \end{cases}$$

We have for $t > 0$:

$$\tilde{u}(t,x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!} I_i(f, c_t)$$

where

$$I_i(f, c_t) = \int_0^\infty f(z) \frac{c_t/2}{\Gamma(i+v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2} dz.$$

Since for $l \leq q$, $|f^{(l)}(z)| \leq A(1+z^{m+q})$, we have

$$\forall i \in \mathbb{N},\ |I_i(f^{(l)}, c_t)| \leq A \left(1 + \left(\frac{2}{c_t}\right)^{m+q} \frac{\Gamma(i+m+q+v/2)}{\Gamma(i+v/2)}\right). \tag{33}$$

Taking $l = 0$, the convergence of the above series is ensured. Derivating successively in x, we get that for $l \leq q$,

$$\forall t \in (0,T], x \in \mathbb{R}^+,\ \partial_x^l \tilde{u}(t,x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!} \Delta_t^l (I_i(f, c_t)) \tag{34}$$

where $\Delta_t : \mathbb{R}^\mathbb{N} \to \mathbb{R}^\mathbb{N}$ is an the operator defined on sequences $(I_i)_{i\geq 0} \in \mathbb{R}^\mathbb{N}$ by $\Delta_t(I_i) = \frac{\lambda_t}{2}(I_{i+1} - I_i) = \frac{e^{-kt}}{2} c_t (I_{i+1} - I_i)$. Let us remark now that, since $f^{(l-1)}(0) = 0$, an integration by part gives for $0 < l \leq q$ and $i \geq 1$

$$\begin{aligned} I_i(f^{(l)}, c_t) &= \int_0^\infty f^{(l-1)}(z) \frac{(c_t/2)^2}{\Gamma(i+v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2} dz \\ &\quad - \int_0^\infty f^{(l-1)}(z) \frac{(c_t/2)^2 (i-1+v/2)}{\Gamma(i+v/2)} \left(\frac{c_t z}{2}\right)^{i-2+v/2} e^{-c_t z/2} dz \\ &= \frac{c_t}{2}(I_i(f^{(l-1)}, c_t) - I_{i-1}(f^{(l-1)}, c_t)). \end{aligned}$$

Therefore, we get that $\Delta_t(I_i(f, c_t)) = e^{-kt} I_{i+1}(f^{(1)}, c_t)$ and finally:

$$\forall t \in (0,T], x \in \mathbb{R}^+,\ \partial_x^l \tilde{u}(t,x) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!} I_{i+l}(f^{(l)}, c_t) e^{-klt}.$$

Using (33), it gives immediately that $|\partial_x^l \tilde{u}(t,x)| \leq A \left(1 + \frac{2^{m+q}}{c_t^{m+q}} \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!} \frac{\Gamma(i+l+m+q+v/2)}{\Gamma(i+l+v/2)}\right)$. The quotient $\frac{\Gamma(i+l+m+q+v/2)}{\Gamma(i+l+p+v/2)}$ is a polynomial of degree $m+q$ in i, and we note $\beta_0, \ldots, \beta_{m+q}$

its coefficients in the basis $\{1, i, i(i-1), \dots, i(i-1)\cdots(i-(m+q)+1)\}$. Thus, we get that $|\partial_x^l \tilde{u}(t,x)| \leq A + \frac{A2^{m+q}}{c_t^{m+q}}(\beta_0 + \beta_1\lambda_t x + \cdots + \beta_{m+q}(\lambda_t x)^{m+q})$ and since $|\lambda_t| \leq c_t e^{|k|T}$,

$$|\partial_x^l \tilde{u}(t,x)| \leq A + Ae^{(m+q)|k|T}\left(|\beta_0|/c_{\min}^{m+q} + |\beta_1|/c_{\min}^{m+q-1}x + \cdots + |\beta_{m+q}|x^{m+q}\right).$$

This allows us to conclude that there is a constant $C > 0$ (that depends only on A, T and the parameters (x_0, k, a, σ)) such that

$$\forall l \leq q,\ \forall t \in (0,T], x > 0,\ |\partial_x^l \tilde{u}(t,x)| \leq C(1 + x^{m+q}).$$

Proof of (23). We deduce from Lemma 2.1 that $\tilde{u}_0(T-t,x)$ and $\tilde{u}_1(T-t,x)$ solve the PDE (23) and it is therefore sufficient to prove the result for functions $f \in \mathcal{C}^2$ that satisfy $|f(x)| \leq A(x^2 + x^m)$. Let us now observe that $\frac{dc_t}{dt} = -\sigma^2 c_t\lambda_t/4$ and $\frac{d\lambda_t}{dt} = -(\sigma^2\lambda_t/4 + k)\lambda_t$. Then, it is no hard to get $\frac{dI_i(f,c_t)}{dt} = (\sigma^2 i/2 + a)\Delta_t(I_i(f,c_t))$ and that for any bounded sequence I_i, $\frac{d}{dt}\sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}I_i = -(\frac{\sigma^2\lambda_t}{4} + k)x\sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}\Delta_t(I_i)$. Combining these results, we get using relation (34):

$$\begin{aligned}
\partial_t \tilde{u}(t,x) &= -(\frac{\sigma^2\lambda_t}{4} + k)x\sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}\Delta_t(I_i(f,c_t)) \\
&\quad + \sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}(\frac{\sigma^2 i}{2} + a)\Delta_t(I_i(f,c_t)) \\
&= (a - kx)\partial_x\tilde{u}(t,x) + \frac{\sigma^2}{2}\sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}\left(-\frac{\lambda_t x}{2} + i\right)\Delta_t(I_i(f,c_t)) \\
&= (a - kx)\partial_x\tilde{u}(t,x) + \frac{\sigma^2}{2}x\sum_{i=0}^{\infty}\frac{e^{-\lambda_t x/2}(\lambda_t x/2)^i}{i!}\frac{\lambda_t}{2}\left(\Delta_t(I_{i+1}(f,c_t)) - \Delta_t(I_i(f,c_t))\right) \\
&= (a - kx)\partial_x\tilde{u}(t,x) + \frac{\sigma^2}{2}x\partial_x^2\tilde{u}(t,x).
\end{aligned}$$

Finally, the continuity of f ensures that $\tilde{u}(t,x) = \mathbb{E}(f(X_t^x)) \to f(x)$ when $t \to 0$ thanks to Lebesgue's theorem.

Let us explain now how to extend the result to a family of functions f_θ and get (25). Let us denote $P_\theta(x) = \sum_{l=0}^{q}\frac{1}{l!}f_\theta^{(l)}(0)x^q$. Condition (24) ensures that the coefficients of P_θ are uniformly bounded in θ. Writing $f_\theta(x) = P_\theta(x) + (f_\theta(x) - P_\theta(x))$, one obtains (25) in the same way as (22). □

References

[1] 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.*

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

[3] Brigo, D. and Alfonsi, A. (2005). Credit default swap calibration and derivatives pricing with the SSRD stochastic intensity model. *Finance and Stochastics, Vol. 9, No. 1, pp 29-42.*

[4] 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.*

[5] 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.*

[6] 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)*

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

[8] Guyon, J. (2005). Euler scheme and tempered distributions, *preprint CERMICS No. 277.*

[9] 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.*

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

[11] Lamberton, D. and Lapeyre, B. (1992). Une introduction au calcul stochastique appliqué à la finance, *Ellipses.* English version (1995): An Introduction to Stochastic Calculus Applied to Finance, *Chapman and Hall.*

[12] Milstein, G. N., Repin, Y. M., and Tretyakov, M. V. (2002). Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal., Vol. 40, No. 4 pp. 1583-1604.*

[13] Rogers, L.C.G. and Williams, D. (2000). Diffusions, Markov Processes and Martingales, 2nd edition. *Cambridge Mathematical Library.*