HIGH ORDER APPROXIMATIONS OF THE COX-INGERSOLL-ROSS PROCESS SEMIGROUP USING RANDOM GRIDS

AURÉLIEN ALFONSI AND EDOARDO LOMBARDO

Abstract. We present new high order approximations schemes for the Cox-Ingersoll-Ross (CIR) process that are obtained by using a recent technique developed by Alfonsi and Bally (2021) for the approximation of semigroups. The idea consists in using a suitable combination of discretization schemes calculated on different random grids to increase the order of convergence. This technique coupled with the second order scheme proposed by Alfonsi (2010) for the CIR leads to weak approximations of order $2k$, for all $k \in \mathbb{N}^*$. Despite the singularity of the square-root volatility coefficient, we show rigorously this order of convergence under some restrictions on the volatility parameters. We illustrate numerically the convergence of these approximations for the CIR process and for the Heston stochastic volatility model and show the computational time gain they give.

1. Introduction

The present paper develops approximations, of any order, of the semigroup $P_t f(x) := \mathbb{E}[f(X^x_t)]$ associated to the following Stochastic Differential Equation (SDE) known as the Cox-Ingersoll-Ross (CIR) process

$$X^x_t = x + \int_0^t (a - kX^x_s)ds + \int_0^t \sigma \sqrt{X^x_s}dW_s, \quad t \geq 0,$$

where $W$ is a Brownian motion, $x, a \geq 0$, $k \in \mathbb{R}$ and $\sigma > 0$. Let us recall that the process (1.1) is nonnegative and the semigroup $(P_t)_{t \geq 0}$ is well defined on the space of functions $f : \mathbb{R} \to \mathbb{R}$ with polynomial growth. The diffusion (1.1) is widely used in financial mathematics, in particular because of its simple parametrisation and the affine property that enables to use numerical methods based on Fourier techniques. We mention here the Cox-Ingersoll-Ross model [8] for the short interest rate and the Heston stochastic volatility model [9], that have been followed by many other ones. Developing efficient numerical methods for the process (1.1) is thus of practical importance.

To deal with the approximation of SDE’s semigroups, a common approach is to consider stochastic approximations and the most standard one is the Euler-Maruyama scheme. The error between the approximated semigroup and the exact one is called the weak error, as opposed to the strong error that quantifies the error "omega by omega" on the probability...
space. The seminal work of Talay and Tubaro [15] shows, under regularity assumptions on the SDE coefficients, that the weak error given by the Euler-Maruyama scheme is of order one, i.e. is proportional to the time step. They also obtain an error expansion that enables to use Richardson-Romberg extrapolations as developed by Pagès [14]. Higher order schemes for SDEs and related extrapolations have been proposed by Kusuoka [10], Ninomiya and Victoir [12], Ninomiya and Ninomiya [11] and Oshima et al. [13] to mention a few. Recently, Alfonsi and Bally [4] have given a method to construct weak approximation of general semigroups of any order by using random time grids.

These general results on weak approximation of SDEs do not apply to the CIR (1.1) process. This is due to the diffusion coefficient, namely the singularity of the square-root at the origin. Besides this, classical schemes such as the Euler-Maruyama scheme are not well-defined for (1.1), and one has to work with dedicated schemes. Under some restrictions on the parameters, the weak convergence of order one for some discretization schemes of the CIR process has been obtained by Alfonsi [1], Bossy and Diop [6], and more recently by Briani et al. [7] who also study the weak convergence of a semigroup approximation for the Heston model. We also mention the earlier work by Altmayer and Neuenkirch [5] that precisely studies the weak error for the Heston model. Adapting ideas from Ninomiya and Victoir [12] who developed a second order scheme for general SDEs, Alfonsi [2] has introduced second order and third order schemes for the CIR and proved their weak order of convergence, without any restriction on the parameters.

The goal of the present paper is to boost the second order scheme developed in [2] and get approximations of any order. To do so, we rely on the method developed recently by Alfonsi and Bally [4] to construct approximation of semigroups of any order. Roughly speaking, this method allows to get, from an elementary weak approximation scheme of order $\alpha > 0$, approximation schemes of any order by computing the elementary scheme on appropriate random grids. The method is illustrated in [4] on the case of the Euler-Maruyama scheme for SDEs, under regularity assumptions on the coefficients that do not hold for the CIR process (1.1). This method is presented briefly in Section 2. It relies on an appropriate choice of a function space endowed with a family of seminorms. Section 3 then presents the second order scheme that is used as an elementary scheme to get higher order approximation. It states in Theorem 3.2 the main result of this paper: we prove, when $\sigma^2 \leq 4a$, that we get weak approximations of any orders for smooth test functions $f$ with derivatives having at most a polynomial growth. Section 4 illustrates the boosting method when considering the space of polynomials function with their usual norm. In this simple case, proofs are quite elementary so that the method can be followed easily. Section 5 is more involved: it first defines the appropriate family of seminorms on the space of smooth functions with derivative of polynomial growth and then proves Theorem 3.2. Last, we illustrate in Section 6 the convergence of the high order approximations for different parameters. It validates our theoretical results and shows important computational gains given by the new approximations. We also test the method on the Heston model and obtain similar convincing results.

2. HIGH ORDER SCHEMES WITH RANDOM GRIDS: THE METHOD IN A NUTSHELL

In this paragraph, we recall briefly the method developed by Alfonsi and Bally in [4] to construct approximations of any order from a family of approximation schemes. We consider $F$ a vector space endowed with a family of seminorms $(|||f|||_k)_{k\in\mathbb{N}}$ such that $||f||_k \leq ||f||_{k+1}$. We
We get by using (2.1) for the weak error of the Euler scheme. From the semigroup property, to achieve this goal, we consider a family of linear operators \((Q_l)_{l \in \mathbb{N}}\) on \(F\). For \(l \in \mathbb{N}\), we note \(Q_l[0] = I\) the identity operator and, for \(j \in \mathbb{N}\), \(Q_l[j] = Q_l^{j-1}Q_l\) the operator obtained by composition. We suppose that the two following conditions are satisfied. The first quantifies how \(Q_l\) approximates \(P_{h_1}\):

\[
\text{there exists } \alpha > 0 \text{ such that for any } l, k \in \mathbb{N}, \text{ there exists } C > 0, \text{ such that }
\| (P_{h_1} - Q_l)f \|_k \leq C\| f \|_{\psi_Q(k)h_1^{1+\alpha}} \text{ for all } f \in F, \tag{\(H_1\)}
\]

where \(\psi_Q : \mathbb{N} \to \mathbb{N}\) is a function. \(^1\) The second one is a uniform bound with respect to all the seminorms:

\[
\max_{0 \leq j \leq n} \|Q_l[j]f\|_k + \sup_{t \leq T} \|P_tf\|_k \leq C\|f\|_k \text{ for all } f \in F. \tag{\(H_2\)}
\]

Then, for any \(\nu \in \mathbb{N}\), Alfonsi and Bally \([4]\) show how one can construct, by mixing the operators \(Q_l\), a linear operator \(\hat{P}_{\nu,n}\) for which there exists \(C > 0\) and \(k \in \mathbb{N}\) such that

\[
\| P_T f - \hat{P}_{\nu,n} f \|_0 \leq C\|f\|_{k n^{-\nu} \alpha} \text{ for all } f \in F. \tag{\(2.2\)}
\]

Let us explain how it works for \(\nu = 1\) and \(\nu = 2\). For \(\nu = 1\), we mainly repeat the proof of Talay and Tubaro \([15]\) for the weak error of the Euler scheme. From the semigroup property, we have

\[
P_T f - Q_1[n] f = P_{n h_1} f - Q_1[n] f = \sum_{k=0}^{n-1} P_{(n-(k+1)) h_1} [P_{h_1} - Q_1] Q_1[k] f. \tag{\(2.3\)}
\]

We get by using (\(H_2\)), then (\(H_1\)) and then again (\(H_2\))

\[
\| P_T f - Q_1[n] f \|_0 \leq \sum_{k=0}^{n-1} C\|P_{h_1} - Q_1\|Q_1[k] f\|_0 \leq \sum_{k=0}^{n-1} C\|Q_1[k] f\|_{\psi_Q(0)h_1^{1+\alpha}} \leq C\| f \|_{\psi_Q(0)n(T/n)^{1+\alpha}} = C\| f \|_{\psi_Q(0)n^{-\alpha}}. \tag{\(2.4\)}
\]

Here, and through the paper, \(C\) denotes a positive constant that may change from one line to another. So, \(\hat{P}_{1,n} = Q_1[n]\) satisfies (\(2.2\)) with \(\nu = 1, k = \psi_Q(0)\). The approximation scheme simply consists in using \(n\) times the scheme \(Q_1\), which can be seen as a scheme on the regular time grid with time step \(h_1\).

We now present the approximation scheme (\(2.2\)) for \(\nu = 2\). To do so, we use again (\(2.3\)) to get

\[
P_{(n-(k+1)) h_1} - Q_1^{n-(k+1)} = \sum_{k'=0}^{n-(k+2)} P_{(n-(k+k'+2)) h_1} [P_{h_1} - Q_1] Q_1[k'] \quad \text{and then expand further (\(2.3\)):
}

\[
P_T f - Q_1[n] f = \sum_{k=0}^{n-1} Q_1^{n-(k+1)} [P_{h_1} - Q_1] Q_1[k] f + R_2^{h_1}(n) f, \tag{\(2.5\)}
\]

with \(R_2^{h_1}(n) = \sum_{k=0}^{n-1} \sum_{k'=0}^{n-(k+2)} P_{(n-(k+k'+2)) h_1} [P_{h_1} - Q_1] Q_1[k'] [P_{h_1} - Q_1] Q_1[k]

\(^1\)Note that in [4], it is taken \(\psi_Q(k) = k + \beta\) for some \(\beta \in \mathbb{N}\), but is can be easily generalized to any function \(\psi_Q\). In this paper, we will work with a doubly indexed norm and take \(\psi_Q(m, L) = (2(m + 3), L - 1)\).
We have
\[ \| P_{h_1}^n f \|_0 \leq C \| f \|_{\psi Q(0)} \frac{n(n - 1)}{2} h_1^{2(1+\alpha)} \leq C \| f \|_{\psi Q(0)} \frac{T^{2(1+\alpha)}}{2} n^{-2\alpha}. \]

Thus, \( Q_1^n + \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [P_{h_1} - Q_1]Q_1^n f \) is an approximation of order \( 2\alpha \), but it still involves the semigroup through \( P_{h_1} \). To get an approximation that is obtained only with the operators \( Q_1 \), we use again (2.3) with time step \( h_2 \) and final time \( h_1 = nh_2 \):
\[
P_{h_1} f - Q_2^n f = \sum_{k=0}^{n-1} P_{(n-(k+1))h_2} [P_{h_2} - Q_2]Q_2^n f.
\]

We have \( \| P_{h_1} f - Q_2^n f \|_0 \leq C \| f \|_{\psi Q(0)} n h_2^{1+\alpha} \) by using again (\( \Pi_1 \)) and (\( \Pi_2 \)). We get from (2.5)
\[
P_T f - Q_1^n f = \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n f + \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [P_{h_1} - Q_2^n]Q_1^n f + R_{h_1}^n(n)f, \quad (2.6)
\]
with \( \| \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [P_{h_1} - Q_2^n]Q_1^n f \|_0 \leq C \| f \|_{\psi Q(0)} n^2 h_2^{1+\alpha} = C \| f \|_{\psi Q(0)} T^{1+\alpha} n^{-2\alpha}. \) Therefore, the approximation
\[
\hat{P}^n f := Q_1^n f + \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n f \quad (2.7)
\]
satisfies (2.2) with \( \nu = 2 \) and is obtained only with the approximating operators \( Q_1 \). The first term \( Q_1^n \) corresponds to apply the scheme \( Q_1 \) on the regular time grid with time step \( h_1 \), while each term \( Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n \) is the difference between this scheme and the one where \( Q_2^n \) is used instead of \( Q_1 \) for the \((k+1)\)-th time step. This amounts to refine this time step and split it into \( n \) time steps of size \( h_2 \), and to use the scheme \( Q_2 \) on this time grid.

In practice, it is inefficient to calculate one by one the terms in \( \hat{P}^n f \). In fact, each term requires a number of calculations that is proportional to \( n \), and the overall computation cost would be of the same order as \( n^2 \). Since the convergence is in \( O(n^{-2\alpha}) \) it would not be better asymptotically than using \( \hat{P}^n f \). To avoid this, we use randomization. We sample a uniform random variable \( \kappa \) on \( \{0, \ldots, n - 1\} \) and calculate \( n E[Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n] f = \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n f \). This amounts to consider the regular time grid with time step \( h_1 \), to select randomly one time step and to refine it, and then to compute the difference between the approximations on the (random) refined time-grid and on the regular time-grid. To be more precise, let us consider the case of an approximation scheme defined by \( \varphi(x, h, V) \) where \( \varphi \) is a measurable function, \( x \) is the starting point, \( h \) the time step and \( V \) a random variable. The associated operators are \( Q_1 f(x) = E[\varphi(x, h, V)], \ l \in N \). For a time-grid \( \Pi = \{0 = t_0 < \cdots < t_n = T\} \), we define \( X_0^\Pi(x) = x \) and \( X_1^\Pi(x) = \varphi(X_1^\Pi(x), t_1 - t_0, V_i) \) for \( 1 \leq i \leq n \), where \( (V_i)_{i \geq 1} \) is an i.i.d. sequence. Thus, we get on the uniform time grid
\[ \Pi^0 = \{ kT/n, 0 \leq k \leq n \} \ E[f(X_1^\Pi(x))] = Q_1^n f(x). \]
By taking the random grid \( \Pi^1 = \Pi^0 \cup \{ \kappa T/n + k'T/n^2, 1 \leq k' \leq n - 1 \} \), we also get \( E[f(X_2^\Pi(x))] = E[Q_2^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n f] \), and then \( E[n(f(X_2^\Pi(x)) - f(X_1^\Pi(x)))] = \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [Q_2^n - Q_1]Q_1^n f(x) \). When using a Monte-Carlo estimator of this identity, one has thus to draw as many \( \kappa \)'s as trajectories.
We have presented here how to construct \( \hat{P}^{\nu,n} \) for \( \nu = 1 \) and \( \nu = 2 \), and it is possible by repeating the same arguments to construct by induction approximations of any order. Unfortunately, the induction is quite involved. It is fully described in [4, Theorem 3.10]. We do not reproduce it in this paper because it would require much more notation, and we will mainly use the scheme (2.7). Here, we give in addition the explicit form of \( \hat{P}^{3,n} \), \( n \geq 2 \):

\[
\hat{P}^{3,n} f := \hat{P}^{2,n} + \sum_{0 \leq k_1 < k_2 < n} Q_1^{n-(k_2+1)}[Q_2^{[n]} - Q_1]Q_1^{[k_1]}[Q_2^{[n]} - Q_1]Q_1^{[k_1]} f
\]

\[
+ \sum_{k=0}^{n-1} Q_1^{n-(k+1)} \left[ \sum_{k'=0}^{n-1} Q_2^{[n-(k'+1)]}[Q_3^{[n]} - Q_2]Q_2^{[k']} \right] Q_1^{[k]} f.
\]

By similar arguments, it satisfies (2.2) with \( \nu = 3 \).

3. SECOND ORDER SCHEMES FOR THE CIR PROCESS AND MAIN RESULT

In this section, we focus on the approximation of the semigroup of the CIR process \( P_t f(x) = E[f(X_t^x)] \), where

\[
X_t^x = x + \int_0^t (a - kX_s^x) ds + \sigma \int_0^t \sqrt{X_s^x} dW_s, \ t \geq 0.
\]

Equation (2.4) shows that, necessarily, approximating operators \( Q_t \) that satisfy both (\( \Pi_1 \)) and (\( \Pi_2 \)) lead to a weak error of order \( \alpha \). Therefore, we are naturally interested in approximation schemes of the CIR for which we know the rate of convergence \( \alpha \) for the weak error. [1, Proposition 4.2] gives a rate \( \alpha = 1 \) for a family of approximation schemes that are basically obtained as a correction of the Euler scheme. Ninomiya and Victoir [12] have developed a generic method to construct second order schemes (\( \alpha = 2 \)) for Stochastic Differential Equations with smooth coefficients. Applied to the Cox-Ingersoll-Ross process, their method leads to the following approximation scheme

\[
X_t^x = \varphi(x, t, \sqrt{t}N),
\]

where \( N \sim \mathcal{N}(0,1) \) and \( \varphi : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+ \) is defined by

\[
\varphi(x, t, w) = e^{-kt/2} \left( \sqrt{(a - \sigma^2/4)\psi_k(t/2) + e^{-kt/2}x + \sigma w/2} \right)^2 + (a - \sigma^2/4)\psi_k(t/2)
\]

\[
= X_0(t/2, X_1(w, X_0(t/2, x))), \quad \text{with}
\]

\[
X_0(t, x) = e^{-kt}x + \psi_k(t)(a - \sigma^2/4), \quad \psi_k(t) = \frac{1 - e^{-kt}}{k},
\]

\[
X_1(t, x) = (\sqrt{x + t\sigma^2/2})^2,
\]

with the convention that \( \psi_0(t) = t \). This scheme corresponds to approximate \( P_t f(x) \) by \( \hat{P}_t f(x) = E[f(X_t^x)] \) for \( x, t \geq 0 \), and then to set \( Q_t = \hat{P}_h \). Its construction comes from the splitting of the infinitesimal generator of the CIR process

\[
\mathcal{L} f(x) = (a - kx)f'(x) + \frac{1}{2} \sigma^2 x f''(x), \ f \in \mathcal{C}^2, x \geq 0,
\]
as \( \mathcal{L} = V_0 + \frac{1}{2} V_1^2 \) with

\[
V_0 f(x) = \left( a - \frac{\sigma^2}{4} - kx \right) f'(x) \quad \text{and} \quad V_1 f(x) = \sigma \sqrt{x} f'(x).
\] (3.6)

The function \( t \mapsto X_0(t,x) \) is the solution of the ODE \( X_0'(t,x) = a - \frac{\sigma^2}{4} - kX_0(t,x) \) such that \( X_0(0,x) = x \), while \( X_1(W_t,x) \) solves the SDE associated to the infinitesimal generator \( V_1^2/2 \).

The scheme (3.1) is well defined for \( \sigma^2 \leq 4a \). Instead, for \( \sigma^2 > 4a \), it is not well defined for any \( x \geq 0 \) since the argument in the square-root is negative when \( x \) is close to zero. To correct this, Alfonsi [2] has proposed the following scheme

\[
\hat{X}^x_t = (1_{x \geq K^Y_2(t)} \psi(x,t,\sqrt{t}Y) + 1_{x < K^Y_2(t)} \hat{X}^x_{t,d}),
\] (3.7)

where \( Y \) is a random variable with compact support on \([-A_Y, A_Y]\) for some \( A_Y > 0 \) such that \( \mathbb{E}[Y^k] = \mathbb{E}[N^k] \) for \( k \leq 5 \), and \( \hat{X}^x_{t,d} \) is a nonnegative random variable such that \( \mathbb{E}[(\hat{X}^x_{t,d})^i] = \mathbb{E}[(X^x_{t,d})^i] \) for \( i \in \{1, 2\} \) and \( K^Y_2(t) \) is a nonnegative threshold defined by

\[
K^Y_2(t) = 1_{\sigma^2 > 4a} \left[ e^{\frac{4t}{\sigma^2}} \left( \frac{\sigma^2}{4} - a \right) \psi_2(t/2) + \left( \sqrt{e^{\frac{4t}{\sigma^2}} \left( \frac{\sigma^2}{4} - a \right) \psi_2(t/2) + \sigma^2} \right)^2 \right].
\] (3.8)

Note that when \( \sigma^2 \leq 4a \), we have \( K^Y_2(t) = 0 \) and that \( \hat{X}^x_t = \varphi(x,t,\sqrt{t}Y) \). In [2], it is taken \( Y \) such that \( \mathbb{P}(Y = \sqrt{3}) = \mathbb{P}(Y = -\sqrt{3}) = 1/6 \) and \( \mathbb{P}(Y = 0) = 2/3 \), and a discrete random variable \( \hat{X}^x_{t,d} \) such that \( \mathbb{P}(\hat{X}^x_{t,d} = \frac{1}{2 \pi(t,x)}) = \pi(t,x), \mathbb{P}(\hat{X}^x_{t,d} = \frac{1}{2(1-\pi(t,x)}) = 1 - \pi(t,x) \) where \( \pi(t,x) = \frac{1 - \sqrt{1 - \mathbb{E}[X^x_t]^2 / \mathbb{E}[(X^x_t)^2]}}{2} \) is close to zero. To handle the analysis, in [2], it is taken \( \hat{X}^x_t = \varphi(x,t,\sqrt{t}Y) \).

We now restate [2, Theorem 2.8] that analyses the weak error. We introduce \( C^k_{\text{pol}}(\mathbb{R}_+) \), the set of \( C^k \) functions \( f : \mathbb{R} \rightarrow \mathbb{R}_+ \) such that all its derivatives have polynomial growth. More precisely, this means that for all \( k' \in \{0, \ldots, k\} \), there exists \( C_{k'}, E_{k'} \in \mathbb{R}_+ \) such that

\[
|f^{(k')}(x)| \leq C_{k'}(1 + x^{E_{k'}}), \quad x \geq 0.
\]

We also set \( C^\infty_{\text{pol}}(\mathbb{R}_+) = \bigcap_{k \in \mathbb{N}} C^k_{\text{pol}}(\mathbb{R}_+) \).

**Theorem 3.1.** Let \( \hat{X}^x_t \) be the scheme defined by (3.1) for \( \sigma^2 \leq 4a \) or by (3.7) for any \( \sigma > 0 \). Then, for all \( f \in C^\infty_{\text{pol}}(\mathbb{R}_+) \), we have \( Q^1_{t,x} f(x) - P_{tf}(x) = O(1/n^2) \) where \( Q_1 f(x) = \mathbb{E}[f(\hat{X}^x_{h_t})] \).

The goal of this paper is to extend this result and prove the estimates \( (\hat{M}_1) \) and \( (\hat{M}_2) \) for a suitable space of functions and a suitable family of seminorms. We are able to prove such results only in the case \( \sigma^2 \leq 4a \): the indicator function in (3.7) creates a singularity that is difficult to handle in the analysis. In Section 4, we first prove \( (\hat{M}_1) \) and \( (\hat{M}_2) \) for polynomial test functions. Then, we deal in Section 5 with the much technical case of smooth test functions with derivatives of polynomial growth. We state here our main result, the proof of which is given in Section 5.

**Theorem 3.2.** Let \( \hat{X}^x_t \) be the scheme defined by (3.1) for \( \sigma^2 \leq 4a \) and \( Q_l f(x) = \mathbb{E}[f(\hat{X}^x_{h_t})] \), for \( l \geq 1 \). Then, for all \( f \in C^\infty_{\text{pol}}(\mathbb{R}_+) \), we have \( \hat{P}^{2,n} f(x) - P_{tf}(x) = O(1/n^4) \) as \( n \rightarrow \infty \).

Besides, for \( f \in C^\infty_{\text{pol}}(\mathbb{R}_+) \), we have \( \hat{P}^{\nu,n} f(x) - P_{tf}(x) = O(1/n^{2\nu}) \).
Let us stress here that Theorem 3.2 gives an asymptotic result as \( n \to \infty \). It thus might happen that for small values of \( n \), \( \hat{P}^{2,n} \) is less accurate than \( \hat{P}^{1,n} = Q^{[n]}_1 \) for some \( f \in C^{\infty}_{\text{pol}}(\mathbb{R}_+) \) and \( x \geq 0 \). In practice, we have always noticed in our numerical experiments that \( \hat{P}^{2,n} \) is more accurate than \( \hat{P}^{1,n} \). However, the estimated rates of convergence obtained from relatively small values of \( n \) may be different from the theoretical asymptotic ones, see Figures 1, 2 and 3 where are given the estimated rates for \( \hat{P}^{1,n} \), \( \hat{P}^{2,n} \) and \( \hat{P}^{3,n} \).

4. The case of polynomial test functions

In this section, we want to illustrate the method and consider test functions that are polynomial test functions. We define for \( L \in \mathbb{N} \)

\[
\mathcal{P}_L(\mathbb{R}_+) = \{ f : \mathbb{R}_+ \to \mathbb{R}, f(x) = \sum_{j=0}^{L} a_j x^j \text{ for some } a_0, \ldots, a_L \in \mathbb{R} \},
\]

the vector space of polynomial functions over \( \mathbb{R}_+ \) with degree less or equal to \( L \). We also define \( \mathcal{P}(\mathbb{R}_+) = \cup_{L \in \mathbb{N}} \mathcal{P}_L(\mathbb{R}_+) \) the space of polynomial functions. We endow \( \mathcal{P}(\mathbb{R}_+) \) with the following norm:

\[
\|f\| = \sum_{j=0}^{L} |a_j|, \text{ for } f(x) = \sum_{j=0}^{L} a_j x^j. \tag{4.1}
\]

We consider the case \( \sigma^2 \leq 4a \) and consider the scheme (3.7) for the CIR process with a time step \( t > 0 \), \( \bar{X}_{t}^x = \varphi(x,t,\sqrt{Y}) \). The approximation scheme \( Q_t \) is then defined by \( Q_t f = E[f(\bar{X}_{h_i}^x)] \). The goal of this section is to prove (\( \mathcal{H}_1 \)) and (\( \mathcal{H}_2 \)) for the norm (4.1). We make the following assumption on \( Y \).

**Assumption (\( \mathcal{H}_Y \)):** \( Y : \Omega \to \mathbb{R} \) is a symmetric random variable such that \( E[|Y|^k] < \infty \) for all \( k \in \mathbb{N} \), and \( E[Y^k] = E[N^k] \) for \( k \in \{2, 4\} \) with \( N \sim \mathcal{N}(0,1) \).

We now state two lemmas that will enable us to prove that (\( \mathcal{H}_2 \)) is satisfied by the scheme (3.7). Lemma 4.1 shows that polynomials functions are preserved by the approximation scheme, and gives short time estimate for the polynomial norm. Lemma 4.2 gives similar results for the CIR diffusion. The proofs of these lemmas are quite elementary and are postponed to Appendix A.

**Lemma 4.1.** Let \( T \geq 0 \), \( t \in [0,T] \), \( f \in \mathcal{P}_L(\mathbb{R}_+) \) and assume (\( \mathcal{H}_Y \)) and \( \sigma^2 \leq 4a \). Then, we have \( f(X_0(t,\cdot)), E[f(X_1(\sqrt{Y},\cdot))] \in \mathcal{P}_L(\mathbb{R}_+) \) where \( X_0 \) and \( X_1 \) are defined by (3.3) and (3.4), and

1. \( \|f(X_0(t,\cdot))\| \leq (1 + e^{-kL_1}t)\|f\|, \)
2. \( \|E[f(X_1(\sqrt{Y},\cdot))]\| \leq (1 + E[Y^{2L_1}]C_{X_1}^L t)\|f\|, \)

for some constants \( C_{X_0}, C_{X_1} \) depending only on \( (a,\sigma,T) \).

**Lemma 4.2.** Let \( (X_t^x, t \geq 0) \) be the CIR process starting from \( x \in \mathbb{R}_+ \). For \( m \in \mathbb{N} \), we define \( \tilde{u}_m(t,x) := E[(X_t^x)^m] \). There exists \( C^{\infty} \) functions \( \tilde{u}_{j,m} : \mathbb{R}_+ \to \mathbb{R} \) that depend on \( (k, a, \sigma) \) such that:

\[
\tilde{u}_m(t,x) = \sum_{j=0}^{m} \tilde{u}_{j,m}(t)x^j. \tag{4.2}
\]
If \( f \in \mathfrak{P}_L(\mathbb{R}_+) \), then we have \( \mathbb{E}[f(X_t)] \in \mathfrak{P}_L(\mathbb{R}_+) \) and for \( t \in [0,T] \),
\[
\|\mathbb{E}[f(X_t)]\| \leq C_{\text{civ}}(L,T)\|f\|,
\] (4.3)
with \( C_{\text{civ}}(L,T) = \max_{t \in [0,T],m \in \{0, \ldots, L\}} \sum_{j=0}^{m} |u_{j,m}(t)|. \)

We are now in position to prove the main result of this section, which is a weaker (but easier to prove) version of our main Theorem 3.2, since it only applies to polynomial test functions. Let us point however that it applies to a larger family of schemes, namely to the schemes \( \varphi(x,t,\sqrt{t}Y) \) with \( Y \) satisfying \( (\mathcal{H}_Y) \), while Theorem 3.2 requires to take \( Y \sim \mathcal{N}(0,1) \).

**Proposition 4.3.** Let \( a^2 \leq 4a \) and assume that \( Y \) satisfies \((\mathcal{H}_Y)\). For any \( L \in \mathbb{N} \), the properties \((\mathcal{H}_1)\) and \((\mathcal{H}_2)\) are satisfied by the scheme (3.7) \( \hat{X}_t^x = \varphi(x,t,\sqrt{t}Y) \) for \( F = \mathfrak{P}_L(\mathbb{R}_+) \) and the norm (4.1). Then, we have for any \( f \in \mathfrak{P}_L(\mathbb{R}_+) \),
\[
\|\mathbb{E}[f(\hat{X}_t^x)] - \hat{P}^{\nu,n}f\| \leq C_L\|f\|n^{-2\nu},
\]
for some constant \( C_L \).

**Proof.** We first prove \((\mathcal{H}_2)\). The property \( \sup_{t \in [0,T]} \|P_tf\| \) is given by Lemma 4.2. Since \( \hat{X}_t^x = X_0(t/2, X_1(\sqrt{t}Y, X_0(t/2, x))) \), we get by Lemma 4.1
\[
\|\mathbb{E}[f(\hat{X}_t^x)]\| \leq (1 + e^{-kLt/2})(1 + C_{X_0}^L t/2)^2(1 + \mathbb{E}[Y^{2L}]C_{\hat{X}_t^x}^L t)\|f\|.
\]
We now use that \( 1 + x \leq e^x \) to get
\[
\|\mathbb{E}[f(\hat{X}_t^x)]\| \leq e^{(-k)^{-L} + C_{X_0}^L + \mathbb{E}[Y^{2L}]C_{\hat{X}_t^x}^L t}\|f\|. \quad (4.4)
\]
Since \( Q_t f(x) = \mathbb{E}[f(\hat{X}_{T/n}^x)] \), this yields to \( \max_{0 \leq j \leq n^l} \|Q_T^j f\| \leq e^{(-k)^{-L} + C_{X_0}^L + \mathbb{E}[Y^{2L}]C_{\hat{X}_t^x}^L T} \|f\|. \)

We now prove \((\mathcal{H}_1)\). Let \( m \in \{0, \ldots, L\} \) and \( 0 < x_0 < \cdots < x_L \) be fixed real numbers (one may take for example \( x_j = j + 1 \)). Lemmas 4.1 and 4.2 give that \( v_m(t,x) = \mathbb{E}[(\hat{X}_t^x)^m] - \mathbb{E}(X_t^x)^m = \sum_{j=0}^{m} v_{j,m}(t)x^j \). By [2, Proposition 2.4], we know that there exists \( C_m^v, E_m^v \) such that for all \( t \in (0,1) \), \( |v_m(t,x)| \leq C_m^v t^3(1 + |x|E_m^v) \). Therefore, there exists \( C_m^v \in \mathbb{R}_+ \) such that for all \( \ell \in \{0, \ldots, L\} \), \( |v_m(t,x_\ell)| \leq \hat{C}_m^v t^3 \). By using the invertibility of the Vandermonde matrix, we get the existence of \( C_m \in \mathbb{R}_+ \) such that
\[
|v_{j,m}(t)| \leq C_m t^3, \quad j \in \{0, \ldots, m\}.
\]
Therefore, we get for \( f \in \mathfrak{P}_L(\mathbb{R}_+) \)
\[
\|\mathbb{E}[f(\hat{X}_t^x)] - \mathbb{E}[f(X_t^x)]\| \leq \sum_{m=0}^{L} |a_m| \sum_{j=0}^{m} C_m t^3 \leq L \max_{m \in \{0, \ldots, L\}} C_m \|f\|t^3,
\]
that gives \((\mathcal{H}_1)\). We conclude by applying [4, Theorem 3.10]. \( \square \)

5. PROOF OF THEOREM 3.2

In Section 4, we have obtained the convergence for test functions that are polynomial functions. For these test functions, the choice of the norm is straightforward and the proofs are not very technical and quite easy. However, one would like to obtain the convergence result for a much larger class of test functions. This is the goal of this section.
We consider test functions that are smooth with polynomial growth, whose derivatives have a polynomial growth. Namely, we introduce for $m, L \in \mathbb{N}$,

$$
C_{\text{pol}}^{m,L}(\mathbb{R}+) = \left\{ f : \mathbb{R}+ \to \mathbb{R} \text{ of class } C^m : \max_{j \in \{0, \ldots, m\}} \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^L} < \infty \right\},
$$

which we endow with the norm

$$
\|f\|_{m,L} = \max_{j \in \{0, \ldots, m\}} \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^L}.
$$

To prove Theorem 3.2, we need to prove the estimates ($H_1$) and ($H_2$) for this family of norms. This is the goal of the two next subsections. More precisely, we will show respectively the estimates

$$
\|(P_t - Q_t)f\|_{m,L+3} \leq C h_t^3 \|f\|_{2(m+3),L}, \ m \leq L + 3, f \in C_{\text{pol}}^{2(m+3),L}(\mathbb{R}+)
$$

in Proposition 5.3 and

$$
\sup_{t \geq T} \|P_t f\|_{m,L} + \max_{0 \leq j \leq n} \|Q_t^j f\|_{m,L} \leq \|f\|_{m,L} C h_t^3, \ m \leq L, f \in C_{\text{pol}}^{m,L}(\mathbb{R}+)
$$

in Proposition 5.9 for $Q_t$ as in Theorem 3.2. Note that $L$ has to be large enough: this is not an issue for our purpose since $C_{\text{pol}}^{m,L}(\mathbb{R}+) \subset C_{\text{pol}}^{m,L+1}(\mathbb{R}+)$, and we can work with $L$ as large as needed. We refer to the proof of Theorem 3.2 in Subsection 5.3 for further details.

Before, we summarize in the next lemma some properties of the norms defined in Equation (5.2) that we will use later on. Its proof is postponed to Appendix B

**Lemma 5.1.** Let $m, L \in \mathbb{N}$. We have the following basic properties:

1. $\|f\|_{m',L} = \max_{j \in \{0, \ldots, m'\}} \|f^{(j)}\|_{0,L}$ for $f \in C_{\text{pol}}^{m,L}(\mathbb{R}+)$ and $m' \in \{0, \ldots, m\}$.
2. $C_{\text{pol}}^{m+1,L}(\mathbb{R}+) \subset C_{\text{pol}}^{m,L}(\mathbb{R}+)$ and $\|f\|_{m,L} \leq \|f\|_{m+1,L}$ for $f \in C_{\text{pol}}^{m+1,L}(\mathbb{R}+)$. 
3. $\|f^{(i)}\|_{m,L} \leq \|f\|_{m+i,L}$ for $i \in \mathbb{N}$ and $f \in C_{\text{pol}}^{m+i,L}(\mathbb{R}+)$. 
4. $C_{\text{pol}}^{m,L}(\mathbb{R}+) \subset C_{\text{pol}}^{m+1,L}(\mathbb{R}+) \subset C_{\text{pol}}^{m,L}(\mathbb{R}+)$ and $\|f\|_{m,L+1} \leq 2 \|f\|_{m,L}$ for $f \in C_{\text{pol}}^{m,L}(\mathbb{R}+)$. 
5. Let $M_1$ be the operator defined by $f \mapsto M_1 f$, $M_1 f(x) = xf(x)$. Then, $M_1 f \in C_{\text{pol}}^{m,L+1}(\mathbb{R}+)$ for $f \in C_{\text{pol}}^{m,L}(\mathbb{R}+)$ and $\|M_1 f\|_{m,L+1} \leq (2m + 3) \|f\|_{m,L}$. 
6. Let $L f(x) = (a - k x)f'(x) + \frac{1}{2} a \sigma^2 x f''(x)$ be the infinitesimal generator of the CIR process. Then, we have for $f \in C_{\text{pol}}^{m+2,L}(\mathbb{R}+)$,

$$
\|L f\|_{m,L+1} \leq (2a + (2m + 3)(|k| + \sigma^2/2)) \|f\|_{m+2,L}.
$$

We also have $\|(V_t^2/2)f\|_{m,L+1} \leq \sigma^2 (m + 2) \|f\|_{m+1,L}$, $\|V_0 f\|_{m,L+1} \leq [2|a - \sigma^2/4| + (2m + 3)|k|] \|f\|_{m+1,L}$, where $V_0$ and $V_1$ are defined by (3.6). 

We also state the following elementary lemma that will be useful to prove both ($H_1$) and ($H_2$).

**Lemma 5.2.** Let $T > 0$, $\sigma^2 \leq 4a$ and $X_0$ be defined by (3.3). Then, there exists a constant $K \geq 0$ such that for any function $f \in C_{\text{pol}}^{m,L}(\mathbb{R}+)$, we have

$$
\|f(X_0(t, \cdot))\|_{m,L} \leq e^{Kt} \|f\|_{m,L}, \ t \in [0, T].
$$
Proposition 5.4. Let \( \nu \in \mathbb{R} \) be the scheme (3.7). For \( i \leq m \), we have:

\[
|\partial_x^i f(X_0(t,x))| = |e^{-ikt} (X_0(t,x))| \leq \frac{e^{-ikt}}{m,L} (1 + X_0(t,x))^L
\]

\[
\leq \frac{\|f\|_{m,L}}{m,L} e^{\frac{C_0(L+m)(-k)^+ t}{2}} (1 + x^L).
\]

This gives \( \|f(X_0(t,\cdot))\|_{m,L} \leq \|f\|_{m,L} e^{\frac{C_0(L+m)(-k)^+ t}{2}}. \)

\[\square\]

5.1. Proof of (H1). In this subsection, we prove the following result which is a direct consequence of Propositions 5.4 (with \( \nu = 2 \)) and 5.8 that are stated below.

**Proposition 5.3.** Let \( Y \) satisfy (H_Y), \( \sigma^2 \leq 4\alpha \) and \( \dot{X}_t = \varphi(x,t,\sqrt{Y}) \) be the scheme (3.7). Let \( m, L \in \mathbb{N} \) such that \( L + 3 \geq m \) and \( f \in C^{2(m+3)}_{\text{pol}}(\mathbb{R}^+). \) Then, there exists a constant \( C \in \mathbb{R}^+ \) such that for \( t \in [0,T] \),

\[
\|\mathbb{E}[f(\dot{X}_t)] - \mathbb{E}[f(X_t)]\|_{m,L+3} \leq C t^3 \|f\|_{2(m+3),L}. \]

To prove this result, we compare each term with the expansion \( f(x) + t\mathcal{L}_f(x) + \frac{t^2}{2} \mathcal{L}^2 f(x) \) of order two. The next proposition analyses the difference between such expansion and the semigroup of the CIR process.

**Proposition 5.4.** Let \( m, \nu, L \in \mathbb{N} \) such that \( L + \nu + 1 \geq m \), \( T > 0 \) and \( f \in C^{\nu+1,m}_{\text{pol}}(\mathbb{R}^+). \) Let \( X^x \) be the CIR process and \( \mathcal{L} \) its infinitesimal generator. Then, for \( t \in [0,T] \), we have

\[
\mathbb{E}[f(X^x_t)] = \sum_{i=0}^{\nu} \frac{t^i}{i!} \mathcal{L}^i f(x) + t^{\nu+1} \int_0^1 \frac{(1-s)^\nu}{\nu!} \mathbb{E}[\mathcal{L}^{\nu+1} f(X^x_{ts})] ds
\]

(5.3)
where the function $x \mapsto \int_0^1 \frac{(1-s)\nu}{\nu!} \mathbb{E}[L^{\nu+1} f(X_s^x)] ds$ belongs to $\mathcal{C}_{\text{pol}}^{m,L}(\mathbb{R}_+)$ and we have the following estimate for all $t \in [0, T]$, 
\[ \left\| \int_0^1 \frac{(1-s)^\nu}{\nu!} \mathbb{E}[L^{\nu+1} f(X_{ts})] ds \right\|_{m,L+\nu+1} \leq C \| f \|_{m+2(\nu+1),L}, \] (5.4)

for some constant $C \in \mathbb{R}_+$ depending on $(a, k, \sigma, m, L, T)$.

**Proof.** Let $f \in \mathcal{C}_{\text{pol}}^{m+2(\nu+1),L}(\mathbb{R}_+)$. Since the coefficients of the CIR SDE have sublinear growth, we have bounds on the moments of $X_t^x$: for any $q \in \mathbb{N}^*$, there exists $C_q > 0$ such that $\mathbb{E}[|X_s^x|^q] \leq C_q (1 + x^q)$ for $s \in [0, T]$. Using iterations of Itô’s formula and a change of variable (in time), we then easily get (5.3) for $t \in [0, T]$. To get the estimate (5.4), we first use Lemma 5.1 and obtain
\[ \|L^{\nu+1} f\|_{m,L+\nu+1} \leq K_{\text{cir}}(m, \nu)^{\nu+1} \| f \|_{m+2(\nu+1),L}, \]

with $K_{\text{cir}}(m, \nu) = 2a + (2m + 4\nu + 3)(|k| + \sigma^2/2)$. By the triangle inequality, we have
\[ \left\| \int_0^1 \frac{(1-s)\nu}{\nu!} \mathbb{E}[L^{\nu+1} f(X_{ts})] ds \right\|_{m,L+\nu+1} \leq \int_0^1 \frac{(1-s)\nu}{\nu!} \mathbb{E}[L^{\nu+1} f(X_{ts})] ds. \]

Since $t \leq T$, we have $\mathbb{E}[L^{\nu+1} f(X_{ts})]_{m,L+\nu+1} \leq C_{\text{cir}}(m, L + \nu + 1, T) \|L^{\nu+1} f\|_{m,L+\nu+1}$ by Proposition 5.10 using that $L + \nu + 1 \geq m$. This gives by Lemma 5.1
\[ \left\| \int_0^1 \frac{(1-s)\nu}{\nu!} \mathbb{E}[L^{\nu+1} f(X_{ts})] ds \right\|_{m,L+\nu+1} \leq \frac{C_{\text{cir}}(m, L + \nu + 1, T)}{(\nu+1)!} K_{\text{cir}}(m, \nu)^{\nu+1} \| f \|_{m+2(\nu+1),L}. \]

We now focus on the approximation scheme. The main difficulty comes from the differentiation of the square-root that may lead to derivatives that blow up at the origin. Here, we exploit the fact that $Y$ is a symmetric random variable to cancel these blowing terms. More precisely, we will then need to differentiate in $x$ the following quantity
\[ g(x) = g(X_1(s\sqrt{t}, x)) + g(X_1(-s\sqrt{t}, x)) = g(x + \sigma s \sqrt{t} \sqrt{x} + \frac{\sigma^2}{4} ts^2) + g(x - \sigma s \sqrt{t} \sqrt{x} + \frac{\sigma^2}{4} ts^2), \]

and the next lemma enables us to have a sharp estimate of the derivatives.

**Lemma 5.5.** Let $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a $\mathcal{C}^{2n}$ function, $\beta \in \mathbb{R}_+$ and $\gamma \geq \beta^2/4$. Then, the function $\psi_g(x) := g(x + \beta \sqrt{x} + \gamma) + g(x - \beta \sqrt{x} + \gamma)$, $x \geq 0$ is $\mathcal{C}^n$ with derivatives
\[ \psi_g^{(n)}(x) = \psi_g^{(n)}(x) + \sum_{j=1}^n \binom{n}{j} \beta^{2j} \int_0^1 g^{(n+j)}(x + \beta(2u - 1) \sqrt{x} + \gamma) \frac{(u - u^2)^{j-1}}{(j-1)!} du \] (5.5)

The proof of this lemma and of the next corollary are postponed to Appendix B.

**Corollary 5.6.** Let $m, L \in \mathbb{N}$, $\beta \geq 0$ and $g \in \mathcal{C}_{\text{pol}}^{2m,L}(\mathbb{R}_+)$. Then, $\psi_g(x) = g((\sqrt{x} + \beta/2)^2) + g((\sqrt{x} - \beta/2)^2)$ belongs to $\mathcal{C}_{\text{pol}}^{m,L}(\mathbb{R}_+)$, and for all $n \in \{0, \ldots, m\}$ we have the following estimates
\[ \| \psi_g \|_{n,L} \leq C_{\beta,m,L} \| g \|_{2n,L}, \] (5.6)

with $C_{\beta,m,L} = (1 + \beta/2)^{2L} + (1 - \beta/2)^{2L} + 2(1 + \beta^2/2)^L (1 + \beta^2/2)^m)$. 


Lemma 5.7. Let $m, \nu, L \in \mathbb{N}$, $T > 0$, $t \in [0, T]$ and $N \sim \mathcal{N}(0, 1)$. Let $Y$ be a symmetric random variable such that $\mathbb{E}[Y^k] = \mathbb{E}[N^k]$ for $k \leq 2\nu$ and $\mathbb{E}[Y^{2k}] < \infty$ for all $k \in \mathbb{N}$. We have, for $f \in C^{m+\nu+1,L}_{\text{pol}}(\mathbb{R}_+)$,

$$f(X_0(t, x)) = \sum_{i=0}^{\nu} \frac{t^i}{i!} V_i f(x) + t^{\nu+1} \int_0^1 \frac{(1-u)^{\nu}}{\nu!} V_{\nu+1} f(X_0(ut, x)) du, \quad (5.7)$$

with $\| \int_0^1 \frac{(1-u)^{\nu}}{\nu!} V_{\nu+1} f(X_0(ut, \cdot)) du \|_{m, L+\nu+1} \leq C_0 \| f \|_{m+\nu+1, L}$; and for $f \in C^{2(m+\nu+1),L}_{\text{pol}}(\mathbb{R}_+)$,

$$\mathbb{E}[f(X_1(\sqrt{t}Y, x))] = \sum_{i=0}^{\nu} \frac{t^i}{i!} \left( \frac{1}{2} V_1^2 \right)^i f(x) + t^{\nu+1} \mathbb{E}\left[ Y^{2\nu+2} \int_0^1 \frac{(1-u)^{2\nu+1}}{(2\nu+1)!} V_{2\nu+2} f(X_1(u\sqrt{t}Y, x)) du \right], \quad (5.8)$$

with $\| \mathbb{E}\left[ Y^{2\nu+2} \int_0^1 \frac{(1-u)^{2\nu+1}}{(2\nu+1)!} V_{2\nu+2} f(X_1(u\sqrt{t}Y, \cdot)) du \right] \|_{m, L+\nu+1} \leq C_1 \| f \|_{2(m+\nu+1), L}$, for some constants $C_0, C_1 \in \mathbb{R}^+$ depending on $a, k, \sigma, T, m, M$ and $\nu$.

Proof. Equation (5.7) holds by using Taylor formula since $\frac{d}{dt} f(X_0(t, x)) = V_0 f(X_0(t, x))$. We have by Property (6) of Lemma 5.1 $\| V_0 f \|_{m, L+1} \leq |a - \frac{a^2}{4}| \| f \|_{m, L+1} + |k|/(2m+3) \| f \|_{m, L} \leq (2|a - \frac{a^2}{4}| + |k|/(2m+3)) \| f \|_{m+1, L}$ and thus $\| V_{\nu+1} f \|_{m, L+\nu+1} \leq C \| f \|_{m+\nu+1, L}$ for some constant $C$ depending on $a, k, \sigma, \nu, u, m$. Using the triangular inequality and Lemma 5.2, we get the result.

We now prove the second part of the claim. We first show Equation (5.8). Since $\frac{d}{dt} f(X_1(t, x)) = V_1 f(X_1(t, x))$, we get by Taylor formula

$$f(X_1(t, x)) = \sum_{i=0}^{2\nu+1} \frac{t^i}{i!} V_i f(x) + t^{2\nu+2} \int_0^1 \frac{(t-s)^{2\nu+1}}{(2\nu+1)!} V^{2\nu+2} f(X_1(s, x)) du$$

$$= \sum_{i=0}^{2\nu+1} \frac{t^i}{i!} V_i f(x) + t^{2\nu+2} \int_0^1 \frac{(1-u)^{2\nu+1}}{(2\nu+1)!} V^{2\nu+2} f(X_1(ut, x)) du, \quad t \in \mathbb{R}.$$ 

We apply this formula at $\sqrt{t}Y$ and take the expectation. Since $\mathbb{E}[Y^{2i+1}] = 0$ by symmetry and $\mathbb{E}[Y^{2i}] = \mathbb{E}[N^{2i}] = (2i)! \sigma^{2i}$ for $i \leq \nu$, we get (5.8). We now analyze the norm of the remainder. We have $\| \frac{1}{2} V_1^2 f \|_{m, L+1} \leq \sigma^2 (m+2) \| f \|_{m+2, L}$ by using Lemma 5.1 (6). Then, we observe that by symmetry of $Y$,

$$\mathbb{E}\left[ Y^{2\nu+2} \int_0^1 \frac{(1-u)^{2\nu+1}}{(2\nu+1)!} V^{2\nu+2} f(X_1(u\sqrt{t}Y, x)) du \right] = \frac{1}{2} \mathbb{E}\left[ Y^{2\nu+2} \int_0^1 \frac{(1-u)^{2\nu+1}}{(2\nu+1)!} \left[ V^{2\nu+2} f(X_1(u\sqrt{t}Y, x)) + V^{2\nu+2} f(X_1(-u\sqrt{t}Y, x)) \right] du \right]$$

By Corollary 5.6, we have

$$\| V^{2\nu+2} f(X_1(u\sqrt{t}Y, \cdot)) + V^{2\nu+2} f(X_1(-u\sqrt{t}Y, \cdot)) \|_{m, L+\nu+1} \leq C_{\sigma u \sqrt{t}, m, L} \| V^{2\nu+2} f \|_{2(m, L+\nu+1)} \leq C' C_{\sigma u \sqrt{t}, m, L} \| f \|_{2(m+\nu+1), L},$$
with \( C' = (4a^2(m + \nu + 1))^{\nu+1} \). The conclusion follows by using the triangle inequality, the polynomial growth of the constant \( C_{\sigma w \sqrt{I}, m, L} \) given by Corollary 5.6 and the finite moments \( \mathbb{E}[Y^{2k}] \) for \( k \) sufficiently large. \( \square \)

We are now in position to prove the estimate for the approximation scheme (3.7). Since this scheme is obtained as the composition of the schemes \( X_0 \) and \( X_1 \), the proof consists of using iteratively the estimates of Lemma 5.7.

**Proposition 5.8.** Let \( Y \) be a symmetric random variable such that \( \mathbb{E}[Y^k] = \mathbb{E}[N^k] \) for \( k \leq 4 \) and \( \mathbb{E}[Y^{2k}] < \infty \) for all \( k \in \mathbb{N} \). Let \( \sigma^2 \leq 4a \) and \( \bar{X}_t^x \) be the scheme (3.7). Let \( m \in \mathbb{N}, L \in \mathbb{N}^* \) and \( f \in C^{2(\nu+3),L}_\text{pol}(\mathbb{R}_+) \). Then, we have for \( t \in [0, T] \),

\[
\mathbb{E}[f(\bar{X}_t^x)] = f(x) + t\mathcal{L}f(x) + \frac{t^2}{2} \mathcal{L}^2f(x) + \bar{R}_f(t, x),
\]

with \( \|\bar{R}_f(t, \cdot)\|_{m, L+3} \leq Ct^3\|f\|_{2(m+3), L} \).

**Proof.** We use \( \bar{X}_t^x = X_0(t/2, X_1(\sqrt{t}Y, X_0(t/2, x))) \) and apply first (5.7):

\[
\mathbb{E}[f(X_0(t/2, X_1(\sqrt{t}Y, X_0(t/2, x))))] = \mathbb{E}\left[ (f + \frac{t}{2}V_0f + \frac{t^2}{8}V_0^2f)(X_1(\sqrt{t}Y, X_0(t/2, x))) \right] + R_{I_1}f(t, x),
\]

with \( R_{I_1}f(t, x) = \left( \frac{t}{2} \right)^3 \int_0^1 \frac{(1-u)^2}{2} \mathbb{E}[V_0^3f(X_0(ut/2, X_1(\sqrt{t}Y, X_0(t/2, x))))]du \) ds.

We get by using Lemma 5.2, Corollary 5.6 (using the symmetry and the finite moments of \( Y \)), again Lemma 5.2 and then Lemma 5.1 (6):

\[
\|\mathbb{E}[V_0^3f(X_0(ut/2, X_1(\sqrt{t}Y, X_0(t/2, \cdot))))]\|_{m, L+3} \leq C\|\mathbb{E}[V_0^3f(X_0(ut/2, X_1(\sqrt{t}Y, \cdot)))]\|_{m, L+3}
\]

\[
\leq C\|V_0^3f(\cdot)\|_{2m, L+3} \leq C\|V_0^3f\|_{2m, L+3} \leq C\|f\|_{2m+3, L}
\]

This gives \( \|R_{I_1}f(t, x)\|_{m, L+3} \leq Ct^3\|f\|_{2m+3, L}, \) for \( t \in [0, T] \).

We now expand again and get from (5.8)

\[
\mathbb{E}\left[ (f + \frac{t}{2}V_0f + \frac{t^2}{8}V_0^2f)(X_1(\sqrt{t}Y, X_0(t/2, x))) \right] = f(X_0(t/2, x)) + \frac{t}{2}V_1f(X_0(t/2, x)) + \frac{t^2}{2}(V_1^2/2)^2f(X_0(t/2, x)) + \frac{t}{2}V_0f(X_0(t/2, x))
\]

\[
+ \frac{t^2}{2}(V_1^2/2)V_0f(X_0(t/2, x)) + \frac{t^2}{8}V_0^2f(X_0(t/2, x)) + R_{II}f(t, x),
\]

with

\[
R_{II}f(t, x) = t^3\mathbb{E}\left[ Y^5 \int_0^u \frac{(1-u)^5}{5!} V_1^4f(X_1(u\sqrt{t}Y, X_0(T/2, x)))du \right]
\]

\[
+ \frac{Y^4}{2} \int_0^u \frac{(1-u)^3}{3!} V_1^4 V_0 f(X_1(u\sqrt{t}Y, X_0(T/2, x)))du
\]

\[
+ \frac{Y^2}{8} \int_0^u (1-u) V_1^2 V_0^2 f(X_1(u\sqrt{t}Y, X_0(T/2, x)))du \right].
\]
We use Lemmas 5.7, 5.2 and 5.1 to get, for \( t \in [0, T]\),
\[
\|R_{II}f(t, \cdot)\|_{m,L+3} \leq C t^3 (\|f\|_{2(m+1),L} + \|V_0 f\|_{2(m+2),L+1} + \|V_0^2 f\|_{2(m+1),L+2}) \leq C t^3 \|f\|_{2(m+3),L}.
\]
Last, we use again (5.7) to get
\[
f(X_0(t/2, x)) + \frac{t}{2} [V_1^2 + V_0] f(X_0(t/2, x)) + \frac{t^2}{2} [(V_1^2/2)^2 + (V_1^2/2)V_0 + V_0^2/4] f(X_0(t/2, x))
= f(x) + \frac{t}{2} V_0 f(x) + \frac{t^2}{8} V_0^2 f(x) + \frac{t}{2} [V_1 + V_0] f(x) + \frac{t^2}{4} [V_0 V_1^2 + V_0^2] f(x)
+ \frac{t^2}{2} [(V_1^2/2)^2 + (V_1^2/2)V_0 + V_0^2/4] f(x) + R_{III} f(t, x),
\]
where again by Lemma 5.7 and 5.1, we have
\[
\|R_{III} f(t, \cdot)\|_{m,L+3} \leq C t^3 (\|f\|_{m+3,L} + \|[V_1^2 + V_0] f\|_{m+2,L+1}
+ \|[(V_1^2/2)^2 + (V_1^2/2)V_0 + V_0^2/4] f\|_{m+1,L+2})
\]
\[
\leq C t^3 \|f\|_{m+5,L}.
\]
Finally, we get \(\|\tilde{R} f(t, \cdot)\|_{m,L+3} \leq C t^3 \|f\|_{2(m+3),L} \) with \(\tilde{R} f := R_I f + R_{II} f + R_{III} f\) and
\[
E[f(X_0(t/2, X_1(\sqrt{4}Y, X_0(t/2, x)))]) = f(x) + t(V_0 + V_1^2/2) f(x)
+ \frac{t^2}{2} [V_0^2 + V_0 V_1^2/2 + (V_1^2/2)V_0 + (V_1^2/2)^2] f(x) + \tilde{R} f(t, x)
= f(x) + t L f(x) + \frac{t^2}{2} L^2 f(x) + \tilde{R} f(t, x).
\]
\(\square\)

5.2. Proof of (\(\mathcal{P}_2\)). In this section, we mainly prove the following result.

**Proposition 5.9.** Let \(\sigma^2 \leq 4\alpha\) and \(\tilde{X}_t^x = \varphi(x, t, \sqrt{N})\) be the scheme (3.1) with \(N \sim \mathcal{N}(0, 1)\). Let \(T > 0\) and \(m, L \in \mathbb{N}\) such that \(L \geq m\). We define for \(n \geq 1\) and \(l \in \mathbb{N}\), \(Q_l f(x) = E[f(\tilde{X}_h^x)]\) with \(h_l = T/n\). Then, there exists a constant \(C \in \mathbb{R}_+\) such that for any \(f \in C_{\text{pol}}^m(\mathbb{R}_+), l \in \mathbb{N}\) and \(t \in [0, T]\),
\[
\|E[f(X_t)]\|_{m,L} + \max_{0 \leq k \leq n} \left\|Q_l^{[k]} f\right\|_{m,L} \leq C \|f\|_{m,L}.
\] (5.9)

We split the proof in two parts. The first one deals with the semigroup of the CIR process, for which the assumption \(\sigma^2 \leq 4\alpha\) is not needed. This is stated in Proposition 5.10, whose proof exploits the particular form of the density of \(X_t^x\). The second part that deals with the approximation scheme is quite technical. We prove in fact in Proposition 5.12 a slightly more general result for the scheme \(\tilde{X}_t^x = \varphi(x, t, \sqrt{N})\), when \(Y\) is a symmetric random variable with a smooth density. However, the conditions needed on the density are quite restrictive. These conditions are satisfied by the standard normal variable by Lemma 5.14. If we want besides to have (5.9) for any \(m\) and in addition to match the moments \(E[Y^2] = E[N^2]\) and \(E[Y^4] = E[N^4]\) – which is required to have a second-order scheme –, then Theorem 5.16 shows that we necessarily have \(Y \sim \mathcal{N}(0, 1)\). This is why we directly state here, for sake of simplicity, Proposition 5.9 with \(Y = N \sim \mathcal{N}(0, 1)\).
5.2.1. **Upper bound for the semigroup.** We first prove the estimate \( \| \Pi_t \|_2 \) for the semigroup of the CIR process. To do so, we take back the arguments of [1, Proposition 4.1] that gives polynomial estimates for \((t, x) \mapsto P_t f(x)\). First we remove the polynomial Taylor expansion of the function \( f \) at 0, which enables then to do an integration by parts and to get the remarkable formula in Eq. (5.13) below for the iterated derivatives of \( P_t f \) that gives then the desired estimate. The polynomial part is analyzed separately in Lemma 5.11 below with standard arguments.

**Proposition 5.10.** Let \( f \in C_{\text{pol}}^{m,L}(\mathbb{R}_+), \ L \geq m, \ T > 0 \) and \( t \in (0,T] \). Let \( X^x \) be the CIR process starting from \( x \geq 0 \). Then, \( \mathbb{E}[f(X_t)] \in C_{\text{pol}}^{m,L}(\mathbb{R}_+) \) and we have the following estimate for some constant \( C_{\text{cir}}(m, L, T) \in \mathbb{R}_+: \)

\[
\| \mathbb{E}[f(X_t)] \|_{m,L} \leq C_{\text{cir}}(m, L, T) \| f \|_{m,L}.
\] (5.10)

**Proof.** Let \( f \in C_{\text{pol}}^{m,L}(\mathbb{R}_+) \) and \( T_m(f)(x) = \sum_{j=0}^{m} \frac{f^{(j)}(0)}{j!} x^j \) its Taylor polynomial expansion at 0 of order \( m \). We define \( \hat{f}_m = f - T_m(f) \in C_{\text{pol}}^{m,L}(\mathbb{R}_+), \) so we have \( f = \hat{f}_m + T_m(f) \). By Lemma 5.11 below, one gets \( \| \mathbb{E}[T_m(f)(X_t)] \|_{m,L} \leq C_{\text{cir}}(m, T) \| T_m(f) \|_{m,L} \) and then

\[
\| \mathbb{E}[T_m(f)(X_t)] \|_{m,L} \leq e C_{\text{cir}}(m, T) \| f \|_{m,L},
\] (5.11)

since for all \( i \in \{0, \ldots, m\} \) and \( x \geq 0 \)

\[
\left| \frac{(T_m(f))(i)(x)}{1 + x^L} \right| = \left| \sum_{j=0}^{m-i} \frac{f^{(i+j)}(0)}{j!} x^j \right| \leq \sum_{j=0}^{m-i} \frac{1}{j!} \| f \|_{m,L} \leq e \| f \|_{m,L}.
\]

We now focus on \( \mathbb{E}[\hat{f}_m(X_t)] \). We recall the density of \( X_t^x \) (see e.g. [3, Proposition 1.2.11])²

\[
p(t, x, z) = \sum_{i=0}^{\infty} \frac{e^{-d_t x/2}}{i!} \frac{c_t/2}{\Gamma(i + v)} \left( \frac{c_t z}{2} \right)^{i+1} e^{-c_t z/2}
\] (5.12)

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

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

We have

\[
\mathbb{E}[\hat{f}_m(X_t^x)] = \sum_{i=0}^{\infty} e^{-d_t x/2} \left( \frac{d_t x}{2} \right)^i I_i(\hat{f}_m, c_t), \ t > 0,
\]

where

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

²In the case \( a = 0 \), \( X_t^x \) is distributed according to the probability measure \( e^{-d_t x/2} \delta_0(dx) + \sum_{i=1}^{\infty} e^{-d_t x/2} \binom{d_t x}{2i} \frac{c_t/2}{\Gamma(i)} \left( \frac{c_t z}{2} \right)^{i-1} e^{-c_t z/2} \). The proof works the same since \( \hat{f}_m(0) = 0 \), so that \( \mathbb{E}[\hat{f}_m(X_t^x)] \) only involves the absolutely continuous part of the distribution.
Differentiating successively, we get that for \( j \leq m \), \( t \in (0, T] \) and \( x \in \mathbb{R}_+ \)

\[
\partial_j^2 \mathbb{E}[\hat{f}(X_t^x)] = \sum_{i=0}^{\infty} \frac{e^{-dx/2(i+1)}}{i!} \Delta^i(I_i(\hat{f}_m, c_t)),
\]

where \( \Delta_t : \mathbb{R}^N \to \mathbb{R}^N \) is an operator defined on sequences \( (I_i)_{i \geq 0} \in \mathbb{R}^N \) by \( \Delta_t(I_i) = \frac{d_t}{2} (I_{i+1} - I_i) = e^{-ct} (I_{i+1} - I_i) \). An integration by parts gives for \( i \geq 1 \)

\[
I_i(\hat{f}_m^{(j)}, c_t) = \int_0^\infty \hat{f}_m^{(j-1)}(z) \left( \frac{c_t}{2} \right)^2 \Gamma(i + 1 + v) \left( \frac{c_t z}{2} \right)^{i+1+v} e^{-ctz/2} dz
\]

\[
- \int_0^\infty \hat{f}_m^{(j-1)}(z) \left( \frac{c_t}{2} \right)^2 (i - 1 + v) \left( \frac{c_t z}{2} \right)^{i-2+v} e^{-ctz/2} dz
\]

\[
= \frac{c_t}{2} (I_i(\hat{f}_m^{(j-1)}, c_t) - I_{i-1}(\hat{f}_m^{(j-1)}, c_t)) = e^{-ct} \Delta_i(I_{i-1}(\hat{f}_m^{(j-1)}, c_t)),
\]

since \( \hat{f}_m^{(j)}(0) = 0 \) for all \( 1 \leq j \leq m \) and \( \hat{f}_m^{(j)} \) has a polynomial growth. By iterating, we get for all \( t \in (0, T] \) and \( x \in \mathbb{R}_+ \),

\[
\partial_j^2 \mathbb{E}[\hat{f}_m(X_t^x)] = \sum_{i=0}^{\infty} \frac{e^{-dx/2(i+1)}}{i!} I_{i+j}(\hat{f}_m^{(j)}, c_t) e^{-kjt}. \tag{5.13}
\]

Note that, since for \( j \leq m, |\hat{f}_m^{(j)}(x)| \leq \|\hat{f}_m\|_{m,L} (1 + zL) \) and using the well known formula for the \( L \)-th raw moment of gamma distribution we have for all \( i \in \mathbb{N} \)

\[
|I_i(\hat{f}_m^{(j)}, c_t)| \leq \|\hat{f}_m\|_{m,L} \left( 1 + \left( \frac{2}{c_t} \right)^L \Gamma(i + L + v) \right). \tag{5.14}
\]

Thus, the derivation of the series \((5.13)\) is valid, and we get that

\[
|\partial_j^2 \mathbb{E}[\hat{f}_m(X_t^x)]| \leq \|\hat{f}_m\|_{m,L} e^{-kJt} \left( 1 + \left( \frac{2}{c_t} \right)^L \sum_{i=0}^{\infty} \frac{e^{-dx/2(i+1)}}{i!} \Gamma(i + j + L + v) \right). \]

The quotient \( \frac{\Gamma(i + j + L + v)}{\Gamma(i + j + v)} \) is a polynomial function of degree \( L \) with respect to \( i \), and we denote \( \beta^1_0, \ldots, \beta^2_L \) its coefficients in the basis \( \{1, i, i(i-1), \ldots, i(i-1) \cdots (i-L+1)\} \). Thus, we get that

\[
|\partial_j^2 \mathbb{E}[\hat{f}_m(X_t^x)]| \leq \|\hat{f}_m\|_{m,L} e^{(-k)^+j} \left( 1 + \left( \frac{2}{c_t} \right)^L \sum_{i=0}^{L} |\beta^1_i| \left( \frac{d_t}{2} \right)^i x^i \right)
\]

\[
\leq \|\hat{f}_m\|_{m,L} e^{(-k)^+m} \left( 1 + \sum_{i=0}^{L} |\beta^1_i| \left( \frac{2}{c_t} \right)^{L-i} e^{-kit} (1 + x^L) \right)
\]

\[
\leq \|\hat{f}_m\|_{m,L} e^{(-k)^+(m+L)} \left( 1 + \sum_{i=0}^{L} |\beta^1_i| \left( \frac{2}{c_{\min}} \right)^{L-i} \right) (1 + x^L).
\]

By the triangular inequality and \((5.11)\), we get \( \|\hat{f}_m\|_{m,L} \leq (1 + e) \|f\|_{m,L} \), so one has for all \( t \in (0, T], j \leq m \)

\[
|\partial_j^2 \mathbb{E}[\hat{f}_m(X_t^x)]| \leq (1 + e) e^{(-k)^+(m+L)} \left( 1 + \sum_{i=0}^{L} |\beta^1_i| \left( \frac{2}{c_{\min}} \right)^{L-i} \right) \|f\|_{m,L} (1 + x^L), \tag{5.15}
\]
and thus for all \( t \in (0, T) \):
\[
\| \mathbb{E}[\hat{f}_m(X_t)] \|_{m,L} \leq \hat{C} \| f \|_{m,L},
\]
(5.16)
where \( \hat{C} := (1 + e)e^{(−k^+(m+L))T} \max_{0 \leq j \leq m} \left( 1 + \sum_{i=0}^{L} |\beta_i^j| \left( \frac{2^{\min}}{e_{\min}} \right)^{L-i} \right) \). Finally, we get the desired estimate by the triangular inequality, (5.11) and Lemma 5.11:
\[
\| \mathbb{E}[f(X_t)] \|_{m,L} \leq \| \mathbb{E}[\hat{f}_m(X_t)] \|_{m,L} + \| \mathbb{E}[T_m(f)(X_t)] \|_{m,L} \leq (\hat{C} + C_{\text{cir}}(m, T)) \| f \|_{m,L}.
\]

**Lemma 5.11.** Let \( P \in \mathfrak{P}_m(\mathbb{R}^+) \) be a polynomial function of degree \( m \in \mathbb{N}^* \) and \( L \in \mathbb{N}^* \) such that \( L \geq m \). Then, for \( t \in [0, T] \) we have the following estimate
\[
\| \mathbb{E}[P(X_t)] \|_{m,L} \leq C_{\text{circ}}(m, T) \| P \|_{m,L},
\]
(5.17)
where \( C_{\text{circ}}(m, T) = \max_{t \in [0, T]} \sum_{j=0}^{m} \sum_{i=0}^{j} |\tilde{u}_{i,j}(t)| \) with \( \tilde{u}_{i,j}(t) \) defined as in Lemma 4.2 by
\[
\mathbb{E}[(X_t^f)^i] = \sum_{i=0}^{j} \tilde{u}_{i,j}(t)x^i.
\]

**Proof.** We consider a polynomial function \( P(y) = \sum_{i=0}^{m} b_i y^i \) of degree \( m \) and \( L \geq m \). For all \( l \in \{0, \ldots, m\} \) one has from Lemma 4.2
\[
|\partial^l_t \mathbb{E}[P(X_t^f)]| \leq \sum_{j=0}^{m} |b_j| \sum_{i=0}^{j} |\tilde{u}_{i,j}(t)| \leq \sum_{j=0}^{m} |b_j| \sum_{i=0}^{j} |\tilde{u}_{i,j}(t)| \leq \max_{t \in [0, T]} \sum_{j=0}^{m} \sum_{i=0}^{j} |\tilde{u}_{i,j}(t)| |b_j|!.
\]

passing to supremum over \( x \geq 0 \), \( l \in \{0, \ldots, m\} \) we get (5.17) observing that \( |b_j|! = |P^{(j)}(0)| \leq \| P \|_{m,L} \). \( \Box \)

### 5.2.2. Upper bound for the approximation scheme.

We now prove the estimate \((\Pi_2)\) for the approximation of the CIR process. The main result of this paragraph is the following.

**Proposition 5.12.** Let \( T > 0, \sigma^2 \leq 4a \), \( m, M \in \mathbb{N}, Y \) be a symmetric random variable with density \( \eta \in \mathcal{C}^M(\mathbb{R}) \) such that for all \( i \in \{0, \ldots, M\} \), \( \eta^{(i)}(y) = o(|y|^{-(2L+i)}) \) for \( |y| \to \infty \), and \( \eta^*_m \) for all \( 1 \leq m \leq M \) (see Lemma 5.13 below for the definition of \( \eta^*_m \)). Let \( Q_t f(x) = \mathbb{E}[f(\tilde{X}^f_{ht})] \) with \( \tilde{X}^f_t = \varphi(t, x, \sqrt{t}Y) \), \( n \geq 1 \), \( l \in \mathbb{N} \) and \( h_t = T/n^l \). Then, for any \( L \in \mathbb{N} \), there exists \( C \in \mathbb{R}^+ \) such that:
\[
\max_{0 \leq j \leq n^l} \| Q_t^{[j]} f \|_{m,L} \leq C \| f \|_{m,L},
\]
\( f \in \mathcal{C}^m_{\text{pol}}(\mathbb{R}^+), l \in \mathbb{N} \).

Note that by Lemma 5.14 below, the assumptions of Proposition 5.12 are satisfied by \( Y \sim \mathcal{N}(0, 1) \). Therefore, \((\Pi_2)\) holds for the scheme of Ninomiya and Victoir (3.1).

**Proof.** We have \( \tilde{X}^f_t = \varphi(t, x, \sqrt{t}Y) = X_0(t/2, X_1(\sqrt{t}Y, X_0(t/2, x))) \). Let \( f \in \mathcal{C}^m_{\text{pol}}(\mathbb{R}^+) \). We apply Lemma 5.2 and Lemma 5.13 below to get:
\[
\| \mathbb{E}[f(X_0(t/2, X_1(\sqrt{t}Y, X_0(t/2, \cdots))))] \|_{m,L} \leq e^{Kt/2} \| \mathbb{E}[f(X_1(\sqrt{t}Y, X_0(t/2, \cdots))))] \|_{m,L} \\
\leq e^{Kt/2+Ct} \| f(X_0(t/2, \cdots)) \|_{m,L} \leq e^{(C+K)t} \| f \|_{m,L}
\]
This gives \( \max_{0 \leq j \leq n^l} \| Q_t^{[j]} f \|_{m,L} \leq e^{(C+K)t} \| f \|_{m,L} \). \( \Box \)
Lemma 5.13. Let $M, L \in \mathbb{N}$. Let $Y$ be a symmetric random variable with density $\eta \in \mathcal{C}^M(\mathbb{R})$ such that for all $i \in \{0, \ldots, M\}$, $|\eta^{(i)}(y)| = o(|y|^{-2L+i})$ for $|y| \to \infty$. Then, for all function $f \in \mathcal{C}^{M,L}_{\text{pol}}(\mathbb{R}_+)$, $m \in \{1, \ldots, M\}$ and $t \in [0, T]$ one has the following representation

$$
\partial_x^n \mathbb{E}[f(X_1(\sqrt{t} Y, x))] = \int_{-\infty}^{\infty} \int_0^1 (u - u^2)^{m-1} f(u, x, y) \eta^*_m(y) dy du
$$

where $w(u, x, y) = x + (2u-1)\sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4$, $\eta^*_m(y) = (-1)^{m-1} \left( \sum_{j=1}^{m} c_{j,m} y^j \eta^{(j)}(y) \right)$, and the coefficients $c_{j,m}$ are defined by induction, starting from $c_{1,1} = -1$, through the following formula

$$
c_{j,m} = \left( \frac{2j}{m-1} - 4 \right) c_{j,m-1} 1_{j < m} + \frac{2}{m-1} c_{j-1,m-1} 1_{j > 1}, \quad j \in \{1, \ldots, m\}, \ m \in \{2, \ldots, M\}.
$$

In particular, $c_{m,m} = -\frac{2^{m-1}}{(m-1)!} < 0$. Furthermore, if the density $\eta$ is such that $\eta^*_m(y) \geq 0$ for all $y \in \mathbb{R}$, and all $m \in \{1, \ldots, M\}$, then there exists $C \in \mathbb{R}_+$ such that

$$
\|\mathbb{E}[f(X_1(\sqrt{t} Y, \cdot))]|m,L \leq (1 + C t) \|f\|_{m,L}, \ t \in [0, T].
$$

Proof. We first consider $m = 1$ and $f \in \mathcal{C}^{M,L}_{\text{pol}}(\mathbb{R}_+)$. From the symmetry of $Y$, we have the equality $\mathbb{E}[f(X_1(\sqrt{t} Y, x))] = \mathbb{E}[f(X_1(\sqrt{t} Y, x)] + f(X_1(-\sqrt{t} Y, x))] / 2$ and using the notation $\psi^\pm_f(x, y) = f(x + \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4) \pm f(x - \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4)$ we can write,

$$
\partial_x \mathbb{E}[f(X_1(\sqrt{t} Y, x))] = \frac{1}{2} \int_{-\infty}^{\infty} \partial_x \psi^+_f(x, y) \eta(y) dy.
$$

One derivation and a little of algebra show that

$$
\partial_x \psi^+_f(x, y) = (1 + \frac{\sigma \sqrt{ty}}{2 \sqrt{x}}) f'(x + \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4) + (1 - \frac{\sigma \sqrt{ty}}{2 \sqrt{x}}) f'(x - \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4)
$$

$$
= \frac{1}{\sigma \sqrt{ty} \sqrt{x}} \left( (\sigma^2 ty/2 + \sigma \sqrt{ty} \sqrt{x}) f'(x + \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4) - (\sigma^2 ty/2 - \sigma \sqrt{ty} \sqrt{x}) f'(x - \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4) \right)
$$

$$
= \frac{1}{\sigma \sqrt{ty} \sqrt{x}} \left( \partial_y [f(x + \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4)] - \partial_y [f(x - \sigma \sqrt{ty} \sqrt{x} + \sigma^2 ty^2/4)] \right)
$$

$$
= \partial_y \psi^-_f(x, y) / \sigma \sqrt{ty} \sqrt{x}.
$$

Integrating by parts in the variable $y$, observing that the boundary term vanishes since $|\eta(y)| = |y| \to \infty o(|y|^{-2L})$ and $f(z) = z \to \infty O(z^L)$, one has

$$
\partial_x \mathbb{E}[f(X_1(\sqrt{t} Y, x))] = -\frac{1}{2} \int_{-\infty}^{\infty} \frac{\psi^-_f(x, y) \eta'(y)}{\sigma \sqrt{ty} \sqrt{x}} dy.
$$
write explicitly the partial derivatives of \( \eta \) the hypothesis on the variable \( m \) that the result holds for (5.18) is true for \( m \) to be intended as the constant map identically equal to 1. So far, we have shown that formula \( \int_{-\infty}^{\infty} \int_{0}^{1} s(u, y)(u^2 - u)^{m-1} f^{(l)}(w(u, x, y))(\sum_{j=1}^{n} c_{j,n,y^{j}} \eta^{(j)}(y)) dudy \), (5.22)

so we can rewrite (5.18) as \( \partial_x^m \mathbb{E}[f(X_1(\sqrt{t}Y, x))] = I_{m,m}^{(m)}(1) \) where the 1 in the argument has to be intended as the constant map identically equal to 1. So far, we have shown that formula (5.18) is true for \( m = 1 \), we take now \( m \geq 2 \) and we prove it by induction over \( m \) assuming that the result holds for \( m - 1 \). We differentiate Eq. (5.18) for \( m - 1 \) and use the second equality of (5.21) to get

\[
\partial_x^m \mathbb{E}[f(X_1(\sqrt{t}Y, x))] = I_{m-1,m-1}^{(m)}(1) + I_{m-1,m-1}^{(m)} \left( \frac{(2u-1)\sigma \sqrt{t}y}{2\sqrt{x}} \right),
\]

(5.23)

Then, from the third equality of (5.21), one has \( \frac{(2u-1)\sigma \sqrt{t}y}{2\sqrt{x}} = \frac{(2u-1)}{2\sqrt{x}} \partial_y w - (2u - 1)^2 \) and so

\[
\partial_x^m \mathbb{E}[f(X_1(\sqrt{t}Y, x))] = I_{m-1,m-1}^{(m)}(1 - (2u - 1)^2) + I_{m-1,m-1}^{(m)} \left( \frac{(2u-1)}{\sigma \sqrt{t}x} \partial_y w(u, x, y) \right)
\]= \-4I_{m-1,m-1}^{(m)}(u^2 - u) + I_{m-1,m-1}^{(m)} \left( \frac{(2u-1)}{\sigma \sqrt{t}x} \partial_y w(u, x, y) \right)
\]= \-4I_{m-1,m-1}^{(m)}(1) + I_{m-1,m-1}^{(m)} \left( \frac{(2u-1)}{\sigma \sqrt{t}x} \partial_y w(u, x, y) \right).
\]

We work on the term \( I_{m-1,m-1}^{(m)}(2u-1) \sigma \sqrt{t}x \partial_y w(u, x, y) \). We use first an integration by parts in the variable \( y \) and subsequently one in the variable \( u \). The boundary terms vanishes by using the hypothesis on \( \eta \) since \( |f^{(m)}(w(u, x, y))| \leq \|f\|_{m,L}(1 + w(u, x, y)L) \to_{|y|\to\infty} O(y^{2L}) \) and to the fact that the function \( u^2 - u \) vanishes in 0 and 1. One gets

\[
\int_{0}^{1} \int_{-\infty}^{\infty} \frac{(2u-1)(u^2 - u)^{m-2}}{\sigma \sqrt{t}x} f^{(m)}(w(u, x, y))\partial_y w(u, x, y) \left( \sum_{j=1}^{m-1} c_{j,m-1,y^{j}} \eta^{(j)}(y) \right) dydu
\]
The double integral

\[= \int_{-\infty}^{\infty} \int_{0}^{1} \frac{(2u-1)(u^2-u)^{m-2}}{\sigma \sqrt{4} \sqrt{x}} f^{(m-1)}(w(u, x, y)) \left( \sum_{j=1}^{m-1} c_{j,m-1}(jy^{j-1} \eta^{(j)}(y) + y^{j+1} \eta^{(j+1)}(y)) \right) dudy \]

\[= \int_{-\infty}^{\infty} \int_{0}^{1} (u^2-u)^{m-1} f^{(m)}(w(u, x, y)) \left( \sum_{j=1}^{m-1} \frac{2}{m-1} c_{j,m-1}(jy^{j} \eta^{(j)}(y) + y^{j+1} \eta^{(j+1)}(y)) \right) dudy \]

\[= \int_{-\infty}^{\infty} \int_{0}^{1} (u^2-u)^{m-1} f^{(m)}(w(u, x, y)) \frac{2}{m-1} \left( \sum_{j=1}^{m-1} j c_{j,m-1} y^{j} \eta^{(j)}(y) + \sum_{j=2}^{m} c_{j-1,m-1} y^{j} \eta^{(j)}(y) \right) dudy. \]

Rewriting the last equality for \( \partial_x^m \mathbb{E}[f(X_1(\sqrt{Y}, x))] \), one has

\[
\partial_x^m \mathbb{E}[f(X_1(\sqrt{Y}, x))] = \int_{-\infty}^{\infty} \int_{0}^{1} (u^2-u)^{m-1} f^{(m)}(w(u, x, y)) \left( -4 \sum_{j=1}^{m-1} c_{j,m-1} y^{j} \eta^{(j)}(y) \right) dudy + \\
+ \int_{-\infty}^{\infty} \int_{0}^{1} (u^2-u)^{m-1} f^{(m)}(w(u, x, y)) \frac{2}{m-1} \left( \sum_{j=1}^{m-1} j c_{j,m-1} y^{j} \eta^{(j)}(y) + \sum_{j=2}^{m} c_{j-1,m-1} y^{j} \eta^{(j)}(y) \right) dudy \\
= \int_{-\infty}^{\infty} \int_{0}^{1} (u^2-u)^{m-1} f^{(m)}(w(u, x, y)) \left( \frac{2}{m-1} - 4 \right) c_{1,m-1} y^{1} \eta^{(1)}(y) + \\
\sum_{j=2}^{m-1} \left( \frac{2j}{m-1} - 4 \right) c_{j,m-1} + \frac{2}{m-1} c_{j-1,m-1} \right) y^{j} \eta^{(j)}(y) + \frac{2}{m-1} c_{m-1,m-1} y^{m} \eta^{(m)}(y) dudy, \tag{5.25}
\]

which proves the representation (5.18). Since \( c_{1,1} = -1 \) and \( c_{m,m} = -\frac{2}{m-1} c_{m-1,m-1} \) for \( m \geq 2 \), we get \( c_{m,m} = -\frac{2m-2}{(m-1)!} \) for \( m \geq 1 \).

We are now able to prove the estimate using this representation. Defining \( \eta_m^*(y) = (-1)^{m-1} \sum_{j=0}^{m} c_{j,m} y^{j} \eta^{(j)}(y) \), that is nonnegative for all \( y \) by hypothesis, one has

\[
|\partial_x^m \mathbb{E}[f(X_1(\sqrt{Y}, x))]| \leq \int_{0}^{1} (u-u^2)^{m-1} \int_{-\infty}^{\infty} |f^{(m)}(w(u, x, y))| \eta_m^*(y) dydu \\
\leq \|f\|_{m,L} \int_{0}^{1} (u-u^2)^{m-1} \int_{-\infty}^{\infty} (1+w(u, x, y)L) \eta_m^*(y) dydu \\
= \|f\|_{m,L} \int_{0}^{1} (u-u^2)^{m-1} \left[ \int_{-\infty}^{\infty} \eta_m^*(y) dydu \right]_A \\
+ \|f\|_{m,L} \int_{0}^{1} (u-u^2)^{m-1} \int_{-\infty}^{\infty} w(u, x, y)L \eta_m^*(y) dydu. \tag{5.26}
\]

The double integral \( A \) can be seen by means of representation (5.18) with \( f(x) = \frac{x^m}{m!} \ (f^{(m)} \equiv 1) \) as

\[
A = \partial_x^m \mathbb{E} \left[ \frac{X_1(\sqrt{Y}, x)^m}{m!} \right] = \frac{1}{m!} \partial_x^m \sum_{j=0}^{m} \frac{2m}{2j} x^{m-j} \left( \frac{\sigma \sqrt{Y}}{2} \right)^j \mathbb{E}[Y^{2j}] = 1,
\]
by using the symmetry of the density $\eta$. In the same way, $B$ can be seen by means of the representation as

$$
B = \partial_x^m \mathbb{E} \left[ \frac{L!}{(L+m)!} X_1(\sqrt{Y}, x)^{L+m} \right]
$$

$$
= \partial_x^m \sum_{j=0}^{L+m} \binom{2(L+m)}{2j} \frac{L!(L+m-j)!}{(L+m)!} \left( \frac{\sigma \sqrt{t}}{2} \right)^{2j} \mathbb{E}[Y^{2j}]
$$

$$
= \sum_{j=0}^{L} \binom{2(L+m)}{2j} \frac{L!(L+m-j)!}{(L-j)!(L+m)!} x^{L-j} \left( \frac{\sigma \sqrt{t}}{2} \right)^{2j} \mathbb{E}[Y^{2j}]
$$

$$
\leq x^L + t \sum_{j=1}^{L} \binom{2(L+m)}{2j} \frac{L!(L+m-j)!}{(L-j)!(L+m)!} x^{L-j} \left( \frac{Ct \sigma}{2} \right)^{2j} t^{j-1} \mathbb{E}[Y^{2j}]
$$

$$
\leq x^L + Ct(1+x^L)
$$

where $C_t = \max(1, T)$ and $C = \frac{1}{2} \left(1 + \frac{\sigma c T}{2}\right)^{2(L+k)} + \left(1 - \frac{\sigma c T}{2}\right)^{2(L+k)} \left(1 + \mathbb{E}[Y^{2L}]\right)$. Putting parts $A$ and $B$ back together one has

$$
\partial_x^m \mathbb{E}[f(X_1(\sqrt{Y}, x))] \leq \|f\|_{m,L}(1+x^L + Ct(1+x^L)) = (1+x^L)(1+Ct)\|f\|_{m,L}, \quad (5.26)
$$

and this proves the desired norm inequality. \hfill \square

**Lemma 5.14.** Let $\eta(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$ be the density of a standard normal variable. Then, we have for $m \geq 1$:

$$
\eta_m(y) := (-1)^{m-1} \sum_{j=1}^{m} c_{j,m} y^j \eta^{(j)}(y) = -c_{m,m} y^{2m} \eta(y), \quad (5.27)
$$

so, in particular $\eta_m^*(y) \geq 0$ for all $y \in \mathbb{R}$.

**Proof.** For $m = 1$, (5.27) is clearly true since $\eta'(y) = -y \eta(y)$. We now take $m \geq 2$, $M \geq m$ and we suppose (5.27) true for $m-1$: for all $f \in C_{\text{pol}}^{M,L}(\mathbb{R}_+)$ and $x \in \mathbb{R}_+$, we have

$$
\int_{-\infty}^{\infty} \int_0^1 (u-u^2)^{m-2} f^{(m-1)}(w(u,x,y)) (\eta_{m-1}'(y) + c_{m-1,m-1} y^{2m-2} \eta(y)) \, du \, dy = 0.
$$

Doing one differentiation step with respect to $x$ like in the proof of Lemma 5.13 and using that $\eta'(y) = -y \eta(y)$, we obtain

$$
\int_{-\infty}^{\infty} \int_0^1 (u-u^2)^{m-1} f^{(m)}(w(u,x,y)) (\eta_m'(y) + c_{m,m} y^{2m} \eta(y)) \, du \, dy = 0.
$$

By choosing $f_L(x) := \frac{L!}{(L+m)!} x^{L+m}$ for $L \in \mathbb{N}$, we get for all $L \in \mathbb{N}, \, x \in \mathbb{R}_+$,

$$
\int_{-\infty}^{\infty} \int_0^1 (u-u^2)^{m-1} w(u,x,y)^L (\eta_m'(y) + c_{m,m} y^{2m} \eta(y)) \, du \, dy = 0.
$$
We now take \( x = 0 \) so that \( w(u, 0, y) = \frac{u^2}{4} y^2 \) and then
\[
\int_{-\infty}^{\infty} y^{2L} (\eta_m^*(y) + c_{m,m} y^{2m} \eta(y)) dy = 0, \quad L \in \mathbb{N}.
\]
We remark also that \( \eta_m^*(y) = (\sum_{j=1}^{m} (-1)^{m+j-1} c_{j,m} y^j H_j(y)) \eta(y) \), where \( H_j \) is the \( j \)th Hermite polynomial function (defined by \( \eta(y) = (-1)^j H_j(y) \eta(y) \)). Thus, \( \eta_m^*(y) + c_{m,m} y^{2m} \eta(y) = P_{2m}(y) \eta(y) \) where \( P_{2m} \) is an even polynomial function of degree \( 2m \). We therefore obtain \( \int_{-\infty}^{\infty} y^l P_{2m}(y) \eta(y) dy = 0 \) for all \( l \in \mathbb{N} \), which gives \( P_{2m} = 0 \) and thus the claim. \( \square \)

**Remark 5.15.** Lemma 5.14 gives a remarkable formula of the monomial of order \( 2m \) \( m \in \mathbb{N}^* \) in terms of the first \( m \) Hermite polynomials multiplied respectively by the first \( m \) monomials
\[
y^{2m} = \sum_{j=1}^{m} (-1)^{m+j} \frac{c_{j,m}}{c_{m,m}} y^j H_j(y).
\]

The next result gives a kind of reciprocal result to Lemma 5.14. It explains why we consider \( (n-1-k+1) \) terms in Theorem 3.2, since we use Proposition 5.12 for any \( M \in \mathbb{N} \).

**Theorem 5.16.** Let \( Y \) be a symmetric random variable with a \( C^\infty \) probability density function \( \eta \) such that \( \mathbb{E}[Y^2] = 1 \), \( \mathbb{E}[Y^4] = 3 \) and \( \eta_m \geq 0 \) for all \( m \geq 1 \). Then, \( Y \sim \mathcal{N}(0,1) \).

**Proof.** By Corollary C.2, there exists a positive Borel measure \( \mu \) such that \( \eta(x) = \int_0^\infty e^{-tx^2} \mu(dt) \).
Since \( \int_{\mathbb{R}} \eta = 1 \), we get \( \int_0^\infty \sqrt{\frac{\pi}{t}} \mu(dt) = 1 \) and then \( \eta(x) = \int_0^\infty \frac{e^{-tx^2}}{\sqrt{\pi/t}} \tilde{\mu}(dt) \) with \( \tilde{\mu}(dt) = \sqrt{\frac{\pi}{t}} \mu(dt) \) being a probability measure on \( \mathbb{R}_+ \). We have \( \mathbb{E}[Y^2] = \int_0^\infty \int_0^\infty x^2 e^{-tx^2} dx \tilde{\mu}(dt) = \int_0^\infty \frac{1}{2t} \tilde{\mu}(dt) \) and \( \mathbb{E}[Y^4] = \int_0^\infty \frac{1}{(1/2t)^2} \tilde{\mu}(dt) \). Therefore, we have
\[
\int_0^\infty \frac{1}{2t} \tilde{\mu}(dt) = \int_0^\infty \left( \frac{1}{2t} \right)^2 \tilde{\mu}(dt) = 1.
\]
The equality condition in the Cauchy-Schwarz inequality implies that \( \tilde{\mu}(dt) = \delta_{1/2}(dt) \), i.e. \( Y \) is a standard normal variable. \( \square \)

### 5.3. Proof of Theorem 3.2

We prove the result for \( \hat{P}^{2,n} \). By assumption, \( f \in C^{18,L}_{pol}(\mathbb{R}_+) \), for \( L \geq 18 \) sufficiently large. From (2.6), we have
\[
Pr f - \hat{P}^{2,n} f = \sum_{k=0}^{n-1} Q_1^{[n-(k+1)]} [P_{h_1} - Q_1^{[n]}] Q_1^{[k]} f
\]
\[
+ \sum_{k=0}^{n-1} \sum_{k' = 0}^{n-(k+k'+2)} \sum_{h_1 = 0}^{n-(k+k'+2)} [P_{h_1} - Q_1] Q_1^{[k']} [P_{h_1} - Q_1] Q_1^{[k]} f
\]
\[
\leq C \|P_{h_1} - Q_1\|_0 \|f\|_{6, L+6} \leq C \|P_{h_1} - Q_1\|_0 \|f\|_{6, L+3}
\]
\[
\leq C h_1^3 \|Q_1^{[k']}\|_{6, L+3} \|P_{h_1} - Q_1\|_0 \|f\|_{6, L+3}
\]

For the other term, we write $P_{h_1} - Q_2^{[n]} = \sum_{k'=0}^{n-1} P_{h_1 - (k'+1)h_2}[P_{h_2} - Q_2]Q_2^{[k']} \text{ and get for } k, k' \in \{0, \ldots, n - 1\}$ by using Proposition 5.9, Proposition 5.3 and Lemma 5.1:

$$\|Q_1^{[n-(k+1)]}P_{h_1 - (k'+1)h_2}[P_{h_2} - Q_2]Q_2^{[k']}Q_1^{[k]}f\|_{0,L+6} \leq C\|P_{h_2} - Q_2]Q_2^{[k']}Q_1^{[k]}f\|_{0,L+6}$$

$$\leq Ch_1^3\|Q_2^{[k']}Q_1^{[k]}f\|_{6,L+3}$$

$$\leq Ch_2^3\|f\|_{6,L+3} \leq Ch_2^3\|f\|_{18,L}.$$ 

This gives

$$\|P_T f - \hat{P}^{2,n} f\|_{0,L+6} \leq C\|f\|_{18,L}n^2(h_1^6 + h_2^6) \leq C\|f\|_{18,L}n^{-4},$$

and in particular $P_T f(x) - \hat{P}^{2,n} f(x) = O(n^{-4})$ for any $x \geq 0$.

We now consider $f \in C^\infty$ with derivatives of polynomial growth. Therefore, for any $m \in \mathbb{N}$, it exists $\nu \in \mathbb{N}^*$ sufficiently large, such that $f \in C_m^\nu([\mathbb{R}^+])$. We can then apply [4, Theorem 3.10] to get that for some functions $m, \ell : \mathbb{N}^* \to \mathbb{N}^*$, we have $\|P_T f - \hat{P}^{m,n} f\|_{0,L+\ell(\nu)} \leq C\|f\|_{m(\nu),L}n^{-2\nu}$ for $L \geq m(\nu)$, which gives the claim.

6. Simulations results

In order to present some numerical test, we first explain how to implement the approximations $\hat{P}^{2,n}$ and $\hat{P}^{3,n}$ defined respectively by (2.7) and (2.8) (let us recall here that $\hat{P}^{1,n}$ is the approximation obtained on the regular time grid $\Pi^0 = \{kT/n, 0 \leq k \leq n\}$). We consider a general case of a scheme that can be written as a function of the starting point, the time step, the Brownian increment and an independent random variable, i.e.

$$Q_1 f(x) = \mathbb{E}[\varphi(x, h_1, W_{h_1}, V)].$$

The second order scheme for the CIR (3.1) falls into this framework as well as the second order scheme for the Heston model (6.5) that we introduce below. As illustrated in [4] the approximation $\hat{P}^{2,n}$ is the simplest case for the implementation. It consists in the simulation of two starting schemes on the uniform time grid $\Pi^0$ and on the random grid : $\Pi^1 = \Pi^0 \cup \{\kappa T/n + kT/n^2, 1 \leq k' \leq n - 1\}$, where $\kappa$ is an independent uniform random variable on $\{0, \ldots, n - 1\}$. We denote by $\hat{X}^{n,0}$ the scheme on $\Pi^0$

$$\hat{X}^{n,0}_0 = x,$$

$$\hat{X}^{n,0}_{(k+1)h_1} = \varphi(\hat{X}^{n,0}_{kh_1}, h_1, W_{(k+1)h_1} - W_{kh_1}, V_k), \quad 0 \leq k \leq n - 1,$$

and by $\hat{X}^{n,1}$ the scheme on $\Pi^1$:

$$\hat{X}^{n,1}_{kh_1} = \hat{X}^{n,0}_{kh_1}, \quad 0 \leq k \leq \kappa,$$

$$\hat{X}^{n,1}_{kh_1+(k'+1)h_2} = \varphi(\hat{X}^{n,1}_{kh_1+k'h_2}, h_2, W_{kh_1+(k'+1)h_2} - W_{kh_1+k'h_2}, V_{n+k'}), \quad 0 \leq k' \leq n - 1,$$

$$\hat{X}^{n,1}_{(k+1)h_1} = \varphi(\hat{X}^{n,1}_{kh_1}, h_1, W_{(k+1)h_1} - W_{kh_1}, V_k), \quad \kappa + 1 \leq k \leq n - 1.$$ 

Here, $(V_k)_{k \geq 0}$ is an i.i.d. sequence with the same law as $V$. Finally, we can give the following probabilistic representation

$$\hat{P}^{2,n} f = Q_1^{[n]} f + n\mathbb{E}[Q_1^{n-(\kappa+1)}][Q_2^{[n]} - Q_1^{[n]} f]$$

$$+ \sum_{k'=0}^{n-1} \mathbb{E}[Q_1^{n-(k'+1)}][Q_2^{[k']} - Q_1^{[k']} f]$$. 
Let us stress here that it is crucial for the Monte-Carlo method to use the same underlying Brownian motion for \( \hat{X}_{T}^{n,0} \) and \( \hat{X}_{T}^{n,1} \). Thus, the variance of \( n \left( f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0}) \right) \) is quite moderate. It is shown in [4, Appendix A] that this variance is bounded when using the Euler scheme for an SDE with smooth coefficients. The theoretical analysis of the variance in our framework is beyond the scope of the paper. We only check numerically how it evolves with respect to \( n \) on our experiments, see Table 3 below.

The approximation \( \hat{P}^{3,n} \) is more involved. Let \( \kappa' \) be an independent uniform random variable on \( \{0, \ldots, n-1\} \). We define the scheme \( \hat{X}_{n,2}^{n,2} \):

\[
\hat{X}_{(k+1)h_1}^{n,2} = \varphi(\hat{X}_{k+1}^{n,2}, h_1, W_{(k+1)h_1} - W_{kh_1}, V_0), \quad \kappa + 1 \leq k \leq n - 1.
\]

This is the scheme obtained on the time grid \( \Pi^1 \cup \{\kappa T/n + \kappa'T/n + k''T/n^3, 1 \leq k'' \leq n - 1\} \). We have

\[
\sum_{k=0}^{n-1} Q_1^{[k]} f = \sum_{k=0}^{n-1} \left[ Q_1^{[k]} - Q_2^{[k]} \right] Q_1^{[k]} f = n^2 \mathbb{E}[f(\hat{X}_{T}^{n,2}) - f(\hat{X}_{T}^{n,1})].
\]

We now explain how to calculate the second term in (2.8). Let \( (\kappa_1, \kappa_2) \) be an independent random variable uniformly distributed on the set \( \{(k_1, k_2) : 0 \leq k_1 < k_2 < n\} \). We define:

\[
\hat{X}_{kh_1}^{n,3} = \hat{X}_{kh_1}^{n,0}, \quad 0 \leq k \leq \kappa_1,
\]

\[
\hat{X}_{\kappa_1 h_1 + (k'+1)h_2}^{n,3} = \varphi(\hat{X}_{\kappa_1 h_1 + k' h_2}^{n,3}, h_2, W_{\kappa_1 h_1 + (k'+1)h_2} - W_{\kappa_1 h_1 + k' h_2}, V_{3n+k'}), \quad 0 \leq k' \leq n - 1,
\]

\[
\hat{X}_{(k+1)h_1}^{n,3} = \varphi(\hat{X}_{kh_1}^{n,3}, h_1, W_{(k+1)h_1} - W_{kh_1}, V_k), \quad \kappa_1 + 1 \leq k \leq n - 1,
\]

and

\[
\hat{X}_{\kappa_2 h_1 + (k'+1)h_2}^{n,4} = \varphi(\hat{X}_{\kappa_2 h_1 + k' h_2}^{n,4}, h_2, W_{\kappa_2 h_1 + (k'+1)h_2} - W_{\kappa_2 h_1 + k' h_2}, V_{4n+k'}), \quad 0 \leq k' \leq n - 1,
\]

\[
\hat{X}_{(k+1)h_1}^{n,4} = \varphi(\hat{X}_{kh_1}^{n,4}, h_1, W_{(k+1)h_1} - W_{kh_1}, V_k), \quad \kappa_2 + 1 \leq k \leq n - 1,
\]

These schemes correspond respectively to the time grids \( \Pi^0 \cup \{\kappa_1 T/n + k'T/n^2, 1 \leq k' \leq n - 1\} \), \( \Pi^0 \cup \{\kappa_2 T/n + k'T/n^2, 1 \leq k' \leq n - 1\} \) and \( \Pi^0 \cup \{\kappa_1 T/n + k'T/n^2, 1 \leq k' \leq n - 1\} \cup \{\kappa_2 T/n + k'T/n^2, 1 \leq k' \leq n - 1\} \). We then get

\[
\hat{P}^{3,n} f = \mathbb{E}[f(\hat{X}_{T}^{n,0})] + n\mathbb{E}[f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0})] + n^2\mathbb{E}[f(\hat{X}_{T}^{n,2}) - f(\hat{X}_{T}^{n,1})].
\]
\[
+ \frac{n(n-1)}{2} \mathbb{E}[f(\hat{X}_T^{n,5}) - f(\hat{X}_T^{n,4}) - f(\hat{X}_T^{n,3}) + f(\hat{X}_T^{n,0})].
\]

6.1. Simulations result for the CIR process. In this subsection, we want to illustrate the convergence of the approximations \( \hat{P}_{2,n} \) and \( \hat{P}_{3,n} \), which together with the use of the second order scheme (3.1) guarantee respectively approximations of order four and six by Theorem 3.2. In order to calculate these approximations, we use Monte-Carlo estimators of (6.2) and (6.3), using independent samples for each expectation. The number of samples (up to \( 10^{11} \)) is such that we can neglect the statistical error. In Figures 1, 2 and 3 we plot the convergence in function of the time step for different parameters choices, taking advantage of the closed formula for the Laplace transform of the CIR process, see e.g. [3, Proposition 1.2.4]. The three numerical experiments test different levels of the ratio \( \sigma^2/4a \) in decreasing order. We observe that the slopes estimated on the log-log plots are close to 2, 4 and 6 respectively, so that they are in accordance with Theorem 3.2. Note however that Theorem 3.2 gives an asymptotic result for \( n \to \infty \), while we are restricted here to rather small values of \( n \) since we are using a large number of samples to kill the statistical error. In all the cases shown, the approximations of higher order outperform the one built with the simple second order scheme (3.1). Talking about accuracies, the fourth order approximation for \( n = 3 \) shows an absolute relative error of about 0.17% in the tests in Figures 1, and 2 and 0.02% in the one in Figure 3; the sixth order approximation already for \( n = 3 \) exhibits a relative error of 0.002% in each case studied.

![Values plot](A) ![Log-log plot](B)

**Figure 1.** Parameters: \( x = 0.0, a = 0.2, k = 0.5, \sigma = 0.65, f(z) = \exp(-10z) \) and \( T = 1 (\frac{\sigma^2}{2a} \approx 1.06) \). Graphic (A) shows the values of \( \hat{P}_{1,n} f, \hat{P}_{2,n} f, \hat{P}_{3,n} f \) as a function of the time step \( 1/n \) and the exact value. Graphic (B) draws \( \log(|\hat{P}_{i,n} f - P_T f|) \) in function of \( \log(1/n) \): the regressed slopes are 1.86, 3.93 and 5.87 for the second, fourth and sixth order respectively.

6.2. Simulations result for the Heston model. In this subsection, we want to test the second order scheme for the Heston model proposed by Alfonsi in [2] along with the approximations of order 4 and 6 obtained with combination of random grids. First, we recall the
Figure 2. Parameters: $x = 0.3$, $a = 0.4$, $k = 1$, $\sigma = 0.4$, $f(z) = \exp(-8z)$ and $T = 1 \left( \frac{\sigma^2}{2a} = 0.2 \right)$. Graphic (A) shows the values of $\hat{P}_1, n f$, $\hat{P}_2, n f$, $\hat{P}_3, n f$ as a function of the time step $1/n$ and the exact value. Graphic (B) draws $\log(|\hat{P}_i, n f - P_T f|)$ in function of $\log(1/n)$: the regressed slopes are 1.90, 3.93 and 5.77 for the second, fourth and sixth order respectively.

Figure 3. Parameters: $x = 10$, $a = 10$, $k = 1$, $\sigma = 0.23$, $f(z) = \exp(-z)$ and $T = 1 \left( \frac{\sigma^2}{2a} \approx 0.0026 \right)$. Graphic (A) shows the values of $\hat{P}_1, n f$, $\hat{P}_2, n f$, $\hat{P}_3, n f$ as a function of the time step $1/n$ and the exact value. Graphic (B) draws $\log(|\hat{P}_i, n f - P_T f|)$ in function of $\log(1/n)$: the regressed slopes are 1.96, 4.00 and 6.02 for the second, fourth and sixth order respectively.

couple of stochastic differential equations describing this model

\[
\begin{align*}
\quad &\quad \begin{cases} dS_t^{(x,s)} = rS_t^{(x,s)}dt + \sqrt{X_t}S_t^{(x,s)}(\rho dW_t + \sqrt{1-\rho^2}dZ_t), \quad S_0^{(x,s)} = s, \\
dX_t^x = (a - kX_t^x)dt + \sigma\sqrt{X_t^x}dW_t, \quad X_0^x = x, \end{cases} \\
\end{align*}
\]
where $W$ and $Z$ are two independent Brownian motions. We define the two following random variables

\[ S_1((x, s), h, Z_h) = (x, s \exp \left( \sqrt{x(1 - \rho^2)} Z_h \right)) \]

\[ S_2((x, s), h, W_h) = \left( \varphi(x, h, W_h), s \exp \left( (r - \frac{\rho}{\sigma}) h + \left( \frac{\rho}{\sigma^2} - \frac{1}{2} \right) x + \varphi(x, h, W_h) h + \frac{\rho}{\sigma} (\varphi(x, h, W_h) - x) \right) \right) \]

where $\varphi$ is defined by (3.2) and corresponds to the second order scheme for the CIR process.

We define as in [2] the second order scheme for (6.4) as follows

\[ \Phi((x, s), h, (W_h, Z_h), B) = \begin{cases} 
  S_2 \left( S_1((x, s), h, Z_h), h, W_h \right), & \text{if } B = 1, \\
  S_1 \left( S_2((x, s), h, W_h), h, Z_h \right), & \text{if } B = 0.
\end{cases} \quad (6.5) \]

where $B$ is an independent Bernoulli random variable of parameter 1/2.

To test the order of the approximations $\hat{P}^{2,n}$ and $\hat{P}^{3,n}$ boosting the second order scheme (6.5), we have calculated European put prices taking advantage of the existence of a semi closed formula for this option, see [9]. In Figure 4 we draw the convergence in function of the time step. Again, we noticed that the slopes obtained on the log-log plot are in line with the expected order of convergence. More importantly, we see that the correction terms of the approximations $\hat{P}^{2,n}$ and $\hat{P}^{3,n}$ really improves the precision. They respectively give relative errors of a 0.035% and 0.0023%, already for $n = 3$.

![Figure 4](image-url)
6.3. Optimized implementation of $\hat{\beta}_2^{2,n}$. The approximations $\hat{\beta}_2^{2,n}$ and $\hat{\beta}_3^{3,n}$ defined respectively by (6.2) and (6.3) involve respectively two and four expectations. The larger is $\nu$ the more expectations are involved in $\hat{\beta}_\nu^{\nu,n}$. Thus, for simplicity, independent samples were used by Alfonsi and Bally [4] to compute each term. However, it may be interesting to reuse some samples in order to spare computation time. This is what we investigate in this subsection.

Namely, Equation (6.2) leads naturally to the two following estimators of $P^{2,n}f$:

$$\Theta_1(M_1, M_2, n) = \frac{1}{M_1} \sum_{j=1}^{M_1} f((\hat{X}_T^{n,0})^{(j)}) + \frac{1}{M_2} \sum_{i=M_1+1}^{M_1+M_2} n \left( f((\hat{X}_T^{n,1})^{(i)}) - f((\hat{X}_T^{n,0})^{(i)}) \right),$$

$$\Theta_D(M_1, M_2, n) = \frac{1}{M_1} \sum_{j=1}^{M_1} f((\hat{X}_T^{n,0})^{(j)}) + \frac{1}{M_2} \sum_{i=1}^{M_2} n \left( f((\hat{X}_T^{n,1})^{(i)}) - f((\hat{X}_T^{n,0})^{(i)}) \right).$$

The first one takes independent samples and we call this estimator $\Theta_1$. This approach is the one used in [4]. In the second case, we reuse the first $M_1 \land M_2$ simulations of $f((\hat{X}_T^{n,0})^{(i)})$ in both sums. We call this estimator $\Theta_D$ to indicate the dependence between samples. In terms of variance, we have

$$\text{Var}(\Theta_1(M_1, M_2, n)) = \frac{\text{Var}(f(\hat{X}_T^{n,0}))}{M_1} + \frac{\text{Var}(n(f(\hat{X}_T^{n,1}) - f(\hat{X}_T^{n,0})))}{M_2},$$

$$\text{Var}(\Theta_D(M_1, M_2, n)) = \frac{\text{Var}(f(\hat{X}_T^{n,0}))}{M_1} + \frac{2 \text{Cov}(f(\hat{X}_T^{n,1}), n(f(\hat{X}_T^{n,1}) - f(\hat{X}_T^{n,0})))}{M_1 \lor M_2} + \frac{\text{Var}(n(f(\hat{X}_T^{n,1}) - f(\hat{X}_T^{n,0})))}{M_2}.$$
of \( f(X^T_{n,0}) \)). Taking the derivative at \( \hat{n} = 0 \), we get that \( M_1 \geq M_2 \) if \( \zeta \frac{\sigma^2_2(n)+2\Gamma(n)}{\sigma^2_4(n)+2\Gamma(n)} \geq 1 \), and \( M_1 < M_2 \) otherwise. When \( M_1 \geq M_2 \), the minimisation of \( g_D \) given \( \text{Var}(\Theta_D(M_1, M_2, n)) = \varepsilon^2 \) leads to

\[
\begin{align*}
M_{1,D} &= \frac{1}{\varepsilon^2} \left( \frac{\sigma^2_2(n) + 2\Gamma(n)}{\sigma^2_4(n) + 2\Gamma(n)} \right), \\
M_{2,D} &= \frac{1}{\varepsilon^2} \left( \sigma^2_2(n) + \sqrt{\frac{(\sigma^2_2(n)+2\Gamma(n))\sigma^2_4(n)}{\zeta-1}} \right).
\end{align*}
\]

(6.11)

We have similar formulas when \( M_1 < M_2 \). In all our numerical experiments below, we are in the case where \( \zeta \frac{\sigma^2_2(n)+2\Gamma(n)}{\sigma^2_4(n)+2\Gamma(n)} \geq 1 \) and thus taking \( M_1 \geq M_2 \) is optimal.

Now, we show the performance of the two estimators (6.6) and (6.7). To do this, we calculate the empirical variances \( \sigma^2_2(n) \), \( \sigma^2_4(n) \) and the empirical covariance \( \Gamma(n) \) on a small sampling, fix a desired precision \( \varepsilon = 1.96\sqrt{\text{Var}(\Theta(M_1, M_2, n))} \) for both the estimators, so that all the terms have roughly the same statistical error with a 95% confidence interval half-width equal to \( \varepsilon \). We show two tables in which we set the precision \( \varepsilon \) to \( 10^{-3} \). In Table 1, we have \( \sigma^2_2(n) \gg \sigma^2_4(n) \), while in Table 2, \( \sigma^2_2(n) \) is still larger than \( \sigma^2_4(n) \), but of the same order of magnitude.

| \( n \) | 2  | 3  | 4  | 5  |
|--------|----|----|----|----|
| \( \Theta_i \) | 63.04 | 96.15 | 131.84 | 165.80 |
| \( \Theta_d \) | 51.61 | 87.24 | 122.76 | 152.32 |

Table 1. Computation time (in seconds) needed by the Estimators \( \Theta_i \) and \( \Theta_d \) for a precision \( \varepsilon = 10^{-3} \). Test function: \( f(x, s) = (K - s)^+ \). Parameters: \( S_0 = 100 \), \( r = 0 \), \( x = 0.4 \), \( a = 0.4 \), \( k = 1 \), \( \sigma = 0.2 \), \( \rho = -0.3 \), \( T = 1 \), \( K = 100 \) \( (\sigma_{2a}^2 = 0.05) \).

| \( n \) | 2  | 3  | 4  | 5  |
|--------|----|----|----|----|
| \( \Theta_i \) | 59.50 | 102.13 | 148.45 | 193.41 |
| \( \Theta_d \) | 37.59 | 70.43 | 100.14 | 136.16 |

Table 2. Computation time (in seconds) needed by the Estimators \( \Theta_i \) and \( \Theta_d \) for a precision \( \varepsilon = 10^{-3} \). Test function: \( f(x, s) = (K - s)^+ \). Parameters: \( S_0 = 100 \), \( r = 0 \), \( x = 0.1 \), \( a = 0.1 \), \( k = 1 \), \( \sigma = 0.63 \), \( \rho = -0.3 \), \( T = 1 \), \( K = 100 \) \( (\sigma_{2a}^2 \approx 1.98) \).

We observe that we do not have a great gain in using \( \Theta_i \) when \( \sigma^2_2(n) \gg \sigma^2_4(n) \) (Table 1), while we save up to 30% of execution time, using \( \Theta_d \) instead of \( \Theta_i \), when \( \sigma^2_2(n) \) is of the same order of magnitude \( \sigma^2_4(n) \) (Table 2). Heuristically, this can be understood as follows: when \( \sigma^2_2(n) \) is of the same magnitude as \( \sigma^2_4(n) \), so are \( M_{1,1} \) and \( M_{2,1} \), which gives an important gain in reusing the simulation of the correction term. In any case, \( \Theta_D \) turns out to be faster for each choice of parameters, and therefore we recommend it at the expense of \( \Theta_i \).
6.4. **Comparison between the second and the fourth order approximation.** Subsections 6.1 and 6.2 have confirmed numerically the theoretical results obtained in this paper. However, they do not compare directly the computation time required by the different methods. We now present numerical tests that allow us to prove the real advantage of using the fourth order approximation $\hat{P}^{2,n}$ instead of the simple second order scheme. Namely, we compare the squared $L^2$ distance of the estimator $\Theta_d$ from the true value with the same distance between the estimator of $\hat{P}^{1,n^2}$ with the true value. We plot these quantities in function of the computation time needed. Note that $\hat{P}^{2,n}$ and $\hat{P}^{1,n^2}$ converges at a rate of $O(n^{-4})$ so that their bias have the same order of magnitude.

![Graphs showing $L^2$-square error in function of the execution time in seconds.](image)

**Figure 5.** $L^2$-square error in function of the execution time in seconds. Test function: $f(x,s) = (K-s)^+$. Parameters in graphic (A): $S_0 = 100$, $r = 0$, $x = 0.4$, $a = 0.4$, $k = 1$, $\sigma = 0.1$, $\rho = -0.3$, $T = 1$, $K = 100$.
Parameters in graphic (B): $S_0 = 100$, $r = 0$, $x = 0.1$, $a = 0.1$, $k = 1$, $\sigma = 0.63$, $\rho = -0.3$, $T = 1$, $K = 100$.

Figure 5 shows the results for the calculation of the price of a European put option in the Heston model with two different sets of parameters. In this numerical experience we set a precision $\varepsilon$ equal to $10^{-3}$. The empirical evidences show that the fourth order estimator $\Theta_d$ is the best choice, especially when the ratio $\frac{\sigma^2}{2a} \ll 1$ (Figure 5 (A)) where the performance of the fourth order estimator is unparalleled. For example, $\hat{P}^{2,3}$ is twice more accurate and more than twice faster than $\hat{P}^{1,9}$. Even in Figure 5 (B), where the ratio $\frac{\sigma^2}{2a}$ is larger and close to 2, the fourth order estimator $\Theta_d$ is more precise than the second order estimator and is faster from $n = 3$ onward. These experiments illustrate the outperformance of the boosted estimator $\hat{P}^{2,n}$ with respect to $\hat{P}^{1,n}$.

6.5. **Numerical experiments for $\sigma^2 > 4a$.** In the previous subsections, we have presented analyses to confirm numerically the theoretical rates of convergence of our approximations, and to assess their computational time. This is why we have only considered parameters such that $\sigma^2 \leq 4a$, since this condition is required in Theorem 3.2. However, it is possible to test numerically the relevance of the boosting technique using random grids when $\sigma^2 > 4a$. This is the purpose of this subsection. We first present the different schemes and then analyse...
numerically the variance of the correcting term. Then, we present the numerical bias of the approximation $\hat{P}^{2,n}$ for the CIR and Heston models.

6.5.1. The approximation schemes. In order to perform the numerical tests for $\sigma^2 > 4a$, we consider two different second order schemes for the CIR process. The first one is the second order scheme (3.7) presented in [2]. More precisely, we define

$$\varphi_A(x, t, \sqrt{t}N) = \varphi^u_A(x, t, \sqrt{t}N) \mathbb{1}_{x \geq K_2^A(t)} + \varphi^d_A(x, t, \sqrt{t}N) \mathbb{1}_{x < K_2^A(t)};$$

(6.12)

with

$$\varphi^u_A(x, t, \sqrt{t}N) = \varphi(x, t, -\sqrt{3}t) \mathbb{1}_{N < N^{-1}(1/6)} + \varphi(x, t, 0) \mathbb{1}_{N^{-1}(1/6) \leq N < \Phi^{-1}(5/6)}$$

$$+ \varphi(x, t, \sqrt{3}t) \mathbb{1}_{N \geq N^{-1}(5/6)};$$

$$\varphi^d_A(x, t, \sqrt{t}N) = \frac{\mathbb{E}[X^x]}{2(1 - \pi(t, x))} \mathbb{1}_{N < N^{-1}(1 - \pi(t, x))} + \frac{\mathbb{E}[X^x]}{2\pi(t, x)} \mathbb{1}_{N \geq N^{-1}(1 - \pi(t, x))},$$

where $\mathcal{N}$ is the cumulative distribution function of the standard normal distribution, $\pi(t, x) = \frac{1 - \frac{\pi(x, t, z)}{\pi(x, t, z)^2}}{\pi^2(x, t, z)^{-1}}$ and $K_2^A(t)$ is the function given by (3.8) with $A_Y = \sqrt{3}$. Here, we have written the scheme $\varphi_A$ as a function of the starting point $x$, the time step $t$ and the Brownian increment $\sqrt{t}N$. When computing $n\mathbb{E}[(f(\hat{X}_T^{n,1}) - f(\hat{X}_T^{n,0}))]$ by Monte-Carlo, we use the same Brownian path to sample $\hat{X}_T^{n,0}$ and $\hat{X}_T^{n,1}$, as explained at the beginning of Section 6. Thus, there is a strong dependence between these schemes.

We present also another scheme that corresponds to other choices of $Y$ and $\hat{X}^{x,d}$ in (3.7). We use a distribution that is pretty similar to a Gaussian distribution over the threshold, and a scaled beta distribution below. Thus, we define

$$\varphi_B(x, t, \sqrt{t}N) = \varphi^u_B(x, t, \sqrt{t}N) \mathbb{1}_{x \geq K_2^B(t)} + \varphi^d_B(x, t, \sqrt{t}N) \mathbb{1}_{x < K_2^B(t)};$$

(6.13)

with

$$\varphi^u_B(x, t, \sqrt{t}N) = \varphi(x, t, -z_2) \mathbb{1}_{N \leq -c_2} + \varphi(x, t, -z_1) \mathbb{1}_{-c_2 \leq N \leq -c_1} + \varphi(x, t, N) \mathbb{1}_{-c_1 \leq N < c_1}$$

$$+ \varphi(x, t, z_1) \mathbb{1}_{c_1 \leq N \leq c_2} + \varphi(x, t, z_2) \mathbb{1}_{N > c_2};$$

$$\varphi^d_B(x, t, \sqrt{t}N) = \frac{\mathbb{E}[X^x]}{2\pi(t, x)} (\mathcal{N}(N)) \mathbb{1}_{N \geq N^{-1}(1 - \pi(t, x)),}$$

where $z_1 = 2.7523451704150586$, $z_2 = 3.5$, $c_1 = 2.58$, $c_2 = 3.106520327375868$, and $K_2^B(t)$ is the function given by (3.8) with $A_Y = 3.5$. Here, we have fixed the values of $c_1$ and $z_2$, and we have numerically calculated $c_2$ and $z_1$ to have $\mathbb{E}[Y^2] = \mathbb{E}[N^2]$ and $\mathbb{E}[Y^4] = \mathbb{E}[N^4]$ with

$$Y = -z_2 \mathbb{1}_{N \leq -c_2} - z_1 \mathbb{1}_{-c_2 \leq N \leq -c_1} + N \mathbb{1}_{-c_1 \leq N \leq c_1} + z_1 \mathbb{1}_{c_1 \leq N \leq c_2} + z_2 \mathbb{1}_{c_2 \leq N}. $$

The random variable $\varphi^d_B(x, t, \sqrt{t}N)$ has the same two first moments as $X^x_t$, and we can prove following the same arguments as [2, Theorem 2.8] that $\varphi_B(x, t, \sqrt{t}N)$ is a second order scheme for the weak error.
6.5.2. Numerical study of the variance of the correcting term $n \left( f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0}) \right)$. We now analyse the variance of the corrections terms of the correcting term $n \left( f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0}) \right)$ in function of the number $n$ of discretization steps, when we use the different schemes (6.12) and (6.13). We start with an example with $\sigma^2 < 4a$ for which $\varphi$ is still defined and $\varphi_A$ (resp. $\varphi_B$) does not use the auxiliary scheme $\varphi_A^d$ (resp. $\varphi_B^d$) since $K_A(t) = K_B(t) = 0$ in this case. We observe in Table 3 that the scheme $\varphi_A$ leads to a value of $\text{Var}(n(f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0})))$ that is more than 20 times as large as that the one obtained using $\varphi$. Besides, the variance given by the scheme $\varphi_A$ increases quite linearly with $n$, while the one obtained with $\varphi$ seems to be bounded and to decrease with $n$. One heuristic explanation for this is that $\varphi_A$ is discrete scheme, which increases the strong error between the scheme on the fine grid $\Pi^1$ and the scheme on the coarse grid $\Pi^0$. Considering the scheme $\varphi_B$ that mixes Gaussian and discrete distributions leads to a much smaller variance than is rather close to the one of the scheme $\varphi$. However, as $n$ gets large, we see that the variance does not decrease in contrast to the scheme $\varphi$.

|       | $n = 2$          | $n = 4$          | $n = 8$          | $n = 16$         | $n = 32$         |
|-------|------------------|------------------|------------------|------------------|------------------|
| $\varphi$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ |
| 95% prec. | 23.86e-4 | 17.43e-4 | 9.35e-4 | 4.85e-4 | 2.49e-4 |
| 95% prec. | 3.2e-6 | 3.7e-6 | 2.8e-6 | 2.1e-6 | 1.5e-6 |

| $\varphi_A$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ |
| 95% prec. | 4.807e-2 | 10.870e-2 | 22.493e-2 | 45.437e-2 | 91.219e-2 |
| 95% prec. | 2.4-5 | 5.2e-5 | 11.1e-5 | 22.9e-5 | 46.3e-5 |

| $\varphi_B$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ | $\sigma_2(n)$ |
| 95% prec. | 24.17e-4 | 18.37e-4 | 11.78e-4 | 10.27e-4 | 13.85e-4 |
| 95% prec. | 3.2e-6 | 3.7e-6 | 2.9e-6 | 3.0e-6 | 4.5e-6 |

Table 3. $\sigma_2^2(n) = \text{Var}(n(f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0})))$ for the different schemes, with $10^8$ samples and 95% confidence interval precision. Test function: $f(x) = \exp(-10x)$. Parameters: $x = 0.2$, $a = 0.2$, $k = 0.5$, $\sigma = 0.5$, $T = 1$ ($\frac{\sigma^2}{2a} = 0.625$).

We now consider a case with $\sigma^2 > 4a$ so that the schemes $\varphi_A$ and $\varphi_B$ switch around their threshold. The scheme $\varphi$ is no longer defined. In Table 4, we observe a huge increase of the variance in time steps with respect to Table 3. We now observe that the variances grow almost linearly with respect to $n$. Again, this can be explained heuristically by the switching that increases the strong error between the schemes on the fine grid $\Pi^1$ and the coarse grid $\Pi^0$. The rather high values of the variance obtained with the scheme $\varphi_A$ makes the boosting technique using random grids less interesting in practice from a computational point of view. In contrast, the scheme $\varphi_B$ produces much lower variances and the Monte-Carlo estimator of $\tilde{P}^{2,n}f$ is more competitive.

We have plotted in Figure 6 the convergence of the estimators of the Monte-Carlo estimators $\tilde{P}^{1,n}f$ and $\tilde{P}^{2,n}f$ for the schemes $\varphi_A$ and $\varphi_B$. We note that in all our experiments, $\tilde{P}^{2,n}f$ gives a better approximation than $\tilde{P}^{1,n}f$, though there is no theoretical guarantee of that. However, the improvement is not as good as for $\sigma^2 \leq 4a$. We know that $\tilde{P}^{1,n}f$ leads to an asymptotic weak error of order 2: the estimated rate of convergence obtained by regression are below since we consider rather small values of $n$ and are not in the asymptotic regime. We have instead no theoretical guarantee that $\tilde{P}^{2,n}f$ gives an asymptotic weak error of order 4.
Table 4. $\sigma^2_4(n) = \text{Var}(n(f(\hat{X}_{T}^{n,1}) - f(\hat{X}_{T}^{n,0})))$ with $10^8$ samples and 95% confidence interval precision. Test function: $f(x) = \exp(-10x)$. Parameters: $x = 0.2$, $a = 0.2$, $k = 0.5$, $\sigma = 1.5$, $T = 1$ ($\sigma^2 = 5.625$).

The estimated rates are quite far from this value, indicating that a fourth order of convergence may not hold. To sum up, even if $\hat{P}^{2,n}f$ is still more accurate than $\hat{P}^{1,n}f$ for $\sigma^2 > 4a$, it does not lead to obvious computational gains.
6.5.4. *Simulations in the Heston model.* We present now some numerical tests for Heston model and consider three different schemes that are well defined for any $\sigma \geq 0$:

- $\Phi_A$ is the scheme (6.5) where $\varphi_A(x,h,W_h)$ is used instead of $\varphi(x,h,W_h)$,
- $\Phi_B$ is the scheme (6.5) where $\varphi_B(x,h,W_h)$ is used instead of $\varphi(x,h,W_h)$,
- $\Phi_E$ is the scheme (6.5) where the exact scheme $X^n_h$ (see, e.g. [3, Proposition 3.1.1]) is used instead of $\varphi(x,h,W_h)$.

We start by comparing the variance of the correcting terms with the different schemes. In Table 5, we consider a case with $\sigma^2 < 4a$ and also include the variance for the scheme $\Phi$ given by (6.5). We remark that the variances of the correction term for the standard scheme $\Phi$ and for the scheme $\Phi_E$ appear to be bounded. In contrast, the variance for the schemes $\Phi_A$ and $\Phi_B$ tends to increase with $n$: the variance is very high for $\Phi_A$ while the one produced by $\Phi_B$ remains close to the one of $\Phi$ and $\Phi_E$. Table 6 deals with a case with $\sigma^2 > 4a$ for which variances are much higher. We observe an approximately linear growth of the variance of the correction term for the schemes $\Phi_A$ and $\Phi_B$. The variance produced by the scheme $\Phi_E$ also increases, but in a much moderate way.

| $n$ ($n=2$) | $n=4$ | $n=8$ | $n=16$ | $n=32$ |
|-------------|--------|--------|--------|--------|
| $\Phi$      | $\sigma^2(n)$ | 33.252 | 41.962 | 46.159 | 48.273 | 49.385 |
|             | 95% prec.      | 0.024  | 0.029  | 0.033  | 0.035  | 0.037  |
| $\Phi_A$    | $\sigma^2(n)$ | 450.95 | 973.82 | 1976.53| 3984.64| 8014.19|
|             | 95% prec.      | 0.20   | 0.40   | 0.83   | 1.70   | 3.47   |
| $\Phi_B$    | $\sigma^2(n)$ | 33.702 | 43.116 | 48.606 | 53.373 | 59.760 |
|             | 95% prec.      | 0.025  | 0.031  | 0.037  | 0.044  | 0.059  |
| $\Phi_E$    | $\sigma^2(n)$ | 51.99  | 53.93  | 52.46  | 51.47  | 50.99  |
|             | 95% prec.      | 0.032  | 0.034  | 0.036  | 0.037  | 0.037  |

Table 5. $\sigma^2(n) = \text{Var}(n(f(\hat{X}^{n,1}_T,\hat{S}^{n,1}_T) - f(\hat{X}^{n,0}_T,\hat{S}^{n,0}_T)))$ with $10^8$ samples and 95% confidence interval precision. Test function: $f(x,s) = (K - s)^+$. Parameters: $S_0 = 100$, $r = 0$, $x = 0.2$, $a = 0.2$, $k = 1.0$, $\sigma = 0.5$, $\rho = -0.7$, $T = 1$, $K = 105$ ($\frac{\sigma^2}{2a} = 0.625$).

We now turn to the convergence of the Monte-Carlo estimators. We have plotted in Figure 7, for the same set of parameters as in Table 6, the behavior of $\hat{P}^{1,n}f$ and $\hat{P}^{2,n}f$ for the schemes $\Phi_B$ and $\Phi_E$. We have discarded the scheme $\Phi_A$ that produces a too large variance for the correcting term. As for the CIR diffusion, we note that $\hat{P}^{2,n}f$ gives a better approximation than $\hat{P}^{1,n}f$ but the bias does not seem to be of order 4. For the scheme $\Phi_B$, the improvement is moderate, and do not really compensate the computational effort of calculating the correcting term. Instead, for the scheme $\Phi_E$, the improvement is rather significant, making the approximation $\hat{P}^{2,n}f$ interesting from a computational point of view with respect to $\hat{P}^{1,n}f$. Also, the estimated rate of convergence is much higher and not so far from 4. A dedicated theoretical study of $\hat{P}^{2,n}f$ with the scheme $\Phi_E$ is left for further research.
Table 6. $\sigma^2(n) = \text{Var}(n(f(\hat{X}^{n,1}_T, \hat{S}^{n,1}_T) - f(\hat{X}^{n,0}_T, \hat{S}^{n,0}_T)))$ with $10^8$ samples and 95% confidence interval precision. Test function: $f(x, s) = (K - s)^+$. Parameters: $S_0 = 100, r = 0, x = 0.2, a = 0.2, k = 1.0, \sigma = 1.5, \rho = -0.7, T = 1, K = 105 (\sigma^2/2a = 5.625)$.

Appendix A. Proofs of Section 4

Proof of Lemma 4.1. (1) Let $f \in \mathcal{P}_L(\mathbb{R}_+)$. We have $X_0(t, x) = \sum_{i=0}^{j} \binom{j}{i} ((a-\sigma^2/4)\psi_k(t)) j-i e^{-kt} x^i$ and thus

$$f(X_0(t, x)) = \sum_{j=0}^{L} a_j \sum_{i=0}^{j} \binom{j}{i} ((a-\sigma^2/4)\psi_k(t)) j-i e^{-kt} x^i.$$ 

Therefore, $f(X_0(t, \cdot)) \in \mathcal{P}_L(\mathbb{R}_+)$ and we have

$$\|f(X_0(t, \cdot))\| \leq \sum_{j=0}^{L} |a_j| \sum_{i=0}^{j} \binom{j}{i} (|a-\sigma^2/4|\psi_k(t)) j-i e^{-kt} = \sum_{j=0}^{L} |a_j| \Xi_0(t)^j,$$

with $\Xi_0(t) = e^{-kt} + |a-\sigma^2/4|\psi_k(t)$. For $k \geq 0$, we have $0 \leq \psi_k(t) \leq t$ and thus $\Xi_0(t) \leq (1+|a-\sigma^2/4|t)$. For $k < 0$, we have $\Xi_0(t) = e^{-kt}(1+|a-\sigma^2/4|\psi_k(t)) \leq e^{-kt}(1+|a-\sigma^2/4|t)$. Since $(1+|a-\sigma^2/4|t)^L \leq 1 + t \sum_{i=1}^{j} \binom{j}{i} |a-\sigma^2/4|(1 \vee T)^i \leq 1 + t(1+|a-\sigma^2/4|(1 \vee T))^L$, we get $\Xi_0(t)^j \leq (1 \vee e^{-kLt})(1 + (1+|a-\sigma^2/4|(1 \vee T))^L)$ for $j \in \{0, \ldots, L\}$ and then

$$\|f(X_0(t, \cdot))\| \leq (1 \vee e^{-kLt})(1 + (1+|a-\sigma^2/4|(1 \vee T))^L)\|f\|,$$

which gives the claim with $C_{X_0} = 1 + |a-\sigma^2/4|(1 \vee T)$.

(2) Since $Y$ is a symmetric random variable, we have

$$\mathbb{E}[f(X_1(\sqrt{t}Y, x))] = \sum_{j=0}^{L} a_j \mathbb{E}[X_1(\sqrt{t}Y, x)^j] = \sum_{j=0}^{L} a_j \sum_{i=0}^{2j} \binom{2j}{i} \left(\frac{\sigma \sqrt{t}}{2}\right)^{2j-i} \mathbb{E}[Y^{2j-i}]x^i/2$$

$$= \sum_{j=0}^{L} a_j \sum_{i=0}^{j} \binom{2j}{2i} \left(\frac{\sigma \sqrt{t}}{4}\right)^{j-i} \mathbb{E}[Y^{2(j-i)}]x^i$$

$$= \sum_{j=0}^{L} a_j x^j + t \sum_{j=0}^{L} a_j \sum_{i=0}^{j-1} \binom{2j}{2i} \left(\frac{\sigma^2}{4}\right)^{j-i} t^{j-i-1}\mathbb{E}[Y^{2(j-i)}]x^i.$$
Figure 7. Test function: \( f(x, s) = (K - s)^+ \). Parameters: \( S_0 = 100, r = 0, x = 0.2, a = 0.2, k = 1, \sigma = 1.5, \rho = -0.7, T = 1, K = 105 (\sigma^2 = 5.625) \). Statistical precision \( \varepsilon = 5e-4 \). Left graphics show the values of \( \hat{P}_{1,n,f} \), \( \hat{P}_{2,n,f} \) as a function of the time step \( 1/n \) and the exact value. Right graphics draw \( \log(|\hat{P}_{i,n,f} - P_{T,f}|) \) in function of \( \log(1/n) \): for the scheme \( \Phi_B \) (resp. \( \Phi_E \)) the regressed slopes are 0.90 (resp. 1.28) and 2.04 (resp. 2.40) for the second and fourth order respectively.

This proves that \( \mathbb{E}[f(X_1(\sqrt{t}Y, \cdot))] \in \mathcal{P}_L(\mathbb{R}_+) \). We note that \( \mathbb{E}[Y^{2j}] \geq 1 \) by Hölder inequality since \( \mathbb{E}[Y^2] = 1 \), and thus \( \mathbb{E}[Y^{2j}] \leq \mathbb{E}[Y^{2L}] \) for \( j \in \{0, \ldots, L\} \). We get

\[
\|f(X_1(\sqrt{t}Y, \cdot))\| \leq \|f\| + t\mathbb{E}[Y^{2L}] \sum_{j=0}^{L} |a_j| \sum_{i=0}^{j-1} \binom{2j}{2i} \left( \frac{\sigma^2}{4} \right)^{j-i} (1 \vee T)^{j-i} \\
\leq \|f\| \left( 1 + t\mathbb{E}[Y^{2L}] \left( 1 + \frac{\sigma}{2} \sqrt{1 \vee T} \right)^{2L} \right),
\]

since \( \sum_{i=0}^{j-1} \binom{2j}{2i} \left( \frac{\sigma^2}{4} \right)^{j-i} (1 \vee T)^{j-i} \leq \left( 1 + \frac{\sigma}{2} \sqrt{1 \vee T} \right)^{2j} \leq \left( 1 + \frac{\sigma}{2} \sqrt{1 \vee T} \right)^{2L} \). This gives the claim with \( C_{X_1} = \left( 1 + \frac{\sigma}{2} \sqrt{1 \vee T} \right)^2 \).
Proof of Lemma 4.2. We have \( \tilde{u}_0(t, x) = 1 \), and in the case \( m = 1 \), we have \( \tilde{u}_1(t, x) = x + \int_0^t (a - k\tilde{u}_1(s, x))ds \) that has the solution:

\[
\tilde{u}_1(t, x) = xe^{-kt} + a\psi_k(t)
\]

where \( \psi_k(t) = \frac{1-e^{-kt}}{k} \) if \( k \neq 0 \) and \( \psi_k(t) = t \) otherwise. This gives the claim for \( m = 1 \) with \( \tilde{u}_{0,1} = a\psi_k(t) \) and \( \tilde{u}_{1,1} = e^{-kt} \). We then prove the result by induction and consider \( m \geq 2 \). Using Itô formula and taking the expected value, one has \( \partial_t\tilde{u}_m(t, x) = (am + \sigma^2 m(m - 1)/2)\tilde{u}_{m-1}(t, x) - km\tilde{u}_m(t, x) \). Hence, we have

\[
\tilde{u}_m(t, x) = (e^{-kt})^m \left( x^m + \int_0^t (am + \sigma^2 m(m - 1)/2)(e^{ks})^m\tilde{u}_{m-1}(s, x)ds \right),
\]

and we get the following induction relations that give us the representation (4.3)

\[
\begin{aligned}
\tilde{u}_{j,m}(t) &= (e^{-kt})^m \int_0^t (am + \sigma^2 m(m - 1)/2)(e^{ks})^m\tilde{u}_{j,m-1}(s, x)ds, \quad 0 \leq j \leq m - 1, \\
\tilde{u}_{m,m}(t) &= (e^{-kt})^m.
\end{aligned}
\]

Let \( f \in \mathfrak{F}_L(\mathbb{R}_+) \). We clearly get from the preceding result that \( \mathbb{E}[f(X_t)] \in \mathfrak{F}_L(\mathbb{R}_+) \) and

\[
\|\mathbb{E}[f(X_t)]\| \leq \sum_{m=0}^{L} |a_m| \sum_{j=0}^{m} |\tilde{u}_{j,m}(t)| \leq C_{\text{cir}}(L, T)\|f\|. \quad \Box
\]

Appendix B. Proofs of Section 5

Proof of Lemma 5.1. Properties (1)–(3) are straightforward, and we prove only (4)–(6).

(4) We use the fact that \( 1 + x^L \leq 2(1 + x^{L+1}) \) for \( x \geq 0 \), hence

\[
\max_{j \in \{0,\ldots,m\}} \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^{L+1}} \leq 2 \max_{j \in \{0,\ldots,m\}} \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^L}.
\]

(5) Let \( f \in C_{\text{pol}}^m(\mathbb{R}_+) \). We will use the fact that for all \( x \geq 0 \), \( (1 + x)(1 + x^L) \leq 3(1 + x^{L+1}) \) so

\[
\sup_{x \geq 0} \frac{x|f^{(j)}(x)|}{1 + x^{L+1}} \leq 3 \sup_{x \geq 0} \frac{x}{1 + x} \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^L} = 3 \sup_{x \geq 0} \frac{|f^{(j)}(x)|}{1 + x^L}.
\]

Now, we use the Leibniz rule on \( M_1f \) and get \( (xf(x))^{(j)} = jf^{(j-1)}(x) + xf^{(j)}(x) \), so

\[
\sup_{x \geq 0} \frac{|(xf(x))^{(j)}|}{1 + x^{L+1}} \leq j \sup_{x \geq 0} \frac{|f^{(j-1)}(x)|}{1 + x^{L+1}} + \sup_{x \geq 0} \frac{x|f^{(j)}(x)|}{1 + x^{L+1}}.
\]

Maximizing both sides on \( j \in \{0,\ldots,m\} \) and using the previous inequality gives \( \|M_1f\|_{m,L+1} \leq m\|f\|_{m-1,L+1} + 3\|f\|_{m,L} \). We get the bound by using properties (2) and (4).

(6) We have \( \|L(f)|_{m,L+1} \leq a\|f\|_{m,L+1} + (2m + 3)|k\|f\|_{m,L} + \frac{\sigma^2}{2}\|f\|_{m,L} \) by using the property (5). We get the estimate by using (3), (4) and (2). The other estimate for \( V_0^2/2 \) is obtained by taking \( a = \sigma^2/4 \) and \( k = 0 \), while the one for \( V_0 \) follows by using the same arguments. \( \Box \)
Proof of Lemma 5.5. For $x > 0$, we have
\[
\psi_g'(x) = \left(1 + \frac{\beta}{2\sqrt{x}}\right) g'(x + \beta\sqrt{x} + \gamma) + \left(1 - \frac{\beta}{2\sqrt{x}}\right) g'(x - \beta\sqrt{x} + \gamma)
\]
\[
= g'(x + \beta\sqrt{x} + \gamma) + g'(x - \beta\sqrt{x} + \gamma) + \beta^2 \int_0^1 g''(x + \beta(2u-1)\sqrt{x} + \gamma) du,
\]
since $\frac{d}{dx}g'(x + \beta(2u-1)\sqrt{x} + \gamma) = 2\beta\sqrt{x}g''(x + \beta(2u-1)\sqrt{x} + \gamma)$. Clearly, this derivative is continuous at 0 which shows that $\psi_g$ is $C^1$.

We are now in position to prove (5.5) by induction on $n$. It is true for $n=0,1$. We assume that it is true for $n$. Then, we get by using the case $n=1$, differentiating (5.5) and an integration by parts for the fourth term:
\[
\psi_g^{(n+1)}(x) = \psi_g^{(n+1)}(x) + \beta^2 \int_0^1 g^{(n+2)}(x + \beta(2u-1)\sqrt{x} + \gamma) du
\]
\[
+ \sum_{j=1}^n \binom{n}{j} \beta^{2j} \left( \int_0^1 g^{(n+j+1)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^{j-1}}{(j-1)!} du \right)
\]
\[
+ \beta^2 \int_0^1 g^{(n+j+2)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^j}{j!} du.
\]
We then reorganize the terms as follows
\[
\psi_g^{(n+1)}(x) = \psi_g^{(n+1)}(x) + \beta^2 \int_0^1 g^{(n+2)}(x + \beta(2u-1)\sqrt{x} + \gamma) du
\]
\[
+ \beta^{2n+2} \int_0^1 g^{(2n+2)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^n}{n!} du
\]
\[
+ \sum_{j=2}^n \binom{n}{j} \beta^{2j} \left( \int_0^1 g^{(n+j+1)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^{j-1}}{(j-1)!} du \right)
\]
\[
+ \sum_{j=1}^{n-1} \binom{n}{j} \beta^{2j+2} \left( \int_0^1 g^{(n+j+2)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^j}{j!} du \right).
\]
The last sum is equal to $\sum_{j=2}^n \binom{n}{j-1} \beta^{2j} \left( \int_0^1 g^{(n+j+1)}(x + \beta(2u-1)\sqrt{x} + \gamma) \frac{(u-u^2)^{j-1}}{(j-1)!} du \right)$ by changing $j$ to $j-1$, and we conclude by using that $\binom{n}{j-1} = \binom{n}{j}$.

Proof of Corollary 5.6. We use (5.5) with $\gamma = \beta^2/4$. We first notice that
\[
|\psi_g^{(n)}(x)| \leq \|g\|_{n,L}(2 + (\sqrt{x} + \beta/2)^{2L} + (\sqrt{x} - \beta/2)^{2L})
\]
\[
= \|g\|_{n,L} \left(2 + 2x^L + 2\sum_{i=1}^L \binom{2L}{2i} (\beta/2)^{2i} x^{L-i} \right).
\]
Using that $x^i \leq 1 + x^L$ for $0 \leq i \leq L - 1$, we get
\[
|\psi_g^{(n)}(x)| \leq 2\|g\|_{n,L}(1 + x^L) \sum_{i=0}^L \binom{2L}{2i} (\beta/2)^{2i} = \|g\|_{n,L}(1 + x^L)((1 + \beta/2)^{2L} + (1 - \beta/2)^{2L}).
\]
For the other terms, we use that for \( u \in [0, 1] \), \( x \geq 0 \) and \( j \in \{1, \ldots, n\} \),
\[
|g^{(n+j)}(x + \beta(2u - 1)\sqrt{x} + \beta^2/4)| \leq \|g\|_{2n,L}(1 + (x + \beta(2u - 1)\sqrt{x} + \beta^2/4)L) \\
\leq \|g\|_{2n,L}(1 + (\sqrt{x} + \beta/2)^{2L}).
\]
We again expand \((\sqrt{x} + \beta/2)^L = x^L + \sum_{i=1}^{2L} \binom{2L}{i} (\beta/2)^i x^{(L-i)/2}\) and use that \(x^{(L-i)/2} \leq 1 + x^L\) to get
\[
|g^{(n+j)}(x + \beta(2u - 1)\sqrt{x} + \beta^2/4)| \leq \|g\|_{2n,L}(1 + \beta/2)^{2L}(1 + x^L).
\]
Besides, we have \(u - u^2 \leq 1/4\) for \( u \in [0, 1] \) and thus \(\int_0^1 (u - u^2) du \leq \frac{1}{16}\), which gives
\[
\sup_{x \geq 0} \frac{|\psi_{n,j}(x)|}{1+x^L} \leq \tilde{C}(\beta) \text{ with}
\]
\[
\tilde{C}(\beta) = \|g\|_{n,L}((1 + \beta/2)^{2L} + (1 - \beta/2)^{2L}) + \|g\|_{2n,L}(1 + \beta/2)^{2L}\sum_{j=1}^{n} \binom{n}{j} \left(\frac{\beta}{4}\right)^j \\
= \|g\|_{n,L}((1 + \beta/2)^{2L} + (1 - \beta/2)^{2L}) + \|g\|_{2n,L}(1 + \beta/2)^{2L}(1 + \beta^2/4)^n \\
\leq \|g\|_{2n,L}((1 + \beta/2)^{2L} + (1 - \beta/2)^{2L} + (1 + \beta/2)^{2L}(1 + \beta^2/4)^m) = C_{\beta,m,L}\|g\|_{2n,L},
\]
which gives the claim.

\[ \square \]

**Appendix C. Assumption (H1) for symmetric random variables**

**Theorem C.1.** Let \( \eta : \mathbb{R} \to \mathbb{R}_+ \) be a \( C^\infty \) even function. Then, \( \eta_m^* \geq 0 \) for all \( m \in \mathbb{N}_* \) if and only if \( \eta(\sqrt{x}) \) is the Laplace transform of a finite positive Borel measure \( \mu \) on \([0, \infty)\), i.e. \( \eta(\sqrt{x}) = \int_0^\infty e^{-tx} \mu(dt) \) for all \( x \in \mathbb{R}_+ \).

\[ \text{Proof.} \] We start to prove that \( \eta_m^* \geq 0 \) for all \( m \in \mathbb{N}_* \) implies \( \eta(\sqrt{x}) = \int_0^\infty e^{-tx} \mu(dt) \) for all \( x \in \mathbb{R} \). To prove this, we use Bernstein’s Theorem for completely monotone functions (see e.g. [16, Theorem 12a p. 160]) and show that for all \( m \in \mathbb{N} \) and \( x \in \mathbb{R}_+ \), \((-1)^m \partial_x^m[\eta(\sqrt{x})] \geq 0\). To do so, we prove by induction on \( m \) the representation
\[
\partial_x^m[\eta(\sqrt{x})] = -\frac{(m-1)!}{2^{2m-1}} x^{-m} \sum_{j=1}^{m} c_{j,m} x^j \eta^{(j)}(\sqrt{x}) = (-1)^m \frac{(m-1)!}{2^{2m-1}} x^{-m} \eta_m^*(\sqrt{x}).
\]
For \( m = 1 \), we have \( \eta_1^*(\sqrt{x}) = c_{1,1} \sqrt{x} \eta(\sqrt{x}) \) and the representation holds from \( \partial_x[\eta(\sqrt{x})] = \frac{1}{2\sqrt{x}} \eta'(\sqrt{x}) = -\frac{1}{2x} \eta(\sqrt{x}) \) using that \( c_{1,1} = -1 \). Now, let \( m \geq 2 \) and suppose the representation is true for \( m - 1 \), so
\[
\partial_x^m[\eta(\sqrt{x})] = \partial_x(\partial_x^{m-1}[\eta(\sqrt{x})]) = \partial_x \left( -\frac{(m-2)!}{2^{2m-3}} x^{-(m-1)} \sum_{j=1}^{m-1} c_{j,m-1} x^j \eta^{(j)}(\sqrt{x}) \right).
\]
Differentiating and using that \( \partial_x(x^j \eta^{(j)}(\sqrt{x})) = \frac{1}{2x} (j x^j \eta^{(j)}(\sqrt{x}) + x^{j+1} \eta^{(j+1)}(\sqrt{x})) \), we get
\[
\partial_x^m[\eta(\sqrt{x})] = -\frac{(m-2)!}{2^{2m-3}} \left( -\frac{m-1}{x^m} \sum_{j=1}^{m-1} c_{j,m-1} x^j \eta^{(j)}(\sqrt{x}) \\
+ \frac{1}{2x^m} \sum_{j=1}^{m-1} c_{j,m-1} (j x^j \eta^{(j)}(\sqrt{x}) + x^{j+1} \eta^{(j+1)}(\sqrt{x})) \right)
\]
\[
\begin{align*}
&= -\frac{(m-2)!}{2^{m-3}} x^{-m} \left( \frac{1}{2} - m - 1 \right) c_{1,m-1} x^{\frac{1}{2}} \eta^{(1)}(\sqrt{x}) \\
&\quad + \sum_{j=1}^{m-1} \left( \left( \frac{j}{2} - m + 1 \right) c_{j,m-1} + \frac{1}{2} c_{j-1,m-1} \right) x^{\frac{j}{2}} \eta^{(j)}(\sqrt{x}) \\
&\quad + \frac{1}{2} c_{m-1,m-1} x^{\frac{m}{2}} \eta^{(m)}(\sqrt{x}) \\
&= -\frac{(m-1)!}{2^{m-1}} x^{-m} \left( \frac{2}{m-1} - 4 \right) c_{1,m-1} x^{\frac{1}{2}} \eta^{(1)}(\sqrt{x}) \\
&\quad + \sum_{j=1}^{m-1} \left( \left( \frac{2j}{m-1} - 4 \right) c_{j,m-1} + \frac{2}{m-1} c_{j-1,m-1} \right) x^{\frac{j}{2}} \eta^{(j)}(\sqrt{x}) \\
&\quad + \frac{2}{m-1} c_{m-1,m-1} x^{\frac{m}{2}} \eta^{(m)}(\sqrt{x})
\end{align*}
\]
and we conclude using the recursion formula (5.19) for \( c_{j,m} \).

We now assume that \( \eta(\sqrt{x}) = \int_0^\infty e^{-tx} \mu(dt) \) and show that \( \eta^*_m \geq 0 \) for all \( m \geq 1 \). We define \( \eta_g(x) = e^{-x^2} \) and consider for all \( t > 0 \) the function \( \eta_t(x) = e^{-tx^2} \). We remark that for all \( t > 0 \), \( \eta_t(x) = \eta_g(h_t(x)) \) with \( h_t(x) = \sqrt{2t}x \) and so we can write by Lemma 5.14

\[
(\eta_t)_m^* = (-1)^{m-1} \sum_{j=1}^{m} c_{j,m} x^j \eta_t^{(j)}(x) = (-1)^{m-1} \sum_{j=1}^{m} c_{j,m}(\sqrt{2t}x)^j \eta_g^{(j)}(\sqrt{2t}x) = (\eta_g)_m^*(\sqrt{2t}x).
\]

Therefore, \( (\eta_t)_m^* \geq 0 \) for all \( t > 0 \) and \( x \in \mathbb{R} \). We now consider an even function \( \eta : \mathbb{R} \to \mathbb{R}_+ \) such that \( \eta(\sqrt{x}) = \int_0^\infty e^{-tx} \mu(dt) \) for some Borel measure \( \mu \) on \([0, \infty)\). We then have for all \( x \in \mathbb{R} \), \( \eta(x) = \int_0^\infty e^{-tx^2} \mu(dt) = \int_0^\infty \eta_t(x) \mu(dt) \) and thus \( \eta^{(j)}(x) = \int_0^\infty \eta_t^{(j)}(x) \mu(dt) \). This gives, for all \( m \in \mathbb{N}^* \),

\[
\eta^*_m(x) = (-1)^{m-1} \sum_{j=1}^{m} c_{j,m} x^j \eta^{(j)}(x) = \int_0^\infty (-1)^{m-1} \sum_{j=1}^{m} c_{j,m} x^j \eta_t^{(j)}(x) \mu(dt) \\
= \int_0^\infty (\eta_t)_m^* \mu(dt) \geq 0
\]

where the last integral is positive for all \( x \in \mathbb{R} \) because is an integral of a positive function against a positive measure. \( \square \)

**Corollary C.2.** All the densities that satisfy the hypothesis of the representation Lemma 5.13 for all \( m \in \mathbb{N}^* \) are such that \( \eta(\sqrt{x}) \) is the Laplace transform of a finite positive Borel measure \( \mu \) over \([0, \infty)\).

**References**

[1] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384, 2005.

[2] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.
HIGH ORDER APPROXIMATIONS FOR THE CIR PROCESS USING RANDOM GRIDS

A. Alfonsi. Affine diffusions and related processes: simulation, theory and applications, volume 6 of Bocconi & Springer Series. Springer, Cham; Bocconi University Press, Milan, 2015.

A. Alfonsi and V. Bally. A generic construction for high order approximation schemes of semigroups using random grids. *Numer. Math.*, 148(4):743–793, 2021.

M. Altmayer and A. Neuenkirch. Discretising the Heston model: an analysis of the weak convergence rate. *IMA J. Numer. Anal.*, 37(4):1930–1960, 2017.

M. Bossy and A. Diop. An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^a$, $a$ in $[1/2, 1)$. (RR-5396):44, 2007. Version 2.

M. Briani, L. Caramellino, and G. Terenzi. Convergence rate of Markov chains and hybrid numerical schemes to jump-diffusion with application to the Bates model. *SIAM J. Numer. Anal.*, 59(1):477–502, 2021.

J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 6(2):327–343, 1993.

S. Kusuoka. Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara ’98*, volume 31 of *Adv. Stud. Pure Math.*, pages 147–165. Math. Soc. Japan, Tokyo, 2001.

M. Ninomiya and S. Ninomiya. A new higher-order weak approximation scheme for stochastic differential equations and the Runge-Kutta method. *Finance Stoch.*, 13(3):415–443, 2009.

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

K. Oshima, J. Teichmann, and D. Velušek. A new extrapolation method for weak approximation schemes with applications. *Ann. Appl. Probab.*, 22(3):1008–1045, 2012.

G. Pagés. Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods Appl.*, 13(1):37–70, 2007.

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

D. V. Widder. *The Laplace Transform*. Princeton Mathematical Series, vol. 6. Princeton University Press, Princeton, N. J., 1941.

Aurélien Alfonsi, CERMICS, Ecole des Ponts, Marne-la-Vallée, France. MathRisk, Inria, Paris, France.

Email address: aurelien.alfonsi@enpc.fr

Edoardo Lombardo, CERMICS, Ecole des Ponts, Marne-la-Vallée, France. MathRisk, Inria, Paris, France. Università degli Studi di Roma Tor Vergata, Rome, Italy.

Email address: edoardo.lombardo@enpc.fr