Recursive Algorithm for Transition Density Approximation and Exact Simulation of SDEs

Samir Ben Hariz, Youssef Esstafa, Helmi Zaatra

Recursive Algorithm for Transition Density Approximation and Exact Simulation of SDEs

Samir Ben Hariz[a], Youssef Esstafa[a], Helmi Zaatra[b]

[a]*Le Mans Université, Laboratoire Manceau de Mathématiques, Avenue Olivier Messiaen, 72085 Le Mans Cedex 09, France*
[b]*Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, P.O.Box 2713, Doha, Qatar*

Abstract

Based on the seminal work by [1], we propose a simple recursive algorithm for approximating the transition density of a stochastic differential equation (SDE). We then introduce a new method for simulating SDEs. Unlike Euler's and other schemes, our discretization step is fixed and does not need to approach zero. We apply our methods through simulations to some well-known financial models, specifically the Vasicek and CIR models.

Keywords: Stochastic differential equations, transition density, Hermite polynomials, exact simulations
2000 MSC: 60H10, 60H35, 65C05

Contents

1. Introduction

In recent decades, stochastic differential equations (see [15] for more details) have been extensively investigated and have proven invaluable across diverse fields such as finance, population dynamics, biology, and telecommunications. While SDEs play a crucial role, their exact solutions are seldom available in explicit form, posing a significant challenge for the applied probability community. In the absence of explicit solutions, numerical approximations of the solutions become indispensable. Classical schemes like Euler and its variants are sensitive to time steps, which makes it important to develop more precise methods, such as exact methods.
Numerous contributions have been made in the literature, with pioneering efforts by [4] and [3] employing a retrospective rejection sampling approach. One can also refer to the work of [6] for improvements on the rejection sampling method.

The objective of this work is twofold: to provide a simple approximation of the transition density around the Gaussian distribution using Hermite polynomials, and to employ this approximation to simulate trajectories of SDEs. Various attempts have been made to approximate the transition density. [13] proposed a numerical resolution of the Kolmogorov partial differential equation satisfied by the transition density. Alternatively, [14] and other authors tackled the problem through finely sampled path simulations. [1] stood out by deriving a closed-form Hermite expansion of the transition density and demonstrating the numerical superiority of his method in terms of accuracy compared to existing ones.

In the same vein as the idea of [1], we aim to achieve a high level of accuracy in approximating the transition density. Aït-Sahalia's work involves developing the transition density of the stochastic differential equation around the Gaussian distribution using Hermite polynomials. The expression for the coefficients in this series expansion is in integral form, which incurs high computational costs. Our main contribution is to address this issue by proposing simple recursive formulas for computing these coefficients, significantly reducing the computational cost.

The transition density is expressed as the product of a Gaussian distribution and a linear combination of Hermite polynomials. Hence, it can be

simulated by rejection from a double exponential distribution, for instance. Consequently, by leveraging the Markovian nature of the stochastic differential equation, an accurate simulation of its trajectory becomes feasible. Unlike classical schemes such as Euler and Milstein (see for example [10]), which require a sufficiently small step to achieve the desired accuracy, our method achieves higher precision with a fixed time step by adding only a few terms in the transition density expansion. It is worth noting that only a few terms are needed to achieve a low approximation error.

The structure of this paper unfolds as follows: Section 2 elucidates the expansion of the transition density using Hermite polynomials and provides the closed-form approximation. Section 3 illustrates the accuracy of our method on two classical models, namely the Vasicek and CIR models. Section 4 is dedicated to the application of the transition density approximation in the development of a new simulation scheme for SDEs.

2. Closed-form approximation of the transition density

Before proceeding further, we introduce the following notations which will be used throughout the paper. Let D_x and D_x^k denote respectively $\frac{\partial}{\partial x}$ and $\frac{\partial^k}{\partial x^k}$. Consider a stochastic differential equation of the form

$$\begin{cases} dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0, \\ X_0 = x \in \mathbb{R}, \end{cases} \tag{1}$$

where μ and σ are respectively the drift and the volatility functions. The process $(W_t)_{t\geq 0}$ stands for a standard Brownian motion.

The fundamental idea in the work of [1] is to develop the transition density of the diffusion process in Equations (1) around the standard normal distribution. Since this density can potentially differ from the standard Gaussian one, a first transformation, known as Lamperti's transformation, is employed to make this possible. Let us be more precise. Define

$$Y_t = \gamma(X_t),$$

where

$$\gamma(x) = \int_{x^*}^{x} \frac{1}{\sigma(u)} du, \tag{2}$$

for some arbitrary x^*. Since $\sigma > 0$, the transformation γ is increasing. Applying Itô's Lemma, we get

$$dY_t = \mu_Y(Y_t)dt + dW_t, \quad Y_0 = y_0, \tag{3}$$

where

$$\mu_Y(y) \coloneqq \frac{\mu(\gamma^{-1}(y))}{\sigma(\gamma^{-1}(y))} - \frac{1}{2}D_x\sigma(\gamma^{-1}(y))$$

with γ^{-1} is the inverse function of γ given by (2). This transformation aims to obtain the unitary diffusion $(Y_t)_t$.

The drift μ and diffusion σ are supposed to satisfy the following set of assumptions:

Assumption 2.1. *The functions μ and σ are infinitely differentiable.*

Assumption 2.2. *Let $\mathcal{S}_X$ denote the domain of X_t.*

1. *If $\mathcal{S}_X = (-\infty, \infty)$: there exists a positive constant C such that $\sigma(x) > C$ for all $x \in \mathcal{S}_X$.*
2. *If $\mathcal{S}_X = (0, \infty)$: near 0, if $\lim_{x\to 0^+} \sigma(x) = 0$, there exists a constant K_0, $\alpha \geq 0$ and $\rho \geq 0$ such that $\sigma(x) \geq \alpha x^\rho$ for all $0 < x < K_0$; away from 0, for each $K > 0$, there exists a constant $C_K > 0$ such that $\sigma(x) \geq C_K$ for all $x \in [K, \infty]$.*

Assumption 2.3. *Let $\lambda_Y(y) = -\frac{1}{2}\{\mu_Y^2(y) + D_y\mu_Y(y)\}$ and denote by $\mathcal{S}_Y = (\underline{y}, \overline{y})$ the domain of the transform diffusion Y_t. The functions μ_Y and $D_y\mu_Y$ have at most polynomial growth near the boundaries and $\lim_{y\to\partial\mathcal{S}_Y} \lambda(y) < \infty$. Near $\overline{y} = \infty$, $\mu(y) \leq -\kappa_1 y^\beta$ for some $\beta > 1$ and $\kappa_1 > 0$; near $\underline{y} = -\infty$, $\mu_Y(y) \geq \kappa_1|y|^\beta$ for some $\beta > 1$ and $\kappa_1 > 0$; near $\underline{y} = 0$, $\mu_Y(y) \geq \kappa_2 y^{-\zeta}$ for some $\zeta > 1$ and $\kappa_2 > 0$; and near $\overline{y} = 0$, $\mu_Y(y) \leq -\kappa_2|y|^{-\zeta}$ for some $\zeta > 1$ and $\kappa_2 > 0$.*

These conditions are similar to the ones of [1] and guarantee the existence and uniqueness of a weak solution of (1) (one can refer to Propositions 1 and 2 in [1] for more details).

Let $\delta > 0$. We further standardize Y_δ by the following operation:

$$Z_\delta = \delta^{-\frac{1}{2}}(Y_\delta - y_0).$$

The previous transformations allow to express the transition density of Z_δ using the standard normal density and Hermite polynomials defined hereafter.

The generalized Hermite polynomials of degree n are defined, for $t \geq 0$, as follows:

$$H_n(x,t) = \begin{cases} (-t)^n e^{\frac{x^2}{2t}} D_x^n e^{-\frac{x^2}{2t}} & \text{if } t > 0, \\ x^n & \text{if } t = 0. \end{cases}$$

In particular, for $t = 1$, we recover the classical Hermite polynomials, still denoted H_n,

$$H_n(x,1) = H_n(x) = (-1)^n \frac{D_z^n \phi(z)}{\phi(z)},$$

where $D_z^n \phi(z)$ is the n-th derivative of the standard Gaussian density $\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$. Moreover, the Hermite polynomials satisfy the following useful properties:

$$\begin{aligned} D_x H_n(x,t) &= n H_{n-1}(x,t), \\ D_t H_n(x,t) &= -\frac{1}{2} D_x^2 H_n(x,t). \end{aligned} \tag{4}$$

[1] has shown the following fundamental result for the development of the transition density of Z_δ using Hermite polynomials.

Theorem 2.4 ([1]). *Let $p_Z(\delta, z|y_0)$ denotes the transition density of Z_δ at z conditionally on y_0 and consider the following expansion:*

$$p_{Z,K}(\delta, z|y_0) := \phi(z) \sum_{k=0}^{K} \eta_{Z,k}(\delta, y_0) H_k(z), \tag{5}$$

where

$$\eta_{Z,k}(\delta, y_0) = \frac{1}{k!} \int_{\mathbb{R}} H_k(z) p_Z(\delta, z|y_0) dz.$$

Under Assumptions 2.1-2.3, we have

$$p_{Z,K}(\delta, z|y_0) \longrightarrow p_Z(\delta, z|y_0), \text{ as } \quad K \to \infty.$$

The proposed approximation depends mainly on the coefficients $(\eta_{Z,k})_{k\geq 0}$. The evaluation of these coefficients requires an integral calculation which risks to be costly. A classical solution to get around this problem is to use a Taylor expansion around 0 of $\eta_{Z,k}$. This is the subject of the following proposition.

Proposition 2.5. *Under Assumptions 2.1-2.3, there exists $\delta^* \in (0, \delta)$ such that*

$$\mathbb{E}\left[H_k\left(Y_{t+\delta} - y_0, \delta\right) | Y_t = y_0\right] = \sum_{j=0}^{J} \frac{\delta^j}{j!} \mathcal{B}^j H_k\left(y - y_0, \delta\right)_{|y=y_0,\ \delta=0} + \mathbb{E}\left[\mathcal{B}^{J+1} H_k\left(Y_{t+\delta^*} - y_0, \delta^*\right) | Y_t = y_0\right] \frac{\delta^{J+1}}{(J+1)!},$$

where $\mathcal{B}$ is the infinitesimal operator given for a smooth function $g := g(t, y)$ by

$$\mathcal{B}g = D_t g + \mu_Y D_y g + \frac{1}{2} D_y^2 g,$$

and $\mathcal{B}^k g$ is the operator applied k times to g.

Proof. We define Ψ: $s \longmapsto \mathbb{E}\left[H_k\left(Y_{t+s} - y_0, s\right) | Y_t = y_0\right]$. This function is differentiable and we have

$$\begin{aligned} D_s \Psi(s) &= \int_{-\infty}^{+\infty} D_s \left\{H_k\left(y - y_0, s\right) p_Y(s, y|y_0)\right\} dy \\ &= \int_{-\infty}^{+\infty} \left\{D_s H_k\left(y - y_0, s\right) p_Y(s, y|y_0) + H_k\left(y - y_0, s\right) D_s p_Y(s, y|y_0)\right\} dy. \end{aligned}$$

The diffusion $(Y_t)_t$ satisfies the following forward Kolmogorov equation (see [9]):

$$D_s p_Y = \frac{1}{2} D_y^2 p_Y - D_y\left(\mu_Y p_Y\right).$$

Therefore, one obtains

$$\begin{aligned} D_s \Psi(s) &= \int_{-\infty}^{+\infty} \Bigg\{ D_s H_k\left(y - y_0, s\right) p_Y(s, y|y_0) \\ &\qquad + H_k\left(y - y_0, s\right) \left(\frac{1}{2} D_y^2 p_Y(s, y|y_0) - D_y\left(\mu_Y p_Y(s, y|y_0)\right)\right) \Bigg\} dy \\ &= \int_{-\infty}^{+\infty} \left\{ D_s H_k\left(y - y_0, s\right) + \mu_Y D_y H_k(y - y_0, s) + \frac{1}{2} D_y^2 H_k(y - y_0, s) \right\} p_Y(s, y|y_0) dy, \end{aligned}$$

where the last equality follows from integration by parts. Consequently, we get

$$D_s \Psi(s) = \mathbb{E}\left[\mathcal{B} H_k\left(Y_{t+s} - y_0, s\right) | Y_t = y_0\right]. \tag{6}$$

By iterating the same steps as above, we have for all $k \geq 0$,

$$D_s^k \Psi(s) = \mathbb{E}\left[\mathcal{B}^k H_k\left(Y_{t+s} - y_0, s\right) | Y_t = y_0\right].$$

The proposition follows from a Taylor development of the function Ψ at $s = 0$. □

Remark 2.6. We can also prove (6) using Itô's Lemma. Indeed, for any regular function f, we have

$$df(Y_t, t) = \mathcal{B}f(Y_t, t)dt + \mu_Y\left(Y_t\right) D_y f(Y_t, t) dW_t.$$

Therefore, writing the previous equation in an integral form and taking the expectation gives

$$\mathbb{E}\left[f(Y_t, t)\right] = f(Y_0, 0) + \int_0^t \mathbb{E}\left[\mathcal{B}\ f(Y_s, s)\right] ds.$$

Differentiating both sides with respect to t ends the claim.

Now, we are in a position to state our main theoretical result.

Theorem 2.7. *For $n \geq 1$, the operator $\mathcal{B}$ applied n times to H_k can be represented as*

$$\mathcal{B}^n H_k = \sum_{j=1}^{n} P_j^n(\mu_Y) D_y^j H_k, \tag{7}$$

where the functions $P_j^n(\mu_Y)$ satisfy

$$\begin{aligned}
&P_1^{n+1}(\mu_Y) = \mu_Y D_y P_1^n(\mu_Y) + \frac{1}{2} D_y^2 P_1^n(\mu_Y), \\
&P_j^{n+1}(\mu_Y) = \mu_Y P_{j-1}^n(\mu_Y) + \mu_Y D_y P_j^n(\mu_Y) + D_y P_{j-1}^n(\mu_Y) + \frac{1}{2} D_y^2 P_j^n(\mu_Y), \quad 2 \leq j \leq n, \\
&P_{n+1}^{n+1}(\mu_Y) = \mu_Y P_n^n(\mu_Y) + D_y P_n^n(\mu_Y),
\end{aligned}$$

with the initial condition $P_1^1(\mu_Y) = \mu_Y$.

The above relations can be written in a matrix form as follows:

$$\begin{pmatrix} P_1^{n+1} \\ P_2^{n+1} \\ P_3^{n+1} \\ \vdots \\ P_n^{n+1} \\ P_{n+1}^{n+1} \end{pmatrix} = \begin{pmatrix} B_1 & 0 & \cdots & \cdots & \cdots & 0 \\ B_2 & B_1 & 0 & \cdots & \cdots & 0 \\ 0 & B_2 & B_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & B_2 & B_1 & 0 \\ 0 & \cdots & 0 & 0 & B_2 & B_1 \end{pmatrix} \begin{pmatrix} P_1^n \\ P_2^n \\ P_3^n \\ \vdots \\ P_n^n \\ 0 \end{pmatrix},$$

where

$$B_1 \coloneqq \mu_Y D_y + \frac{1}{2} D_y^2$$

and

$$B_2 \coloneqq \mu_Y D_y^0 + D_y.$$

Proof. We proceed by induction on n. For $n = 1$, we have by definition

$$\mathcal{B} H_k = D_t H_k + \mu_Y D_y H_k + \frac{1}{2} D_y^2 H_k.$$

In view of the identities in (4), one has $D_t H_k + \frac{1}{2} D_y^2 H_k = 0$. It follows that

$$\mathcal{B} H_k = \mu_Y D_y H_k = P_1^1(\mu_Y) D_y H_k.$$

Now we assume that the property (7) is true for some n and we shall deduce that it remains valid for $n + 1$.

$$\begin{aligned}
\mathcal{B}^{n+1} H_k &= \mathcal{B}(\mathcal{B}^n H_k) \\
&= \mathcal{B} \left(\sum_{j=1}^{n} P_j^n(\mu_Y) D_y^j H_k \right) \\
&= \sum_{j=1}^{n} P_j^n(\mu_Y) D_t D_y^j H_k \\
&\quad + \mu_Y \sum_{j=1}^{n} \left\{ D_y P_j^n(\mu_Y) D_y^j H_k + P_j^n(\mu_Y) D_y^{j+1} H_k \right\} \\
&\quad + \frac{1}{2} \sum_{j=1}^{n} \left\{ D_y^2 P_j^n(\mu_Y) D_y^j H_k + 2 D_y P_j^n(\mu_Y) D_y^{j+1} H_k + P_j^n(\mu_Y) D_y^{j+2} H_k \right\}.
\end{aligned}$$

Again, we use the properties in (4) to obtain

$$\begin{aligned}
\mathcal{B}^{n+1}H_k &= \mu_Y \sum_{j=1}^{n} D_y P_j^n(\mu_Y) D_y^j H_k + P_j^n(\mu_Y) D_y^{j+1} H_k \\
&\quad + \frac{1}{2}\sum_{j=1}^{n} D_y^2 P_j^n(\mu_Y) D_y^j H_k + \sum_{j=1}^{n} D_y P_j^n(\mu_Y) D_y^{j+1} H_k \\
&= \sum_{j=1}^{n} \left[\mu_Y D_y P_j^n(\mu_Y) + \frac{1}{2} D_y^2 P_j^n(\mu_Y)\right] D_y^j H_k \\
&\quad + \sum_{j=2}^{n+1} \left[\mu_Y P_{j-1}^n(\mu_Y) + D_y P_{j-1}^n(\mu_Y)\right] D_y^j H_k \\
&= \sum_{j=1}^{n+1} P_j^{n+1}(\mu_Y) D_y^j H_k.
\end{aligned}$$

Therefore, for any $n \geq 1$, we get

$$\mathcal{B}^n H_k = \sum_{j=1}^{n} P_j^n(\mu_Y) D_y^j H_k,$$

where

$$\begin{aligned}
&P_1^{n+1}(\mu_Y) = \mu_Y D_y P_1^n(\mu_Y) + \frac{1}{2} D_y^2 P_1^n(\mu_Y) \coloneqq B_1 P_1^n(\mu_Y), \\
&P_{n+1}^{n+1}(\mu_Y) = \mu_Y P_n^n(\mu_Y) + D_y P_n^n(\mu_Y) \coloneqq B_2 P_n^n(\mu_Y), \\
&P_j^{n+1}(\mu_Y) = B_1 P_j^n(\mu_Y) + B_2 P_{j-1}^n(\mu_Y), \text{ for } j = 2, \ldots, n.
\end{aligned}$$

□

In particular, for $j = 1$, one has

$$P_1^n(\mu_Y) = B_1 P_1^{n-1}(\mu_Y) = \cdots = B_1^{n-1}(\mu_Y).$$

Now, for $j = 2$ and $n \geq 2$, we can express

$$\begin{aligned}
P_2^2(\mu_Y) &= B_2 P_1^1(\mu_Y) = B_2(\mu_Y), \\
P_2^{n+1}(\mu_Y) &= B_1 P_2^n(\mu_Y) + B_2 P_1^n(\mu_Y).
\end{aligned}$$

Remark 2.8. The coefficients $\eta_{Z,k}(\delta, y_0)$ are given by

$$\eta_{Z,k}(\delta, y_0) = \frac{1}{k!}\mathbb{E}\left[H_k\left(\frac{Y_{t+\delta} - y_0}{\sqrt{\delta}}\right)\right].$$

Since $H_n(x,t) = t^{n/2}H_n\left(x/\sqrt{t}\right)$, we can write

$$\eta_{Z,k}(\delta, y_0) = \frac{\delta^{-k/2}}{k!}\mathbb{E}\left[H_k\left(Y_{t+\delta} - y_0, \delta\right)|Y_t = y_0\right].$$

The following corollary is a consequence of Proposition 2.5 and Theorem 2.7.

Corollary 2.9. *For any $j \geq 1$, we have*

$$\eta_{Z,k}(\delta, y_0) = \sum_{n=k}^{J} P_k^n(\mu_Y)\frac{\delta^{n-k/2}}{j!} + R_k^{J+1}\frac{\delta^{J+1}}{(J+1)!}, \tag{8}$$

where

$$R_k^{J+1} = \mathbb{E}\left[\mathcal{B}^{J+1}H_k\left(Y_{\theta\delta} - y_0, \delta\right)|Y_0 = y_0\right].$$

Proof. We apply Theorem 2.7 to H_k when $y = y_0$ and $t = 0$ to deduce the following. From Proposition 2.5 we have

$$\begin{aligned}
\eta_{Z,k}(\delta, y_0) &= \frac{\delta^{-k/2}}{k!}\sum_{n=0}^{J}\mathcal{B}^n H_k\left(y - y_0, \delta\right)_{|y=y_0,\delta=0}\frac{\delta^n}{n!} + R_k^{J+1}\frac{\delta^{J+1}}{(J+1)!} \\
&= \frac{\delta^{-k/2}}{k!}\sum_{n=0}^{k-1}\mathcal{B}^n H_k\left(y - y_0, \delta\right)_{|y=y_0,\delta=0}\frac{\delta^n}{n!} \\
&\quad + \frac{\delta^{-k/2}}{k!}\sum_{n=k}^{J}\mathcal{B}^n H_k\left(y - y_0, \delta\right)_{|y=y_0,\delta=0}\frac{\delta^n}{n!} + R_k^{J+1}\frac{\delta^{J+1}}{(J+1)!}.
\end{aligned}$$

To simplify $\mathcal{B}^n H_k$ we use (7) and distinguish two cases:

$\diamond$ *For* $\mathbf{n < k}$*:.* Since $D_y H_k = kH_{k-1}$ and $H_k(0,0) = 0$ for $k \geq 1$, we have

$$\begin{aligned}
\mathcal{B}^n H_k(y - y_0, t)_{|y=y_0,t=0} &= \sum_{j=1}^{n} P_j^n(\mu_Y)D_y^j H_k(y - y_0, t)_{|y=y_0,t=0} \\
&= \sum_{j=1}^{n} P_j^n(\mu_Y)\frac{k!}{(k-j)!}H_{k-j}(0,0) = 0.
\end{aligned}$$

$\diamond$ *For* $\mathbf{n} \geq \mathbf{k}$*:.* The polynomial H_k is of order k with respect to y. Then, it follows

$$\begin{aligned}\mathcal{B}^n H_k(y-y_0,t)_{|y=y_0,t=0} &= \sum_{j=1}^{k} P_j^n(\mu_Y) D_y^j H_k(y-y_0,t)_{|y=y_0,t=0} \\ &= k! P_k^n(\mu_Y).\end{aligned}$$

Hence, we only need to evaluate $P_k^n(\mu_Y)$. Finally we conclude

$$\eta_{Z,k}(\delta, y_0) = \sum_{n=k}^{J} P_k^n(\mu_Y) \frac{\delta^{n-k/2}}{n!} + R_k^{J+1} \frac{\delta^{J+1}}{(J+1)!}.$$

Remark 2.10. It is worth noting from the previous relation (8) that the leading term in the development of $\eta_{Z,k}(\delta, y_0)$ is of order $\delta^{k/2}$. Finally, plugged into Equation (5) we obtain the following approximation

$$\begin{aligned}p_{Z,K}(\delta, z|y_0) &= \phi(z) \sum_{k=0}^{K} \eta_{Z,k}(\delta, y_0) H_k(z) \\ &\approx \phi(z) \sum_{k=0}^{K} \sum_{n=k}^{J} P_k^n(\mu_Y) \frac{\delta^{n-k/2}}{n!} H_k(z). \end{aligned} \tag{9}$$

Example 2.11 (Ornstein–Uhlenbeck (O.-U.) process)**.** Consider the following SDE: $dY_t = -Y_t dt + dW_t, Y_0 = y_0$. Itô's Lemma applied to $\exp(t)Y_t$ yields: $Y_t = y_0 \exp(-t) + \int_0^t \exp(s-t) dW_s$. Therefore Y_t is Gaussian with $\mathbb{E}[Y_t] = y_0 \exp(-t)$ and $\mathrm{Var}(Y_t) = \int_0^t \exp(2(s-t)) ds = \frac{1}{2}(1 - \exp(-2t))$. The evaluation of the first coefficient $\eta_{Z,1}$ corresponding to O.-U. process is given by

$$\eta_{Z,1}(\delta, y_0) = \delta^{-\frac{1}{2}} \mathbb{E}[Y_\delta - y_0 | Y_0 = y_0] = y_0 \delta^{-\frac{1}{2}} [\exp(-\delta) - 1] = \delta^{-1/2} y_0 \sum_{n=1}^{+\infty} \frac{(-\delta)^n}{n!}.$$

Now, simple calculation gives $P_1^n(\mu_Y) = (-1)^n y$ and

$$\eta_{Z,1}(\delta, y_0) \simeq \sum_{n=1}^{J} P_1^n(\mu_Y) \frac{\delta^{n-1/2}}{n!} = \delta^{-1/2} y_0 \sum_{n=1}^{J} \frac{(-\delta)^n}{n!}.$$

Regarding $\eta_{Z,2}(\delta, y_0)$, we have

$$
\begin{aligned}
\eta_{Z,2}(\delta, y_0) &= \frac{\delta^{-1}}{2}\,\mathbb{E}\left[(Y_\delta - y_0)^2 - \delta\right] \\
&= \frac{\delta^{-1}}{2}\left\{\mathrm{Var}\left(Y_\delta\right) + \left(\mathbb{E}\left[Y_\delta\right] - y_0\right)^2 - \delta\right\} \\
&= \frac{\delta^{-1}}{4}\left\{1 - \exp(-2\delta) + 2y_0^2\left(\exp(-\delta) - 1\right)^2 - 2\delta\right\} \\
&= \frac{2y_0^2 - 1}{4\delta}\sum_{k\geq 2}\frac{(-2\delta)^k}{k!} - \frac{y_0^2}{\delta}\sum_{k\geq 2}\frac{(-\delta)^k}{k!} \\
&= \delta\left(\frac{1}{2}y_0^2 - \frac{1}{2}\right) - \delta^2\left(\frac{1}{2}y_0^2 - \frac{1}{3}\right) + \delta^3\left(\frac{7}{24}y_0^2 - \frac{1}{6}\right) + O\left(\delta^4\right).
\end{aligned}
$$

Similarly, the approximation of $\eta_{Z,2}(\delta, y_0)$ provide the same previous development.

3. Numerical approximation for transition densities

In this section, we illustrate the theoretical results obtained for the approximation of the transition density for the Vasicek and CIR models. These densities are available in closed form. We assess the accuracy of the transition density approximation for various values of the time step δ and different truncation thresholds in the approximation (9).

3.1. Vasicek model

The dynamics is given by the following SDE:

$$
dX_t = \kappa(\alpha - X_t)dt + \sigma dW_t, \quad X_0 = x_0, \tag{10}
$$

where κ, α and σ are positive constants. In this model, an explicit formula for X_t is given by

$$
X_t = X_0\exp(-\kappa t) + \alpha(1 - \exp(-\kappa t)) + \sigma\int_0^t \exp(-\kappa(t-s))dW_s. \tag{11}
$$

Clearly, from (11), the process $(X_t)_t$ is Gaussian with the following transition density:

$$
p_X(\delta, x|x_0) = \left(\frac{\pi\sigma^2(1 - e^{-2\kappa\delta})}{\kappa}\right)^{-\frac{1}{2}}\exp\left(\frac{-\kappa(x - \alpha - (x_0 - \alpha)e^{-\kappa\delta})^2}{\sigma^2(1 - e^{-2\kappa\delta})}\right). \tag{12}
$$

In this case, Lamperti's transformation gives $Y_t = X_t/\sigma$. This process satisfies the following unitary diffusion $dY_t = \mu_Y(Y_t)dt + dW_t, Y_0 = y_0$, with $\mu_Y(y) = \sigma^{-1}\kappa\alpha - \kappa y$.

Figure 3.1 clearly shows that the absolute uniform error of the approximation decreases rapidly with δ for every truncation order K. For example, for $K = 3$, the error ranges from 10^{-2} to 10^{-6} for δ between 0.5 and 0.01. Even better, for a fixed δ, the error decreases exponentially with K. For $\delta = 0.1$, the error drops from 10^{-2} for $K = 1$ to 10^{-5} for $K = 5$.

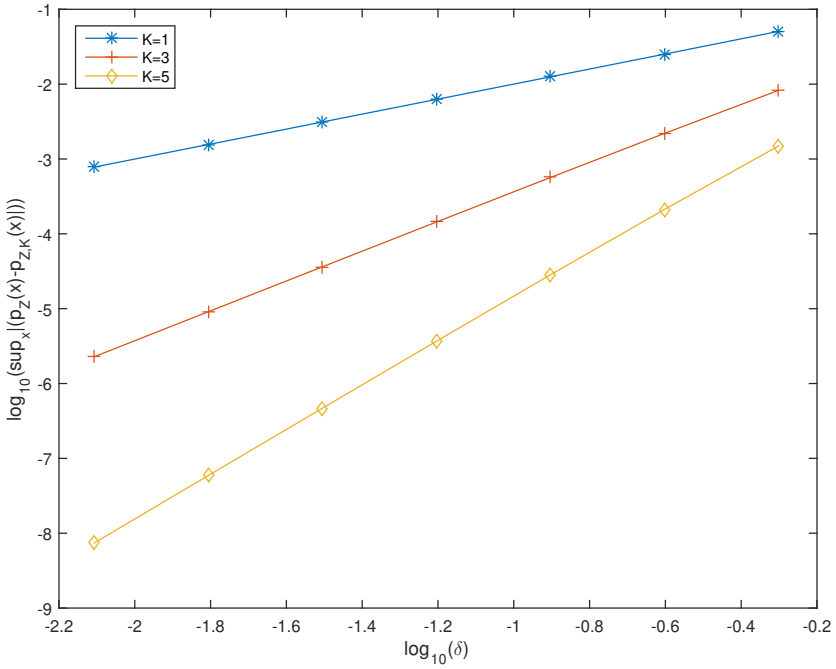

Figure 1: The evolution of the absolute uniform error between the transition density and its approximation for the Vasicek model with $(\kappa, \alpha, \sigma) = (0.5, 0.05, 0.1)$.

3.2. Cox-Ingersoll-Ross (CIR) model

[5] have proposed a model for describing the evolution of interest rates (one can refer to [2] and [8] for more details). The CIR model is given by

$$dX_t = \kappa(\alpha - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 = x_0,$$

where κ, α and σ are positive constants. When $q = 2\kappa\alpha/\sigma^2 - 1 > 0$, the

solution X_t remains positive and its transition density can be expressed as

$$p_X(\delta, x|x_0) = ce^{-u-v}\left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}),$$

where $c = 2\kappa/\left(\sigma^2(1-e^{-\kappa\delta})\right)$, $u = cx_0e^{-\kappa\delta}$, $v = cx$, and I_q is the modified Bessel function of the first kind of order q. The first transformation $Y_t = 2\sqrt{X_t}/\sigma$ satisfies the unitary diffusion

$$dY_t = \mu_Y(Y_t)dt + dW_t, \quad Y_0 = y_0,$$

with drift function $\mu_Y(y) = \frac{q+\frac{1}{2}}{y} - \frac{\kappa y}{2}$.

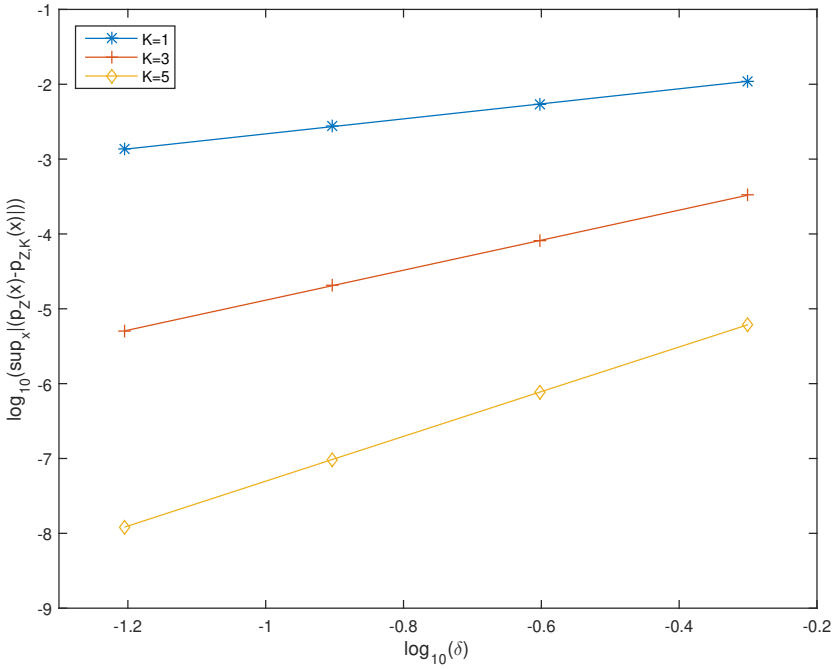

Figure 2: The evolution of the absolute uniform error between the transition density and its approximation for the CIR model with $(\kappa, \alpha, \sigma) = (0.1, 0.1, 0.1)$.

Figure 2 confirms the conclusions drawn regarding the behavior of the error with respect to K and δ for the Vasicek model. It is clear that the proposed approximation in (9) outperforms the Euler scheme in terms of accuracy, which can be viewed as a zero-order expansion ($K = 0$) of the density. Consequently, the proposed method can be used, for example, for the pricing of European options, as was done in the work of [18].

4. Simulation of SDEs

In this section, we present a new simulation scheme for trajectories of SDEs of the form (1). The idea is to exploit the approximation of the transition density of the normalized process $(Z_t)_t$ to simulate the different increments of $(Y_t)_t$. Starting from y_0, we simulate Z_δ by rejection sampling from a double exponential distribution. We then set Y_δ as $y_0 + \sqrt{\delta} Z_\delta$. We repeat this process to simulate $Y_{2\delta}$, $Y_{3\delta}$ and so on. Finally, we can use the inverse transformation to go back to $(X_t)_t$.

We recall that the transition density of Z_δ at z conditionally to y_0 is given by

$$\begin{aligned} p_Z(\delta, z|y_0) &= \lim_{K\to\infty} p_{Z,K}(\delta, z|y_0), \\ &= \phi(z) \sum_{k=0}^{+\infty} \eta_{Z,k}(\delta, y_0) H_k(z), \\ &= \phi(z) \left\{ 1 + \eta_{Z,1}(\delta, y_0) z + \eta_{Z,2}(\delta, y_0)(z^2 - 1) + \cdots \right\}, \end{aligned}$$

where $\eta_{Z,k}(\delta, y_0) = \frac{1}{k!} \mathbb{E}\left[H_k\left(\delta^{-\frac{1}{2}} (Y_\delta - y_0) \right) | Y_0 = y_0 \right]$. The transition density appears as a polynomial in z factor the standard normal one. Hence, it is possible to sample by rejection from the double exponential distribution, given by

$$q_Z(z|z_0) = \frac{\lambda}{2} \exp(-\lambda |z - z_0|),$$

where λ and z_0 are chosen to minimize the rejection probability. We choose z_0 such that $\int_{\mathbb{R}} z q_Z(z|z_0) dz = \int_{\mathbb{R}} z p_Z(\delta, z|y_0) dz$. That is $z_0 = \eta_{Z,1}(\delta, y_0) \simeq \sqrt{\delta} \mu_Y(y_0)$.

We can summarize the simulation algorithm for $(Y_t)_t$ as follows:

Algorithm 1 The new algorithm for simulation of $(Y_t)_t$

Require: $y_0, \mu_Y, \delta, N, K, J$
  Initialize Y_0 to y_0;
  Approximate $p_Z(\delta, z|y_0)$ by $p_{Z,K}(\delta, z|y_0)$ as in (9);
  Simulate Z_δ from $q_Z(\cdot|y_0)$;
  Set $Y_\delta = y_0 + \sqrt{\delta} Z_\delta$.
  for k from 1 to $N-1$ **do**
    Simulate $Z_{k\delta}$ from $q_Z(\cdot|Y_{(k-1)\delta})$;
    Set $Y_{(k+1)\delta} = Y_{k\delta} + \sqrt{\delta} Z_{k\delta}$.
  end for
  Return Y_{t_k} for $t_k = k\delta,\ k = 0, \ldots, N$.

4.1. Comparative analysis

We will compare three simulation techniques for stochastic differential equations: the Euler scheme, the Beskos and Robert method (BR) and the new algorithm (NA). Beskos and Robert's method is an exact scheme to simulate unitary diffusions (see [4] for more details).

We choose the following SDE to achieve our comparative analysis:

$$dY_t = \sin(Y_t)dt + dW_t, \quad Y_0 = 0. \tag{13}$$

The drift function satisfies all the required conditions in [4] as well as in [1].

We fix the time step δ and the terminal time T. Then, for each method, we simulate the trajectory of the solution at points $k\delta$, $k = 0, \ldots, N = \lfloor T/\delta \rfloor$. We repeat this process 10,000 times to obtain samples of discretized trajectories. Finally, we perform Kolmogorov-Smirnov tests to check the equality of the empirical distributions for $Y_T = Y_{\lfloor N\delta \rfloor}$. The comparison of distributions across all trajectories is presented in the following subsection using an appropriate method.

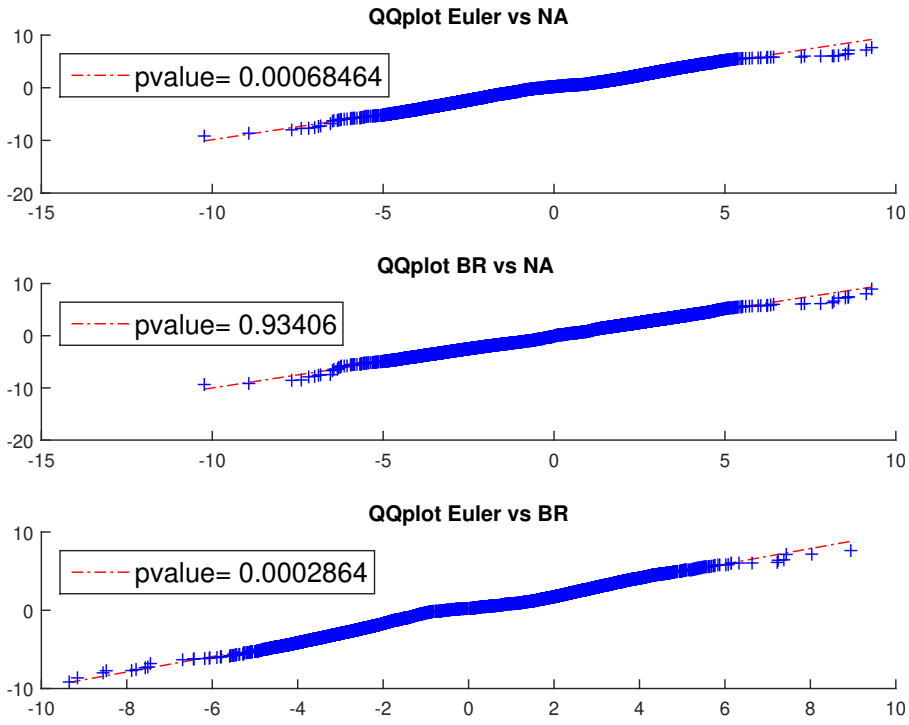

Figure 3: The QQ-plot comparison of NA, Euler and BR methods for simulation of $Y_T, T = 2$ and $\delta = 0.5$ following the Sine model (13) with the associated Kolmogorov Smirnov p-value test for equality of distributions.

Clearly, Figure 3 shows an excellent adequation between the two exact methods (BR and NA) at $\delta = 0.5$. However, the comparison of these two methods with the Euler scheme yields a very low p-value, suggesting a poor fit. This is due to error propagation in the Euler scheme when δ is large.

4.2. Energy distance

The Kolmogorov-Smirnov test allows for the comparison of one-dimensional distributions. In order to take into account multidimensional distributions, we will use the Energy distance and its associated test to compare our approach with the Euler's scheme and Beskos & Robert's exact algorithm.

Energy distance is a statistical measure that quantifies the discrepancy or dissimilarity between two probability distributions. It provides a way to compare and assess the similarity or difference between the underlying distributions of two sets of observations or data samples. Computationally,

the energy distance can be estimated using sample data by replacing the true distributions with their empirical counterparts based on the observed data samples. We will use this measure to test equality of distributions between two discretized paths of an SDE.

The energy Distance was introduced by Székely and Rizzo (see [16] and the references therein for further details). Given two samples $S_1 = (x_1, x_2, ..., x_n)$ and $S_2 = (y_1, y_2, ..., y_m)$ from two distributions, the E-statistic for testing equality in distributions, is defined as follows:

$$\mathcal{E}_{n,m}(S_1, S_2) = \frac{2}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \|x_i - y_j\| - \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \|x_i - x_k\| - \frac{1}{m^2} \sum_{l=1}^{n} \sum_{j=1}^{m} \|y_l - y_j\| .$$

The associated test uses a permutation test for checking the equality in distribution between the two samples.

For each of the three methods introduced above, we simulate $10,000$ trajectories of the Sine model (13) over $[0, T]$ for different discretization steps δ. We test the equality of distributions between the trajectories generated by the three methods using the test associated with the energy distance. More precisely, for a fixed time T and δ, we generate n trajectories of $(Y_t)_{t\in[0,T]}$ using the three methods. Thus, we obtain n i.i.d. vectors of size $\lfloor T/\delta \rfloor$ for each method. We perform pairwise tests between the samples and calculate the p-values. This process is repeated $M = 1,000$ times.

Figure 4 shows the average p-values as a function of δ for pairwise method comparisons. It is clear that the p-values for the comparison of the two exact methods are roughly centered around 0.5, regardless of δ. This confirms the equality of the distributions of the trajectories of the two methods. However, the comparison test between one of the exact methods and the Euler scheme returns average p-values that decrease drastically with δ, approaching 0. These results confirm the conclusions drawn from the comparison test on the terminal time distribution.

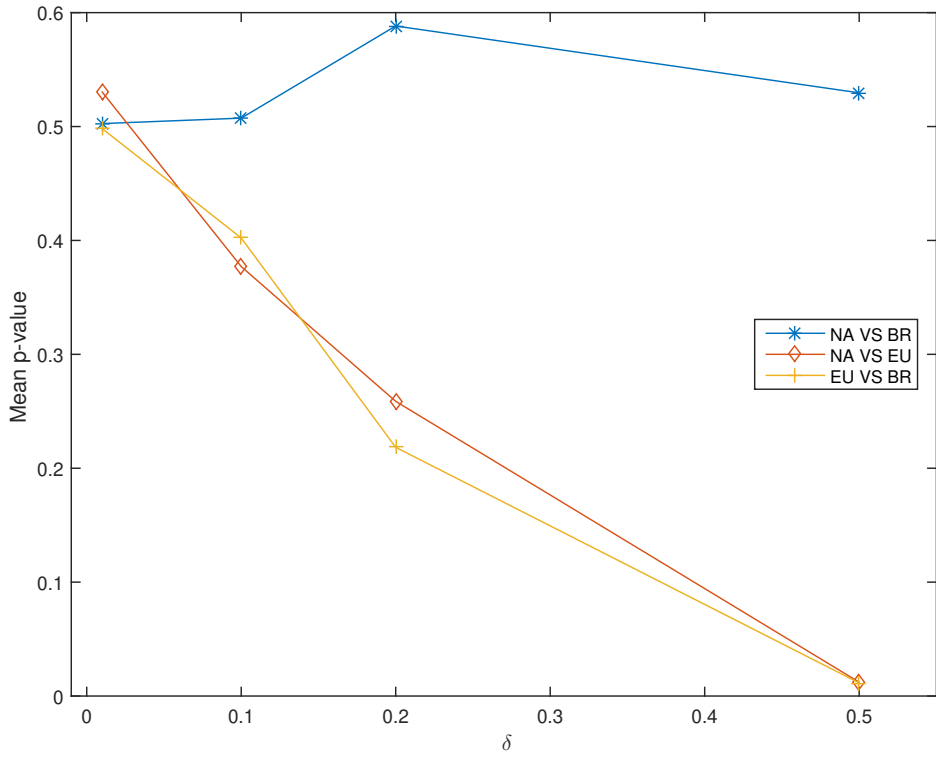

Figure 4: The average p-values of the tests comparing simulated trajectories of the Sine model (13) with the NA, BR, and Euler methods. The terminal time is $T = 2$ and the discretization step δ takes its values in $\{0.01, 0.1, 0.2, 0.5\}$.

References

[1] Aït-Sahalia, Y. (2002). Maximum-likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica*, 70(1):223–262.

[2] Aït-Sahalia, Y. (1999). Transition densities for interest rate and other nonlinear diffusions. *The Journal of Finance*, 54(4):1361–1395.

[3] Beskos, A., Papaspiliopoulos, O. and Roberts, G.O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, 12(6):1077–1098.

[4] Beskos A. and Roberts G.O. (2005). Exact simulation of diffusions. *The Annals of Applied Probability*, 15(4):2422–2444.

[5] Cox, J.C., Ingersoll Jr., J.E. and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407.

[6] Dai, H. (2017). A new rejection sampling method without using hat function. *Bernoulli*, 23(4A):2434–2465.

[7] Devroye, L. (2013). Non-Uniform Random Variate Generation. *Springer New York*.

[8] Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering. *Springer New York*.

[9] Karatzas, I. and Shreve, S.E. (1996). Brownian Motion and Stochastic Calculus. *Springer New York*.

[10] Kloeden, P.E. and Platen, E. (1992). Numerical Solution of Stochastic Differential Equations. *Springer Berlin, Heidelberg*.

[11] Kuo, H.-H., (2006). Introduction to Stochastic Integration. *Springer New York*.

[12] Lee, Y.D., Song, S. and Lee, E.-K. (2014). The delta expansion for the transition density of diffusion models. *Journal of Econometrics*, 178(3):694–705.

[13] Lo, A.W. (1988). Maximum likelihood estimation of generalized Itô processes with discretely sampled data. *Econometric Theory*, 4(2):231–247.

[14] Pedersen, A.R. (1995). A new approach to maximum-likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*, 22(1):55–71.

[15] Särkkä, S. and Solin A. (2019). Applied Stochastic Differential Equations. *Cambridge University Press*.

[16] Székely, G.J. and Rizzo, M.L. (2013). Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, 143(8):1249–1272.

[17] Wilk, M.B. and Gnanadesikan, R. (1968). Probability Plotting Methods for the Analysis of Data. *Biometrika*, 55(1):1–17.

[18] Xiu, D. (2014), Hermite polynomial based expansion of European option prices. *Journal of Econometrics*, 179(2):158–177.