Frentz, Marie; Nyström, Kaj

Adaptive stochastic weak approximation of degenerate parabolic equations of Kolmogorov type

Journal of Computational and Applied Mathematics, 2010, Vol. 234, Issue 1: 146-164

URL: http://dx.doi.org/10.1016/j.cam.2009.12.011

Adaptive Stochastic Weak Approximation of Degenerate Parabolic Equations of Kolmogorov type

Marie Frentz*
Department of Mathematics, Umeå University
S-90187 Umeå, Sweden
Kaj Nyström†
Department of Mathematics, Umeå University
S-90187 Umeå, Sweden

August 27, 2009

Abstract

Degenerate parabolic equations of Kolmogorov type occur in many areas of analysis and applied mathematics. In their simplest form these equations were introduced by Kolmogorov in 1934 to describe the probability density of the positions and velocities of particles but the equations are also used as prototypes for evolution equations arising in the kinetic theory of gases. More recently equations of Kolmogorov type have also turned out to be relevant in option pricing in the setting of certain models for stochastic volatility and in the pricing of Asian options. The purpose of this paper is to numerically solve the Cauchy problem, for a general class of second order degenerate parabolic differential operators of Kolmogorov type with variable coefficients, using a posteriori error estimates and an algorithm for adaptive weak approximation of stochastic differential equations. Furthermore, we show how to apply these results in the context of mathematical finance and option pricing. The approach outlined in this paper circumvents many of the problems confronted by any deterministic approach based on, for example, a finite-difference discretization of the partial differential equation in itself. These problems are caused by the fact that the natural setting for degenerate parabolic differential operators of Kolmogorov type is that of a Lie group much more involved than the standard Euclidean Lie group of translations, the latter being relevant in the case of uniformly elliptic parabolic operators.

2000 *Mathematics Subject Classification: 35K65, 60H07, 60H30, 65C05, 91B28.*
Keywords and phrases: degenerate parabolic, weak approximation, adaptivity, Malliavin calculus, option pricing.

*email: marie.frentz@math.umu.se
†email: kaj.nystrom@math.umu.se

1 Introduction

The simplest form of an operator of Kolmogorov type is the following degenerate parabolic operator in $\mathbb{R}^{2n+1}$,

$$\sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} + \sum_{j=n+1}^{2n} x_{j-n} \frac{\partial}{\partial x_j} - \partial_t. \tag{1.1}$$

The operator in (1.1) was introduced by Kolmogorov in 1934 in order to describe the density of a system with $2n$ degrees of freedom. In particular, here $\mathbb{R}^{2n}$ represents the phase space where $(x_1, ..., x_n)$ and $(x_{n+1}, ..., x_{2n})$ are, respectively, the velocity and position of the system, see [17]. An area of applied mathematics where operators of Kolmogorov type recently have turned out to be relevant is that of mathematical finance and option pricing. Degenerate equations of Kolmogorov type arise naturally in the problem of pricing path-dependent contingent claims referred to as Asian-style derivatives, see [7, 8, 9] and the references therein. In particular, after some manipulations the pricing of a geometric average Asian option in the standard Black-Scholes model is equivalent to solving the Cauchy problem for the operator (1.1), in this case $n = 1$, in $\mathbb{R}^2 \times [0, T]$ with Cauchy data, also called terminal data, defined by the pay-off of the contract. Moreover, the Cauchy problem for operators of Kolmogorov type, more general than the one stated in (1.1) and with variable coefficients, also appear in the pricing of general European derivatives in the framework of the stochastic volatility model suggested by Hobson and Rogers, see [8, 11].

The purpose of this paper is to apply and work out, for the backward in time Cauchy problem for a general class of second order degenerate parabolic partial differential operators of Kolmogorov type, the approach concerning a posteriori error estimates and adaptive weak approximations of stochastic differential equation due to Szepessy, Tempone and Zouraris, see [27]. Furthermore, we show how this approach can be applied to problems in mathematical finance and option pricing where degenerate parabolic operators of Kolmogorov type occur. In particular, we consider operators of the form

$$L = \frac{1}{2} \sum_{i,j=1}^{m} [\sigma\sigma^*]_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{m} b_i(x,t) \frac{\partial}{\partial x_i} + \sum_{i,j=1}^{n} c_{ij} x_i \frac{\partial}{\partial x_j} + \frac{\partial}{\partial t} \tag{1.2}$$

where $(x, t) \in \mathbb{R}^n \times \mathbb{R}$, m is a positive integer satisfying $m \leq n$, $\sigma(x,t) = \{\sigma_{ij}(x,t)\} : \mathbb{R}^n \times \mathbb{R}_+ \to M(n, m, \mathbb{R})$, $M(n, m, \mathbb{R})$ is the set of all $n \times m$-matrices with real valued entries and σ^* is the transpose of the matrix σ. $[\sigma\sigma^*]_{ij}(x,t)$ denotes the (i,j) entry of the matrix $[\sigma\sigma^*](x,t)$. The functions $\{\sigma_{ij}(\cdot,\cdot)\}$ and $\{b_i(\cdot,\cdot)\}$ are continuous with bounded derivatives and the matrix $C := \{c_{ij}\}$ is a matrix of constant real numbers. Note that we are particularly interested in the case $m < n$. Given $T > 0$ we consider the problem

$$\begin{cases} Lu(x,t) = 0 & \text{whenever } (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,T) \; = g(x) & \text{whenever } x \in \mathbb{R}^n, \end{cases} \tag{1.3}$$

where g is a given function. The problem in (1.3) represents the backward in time Cauchy problem for the operator L with terminal data g. Concerning structural assumptions on the operator L we assume that

$$A(x,t) = \{a_{ij}(x,t)\}, \;\; a_{ij}(x,t) := [\sigma\sigma^*]_{ij}(x,t), \;\; \text{is symmetric,} \tag{1.4}$$

and that there exists a $\epsilon \in [1,\infty)$ such that

$$\epsilon^{-1}|\xi|^2 \leq \sum_{i,j=1}^{m} a_{ij}(x,t)\xi_i\xi_j \leq \epsilon|\xi|^2 \text{ whenever } (x,t) \in \mathbb{R}^{n+1},\ \xi \in \mathbb{R}^m. \tag{1.5}$$

Note that in (1.5) we are only assuming ellipticity in m out of n spatial directions. Let $\bar{A}(x,t) = \{\bar{a}_{ij}(x,t)\}$ denote, whenever $(x,t) \in \mathbb{R}^{n+1}$, the unique $m \times m$-matrix which satisfies $\bar{A}(x,t)\bar{A}(x,t) = A(x,t)$. For $(x_0,t_0) \in \mathbb{R}^{n+1}$, fixed but arbitrary, we introduce the differential operators

$$X_0 = \sum_{i,j=1}^{n} c_{ij}x_i\frac{\partial}{\partial x_j} + \frac{\partial}{\partial t},\ X_i = \frac{1}{\sqrt{2}}\sum_{j=1}^{m} \bar{a}_{ij}(x_0,t_0)\frac{\partial}{\partial x_j},\ i \in \{1,..,m\}, \tag{1.6}$$

as well as the operator

$$\tilde{L} = \tilde{L}_{(x_0,t_0)} := \sum_{i=1}^{m} X_i^2 + X_0 = \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(x_0,t_0)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i,j=1}^{n} c_{ij}x_i\frac{\partial}{\partial x_j} + \frac{\partial}{\partial t}. \tag{1.7}$$

To compensate for the lack of ellipticity, see (1.5), we assume that

$$\tilde{L} = \tilde{L}_{(x_0,t_0)} \text{ is hypoelliptic for every fixed } (x_0,t_0) \in \mathbb{R}^n \times \mathbb{R}_+. \tag{1.8}$$

Let $\mathrm{Lie}(X_0,X_1,..,X_m)$ denote the Lie algebra generated by the vector fields $X_0,X_1,..,X_m$. It is well-known that (1.8) can be stated in terms of the following Hörmander condition:

$$\text{rank Lie}(X_0,X_1,..,X_m) = n+1 \text{ at every point } (x,t) \in \mathbb{R}^{n+1}. \tag{1.9}$$

Another condition, equivalent to (1.8) and (1.9), is that there exists a basis for $\mathbb{R}^n$ such that the matrix C has the form

$$\begin{pmatrix} * & C_1 & 0 & \cdots & 0 \\ * & * & C_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & C_l \\ * & * & * & \cdots & * \end{pmatrix} \tag{1.10}$$

where C_j, for $j \in \{1,..,l\}$, is a $m_{j-1} \times m_j$ matrix of rank m_j, $1 \leq m_l \leq ... \leq m_1 \leq m_0$ and $m_0 + m_1 + ... + m_l = n$ while $*$ represents arbitrary matrices with constant entries. For a proof of the equivalence between the conditions stated above we refer to [19].

The problem in (1.3) can be approached using either techniques from the area of partial differential equations (PDEs) or from the area of stochastic differential equations (SDEs). Focusing, to start with, on the PDE-perspective we note that the problem in (1.3) is very well understood, both from a theoretical as well from a numerical perspective, in the case $m = n$ as in this case, the operator in (1.2) is uniformly elliptic. In particular, Cauchy problems for uniformly elliptic parabolic operators is a classical topic in the numerical analysis of partial differential equations and we refer to [20] for the finite-difference method and to [3, 5, 20] for the finite element method. In the case $m < n$, the problem in (1.3) is less developed, in particular from a numerical perspective and concerning the theoretical aspects of the Cauchy problem in (1.3), for the operators in (1.2) in the case $m < n$, we refer to [9, 12] and the references therein. Concerning numerical methods based on partial differential equations and finite-difference schemes we are aware of a few papers focusing on degenerate parabolic operators of Kolmogorov type and in these works the authors attempt to develop appropriate finite-difference schemes for the problem at hand, see [1, 8, 7, 23]. To understand the difficulties involved when discretizing the problem (1.3), in the case $m < n$ and using finite-differences, and this is in contrast to the case $m = n$, we recall that the natural setting for operators satisfying a Hörmander condition is that of the, to the Lie algebra, associated Lie group. In particular, as shown in [19] the relevant Lie group related to the operator $\tilde{L}$ in (1.7), and hence to degenerate parabolic operators of Kolmogorov type, is defined using the group law

$$(x,t) \circ (y,s) = (y + E(s)x, t+s),\ \ E(s) = \exp(-sC^*),\ \ (x,t),(y,s) \in \mathbb{R}^{n+1}. \tag{1.11}$$

Moreover, based on the block structure of C defined in (1.10) there is a natural family of dilations

$$\delta_r = \mathrm{diag}(r^{q_0} I_m, r^{q_1} I_{m_1}, .., r^{q_l} I_{m_l}, r^2),\ \ r > 0,\ \ q_k = 2k+1,\ \ k \in \{0,1,...,l\}, \tag{1.12}$$

associated to the Lie group. In (1.12) I_k, $k \in \mathbb{Z}_+$, is the k-dimensional unit matrix and δ_r is by definition a diagonal matrix. Moreover,

$$q+2,\ \ q := q_0 m + q_1 m_1 + ... + q_l m_l, \tag{1.13}$$

is said to be the homogeneous dimension of $\mathbb{R}^{n+1}$ defined with respect to the dilations $\{\delta_r\}_{r>0}$. Furthermore, we split the coordinate $(x,t) \in \mathbb{R}^{n+1}$ as $(x,t) = (x^{(0)}, x^{(1)}, ..., x^{(l)}, t)$ where $x^{(0)} \in \mathbb{R}^m$ and $x^{(j)} \in \mathbb{R}^{m_j}$ for all $j \in \{1,..,l\}$. Based on this we define

$$||(x,t)|| = \sum_{j=0}^{l} |x^{(j)}|^{\frac{1}{q_j}} + |t|^{\frac{1}{2}} \tag{1.14}$$

and we note that $||\delta_r(x,t)|| = r||(x,t)||$. The problem when discretizing the problem in (1.3) using finite-differences, in the case $m < n$, stems from the fact that the discretization has to respect the more involved Lie group structure as well as the anisotropic dilations $\{\delta_r\}_{r>0}$. In particular, standard rectangular grids used for elliptic problems can be proved to not perform optimal in the case $m < n$, see [7] for instance. Concerning the potential use of the finite element method, in the case $m < n$, we refer to [13] and the references therein. Next, focusing on the SDE-perspective we note that there are several algorithms for solving the problem in (1.3) using the Feynman-Kac formula, stochastic representation formulas and Euler schemes for the underlying system of stochastic differential equations and we refer to [2, 16, 27, 28] for details. In particular, the rate of convergence, in the setting of the Euler scheme, for smooth functions g and uniform time-steps, as well as a priori error expansion, are presented in [2, 28]. The a priori error expansion established in [28] is proved to be valid, assuming in addition that the underlying partial differential operator fulfills a Hörmander condition, also in the case when g is only measurable and bounded. Another approach is presented in [21] and uses cubature formulas on Wiener spaces. In this case there is no need to assume ellipticity for the underlying partial differential operator, instead the usefulness of the method depends on how well the coefficients can be approximated with polynomials. Finally, in [27] a method based on a posteriori error estimates and adaptive weak approximations of SDEs is presented in the case $m = n$. In this paper we focus on the case $m < n$ and we show, by applying and building on the pioneering work in [27], that one can circumvent all of the problems described above in the context of finite-difference schemes, by using a posteriori error estimates and adaptive weak approximations of SDEs. We are convinced that the approached presented in this paper will be useful in many areas where degenerate parabolic operators of Kolmogorov type occur. Moreover, to our knowledge the analysis developed in this paper has previously not been discussed in the literature in the setting of degenerate parabolic operators of Kolmogorov type even though the case $m = n$, i.e., the case of uniformly elliptic operators, is developed in [27]. Finally, for more general descriptions of the Monte Carlo method we refer to [6, 10, 24].

To briefly outline the way we proceed we first note that we throughout the paper consider the problem in (1.3) assuming that (1.5) holds and that C satisfies (1.10). Moreover, concerning regularity the appropriate regularity assumptions on $A_{i,j}$, b_i and g are defined and discussed in the bulk of the paper. We approach the problem in (1.3) using stochastic differential equations and we let

$$X_i(t) = x_i + \int_0^t \left(b_i(X(s), s) + \sum_{j=1}^n c_{ij} X_i(s) \right) ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(X(s), s) dW_j(s), \tag{1.15}$$

for $i \in \{1, ..., n\}$. We let $X(t) = (X_1(t), ..., X_n(t))^*$ denote the corresponding vector. In (1.15) $(W(t))_{0 \leq t \leq T}$, $W(t) = (W_1(t), ...W_m(t))^*$, is a standard Brownian motion in $\mathbb{R}^m$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ with the usual assumptions on $(\mathcal{F}_t)_{0 \leq t \leq T}$. By a standard Brownian motion in $\mathbb{R}^m$ we mean that the components are independent one-dimensional Brownian motions. Note that the vector $b(x, t) = (b_1(x, t), ..., b_n(x, t))^*$ satisfies $b_{m+1} \equiv ... \equiv b_n \equiv 0$. Assuming appropriate regularity conditions on the coefficients b_i, σ_{ij}, this will be discussed in detail below, one can combine results in [9, 26] to ensure existence and uniqueness, assuming that (1.5) holds and that C is as in (1.10), of a solution to the system in (1.15). For simplicity in the following

$$\mu_i(x, s) := b_i(x, s) + \sum_{j=1}^n c_{ij} x_i \text{ for } i \in \{1, ..., n\} \tag{1.16}$$

and we rewrite (1.15) as

$$X_i(t) = x_i + \int_0^t \mu_i(X(s), s) ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(X(s), s) dW_j(s). \tag{1.17}$$

Moreover, assuming that g is sufficiently regular one can use the Feynman-Kac formula to conclude that the unique solution to the problem in (1.3) is given as

$$u(x, t) = E[g(X(T)) | X(t) = x]. \tag{1.18}$$

Based on (1.18) we construct an approximation of the solution to (1.3) using the Euler scheme associated to the system in (1.15). In particular, given a time horizon of T we let $\{t_k\}_{k=0}^N$ define a partition Δ of the interval $[0, T]$, i.e., $0 = t_0 < t_1 < .... < t_{N-1} < t_N = T$, and we let

$\Delta t_k = t_{k+1} - t_k$ for $k \in \{0, ..., N-1\}$. Let $\{X(t), t \in [0,T]\}$ solve (1.15). In the following we let $\{X^\Delta(t), t \in [0,T]\}$ denote the continuous Euler approximation of $\{X(t), t \in [0,T]\}$ defined as follows. $X^\Delta(t)$ satisfies, for $k \in \{0, ..., N-1\}$, the difference equation

$$\begin{aligned} X^\Delta(t_{k+1}) &= X^\Delta(t_k) + \mu(X^\Delta(t_k), t_k)\ \Delta t_k + \sum_{j=1}^{m} \sigma_j(X^\Delta(t_k), t_k)\ \Delta W_j(t_k), \\ X^\Delta(t_0) &= x. \end{aligned} \tag{1.19}$$

In (1.19) $\Delta W_j(t_k) = W_j(t_{k+1}) - W_j(t_k)$. Moreover, $\{X^\Delta(t), t \in \{t_0, ..., t_N\}\}$ is often referred to as the associated discrete Euler approximation. In the following we will also make use of the function $\varphi(t) = \sup\{t_i : t_i \leq t\}$ which is defined whenever $t \in [0,T]$. Using this notation we define $\{X^\Delta(t), t \in [0,T]\}$ through the relation

$$X^\Delta(t) = X^\Delta(\varphi(t)) + \int_{\varphi(t)}^{t} \mu(X^\Delta(\varphi(s)), \varphi(s))ds + \sum_{j=1}^{m} \int_{\varphi(t)}^{t} \sigma_j(X^\Delta(\varphi(s)), \varphi(s))dW_j(s). \tag{1.20}$$

$\{X^\Delta(t), t \in [0,T]\}$ is referred to as the continuous Euler approximation. Let

$$u^\Delta(x, t_k) = E[g(X^\Delta(T))|X^\Delta(t_k) = x] \text{ for } k \in \{0, ..., N-1\},\ x \in \mathbb{R}^n. \tag{1.21}$$

The standard stochastic method for determining $u^\Delta(x) = u^\Delta(x, 0)$ is to use the Monte Carlo estimator

$$u^{\Delta,M}(x) = \frac{1}{M} \sum_{l=1}^{M} g(X^\Delta(T, \omega_l)), \tag{1.22}$$

where M is some positive integer and $\{\omega_l\}_{l=1}^M$ represents M realizations of the discrete Euler approximation of $\{X(t) : t \in [0,T]\}$. In particular, we see that

$$u(x) = u^{\Delta,M}(x) + \underbrace{u(x) - u^\Delta(x)}_{E_d^\Delta(x)} + \underbrace{u^\Delta(x) - u^{\Delta,M}(x)}_{E_s^{\Delta,M}(x)}, \tag{1.23}$$

where $E_d^\Delta(x)$ and $E_s^{\Delta,M}(x)$ represent, respectively, the time-discretization error and the statistical error. While $E_s^{\Delta,M}(x)$ can be controlled using the central limit theorem and the Berry-Esséen theorem, see Section 4, $E_d^\Delta(x)$ can be expressed as

$$E_d^\Delta(x) = \frac{1}{M} \sum_{l=1}^{M} \sum_{k=0}^{N-1} \rho_k(x, \omega_l) + R^{\Delta,M}(x) \tag{1.24}$$

where the error indicator $\rho_k(x, \omega_l)$ is, as indicated, computable based on the scenarios $\{\omega_l\}_{l=1}^M$ while the reminder $R^{\Delta,M}(x)$ is of lower order compared to the first term to the right in (1.24). In particular, (1.24) is an expansion of $E_d^\Delta(x)$ which is computable in a posteriori form. Based on (1.24) one can then proceed as [27] to define an adaptive algorithm based on which one can ensure, with high probability, that the time-discretization errors, as well as the statistical errors, are within a user defined error tolerance.

The rest of the paper is organized as follows. In Section 2, which is of preliminary nature, we introduce notation and briefly review a few fact from the Malliavin calculus, the latter being an important tool in the forthcoming sections. In this section we also derive a priori estimates for degenerate parabolic equations. In Section 3 we derive the expansion of the time-discretization error, $E_d^{\Delta}(x)$, described above. In Section 4 we briefly discuss how to control $E_s^{\Delta,M}(x)$. In Section 5 we apply the method developed to the problem of pricing European derivatives in the framework of the stochastic volatility model suggested by Hobson and Rogers, see [11]. Moreover, we compare the efficiency of the method outlined to the recently developed methods in [7] which are based on finite-differences. We emphasize that one important advantage of the method outlined in this paper, compared to others, and in particular to the method developed in [7], is that one can ensure, with high probability, that the method presented here produces a result, given a user defined error tolerance, which is within the error tolerance of the correct value.

2 Preliminaries and notation

Throughout the paper we will write $\partial_i f$ for $\frac{\partial f}{\partial x_i}$, $\partial_{ij} f$ for $\frac{\partial^2 f}{\partial x_i \partial xj}$ and so on. If $f = f(x,t)$, $(x,t) \in \mathbb{R}^n \times \mathbb{R}_+$, then in general $\partial_i, \partial_{ij}$ and so forth will only refer to differentiation with respect to the space variable x. For a multiindex $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, $\alpha_i \in \mathbb{Z}_+$, we define $|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_n$ and we let ∂_α denote differentiation with respect to the space variables according to the multiindex α. Given an open set $O \subset \mathbb{R}^n$ we let $C_b^k(O)$ denote functions $f : O \to \mathbb{R}$ which are k times continuously differentiable whose derivatives are all bounded. Similarly, we let $C_p^k(\mathbb{R}^n)$ denote k times continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$, for which there exist constants c_α, $q_\alpha \in \mathbb{Z}_+$, such that

$$|\partial_\alpha f(x)| \leq c_\alpha (1 + |x|^{q_\alpha}) \text{ whenever } x \in \mathbb{R}^n,\ |\alpha| \leq k. \tag{2.1}$$

Furthermore, we let $C_b^k(\mathbb{R}^n \times \mathbb{R}_+)$, $k \in \mathbb{Z}_+$, denote the space of all functions defined on $\mathbb{R}^n \times \mathbb{R}_+$ which have continuous and bounded partial derivatives, in both space and time, up to order k. Note that to be in the space $C_b^k(\mathbb{R}^n \times \mathbb{R}_+)$ the function it self does not have to be bounded. We also let $C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+) = \bigcap_{k \geq 1} C_b^k(\mathbb{R}^n \times \mathbb{R}_+)$. Furthermore, if $X(t)$, $t \in [0,T]$, is a stochastic process satisfying (1.17) then we, at some instances, let $X_t(x)$ denote the process $X(t)$ with initial data $X(0) = x$. For a random variable X we denote the variance by $\mathrm{Var}[X]$. Finally, given a set $I \subset \mathbb{R}$ we let χ_I denote the indicator function of the set I.

In the following we will briefly introduce some basic facts from Malliavin calculus which we will use in the forthcoming sections and we will, in particular, prove an a priori estimate for degenerate parabolic equations. For expositions of the Malliavin calculus we refer to [14, 22] and we refer to [18, 25] for a somewhat different approach to stochastic flows. Below we follow [22] and we will use the notation introduced in [22]. In particular, to proceed we let $(\Omega, \mathcal{F}, (\mathcal{F})_t, P)$ be a probability space with a filtration generated by a Wiener process $W(t) \in \mathbb{R}^m$. For functions $h : \mathbb{R}_+ \to \mathbb{R}^m$ which belong to the space $L^2([0,T])$, the space of functions defined on $[0,T]$ which are square integrable with respect to the Lebesgue measure dt, we define $W(h) := \int_0^T \langle h(t), dW(t) \rangle$ where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product on $\mathbb{R}^m$. Let $\mathcal{S}$ be the class of Wiener polynomials, i.e., the class of smooth random variables F such that

$$F = f(W(h_1), W(h_2), ..., W(h_n)) \tag{2.2}$$

for some functions $f \in C_p^\infty(\mathbb{R}^n)$ and $h_i \in L^2[0,T],\ i \in \{1,2,...,n\}$. The first order derivative of a smooth random variable $F \in \mathcal{S}$ is a stochastic process $DF = \{D_t F\}_{t\in[0,T]}$ given by

$$D_t F = \frac{\partial}{\partial x_i} f(W(h_1), W(h_2), ..., W(h_n))h_i(t). \tag{2.3}$$

We consider DF as an element of $L^2([0,T]\times\Omega, \mathcal{B}(T)\times\mathcal{F}, dt\times P)$ where $\mathcal{B}$ denotes the σ-algebra of Borel sets on $\mathbb{R}$. We let

$$||F||_{1,p} := \left[E\left[|F|^p\right] + E\left[||DF||^p_{L^2[0,T]}\right]\right]^{1/p}, \tag{2.4}$$

and we let $\mathbb{D}^{1,p}$ be the closure of $\mathcal{S}$ with respect to this norm. The domain of the derivative operator D, in $L^p(\Omega)$, is $\mathbb{D}^{1,p}$. The k-th order derivative of F is defined as

$$D^k_{t_1,t_2,...,t_k} F = D_{t_1} D_{t_2} \cdots D_{t_k} F \tag{2.5}$$

and we let $\mathbb{D}^{k,p}$ be the closure of $\mathcal{S}$ with respect to

$$||F||_{k,p} := \left[E\left[|F|^p\right] + \sum_{j=1}^{k} E\left[||D^j F||^p_{L^2((0,T)^j)}\right]\right]^{1/p}. \tag{2.6}$$

Finally, we let $\mathbb{D}^\infty = \cap_{k\geq 1} \cap_{p\geq 1} \mathbb{D}^{k,p}$. For the proof of the following two theorems we refer to Theorem 2.2.2 and Theorem 1.5.1 in [22].

Theorem 2.1 *Let $X(t),\ t \in [0,T]$, be a stochastic process satisfying (1.17) and assume that $\mu_i(x,t), \sigma_{ij}(x,t) \in C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+)$. Assume, in addition that $\mu_i(0,t)$ and $\sigma_{ij}(0,t)$ are bounded for all $t \in [0,T]$. Then $X_i(t) \in \mathbb{D}^\infty$ for all $t \in [0,T]$.*

Theorem 2.2 *Suppose that $F = (F^1, F^2, ..., F^m) \in (\mathbb{D}^\infty)^m$ and that $\psi \in C_p^\infty(\mathbb{R}^m)$. Then $\psi(F) \in \mathbb{D}^\infty$ and*

$$D(\psi(F)) = \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \psi(F) DF^i. \tag{2.7}$$

In this section we prove the following lemma.

Lemma 2.3 *Consider the Cauchy problem in (1.3) with terminal data $g \in C_p^\infty(\mathbb{R}^n)$. Assume that (1.5) holds and that C satisfies (1.10). Let μ_i be defined as in (1.16) and assume that $\mu_i(x,t), \sigma_{ij}(x,t) \in C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+)$. Let $X(t),\ t \in [0,T]$, be a stochastic process satisfying (1.17) with deterministic initial value. Then*

$$u(x,t) = E[g(X(T)) \mid X(t) = x] \tag{2.8}$$

is a solution to (1.3). Furthermore, $u(x,t)$ is infinitely differentiable and there exists for every multiindex α, constants c_α, $q_\alpha \in \mathbb{Z}_+$, such that

$$\sup_{0\leq t\leq T} |\partial_\alpha u(x,t)| \leq c_\alpha(1 + |x|^{q_\alpha}). \tag{2.9}$$

In particular, u is the unique infinitely differentiable solution satisfying the growth condition in (2.9).

Proof. Using the assumptions on μ_i, σ_{ij} and $X(t)$ together with Theorem 2.1 it follows that $X_i(t) \in \mathbb{D}^\infty$ for all $t \in [0,T]$. Actually, following [18, 25], there exists a smooth version of the stochastic flow $x \longmapsto X_t(x)$. For $k \in \mathbb{Z}_+$ the family of processes $\{\partial_\alpha X_t(x) : |\alpha| \leq k\}$ solves a system of SDEs with Lipschitz coefficients. Moreover, assuming $g \in C_p^\infty(\mathbb{R}^n)$ and using Theorem 2.2, we see that $g(X(t)) \in \mathbb{D}^\infty$ for all $t \in [0,T]$. As a consequence there exist constants c and q such that

$$\partial_i E[g(X_t(x))] \leq c(1+|x|^q). \tag{2.10}$$

Above we used that when we have Lipschitz coefficients

$$E\left[\sup_{0\leq t\leq T} |X_t(x)|^{2\beta}\right] \leq K(T)\left(1+|X_t(x)|^{2\beta}\right), \tag{2.11}$$

for some increasing function $K(T)$, see [16]. By induction it can be shown that for $|\alpha| \leq k$ there exist constants c_α and q_α such that

$$\partial_\alpha E[g(X_t(x))] \leq c_\alpha(1+|x|^{q_\alpha}). \tag{2.12}$$

The existence of a classical solution to (1.3) was proved in [9] and that u given by (2.8) is indeed a solution to (1.3) follows from an application of Itô´s lemma to $u(t, X_t)$ in analogue with the proof of the Feynman-Kac formula, see Theorem 5.7.6 in [15]. Finally, note that uniqueness follows immediately. Indeed assuming we have two smooth solutions u and v with polynomial growth, we may apply Itô´s lemma to u and v. Then

$$(u(T,X(T))-u(x,t))-(v(T,X(T))-v(x,t)) = \int_t^T (Lu-Lv)dt + \int_t^T \sum_{i=1}^n \frac{\partial}{\partial x_i}(u-v)dW_i. \tag{2.13}$$

Now, since the terminal data coincide and since $Lu = Lv = 0$, we may take expected values to obtain $u(x,t) - v(x,t) = -E[\int_t^T \sum_{i=1}^n \frac{\partial}{\partial x_i}(u-v)dW_i] = 0$. □

3 An a posteriori error expansion

In this section we show how to derive, by using the regularity theorem of the previous section and by proceeding as in [27], the a posteriori error expansion for the time-discretization error, $E_d^\Delta(x)$, referred to in the introduction. Throughout the section we will impose the following assumption.

Assumption 3.1 *Let L be the operator in (1.2), assume that (1.5) holds and that C satisfies (1.10). Let μ_i be defined as in (1.16), let $X(t)$, $t \in [0,T]$, satisfy (1.17). Assume that μ_i, $\sigma_{ij} \in C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+)$, for all $i \in \{1,..,n\}$, $j \in \{1,..,m\}$, and that $g \in C_p^\infty(\mathbb{R}^n)$.*

The assumptions above differs from the assumptions made in [27] in that we consider a wider class of operators L at the expense of more regularity assumptions on μ_i and σ_{ij}. In [27] they only need to assume that μ_i, $\sigma_{ij} \in C_b^{m_0}(\mathbb{R}^n \times \mathbb{R}_+)$ for some $m_0 > [n/2] + 10$ and in contrast with our approach they have a stochastic initial datum $X(0)$ (see Lemma 2.1 on p.6 in [27]). In the following we reuse the notation introduced in the introduction and in particular we let $\{X^\Delta(t), t \in [0,T]\}$ denote the continuous Euler approximation introduced in (1.20) and associated to the partition Δ defined by $\{t_k\}_{k=0}^N$, $0 = t_0 < t_1 < ... < t_{N-1} < t_N = T$. Recall that $\varphi(t) := \sup\{t_i;\ t_i \le t\}$. In the following, let

$$\begin{aligned} \mu_i^\Delta(X^\Delta(t),t) &= \mu_i(X^\Delta(\varphi(t)),\varphi(t)), \quad \sigma_i^\Delta(X^\Delta(t),t) = \sigma_i(X^\Delta(\varphi(t)),\varphi(t)), \\ a_{ij}^\Delta(X^\Delta(t),t) &= a_{ij}(X^\Delta(\varphi(t)),\varphi(t)), \end{aligned} \tag{3.1}$$

whenever $t \in [0,T]$. We will derive the appropriate error expansion based on Assumption 3.1. To proceed we first introduce, in analogue with [27], appropriate dual functions. In particular, we define, whenever $i \in \{1,...,n\}$ and $x \in \mathbb{R}^n$,

$$c_i(x,t_k) = x_i + \mu_i(x,t_k)\Delta t_k + \sum_{j=1}^m \sigma_{ij}(x,t_k)\Delta W_j(t_k) \text{ for } k \in \{0,...,N-1\}. \tag{3.2}$$

The discrete dual functions, associated to $g(X(T))$, are defined, whenever $i \in \{1,...,n\}$, recursively as follows

$$\begin{aligned} \phi_i(t_N) &= \partial_i g(X^\Delta(t_N)), \\ \phi_i(t_k) &= \sum_{l=1}^n \partial_i c_l(X^\Delta(t_k),t_k)\phi_l(t_{k+1}) \text{ whenever } k \in \{0,...,N-1\}. \end{aligned} \tag{3.3}$$

The first variations of the dual functions are defined as

$$\phi'_{ij}(t_k;\omega) := \partial_{x_j(t_k)}\phi_i(t_k;\omega) \equiv \frac{\partial \phi_i(t_k; X^\Delta(t_k) = x)}{\partial x_j} \tag{3.4}$$

and they satisfy, whenever $i,j \in \{1,...,n\}$ and $k \in \{0,...,N-1\}$,

$$\begin{aligned} \phi'_{ij}(t_N) &= \partial_{ij} g(X^\Delta(t_N)), \\ \phi'_{ij}(t_k) &= \sum_{r=1}^n \sum_{l=1}^n \partial_i c_r(X^\Delta(t_k),t_k)\partial_j c_l(X^\Delta(t_k),t_k)\phi'_{rl}(t_{k+1}) \\ &+ \sum_{r=1}^n \partial_{ij} c_r(X^\Delta(t_k),t_k)\phi_r(t_{k+1}). \end{aligned} \tag{3.5}$$

We also introduce, for $k \in \{1,..,N-1\}$, the variance

$$\begin{aligned} \sigma_k^2 &= \operatorname{Var}\left[\sum_{i=1}^n \left[\mu_i(X^\Delta(t_{k+1}),t_{k+1}) - \mu_i(X^\Delta(t_k),t_k)\right]\phi_i(t_{k+1})\right] \\ &+ \operatorname{Var}\left[\sum_{i=1}^m \sum_{j=1}^m \left[a_{ij}(X^\Delta(t_{k+1}),t_{k+1}) - a_{ij}(X^\Delta(t_k),t_k)\right]\phi'_{ij}(t_{k+1})\right]. \end{aligned} \tag{3.6}$$

The purpose of the section is to derive the following a posteriori error expansion of the discretization error.

Theorem 3.2 *Assume that $X(t) = (X_1(t), ..., X_n(t)) \in \mathbb{R}^n$, $t \in [0, T]$, solves (1.17) and that Assumption 3.1 holds. Then $E_d^\Delta - \bar{E}_d^{\Delta,M}$ equals*

$$\begin{aligned} &\sum_{i=1}^{n} \sum_{l=1}^{M} \sum_{k=0}^{N-1} \left[\mu_i(X^\Delta(t_{k+1}, \omega_l), t_{k+1}) - \mu_i(X^\Delta(t_k, \omega_l), t_k)\right] \phi_i(t_{k+1}, \omega_l) \frac{\Delta t_k}{2M} \\ + \quad &\sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{l=1}^{M} \sum_{k=0}^{N-1} \left[a_{ij}(X^\Delta(t_{k+1}, \omega_l), t_{k+1}) - a_{ij}(X^\Delta(t_k, \omega_l), t_k)\right] \phi'_{ij}(t_{k+1}, \omega_l) \frac{\Delta t_k}{2M}, \end{aligned} \tag{3.7}$$

where ϕ and ϕ' are the discrete dual functions satisfying (3.3) and (3.5) and

$$\bar{E}_d^{\Delta,M} = \sum_{k=0}^{N-1} (\Delta t_k)^2 \left\{ \mathcal{O}(\Delta t_k) + \sum_{r=k}^{N-1} \mathcal{O}((\Delta t_r)^2) \right\} + \sum_{k=0}^{N-1} \int_0^T I_{k,M} dt. \tag{3.8}$$

The random variable $\sqrt{M} I_{k,M}$ converges, as $M \to \infty$, for each $k \in \{0, 1, .., N-1\}$, to a normally distributed random variable with zero mean and variance σ_k^2, see (3.6).

By Assumption 3.1 we are assuming that $g \in C_p^\infty(\mathbb{R}^n)$. However, for the proof of Theorem 3.2 we only have to assume that

$$|\partial_\alpha g(x)| \leq c_\alpha (1 + |x|^{q_\alpha}), \tag{3.9}$$

for some constants $c_\alpha \in [1, \infty)$, $q_\alpha \in \mathbb{Z}_+$, for all multiindices α, $|\alpha| \leq 6$. Furthermore, we emphasize that Theorem 3.2 was proved in [27] in the case $m = n$, i.e., in the uniformly elliptic case, and, in fact, in the case $m < n$ one can proceed along the same lines once the appropriate regularity theory for u is established.

3.1 Proof of Theorem 3.2

We divide the proof of Theorem 3.2 into a number of lemmas. We emphasize that we proceed along the same lines as in [27], the only changes being in motivating regularity and boundedness. In doing so we note that the variational processes described in [27] are nothing but stochastic flows and we use results from Malliavin calculus to complete the proofs.

Lemma 3.3 *Assume that the assumptions in Theorem 3.2 are fulfilled. Then $E_d^\Delta = E_1^\Delta + E_2^\Delta$ where*

$$\begin{aligned} E_1^\Delta &= \sum_{i=1}^{n} \int_0^T E\Big[[\mu_i(X^\Delta(t), t) - \mu_i^\Delta(X^\Delta(t), t)] \partial_i u(X^\Delta(t), t) \Big] dt, \\ E_2^\Delta &= \sum_{i=1}^{m} \sum_{j=1}^{m} \int_0^T E\Big[[a_{ij}(X^\Delta(t), t) - a_{ij}^\Delta(X^\Delta(t), t)] \partial_{ij} u(X^\Delta(t), t) \Big] dt. \end{aligned} \tag{3.10}$$

Proof. As a consequence of Lemma 2.3, $u(x, t) = E[g(X(T)|X(t) = x]$ is sufficiently smooth to allow us to apply Ito's lemma to the function $u(X^\Delta(t), t)$. One can now proceed as in the second half of the proof of Lemma 2.1. in [27] □

We next prove the following lemma.

Lemma 3.4 *Assume that the assumptions in Theorem 3.2 are fulfilled and let E_1^Δ and E_2^Δ be as in the statement of Lemma 3.3. Then*

$$\begin{aligned} E_1^\Delta &- \sum_{i=1}^{n}\sum_{k=0}^{N-1} E\left[\left(\mu_i(X^\Delta(t_{k+1}),t_{k+1}) - \mu_i(X^\Delta(t_k),t_k)\right)\partial_i u(X^\Delta(t_{k+1}),t_{k+1})\right]\frac{\Delta t_k}{2} \\ &= \sum_{k=0}^{N-1}\mathcal{O}\left((\Delta t_k)^3\right) \end{aligned} \tag{3.11}$$

$$\begin{aligned} E_2^\Delta &- \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=0}^{N-1} E\left[\left(a_{ij}(X^\Delta(t_{k+1}),t_{k+1}) - a_{ij}(X^\Delta(t_k),t_k)\right)\partial_{ij} u(X^\Delta(t_{k+1}),t_{k+1})\right]\frac{\Delta t_k}{2} \\ &= \sum_{k=0}^{N-1}\mathcal{O}\left((\Delta t_k)^3\right). \end{aligned} \tag{3.12}$$

Proof. We let

$$\begin{aligned} f(X^\Delta(t),t) &= (\mu_i(X^\Delta(t),t) - \mu_i^\Delta(X^\Delta(t),t))\partial_i u(X^\Delta(t),t), \\ \hat{f}(t) &= (\mu_i(X^\Delta(t_{k+1}),t_{k+1}) - \mu_i(X^\Delta(t_k),t_k))\partial_i u(X^\Delta(t_{k+1}),t_{k+1})\frac{t-t_k}{\Delta t_k}, \end{aligned} \tag{3.13}$$

whenever $t_k \le t < t_{k+1}$, $k \in \{0,..,N-1\}$, and we let $\hat{f}(t)$ be a piecewise linear function such that

$$\hat{f}(t_k) = f(X^\Delta(t_k),t_k) \text{ for every } k \in \{0,..,N\}. \tag{3.14}$$

Moreover, as

$$\int_0^T E[f(X^\Delta(t),t) - \hat{f}(t)]dt \tag{3.15}$$

equals

$$E_1^\Delta - \sum_{i=1}^{n}\sum_{k=0}^{N-1} E\left[\left(\mu_i(X^\Delta(t_{k+1}),t_{k+1}) - \mu_i(X^\Delta(t_k),t_k)\right)\partial_i u(X^\Delta(t_{k+1}),t_{k+1})\right]\frac{\Delta t_k}{2}, \tag{3.16}$$

we see that to prove the first estimate in Lemma 3.4 it is enough to estimate the integral in (3.15). Let $k \in \{0,..,N-1\}$. Then, using integration by parts and (3.14)

$$\left|\int_{t_k}^{t_{k+1}} E[f(X^\Delta(t),t) - \hat{f}(t)]dt\right| \le \frac{(\Delta t_k)^2}{8}\int_{t_k}^{t_{k+1}}\left|\frac{d^2}{dt^2}E[f(X^\Delta(t),t)]dt\right|. \tag{3.17}$$

Next using Itô's lemma, the conditions on μ_i, σ_{ij} stated in Assumption 3.1, Lemma 2.3 as well as the differentiability of f, we see that

$$\frac{d^2}{dt^2}E[f(X^\Delta(t),t)] = E\left[\frac{d^2}{dt^2}f(X^\Delta(t),t)\right] = E\left[L^2 f(X^\Delta(t),t)\right], \tag{3.18}$$

where L is defined as in (1.2). Moreover, $L^2 f(X^\Delta(t), t)$ is a sum of terms, each consisting of products of μ_i, μ_i^Δ, a_{kl}, a_{kl}^Δ and derivatives of order less than or equal to four of these functions as well as derivatives of u of order less than or equal to five. Furthermore, as derivatives of μ_i, a_{kl} are bounded and as for all multiindices α there exist constants $\tilde{c}_\alpha$, $\tilde{q}_\alpha \in \mathbb{Z}_+$, see Lemma 2.3, such that $|\partial_\alpha u(x,t)| \leq \tilde{c}_\alpha(1 + |x|^{\tilde{q}_\alpha})$ it follows that

$$\left|L^2 f(X^\Delta(t), t)\right| \leq \widetilde{c}(1 + |X^\Delta(t)|^{\widetilde{q}}) \tag{3.19}$$

for some constants $\widetilde{c} \in \mathbb{R}_+$, $\widetilde{q} \in \mathbb{Z}_+$. As the initial condition is deterministic and μ_i, $\sigma_{ij} \in C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+)$, it follows that $E[|X(t)|^p + |X^\Delta(t)|^p] \leq \widehat{c}$ for some constant $\widehat{c} \in \mathbb{R}_+$, depending on $\mu_i, \sigma_{ij}, n, x, p, t$ and Δ. In particular, for the proof of this in case of $X(t)$ we refer to Theorem 4.5.4 in [16] and we note that the result for the Euler discretization can be proved similarly. Put together, we see that there exist a constant $c \in \mathbb{R}_+$ such that

$$\left|\frac{d^2}{dt^2} E[f(X^\Delta(t), t)]\right| \leq c \tag{3.20}$$

whenever $t \in [t_k, t_{k+1})$ and hence the first conclusion of the lemma follows. The second conclusion of the lemma can be proved in a similar manner and in this case the only difference is that we now have to handle derivatives of $u(x,t)$ of order less than or equal to six. We omit the details. $\square$

Note that when proving this lemma we did follow the lines of the proof of Lemma 2.3 in [27], the difference is our motivation of the conclusion $\left|\frac{d^2}{dt^2} E[f(X^\Delta(t), t)]\right| \leq c$. Furthermore, it should be clear that the assumptions were necessary for Lemma 3.4 to hold.

Let

$$u^\Delta(x,t) = E[g(X^\Delta(T)) \,|X^\Delta(t) = x]. \tag{3.21}$$

The next lemma concerns u^Δ as an approximation of u. The proof of the lemma uses stochastic flows and Malliavin calculus.

Lemma 3.5 *Assume that the assumptions in Theorem 3.2 are fulfilled. Let $t_k \leq t < t_{k+1}$, and define $\Delta t_k(s) := \sum_{i=0}^{N-1} \Delta t_i \chi_{[t_i, t_{i+1})}(s)$. Then, for $|\alpha| \leq 4$,*

$$\partial_\alpha(u - u^\Delta)(X^\Delta(t), t) = \int_t^T \mathcal{O}(\Delta t_k(s)) ds, \tag{3.22}$$

and

$$\begin{aligned} &E\left[\left(\mu_i(X^\Delta(t), t) - \mu_i(X^\Delta(t_k), t_k)\right) \cdot \right. \\ &\qquad \left.\left(\partial_i u(X^\Delta(t), t) - \partial_i u^\Delta(X^\Delta(t), t)\right)\right] = \Delta t_k \int_t^T \mathcal{O}(\Delta t_k(s)) ds, \end{aligned} \tag{3.23}$$

$$\begin{aligned} &E\left[\left(a_{ij}(X^\Delta(t), t) - a_{ij}(X^\Delta(t_k), t_k)\right) \cdot \right. \\ &\qquad \left.\left(\partial_{ij} u(X^\Delta(t), t) - \partial_{ij} u^\Delta(X^\Delta(t), t)\right)\right] = \Delta t_k \int_t^T \mathcal{O}(\Delta t_k(s)) ds. \end{aligned} \tag{3.24}$$

Proof. We will use the same techniques as in the proof of Lemma 2.4 in [27]. To be able to do that we will note that the so called variational processes are in fact stochastic flows, and our assumptions imply a nice behavior of these flows, to be specified below. In this setting it is reasonable to consider two separate cases, namely when μ and σ are independent respectively dependent on t. We only supply the proof of the lemma assuming that μ, σ are independent of t. The general case then follows by introducing the additional variable $X_{n+1} = t$. The idea of the proof is to write down explicit expressions for $\partial_\alpha(u - u^\Delta)$ in terms of derivatives of g and certain processes associated to $X(t)$ and $X^\Delta(t)$ and then to use Itô's lemma repeatedly. In particular, the stochastic representation formulas for u and u^Δ can be expressed as

$$u(x,t) = E[g(X_{T-t}(x))],\ u^\Delta(x,t) = E[g(X^\Delta_{T-t}(x))], \tag{3.25}$$

where $X_{T-t}(x)$ is the stochastic process $X(T-t)$ which solves (1.17) but with initial condition $X(0) = x$, $X^\Delta_{T-t}(x)$ is defined similarly. $X_{T-t}(x)$ can, as we are assuming that μ_i and σ_{ij} are independent of t, also be interpreted as the stochastic process $X(T)$ with initial datum $X(t) = x$ resulting in the above expressions for u and u^Δ respectively. Using the assumption that $\mu_i, \sigma_{ij} \in C_b^\infty(\mathbb{R}^n)$ it follows, see Theorem 2.1, that $X(t) \in (\mathbb{D}^\infty)^n$. Moreover, the to $X(t)$ associated first variation process $X^{(1)}(t) = \left(\frac{\partial X_t(x)}{\partial x_i}\right)_{i=1}^n = \{X^{(1)}_{ij}(t)\}$, which is a stochastic flow, see Section 2.3 in [22], is in $(\mathbb{D}^\infty)^{n\times n}$ and satisfies, for $t > 0$, the stochastic differential equation

$$\begin{aligned} dX^{(1)}_{ij}(t) &= \sum_{k=1}^n \partial_k\mu_i(X_t(x))X^{(1)}_{kj}(t)dt + \sum_{k=1}^n\sum_{l=1}^m \partial_k\sigma_{il}(X_t(x))X^{(1)}_{kj}(t)dW_l(t),\\ X^{(1)}_{ij}(0) &= \delta_{ij}, \end{aligned} \tag{3.26}$$

where δ_{ij} is the Kronecker delta. Similarly, the to $X(t)$ associated second variational process $X^{(2)}(t) = \{X^{(2)}_{ijk}(t)\}$, satisfies, for $t > 0$, the stochastic differential equation

$$\begin{aligned} dX^{(2)}_{ijk}(t) &= \left(\sum_{r=1}^n \partial_r\mu_i(X_t(x))X^{(2)}_{rjk}(t) + \sum_{r=1}^n\sum_{s=1}^n \partial_{rs}\mu_i(X_t(x))X^{(1)}_{rj}(t)X^{(1)}_{sk}(t)\right)dt\\ &\quad + \sum_{l=1}^m\left(\sum_{r=1}^n \partial_r\sigma_{il}(X_t(x))X^{(2)}_{rjk}(t) + \sum_{r=1}^n\sum_{s=1}^n \partial_{rs}\sigma_{ij}(X_t(x))X^{(1)}_{rj}(t)X^{(1)}_{sk}(t)\right)dW_l(t),\\ X^{(2)}_{ijk}(0) &= 0. \end{aligned} \tag{3.27}$$

Finally, the to $X(t)$ associated third and fourth variational processes, $X^{(3)}(t) = \{X^{(3)}_{ijkl}(t)\}$ and $X^{(4)}(t) = \{X^{(4)}_{ijklm}(t)\}$, satisfy, for $t > 0$, the stochastic differential equations

$$\begin{aligned} dX^{(3)}_{ijkl} &= \partial_l(\text{right hand side of the main equation in (3.27)}),\\ X^{(3)}_{ijkl}(0) &= 0, \end{aligned} \tag{3.28}$$

and

$$\begin{aligned} dX^{(4)}_{ijklm} &= \partial_m\partial_l(\text{right hand side of the main equation in (3.27)}),\\ X^{(4)}_{ijklm}(0) &= 0, \end{aligned} \tag{3.29}$$

respectively. In particular. $X^{(1)}, X^{(2)}, X^{(3)}$ and $X^{(4)}$ are matrix valued processes of dimension $n \times n$, $n \times n \times n$, $n \times n \times n \times n$ and $n \times n \times n \times n \times n$ respectively. Moreover, by our assumption on μ_i, σ_{ij}, all components of $X^{(1)}, X^{(2)}, X^{(3)}$ and $X^{(4)}$ belong to $\mathbb{D}^\infty$, see [22]. Let $Y(t) := (X(t), X^{(1)}(t), X^{(2)}(t), X^{(3)}(t), X^{(4)}(t))$. Then there exist, as μ_i, $\sigma_{ij} \in C_b^\infty(\mathbb{R}^n)$, matrix valued functions $M, S \in C^\infty(\mathbb{R}^n \times \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n} \times \mathbb{R}^{n\times n\times n\times n} \times \mathbb{R}^{n\times n\times n\times n\times n})$ such that

$$dY(t) = M(Y(t))dt + S_j(Y(t))dW_j(t), \; Y(0) = (x, I_n, 0, 0, 0) \tag{3.30}$$

where I_n denotes the $n \times n$ identity matrix. In particular, $Y(t)$ is a vector of $d(n) := n + n^2 + ... + n^5 = n(n^5 - 1)/(n - 1)$ elements. As all components of Y belong to $\mathbb{D}^\infty$ we see that

$$\begin{aligned}
\partial_i u(x,t) &= E\left[\sum_{j=1}^n \partial_j g(X(T-t))X^{(1)}_{ji}(T-t)\right] \\
&=: E[f_i(Y(T-t))] \\
\partial_{ij} u(x,t) &= E\left[\sum_{k=1}^n\sum_{l=1}^n\sum_{m=1}^n \partial_{kl} g(X(T-t))X^{(1)}_{km}(T-t)\, X^{(1)}_{lj}(T-t)\right. \\
&\qquad \left. + \sum_{k=1}^n\sum_{l=1}^n\sum_{m=1}^n \partial_k g(X(T-t))X^{(2)}_{kmj}(T-t)\right] \\
&=: E[f_{ij}(Y(T-t))]
\end{aligned} \tag{3.31}$$

where

$$\begin{aligned}
f_i(Y(T-t)) &= \sum_{j=1}^n \partial_j g(X(T-t))X^{(1)}_{ji}(T-t), \\
f_{ij}(Y(T-t)) &= \sum_{k=1}^n\sum_{l=1}^n\sum_{m=1}^n \partial_{kl} g(X(T-t))X^{(1)}_{km}(T-t)\, X^{(1)}_{lj}(T-t) \\
&\qquad + \sum_{k=1}^n\sum_{l=1}^n\sum_{m=1}^n \partial_k g(X(T-t))X^{(2)}_{kmj}(T-t).
\end{aligned} \tag{3.32}$$

We have omitted to explicitly write down the dependency on the initial condition $X(0) = x$ although it should be clear that $Y(t)$ indeed depends on x. For a general multiindex α we have

$$\partial_\alpha u(x,t) = E[f_\alpha(Y(T-t))]$$

for an appropriate function f_α. It is worth noting that the Euler approximation of the variational process is equal to the variational process of the Euler approximation, i.e.,

$$\left\{\frac{\partial X^\Delta(x)}{\partial x_i}\right\}_{i=1}^n = X^{(1),\Delta}(T-t) \tag{3.33}$$

where $X^{(1),\Delta}(T-t)$ is the continuous Euler approximation of $X^{(1)}(T-t)$. To proceed we let, for $|\alpha| \le 4$, v^α solve the problem

$$\begin{aligned}
\frac{\partial}{\partial t} v^\alpha(y,t) + \sum_{i=1}^{d(n)} M_i(y,t)\partial_i v^\alpha(y,t) + \frac{1}{2}\sum_{i=1}^{d(n)}\sum_{j=1}^{n}\sum_{k=1}^{d(n)} S_{ij}(y,t)S^*_{kj}(y,t)\partial_{ik} v^\alpha(y,t) &= 0, \\
v^\alpha(\cdot, T) &= f_\alpha,
\end{aligned} \tag{3.34}$$

and we let

$$A_{ij}(y,t) = \frac{1}{2}\left[SS^*\right]_{ij}(y,t) \text{ whenever } (y,t) \in \mathbb{R}^{n(n^5-1)/(n-1)} \times \mathbb{R}_+. \tag{3.35}$$

Furthermore, given a partition we let Y^Δ be the to the vector valued process Y, see (3.30), associated continuous Euler approximation and we let $M^\Delta(Y^\Delta(t),t)$ and $A^\Delta(Y^\Delta(t),t)$ be defined in analogue with the definitions in (3.1). Arguing as in Lemma 3.3

$$\begin{aligned} \partial_\alpha(u-u^\Delta)(X^\Delta(t),t) &= \int_t^T E\left[\sum_{i=1}^{d(n)}(M_i - M_i^\Delta)\partial_i v^\alpha(Y^\Delta(s),s)|\mathcal{F}_t\right] ds \\ &+ \int_t^T E\left[\sum_{i=1}^{d(n)}\sum_{j=1}^{d(n)}(A_{ij} - A_{ij}^\Delta)\partial_{ij} v^\alpha(Y^\Delta(s),s)|\mathcal{F}_t\right] ds. \end{aligned} \tag{3.36}$$

In particular, we introduce the short notation

$$\partial_\alpha(u-u^\Delta)(X^\Delta(t),t) = \int_t^T E\left[\widehat{f}_\alpha(Y^\Delta(s),s)|\mathcal{F}_t\right] ds \tag{3.37}$$

for the formula derived in the last display and for an appropriate functions $\widehat{f}_\alpha$. $\mathcal{F}_t$ denotes the σ-algebra generated by $\{W(s) : 0 \leq s \leq t\}$. Let L^Δ denote the operator

$$L^\Delta \widehat{f}_\alpha(Y^\Delta(t),t) = \left(\frac{\partial}{\partial t}\widehat{f}_\alpha + \sum_{i=1}^{d(n)} M_i^\Delta \partial_i \widehat{f}_\alpha + \sum_{i=1}^{d(n)}\sum_{j=1}^{d(n)} A_{ij}^\Delta \partial_{ij}\widehat{f}_\alpha\right)(Y^\Delta(t),t). \tag{3.38}$$

Then, again using Itô's lemma we get, for $t_k \leq s < t_{k+1}$,

$$E\left[\widehat{f}_\alpha(Y^\Delta(s),s)|\mathcal{F}_t\right] = \int_{t_k}^s E\left[L^\Delta \widehat{f}_\alpha(Y^\Delta(u),u)|\mathcal{F}_t\right] du = \mathcal{O}(\Delta t_k). \tag{3.39}$$

The equality in the last display follows from the fact that $M, S \in C_b^\infty$. The proof of the first statement in the lemma is therefore complete. To prove the second statement, we define

$$\widetilde{f}(Y^\Delta(t),t) := \widetilde{f}(X^\Delta(t),t) := (\mu_i - \mu_i^\Delta)\partial_i(u-u^\Delta)(X^\Delta(t),t). \tag{3.40}$$

Again using itô's lemma we see, for $t_k \leq s < t_{k+1}$, that

$$E[\widetilde{f}(Y^\Delta(s),s)] = \int_{t_k}^s E[L^\Delta \widetilde{f}(Y^\Delta(t),t)]dt. \tag{3.41}$$

Note that $L^\Delta \widetilde{f}$ splits into two parts $f_1 := (\mu_i - \mu_i^\Delta)v$ and $f_2 := -v\partial_\alpha(u-u^\Delta)$, with v being a smooth, polynomially bounded, function. Terms of type f_1 equals zero for $t = t_k$ and using Itô's lemma and the fact that each component of Y^Δ belongs to $\mathbb{D}^\infty$ on $[t_k, t_{k+1})$,

$$E[f_1](t) = \int_{t_n}^t E[L^\Delta f_1](\tau)d\tau = \mathcal{O}(\Delta t_k). \tag{3.42}$$

Terms of type f_2 can be treated by using (3.22) and we get

$$E[f_2](t) = \int_t^T \mathcal{O}(\Delta t_k(s))ds. \tag{3.43}$$

Put together these estimates complete the proof of the second statement in the lemma. The third statement in the lemma follows similarly. We omit the details. □

Finally, we note that to prove the following two lemmas one can proceed as in the proof of Lemma 2.5 in [27].

Lemma 3.6 *Assume that the assumptions in Theorem 3.2 are fulfilled. Let $\mathcal{F}_t$ denote the σ-algebra generated by $\{W(s) : 0 \le s \le t\}$ and let the dual functions ϕ and ϕ' satisfy (3.3) and (3.5). Then*

$$\partial_i u^\Delta(X^\Delta(t_k), t_k) = E[\phi_i(X^\Delta(t_k), t_k)|\mathcal{F}_{t_k}], \;\; \partial_{ij} u^\Delta(X^\Delta(t_k), t_k) = E[\phi'_{ij}(X^\Delta(t_k), t_k)|\mathcal{F}_{t_k}]. \tag{3.44}$$

Lemma 3.7 *Assume that the assumptions in Theorem 3.2 are fulfilled. Let $\mathcal{F}_t$ denote the σ-algebra generated by $\{W(s) : 0 \le s \le t\}$ and let the dual functions ϕ and ϕ' satisfy (3.3) and (3.5). Then*

$$\begin{aligned} &E\left[\left(\mu_i(X^\Delta(t_{k+1}), t_{k+1}) - \mu_i(X^\Delta(t_k), t_k)\right) E[\phi_i(t_{k+1})|\mathcal{F}_{t_{k+1}}]\right] = \\ &\qquad E\left[\left(\mu_i(X^\Delta(t_{k+1}), t_{k+1}) - \mu_i(X^\Delta(t_k), t_k)\right) \phi_i(t_{k+1})\right] \qquad (3.45)\\ &E\left[\left(a_{ij}(X^\Delta(t_{k+1}), t_{k+1}) - a_{ij}(X^\Delta(t_k), t_k)\right) E[\phi'_{ij}(t_{k+1})|\mathcal{F}_{t_{k+1}}]\right] = \\ &\qquad E\left[\left(a_{ij}(X^\Delta(t_{k+1}), t_{k+1}) - a_{ij}(X^\Delta(t_k), t_k)\right) \phi'_{ij}(t_{k+1})\right]. \qquad (3.46) \end{aligned}$$

Combining the lemmas above we see that

$$\begin{aligned} E_d^\Delta(x) &= \sum_{i=1}^{n}\sum_{k=0}^{N-1} E\left[\left(\mu_i(X^\Delta(t_{k+1}), t_{k+1}) - \mu_i(X^\Delta(t_k), t_k)\right) \phi_i(t_{k+1})\right] \frac{\Delta t_k}{2} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=0}^{N-1} E\left[\left(a_{ij}(X^\Delta(t_{k+1}), t_{k+1}) - a_{ij}(X^\Delta(t_k), t_k)\right) \phi'_{ij}(t_{k+1})\right] \frac{\Delta t_k}{2} \\ &+ \sum_{k=0}^{N-1} (\Delta t_k)^2 \left\{ \mathcal{O}(\Delta t_k) + \sum_{r=k}^{N-1} \mathcal{O}((\Delta t_r)^2) \right\}. \end{aligned} \tag{3.47}$$

Theorem 3.2 now follows readily from (3.47).

3.2 Controlling the discretization error

Using the notation in (3.47) we let, for $k \in \{0, ..., N-1\}$ and $l \in \{1, ..., M\}$,

$$\begin{aligned} \rho_k(\omega_l) &= \sum_{i=1}^{n} \left[\mu_i(X^\Delta(t_{k+1}, \omega_l), t_{k+1}) - \mu_i(X^\Delta(t_k, \omega_l), t_k)\right] \phi_i(t_{k+1}, \omega_l) \frac{1}{2\Delta t_k} \\ &+ \sum_{i=1}^{m}\sum_{j=1}^{m} \left[a_{ij}(X^\Delta(t_{k+1}, \omega_l), t_{k+1}) - a_{ij}(X^\Delta(t_k, \omega_l), t_k)\right] \phi'_{ij}(t_{k+1}, \omega_l) \frac{1}{2\Delta t_k}. \end{aligned} \tag{3.48}$$

Furthermore, we introduce

$$E_d^{\Delta,M} = \frac{1}{M}\sum_{l=1}^{M}\sum_{k=0}^{N-1}\rho_k(\omega_l)(\Delta t_k)^2 \tag{3.49}$$

and

$$E_{ds}^{\Delta,M} = E\left[\sum_{k=0}^{N-1}\rho_k(\Delta t_k)^2\right] - \frac{1}{M}\sum_{l=1}^{M}\sum_{k=0}^{N-1}\rho_k(\omega_l)(\Delta t_k)^2. \tag{3.50}$$

Then

$$\begin{aligned} E_d^{\Delta} &= E_d^{\Delta,M} + \bar{E}_d^{\Delta,M}, \\ \bar{E}_d^{\Delta,M} &= \sum_{k=0}^{N-1}(\Delta t_k)^2\left\{\mathcal{O}(\Delta t_k)+\sum_{r=k}^{N-1}\mathcal{O}((\Delta t_r)^2)\right\} + E_{ds}^{\Delta,M}. \end{aligned} \tag{3.51}$$

We note that $E_{ds}^{\Delta,M}$ can be handled using the techniques briefly described in Section 4 below and therefore we here discuss, following [27], how to control the error $E_d^{\Delta,M}$ in (3.49) and in particular how to use iterative refinements of the mesh $\Delta t = \{t_0, t_1, ..., t_N\}$ in order ensure that $E_d^{\Delta,M}$ is below a pre-specified error tolerance denoted by TOL_d. In particular, let at step j in the refinement procedure, a time discretization $\Delta t[j] = \{t_0, t_1, ..., t_{N[j]}\}$ of $[0, T]$ be given and assume that we have generated $M[j]$ trajectories from the underlying model. Let $\rho[j](\omega_l) = \sum_{k=0}^{N[j]-1}\rho_k(\omega_l)$, for $j \in \{0, 1, ...\}$, $l \in \{1, ..., M[j]\}$, and let

$$\bar{\rho}[j] = \frac{1}{M[j]}\sum_{l=1}^{M[j]}\rho[j](\omega_l). \tag{3.52}$$

Moreover, let for $\tau \in [t_k, t_{k+1})$, $t_k, t_{k+1} \in \Delta t[j]$, $\bar{\rho}[j](\tau) = (1/M[j])\sum_{l=1}^{M[j]}\rho_k[j](\omega_l)$ and likewise, $\Delta t[j](\tau) = t_{k+1} - t_k$. Then, for a given tolerance TOL_d the idea is to solve the following minimization problem

$$\min_{\Delta t \in \mathcal{K}[j]} \mathcal{N}(\Delta t) \tag{3.53}$$

where

$$\begin{aligned} \mathcal{K}[j] \; = \; &\{\Delta t \in L^2[0,T] : \Delta t \text{ is positive and piecewise} \\ &\text{constant on } \Delta t[j] \text{ and } \int_0^T |\bar{\rho}[j](\tau)|\Delta t(\tau)d\tau \leq TOL_d\}, \end{aligned} \tag{3.54}$$

and where

$$\mathcal{N}(\Delta t) := \int_0^T \frac{1}{\Delta t(\tau)}d\tau = \sum_{k=0}^{N[j]-1}\frac{\Delta t_k[j]}{\Delta t(t_k)} \tag{3.55}$$

is the number of steps of the partition Δt. In particular, the idea is to minimize the size of $\Delta t[j]$ in order to have as few time-steps as possible, i.e., to minimize $N[j]$ while the error

$E_d^{\Delta,M}(\Delta t[j])$, defined in (3.49), is below a given threshold, TOL_d, as in the definition (3.54) of $\mathcal{K}[j]$. Furthermore, using Lagrange multipliers one can prove that the minimizer equals

$$\Delta t^* = \frac{TOL_d}{\sqrt{|\bar{\rho}[j]|}\left(\int_0^T \sqrt{|\bar{\rho}[j](s)|}ds\right)}. \tag{3.56}$$

4 The statistical error

Let Y be a random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ and let $\{Y(\omega_l)\}_{l=1}^M$, $\omega_j \in \Omega$, denote M independent samples of Y. We let

$$\mathcal{M}(M;Y) = \frac{1}{M}\sum_{l=1}^{M} Y(\omega_l),\ \mathcal{S}(M;Y) = \left(\mathcal{M}(M;Y^2) - (\mathcal{M}(M;Y))^2\right)^{1/2}. \tag{4.1}$$

Then $\mathcal{M}(M;Y)$ and $\mathcal{S}(M;Y)$ denote the sample average and sample standard deviation respectively. Moreover, let $\sigma = (E(|Y - E(Y)|^2))^{1/2}$ and assume that $\kappa = \frac{1}{\sigma}(E(|Y - E(Y)|^3))^{1/3} < \infty$. Let

$$Z_M = \frac{\mathcal{M}(M;Y) - E(Y)}{\sigma/\sqrt{M}}. \tag{4.2}$$

Let $\Phi(z)$, $z \in \mathbb{R}$, be the cumulative distribution function for a standard normal random variable with mean 0 and standard deviation equal to 1 and let $F_{Z_M}(z) = P(Z_M \leq z)$. Then using the Berry-Esséen theorem, see for example Theorem 2.4.10 in [4],

$$\sup_{z\in\mathbf{R}} |F_{Z_M}(z) - \Phi(z)| \leq \frac{3\kappa^3}{\sqrt{M}}. \tag{4.3}$$

In particular, if we introduce the error $\mathcal{E}_S(M;Y) := E(Y) - \mathcal{M}(M;Y)$, then

$$P\left(|\mathcal{E}_S(M;Y)| \leq c_0\frac{\sigma}{\sqrt{M}}\right) \geq 2\Phi(c_0) - 1 - 2\sup_{z\in\mathbb{R}}|F_{Z_M}(x) - \Phi(z)| \tag{4.4}$$

Let $M = \beta^2\kappa^6$ where $\beta \gg 1$ and let α be defined through the relation $\Phi(c_0) = \alpha$. Combining the estimates in the last two displays we see that

$$P\left(|\mathcal{E}_S(M;Y)| \leq c_0\frac{\sigma}{\sqrt{M}}\right) \geq 2\alpha - 1 - 6\beta^{-1}. \tag{4.5}$$

In particular, if we let $\beta^2 \gg 14400$ and $c_0 \geq 1.96$ then $P\left(|\mathcal{E}_S(M;Y)| \leq c_0\frac{\sigma}{\sqrt{M}}\right) \geq 0.90$. Moreover, we can ensure, with high probability, that

$$|\mathcal{E}_S(M;Y)| \leq E_S(M;Y) := c_0\frac{\mathcal{S}(M;Y)}{\sqrt{M}}, \tag{4.6}$$

where we have used $\mathcal{S}(M;Y)$ as an approximation of σ. More details on when this can be done (i.e. the size of M) can be found in [6] chapter 2 or [24] Section 14.1 and the references therein.

5 An application: pricing European options in the Hobson-Rogers model for stochastic volatility

In this section we apply the approach outlined in the previous sections to the problem of pricing European options in the setting of the stochastic volatility model proposed by Hobson and Rogers in [11]. In this model the underlying partial differential equation is a degenerate parabolic equation of Kolmogorov type. For comparison we note that the numerical aspects of the pricing of European options in this model have recently also been investigated by Di Francesco, Foschi and Pascucci in [7] and therein the authors develop, in particular, a finite-difference scheme for the underlying operator tailored to the specified contract. What makes this task more complicated, compared to the case of uniformly elliptic operators, i.e., the case $m = n$, is that in the case of degenerate parabolic operators of Kolmogorov type, the grids used can no longer be chosen with respect to the Euclidean geometry. Instead one has to investigate the geometry induced by the underlying Lie-group structure. As stated in the introduction, by going down the stochastic route one can circumvent all explicit problems related to the presence of a more involved Lie-group structure.

To proceed we next briefly outline the Hobson-Rogers model but for full details we refer the reader to [11]. Let $T \in \mathbb{R}_+$, the time to maturity, be fixed and let $W(t)$, $t \in [0,T]$, be a standard one-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$. Let $S(t)$ denote the price of an asset at time $t \in [0,T]$ and let r denote the risk-free short rate. For $t \in [0,T]$ we introduce the discounted log-price $Z(t) = \log(e^{-rt}S(t))$ and, for fixed $\lambda > 0$,

$$D(t) = Z(t) - \int_0^\infty \lambda e^{-\lambda t} Z(t-\tau)d\tau. \tag{5.1}$$

$D(t)$ can be thought of as a measure of the deviation of the discounted log-price $Z(t)$ from a trend and λ is the rate at which information from the past is discounted. Moreover we assume that

$$dZ(t) = \mu(D(t))dt + \sigma(D(t))dW(t) \tag{5.2}$$

where μ and σ are deterministic functions and σ is positive. Using a Girsanov transformation one can prove, see Section 4.1 in [11], that there exists a new probability space on which

$$\begin{aligned} dD(t) &= -\left(\frac{1}{2}\sigma^2(D(t)) + \lambda D(t)\right)dt + \sigma(D(t))dW(t), \\ dZ(t) &= -\frac{1}{2}\sigma^2(D(t))dt + \sigma(D(t))dW(t). \end{aligned} \tag{5.3}$$

In (5.3), $W(t)$, $t \in [0,T]$, now denotes a Brownian motion on the new probability space. We next introduce $U(t) = Z(t) - D(t)$, $t \in [0,T]$, and we note that

$$dU(t) = \lambda(Z(t) - U(t))dt. \tag{5.4}$$

For simplicity we in the following assume that $r = 0$. Using this notation, and after these transformations, we see that the pay-off for a European option, originally written on S, become $g(z, \upsilon)$ in the variables z, υ. Hence, the price of this option, at time $t \in [0,T]$, is, as $r = 0$, given by $u(Z(t), U(t), t) = E[g(Z(T), U(T))|\mathcal{F}_t]$. Furthermore, using the Feynman-Kac's formula we can conclude, if we assume appropriate regularity on σ and g, that u is the solution to

$$\begin{aligned}\frac{1}{2}\sigma^2(z-u)\,(f_{zz}-f_z)+\lambda(z-u)f_u+f_t &= 0\\ f(z,u,T) &= g(z,u). \end{aligned}\tag{5.5}$$

Recall that a degenerate parabolic operator of Kolmogorov type is an operator of the form

$$L = \frac{1}{2}\sum_{i,j=1}^{m} a_{i,j}(x,t)\partial_{ij} + \sum_{i=1}^{m} b_i(x,t)\partial_i + \sum_{i,j=1}^{n} c_{i,j}x_i\partial_j + \partial_t, \tag{5.6}$$

and if we let $\partial_1 = \frac{\partial}{\partial z}$, $\partial_2 = \frac{\partial}{\partial u}$ then we see that (5.5) is a degenerate parabolic operator of Kolmogorov type with

$$\begin{aligned} m &= 1,\ n=2,\ A_{1,1} = -\sigma^2(z-u),\ b_1 = \frac{1}{2}\sigma^2(z-u),\\ C &= \{c_{i,j}\} = \begin{pmatrix} 0 & 0\\ -\lambda & \lambda\end{pmatrix}. \end{aligned}\tag{5.7}$$

Furthermore, for (z, u) fixed we see that the operator in (5.5) satisfies the condition in (1.9) which ensure that the operator is hypoelliptic. In particular, if σ and g are such that Assumption 3.1 is satisfied, then we can apply the methodology outlined in the previous sections.

To completely specify our numerical application we in the following focus on a particular problem assuming that the Hobson-Rogers model is valid. In particular, as in [11] and [7] we assume, for some large positive constant M_σ, that

$$\sigma(x) = \eta\sqrt{1+\varepsilon x^2} \wedge M_\sigma, \tag{5.8}$$

where $\eta > 0$, $\epsilon > 0$. The cutoff is necessary to avoid Z and U from exploding. By using a smooth approximation of σ in a neighborhood of the cutoff we are still able to use the algorithm. We set $g(z, u) = (K - e^z)^+$ which corresponds to a European put option with strike K. Likewise we use a smooth approximation of g in a neighborhood of the cusp[1]. In both our examples we have set

$$\lambda = 1,\ \varepsilon = 5,\ S(0) = K(0) = 1,\ D(0) = 0.1,\ NCH = 3, \tag{5.9}$$

$$MCH = 10,\ c_0 = 1.96 \text{ and } M_\sigma = 10.000. \tag{5.10}$$

Thus $Z(0) = 0$, and $U(0) = -0.1$. We will consider two cases, in the first case $\eta = 0.2$ and $T = 0.25$ and in the second case $\eta = 0.7$ and $T = 0.75$. Our goal is to approximate $E[g(Z(T))] := E[(1-e^{Z(T)})^+]$ with a prescribed accuracy TOL and the results are presented in table 1 and table 2. Recall the dual functions ϕ and ϕ' defined in (3.3)-(3.5) which in turn are used to compute the error functions ρ as in (3.48). To approximate $E[g(Z(T))]$ we perform a Monte Carlo simulation, using antithetic variates, see [10]. In particular, assume that we have

[1]This will however contribute to an error when pricing options. Luckily the error introduced when smoothening g is controlled. I.e. if $\widehat{g}$ is a smooth version of g $E[g(Z(T))] - E[\widehat{g}(Z(T))] \le P(\left|K - e^{Z(T)}\right| < \varepsilon) \max|g(z) - \widehat{g}(z)|$ and by an appropriate choice of ε and $\widehat{g}$ this can be arbitrarily small.

a discretization $\Delta = \{t_k\}_{k=1}^N$ of $[0,T]$ and that we have computed samples $\left\{Z^\Delta(t_k,\omega_l)\right\}_{l=1}^M$ from $Z^\Delta(t_k)$, for $k = 1,...,N$, using the discrete Euler approximation, see (1.19). Then

$$E[(1-e^{Z(T)})^+] \approx \frac{1}{M}\sum_{l=1}^{M}(1-e^{Z^\Delta(T,\omega_l)})^+. \tag{5.11}$$

As outlined in the previous sections, this approximation induces the three errors,

$$\begin{aligned} E_d^{\Delta,M} &= \frac{1}{M}\sum_{l=1}^{M}\sum_{k=0}^{N-1}\rho_k(\omega_l)(\Delta t_k)^2, \\ E_{ds}^{\Delta,M} &= E\left[\sum_{k=0}^{N-1}\rho_k(\Delta t_k)^2\right] - \frac{1}{M}\sum_{l=1}^{M}\sum_{k=0}^{N-1}\rho_k(\omega_l)(\Delta t_k)^2, \\ E_s^{\Delta,M} &= E[g(Z^\Delta(T))] - \frac{1}{M}\sum_{l=1}^{M} g(Z^\Delta(T,\omega_l)). \end{aligned} \tag{5.12}$$

When approximating the discretization error E_d^Δ in Theorem 3.2 with $E_d^{\Delta,M}$ a statistical error $E_{ds}^{\Delta,M}$ is introduced which gives us the a posteriori controllable discretization error $E_d^{\Delta,M}+E_{ds}^{\Delta,M}$, see Section 3.2 and Section 4. Finally we have a statistical error $E_s^{\Delta,M}$ due to the use of finite samples of $g(Z^\Delta(T,\omega_l))$. We wish to get below a pre-specified accuracy TOL and we divide our tolerance into three pieces, one for each error above. That is, we set $TOL = TOL_{ds} + TOL_d + TOL_s$. To be more precise we set $TOL_{ds} = TOL/9$, $TOL_d = 2TOL/9$ and $TOL_s = 2TOL/3$.

We are now ready to set up an algorithm for this problem. In the following we have chosen to formulate the algorithm using a number of sub-algorithms. The exact details of these sub-algorithms can be found in the appendix. The results from applying the algorithm are shown in table 1 and table 2. Error % is the tolerance level TOL which we have as input in algorithm **Adaptive,** given in percent of the true value of the price. The reference value, or 'true value', of the price was obtained by repeatedly using Monte Carlo simulations with antithetic variates simulating $2 \cdot 10^9$ realizations. The discretization of the time interval was unifom with $N = 50$. It seemed to produce an estimate of the price, good enough, for us to talk about a relative error of 0.1%. $M(0)$ and $N(0)$ are the starting values for M- the number of realizations $\{\omega_l\}_{l=1}^M$, and N- the number of points in the discretization $\Delta = \{t_k\}_{k=1}^N$. We emphasize the importance of not choosing $M(0)$ and $N(0)$ at random. As pointed out in Remark A.2 we choose $M(0) \geq 14400\kappa^6$ where κ is the third central moment. To do this we first simply approximate κ using a Monte Carlo simulation. Since we do not control the error we might, for example, double this value, or in some other way make sure we do not underestimate κ, when choosing $M(0)$. Regarding the choice of N, Theorem 3.2 forces us to choose $N \geq T/\sqrt{TOL}$. The constant c_0 in Remark A.2 equals 1.96 in both examples which, together with our choice of $M(0)$ and $N(0)$, assures that the algorithm works with probability $\gg 0.9$. Continuing, $M(E_{ds}^{\Delta,M})$ is the number of replications needed to assure that $E_{ds}^{\Delta,M} < TOL_{ds}$. This is achieved in algorithm **Adaptive** by calling algorithm **ChangeM**. $N(E_d^{\Delta,M})$ is the number of time steps in the discretization of $[0,T]$ needed to assure that $E_d^{\Delta,M} < TOL_d$. This is controlled in algorithm **Adaptive** by calling algorithm **Refine.** $M(E_s^{\Delta,M})$ is the number of replications needed to assure that $E_s^{\Delta,M} < TOL_s$ which is controlled in algorithm **Adaptive** by calling algorithm **StatisticalError**. The errors

Error %	$M(0)$	$N(0)$	$M(E_{ds})$	$N(E_d)$	$M(E_s)$	Price	True Error %
5	4000	12	4000	12	4000	0.0413	1.90
3	4000	16	4000	16	4000	0.0410	1.04
1	4000	27	4000	27	4000	0.0409	0.69
0.1	4000	84	4000	84	54.942	0.0406	0.02

Table 1: Price of a European put option with $\eta = 0.2,\ T = 0.25$. Reference value: 0.0406.

Error %	$M(0)$	$N(0)$	$M(E_{ds})$	$N(E_d)$	$M(E_s)$	Price	True error %
5	4.000	14	4.000	14	4.000	0.2674	0.67
1	4.000	18	4.000	18	4.000	0.2705	1.82
0.5	4.000	31	4.000	35	9.114	0.2679	0.86
0.1	4.000	98	4.000	477	929.174	0.2655	0.02

Table 2: Price of a European put option with $\eta = 0.7,\ T = 0.75$. Reference value: 0.2656.

are controlled in the order indicated above. The rightmost column finally shows the true error in percent of the true value. Table 3 then visualizes how the algorithm iterates to bound the errors for $\eta = 0.7$, $T = 0.75$ and for a relative error tolerance TOL of 0.1%. Furthermore, the refinement of the mesh is visualized in Figure 1, also in this case for $\eta = 0.7$, $T = 0.75$ and for a relative error tolerance TOL of 0.1%. The horizontal axis indicates the time interval $[0, T]$. Each bar should be associated to one time step in the original uniform discretization $\{t_k\}_{k=1}^{N(0)}$ of $[0, T]$. The vertical axis shows the emerging number of time steps in the final mesh in Algorithm **Adaptive** for each of the original time steps. Hence, if we use uniform time steps, instead of using algorithm **Refine** in algorithm **Adaptive,** all bars would be of equal height.

Iter.	M	N	E_{ds}	E_d	E_s
1	4.000	98	$7.5548 \cdot 10^{-8}$	$2.0574 \cdot 10^{-4}$	$2.4978 \cdot 10^{-3}$
2	4.000	263	$1.5580 \cdot 10^{-8}$	$6.6230 \cdot 10^{-5}$	$2.6530 \cdot 10^{-3}$
3	4.000	477	$4.2402 \cdot 10^{-9}$	$5.5577 \cdot 10^{-5}$	$2.5853 \cdot 10^{-3}$
4	40.000	477	–	–	$2.5808 \cdot 10^{-3}$
5	400.000	477	–	–	$8.0608 \cdot 10^{-4}$
6	929.174	477	–	–	$1.6755 \cdot 10^{-4}$

Table 3: The iteration procedure for $\eta = 0.7,\ T = 0.75$ when the input tolerance is 0.1%.

Finally, we briefly emphasize that the algorithm we have presented is applicable in situations far more general than the one considered in the example above and this is different compared to [7] were the finite-difference schemes have to be sort of tailored to the operator at hand. The drawback of the algorithm outlined, in the case of option pricing when the pay-off usually is only Lipschitz, is that in Theorem 3.2 we assume $\mu_i,\ \sigma_{ij} \in C_b^\infty(\mathbb{R}^n \times \mathbb{R}_+)$ for all $i \in \{1, .., n\},\ j \in \{1, .., m\}$ and that $g \in C_p^\infty(\mathbb{R}^n)$. As previously pointed out, we can in some cases circumvent this problem. The real problem is when we have a jump discontinuity as we then have to smooth out the pay-off in an εneighborhood at the expense of creating large partial derivatives. Whether or not these derivatives can be suppressed, using ε, depends on the problem at hand.

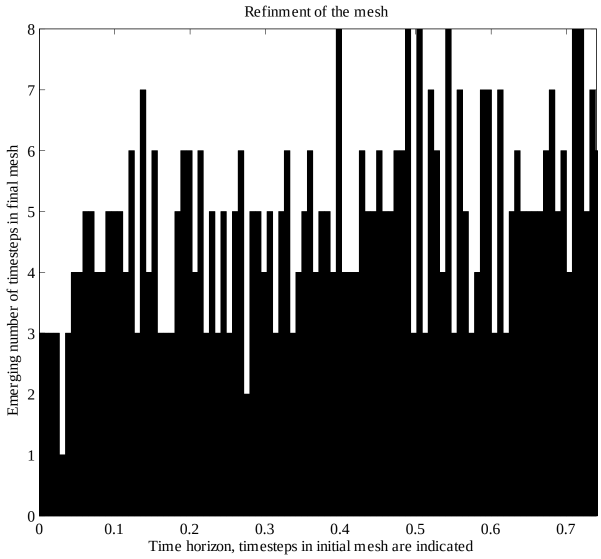

Figure 1: Each box corresponds to one interval $\Delta t_k = [t_k, t_{k+1})$ in the starting mesh. The height of the box indicates the number of time steps contained within Δt_k after the refinement procedure. Here $\eta = 0.7$ and $T = 0.75$. The relative error is 0.1%.

6 Concluding remarks

In this paper we have shown how to use a posteriori error estimates and adaptive weak approximations of stochastic differential equations to numerically solve, with control of the time-discretization error and the statistical errors, the backward in time Cauchy problem for a general class of second order degenerate parabolic partial differential operators of Kolmogorov type. Moreover, we have demonstrated the effectiveness of the methodology outlined when pricing European derivatives in the framework of the stochastic volatility model suggested by Hobson and Rogers, see [11]. In particular, we have compared the efficiency of our method to finite-difference methods recently developed in [7]. This comparison highlights an important advantage of the method outlined in this paper, compared to several other techniques, and that is that one can ensure, with high probability, that the method presented here produces a result, given a user defined error tolerance, which is within the error tolerance of the correct value. Furthermore, the algorithms considered are generally applicable to the large class of second order degenerate parabolic partial differential operators of Kolmogorov type.

The interested reader might have noted that in [27] two algorithms are presented, one with deterministic time steps and one with stochastic time steps. The adaptive algorithm with stochastic time steps can also be proved for degenerate parabolic equations of Kolmogorov type. We have chosen not to include it in our paper, since the expected number of time steps is larger in the stochastic time step algorithm at the same time as the error functions ρ are more complicated than in the deterministic algorithm. However, if the coefficients $\mu_i(x,t)$ and $\sigma_{ij}(x,t)$ has a singularity at x_0 the stochastic time step algorithm is likely to perform better. The benefit of using the stochastic time step algorithm comes from the fact that if $X^{\Delta}(t_k, \omega_l)$ is near x_0 one only has to refine the mesh at t_k for the realization ω_l instead of refining the mesh at t_k for all realizations.

A Algorithms

For the convenience of the reader we include, in detail, the algorithms used in the numerical example. These algorithms were introduced in [27] and while we have made some changes concerning how to choose $M[0]$ and κ, see Remark A.2, they are essentially the same. Finally, we note that to speed up computations various variance reduction techniques might be used. In our example we used antithetic variates. More information on this are available in, for example, [6, 10] and the references therein.

A.1 Auxiliary algorithm

Our first task is to use the Euler scheme defined in (1.19) to approximate $X(t)$. This is made in algorithm **Euler.**

Algorithm Euler

Purpose: Compute M realizations of the discrete Euler approximation $X^{\Delta}(t)$ of $X(t) \in \mathbb{R}^n$.

Input: M- number of realizations
m- dimension of the Wiener process driving X
$x = X(0)$- intitial value
$t = [t_0, ..., t_N]$- a discretization of $[0, T]$

Output: $X_i^{\Delta}(t_k, \omega_l)$, $i = 1, .., n$, $k = 1, ..., N$ and $l = 1, ..., M$

Method: $\Delta t_i = t_{i+1} - t_i$,
Compute an $N \times M \times m$ matrix ΔW whose components $\Delta W_{ik} \in \mathbb{R}^m$ are independent samples from $N^m(0, I_m \Delta t_k)$.
Use the Euler scheme (1.19) to compute $X_i^{\Delta}(t_k, \omega_l)$ for $i = 1, .., n$, $k = 1, ..., N$ and $l = 1, ..., M$.

A.2 Reliability of estimates - statistical error control

We propose two algorithms, algorithm **StatisticalError** and algorithm **ChangeM**, which combined will control the statistical error in accordance with Section 4.

Algorithm StatisticalError

Purpose: Determine the number of realizations necessary to assure, with probability $0 < p < 1$, that the statistical error is below a prescribed tolerance when approximating a random variable Y.

Input: $M[0]$- initial number of realizations, see Remark A.2.
m- dimension of the brownian motion driving Y
$y = Y(0)$- initial value
TOL- error tolerance
c_0- probability constant, see equation (4.4)
$t = [t_0, ..., t_N]$- a discretization of $[0, T]$.

Output: $E[Y]$- expectation of Y
M- number of realizations needed

Method: Set $j = 0$, $E_S[0] = 2TOL$
while $E_S[j] > TOL$
Compute $M[j]$ new samples of Y using algorithm **Euler**.
Compute the sample average $E[Y] = \mathcal{A}(M[j], Y)$ and variance $\mathcal{S}[j] = \mathcal{S}(M[j], Y)$, as in (4.1), and the deviation $E_S[j+1] := E_S(M[j], Y)$ as in (4.6)
Compute $M[j+1]$ by calling algorithm **ChangeM**$(M[j], \mathcal{S}[k], TOL)$
$j = j + 1$
end while
Return $E[Y]$, $M[j-1]$

Algorithm ChangeM

Purpose: Determine the number of realizations in the next loop in algorithm **StatisticalError**.

Input: $M[j]$- number of realizations, see Remark A.2.
$\mathcal{S}[j]$- variance
TOL- error tolerance
c_0- probability constant, see equation (4.4)
MCH- maximal ratio between $M[k+1]$ and $M[k]$, see remark A.1.

Output: $M[j+1]$- number of realizations

Method: $M[j+1] = \min\left\{\text{integer part}\left(\frac{c_0\mathcal{S}[j]}{0.95TOL}\right)^2, MCH \times M[j]\right\}$

Remark A.1 *The factor 0.95 above is used to avoid heavy oscillations. The factor MCH is used to avoid a large number of realizations due to one, possibly inaccurate, extreme event. The algorithm is due to equation (4.6) and is a direct consequence of our ambition to keep $E_S(M;Y) < TOL$.*

Remark A.2 *The probability coefficient c_0 in the algorithm **StatisticalError** and in algorithm **ChangeM** and $M[0]$ in algorithm **StatisticalError** must be chosen so that we can assure that*

$$P(|E[Y] - \mathcal{A}(M;Y)| \leq TOL) \geq p \tag{A.1}$$

This is done by choosing c_0 and $M[0]$ so that $2\Phi(c_0) - 1 + 2\frac{3\kappa^3}{\sqrt{M[0]}} \geq p$, see (4.3)-(4.6). To *achieve $p = 0.90$ we might choose $c_0 = 1.96$ and $M[0] = 14400\kappa^6$, where κ is the third moment of Y.*

A.3 Mesh refinement

We refine the mesh in accordance with Section 3.2 in algorithm **Refine.** Repeated use then assures that $E_d^{\Delta t[j],M[j]} < TOL_d$.

Algorithm Refine

Purpose: Refine the mesh Δt to assure that $E_d^{\Delta t[j],M[j]}(M[j], \Delta t[j]) < TOL$.

Input: $\Delta t[j] = \{t_0, t_1, ..., t_{N[j]}\}$- initial partition of $[0, T]$
$\bar{\rho}[j]$- error density
TOL- error tolerance
NCH- maximal splitting rate, see remark A.3.

Output: $\Delta t[j+1]$- new partition

Method: **for** $1 \leq n \leq N[j]$
  Compute t^* as in (3.56) and let
  $$m_k = \min\left\{\max\left\{\text{integer part}\left(\frac{\Delta t_k[j]}{\Delta t_k^*}\right), 1\right\}, NCH\right\}$$
  Divide $[t_k, t_{k+1}]$ into m_k uniform subintervals
end for
Let $\Delta t[j+1]$ be the partition merging above.

Remark A.3 *The number NCH describing the maximal splitting rate is set to avoid the time interval $\Delta t_k[j]$ to split into more than NCH subintervals during one iteration of algorithm **Refine** in the same manner as MCH is set to avoid the number of realizations to explode.*

A.4 The adaptive algorithm

In algorithm **Adaptive** we combine the previously stated algorithms and the emerging algorithm assures that the error is less than the pre-specified tolerance TOL.

Algorithm Adaptive

Purpose: Calculate $E[g(X(T))]$ with a given accuracy TOL.
Input: $M[0]$- initial number of realizations, must be chosen in accordance with remark A.2
$\Delta t[0]$- initial coarse mesh of $[0, T]$
TOL- a preassigned tolerance level
c_0- probability constant, see (4.6)
MCH- maximum ratio between $M[j+1]$ and $M[j]$, see algorithm **ChangeM**
NCH- maximum splitting rate, see algorithm **Refine**
Output: $E[g(X(T))]$ with error less than TOL with probability p
Method: Set $TOL_s = 2TOL/3$, $TOL_d = 2TOL/9$ and $TOL_{ds} = TOL/9$
Set $E_d[0] = 2TOL_d$, $E_{ds} = 2TOL_{ds}$
let $j = 0$
while $E_d[j] + E_{ds}[j] > TOL_d + TOL_{ds}$
  Compute $X^\Delta(t_k)$ by calling algorithm **Euler**.
  Use this to compute $\bar\rho[j]$, $E_d[j]$ and $E_{ds}[j]$ according to (3.48) and (3.52), (3.49) respectively (3.50) and (4.6).
  if $E_{ds}[j] > TOL_{ds}$
    Compute $M[j+1]$ by calling **ChangeM**$(M[j], \mathcal{S}[j], TOL_{ds}, MCH)$
    where $\mathcal{S}[j]$ is calculated as in (4.1) with $Y = \rho[j]$
    Let $j = j+1$
  else if $E_d[j] > TOL_d$
    Compute $\Delta t[j+1]$ by calling **Refine**$(\Delta t[j], \bar\rho[j], TOL_d, NCH\ )$
    Let $j = j+1$
  end if
end while
Call **StatisticalError**($M[j]$, TOL_s, $\Delta t[j]$) with $Y = g(X^\Delta(T))$.
Accept $E[g(X^\Delta(T))]$ as an approximation of $E[g(X_T)]$

References

[1] G. Akrivis, M. Crouzeix, V. Thomée, Numerical methods for ultraparabolic equations, Calcolo **31** (3-4) (1994) 179-190.

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

[3] D. Braess, Finite Elements, 2nd Ed., Cambridge University Press, New York, 2001.

[4] R. Durrett, Probability: Theory and Examples, Wadsworth & Brooks/Cole, Pacific Grove, 1991.

[5] K. Eriksson, C. Johnson, Adaptive Finite Element Methods for Parabolic Problems IV: Nonlinear Problems, SIAM J. Numer. Anal. **32** (6) (1995) 1729-1749.

[6] G. S. Fishman, Monte Carlo. Concepts, Algorithms and Applications, Springer-Verlag, New York, 1996

[7] M. Di Francesco, P. Foschi, A. Pascucci, Analysis of an Uncertain Volatility Model, J. Appl. Math. Decis. Sci. **2006** (2006), Article ID 15609, 1-17.

[8] M. Di Francesco, A. Pascucci, On the complete model with stochastic volatility by Hobson and Rogers, Proc. Roy. Soc. London Ser. A **460** (2051) (2004) 3327-3338.

[9] M. Di Francesco, A. Pascucci, On a Class of Degenerate Parabolic Equations of Kolmogorov Type, Appl. Math. Res. eXpress **3** (2005) 77-116.

[10] P. Glasserman, Monte Carlo Methods in Financial Engineering, Springer-Verlag, New York, 2003.

[11] D. G. Hobson, L. C. G. Rogers, Complete models with stochastic volatility, Math. Finance **8** (1) (1998) 27-48.

[12] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. **119** (1967) 147-171.

[13] O. Houston, E. Süli, Stabilized *hp*-finite element approximation partial differential equations with nonnegative characteristic form, Computing **66** (2) (2001) 99-119.

[14] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.

[15] I. Karatzas, S. E. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, New York, 1988.

[16] P. E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Berlin, 1992.

[17] A. Kolmogorov, Zufällige Bewegungen, Ann. of Math. **35** (2) (1934) 116-117.

[18] H. Kunita, Stochastic flows and stochastic differential equations, Cambridge University Press, Cambridge, 1997.

[19] E. Lanconelli, S. Polidoro, On a class of hypoelliptic evolution operators, Rend. Sem. Mat. Univ. Politec. Torino **52** (1994) 29-63.

[20] S. Larsson, V. Thomée, Partial Differential Equations with Numerical Methods, Springer-Verlag, New York, 2005.

[21] T. Lyons, N. Victoir, Cubature on Wiener Space, Proc. R. Soc. Lond. A **460** (2004) 169-198.

[22] D. Nualart, The Malliavin Calculus and Related Topics, Springer-Verlag, New York, 1995.

[23] C. Mogavero, S. Polidoro, A finite difference method for a boundary value problem related to the Kolmogorov equation, Calcolo **32** (3-4) (1995) 193-205.

[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C, Cambridge University Press, Cambridge, 1992.

[25] P. Protter, Stochastic Integration and Differential Equations, Springer-Verlag, New York, 1990.

[26] D. W. Stroock, S. R. S. Varadhan, Multidimension Diffusion Processes, Springer-Verlag, Berlin, 1979

[27] A. Szepessy, R. Tempone, G.E. Zouraris, Adaptive Weak Approximation of Stochastic Differential Equations, Comm. Pure Appl. Math. **54** (10) (2001) 1169-1214.

[28] D. Talay, L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, Stochastic Anal. Appl. **8** (4) (1990) 94-120.