



Contents lists available at ScienceDirect

# Journal of Computational and Applied Mathematics

journal homepage: [www.elsevier.com/locate/cam](http://www.elsevier.com/locate/cam)

## Pricing and hedging of financial derivatives using a posteriori error estimates and adaptive methods for stochastic differential equations

Kaj Nyström, Thomas Önskog\*

Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden

### ARTICLE INFO

#### Article history:

Received 25 August 2009

Received in revised form 10 June 2010

#### MSC:

49Q12

62P05

65C

#### Keywords:

Sensitivity analysis

Parabolic partial differential equations

Stochastic differential equations

Euler scheme

A posteriori error estimate

Adaptive algorithms

Hedging

Financial derivatives

### ABSTRACT

The efficient and accurate calculation of sensitivities of the price of financial derivatives with respect to perturbations of the parameters in the underlying model, the so-called ‘Greeks’, remains a great practical challenge in the derivative industry. This is true regardless of whether methods for partial differential equations or stochastic differential equations (Monte Carlo techniques) are being used. The computation of the ‘Greeks’ is essential to risk management and to the hedging of financial derivatives and typically requires substantially more computing time as compared to simply pricing the derivatives. Any numerical algorithm (Monte Carlo algorithm) for stochastic differential equations produces a time-discretization error and a statistical error in the process of pricing financial derivatives and calculating the associated ‘Greeks’. In this article we show how a posteriori error estimates and adaptive methods for stochastic differential equations can be used to control both these errors in the context of pricing and hedging of financial derivatives. In particular, we derive expansions, with leading order terms which are computable in a posteriori form, of the time-discretization errors for the price and the associated ‘Greeks’. These expansions allow the user to simultaneously first control the time-discretization errors in an adaptive fashion, when calculating the price, sensitivities and hedging parameters with respect to a large number of parameters, and then subsequently to ensure that the total errors are, with prescribed probability, within tolerance.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

It is fair to say that it is still a great practical challenge in the derivative industry to efficiently and accurately calculate the so-called ‘Greeks’, that is sensitivities of the price of financial derivatives with respect to perturbations of the parameters in the underlying model. Focusing on methods based on stochastic differential equations, the calculation of these sensitivities remains a particularly topical area of current research and the prevailing techniques include finite difference approximations, pathwise derivative estimates, the likelihood ratio method and its generalizations using the Malliavin calculus. We refer to [1] for an excellent account of these methods and their advantages and disadvantages. Although [2,1,3] contain most of the relevant references on these topics we here still would like to suggest [4–14] as additional references for the interested reader. We emphasize that while these articles are almost exclusively devoted to financial applications, the techniques developed are also useful in many other contexts. Moreover, we note that a key feature of the techniques in many of these articles is, heuristically, that the computations tend to be organized in a forward looking way where the calculations in the next step depend on the calculations up to the present. However, in [3] an adjoint formulation for the calculation of sensitivities is suggested and it is shown, numerically, that this formulation can be used to accelerate the calculation of

\* Corresponding author. Tel.: +46 90 786 93 62; fax: +46 90 786 52 22.

E-mail addresses: [kaj.nystrom@math.umu.se](mailto:kaj.nystrom@math.umu.se) (K. Nyström), [thomas.onskog@math.umu.se](mailto:thomas.onskog@math.umu.se) (T. Önskog).

the ‘Greeks’. The method outlined in [3] is particularly well suited in applications requiring sensitivities to a large number of parameters and particular examples of such applications include interest rate derivatives requiring sensitivities with respect to all initial forward rates and equity derivatives requiring sensitivities with respect to all points on a volatility surface. Furthermore, as emphasized in [3] the adjoint method has its advantages, compared to competing methods with forward looking features, when calculating the sensitivities of a small number of securities with respect to a large number of parameters. On the contrary, competing methods with forward looking features are advantageous when calculating the sensitivities of many securities with respect to a small number of parameters. The notion of ‘small number of securities’ can here be an entire book, consisting of many individual securities, as long as the sensitivities to be calculated are for the book as a whole and not for the constituent securities.

In this article we further develop the adjoint method suggested in [3] by outlining how a posteriori error estimates and adaptive methods for stochastic differential equations can be used to adaptively first control the time-discretization errors in these calculations and then to ensure that the total error, defined as sum of the time-discretization error and the statistical error, is, with prescribed probability, within tolerance. In particular, we give a theoretically sound base for the adjoint method suggested in [3]. Our results concerning a posteriori error estimates and adaptive methods for stochastic differential equations build and expand on the work by Szepeszy et al. [15] concerning adaptive weak approximations of stochastic differential equations and, to our knowledge, a posteriori error estimates for stochastic differential equations applied to the pricing of financial derivatives and, in particular, applied to the calculation of hedging parameters for financial derivatives, have previously not been discussed in the literature. Hence, we claim to give a novel contribution to the literature concerning the numerical aspects of pricing and hedging of financial derivatives, as well as to the general problem of conducting sensitivity analysis for solutions of second order parabolic partial differential equations using stochastic techniques. Finally, this article is based on the results developed in the thesis of the second author, see [16].

To more thoroughly describe the methodology outlined in this article we first have to introduce some notation. Let  $(t, x) = (t, x_1, \dots, x_n) \in \mathbb{R}_+ \times \mathbb{R}^n$  and let  $M(n, \mathbb{R})$  be the set of all  $n \times n$ -matrices with real valued entries. Given a matrix  $\sigma \in M(n, \mathbb{R})$  its transpose is denoted by  $\sigma^*$ . Let

$$\begin{aligned}\mu(t, x) &= \mu(t, x, \theta_\mu) = \bar{\mu}(t, x) + \theta_\mu \tilde{\mu}(t, x), \\ \sigma(t, x) &= \sigma(t, x, \theta_\sigma) = \bar{\sigma}(t, x) + \theta_\sigma \tilde{\sigma}(t, x),\end{aligned}\tag{1.1}$$

where  $\bar{\mu}, \tilde{\mu} : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\bar{\sigma}, \tilde{\sigma} : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow M(n, \mathbb{R})$ ,  $\theta_\mu \in \mathbb{R}$ ,  $\theta_\sigma \in \mathbb{R}$ , and  $|\theta_\mu| \leq \epsilon$ ,  $|\theta_\sigma| \leq \epsilon$ , for some small  $\epsilon > 0$ .  $\tilde{\mu}$  and  $\tilde{\sigma}$  represent perturbations of  $\bar{\mu}$  and  $\bar{\sigma}$ . In the following we assume that there exists  $\eta > 0$  such that the following ellipticity condition is satisfied,

$$\xi^*(\bar{\sigma}(t, x) + \theta_\sigma \tilde{\sigma}(t, x))(\bar{\sigma}(t, x) + \theta_\sigma \tilde{\sigma}(t, x))^* \xi \geq \eta |\xi|^2,\tag{1.2}$$

whenever  $|\theta_\sigma| \leq \epsilon$ ,  $\xi \in \mathbb{R}^n$  and  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$ . The ellipticity condition in (1.2) is not crucial to the analysis outlined in this article. In fact, the more general assumption of hypoellipticity suffices as discussed at the end of the article. Define  $\theta = (\theta_\mu, \theta_\sigma)$  and let, for  $i \in \{1, \dots, n\}$ ,

$$X_i(t) = X_i(t, \theta) = x_i + \int_0^t \mu_i(s, X(s), \theta) ds + \sum_{j=1}^n \int_0^t \sigma_{ij}(s, X(s), \theta) dW_j(s).\tag{1.3}$$

Let  $X(t) = (X_1(t), \dots, X_n(t))^*$  denote the corresponding vector. Here  $(W(t))_{0 \leq t \leq T}$ ,  $W(t) = (W_1(t), \dots, W_n(t))^*$ , is a standard Brownian motion in  $\mathbb{R}^n$  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  $\mathbb{R}^n$  we mean a process whose components are independent one-dimensional Brownian motions. In the definition of  $X_i(t) = X_i(t, \theta)$  we have indicated the dependence on the parameter vector  $\theta = (\theta_\mu, \theta_\sigma)$ . Assuming appropriate growth and regularity conditions on the coefficients  $\mu_i$  and  $\sigma_{ij}$ , as will be discussed in detail below, the system in (1.3) has a unique strong solution for all parameters  $\theta = (\theta_\mu, \theta_\sigma)$ ,  $|\theta_\mu| \leq \epsilon$ ,  $|\theta_\sigma| \leq \epsilon$ . We recall that there is a well-known close connection between stochastic differential equations and second order parabolic partial differential equations. We therefore introduce the second order parabolic operator

$$L = \frac{1}{2} \sum_{i,j=1}^n [\sigma \sigma^*]_{ij}(t, x) \partial_{ij} + \sum_{i=1}^n \mu_i(t, x) \partial_i,\tag{1.4}$$

and we note that the structural assumption on the operator  $L$ , imposed by (1.2), is that the operator  $\partial_t + L$  is uniformly elliptic–parabolic. Let  $T > 0$  and let the function  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  be given. Define

$$u(t, x) = u(t, x, \theta) = u(t, x, (\theta_\mu, \theta_\sigma)) = E[g(X(T, \theta)) | X(t, \theta) = x].\tag{1.5}$$

Then, under appropriate smoothness and growth conditions on  $\mu_i, \sigma_{ij}$  and  $g$ , the Feynman–Kac formula asserts that  $u$  in (1.5) is the unique solution to the Cauchy problem

$$\begin{cases} \partial_t u(t, x) + Lu(t, x) = 0, & \text{whenever } (t, x) \in (0, T) \times \mathbb{R}^n, \\ u(T, x) = g(x), & \text{whenever } x \in \mathbb{R}^n, \end{cases}\tag{1.6}$$

where  $L$  is defined in (1.4). In this article we focus on numerical algorithms for stochastic differential equations, with control of the errors, using which we can calculate, simultaneously, the quantities

$$u(t, x, (0, 0)), \quad (\partial_{\theta_\mu} u)(t, x, (0, 0)), \quad (\partial_{\theta_\sigma} u)(t, x, (0, 0)). \tag{1.7}$$

In fact, when calculating the quantities in (1.7) we also, as part of our analysis, calculate all derivatives of  $u(t, x, (0, 0))$ , with respect to the spatial variables, up to fourth order. Furthermore, as  $u$  in (1.5) solves the Cauchy problem in (1.6) this article can also be considered to be devoted to the numerical aspects of a sensitivity analysis for the Cauchy problem, for the operator  $\partial_t + L$ , using stochastic differential equations and the stochastic representation formula in (1.5). To proceed, we note, as a general motivation which is not only limited to the hedging of financial derivatives, that in many applications it is important not only to solve for  $u(t, x, (0, 0))$  but also to quantify the effect, on the solution, of misspecifications of  $\mu$  and  $\sigma$ . Moreover, one would often like to do this for many different perturbations, i.e. for many different choices of the pair  $(\tilde{\mu}, \tilde{\sigma})$ , without too much additional computational effort compared to the calculation of the solution itself. Naturally one would also like to control the error, relative to the true theoretical value, produced by the numerical and computational scheme. In particular, assuming that the coefficients are given through measurements, any such measurement should be seen as a sample from a distribution and hence the coefficients are not known with certainty. One way to account for this is to calculate the sensitivity of the solution to perturbations of the parameters in the underlying model and then to derive an approximative distribution of  $u$  based on distributional assumptions on the parameters.

To outline the actual numerical approximation of (1.5), and to formulate the main results, we next describe the Euler scheme associated to the system in (1.3). In particular, given a time horizon of  $T$  we let  $\{t_k\}_{k=0}^N$  define a partition  $\Delta$  of the interval  $[0, T]$ , i.e.  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$ , and we let  $\Delta t_k = t_{k+1} - t_k$  for  $k \in \{0, \dots, N - 1\}$ . Let  $\{\bar{X}(t) : t \in [0, T]\}$  solve (1.3) for parameter values  $(\theta_\mu, \theta_\sigma) = (0, 0)$ . In the following we let  $\{\bar{X}^\Delta(t) : t \in [0, T]\}$  denote the continuous Euler approximation of  $\{\bar{X}(t) : t \in [0, T]\}$  defined as follows.  $\bar{X}^\Delta(t)$  satisfies the initial condition  $\bar{X}_i^\Delta(t_0) = x_i$  and, for  $k \in \{0, \dots, N - 1\}$ , the difference equation

$$\bar{X}_i^\Delta(t_{k+1}) = \bar{X}_i^\Delta(t_k) + \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k))\Delta t_k + \sum_{j=1}^n \bar{\sigma}_{ij}(t_k, \bar{X}^\Delta(t_k))\Delta W_j(t_k), \tag{1.8}$$

where  $\Delta W_j(t_k) = W_j(t_{k+1}) - W_j(t_k)$  represents the Wiener increment during the time step  $[t_k, t_{k+1}]$ . In the following we make use of the function  $\phi(t) = \sup\{t_k : t_k \leq t\}$ , which is defined whenever  $t \in [0, T]$ . Using this notation, we define the continuous Euler approximation  $\{\bar{X}^\Delta(t) : t \in [0, T]\}$  through the relation

$$\bar{X}_i^\Delta(t) = \bar{X}_i^\Delta(\phi(t)) + \int_{\phi(t)}^t \bar{\mu}_i(\phi(s), \bar{X}^\Delta(\phi(s)))ds + \sum_{j=1}^n \int_{\phi(t)}^t \bar{\sigma}_{ij}(\phi(s), \bar{X}^\Delta(\phi(s)))dW_j(s). \tag{1.9}$$

The set  $\{\bar{X}^\Delta(t) : t \in \{t_0, \dots, t_N\}\}$  is often referred to as the associated discrete Euler approximation. We let  $\bar{u}$  denote  $u(t, x, (0, 0))$  and introduce

$$\bar{u}^\Delta(t_k, x) = E[g(\bar{X}^\Delta(T)) | \bar{X}^\Delta(t_k) = x], \tag{1.10}$$

for  $k \in \{0, \dots, N - 1\}$ ,  $x \in \mathbb{R}^n$ , as an approximation of  $\bar{u}$ . Furthermore, we let  $M$  be an integer and we let  $\{\omega_r\}_{r=1}^M$  represent  $M$  realizations of the discrete Euler approximation of  $\{\bar{X}(t) : t \in [0, T]\}$ . Then, focusing on the calculation of  $(\partial_{\theta_\mu} u)(t, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(t, x, (0, 0))$ , we prove, by proceeding similar to [15], that

$$\begin{aligned} (\partial_{\theta_\mu} u)(0, x, (0, 0)) &= \bar{u}_\mu^{\Delta, M}(x) + \bar{E}_{\mu, s}^{\Delta, M} + \bar{E}_{\mu, d}^{\Delta, M} + \bar{E}_{\mu, d, s}^{\Delta, M} + \bar{R}_{\mu, d}^\Delta, \\ (\partial_{\theta_\sigma} u)(0, x, (0, 0)) &= \bar{u}_\sigma^{\Delta, M}(x) + \bar{E}_{\sigma, s}^{\Delta, M} + \bar{E}_{\sigma, d}^{\Delta, M} + \bar{E}_{\sigma, d, s}^{\Delta, M} + \bar{R}_{\sigma, d}^\Delta, \end{aligned} \tag{1.11}$$

where

$$\begin{aligned} \bar{u}_\mu^{\Delta, M}(x) &= \sum_{r=1}^M \sum_{k=0}^{N-1} \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k, \omega_r)) \bar{\psi}_i(t_k, \omega_r) \frac{\Delta t_k}{M}, \\ \bar{u}_\sigma^{\Delta, M}(x) &= \frac{1}{2} \sum_{r=1}^M \sum_{k=0}^{N-1} [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k, \omega_r)) \bar{\psi}_{ij}^{(1)}(t_k, \omega_r) \frac{\Delta t_k}{M}. \end{aligned} \tag{1.12}$$

In particular,  $\bar{u}_\mu^{\Delta, M}(x)$  and  $\bar{u}_\sigma^{\Delta, M}(x)$  are to be used as Monte Carlo estimators of  $(\partial_{\theta_\mu} u)(0, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$  respectively. The functions  $\bar{\psi}_i$  and  $\bar{\psi}_{ij}^{(1)}$ , appearing in (1.12), are first and second order dual functions associated to the underlying system. These functions solve certain backwards in time stochastic differential equations which facilitate their computation, as outlined in the bulk of the article, and make no reference to the perturbations  $\tilde{\mu}$  and  $\tilde{\sigma}$ . Furthermore (1.11) gives an expansion of the errors produced when  $\bar{u}_\mu^{\Delta, M}(x)$  and  $\bar{u}_\sigma^{\Delta, M}(x)$  are used as approximations of  $(\partial_{\theta_\mu} u)(0, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$ , respectively. The expansions in (1.11) makes it possible to control the errors  $|( \partial_{\theta_\mu} u)(0, x, (0, 0)) - \bar{u}_\mu^{\Delta, M}(x)|$  and  $|( \partial_{\theta_\sigma} u)(0, x, (0, 0)) - \bar{u}_\sigma^{\Delta, M}(x)|$ . To describe the characteristics of the remaining terms in (1.11), we first note

that  $\bar{E}_{\mu,s}^{\Delta,M}$ ,  $\bar{E}_{\mu,d,s}^{\Delta,M}$ ,  $\bar{E}_{\sigma,s}^{\Delta,M}$  and  $\bar{E}_{\sigma,d,s}^{\Delta,M}$  all represent statistical errors, resulting from the use of the finite set  $\{\omega_r\}_{r=1}^M$ , while  $\bar{E}_{\mu,d}^{\Delta,M}$ ,  $\bar{R}_{\mu,d}^{\Delta,M}$ ,  $\bar{E}_{\sigma,d}^{\Delta,M}$  and  $\bar{R}_{\sigma,d}^{\Delta,M}$  represent time-discretization errors due to the discrete Euler scheme. In general, the statistical errors can be controlled using the central limit theorem and in this article we focus mainly on the time-discretization error term. In particular,  $\bar{E}_{\mu,d}^{\Delta,M}$ ,  $\bar{R}_{\mu,d}^{\Delta,M}$ ,  $\bar{E}_{\sigma,d}^{\Delta,M}$  and  $\bar{R}_{\sigma,d}^{\Delta,M}$  have the following important features. Let  $\Delta_N^* = \max\{\Delta t_0, \Delta t_1, \dots, \Delta t_{N-1}\}$ . Then,

- (i)  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  are of order  $\mathcal{O}(\Delta_N^*)$ ,
- (ii)  $\bar{R}_{\mu,d}^{\Delta,M}$  and  $\bar{R}_{\sigma,d}^{\Delta,M}$  are of order  $\mathcal{O}((\Delta_N^*)^2)$ ,
- (iii)  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  are computable a posteriori form. (1.13)

Hence  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  are the leading order terms in the expansions of the time-discretization errors and these leading order terms are computable a posteriori form. The exact form of  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  are derived in the bulk of the article, but a few additional and important features of  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  should be mentioned here. In particular,  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  depend on the first, second, third and fourth order dual functions,  $(\bar{\psi}, \bar{\psi}^{(1)}, \bar{\psi}^{(2)}, \bar{\psi}^{(3)})$  as well as on the Euler discretization, here denoted  $(\bar{X}^\Delta, \bar{X}^{(1),\Delta}, \bar{X}^{(2),\Delta}, \bar{X}^{(3),\Delta}, \bar{X}^{(4),\Delta})$ , of  $(\bar{X}, \bar{X}^{(1)}, \bar{X}^{(2)}, \bar{X}^{(3)}, \bar{X}^{(4)})$ , the latter being a high-dimensional vector containing the variation processes of  $\bar{X}(t)$  up to fourth order. Furthermore, the expressions for  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  also contain derivatives of  $\bar{\mu}$ ,  $\bar{\mu}$ ,  $\bar{\sigma}$  and  $\bar{\sigma}$  up to second order. However,

$$\text{neither } (\bar{\psi}, \bar{\psi}^{(1)}, \bar{\psi}^{(2)}, \bar{\psi}^{(3)}) \text{ nor } (\bar{X}^\Delta, \bar{X}^{(1),\Delta}, \bar{X}^{(2),\Delta}, \bar{X}^{(3),\Delta}, \bar{X}^{(4),\Delta})$$

make any reference to the perturbations  $\bar{\mu}$  and  $\bar{\sigma}$  and these perturbations enter in  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  in a linear fashion. We emphasize that while (1.11)–(1.13) are theoretically rigorous results, their practical use may, at first glance, seem unclear. In particular, due to the potentially very costly computations of  $(\bar{\psi}, \bar{\psi}^{(1)}, \bar{\psi}^{(2)}, \bar{\psi}^{(3)})$  and  $(\bar{X}^\Delta, \bar{X}^{(1),\Delta}, \bar{X}^{(2),\Delta}, \bar{X}^{(3),\Delta}, \bar{X}^{(4),\Delta})$ , the result may be of limited practical use in some applications and we emphasize that the approach described in this article will not, from an efficiency perspective, outperform competing approaches if the aim is to calculate sensitivities with respect to only one set of perturbations  $(\bar{\mu}, \bar{\sigma})$ . Instead, the methodology outlined has its potential merits in case one wants to calculate sensitivities with respect to a set of perturbations  $\{(\bar{\mu}_i, \bar{\sigma}_i)\}_{i=1}^K$  where  $K$  is large. Still, if there is a need to rigorously control the errors, and in particular the time-discretization errors, in a calculation of  $(\partial_{\theta_\mu} u)(0, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$  our result gives, in a posteriori form, the fundamentals and details for such an implementation. Furthermore, as the leading order terms in the expansions of the time-discretization errors are in a posteriori form these terms can also be computed simply to get an estimate of the magnitude of the time-discretization errors. The numerical examples in Sections 5 and 6 illustrate the advantages and disadvantages of the methodology outlined. Moreover, we stress that in many application the statistical errors will, in general, be much larger compared to the time-discretization error. Thus, to obtain computational efficiency and accuracy within a limited computational budget, the use of variance reductions techniques is strongly recommended. We refer to [1] for the fundamentals on variance reductions techniques for financial applications.

Returning to the adjoint method proposed in [3], we note that this is simply an efficient way to organize the calculation of the estimators  $\bar{u}_{\mu}^{\Delta,M}(x)$  and  $\bar{u}_{\sigma}^{\Delta,M}(x)$  in (1.12), when calculating the ‘Greeks’, and, as such, this computational scheme comes about naturally from the very definition of the discrete dual functions. We claim that our results take the approach in [3] one step further as we rigorously derive theoretical expansions of the time-discretization errors with leading order terms in a posteriori form. This renders the possibility of actually controlling, using adaptive type algorithms as proposed in [15], that the errors produced by the method in [3] are, with given probability, within a given tolerance.

The rest of the article is organized as follows. Section 2 is of preliminary nature. In Section 3 we present a general error expansion for the solution to problem (1.6) allowing also for a nonzero right-hand side. The results of Section 3 are utilized in Section 4 as we present and derive the details of (1.11)–(1.13). In Sections 5 and 6 we illustrate the use of (1.11)–(1.13) in two applications. The first application, outlined in Section 5, is a simple benchmark example, for which  $u$  and the sensitivities can be explicitly calculated. This example serves as a stylized illustration of the techniques involved and, in particular, we use this example to evaluate the performance of the estimators in (1.12). The second application, outlined in Section 6, concerns the problem of pricing and hedging of interest rate derivatives in LIBOR market models. This choice is based on the fact that in a fixed-income setting one is often interested in understanding the change of the value of a portfolio of derivative instruments with respect to multi-dimensional structures and hence the approach of this article is attractive in this setting. Moreover, concerning numerical evaluations we note that in [3] the adjoint method is numerically illustrated in the setting of LIBOR market models and is found to be very fast for this application. Section 7 contains a brief summary and discussion.

## 2. Preliminaries

In this section we introduce notation, state representation formulas for solutions to second order parabolic partial differential equations and introduce the Euler scheme.

### 2.1. Notation

Throughout the article we write  $\partial_i f$  for  $\frac{\partial f}{\partial x_i}$ ,  $\partial_{ij} f$  for  $\frac{\partial^2 f}{\partial x_i \partial x_j}$  and so on. If  $f = f(t, x)$ ,  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$ , then  $\partial_t, \partial_{ij}$  and so on will refer to differentiation with respect to the space variable  $x$ . For a multi-index  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ ,  $\alpha_i \in \mathbb{Z}^+$ , we define  $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$  and let  $\partial_\alpha$  denote differentiation with respect to the space variables according to the multi-index  $\alpha$ .

**Remark 2.1.** Throughout the article we use the Einstein summation convention for indices representing spatial directions, implying that when an index occurs more than once in the same expression, the expression is implicitly summed over all possible values for that index.

Given an open set  $U \subset \mathbb{R}^n$  we let  $C_b^k(U)$  denote functions  $f : U \rightarrow \mathbb{R}$  which are  $k$  times continuously differentiable with bounded derivatives. Similarly, we let  $C_p^k(\mathbb{R}^n)$  denote functions  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  which are  $k$  times continuously differentiable and satisfies

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

for some constants  $c_\alpha, q_\alpha \in \mathbb{Z}^+$ . Furthermore, we let  $C_0^\infty(\mathbb{R}^n)$  be the set of infinitely differentiable functions with compact support and let  $C_b^k(\mathbb{R}_+ \times \mathbb{R}^n)$ ,  $k \in \mathbb{Z}^+$ , be the space of all functions defined on  $\mathbb{R}_+ \times \mathbb{R}^n$ , which are continuous and bounded and have continuous and bounded partial derivatives, in both space and time, up to order  $k$ . We also let  $C_b^\infty(\mathbb{R}_+ \times \mathbb{R}^n) = \bigcap_{k \geq 1} C_b^k(\mathbb{R}_+ \times \mathbb{R}^n)$ .

### 2.2. Representation formulas

Let  $\mu$  and  $\sigma$  be defined as in (1.1). We assume (1.2) and that

$$\tilde{\mu}_i, \tilde{\mu}_i, \tilde{\sigma}_{ij}, \tilde{\sigma}_{ij} \in C_b^\infty(\mathbb{R}_+ \times \mathbb{R}^n). \tag{2.2}$$

We let  $A = A(t, x) = (a_{ij}(t, x))_{i,j=1}^n$  and  $\bar{A} = \bar{A}(t, x) = (\bar{a}_{ij}(t, x))_{i,j=1}^n$  denote the  $n \times n$ -matrices defined as

$$a_{ij}(t, x) = \frac{1}{2} [\sigma \sigma^*]_{ij}(t, x), \quad \bar{a}_{ij}(t, x) = \frac{1}{2} [\bar{\sigma} \bar{\sigma}^*]_{ij}(t, x), \tag{2.3}$$

whenever  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^n$ . In the following we let  $L^*$  denote the adjoint operator to  $L$ , i.e.

$$L^* = (\partial_{ij} a_{ij}) + 2(\partial_i a_{ij}) \partial_j + a_{ij} \partial_{ij} - ((\partial_i \mu_i) + \mu_i \partial_i). \tag{2.4}$$

Then, for fixed  $(t, x) \in [0, T] \times \mathbb{R}^n$ , the Green function  $\Gamma(s, y) = \Gamma(t, x, s, y) = \Gamma(t, x, s, y, \theta)$  solves the problem

$$(-\partial_s + L^*) \Gamma(s, y) = \delta_{t,x}(s, y), \quad \text{whenever } (s, y) \in [t, T] \times \mathbb{R}^n, \tag{2.5}$$

where  $\delta_{t,x}(\cdot, \cdot)$  is the Dirac delta function with mass at  $(t, x)$ . Moreover, formally

$$u(t, x) = \langle u, \delta_{t,x} \rangle = \int_{\mathbb{R}^n} \int_t^T u(s, y) (-\partial_s + L^*) \Gamma(s, y) ds dy = \int_{\mathbb{R}^n} g(y) \Gamma(t, x, T, y) dy. \tag{2.6}$$

In particular, (2.6) gives a representation formula for  $u(t, x)$  in terms of the function  $\Gamma(s, y) = \Gamma(t, x, s, y)$  which solves the dual problem in (2.5). The following theorem makes this formal calculation rigorous.

**Theorem 2.2.** Assume (1.2) and (2.2). Let  $T > 0$  be given and let  $f \in C_p^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_p^\infty(\mathbb{R}^n)$ . Then there exists a fundamental solution  $\Gamma = \Gamma(s, y) = \Gamma(t, x, s, y)$  to the operator  $\partial_t + L^*$  in the sense of (2.5). Furthermore, a classical solution  $v \in C_p^\infty((0, T) \times \mathbb{R}^n)$  to the Cauchy problem

$$\begin{cases} \partial_t v(t, x) + Lv(t, x) = f(t, x), & \text{whenever } (t, x) \in (0, T) \times \mathbb{R}^n, \\ v(T, x) = g(x), & \text{whenever } x \in \mathbb{R}^n, \end{cases} \tag{2.7}$$

is given by

$$v(t, x) = \int_{\mathbb{R}^n} \Gamma(t, x, T, y) g(y) dy - \int_t^T \int_{\mathbb{R}^n} \Gamma(t, x, s, y) f(s, y) dy ds. \tag{2.8}$$

The formula in (2.8) holds whenever there exists a positive constant  $M$  such that

$$|f(t, x)| \leq M e^{M|x|^2}, \quad |g(x)| \leq M e^{M|x|^2}, \quad (t, x) \in (0, T) \times \mathbb{R}^n. \tag{2.9}$$

Furthermore,  $v$  in (2.8) is the unique solution to the problem in (2.7) in the class of all functions satisfying (2.9).

**Theorem 2.3.** Assume (1.2) and (2.2). Let  $T > 0$  be given and let  $f \in C_b^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_b^\infty(\mathbb{R}^n)$ . Let  $X(t) = (X_1(t), \dots, X_n(t))^*$  be the stochastic process introduced in (1.3) and let  $v$  be given as in Theorem 2.2. Then  $v$  is uniquely determined by

$$v(t, x) = E \left[ g(X(T)) - \int_t^T f(s, X(s)) ds \mid X(t) = x \right]. \tag{2.10}$$

**Proof of Theorems 2.2 and 2.3.** These results are classical. For the theory of partial differential equations we refer to [17] and for the derivation of the stochastic representation formula in (2.10) we refer to Theorem 5.7.6 in [18].  $\square$

2.3. The Euler scheme for an extended system

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$  be a probability space with a filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$  generated by a Wiener process  $W(t) \in \mathbb{R}^n$  and let  $\mathbb{D}^\infty = \bigcap_{k \geq 1} \bigcap_{p \geq 1} \mathbb{D}^{k,p}$ , where  $\mathbb{D}^{k,p}$  are standard stochastic Sobolev spaces as introduced on page 27 in [19]. For the proof of the following theorem we refer to Theorem 2.2.2 in [19].

**Theorem 2.4.** Let  $X(t)$ ,  $t \in [0, T]$ , be a stochastic process satisfying (1.3) and assume that  $\mu_i(t, x, \theta_\mu)$ ,  $\sigma_{ij}(t, x, \theta_\sigma) \in C_b^\infty(\mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R})$ . Then  $X_i(t) \in \mathbb{D}^\infty$  for all  $t \in [0, T]$ .

In the following analysis we make use of variations of the process  $X(t)$  up to fourth order and associated Euler approximations. Theorem 2.4 states that  $X_i(t) \in \mathbb{D}^\infty$ , for all  $t \in [0, T]$ , and, in particular, this implies that the  $p$ -th variation of the process  $X(t) = (X_1(t), \dots, X_n(t))^*$ , with respect to  $x$ , exists for any  $p \in \mathbb{Z}^+$ . We denote, for  $t \in [0, T]$ , the  $p$ -th variation of  $X(t)$  by  $X^{(p)}(t)$  and, for  $0 \leq t \leq s \leq T$ , the first variation process  $X^{(1)}(s) = X^{(1)}(t, s) = (\partial X(s) \mid X(t) = x) / \partial x_i|_{i=1}^n$ , see Section 2.3 in [19], solves the stochastic differential equation

$$\begin{aligned} dX_{ij}^{(1)}(s) &= \partial_\beta \mu_i(s, X(s)) X_{\beta j}^{(1)}(s) ds + \partial_\beta \sigma_{i\gamma}(s, X(s)) X_{\beta j}^{(1)}(s) dW_\gamma(s), \\ X_{ij}^{(1)}(t) &= \delta_{ij}, \end{aligned} \tag{2.11}$$

where  $\delta_{ij}$  is the Kronecker delta. In particular,  $X^{(1)}(\cdot)$  is an  $n \times n$ -matrix and  $X^{(1)}(s) \in (\mathbb{D}^\infty)^{n \times n}$ . Similarly, for  $0 \leq t \leq s \leq T$ , the second variation solves the stochastic differential equation

$$\begin{aligned} dX_{ijl}^{(2)}(s) &= \left[ \partial_\beta \mu_i(s, X(s)) X_{\beta jl}^{(2)}(s) + \partial_{\beta\gamma} \mu_i(s, X(s)) X_{\beta j}^{(1)}(s) X_{\gamma l}^{(1)}(s) \right] ds \\ &\quad + \left[ \partial_\beta \sigma_{i\gamma}(s, X(s)) X_{\beta jl}^{(2)}(s) + \partial_{\beta\varphi} \sigma_{i\gamma}(s, X(s)) X_{\beta j}^{(1)}(s) X_{\varphi l}^{(1)}(s) \right] dW_\gamma(s), \\ X_{ijm}^{(2)}(t) &= 0. \end{aligned} \tag{2.12}$$

$X^{(2)}(\cdot)$  is an  $n \times n \times n$ -matrix and  $X^{(2)}(s) \in (\mathbb{D}^\infty)^{n \times n \times n}$ . We let

$$Z(s) = (X(s), X^{(1)}(s), X^{(2)}(s), X^{(3)}(s), X^{(4)}(s))^*, \quad 0 \leq t \leq s \leq T, \tag{2.13}$$

be a vector containing the process  $X$  and its variations up to fourth order. As above we note that  $X^{(3)}(\cdot)$  is an  $n \times n \times n \times n$ -matrix and that  $X^{(4)}(\cdot)$  is an  $n \times n \times n \times n \times n$ -matrix. Hence the vector  $Z(s)$  contains  $\tilde{n} := n + n^2 + \dots + n^5 = n(n^5 - 1)/(n - 1)$  elements. Moreover, the vector of variations  $Z(s)$  satisfies the following system of stochastic differential equations

$$\begin{aligned} dZ(s) &= \Lambda(s, Z(s)) ds + \Sigma_j(s, Z(s)) dW_j(s), \quad 0 \leq t \leq s \leq T, \\ Z(t) &= (x, I_n, 0, 0, 0)^* := z(x), \end{aligned} \tag{2.14}$$

where  $\Lambda$  and  $\Sigma$  are matrix valued functions. In (2.14)  $I_n$  denotes the  $n \times n$  unit matrix. Note that  $\Lambda(s, z) = (\Lambda_1(s, z), \dots, \Lambda_{\tilde{n}}(s, z))^*$  and that the matrix  $\Sigma$  has dimension  $\tilde{n} \times n$ .

We next introduce the Euler scheme for the system defined in (2.14). Given a time horizon of  $T$ , we let  $\{t_k\}_{k=0}^N$  define a partition  $\Delta$  of the interval  $[0, T]$ , i.e.  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$ , and we define  $\Delta t_k = t_{k+1} - t_k$  for  $k \in \{0, \dots, N - 1\}$ . Recall that the Euler scheme associated to the system in (1.3) was introduced in (1.8)–(1.9). In analogy with (1.8)–(1.9), we define, for  $m \in \{0, \dots, N - 1\}$ , the continuous Euler approximation of  $\{Z(s) : 0 \leq t_m \leq s \leq T\} = \{Z(t_m, s) : 0 \leq t_m \leq s \leq T\}$ , denoted  $\{Z^\Delta(s) : 0 \leq t_m \leq s \leq T\} = \{Z^\Delta(t_m, s) : 0 \leq t_m \leq s \leq T\}$ , through the difference equation

$$\begin{aligned} Z^\Delta(t_{k+1}) &= Z^\Delta(t_k) + \Lambda(t_k, Z^\Delta(t_k)) \Delta t_k + \Sigma_j(t_k, Z^\Delta(t_k)) \Delta W_j(t_k), \\ Z^\Delta(t_m) &= (x, I_n, 0, 0, 0)^* = z(x) \end{aligned} \tag{2.15}$$

for  $k \in \{m, \dots, N - 1\}$ . Furthermore, for general  $0 \leq t_m \leq s \leq T$ , we have

$$Z^\Delta(s) = Z^\Delta(\phi(s)) + \int_{\phi(s)}^s \Lambda(\phi(r), Z^\Delta(\phi(r))) dr + \int_{\phi(s)}^s \Sigma_j(\phi(r), Z^\Delta(\phi(r))) dW_j(r), \tag{2.16}$$

where  $\phi(s) = \sup\{t_k : t_k \leq s\}$ . For  $0 \leq t_m \leq t_k \leq T$ , we write  $X_{ij}^{(1),\Delta}(t_k) = X_{ij}^{(1),\Delta}(t_m, t_k)$  for the first variation of  $X^\Delta$  with initial data  $X_{ij}^{(1),\Delta}(t_m, t_m) = \delta_{ij}$  and  $X_{ijl}^{(2),\Delta}(t_k) = X_{ijl}^{(2),\Delta}(t_m, t_k)$  for the second variation of  $X^\Delta$  with initial data  $X_{ijl}^{(2),\Delta}(t_m, t_m) = 0$ . A similar notation will be used for the higher order variations of  $X^\Delta$ . Note also that  $Z^\Delta(t_k) = (X^\Delta(t_k), X^{(1),\Delta}(t_k), X^{(2),\Delta}(t_k), X^{(3),\Delta}(t_k), X^{(4),\Delta}(t_k))$  for  $k \in \{m, \dots, N\}$ .

### 3. An error expansion for the Cauchy problem

Assume that the assumptions of Theorem 2.3 are satisfied, let  $v$  be as in Theorem 2.3 and let  $\alpha, |\alpha| \leq 4$ , be a multi-index. Recall that the vector of variations  $\{Z(s) : 0 \leq t \leq s \leq T\}$  was introduced in (2.14). Then, using the notation in (2.13), we have

$$\partial_\alpha v(t, x) = E \left[ g_\alpha(Z(T)) - \int_t^T f_\alpha(s, Z(s)) ds | Z(t) = z \right], \tag{3.1}$$

where  $z = z(x)$  and the functions  $g_\alpha$  and  $f_\alpha$  are composed of partial derivatives of  $g$  and  $f$ , respectively, up to order  $|\alpha|$  multiplied by polynomials, of degree at most  $|\alpha|$ , defined using the components of the vector  $Z$  as coordinates. In particular, by an explicit calculation,

$$\partial_i v(t, x) = E \left[ \partial_{\beta i} g(X(T)) X_{\beta i}^{(1)}(T) \right] - E \left[ \int_t^T \partial_{\beta f}(s, X(s)) X_{\beta i}^{(1)}(s) ds \right], \tag{3.2}$$

conditioned on  $X_{\beta i}^{(1)}(t) = \delta_{\beta i}, X(t) = x$ , and

$$\begin{aligned} \partial_{ij} v(t, x) = E & \left[ \partial_{\beta \gamma} g(X(T)) X_{\beta ij}^{(2)}(T) + \partial_{\beta \gamma} g(X(T)) X_{\beta i}^{(1)}(T) X_{\gamma j}^{(1)}(T) \right] \\ & - E \left[ \int_t^T \left( \partial_{\beta f}(s, X(s)) X_{\beta ij}^{(2)}(s) + \partial_{\beta \gamma} f(s, X(s)) X_{\beta i}^{(1)}(s) X_{\gamma j}^{(1)}(s) \right) ds \right], \end{aligned} \tag{3.3}$$

conditioned on  $X_{\beta ij}^{(2)}(t) = 0, X_{\beta i}^{(1)}(t) = \delta_{\beta i}, X(t) = x$ . Hence

$$\begin{aligned} g_i(Z) &= \partial_{\beta i} g(X) X_{\beta i}^{(1)}, & g_{ij}(Z) &= \partial_{\beta \gamma} g(X) X_{\beta ij}^{(2)} + \partial_{\beta \gamma} g(X) X_{\beta i}^{(1)} X_{\gamma j}^{(1)}, \\ f_i(t, Z) &= \partial_{\beta f}(t, X) X_{\beta i}^{(1)}, & f_{ij}(t, Z) &= \partial_{\beta f}(t, X) X_{\beta ij}^{(2)} + \partial_{\beta \gamma} f(t, X) X_{\beta i}^{(1)} X_{\gamma j}^{(1)}. \end{aligned} \tag{3.4}$$

Next we define

$$\Delta_\alpha v(t_k, z) := E \left[ g_\alpha(Z^\Delta(T)) - \int_{t_k}^T f_\alpha(\phi(s), Z^\Delta(\phi(s))) ds | Z^\Delta(t_k) = z \right], \tag{3.5}$$

whenever  $k \in \{0, 1, \dots, N-1\}$ . We let  $G(z)$  and  $F(t, z)$  denote, respectively, the column vectors having an enumeration of  $\{g_\alpha(z)\}$  and  $\{f_\alpha(t, z)\}$  as their components. Similarly, for  $k \in \{0, 1, \dots, N-1\}$  and  $z \in \mathbb{R}^{\tilde{n}}$ , we let  $V(t_k, z) = V_{G,F}(t_k, z)$  and  $V^\Delta(t_k, z) = V_{G^\Delta, F^\Delta}^\Delta(t_k, z)$  be the column vectors having an enumeration of  $\{\partial_\alpha v(t_k, z)\}$  and  $\{\Delta_\alpha v(t_k, z)\}$ , respectively, as their components. Moreover, we let

$$\Delta_{G,F}^\Delta(t_k, z) := V(t_k, z) - V^\Delta(t_k, z), \quad \text{whenever } k \in \{0, 1, \dots, N-1\}, z \in \mathbb{R}^{\tilde{n}}, \tag{3.6}$$

and we note that  $\Delta_{G,F}^\Delta(t_k, z)$  is a column vector of the same dimensions as  $V(t_k, z)$  and  $V^\Delta(t_k, z)$ . To continue we let

$$D_{ij}(t, z) = \frac{1}{2} [\Sigma \Sigma^*]_{ij}(t, z), \quad \text{whenever } (t, z) \in \mathbb{R}_+ \times \mathbb{R}^{\tilde{n}}, \tag{3.7}$$

and

$$\Lambda^\Delta(t, z) = \Lambda(\phi(t), z), \quad \Sigma_j^\Delta(t, z) = \Sigma_j(\phi(t), z), \quad D_{ij}^\Delta(t, z) = \frac{1}{2} [\Sigma^\Delta(\Sigma^\Delta)^*]_{ij}(t, z), \tag{3.8}$$

whenever  $(t, z) \in \mathbb{R}_+ \times \mathbb{R}^{\tilde{n}}$ . Furthermore, we write

$$\begin{aligned} \tilde{L} &= A_i(t, z) \partial_i + D_{ij}(t, z) \partial_{ij}, \\ \tilde{L}^\Delta &= A_i^\Delta(t, z) \partial_i + D_{ij}^\Delta(t, z) \partial_{ij}. \end{aligned} \tag{3.9}$$

In this general setting we first derive the following representation formula for the time-discretization error  $\Delta_{G,F}^\Delta(t_k, z)$ .

**Lemma 3.1.** Let  $T > 0$  be given, let  $f \in C_p^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_p^\infty(\mathbb{R}^n)$  and assume that (1.2) and (2.2) hold. Let  $\Delta_{G,F}^\Delta(t_k, z)$  be defined as in (3.6) whenever  $k \in \{0, 1, \dots, N - 1\}$  and  $z \in \mathbb{R}^{\tilde{n}}$ . Then  $\Delta_{G,F}^\Delta(t_k, z)$  equals

$$\int_{t_k}^T E \left[ (\tilde{L} - \tilde{L}^\Delta)V(t, Z^\Delta(t)) + F(\phi(t), Z^\Delta(\phi(t))) - F(t, Z^\Delta(t)) \mid Z^\Delta(t_k) = z \right] dt. \tag{3.10}$$

**Proof.** To prove the error representation, we first apply Itô’s formula to  $dV(t, Z^\Delta(t))$

$$dV(t, Z^\Delta(t)) = (\partial_t V + \tilde{L}^\Delta V)(t, Z^\Delta(t))dt + \Sigma_{ij}^\Delta(t, Z^\Delta(t))\partial_{ij}V(t, Z^\Delta(t))dW_j(t). \tag{3.11}$$

By Theorem 2.3,  $V$  solves the equation  $\partial_t V + \tilde{L}V = F$  and hence we can eliminate the term  $\partial_t V(t, Z^\Delta(t))$  in (3.11) and obtain

$$dV(t, Z^\Delta(t)) - F(t, Z^\Delta(t))dt = (\tilde{L}^\Delta - \tilde{L})V(t, Z^\Delta(t))dt + \Sigma_{ij}^\Delta(t, Z^\Delta(t))\partial_{ij}V(t, Z^\Delta(t))dW_j(t). \tag{3.12}$$

Integrating both sides of (3.12) and taking expectations we see, using standard arguments, that

$$\begin{aligned} E[V(T, Z^\Delta(T)) \mid Z^\Delta(t_k) = z] - E[V(t_k, Z^\Delta(t_k)) \mid Z^\Delta(t_k) = z] &= \int_{t_k}^T E[F(t, Z^\Delta(t)) \mid Z^\Delta(t_k) = z] dt \\ &= \int_{t_k}^T E[(\tilde{L}^\Delta - \tilde{L})V(t, Z^\Delta(t)) \mid Z^\Delta(t_k) = z] dt. \end{aligned} \tag{3.13}$$

Furthermore,

$$\begin{aligned} E[V(T, Z^\Delta(T)) \mid Z^\Delta(t_k) = z] - E[V(t_k, Z^\Delta(t_k)) \mid Z^\Delta(t_k) = z] &= E[G(Z^\Delta(T)) \mid Z^\Delta(t_k) = z] - V(t_k, z) \\ &= V^\Delta(t_k, z) + E \left[ \int_{t_k}^T F(\phi(s), Z^\Delta(\phi(s))) ds \mid Z^\Delta(t_k) = z \right] - V(t_k, z). \end{aligned} \tag{3.14}$$

Now combining (3.13) and (3.14), the lemma follows readily.  $\square$

To proceed, we define

$$\begin{aligned} \Gamma_i^{(k)} &= A_i(t_{k+1}, Z^\Delta(t_{k+1})) - A_i(t_k, Z^\Delta(t_k)), \\ \Gamma_{ij}^{(k)} &= D_{ij}(t_{k+1}, Z^\Delta(t_{k+1})) - D_{ij}(t_k, Z^\Delta(t_k)), \end{aligned} \tag{3.15}$$

whenever  $k \in \{0, 1, \dots, N - 1\}$ ,  $i, j \in \{1, \dots, \tilde{n}\}$ . Next, we introduce

$$\begin{aligned} \tilde{\Delta}_{G,F}^\Delta(t_k, z) &= \sum_{h=k}^{N-1} E \left[ \Gamma_i^{(h)} \partial_i V(t_{h+1}, Z^\Delta(t_{h+1})) + \Gamma_{ij}^{(h)} \partial_{ij} V(t_{h+1}, Z^\Delta(t_{h+1})) \right] \frac{\Delta t_h}{2} \\ &\quad + \sum_{h=k}^{N-1} E [F(t_h, Z^\Delta(t_h)) - F(t_{h+1}, Z^\Delta(t_{h+1}))] \frac{\Delta t_h}{2}, \end{aligned} \tag{3.16}$$

conditioned on the event  $Z^\Delta(t_k) = z$  and, similarly,

$$\begin{aligned} \bar{\Delta}_{G,F}^\Delta(t_k, z) &= \sum_{h=k}^{N-1} E \left[ \Gamma_i^{(h)} \partial_i V^\Delta(t_{h+1}, Z^\Delta(t_{h+1})) + \Gamma_{ij}^{(h)} \partial_{ij} V^\Delta(t_{h+1}, Z^\Delta(t_{h+1})) \right] \frac{\Delta t_h}{2} \\ &\quad + \sum_{h=k}^{N-1} E [F(t_h, Z^\Delta(t_h)) - F(t_{h+1}, Z^\Delta(t_{h+1}))] \frac{\Delta t_h}{2}, \end{aligned} \tag{3.17}$$

conditioned on the event  $Z^\Delta(t_k) = z$ . Equipped with this notation we derive the following lemma, which provides an expansion of the error  $\Delta_{G,F}^\Delta(t_k, z)$  with  $\tilde{\Delta}_{G,F}^\Delta(t_k, z)$  as the leading order term.

**Lemma 3.2.** Let  $T > 0$  be given, let  $f \in C_p^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_p^\infty(\mathbb{R}^n)$  and assume that (1.2) and (2.2) hold. Let  $\Delta_{G,F}^\Delta(t_k, z)$ ,  $\tilde{\Delta}_{G,F}^\Delta(t_k, z)$  and  $\bar{\Delta}_{G,F}^\Delta(t_k, z)$  be defined as in (3.6) and (3.16) and (3.17), respectively, whenever  $k \in \{0, 1, \dots, N - 1\}$ ,  $z \in \mathbb{R}^{\tilde{n}}$ . Then

$$\Delta_{G,F}^\Delta(t_k, z) - \tilde{\Delta}_{G,F}^\Delta(t_k, z) = \sum_{h=k}^{N-1} \mathcal{O}((\Delta t_h)^3), \tag{3.18}$$

$$\tilde{\Delta}_{G,F}^\Delta(t_k, z) - \bar{\Delta}_{G,F}^\Delta(t_k, z) = \sum_{h=k}^{N-1} \left( (\Delta t_h)^2 \sum_{p=h}^{N-1} \mathcal{O}((\Delta t_p)^2) \right), \tag{3.19}$$

which immediately implies that

$$\Delta_{G,F}^\Delta(t_k, z) - \bar{\Delta}_{G,F}^\Delta(t_k, z) = \mathcal{O}((\Delta_N^*)^2), \tag{3.20}$$

where  $\Delta_N^* = \max \{ \Delta t_0, \Delta t_1, \dots, \Delta t_{N-1} \}$ .

**Proof.** Since all components of  $\Lambda_i, D_{ij}, V, G$  and  $F$  are smooth and have polynomial growth, (3.18) follows by a standard interpolation estimate, analogous to Lemma 2.3 in [15]. To prove (3.19), we first conclude, proceeding as in the proof of Lemma 3.1, that

$$V(t_h, Z^\Delta(t_k)) - V^\Delta(t_h, Z^\Delta(t_k)) = \int_{t_h}^T E [H(t, Z^\Delta(t)) | \mathcal{F}_{t_h}] dt, \tag{3.21}$$

where

$$H(t, Z^\Delta(t)) = (\tilde{L} - \tilde{L}^\Delta)V(t, Z^\Delta(t)) + F(\phi(t), Z^\Delta(\phi(t))) - F(t, Z^\Delta(t)). \tag{3.22}$$

The generator of the process  $Z^\Delta(t)$  is  $\tilde{L}^\Delta$  and hence, for  $t_h \leq t_p \leq t < t_{p+1}$ ,

$$E [H(t, Z^\Delta(t)) | \mathcal{F}_{t_h}] = \int_{t_p}^t E [\tilde{L}^\Delta H(s, Z^\Delta(s)) | \mathcal{F}_{t_h}] ds. \tag{3.23}$$

Since all components of  $\Lambda_i, \Lambda_i^\Delta, D_{ij}, D_{ij}^\Delta, V, G$  and  $F$  are smooth and have polynomial growth, the integrand in (3.23) is bounded and, hence, the right-hand side of (3.23) is of order  $\mathcal{O}(\Delta t_p)$ . Inserting this estimate into (3.21), we immediately obtain

$$V(t_h, Z^\Delta(t_h)) - V^\Delta(t_h, Z^\Delta(t_h)) = \sum_{p=h}^{N-1} \mathcal{O}((\Delta t_p)^2). \tag{3.24}$$

The remainder of the proof of (3.19) now follows directly along the lines of the proof of Lemma 2.4 in [15].  $\square$

Finally we note that if  $F \equiv 0$ , then

$$V_{G,0}^\Delta(t_k, z) = E [G(Z^\Delta(T)) | Z^\Delta(t_k) = z], \tag{3.25}$$

$$\Delta_{G,0}^\Delta(t_k, z) = E [G(Z(T)) - G(Z^\Delta(T)) | Z(t_k) = Z^\Delta(t_k) = z], \tag{3.26}$$

and

$$\bar{\Delta}_{G,0}^\Delta(t_k, z) = \sum_{h=k}^{N-1} E \left[ \Gamma_i^{(h)} \partial_i V^\Delta(t_{h+1}, Z^\Delta(t_{h+1})) + \Gamma_{ij}^{(h)} \partial_{ij} V^\Delta(t_{h+1}, Z^\Delta(t_{h+1})) \right] \frac{\Delta t_h}{2}, \tag{3.27}$$

conditioned on the event  $Z^\Delta(t_k) = z$ , with  $V^\Delta = V_{G,0}^\Delta$ .

#### 4. Error expansions in a posteriori form

Let  $T > 0$  be given, let  $f \in C_p^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_p^\infty(\mathbb{R}^n)$  and assume that (1.2) and (2.2) hold. Let  $\bar{v}_{g,f}$  be the solution to (2.7) for data given by  $g, f$  and parameter values  $(\theta_\mu, \theta_\sigma) = (0, 0)$ . Furthermore, we let, as in the introduction,  $\{\bar{X}(t) : t \in [0, T]\}$  solve (1.3) for parameter values  $(\theta_\mu, \theta_\sigma) = (0, 0)$  and we let  $\{\bar{X}^\Delta(t) : t \in [0, T]\}$  be the corresponding continuous Euler approximation introduced in (1.8)–(1.9). Moreover, in the following, we let

$$\begin{aligned} \bar{\mu}_i^\Delta(t, \bar{X}^\Delta(t)) &= \bar{\mu}_i(\phi(t), \bar{X}^\Delta(\phi(t))), & \bar{\sigma}_{ij}^\Delta(t, \bar{X}^\Delta(t)) &= \bar{\sigma}_{ij}(\phi(t), \bar{X}^\Delta(\phi(t))), \\ \bar{a}_{ij}^\Delta(t, \bar{X}^\Delta(t)) &= \bar{a}_{ij}(\phi(t), \bar{X}^\Delta(\phi(t))), \end{aligned} \tag{4.1}$$

whenever  $t \in [0, T]$ . Recall that  $\bar{a}_{ij}$  was introduced in (2.3). Let  $\bar{Z}(t)$  be the vector defined in (2.13) based on  $\bar{X}(t)$  and let  $\bar{Z}^\Delta(t)$  be the Euler discretization of  $\bar{Z}(t)$ . Note that the first  $n$  components of the vector  $\bar{Z}(t)$  introduced in (2.13) equal  $\bar{X}(t)$  and also that  $\bar{Z}^\Delta(t) = (\bar{X}^\Delta(t), \bar{X}^{(1),\Delta}(t), \bar{X}^{(2),\Delta}(t), \bar{X}^{(3),\Delta}(t), \bar{X}^{(4),\Delta}(t))$ . We define

$$\bar{v}_{g,f}^\Delta(t_k, x) = E \left[ g(\bar{X}^\Delta(T)) - \int_{t_k}^T f(\phi(t), \bar{X}^\Delta(\phi(t))) dt | \bar{X}^\Delta(t_k) = x \right], \tag{4.2}$$

whenever  $k \in \{0, 1, \dots, N - 1\}$  and we let

$$\Delta_{g,f}^\Delta(t_k, x) := \bar{v}_{g,f}^\Delta(t_k, x) - \bar{v}_{g,f}(t_k, x), \quad \text{whenever } k \in \{0, 1, \dots, N - 1\}. \tag{4.3}$$

Moreover, we let  $\Delta_{g,f}^\Delta(x) := \Delta_{g,f}^\Delta(0, x)$  and write

$$\begin{aligned} \bar{L} &= \bar{\mu}_i(t, x) \partial_i + \bar{a}_{ij}(t, x) \partial_{ij}, \\ \bar{L}^\Delta &= \bar{\mu}_i^\Delta(t, x) \partial_i + \bar{a}_{ij}^\Delta(t, x) \partial_{ij}. \end{aligned} \tag{4.4}$$

Then, arguing as in the proof of Lemma 3.1 it follows that  $\Delta_{g,f}^\Delta(x)$  equals

$$\int_0^T E [\bar{L} - \bar{L}^\Delta] \bar{v}_{g,f}(t, \bar{X}^\Delta(t)) + f(\phi(t), \bar{X}^\Delta(\phi(t))) - f(t, \bar{X}^\Delta(t)) | \bar{X}^\Delta(0) = x \, dt. \tag{4.5}$$

To proceed, we let

$$\begin{aligned} \bar{\epsilon}_i^{(k)} &= \bar{\mu}_i(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k)), \\ \bar{\epsilon}_{ij}^{(k)} &= \bar{a}_{ij}(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - \bar{a}_{ij}(t_k, \bar{X}^\Delta(t_k)), \end{aligned} \tag{4.6}$$

whenever  $k \in \{0, 1, \dots, N - 1\}$ ,  $i, j \in \{1, \dots, n\}$ , and we introduce

$$\begin{aligned} \bar{\Delta}_{g,f}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{\epsilon}_i^{(k)} \partial_i \bar{v}_{g,f}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) + \bar{\epsilon}_{ij}^{(k)} \partial_{ij} \bar{v}_{g,f}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \right] \frac{\Delta t_k}{2} \\ &\quad + \sum_{k=0}^{N-1} E [f(t_k, \bar{X}^\Delta(t_k)) - f(t_{k+1}, \bar{X}^\Delta(t_{k+1}))] \frac{\Delta t_k}{2}, \end{aligned} \tag{4.7}$$

conditioned on the event  $\bar{X}^\Delta(0) = x$ . Now, by reproducing the proof of Lemma 3.2, we arrive at the following lemma.

**Lemma 4.1.** *Let  $T > 0$  be given, let  $f \in C_p^\infty((0, T) \times \mathbb{R}^n)$ ,  $g \in C_p^\infty(\mathbb{R}^n)$  and assume that (1.2) and (2.2) hold. Let  $\bar{v}_{g,f}$  be the solution to (2.7) for data given by  $g, f$  and parameter values  $(\theta_\mu, \theta_\sigma) = (0, 0)$ , let  $\bar{v}_{g,f}^\Delta$  and  $\Delta_{g,f}^\Delta$  be defined as in (4.2) and (4.3) respectively. Furthermore, let  $\bar{\Delta}_{g,f}^\Delta$  be defined as in (4.7). Then,*

$$\Delta_{g,f}^\Delta(x) - \bar{\Delta}_{g,f}^\Delta(x) = \mathcal{O}((\Delta_N^*)^2), \tag{4.8}$$

where  $\Delta_N^* = \max \{\Delta t_0, \Delta t_1, \dots, \Delta t_{N-1}\}$ .

#### 4.1. Discrete dual functions

Recall that

$$\bar{u}^\Delta(t_k, x) = E [g(\bar{X}^\Delta(T)) | \bar{X}^\Delta(t_k) = x]. \tag{4.9}$$

In the following we introduce appropriate dual functions associated to  $g, \bar{X}^\Delta$  and related to  $\bar{u}^\Delta$ . In particular, a simple generalization of Lemma 2.5 in [15] shows that

$$\begin{aligned} \partial_i \bar{u}^\Delta(t_k, \bar{X}^\Delta(t_k)) &= E[\bar{\psi}_i(t_k) | \mathcal{F}_{t_k}], \\ \partial_{ij} \bar{u}^\Delta(t_k, \bar{X}^\Delta(t_k)) &= E[\bar{\psi}_{ij}^{(1)}(t_k) | \mathcal{F}_{t_k}], \\ \partial_{ijl} \bar{u}^\Delta(t_k, \bar{X}^\Delta(t_k)) &= E[\bar{\psi}_{ijl}^{(2)}(t_k) | \mathcal{F}_{t_k}], \\ \partial_{ijlq} \bar{u}^\Delta(t_k, \bar{X}^\Delta(t_k)) &= E[\bar{\psi}_{ijlq}^{(3)}(t_k) | \mathcal{F}_{t_k}], \end{aligned} \tag{4.10}$$

for  $i, j, l, q \in \{1, \dots, n\}$  and  $k \in \{0, \dots, N\}$ . The functions  $\bar{\psi}, \bar{\psi}^{(1)}, \bar{\psi}^{(2)}$  and  $\bar{\psi}^{(3)}$  are referred to as dual functions and can be explicitly calculated by means of certain backwards in time difference equations. In particular, let

$$c_i(t_k, x) = x_i + \bar{\mu}_i(t_k, x) \Delta t_k + \bar{\sigma}_{i\beta}(t_k, x) \Delta W_\beta(t_k), \tag{4.11}$$

whenever  $i \in \{1, \dots, n\}$ ,  $k \in \{0, \dots, N - 1\}$  and  $x \in \mathbb{R}^n$ . The discrete dual function  $\bar{\psi}$ , associated to  $g$  and  $\bar{X}^\Delta$ , is then recursively defined as

$$\begin{aligned} \bar{\psi}_i(t_N) &= \partial_i g(\bar{X}^\Delta(t_N)), \\ \bar{\psi}_i(t_k) &= \partial_i c_\beta(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_\beta(t_{k+1}), \end{aligned} \tag{4.12}$$

whenever  $i \in \{1, \dots, n\}$  and  $k \in \{0, \dots, N - 1\}$ . The second dual function  $\bar{\psi}^{(1)}$ , which is the first variation of  $\bar{\psi}$ , satisfies

$$\begin{aligned} \bar{\psi}_{ij}^{(1)}(t_N) &= \partial_{ij} g(\bar{X}^\Delta(t_N)), \\ \bar{\psi}_{ij}^{(1)}(t_k) &= \partial_i c_\beta(t_k, \bar{X}^\Delta(t_k)) \partial_j c_\gamma(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) + \partial_{ij} c_\beta(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_\beta(t_{k+1}), \end{aligned} \tag{4.13}$$

for  $i, j \in \{1, \dots, n\}$  and  $k \in \{0, \dots, N - 1\}$ . Analogously we can, by differentiating  $\bar{\psi}^{(1)}$ , derive the following recursive relations for the third dual function  $\bar{\psi}^{(2)}$ , for  $i, j, l \in \{1, \dots, n\}$  and  $k \in \{0, \dots, N - 1\}$ ,

$$\begin{aligned} \bar{\psi}_{ijl}^{(2)}(t_N) &= \partial_{ijl}g(\bar{X}^\Delta(t_N)), \\ \bar{\psi}_{ijl}^{(2)}(t_k) &= \partial_i c_\beta(t_k, \bar{X}^\Delta(t_k)) \partial_j c_\gamma(t_k, \bar{X}^\Delta(t_k)) \partial_l c_\varphi(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{\beta\gamma\varphi}^{(2)}(t_{k+1}) \\ &\quad + \partial_{il} c_\beta(t_k, \bar{X}^\Delta(t_k)) \partial_j c_\gamma(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) + \partial_i c_\beta(t_k, \bar{X}^\Delta(t_k)) \partial_{jl} c_\gamma(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) \\ &\quad + \partial_{ij} c_\beta(t_k, \bar{X}^\Delta(t_k)) \partial_l c_\gamma(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) + \partial_{ijl} c_\beta(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_\beta(t_{k+1}). \end{aligned} \tag{4.14}$$

Finally, a similar calculation also yields a recursive scheme for the calculation of the fourth dual function  $\bar{\psi}^{(3)}$ , but we omit the details.

#### 4.2. Calculation of $\bar{u}$

In this section we apply the general theory above to the special case of calculating  $\bar{u}$ . This corresponds to setting  $f = 0$  in the above deductions and we emphasize that the results in this section has previously been derived in [15]. To proceed we see, by applying the results above to the case  $f = 0$ , that

$$\bar{u}(x) = \bar{u}(0, x) = u(0, x, (0, 0)) = \bar{u}^\Delta(x) + \bar{\Delta}_{g,0}^\Delta(x) + \bar{R}_d^\Delta \tag{4.15}$$

where

$$\bar{\Delta}_{g,0}^\Delta(x) = \sum_{k=0}^{N-1} E \left[ \bar{\epsilon}_i^{(k)} \bar{\psi}_i(t_{k+1}) + \bar{\epsilon}_{ij}^{(k)} \bar{\psi}_{ij}^{(1)}(t_{k+1}) \right] \frac{\Delta t_k}{2}, \tag{4.16}$$

and  $\bar{R}_d^\Delta = \mathcal{O}((\Delta_N^*)^2)$ . It is standard to determine  $\bar{u}^\Delta(x)$  by means of the Monte Carlo estimator

$$\bar{u}^{\Delta,M}(x) = \frac{1}{M} \sum_{m=1}^M g(\bar{X}^\Delta(T, \omega_m)), \tag{4.17}$$

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

$$\bar{u}(x) = \bar{u}^{\Delta,M}(x) + \underbrace{\bar{u}(x) - \bar{u}^\Delta(x)}_{\bar{E}_d^\Delta(x)} + \underbrace{\bar{u}^\Delta(x) - \bar{u}^{\Delta,M}(x)}_{\bar{E}_s^{\Delta,M}(x)}, \tag{4.18}$$

where  $\bar{E}_d^\Delta(x)$  and  $\bar{E}_s^{\Delta,M}(x)$  represent the time-discretization error and the statistical error, respectively. For  $k \in \{0, \dots, N - 1\}$  and  $m \in \{1, \dots, M\}$ , we define

$$\bar{\rho}_k(\omega_m) = \frac{\bar{\epsilon}_i^{(k)}(\omega_m) \bar{\psi}_i(t_{k+1}, \omega_m) + \bar{\epsilon}_{ij}^{(k)}(\omega_m) \bar{\psi}_{ij}^{(1)}(t_{k+1}, \omega_m)}{2\Delta t_k}. \tag{4.19}$$

Then, using (4.16) and (4.18), we see that

$$\bar{E}_d^\Delta(x) = \bar{E}_d^{\Delta,M}(x) + \bar{E}_{d,s}^{\Delta,M}(x) + \bar{R}_d^\Delta, \tag{4.20}$$

where

$$\bar{E}_d^{\Delta,M}(x) = \frac{1}{M} \sum_{k=0}^{N-1} \sum_{m=1}^M \bar{\rho}_k(\omega_m) (\Delta t_k)^2, \tag{4.21}$$

and

$$\bar{E}_{d,s}^{\Delta,M}(x) = \sum_{k=0}^{N-1} E [\bar{\rho}_k(\cdot)] (\Delta t_k)^2 - \frac{1}{M} \sum_{k=0}^{N-1} \sum_{m=1}^M \bar{\rho}_k(\omega_m) (\Delta t_k)^2. \tag{4.22}$$

Furthermore, by the central limit theorem, we have

$$\bar{E}_{d,s}^{\Delta,M} = \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} I_{k,M} dt, \tag{4.23}$$

where, for each  $k \in \{0, 1, \dots, N - 1\}$ , the random variable  $\sqrt{M}I_{k,M}$  converges as  $M \rightarrow \infty$  to a normally distributed random variable with zero mean and variance

$$\sigma_k^2 = \text{Var} \left[ \bar{\epsilon}_i^{(k)} \bar{\psi}_i(t_{k+1}) \right] + \text{Var} \left[ \bar{\epsilon}_{ij}^{(k)} \bar{\psi}_{ij}^{(1)}(t_{k+1}) \right]. \tag{4.24}$$

4.3. Calculation of  $\partial_{\theta_\mu} u$  and  $\partial_{\theta_\sigma} u$

Assume (1.2) and (2.2), let  $T > 0$  be given, let  $g \in C_p^\infty(\mathbb{R}^n)$  and let  $u = u(t, x) = u(t, x, \theta)$  be the unique solution to (1.6) defined with respect to  $\mu$  and  $\sigma$ . Similarly, let  $\bar{u} = \bar{u}(t, x)$  be the unique solution to (1.6) defined with respect to  $\bar{\mu}$  and  $\bar{\sigma}$ . Then, formally differentiating (1.6) with respect to  $\theta_\mu$  and  $\theta_\sigma$ , we see that  $\partial_{\theta_\mu} u(t, x, (0, 0))$  and  $\partial_{\theta_\sigma} u(t, x, (0, 0))$  satisfy

$$\begin{aligned} \partial_t(\partial_{\theta_\mu} u)(t, x) + \bar{L}(\partial_{\theta_\mu} u)(t, x) &= -\tilde{\mu}_i(t, x) \partial_i \bar{u}(t, x), \\ \partial_t(\partial_{\theta_\sigma} u)(t, x) + \bar{L}(\partial_{\theta_\sigma} u)(t, x) &= -\frac{1}{2}[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t, x) \partial_{ij} \bar{u}(t, x), \end{aligned} \tag{4.25}$$

whenever  $(t, x) \in (0, T) \times \mathbb{R}^n$ , and

$$\partial_{\theta_\mu} u(T, x) = \partial_{\theta_\sigma} u(T, x) = 0, \tag{4.26}$$

whenever  $x \in \mathbb{R}^n$ . By (4.25) and (4.26), it is clear that  $(\partial_{\theta_\mu} u)(t, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(t, x, (0, 0))$  solve, respectively, the problem stated in Theorem 2.2 with  $g \equiv 0$  and with right-hand sides

$$\begin{aligned} F_\mu(t, x) := f(t, x) &= -\tilde{\mu}_i(t, x) \partial_i \bar{u}(t, x), \\ F_\sigma(t, x) := f(t, x) &= -\frac{1}{2}[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t, x) \partial_{ij} \bar{u}(t, x), \end{aligned} \tag{4.27}$$

respectively. Recall that  $u$  solves the same problem but with Cauchy data defined by  $g$  and with  $f \equiv 0$ . An application of Theorem 2.3 now yield the following stochastic representation formulas for  $(\partial_{\theta_\mu} u)(t, x, (0, 0))$  and  $(\partial_{\theta_\sigma} u)(t, x, (0, 0))$ .

$$\begin{aligned} (\partial_{\theta_\mu} u)(t, x, (0, 0)) &= E \left[ -\int_t^T F_\mu(s, X(s)) ds | X(t) = x \right], \\ (\partial_{\theta_\sigma} u)(t, x, (0, 0)) &= E \left[ -\int_t^T F_\sigma(s, X(s)) ds | X(t) = x \right]. \end{aligned} \tag{4.28}$$

Hence, by (4.3) and (4.8) we deduce that, in our context,

$$\begin{aligned} (\partial_{\theta_\mu} u)(0, x, (0, 0)) &= \bar{v}_{0, F_\mu}^\Delta(x) + \bar{\Delta}_{0, F_\mu}^\Delta(x) + \bar{R}_{\mu, d}^\Delta, \\ (\partial_{\theta_\sigma} u)(0, x, (0, 0)) &= \bar{v}_{0, F_\sigma}^\Delta(x) + \bar{\Delta}_{0, F_\sigma}^\Delta(x) + \bar{R}_{\sigma, d}^\Delta, \end{aligned} \tag{4.29}$$

where  $\bar{R}_{\mu, d}^\Delta + \bar{R}_{\sigma, d}^\Delta = \mathcal{O}((\Delta_N^*)^2)$ , and furthermore, based on (4.2), (4.7) and (4.27), we have

$$\begin{aligned} \bar{v}_{0, F_\mu}^\Delta(x) &= \sum_{k=0}^{N-1} E [\tilde{\mu}_i(t_k, \bar{X}^\Delta(t_k)) \partial_i \bar{u}(t_k, \bar{X}^\Delta(t_k))] \Delta t_k, \\ \bar{v}_{0, F_\sigma}^\Delta(x) &= \frac{1}{2} \sum_{k=0}^{N-1} E [[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \partial_{ij} \bar{u}(t_k, \bar{X}^\Delta(t_k))] \Delta t_k, \end{aligned} \tag{4.30}$$

and

$$\begin{aligned} \bar{\Delta}_{0, F_\mu}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{\epsilon}_i^{(k)} \partial_i \bar{v}_{0, F_\mu}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) + \bar{\epsilon}_{ij}^{(k)} \partial_{ij} \bar{v}_{0, F_\mu}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \right] \frac{\Delta t_k}{2} \\ &\quad + \sum_{k=0}^{N-1} E [[\tilde{\mu}_i \partial_i \bar{u}](t_{k+1}, \bar{X}^\Delta(t_{k+1})) - [\tilde{\mu}_i \partial_i \bar{u}](t_k, \bar{X}^\Delta(t_k))] \frac{\Delta t_k}{2}, \\ \bar{\Delta}_{0, F_\sigma}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{\epsilon}_i^{(k)} \partial_i \bar{v}_{0, F_\sigma}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) + \bar{\epsilon}_{ij}^{(k)} \partial_{ij} \bar{v}_{0, F_\sigma}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \right] \frac{\Delta t_k}{2} \\ &\quad + \sum_{k=0}^{N-1} E [[[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij} \partial_{ij} \bar{u}](t_{k+1}, \bar{X}^\Delta(t_{k+1})) - [[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij} \partial_{ij} \bar{u}](t_k, \bar{X}^\Delta(t_k))] \frac{\Delta t_k}{4}, \end{aligned} \tag{4.31}$$

conditioned on the event  $\bar{X}^\Delta(0) = x$ . For  $v \in \{\bar{u}, \bar{u}^\Delta\}$ , we introduce the notation

$$\begin{aligned} \bar{A}_{F_\mu, k}^\Delta(v) &= \tilde{\mu}_i(t_k, \bar{X}^\Delta(t_k)) \partial_i v(t_k, \bar{X}^\Delta(t_k)), \\ \bar{A}_{F_\sigma, k}^\Delta(v) &= [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \partial_{ij} v(t_k, \bar{X}^\Delta(t_k)), \end{aligned} \tag{4.32}$$

and, as a consequence, (4.30) can be neatly rewritten as

$$\begin{aligned} \bar{v}_{0,F\mu}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{A}_{F\mu,k}^\Delta(\bar{u}) | \bar{X}^\Delta(0) = x \right] \Delta t_k, \\ \bar{v}_{0,F\sigma}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u}) | \bar{X}^\Delta(0) = x \right] \frac{\Delta t_k}{2}. \end{aligned} \tag{4.33}$$

Moreover, arguing as in (3.2)–(3.4), we can calculate the first and second order derivatives of  $\bar{v}_{0,F\mu}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1}))$  and  $\bar{v}_{0,F\sigma}^\Delta(t_{k+1}, \bar{X}^\Delta(t_{k+1}))$  explicitly, conditioned on the event  $\bar{X}^\Delta(0) = x$ . Indeed, introducing, for  $v \in \{\bar{u}, \bar{u}^\Delta\}$ , the notation

$$\begin{aligned} \bar{B}_{F\mu,k}^\Delta(v) &= \bar{\mu}_i(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \partial_i v(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k)) \partial_i v(t_k, \bar{X}^\Delta(t_k)), \\ \bar{B}_{F\sigma,k}^\Delta(v) &= [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \partial_{ij} v(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \partial_{ij} v(t_k, \bar{X}^\Delta(t_k)), \\ \bar{C}_{F\mu,k,h}^\Delta(v) &= \bar{\epsilon}_{ij}^{(k)} \partial_i [\bar{\mu}_{r_1} \partial_{r_1} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(1),\Delta}(t_{k+1}, t_h) + \bar{\epsilon}_{ij}^{(k)} \partial_i [\bar{\mu}_{r_1} \partial_{r_1} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(2),\Delta}(t_{k+1}, t_h) \\ &\quad + \bar{\epsilon}_{ij}^{(k)} \partial_{im} [\bar{\mu}_{r_1} \partial_{r_1} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(1),\Delta}(t_{k+1}, t_h) \bar{X}_{mj}^{(1),\Delta}(t_{k+1}, t_h), \end{aligned} \tag{4.34}$$

$$\begin{aligned} \bar{C}_{F\sigma,k,h}^\Delta(v) &= \bar{\epsilon}_{ij}^{(k)} \partial_i [[\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{r_1 r_2} \partial_{r_1 r_2} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(1),\Delta}(t_{k+1}, t_h) \\ &\quad + \bar{\epsilon}_{ij}^{(k)} \partial_i [[\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{r_1 r_2} \partial_{r_1 r_2} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(2),\Delta}(t_{k+1}, t_h) \\ &\quad + \bar{\epsilon}_{ij}^{(k)} \partial_{im} [[\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{r_1 r_2} \partial_{r_1 r_2} v](t_h, \bar{X}^\Delta(t_h)) \bar{X}_{ij}^{(1),\Delta}(t_{k+1}, t_h) \bar{X}_{mj}^{(1),\Delta}(t_{k+1}, t_h), \end{aligned} \tag{4.35}$$

we deduce that (4.31) can be rewritten as

$$\begin{aligned} \bar{\Delta}_{0,F\mu}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{B}_{F\mu,k}^\Delta(\bar{u}) \right] \frac{\Delta t_k}{2} + \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} E \left[ E \left[ \bar{C}_{F\mu,k,h}^\Delta(\bar{u}) | \bar{Z}^\Delta(t_{k+1}) = z(\bar{X}^\Delta(t_{k+1})) \right] \right] \frac{\Delta t_h \Delta t_k}{2}, \\ \bar{\Delta}_{0,F\sigma}^\Delta(x) &= \sum_{k=0}^{N-1} E \left[ \bar{B}_{F\sigma,k}^\Delta(\bar{u}) \right] \frac{\Delta t_k}{4} + \sum_{h=k+1}^{N-1} E \left[ E \left[ \bar{C}_{F\sigma,k,h}^\Delta(\bar{u}) | \bar{Z}^\Delta(t_{k+1}) = z(\bar{X}^\Delta(t_{k+1})) \right] \right] \frac{\Delta t_h \Delta t_k}{4}, \end{aligned} \tag{4.36}$$

conditioned on  $\bar{X}^\Delta(0) = x$ . For the remainder of this section we consider, for the sake of brevity, only the calculation of  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$ . The calculation of  $(\partial_{\theta_\mu} u)(0, x, (0, 0))$  proceeds analogously. Then, to start with, combining (4.29), (4.33) and (4.36), we see that  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$  equals

$$\begin{aligned} &\sum_{k=0}^{N-1} E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u}) | \bar{X}^\Delta(0) = x \right] \frac{\Delta t_k}{2} + \sum_{k=0}^{N-1} E \left[ \bar{B}_{F\sigma,k}^\Delta(\bar{u}) | \bar{X}^\Delta(0) = x \right] \frac{\Delta t_k}{4} \\ &\quad + \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} E \left[ E \left[ \bar{C}_{F\sigma,k,h}^\Delta(\bar{u}) | \bar{Z}^\Delta(t_{k+1}) = z(\bar{X}^\Delta(t_{k+1})) \right] | \bar{X}^\Delta(0) = x \right] \frac{\Delta t_h \Delta t_k}{4} + \mathcal{O}((\Delta_N^*)^2). \end{aligned} \tag{4.37}$$

To find a computable expansion, in a posteriori form, of  $(\partial_{\theta_\sigma} u)(0, x, (0, 0))$  and the time-discretization error produced, we have to replace  $\bar{u}$  by  $\bar{u}^\Delta$  in (4.37). To do this we first note, simply by using linearity, that

$$E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u}) | \bar{X}^\Delta(0) = x \right] = E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u}^\Delta) | \bar{X}^\Delta(0) = x \right] + E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u} - \bar{u}^\Delta) | \bar{X}^\Delta(0) = x \right], \tag{4.38}$$

and analogously for the second and third term in (4.37). Hence the three terms in (4.37) can be written as a sum of three terms containing  $\bar{u}^\Delta$  and three terms containing  $\bar{u} - \bar{u}^\Delta$ . The terms containing  $\bar{u}^\Delta$  are computable in a posteriori form and can, by means of (4.10), easily be expressed using the discrete dual functions  $\bar{\psi}, \bar{\psi}^{(1)}, \bar{\psi}^{(2)}$  and  $\bar{\psi}^{(3)}$  defined in (4.12)–(4.14). Concerning the terms containing  $\bar{u} - \bar{u}^\Delta$ , we first note that the  $\bar{\epsilon}$ -factors in  $\bar{C}_{F\sigma,k,h}^\Delta$  give rise to a term of order  $\mathcal{O}(\Delta_N^*)$ , implying that the term including  $\bar{C}_{F\sigma,k,h}^\Delta(\bar{u} - \bar{u}^\Delta)$  is of order  $\mathcal{O}((\Delta_N^*)^2)$ . Similarly, since  $\bar{B}_{F\sigma,k}^\Delta$  can be expanded as

$$\begin{aligned} &[\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \partial_{ij} v(t_{k+1}, \bar{X}^\Delta(t_{k+1})) \\ &\quad + [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \left[ \partial_{ij} v(t_{k+1}, \bar{X}^\Delta(t_{k+1})) - \partial_{ij} v(t_k, \bar{X}^\Delta(t_k)) \right], \end{aligned} \tag{4.39}$$

and since the regularity assumptions assert that the differences in brackets on the right-hand side of (4.39) give rise to terms which are of order  $\mathcal{O}(\Delta_N^*)$ , we conclude that the term including  $\bar{B}_{F\sigma,k}^\Delta(\bar{u} - \bar{u}^\Delta)$  is of order  $\mathcal{O}((\Delta_N^*)^2)$  as well. It remains to consider the term involving  $\bar{A}_{F\sigma,k}^\Delta(\bar{u} - \bar{u}^\Delta)$ . However, this term can be handled as follows. By elementary properties of conditional expectations, we obtain

$$\begin{aligned} E \left[ \bar{A}_{F\sigma,k}^\Delta(\bar{u} - \bar{u}^\Delta) | \bar{X}^\Delta(0) = x \right] &= E \left[ [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \partial_{ij}(\bar{u} - \bar{u}^\Delta)(t_k, \bar{X}^\Delta(t_k)) | \bar{X}^\Delta(0) = x \right] \\ &= E \left[ [\bar{\sigma} \bar{\sigma}^* + \bar{\sigma} \bar{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) E \left[ \partial_{ij}(\bar{u} - \bar{u}^\Delta)(t_k, \bar{X}^\Delta(t_k)) | \bar{Z}^\Delta(t_k) = z(\bar{X}^\Delta(t_k)) \right] | \bar{X}^\Delta(0) = x \right]. \end{aligned} \tag{4.40}$$

Moreover, using (3.26) and (3.27), we have

$$\begin{aligned}
 E \left[ \partial_{ij}(\bar{u} - \bar{u}^\Delta) (t_k, \bar{X}^\Delta(t_k)) \mid \bar{Z}^\Delta(t_k) = z(\bar{X}^\Delta(t_k)) \right] \\
 = \mathcal{O}((\Delta_N^*)^2) + \sum_{h=k}^{N-1} E \left[ \bar{\Gamma}_l^{(h)} \partial_l \bar{V}_{ij}^{k,\Delta} (t_{h+1}, \bar{Z}^\Delta(t_{h+1})) \mid \bar{Z}^\Delta(t_k) = z(\bar{X}^\Delta(t_k)) \right] \frac{\Delta t_h}{2} \\
 + \sum_{h=k}^{N-1} E \left[ \bar{\Gamma}_{lm}^{(h)} \partial_{lm} \bar{V}_{ij}^{k,\Delta} (t_{h+1}, \bar{Z}^\Delta(t_{h+1})) \mid \bar{Z}^\Delta(t_k) = z(\bar{X}^\Delta(t_k)) \right] \frac{\Delta t_h}{2},
 \end{aligned} \tag{4.41}$$

where

$$\bar{V}_{ij}^{k,\Delta} (t_{h+1}, \bar{Z}^\Delta(t_{h+1})) = \partial_{\beta\gamma} g(\bar{X}^\Delta(t_{h+1})) \bar{X}_{\beta ij}^{(2),\Delta} (t_k, t_{h+1}) + \partial_{\beta\gamma} g(\bar{X}^\Delta(t_{h+1})) \bar{X}_{\beta i}^{(1),\Delta} (t_k, t_{h+1}) \bar{X}_{\gamma j}^{(1),\Delta} (t_k, t_{h+1}), \tag{4.42}$$

and the terms  $\bar{\Gamma}_l^{(h)}$  and  $\bar{\Gamma}_{lm}^{(h)}$  are determined by replacing  $Z^\Delta$  by  $\bar{Z}^\Delta$  in (3.15). The derivatives of  $\bar{V}_{ij}^{k,\Delta}$  can be computed by means of first and second order dual function for the extended system  $\bar{Z}^\Delta$ . As we are only interested in first and second order derivatives of the function  $\bar{V}_{ij}^{k,\Delta}$  it suffices to consider the subset of components of the vector

$$\bar{Z}^\Delta(t_k, t_h) = (\bar{X}^\Delta(t_h), \bar{X}^{(1),\Delta}(t_k, t_h), \bar{X}^{(2),\Delta}(t_k, t_h), \bar{X}^{(3),\Delta}(t_k, t_h), \bar{X}^{(4),\Delta}(t_k, t_h))^*$$

which coincides with  $(\bar{X}^\Delta(t_h), \bar{X}^{(1),\Delta}(t_k, t_h), \bar{X}^{(2),\Delta}(t_k, t_h))^*$ . Hence, with a slight abuse of notation, we shall, in the following, let  $\bar{Z}^\Delta(t_h)$  denote the vector

$$(\bar{X}^\Delta(t_h), \bar{X}^{(1),\Delta}(t_k, t_h), \bar{X}^{(2),\Delta}(t_k, t_h))^*$$

and we note that this vector contains  $\hat{n} = n + n^2 + n^3 = n(n^3 - 1)/(n - 1)$  elements. Naturally  $\bar{Z}^\Delta(t_h)$  can be considered as consisting of three blocks of components and in the following we, for clarity, treat the three blocks of components of  $\bar{Z}^\Delta$  separately in the definition of the dual functions. Let  $i, j$  and  $k$  be the indices in the definition of  $\bar{V}_{ij}^{k,\Delta}$ . Then, for  $i, j$  and  $k$  fixed,  $\bar{V}_{ij}^{k,\Delta}$  gives rise to three sets of first order dual functions

$$((\bar{\xi}_{ij}^k)_a^0 \quad (\bar{\xi}_{ij}^k)_{ab}^1 \quad (\bar{\xi}_{ij}^k)_{abc}^2) \tag{4.43}$$

where  $a, b, c \in \{1, \dots, n\}$ . In particular, the sets  $(\bar{\xi}_{ij}^k)_a^0$ ,  $(\bar{\xi}_{ij}^k)_{ab}^1$  and  $(\bar{\xi}_{ij}^k)_{abc}^2$  relate to  $\bar{X}^\Delta$ ,  $\bar{X}^{(1),\Delta}$ , and  $\bar{X}^{(2),\Delta}$  respectively. Similarly, for  $i, j$  and  $k$  fixed,  $\bar{V}_{ij}^{k,\Delta}$  gives rise to nine sets of second order dual functions,

$$\left( \begin{array}{ccc} (\bar{\xi}_{ij}^k)_{a,d}^{00} & (\bar{\xi}_{ij}^k)_{a,de}^{01} & (\bar{\xi}_{ij}^k)_{a,def}^{02} \\ (\bar{\xi}_{ij}^k)_{ab,d}^{10} & (\bar{\xi}_{ij}^k)_{ab,de}^{11} & (\bar{\xi}_{ij}^k)_{ab,def}^{12} \\ (\bar{\xi}_{ij}^k)_{abc,d}^{20} & (\bar{\xi}_{ij}^k)_{abc,de}^{21} & (\bar{\xi}_{ij}^k)_{abc,def}^{22} \end{array} \right), \tag{4.44}$$

for  $a, b, c, d, e, f \in \{1, \dots, n\}$ . Note that in (4.43) and (4.44) the number of elements are of the order  $n^3$  and  $n^6$ , respectively, and the calculation of the second order dual functions in (4.44) may seem prohibitively extensive and expensive for large  $n$ . However, in many application the ‘matrix’ of dual functions in (4.44) turns out to be very sparse in the sense that many entries are zero. In particular, in our example in Section 6 concerning the LIBOR market models the number of nonzero entries turns out to be of the order  $n^2$  instead of  $n^6$ . To get a more thorough understanding of the complexity of the set of elements in (4.43) and (4.44) we refer the reader to the examples in Sections 5 and 6. To continue we, in the following, let  $(x, x^1, x^2)$  denote an  $\hat{n}$ -dimensional vector describing the components of the vector  $(\bar{X}^\Delta, \bar{X}^{(1),\Delta}, \bar{X}^{(2),\Delta})^*$ . Equipped with this notation we note that in the extended system the counterpart of (4.11) are the three equations

$$w_r^0(t_h, x, x^1, x^2) = x_r + \bar{\mu}_r(t_h, x) \Delta t_h + \bar{\sigma}_{r\beta}(t_h, x) \Delta W_\beta(t_h), \tag{4.45}$$

$$w_{rs}^1(t_h, x, x^1, x^2) = x_{rs}^1 + \partial_\gamma \bar{\mu}_r(t_h, x) x_{\gamma s}^1 \Delta t_h + \partial_\gamma \bar{\sigma}_{r\beta}(t_h, x) x_{\gamma s}^1 \Delta W_\beta(t_h), \tag{4.46}$$

$$\begin{aligned}
 w_{rst}^2(t_h, x, x^1, x^2) &= x_{rst}^2 + (\partial_\gamma \bar{\mu}_r(t_h, x) x_{\gamma st}^2 + \partial_{\gamma\varphi} \bar{\mu}_r(t_h, x) x_{\gamma s}^1 x_{\varphi t}^1) \Delta t_h \\
 &\quad + (\partial_\gamma \bar{\sigma}_{r\beta}(t_h, x) x_{\gamma st}^2 + \partial_{\gamma\varphi} \bar{\sigma}_{r\beta}(t_h, x) x_{\gamma s}^1 x_{\varphi t}^1) \Delta W_\beta(t_h).
 \end{aligned} \tag{4.47}$$

Now, the first order dual function  $(\bar{\xi}_{ij}^k)_a^0$  is defined by the recursive relation

$$\begin{aligned}
 (\bar{\xi}_{ij}^k)_a^0(T) &= \partial_a \bar{V}_{ij}^{k,\Delta}(T, \bar{Z}^\Delta(T)), \\
 (\bar{\xi}_{ij}^k)_a^0(t_h) &= \partial_a w_r^0(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)_r^0(t_{h+1}) + \partial_a w_{rs}^1(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)_{rs}^1(t_{h+1}) \\
 &\quad + \partial_a w_{rst}^2(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)_{rst}^2(t_{h+1}),
 \end{aligned} \tag{4.48}$$

for  $h < N$ . Moreover, analogous relations hold for  $(\bar{\xi}_{ij}^k)^1_{ab}$  and  $(\bar{\xi}_{ij}^k)^2_{abc}$ . To define the second order dual function  $(\bar{\xi}_{ij}^k)^{00}_{a,d}$ , we note that

$$(\bar{\xi}_{ij}^k)^{00}_{a,d}(T) = \partial_{ad} \bar{V}_{ij}^{k,\Delta}(T, \bar{Z}^\Delta(T)), \tag{4.49}$$

and that  $(\bar{\xi}_{ij}^k)^{00}_{a,d}(t_h)$ , for  $h < N$ , equals

$$\begin{aligned} &\partial_a w_r^0(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_u^0(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{00}_{r,u}(t_{h+1}) + \partial_a w_r^0(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uv}^1(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{01}_{r,uv}(t_{h+1}) \\ &+ \partial_a w_r^0(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uv}^2(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{02}_{r,uvw}(t_{h+1}) + \partial_a w_{rs}^1(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_u^0(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{10}_{rs,u}(t_{h+1}) \\ &+ \partial_a w_{rs}^1(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uv}^1(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{11}_{rs,uv}(t_{h+1}) + \partial_a w_{rs}^1(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uvw}^2(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{12}_{rs,uvw}(t_{h+1}) \\ &+ \partial_a w_{rst}^2(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_u^0(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{20}_{rst,u}(t_{h+1}) + \partial_a w_{rst}^2(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uv}^1(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{21}_{rst,uv}(t_{h+1}) \\ &+ \partial_a w_{rst}^2(t_h, \bar{Z}^\Delta(t_h)) \partial_d w_{uvw}^2(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^{22}_{rst,uvw}(t_{h+1}) + \partial_a \partial_d w_r^0(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^0_r(t_{h+1}) \\ &+ \partial_a \partial_d w_{rs}^1(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^1_{rs}(t_{h+1}) + \partial_a \partial_d w_{rst}^2(t_h, \bar{Z}^\Delta(t_h)) (\bar{\xi}_{ij}^k)^2_{rst}(t_{h+1}). \end{aligned} \tag{4.50}$$

Furthermore, analogous relations hold for the other eight blocks of dual functions of second order. Using the counterpart of (4.10) associated to the dual functions of the extended system, we obtain

$$\begin{aligned} &E[\partial_{ij}(\bar{u} - \bar{u}^\Delta)(t_k, \bar{X}^\Delta(t_k)) | \bar{Z}^\Delta(t_k) = z(\bar{X}^\Delta(t_k))] \\ &= \sum_{h=k}^{N-1} (\bar{\mu}_a(t_{h+1}, \bar{X}^\Delta(t_{h+1})) - \bar{\mu}_a(t_h, \bar{X}^\Delta(t_h))) (\bar{\xi}_{ij}^k)^0_a(t_{h+1}) \frac{\Delta t_h}{2} + \{ \text{two sums containing } (\bar{\xi}_{ij}^k)^1_{ab} \text{ and } (\bar{\xi}_{ij}^k)^2_{abc} \} \\ &+ \sum_{h=k}^{N-1} (\bar{a}_{ad}(t_{h+1}, \bar{X}^\Delta(t_{h+1})) - \bar{a}_{ad}(t_h, \bar{X}^\Delta(t_h))) (\bar{\xi}_{ij}^k)^{00}_{a,d}(t_{h+1}) \frac{\Delta t_h}{2} + \{ \text{eight sums containing } (\bar{\xi}_{ij}^k)^{01}_{a,de}, \\ &(\bar{\xi}_{ij}^k)^{02}_{a,def}, (\bar{\xi}_{ij}^k)^{10}_{ab,d}, (\bar{\xi}_{ij}^k)^{11}_{ab,de}, (\bar{\xi}_{ij}^k)^{12}_{ab,def}, (\bar{\xi}_{ij}^k)^{20}_{abc,d}, (\bar{\xi}_{ij}^k)^{21}_{abc,de} \text{ and } (\bar{\xi}_{ij}^k)^{22}_{abc,def} \}, \end{aligned} \tag{4.51}$$

In particular, inserting (4.51) into (4.40), we obtain an expression for (4.40) which is computable a posteriori form. To conclude, we have shown that

$$(\partial_{\theta_\sigma} u)(0, x, (0, 0)) = \bar{u}_\sigma^\Delta(x) + \bar{E}_{\sigma,d}^\Delta(x) + \bar{R}_{\sigma,d}^\Delta, \tag{4.52}$$

where

$$\bar{u}_\sigma^\Delta(x) = \sum_{k=0}^{N-1} E[\bar{A}_{F_\sigma,k}^\Delta(\bar{u}^\Delta) | \bar{X}^\Delta(0) = x] \frac{\Delta t_k}{2} = \sum_{k=0}^{N-1} E[[\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{ij}(t_k, \bar{X}^\Delta(t_k)) \bar{\psi}_{ij}^{(1)}(t_k) | \bar{X}^\Delta(0) = x] \frac{\Delta t_k}{2}, \tag{4.53}$$

$\bar{E}_{\sigma,d}^\Delta(x)$  equals

$$\begin{aligned} &\sum_{k=0}^{N-1} E[2\bar{A}_{F_\sigma,k}^\Delta(\bar{u} - \bar{u}^\Delta) + \bar{B}_{F_\sigma,k}^\Delta(\bar{u}^\Delta) | \bar{X}^\Delta(0) = x] \frac{\Delta t_k}{4} \\ &+ \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} E[E[\bar{C}_{F_\sigma,k,h}^\Delta(\bar{u}^\Delta) | \bar{Z}^\Delta(t_{k+1}) = z(\bar{X}^\Delta(t_{k+1}))] | \bar{X}^\Delta(0) = x] \frac{\Delta t_h \Delta t_k}{4}. \end{aligned} \tag{4.54}$$

and  $\bar{R}_{\sigma,d}^\Delta = \mathcal{O}((\Delta_N^*)^2)$ . As noted above  $\bar{E}_{\sigma,d}^\Delta$  can be written down explicitly in a posteriori form. An analogous representation holds for  $(\partial_{\theta_\mu} u)(0, x, (0, 0))$  as well. Finally replacing the expectations with averages over a finite set of simulations  $\{\omega_m\}_{m=1}^M$  we obtain (1.11). This completes the proof of (1.11)–(1.13).

#### 4.4. Controlling statistical errors

To discuss how to control the statistical errors in the calculation above we first consider a general random variable  $Y$  defined on a probability space  $(\Omega, \mathcal{F}, P)$  and we let  $\{Y(\omega_m)\}_{m=1}^M, \omega_m \in \Omega$ , denote  $M$  independent samples of  $Y$ . Let  $\mathcal{A}(M, Y)$  and  $\mathcal{S}(M, Y)$  denote the sample average and the sample standard deviation, respectively,

$$\mathcal{A}(M, Y) = \frac{1}{M} \sum_{j=1}^M Y(\omega_j), \quad \mathcal{S}(M, Y) = (\mathcal{A}(M, Y^2) - (\mathcal{A}(M, Y))^2)^{1/2}. \tag{4.55}$$

Moreover, let  $\sigma = (E(|Y - E(Y)|^2))^{1/2}$  and assume that  $\lambda = \frac{1}{\sigma} (E(|Y - E(Y)|^3))^{1/3} < \infty$ . Define

$$Z_M = \frac{\mathcal{A}(M, Y) - E(Y)}{\sigma/\sqrt{M}}, \tag{4.56}$$

and let  $F_{Z_M}(z) = P(Z_M \leq z)$  be the cumulative distribution function of  $Z_M$ . Similarly let  $\Phi(z)$ ,  $z \in \mathbb{R}$ , be the cumulative distribution function of a standard normal random variable with zero mean and unit variance. By the Berry–Esséen theorem, see for example Theorem 2.4.10 in [20], we see that

$$\sup_{z \in \mathbb{R}} |F_{Z_M}(z) - \Phi(z)| \leq \frac{3\lambda^3}{\sqrt{M}}. \quad (4.57)$$

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

$$P\left(|\mathcal{E}_S(M, Y)| \leq c_0 \frac{\sigma}{\sqrt{M}}\right) \geq 2\Phi(c_0) - 1 - 2 \sup_{z \in \mathbb{R}} |F_{Z_M}(z) - \Phi(z)|. \quad (4.58)$$

Let  $M = \beta^2 \lambda^6$  where  $\beta \gg 1$  and let  $\zeta$  be defined through the relation  $\Phi(c_0) = \zeta$ . Then, combining the estimates in the last two displays we see that

$$P\left(|\mathcal{E}_S(M, Y)| \leq c_0 \frac{\sigma}{\sqrt{M}}\right) \geq 2\zeta - 1 - 6\beta^{-1}. \quad (4.59)$$

In particular, if we let  $\beta^2 \gg 14400$  and  $c_0 \geq 1.96$  then  $P\left(|\mathcal{E}_S(M, Y)| \leq c_0 \frac{\sigma}{\sqrt{M}}\right) \geq 0.90$ . Finally, using  $\mathcal{J}(M, Y)$  as an approximation of  $\sigma$  we can ensure that

$$|\mathcal{E}_S(M, Y)| \leq E_S(M, Y) := c_0 \frac{\mathcal{J}(M, Y)}{\sqrt{M}} \quad (4.60)$$

with probability close to one. To apply this general theory to the calculation of  $\bar{u}$  we let  $Y(\omega_m) = g(X^\Delta(T, \omega_m))$ , for  $j = 1, \dots, M$ , where  $M$  is sufficiently large. Then, from the discussion above it follows that the statistical error has the upper bound

$$\bar{E}_S^{\Delta, M}(x) \leq \frac{c_0}{\sqrt{M}} \mathcal{J}\left(M, g(X^\Delta(T, \cdot))\right), \quad (4.61)$$

with probability close to one. The same argument can be applied to  $\bar{E}_{\mu, s}^{\Delta, M}(x)$  and  $\bar{E}_{\sigma, s}^{\Delta, M}(x)$ . Moreover, due to (4.22),  $\bar{E}_{d, s}^{\Delta, M}$  can be handled by setting  $Y(\omega_m) = \sum_{k=0}^{N-1} \bar{\rho}_k(\omega_m) (\Delta t_k)^2$ . Then,

$$\bar{E}_{d, s}^{\Delta, M} \leq \frac{c_0}{\sqrt{M}} \mathcal{J}\left(M, \sum_{k=0}^{N-1} \bar{\rho}_k(\cdot) (\Delta t_k)^2\right). \quad (4.62)$$

#### 4.5. An adaptive algorithm to control errors

Firstly, focusing on the calculation of  $\bar{u}$  we now outline, equipped with the error expansions for  $\bar{E}_d^{\Delta, M}$ ,  $\bar{E}_{d, s}^{\Delta, M}$  and  $\bar{E}_s^{\Delta, M}$  in (4.21), (4.62) and (4.61), respectively, the adaptive algorithm for deterministic time steps proposed in [15]. We begin by calculating  $M_{init}$  trajectories of (1.8) using the standard Euler approximation on a uniform mesh of  $N_{init}$  time steps. Using these trajectories, we calculate the errors  $\bar{E}_d^{\Delta, M}$  and  $\bar{E}_{d, s}^{\Delta, M}$  by means of (4.21) and (4.62). If  $\bar{E}_{d, s}^{\Delta, M}$  is larger than some given tolerance, we discard the  $M_{init}$  trajectories and generate  $M \gg M_{init}$  new trajectories on the same mesh and if  $\bar{E}_d^{\Delta, M}$  is larger than some given tolerance, we refine the time steps where the time-discretization error is too large. This process is then iterated until  $\bar{E}_d^{\Delta, M}$  and  $\bar{E}_{d, s}^{\Delta, M}$  are sufficiently small. As the final step, (4.61) is used to estimate the statistical error. If the statistical error exceeds some given tolerance, the  $M$  trajectories are discarded and  $M' \gg M$  new trajectories are generated on the refined mesh. This process is then iterated until the statistical error is sufficiently low. For details we refer to Section 5.3 in [15]. Secondly, focusing on the calculation of  $\partial_{\theta_\mu} u$  and  $\partial_{\theta_\sigma} u$  we see, by analogy, that the same algorithm, together with (1.11)–(1.13), can be used to adaptively control the error in the calculation of the sensitivities.

### 5. A numerical benchmark example

In this section we supply a simple one-dimensional benchmark example for which  $u$  and the sensitivities can be explicitly calculated. This example serves as a stylized illustration of the techniques involved and we use this example, in particular,

to evaluate the performance of the estimators in (1.12). To outline the example we let  $T > 0$  and consider

$$\sigma(t) = \sigma(t, \theta_\sigma) = \bar{\sigma}(t) + \theta_\sigma \tilde{\sigma}(t), \quad \text{where } \bar{\sigma}(t) = \frac{1+t}{10}, \tilde{\sigma}(t) = t^2. \tag{5.1}$$

Then  $\sigma(t)$  satisfies (1.2) whenever  $(t, x) \in [0, T] \times \mathbb{R}$ . Using  $\sigma(t)$ , we let  $X$  solve the stochastic differential equation

$$dX(t) = X(t)dt + \sigma(t)dW(t) \tag{5.2}$$

and the corresponding differential operator is

$$L = \frac{1}{2} (\sigma(t))^2 \partial_{11} + x\partial_1. \tag{5.3}$$

Let

$$u(t, x) = u(t, x, (0, \theta_\sigma)) = E[(X(T))^2 | X(t) = x]. \tag{5.4}$$

Then, using the Feynman–Kac formula, we see that  $u$  solves the Cauchy problem

$$\begin{cases} \partial_t u(t, x) + Lu(t, x) = 0, & \text{whenever } (t, x) \in (0, T) \times \mathbb{R}, \\ u(T, x) = x^2, & \text{whenever } x \in \mathbb{R}. \end{cases} \tag{5.5}$$

Moreover, using Itô calculus, we deduce that  $u(t, x, (0, \theta_\sigma))$  equals

$$\begin{aligned} &x^2 e^{2T} + \frac{1}{400} ((5 + 6t + 2t^2) e^{2(T-t)} - (5 + 6T + 2T^2)) + \theta_\sigma \frac{1}{40} ((5 + 10t + 10t^2 + 4t^3) e^{2(T-t)} \\ &- (5 + 10T + 10T^2 + 4T^3)) + \theta_\sigma^2 \frac{1}{4} ((3 + 6t + 6t^2 + 4t^3 + 2t^4) e^{2(T-t)} - (3 + 6T + 6T^2 + 4T^3 + 2T^4)), \end{aligned} \tag{5.6}$$

and, hence,

$$u(t, x, (0, 0)) = x^2 e^{2T} + \frac{1}{400} ((5 + 6t + 2t^2) e^{2(T-t)} - (5 + 6T + 2T^2)), \tag{5.7}$$

and

$$(\partial_{\theta_\sigma} u)(t, x, (0, 0)) = \frac{1}{40} ((5 + 10t + 10t^2 + 4t^3) e^{2(T-t)} - (5 + 10T + 10T^2 + 4T^3)). \tag{5.8}$$

We intend to demonstrate how the methodology outlined in the previous sections can be used to find  $u(t, x, (0, 0))$ ,  $(\partial_{\theta_\sigma} u)(t, x, (0, 0))$  and the associated time-discretization errors numerically. In the following, we let  $T = 1$  and we consider  $t = 0, x = 1$ . Then,

$$\begin{aligned} u(0, 1, (0, 0)) &= \frac{81}{80} e^2 - \frac{13}{400} \approx 7.44892, \\ (\partial_{\theta_\sigma} u)(0, 1, (0, 0)) &= \frac{5e^2 - 29}{40} \approx 0.19863. \end{aligned} \tag{5.9}$$

Let  $\{t_k\}_{k=0}^N$  define a partition  $\Delta$  of the interval  $[0, 1]$ , i.e.  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = 1$  with  $\Delta t_k = t_{k+1} - t_k$  for  $k \in \{0, \dots, N-1\}$ . Let  $\bar{X}^\Delta$  be the Euler approximation of  $X$  for  $\theta_\sigma = 0$  and note that

$$\bar{X}^\Delta(t_{k+1}) = \bar{X}^\Delta(t_k) (1 + \Delta t_k) + \bar{\sigma}(t_k) \Delta W(t_k), \tag{5.10}$$

with initial condition  $\bar{X}^\Delta(t_0) = 1$ . Furthermore, in this particular case (4.11) reduces to

$$c(t_k, x) = x(1 + \Delta t_k) + \bar{\sigma}(t_k) \Delta W(t_k), \tag{5.11}$$

for all  $x \in \mathbb{R}^n$ . Hence

$$\partial_1 c(t_k, x) = 1 + \Delta t_k, \tag{5.12}$$

while all higher order derivatives of  $c$  with respect to  $x$  are zero. Given the data  $g(x) = x^2$ , the first and second dual functions  $\bar{\psi}, \bar{\psi}^{(1)}$ , associated to  $g(\bar{X}^\Delta(t_N))$ , are recursively defined as follows

$$\begin{aligned} \bar{\psi}(t_N) &= 2\bar{X}^\Delta(t_N), & \bar{\psi}^{(1)}(t_N) &= 2, \\ \bar{\psi}(t_k) &= (1 + \Delta t_k) \bar{\psi}(t_{k+1}), & \bar{\psi}^{(1)}(t_k) &= (1 + \Delta t_k)^2 \bar{\psi}^{(1)}(t_{k+1}). \end{aligned} \tag{5.13}$$

Now, using the deductions in Section 4.2, we see that

$$\bar{u}(0, 1) = \bar{u}^{\Delta, M}(x) + \bar{E}_d^{\Delta, M} + \bar{E}_{d,s}^{\Delta, M} + \bar{E}_s^{\Delta, M} + \mathcal{O}((\Delta_N^*)^2), \tag{5.14}$$

where

$$\bar{u}^{\Delta,M}(x) = \frac{1}{M} \sum_{m=1}^M (\bar{X}^{\Delta}(t_N, \omega_m))^2, \quad (5.15)$$

and the time-discretization error  $\bar{E}_d^{\Delta,M}$  is given by

$$\begin{aligned} \bar{E}_d^{\Delta,M} &= \sum_{m=1}^M \sum_{k=0}^{N-1} [(\bar{X}^{\Delta}(t_{k+1}, \omega_m) - \bar{X}^{\Delta}(t_k, \omega_m)) \bar{\psi}(t_{k+1}, \omega_m)] \frac{\Delta t_k}{2M} \\ &\quad + \sum_{m=1}^M \sum_{k=0}^{N-1} [((\bar{\sigma}(t_{k+1}))^2 - (\bar{\sigma}(t_k))^2) \bar{\psi}^{(1)}(t_{k+1}, \omega_m)] \frac{\Delta t_k}{4M}. \end{aligned} \quad (5.16)$$

Similarly, to find  $\partial_{\theta_\sigma} u$ , we use the results deduced in Section 4.3. In particular, (4.52)–(4.54) reduce to

$$(\partial_{\theta_\sigma} u)(0, 1, (0, 0)) = \bar{u}_{\sigma}^{\Delta,M}(x) + \bar{E}_{\sigma,d}^{\Delta,M} + \bar{E}_{\sigma,d,s}^{\Delta,M} + \bar{E}_s^{\Delta,M} + \mathcal{O}((\Delta_N^*)^2), \quad (5.17)$$

where

$$\bar{u}_{\sigma}^{\Delta,M}(x) = \frac{1}{M} \sum_{m=1}^M \sum_{k=0}^{N-1} \bar{\sigma}(t_k) \tilde{\sigma}(t_k) \bar{\psi}^{(1)}(t_k, \omega_m) \Delta t_k. \quad (5.18)$$

Regarding  $\bar{E}_{\sigma,d}^{\Delta,M}$ , which is the sample mean of  $\bar{E}_{\sigma,d}^{\Delta}$ , we first note the term in (4.54) containing  $\bar{C}_{F_{\sigma,k,h}}^{\Delta}(\bar{u}^{\Delta})$  is zero. Indeed, this easily follows from the definition of the dual functions and that the facts that  $\partial_{\alpha} c = 0$ , for  $|\alpha| \geq 2$ , and  $\partial_{\alpha} g = 0$ , for  $|\alpha| \geq 3$ , imply that the third and fourth order dual functions are identically zero. The term in (4.54) containing  $\bar{B}_{F_{\sigma,k}}^{\Delta}(\bar{u}^{\Delta})$  can be rewritten as

$$\sum_{k=0}^{N-1} E [\bar{\sigma}(t_{k+1}) \tilde{\sigma}(t_{k+1}) \bar{\psi}^{(1)}(t_{k+1}) - \bar{\sigma}(t_k) \tilde{\sigma}(t_k) \bar{\psi}^{(1)}(t_k) | \bar{X}^{\Delta}(0) = 1] \frac{\Delta t_k}{2}. \quad (5.19)$$

Hence it remains to consider the term in (4.54) containing  $\bar{A}_{F_{\sigma,k}}^{\Delta}(\bar{u} - \bar{u}^{\Delta})$  and, as outlined in Section 4.3, we use dual functions for the extended system. However, by the recursive relation (2.12), and the fact that spatial second order derivatives of the drift and diffusion coefficients of the stochastic differential equation (5.2) are zero, we can conclude that the second variation of this system is identically zero. Hence, the extended system reduces to  $\bar{Z}^{\Delta}(t_h) = (\bar{X}^{\Delta}(t_h), \bar{X}^{(1),\Delta}(t_k, t_h))$ , where the Euler approximation  $\bar{X}^{(1),\Delta}$  of the first variation is defined through the recursive relation

$$\bar{X}^{(1),\Delta}(t_k, t_{h+1}) = (1 + \Delta t_h) \bar{X}^{(1),\Delta}(t_k, t_h), \quad \bar{X}^{(1),\Delta}(t_k, t_k) = 1. \quad (5.20)$$

Furthermore, in this case (4.42) and (4.45)–(4.47) reduce to

$$\begin{aligned} \bar{V}^{k,\Delta}(t_{h+1}, \bar{Z}^{\Delta}(t_{h+1})) &= 2 (\bar{X}^{(1),\Delta}(t_k, t_{h+1}))^2, \\ w^0(t_h, x, x^1) &= x(1 + \Delta t_h) + \frac{1+t}{10} \Delta W(t_h), \\ w^1(t_h, x, x^1) &= x^1(1 + \Delta t_h), \end{aligned} \quad (5.21)$$

and  $w^2(t_h, x, x^1) \equiv 0$ . By induction,  $(\bar{\xi}_{ij}^k)^0$  is identically zero and, as a consequence, the only nonzero first order dual function is  $(\bar{\xi}_{ij}^k)^1$  which satisfies,

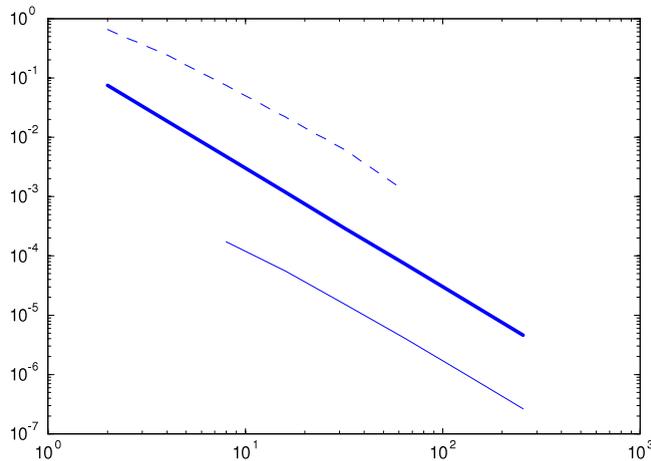
$$\begin{aligned} (\bar{\xi}_{ij}^k)^1(T) &= 4 \bar{X}^{(1),\Delta}(t_k, T), \\ (\bar{\xi}_{ij}^k)^1(t_h) &= (\partial w^1(t_h, \bar{Z}^{\Delta}(t_h)) / \partial x^1) (\bar{\xi}_{ij}^k)^1(t_{h+1}) = (1 + \Delta t_h) (\bar{\xi}_{ij}^k)^1(t_{h+1}), \end{aligned} \quad (5.22)$$

for  $h < N$ . Similarly, by induction it follows that all second order dual functions, except  $(\bar{\xi}_{ij}^k)^{11}$ , vanish. However, since  $\bar{\Gamma}_{lm}^{(h)}$  is nonzero only for coordinates corresponding to  $\bar{X}^{\Delta}(t_h)$ , the dual function  $(\bar{\xi}_{ij}^k)^{11}$  does not contribute to  $\partial_{11}(\bar{u} - \bar{u}^{\Delta})$  and can be omitted. In particular, we obtain

$$\begin{aligned} E [\partial_{11}(\bar{u} - \bar{u}^{\Delta})(t_k, \bar{X}^{\Delta}(t_k)) | \bar{Z}(t_k) = z(\bar{X}^{\Delta}(t_k))] \\ = \sum_{h=k}^{N-1} ((\bar{X}^{(1),\Delta}(t_k, t_{h+1}) - \bar{X}^{(1),\Delta}(t_k, t_h)) (\bar{\xi}_{ij}^k)^1(t_{h+1})) \frac{\Delta t_h}{2}. \end{aligned} \quad (5.23)$$

**Table 1**  
Simulation results for the benchmark example.

$N$	$\bar{u}^{\Delta,M}$	$\bar{E}_d^{\Delta,M}$	$\bar{E}_{d,s}^{\Delta,M}$	$\bar{E}_s^{\Delta,M}$	$\bar{u}_\sigma^\Delta$	$\bar{E}_{\sigma,d}^{\Delta,M}$
2	5.0851	1.7121	$6.099 \cdot 10^{-5}$	$1.324 \cdot 10^{-4}$	0.08438	0.11406
4	5.9958	1.2113	$4.812 \cdot 10^{-5}$	$1.806 \cdot 10^{-4}$	0.13758	0.06066
8	6.6288	0.7440	$3.128 \cdot 10^{-5}$	$2.152 \cdot 10^{-4}$	0.16708	0.03138
16	7.0107	0.4166	$1.803 \cdot 10^{-5}$	$2.365 \cdot 10^{-4}$	0.18259	0.01599
32	7.2209	0.2211	$9.712 \cdot 10^{-6}$	$2.484 \cdot 10^{-4}$	0.19054	0.00807
64	7.3337	0.1140	$5.044 \cdot 10^{-6}$	$2.547 \cdot 10^{-4}$	0.19457	0.00406
128	7.3909	0.0579	$2.571 \cdot 10^{-6}$	$2.579 \cdot 10^{-4}$	0.19660	0.00203
256	7.4200	0.0292	$1.298 \cdot 10^{-6}$	$2.596 \cdot 10^{-4}$	0.19761	0.00102



**Fig. 1.** Plot of the differences  $|(\partial_{\theta_\sigma} u)(0, 1, (0, 0)) - (\bar{u}^{\Delta,M} + \bar{E}_{\sigma,d}^{\Delta,M})|$  (solid, thin) and  $|u(0, 1, (0, 0)) - (\bar{u}^{\Delta,M} + \bar{E}_d^{\Delta,M})|$  (dash) and as a function of the number of time steps for  $M = 10^8$  trajectories. Reference line representing order of convergence equal to two (solid, thick).

and, combining (5.23) with the deductions above, we can conclude that  $\bar{E}_{\sigma,d}^\Delta$  equals

$$\begin{aligned} & \sum_{k=0}^{N-1} E[\bar{\sigma}(t_{k+1}) \tilde{\sigma}(t_{k+1}) \bar{\psi}^{(1)}(t_{k+1}) - \bar{\sigma}(t_k) \tilde{\sigma}(t_k) \bar{\psi}^{(1)}(t_k)] \frac{\Delta t_k}{2}, \\ & + \sum_{k=0}^{N-1} \sum_{h=k}^{N-1} E[\bar{\sigma}(t_k) \tilde{\sigma}(t_k) (\bar{X}^{(1),\Delta}(t_k, t_{h+1}) - \bar{X}^{(1),\Delta}(t_k, t_h)) (\bar{\xi}_{ij}^k)^{-1}(t_{h+1})] \frac{\Delta t_h \Delta t_k}{2}, \end{aligned} \tag{5.24}$$

conditioned on  $\bar{X}^\Delta(0) = 1$ .

We now present the results of our simulation study. In Table 1 we have collected the simulation results for  $\bar{u}^{\Delta,M}$ ,  $\bar{E}_d^{\Delta,M}$ ,  $\bar{E}_{d,s}^{\Delta,M}$ ,  $\bar{E}_s^{\Delta,M}$ ,  $\bar{u}_\sigma^\Delta$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  with  $M = 10^8$  trajectories. The statistical error and statistical time-discretization error for the sensitivity is zero in this particular case as can be deduced from the equations for  $\bar{u}_\sigma^\Delta$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$ . It is clear from Table 1 that the time-discretization errors  $\bar{E}_d^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  have asymptotic orders of convergence close to one. Fig. 1 displays the rest terms  $\bar{R}_{\sigma,d}^\Delta$  and  $\bar{R}_{\sigma,d}^\Delta$  as a function of the number of time steps. A least square approximation shows that, in the limit of vanishing statistical error, these terms are of order  $\mathcal{O}(N^{-\alpha})$ , with  $\alpha \approx 1.8$ . Note, in this context, that the error expansion in this article can be used to construct extrapolation methods with order of convergence close to two for the sensitivity. Moreover, we stress that, in this and many other examples, the approach described in this article is much more efficient than the finite difference method, which suggests that  $\partial_{\theta_\sigma} u(0, 1)$  is approximated by means of

$$\frac{u(0, 1, (0, \theta_\sigma)) - u(0, 1, (0, 0))}{\theta_\sigma}, \tag{5.25}$$

for some  $\theta_\sigma$ . In particular, we conclude from (5.6) that  $\theta_\sigma$  must be very small in order to assert that higher order terms in  $\theta_\sigma$  do not influence the estimate of the sensitivity. Furthermore, as  $\theta_\sigma$  appears in the denominator of (5.25), the statistical error in the calculation of  $u$  is magnified by a factor  $2\theta_\sigma^{-1}$ . Hence, to obtain the same statistical error in the estimate of the sensitivity as in the estimate of  $u$ , we need to multiply the number of trajectories used in the estimate of  $u$  by a factor  $4\theta_\sigma^{-2}$ , which, in most cases, will result in an immense number of trajectories.

## 6. Pricing and hedging of financial derivatives

In this section we illustrate the method outlined in this article in the context of pricing and hedging of interest rate derivatives in LIBOR market models. In particular, we first show how to calculate the value of European swaptions, with control of the errors, using the analysis outlined in Section 4.2. We then illustrate and evaluate (1.11) and the estimators in (1.12) as we perturb the underlying volatility structure. Note that the basic articles on LIBOR market models and swap market models are [21–23]. However, today there exists an extensive literature on the subject and we refer to [24] for a thorough outline of this type of models.

**LIBOR market models.** To outline these models we let  $T_i, i \in \{1, 2, \dots, n+1\}$ , denote a fixed set of  $n+1$  bond maturities with equal spacings  $T_{i+1} - T_i = \delta$  for some  $\delta > 0$ . We let  $L_i(t), i \in \{1, 2, \dots, n\}$ , denote the forward LIBOR rate, contracted at  $0 \leq t \leq T_i$ , for the interval  $[T_i, T_{i+1})$  and we let  $L(t) = (L_1(t), \dots, L_n(t))$ . Furthermore, we let  $\eta(t)$  denote the index of the next maturity date, at time  $t$ , and we note that  $T_{\eta(t)-1} < t \leq T_{\eta(t)}$ . In LIBOR market models the arbitrage free dynamics of the forward rates are given by

$$dL_i(t) = \hat{\mu}_i(t, L(t))L_i(t)dt + (\hat{\sigma}_i(t, L(t)))^*L_i(t)dW(t), \quad (6.1)$$

whenever  $0 \leq t \leq T_i, i \in \{1, \dots, n\}$ , where  $W$  is a standard  $n$ -dimensional Brownian motion under the risk-adjusted measure and

$$\hat{\mu}_i(t, L(t)) = \sum_{j=\eta(t)}^i \frac{(\hat{\sigma}_i(t, L(t)))^* \hat{\sigma}_j(t, L(t)) \delta L_j(t)}{1 + \delta L_j(t)}, \quad (6.2)$$

is the drift under the forward measure. In the following we let  $X_i(t) = \log L_i(t), X(t) = (X_1(t), \dots, X_n(t))$ , and we note that

$$dX_i(t) = \left( \hat{\mu}_i(t, L(t)) - \frac{1}{2} \sum_{j=1}^n (\hat{\sigma}_{ij}(t, L(t)))^2 \right) dt + (\hat{\sigma}_i(t, L(t)))^* dW(t), \quad (6.3)$$

whenever  $0 \leq t \leq T_i, i = 1, \dots, n$ . Moreover, we introduce

$$\begin{aligned} \mu_i(t, X(t)) &= \hat{\mu}_i(t, L(t)) - \frac{1}{2} \sum_{j=1}^n \hat{\sigma}_{ij}(t, L(t))^2, \\ \sigma_i(t, X(t)) &= \hat{\sigma}_i(t, L(t)), \end{aligned} \quad (6.4)$$

and, using the notation in (6.4), we can rewrite (6.3) as

$$dX_i(t) = \mu_i(t, X(t))dt + (\sigma_i(t, X(t)))^* dW(t), \quad (6.5)$$

whenever  $0 \leq t \leq T_i, i \in \{1, \dots, n\}$ . In the following we assume, in order to limit the complexity in the example, that

$$\sigma_i(t, X(t)) = \hat{\sigma}_i(t, L(t)) = \sigma_i(t), \quad \text{for } i \in \{1, \dots, n\}. \quad (6.6)$$

In particular, we assume that the volatility structure is independent of  $X(t)$  (and  $L(t)$ ) but depends on  $t$ . To summarize, we consider

$$dX_i(t) = \mu_i(t, X(t))dt + (\sigma_i(t))^* dW(t), \quad (6.7)$$

whenever  $0 \leq t \leq T_i, i \in \{1, \dots, n\}$ , where the drift coefficient can be specified according to

$$\mu_i(t, X(t)) = \delta \sum_{j=\eta(t)}^i \frac{(\sigma_i(t))^* \sigma_j(t) e^{X_j(t)}}{1 + \delta e^{X_j(t)}} - \frac{1}{2} \sum_{j=1}^n (\sigma_{ij}(t))^2. \quad (6.8)$$

**Perturbations of the volatility structure.** In the example we will perturb the volatility structure as in (1.1). In particular, we let

$$\sigma_i(t) = \sigma_i(t, \theta_\sigma) = \bar{\sigma}_i(t) + \theta_\sigma \tilde{\sigma}_i(t), \quad \text{for } i \in \{1, \dots, n\}, \quad (6.9)$$

where  $\bar{\sigma}, \tilde{\sigma} : \mathbb{R}_+ \rightarrow M(n, \mathbb{R}), \theta_\sigma \in \mathbb{R}$ , and  $|\theta_\sigma| \leq \bar{\epsilon}$ , for some small  $\bar{\epsilon} > 0$ . Combining (6.8) and (6.9) we see that the perturbations in (6.9) give rise to perturbations of  $\mu$ . In particular, inserting (6.9) in (6.8) we see, for  $i \in \{1, \dots, n\}$ , that

$$\mu_i(t) = \mu_i(t, \theta_\mu) = \bar{\mu}_i(t) + \theta_\mu \tilde{\mu}_i(t) + \text{a term of second order in } \theta_\mu, \quad (6.10)$$

where  $\theta_\mu = \theta_\sigma$  and

$$\begin{aligned} \bar{\mu}_i(t) &= \delta \sum_{j=\eta(t)}^i \frac{(\bar{\sigma}_i(t))^* \bar{\sigma}_j(t) e^{X_j(t)}}{1 + \delta e^{X_j(t)}} - \frac{1}{2} \sum_{j=1}^n (\bar{\sigma}_{ij}(t))^2, \\ \tilde{\mu}_i(t) &= \delta \sum_{j=\eta(t)}^i \frac{((\bar{\sigma}_i(t))^* \tilde{\sigma}_j(t) + (\tilde{\sigma}_i(t))^* \bar{\sigma}_j(t)) e^{X_j(t)}}{1 + \delta e^{X_j(t)}} - \sum_{j=1}^n \bar{\sigma}_{ij}(t) \tilde{\sigma}_{ij}(t). \end{aligned} \quad (6.11)$$

**Financial derivatives.** We illustrate the method outlined in this article by applying it to the pricing and hedging of European swaptions. In the following we let

$$B_{m+1}(T_p) = \prod_{i=p}^m \frac{1}{1 + \delta L_i(T_p)}, \quad \text{whenever } 1 \leq p \leq m \leq n, \tag{6.12}$$

be the value at  $T_p$  of a zero-coupon bond maturing at  $T_{m+1}$ . In our setting the forward swap rate at  $T_p$ ,  $p \in \{1, \dots, n\}$ , for a swap with payment dates  $T_{p+1}, \dots, T_{n+1}$  equals

$$S_p(T_p) = \frac{1 - B_{n+1}(T_p)}{\delta \sum_{m=p}^n B_{m+1}(T_p)}. \tag{6.13}$$

We recall that a European (payer) swaption grants the holder the right, expiring at  $T_p$ , to enter into a swap with payment dates  $T_{p+1}, \dots, T_{n+1}$ , where the holder pays the fixed leg and receives the floating leg on a principle of 1. Let  $\chi_p$  denote the pay-off of this option and let  $R$  be the fixed rate specified in the underlying swap. Then

$$\chi_p = \delta \sum_{m=p}^n B_{m+1}(T_p) (S_p(T_p) - R)^+ := F_p \left( \frac{H_p}{F_p} - R \right)^+, \tag{6.14}$$

where

$$F_p := \delta \sum_{m=p}^n B_{m+1}(T_p), \quad H_p := 1 - B_{n+1}(T_p). \tag{6.15}$$

We next note that we, in analogue with most other approaches to the calculation of the ‘Greeks’, somehow have to adjust to the fact that the function  $x^+$  is not differentiable at  $x = 0$ . In particular, as derivatives of the pay-off up to fourth order is required in order to define the dual functions, we have to work with a smooth approximation of  $x^+$  instead of  $x^+$  itself. To proceed we let, for fixed  $\epsilon > 0$ ,  $\phi_\epsilon = \phi_\epsilon(x) : \mathbb{R} \rightarrow \mathbb{R}$  be a sufficiently smooth approximation of the function  $x^+$  such that  $\phi_\epsilon(x) \rightarrow x^+$ , for every  $x \in \mathbb{R}$ , as  $\epsilon \rightarrow 0$ . As we, in order to determine the dual functions, need to calculate derivatives of  $\phi_\epsilon$  up to fourth order one possible choice for  $\phi_\epsilon$  is the four times continuously differentiable piecewise polynomial

$$\phi_\epsilon(x) = \frac{1}{256\epsilon^7} (-5x^8 + 28x^6\epsilon^2 - 70x^4\epsilon^4 + 140x^2\epsilon^6 + 128x\epsilon^7 + 35\epsilon^8) \chi_{(-\epsilon, \epsilon)}(x) + x\chi_{[\epsilon, \infty)}(x), \tag{6.16}$$

where  $\chi_I(x)$  denotes the indicator function for the interval  $I \subset \mathbb{R}$ . Recall that  $X(T_p)$  is the vector of solutions to (6.7) at time  $T_p$ . It is clear, by (6.12), (6.15) and the definition of  $X$ , that  $F_p$  and  $H_p$  are completely determined by the values of  $X(T_p)$ . We let

$$g_p(X(T_p)) := F_p \left( \frac{H_p}{F_p} - R \right)^+, \quad g_p^\epsilon(X(T_p)) := F_p \phi_\epsilon \left( \frac{H_p}{F_p} - R \right). \tag{6.17}$$

Then, our ambition is to calculate

$$u_p(0, x) = E[g_p(X(T_p)) | X(0) = x], \tag{6.18}$$

where  $x$  will be specified below, but, for the reasons discussed above, we instead focus on the calculation of

$$u_p^\epsilon(0, x) := E[g_p^\epsilon(X(T_p)) | X(0) = x], \tag{6.19}$$

for small  $\epsilon$ . In particular,  $u_p^\epsilon(0, x)$  is an approximation of the expectation in (6.18). Furthermore, based on (6.9)–(6.11) and (6.19), we see that the approximation of the sensitivity of the expectation in (6.18), with respect to perturbations of the volatility structure, based on (6.19) and at  $(0, x)$ , equals

$$(\partial_{\theta_\mu} u_p^\epsilon)(0, x, (0, 0)) + (\partial_{\theta_\sigma} u_p^\epsilon)(0, x, (0, 0)). \tag{6.20}$$

**Parameterization and model reductions of the LIBOR market model.** We introduce the parameters

$$\begin{aligned} \gamma_i(t) &= \|\sigma_i(t)\| := \sqrt{\sum_{j=1}^n (\sigma_{ij}(t))^2}, \quad \text{whenever } 0 \leq t \leq T_i, i \in \{1, \dots, n\}, \\ \rho_{ij}(t) &= \frac{(\sigma_i(t))^* \sigma_j(t)}{\|\sigma_i(t)\| \|\sigma_j(t)\|}, \quad \text{whenever } 0 \leq t \leq \min\{T_i, T_j\}, i, j \in \{1, \dots, n\}, \end{aligned} \tag{6.21}$$

where  $\gamma_i(t)$  and  $\rho_{ij}(t)$  represent, respectively, the instantaneous volatility of  $\log L_i(t)$  and the instantaneous correlation between  $\log L_i(t)$  and  $\log L_j(t)$ . Using this notation it is clear that we have to specify  $\gamma_i(t)$  and  $\rho_{ij}(t)$  in order to specify the

model. In reality the calibration of these quantities is a non-trivial problem, e.g. see [25], and there are many suggested approaches to the reduction of the effective number of parameters. One frequently used parameterization of the LIBOR market model is

$$\begin{aligned}\gamma_i(t) &= c_i h(T_i - t), \quad \text{whenever } 0 \leq t \leq T_i, i \in \{1, \dots, n\}, \\ \rho_{ij}(t) &= \rho(T_i - t, T_j - t), \quad \text{whenever } 0 \leq t \leq \min\{T_i, T_j\}, i, j \in \{1, \dots, n\},\end{aligned}\quad (6.22)$$

where  $h$  is a positive real valued function,  $\{c_i\}$  are positive real numbers and  $\rho : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow M(n, \mathbb{R})$  satisfies the characteristics of a correlation matrix. This model reduction is thoroughly described in [25]. In the following we let  $\rho$  be a piecewise constant function with the property

$$\rho(T_i - t, T_j - t) = \widehat{\rho}_{i-\eta(t), j-\eta(t)}, \quad \text{whenever } 0 \leq t \leq \min\{T_i, T_j\}, i, j \in \{1, \dots, n\}, \quad (6.23)$$

and where the matrix  $\widehat{\rho} \in M(n, \mathbb{R})$  can be decomposed

$$\widehat{\rho}_{i,j} = e_i^* e_j, \quad (6.24)$$

for  $i, j \in \{1, \dots, n\}$ , for some set of (constant) unit vectors  $\{e_i\}_{i=1}^n$  in  $\mathbb{R}^n$ . Using (6.22)–(6.24), we see that

$$\sigma_i(t) = d_i h(T_i - t) e_{i-\eta(t)}, \quad \text{whenever } 0 \leq t \leq T_i, i \in \{1, \dots, n\} \quad (6.25)$$

for some positive real numbers  $\{d_i\}$ . Hence it remains to fix the function  $h$  and one possible form of  $h(t)$ , suggested in [26], is

$$h(t) = h^{a,b,c}(t) = c + (1 - c + at)e^{-bt}, \quad \text{where } a, b, c > 0. \quad (6.26)$$

This class of functions is, according to [26], designed to replicate volatility structures that are often observed in real markets. Finally, to complete the model we have to specify the parameters to be used in the numerical simulation below. In particular, we let

$$n = 21, \quad \delta = 1/2, \quad T_p = 5, \quad T = 10, \quad R = 0.04,$$

and we consider  $x = \log L(0)$  where<sup>1</sup>

$$\begin{aligned}L(0) &= (0.03691, 0.04425, 0.04984, 0.05292, 0.05268, 0.05426, 0.05565, \dots \\ &\dots, 0.05664, 0.05825, 0.05920, 0.05965, 0.06029, 0.05981, 0.05964, \dots \\ &\dots, 0.06062, 0.06048, 0.06278, 0.06238, 0.06021, 0.06004, 0.06136).\end{aligned}\quad (6.27)$$

Furthermore, concerning the unperturbed model volatility structure  $\bar{\sigma}$  we use parameters calibrated by means of the methods in [25]. In particular, we let

$$\bar{\sigma}_i(t) = d_i (c + (1 - c + a(T_i - t)) e^{-b(T_i - t)}) e_{i-\eta(t)}, \quad (6.28)$$

where

$$a = 0, \quad b = \frac{1}{\delta}, \quad c = 0.9, \quad d_i = 0.15, \quad \text{for } i \in \{1, \dots, n\}. \quad (6.29)$$

Finally, the unit vector  $e_{i-\eta(t)}$  in (6.28) is determined as in (6.24) based on the matrix  $\{\widehat{\rho}_{i,j}\}_{i,j=1}^n$  where

$$\widehat{\rho}_{i,j} = \exp\left(\frac{|i-j|}{n-1} \left(\log 0.13 - 1.76 \frac{i^2 + j^2 + ij + 3(i+j)(1-n) + 2n^2 - n - 4}{(n-2)(n-3)}\right)\right). \quad (6.30)$$

This specification can be found in [25]. To specify the perturbations, we let, for  $\nu \in \{1, \dots, n\}$ , a perturbation  $\bar{\sigma}^{(\nu)}(t)$  of the volatility structure be given by

$$\bar{\sigma}_i^{(\nu)}(t) = d_i (c + (1 - c + a(T_i - t)) e^{-b(T_i - t)}) e_{i-\eta(t)} \delta_{i\nu}, \quad (6.31)$$

where  $\delta_{i\nu}$  is the Kronecker delta and  $a, b, c, \{d_i\}$  are defined in (6.29). This perturbation corresponds to multiplicative noise in the  $\nu$ -th coordinate of  $\bar{\sigma}(t)$ .

**Discrete dual functions.** Let  $p \in \{1, \dots, n\}$  be given, let  $i, j \in \{1, \dots, n\}$  and let  $\{t_k\}_{k=0}^N$  define a partition  $\Delta$  of the interval  $[0, T_p]$ , i.e.  $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T_p$  with  $\Delta t_k = t_{k+1} - t_k$  for  $k \in \{0, \dots, N-1\}$ . Let  $\bar{X}^\Delta$  be the Euler approximation of  $\bar{X}$ , i.e.,

$$\begin{aligned}\bar{X}_i^\Delta(t_{k+1}) &= \bar{X}_i^\Delta(t_k) + \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k)) \Delta t_k + (\bar{\sigma}_i(t_k))^* \Delta W(t_k), \\ \bar{X}_i^\Delta(0) &= x = \log L(0),\end{aligned}\quad (6.32)$$

<sup>1</sup>  $L(0)$  are the Libor market rates as of the 14th of May 2002.

for  $i \in \{1, \dots, n\}$ . Next we note that in this case (4.11) becomes

$$c_i(t_k, x) = x_i + \bar{\mu}_i(t_k, x) \Delta t_k + \bar{\sigma}_{i\beta}(t_k) \Delta W_\beta(t_k), \quad \text{whenever } x \in \mathbb{R}^n. \tag{6.33}$$

Hence

$$\begin{aligned} \partial_\beta c_i(t_k, x) &= \delta_{\beta i} + \partial_\beta \bar{\mu}_i(t_k, x) \Delta t_k, \\ \partial_{\beta\gamma} c_i(t_k, x) &= \partial_{\beta\gamma} \bar{\mu}_i(t_k, x) \Delta t_k, \end{aligned} \tag{6.34}$$

with similar expressions for the third and fourth order derivatives of  $c_i$ . Furthermore

$$\partial_\beta \bar{\mu}_i(t_k, x) = \delta \frac{(\bar{\sigma}_i(t_k))^* \bar{\sigma}_\beta(t_k) e^{x\beta}}{(1 + \delta e^{x\beta})^2}, \quad \text{for } \eta(t_k) \leq \beta \leq i, \tag{6.35}$$

and

$$\partial_{\beta\gamma} \bar{\mu}_i(t_k, x) = \delta \frac{(\bar{\sigma}_i(t_k))^* \bar{\sigma}_\beta(t_k) e^{x\beta} (1 - \delta e^{x\beta})}{(1 + \delta e^{x\beta})^3} \delta_{\beta\gamma}, \quad \text{for } \eta(t_k) \leq \beta, \gamma \leq i. \tag{6.36}$$

Moreover, the third and fourth order derivatives of  $\bar{\mu}$  can also be derived in a similar fashion. In particular, we note that due to the presence of the Kronecker delta  $\delta_{\beta\gamma}$  in (6.36), and similarly for the higher order derivatives, it follows that all mixed derivatives of  $\bar{\mu}_i$  (and hence of  $c_i$ ), of order two and higher, are zero and this fact reduces the computational complexity considerably.

We are now ready to define the dual functions. In particular, given the data  $g_p^\epsilon$ , the first, second and third dual functions  $\bar{\psi}$ ,  $\bar{\psi}^{(1)}$  and  $\bar{\psi}^{(2)}$ , associated to  $g_p^\epsilon(\bar{X}^\Delta(t_N))$ , are recursively defined as follows. To start with, we have

$$\bar{\psi}_i(t_N) = \partial_i g_p^\epsilon(\bar{X}^\Delta(t_N)), \quad \bar{\psi}_{ij}^{(1)}(t_N) = \partial_{ij} g_p^\epsilon(\bar{X}^\Delta(t_N)), \quad \bar{\psi}_{ijl}^{(2)}(t_N) = \partial_{ijl} g_p^\epsilon(\bar{X}^\Delta(t_N)). \tag{6.37}$$

Then for  $k < N$ ,

$$\begin{aligned} \bar{\psi}_i(t_k) &= (\delta_{i\beta} + \partial_i \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_\beta(t_{k+1}), \\ \bar{\psi}_{ij}^{(1)}(t_k) &= (\delta_{i\beta} + \partial_i \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) (\delta_{j\gamma} + \partial_j \bar{\mu}_\gamma(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) + \delta_{ij} \partial_{ii} \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k \bar{\psi}_\beta(t_{k+1}), \\ \bar{\psi}_{ijl}^{(2)}(t_k) &= (\delta_{i\beta} + \partial_i \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) (\delta_{j\gamma} + \partial_j \bar{\mu}_\gamma(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \cdot (\delta_{l\varphi} + \partial_l \bar{\mu}_\varphi(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_{\beta\gamma\varphi}^{(2)}(t_{k+1}) \\ &\quad + \delta_{ij} \partial_{ii} \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k (\delta_{j\gamma} + \partial_j \bar{\mu}_\gamma(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) \\ &\quad + \delta_{ji} \partial_{jj} \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k (\delta_{i\gamma} + \partial_i \bar{\mu}_\gamma(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) \\ &\quad + \delta_{ii} \partial_{ii} \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k (\delta_{j\gamma} + \partial_j \bar{\mu}_\gamma(t_k, \bar{X}^\Delta(t_k)) \Delta t_k) \bar{\psi}_{\beta\gamma}^{(1)}(t_{k+1}) \\ &\quad + \delta_{ij} \partial_{iii} \bar{\mu}_\beta(t_k, \bar{X}^\Delta(t_k)) \Delta t_k \bar{\psi}_\beta(t_{k+1}). \end{aligned} \tag{6.38}$$

The fourth order dual function  $\bar{\psi}^{(3)}$  is defined similarly but we omit further details. The calculation of the dual functions requires that derivatives of  $g_p^\epsilon$  up to fourth order are determined and we note that these derivatives can be calculated explicitly from (6.15)–(6.17).

**Calculation of  $\bar{u}$ .** Using the deductions in Section 4.2, we have that

$$u_p^\epsilon(0, x) = \bar{u}_p^{\Delta, M}(x) + \bar{E}_d^{\Delta, M} + \bar{E}_{d,s}^{\Delta, M} + \bar{E}_s^{\Delta, M} + \bar{R}_d^\Delta, \tag{6.39}$$

where

$$\bar{u}_p^{\Delta, M}(x) = \frac{1}{M} \sum_{m=1}^M g_p^\epsilon(\bar{X}^\Delta(t_N, \omega_m)), \tag{6.40}$$

and the time-discretization error is given by

$$\begin{aligned} \bar{E}_d^{\Delta, M} &= \sum_{m=1}^M \sum_{k=0}^{N-1} [ [\bar{\mu}_i(t_{k+1}, \bar{X}^\Delta(t_{k+1}, \omega_m)) - \bar{\mu}_i(t_k, \bar{X}^\Delta(t_k, \omega_m))] \bar{\psi}_i(t_{k+1}, \omega_m) ] \frac{\Delta t_k}{2M} \\ &\quad + \sum_{k=0}^{N-1} \sum_{m=1}^M [ [\bar{a}_{ij}(t_{k+1}) - \bar{a}_{ij}(t_k)] \bar{\psi}_{ij}^{(1)}(t_{k+1}, \omega_m) ] \frac{\Delta t_k}{2M}. \end{aligned} \tag{6.41}$$

Using (6.39)–(6.41) and the method for controlling the statistical error described in Section 4.4, it is straightforward to construct an adaptive algorithm, as outlined in [15] and described in Section 4.5, to calculate  $u_p^\epsilon(0, x)$  such that the error with high probability is within a predefined tolerance. In this example we omit the details and instead we focus on the problem of calculating the sensitivities.

**Discrete dual functions for the extended system.** To explicitly calculate the time-discretization error, arising in the calculation of the sensitivities below, we must, as outlined in Section 4.3, determine the first and second order dual functions of the extended system  $\bar{Z}^\Delta(t_h) = (\bar{X}^\Delta(t_h), \bar{X}^{(1),\Delta}(t_k, t_h), \bar{X}^{(2),\Delta}(t_k, t_h))$  associated with the functional

$$\bar{V}_{ij}^{k,\Delta}(t_N, \bar{Z}^\Delta(t_N)) = \partial_{\beta g^\epsilon}(\bar{X}^\Delta(t_N)) \bar{X}_{\beta ij}^{(2),\Delta}(t_k, t_N) + \partial_{\beta \gamma g^\epsilon}(\bar{X}^\Delta(t_N)) \bar{X}_{\beta i}^{(1),\Delta}(t_k, t_N) \bar{X}_{\gamma j}^{(1),\Delta}(t_k, t_N). \tag{6.42}$$

We here note that in this case the Euler approximations of the first and second variation processes are defined recursively as

$$\bar{X}_{uv}^{(1),\Delta}(t_k, t_{h+1}) = \bar{X}_{uv}^{(1),\Delta}(t_k, t_h) + \partial_{\beta \bar{\mu}_u}(t_h, \bar{X}^\Delta(t_h)) \bar{X}_{\beta v}^{(1),\Delta}(t_k, t_h) \Delta t_h, \tag{6.43}$$

with initial condition  $X_{uv}^{(1),\Delta}(t_k, t_k) = \delta_{uv}$  and

$$\begin{aligned} \bar{X}_{uvw}^{(2),\Delta}(t_k, t_{h+1}) &= \bar{X}_{uvw}^{(2),\Delta}(t_k, t_h) + \partial_{\beta \bar{\mu}_u}(t_h, \bar{X}^\Delta(t_h)) \bar{X}_{\beta vw}^{(2),\Delta}(t_k, t_h) \Delta t_h \\ &\quad + \partial_{\beta \beta \bar{\mu}_u}(t_h, \bar{X}^\Delta(t_h)) \bar{X}_{\beta v}^{(1),\Delta}(t_k, t_h) \bar{X}_{\beta w}^{(1),\Delta}(t_k, t_h) \Delta t_h, \end{aligned} \tag{6.44}$$

with initial condition  $\bar{X}_{uvw}^{(2),\Delta}(t_k, t_k) = 0$ . To perform the explicit calculation of the dual functions for the extended system associated to the numerical example at hand we first note that in this case, (4.45)–(4.47) can be written

$$w_r^0(t_h, x, x^1, x^2) = x_r + \bar{\mu}_r(t_h, x) \Delta t_h + \bar{\sigma}_{r\beta}(t_h) \Delta W_\beta(t_h), \tag{6.45}$$

$$w_{rs}^1(t_h, x, x^1, x^2) = x_{rs}^1 + \partial_{\beta \bar{\mu}_r}(t_h, x) x_{\beta s}^1 \Delta t_h, \tag{6.46}$$

$$w_{rst}^2(t_h, x, x^1, x^2) = x_{rst}^2 + \partial_{\beta \bar{\mu}_r}(t_h, x) x_{\beta st}^2 \Delta t_h + \partial_{\beta \beta \bar{\mu}_r}^2(t_h, x) x_{\beta s}^1 x_{\beta t}^1 \Delta t_h. \tag{6.47}$$

We emphasize that in the following we consider  $i, j$  and  $k$  as fixed and use the Einstein summation convention (see Remark 2.1) for the variables  $r, s, t, u, v, w, \beta$  and  $\gamma$ .

**Discrete dual functions of first order for the extended system.** Using (6.45)–(6.47) in (4.48) we first see that

$$\begin{aligned} (\bar{\xi}_{ij}^k)_a^0(t_N) &= \partial_{a\beta g^\epsilon}(\bar{X}^\Delta(t_N)) \bar{X}_{\beta ij}^{(2),\Delta}(t_N) + \partial_{a\beta \gamma g^\epsilon}(\bar{X}^\Delta(t_N)) \bar{X}_{\beta i}^{(1),\Delta}(t_N) \bar{X}_{\gamma j}^{(1),\Delta}(t_N), \\ (\bar{\xi}_{ij}^k)_a^0(t_h) &= (\delta_{ar} + \partial_a \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h) (\bar{\xi}_{ij}^k)_r^0(t_{h+1}) \\ &\quad + \partial_{aa} \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h (\bar{X}_{ai}^{(1),\Delta}(t_h) (\bar{\xi}_{ij}^k)_{ri}^1(t_{h+1}) + \bar{X}_{aj}^{(1),\Delta}(t_h) (\bar{\xi}_{ij}^k)_{rj}^1(t_{h+1})) \\ &\quad + \partial_{aaa} \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h \bar{X}_{aij}^{(2),\Delta}(t_h) (\bar{\xi}_{ij}^k)_{rj}^2(t_{h+1}) \\ &\quad + \partial_{aaa} \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h \bar{X}_{ai}^{(1),\Delta}(t_h) \bar{X}_{aj}^{(1),\Delta}(t_h) (\bar{\xi}_{ij}^k)_{rj}^2(t_{h+1}). \end{aligned} \tag{6.48}$$

Furthermore, the dual function  $(\bar{\xi}_{ij}^k)_{ab}^1$  is zero unless  $b = i$  or  $b = j$ . For  $b = i$  we get

$$\begin{aligned} (\bar{\xi}_{ij}^k)_{ai}^1(t_N) &= \partial_{a\beta g^\epsilon}(\bar{X}^\Delta(t_N)) \bar{X}_{\beta j}^{(1),\Delta}(t_N), \\ (\bar{\xi}_{ij}^k)_{ai}^1(t_h) &= (\delta_{ar} + \partial_a \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h) (\bar{\xi}_{ij}^k)_{ri}^1(t_{h+1}) + \partial_{aa} \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h \bar{X}_{aj}^{(1),\Delta}(t_h) (\bar{\xi}_{ij}^k)_{rj}^2, \end{aligned} \tag{6.49}$$

and similarly for  $(\bar{\xi}_{ij}^k)_{aj}^1$ . Finally, we conclude, by induction, that  $(\bar{\xi}_{ij}^k)_{abc}^2$  is nonzero only if  $b = i$  and  $c = j$ . Moreover,  $(\bar{\xi}_{ij}^k)_{aj}^2$  satisfies

$$\begin{aligned} (\bar{\xi}_{ij}^k)_{ajj}^2(t_N) &= \partial_a g^\epsilon(\bar{X}^\Delta(t_N)), \\ (\bar{\xi}_{ij}^k)_{ajj}^2(t_h) &= (\delta_{ar} + \partial_a \bar{\mu}_r(t_h, \bar{X}^\Delta(t_h)) \Delta t_h) (\bar{\xi}_{ij}^k)_{rjj}^2(t_{h+1}). \end{aligned} \tag{6.50}$$

**Discrete dual functions of second order for the extended system.** To calculate the dual functions of second order we use (6.45)–(6.47) in (4.50). However, for the sake of brevity, we refer the reader to the appendix, Section 8, in [16] for the explicit calculations. Based on the calculations in Section 8 in [16] we conclude, see [16], that there are, for every choice of  $(i, j)$ , a total of  $4n + 9n^2$  nonzero distinct dual functions to be calculated for the extended system. In particular, the complexity is much lower compared to the upper theoretical bound, see (4.44) and the subsequent discussion, on the number of dual functions which is of the order  $n^6$ .

**Calculation of sensitivities.** We here calculate the sensitivity in (6.20) for the perturbations  $\tilde{\sigma} = \tilde{\sigma}^v$  in (6.31). We will accomplish this by means of the results established in Section 4.3. In particular, applying (4.52)–(4.54) in our case we see that

$$\begin{aligned} (\partial_{\theta_\mu} u_p^\epsilon)(x, (0, 0)) &= \bar{u}_\mu^{\Delta, M}(x) + \bar{E}_{\mu, d}^{\Delta, M} + \bar{E}_{\mu, d, s}^{\Delta, M} + \bar{E}_{\mu, s}^{\Delta, M} + \mathcal{O}((\Delta_N^*)^2), \\ (\partial_{\theta_\sigma} u_p^\epsilon)(x, (0, 0)) &= \bar{u}_\sigma^{\Delta, M}(x) + \bar{E}_{\sigma, d}^{\Delta, M} + \bar{E}_{\sigma, d, s}^{\Delta, M} + \bar{E}_{\sigma, s}^{\Delta, M} + \mathcal{O}((\Delta_N^*)^2), \end{aligned} \tag{6.51}$$

where

$$\begin{aligned} \bar{u}_{\mu}^{\Delta,M}(x) &= \sum_{m=1}^M \sum_{k=0}^{N-1} \tilde{\mu}_i(t_k, \bar{X}^{\Delta}(t_k, \omega_m)) \bar{\psi}_i(t_k, \omega_m) \frac{\Delta t_k}{M}, \\ \bar{u}_{\sigma}^{\Delta,M}(x) &= \sum_{m=1}^M \sum_{k=0}^{N-1} \left[ [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \bar{\sigma}^*]_{ij}(t_k) \bar{\psi}_{ij}^{(1)}(t_k, \omega_m) \right] \frac{\Delta t_k}{2M}. \end{aligned} \tag{6.52}$$

Furthermore,  $\bar{E}_{\mu,d}^{\Delta,M}$  and  $\bar{E}_{\sigma,d}^{\Delta,M}$  are the naturally defined Monte Carlo estimators associated to the time-discretization errors  $\bar{E}_{\mu,d}^{\Delta}$  and  $\bar{E}_{\sigma,d}^{\Delta}$ . With the notation introduced in Section 4.3,  $\bar{E}_{\mu,d}^{\Delta}$  equals

$$\begin{aligned} &\sum_{k=0}^{N-1} E \left[ 2\bar{A}_{F_{\mu},k}^{\Delta}(\bar{u} - \bar{u}^{\Delta}) + \bar{B}_{F_{\mu},k}^{\Delta}(\bar{u}^{\Delta}) | \bar{X}^{\Delta}(0) = x \right] \frac{\Delta t_k}{4} \\ &+ \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} E \left[ E \left[ \bar{C}_{F_{\mu},k,h}^{\Delta}(\bar{u}^{\Delta}) | \bar{Z}(t_{k+1}) = z(\bar{X}^{\Delta}(t_{k+1})) \right] | \bar{X}^{\Delta}(0) = x \right] \frac{\Delta t_h \Delta t_k}{4}, \end{aligned} \tag{6.53}$$

and  $\bar{E}_{\sigma,d}^{\Delta}$  equals

$$\begin{aligned} &\sum_{k=0}^{N-1} E \left[ 2\bar{A}_{F_{\sigma},k}^{\Delta}(\bar{u} - \bar{u}^{\Delta}) + \bar{B}_{F_{\sigma},k}^{\Delta}(\bar{u}^{\Delta}) | \bar{X}^{\Delta}(0) = x \right] \frac{\Delta t_k}{4} \\ &+ \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} E \left[ E \left[ \bar{C}_{F_{\sigma},k,h}^{\Delta}(\bar{u}^{\Delta}) | \bar{Z}(t_{k+1}) = z(\bar{X}^{\Delta}(t_{k+1})) \right] | \bar{X}^{\Delta}(0) = x \right] \frac{\Delta t_h \Delta t_k}{4}. \end{aligned} \tag{6.54}$$

In Section 4.3 we proved that all terms in  $\bar{E}_{\mu,d}^{\Delta}$  and  $\bar{E}_{\sigma,d}^{\Delta}$  are computable a posteriori form and we here need to derive explicit expressions for these quantities. In the following we again, for the sake of brevity, only supply the details for  $\bar{E}_{\sigma,d}^{\Delta}$ . We consider the terms in (6.54) one at a time. First, the term involving  $\bar{A}_{F_{\sigma},k}^{\Delta}$  equals, due to (4.40),

$$\sum_{k=0}^{N-1} E \left[ [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \bar{\sigma}^*]_{ij}(t_k) \partial_{ij}(\bar{u} - \bar{u}^{\Delta})(t_k, \bar{X}^{\Delta}(t_k)) | \bar{X}^{\Delta}(0) = x \right] \frac{\Delta t_k}{2}. \tag{6.55}$$

Moreover, introducing the notation  $\bar{\mu}_a^h := \bar{\mu}_a(t_h, \bar{X}^{\Delta}(t_h))$ ,  $\partial_r \bar{\mu}_a^h := \partial_r \bar{\mu}_a(t_h, \bar{X}^{\Delta}(t_h))$ ,  $\partial_{rr} \bar{\mu}_a^h := \partial_{rr} \bar{\mu}_a(t_h, \bar{X}^{\Delta}(t_h))$ ,  $\bar{X}_{ri}^{(1),\Delta}(t_h) := \bar{X}_{ri}^{(1),\Delta}(t_k, t_h)$ ,  $\bar{X}_{rij}^{(2),\Delta}(t_h) := \bar{X}_{rij}^{(2),\Delta}(t_k, t_h)$ , we conclude, by (4.51) and the definition of the extended system, that the derivative  $\partial_{ij}(\bar{u} - \bar{u}^{\Delta})(t_k, \bar{X}^{\Delta}(t_k))$  can be rewritten as

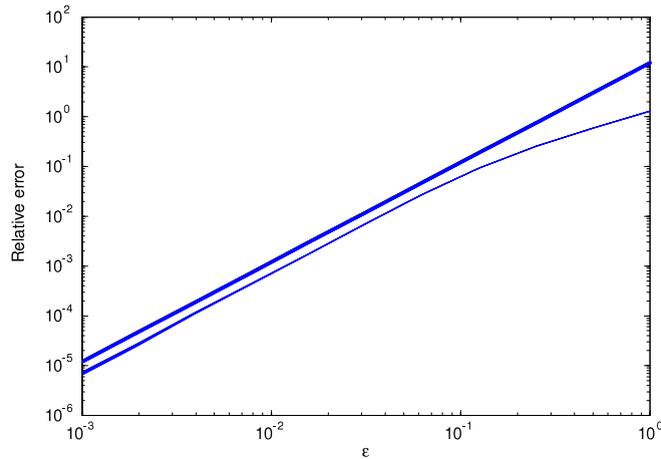
$$\begin{aligned} &\sum_{h=k}^{N-1} \left( \left( \partial_r \bar{\mu}_a^{h+1} \bar{X}_{rij}^{(2),\Delta}(t_{h+1}) - \partial_r \bar{\mu}_a^h \bar{X}_{rij}^{(2),\Delta}(t_h) \right) (\bar{\xi}_{ij}^k)_{aj}^2(t_{h+1}) \right. \\ &+ \left( \partial_{rr} \bar{\mu}_a^{h+1} \bar{X}_{ri}^{(1),\Delta}(t_{h+1}) \bar{X}_{rj}^{(1),\Delta}(t_{h+1}) - \partial_{rr} \bar{\mu}_a^h \bar{X}_{ri}^{(1),\Delta}(t_h) \bar{X}_{rj}^{(1),\Delta}(t_h) \right) (\bar{\xi}_{ij}^k)_{aj}^2(t_{h+1}) \\ &+ \left( \partial_r \bar{\mu}_a^{h+1} \bar{X}_{ri}^{(1),\Delta}(t_{h+1}) - \partial_r \bar{\mu}_a^h \bar{X}_{ri}^{(1),\Delta}(t_h) \right) (\bar{\xi}_{ij}^k)_{ai}^1(t_{h+1}) \\ &+ \left( \partial_r \bar{\mu}_a^{h+1} \bar{X}_{rj}^{(1),\Delta}(t_{h+1}) - \partial_r \bar{\mu}_a^h \bar{X}_{rj}^{(1),\Delta}(t_h) \right) (\bar{\xi}_{ij}^k)_{aj}^1(t_{h+1}) \\ &+ \left( \mu_a^{h+1} - \mu_a^h \right) (\bar{\xi}_{ij}^k)_a^0(t_{h+1}) + \left( \bar{a}_{ad}(t_{h+1}) - \bar{a}_{ad}(t_h) \right) (\bar{\xi}_{ij}^k)_{a,d}^{00}(t_{h+1}) \right) \frac{\Delta t_h}{2}. \end{aligned} \tag{6.56}$$

Note that the conditional expectation with respect to  $\bar{Z}^{\Delta}(t_k) = (\bar{X}^{\Delta}(t_k), I_n, 0)$ , occurring in (4.51), can be removed in this case as the randomness in (6.43)–(6.44) only enters through  $\bar{X}^{\Delta}$ . Next, for term containing  $\bar{B}_{F_{\sigma},k}^{\Delta}(\bar{u}^{\Delta})$  we obtain

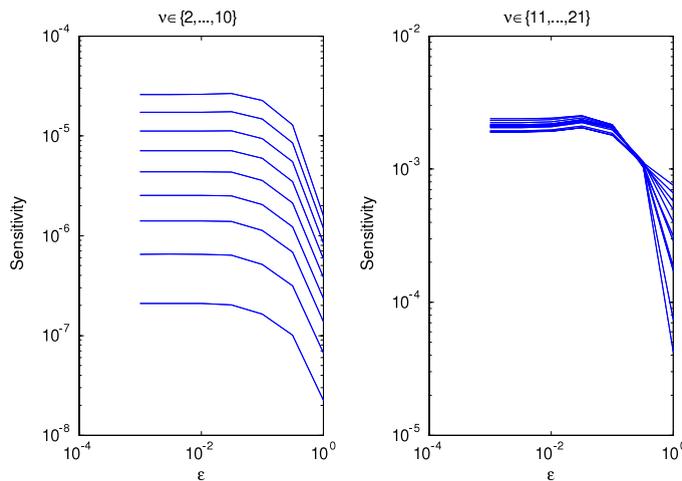
$$\bar{B}_{F_{\sigma},k}^{\Delta}(\bar{u}^{\Delta}) = [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \bar{\sigma}^*]_{ij}(t_{k+1}) \bar{\psi}_{ij}^{(1)}(t_{k+1}) - [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \bar{\sigma}^*]_{ij}(t_k) \bar{\psi}_{ij}^{(1)}(t_k). \tag{6.57}$$

Finally, focusing on the term containing  $\bar{C}_{F_{\sigma},k,h}^{\Delta}$  we first note that the conditional expectation can be removed for the same reason as in (6.56). Furthermore, recalling (4.6) and (4.35), we see that

$$\sum_{h=k+1}^{N-1} E \left[ \bar{C}_{F_{\sigma},k,h}^{\Delta}(\bar{u}^{\Delta}) | \bar{Z}(t_{k+1}) = z(\bar{X}^{\Delta}(t_{k+1})) \right] \Delta t_h$$



**Fig. 2.** Plot of the dependence on  $\epsilon$  of the relative error  $|u_p^\epsilon(0, x) - u_p(0, x)| / u_p(0, x)$  based on  $M = 10^4$  trajectories and  $N = 10$  time steps. Legend: Relative error (solid thin); reference line with slope 2 (solid thick). Note that the gradual widening of the thin line represents the effect of the statistical error.



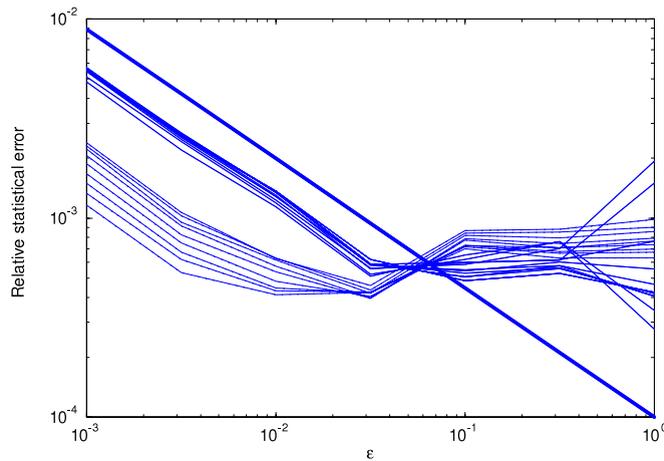
**Fig. 3.** Plot of the sensitivities as a function of the number of time steps based on  $M = 10^6$  trajectories and  $N = 10$  time steps. The left plot corresponds to  $v \in \{2, \dots, 10\}$  (the sensitivities increase with  $v$ ) and the right plot corresponds to  $v \in \{11, \dots, 21\}$ .

$$\begin{aligned}
 &= \sum_{h=k+1}^{N-1} (\bar{\mu}_i^{k+1} - \bar{\mu}_i^k) [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{uv}(t_h) \bar{\psi}_{luv}^{(2)}(t_h) \bar{X}_{li}^{(1),\Delta}(t_{k+1}, t_h) \Delta t_h \\
 &+ \sum_{h=k+1}^{N-1} (\bar{a}_{ij}(t_{k+1}) - \bar{a}_{ij}(t_k)) [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{uv}(t_h) \bar{\psi}_{luv}^{(2)}(t_h) \bar{X}_{ij}^{(2),\Delta}(t_{k+1}, t_h) \Delta t_h \\
 &+ \sum_{h=k+1}^{N-1} (\bar{a}_{ij}(t_{k+1}) - \bar{a}_{ij}(t_k)) [\bar{\sigma} \bar{\sigma}^* + \tilde{\sigma} \tilde{\sigma}^*]_{uv}(t_h) \bar{\psi}_{lmuv}^{(3)}(t_h) \bar{X}_{li}^{(1),\Delta}(t_{k+1}, t_h) \bar{X}_{mj}^{(1),\Delta}(t_{k+1}, t_h) \Delta t_h.
 \end{aligned} \tag{6.58}$$

Combining (6.54)–(6.57), we arrive at a computable expression, in a posteriori form, for the time-discretization error  $\bar{E}_{\sigma,d}^{\Delta,M}$ . Moreover, to handle  $\bar{E}_{\mu,d}^{\Delta,M}$  we can argue similarly.

**Numerical results from the simulations.** In order to be able to estimate the price of European swaptions in LIBOR market models and the corresponding sensitivities with respect to the underlying volatility structure we must first choose an appropriate value of  $\epsilon$ . Fig. 2 shows that the error in the swaption price that is due to  $u_p$  being approximated by  $u_p^\epsilon$  decreases as  $\epsilon^2$  as  $\epsilon$  tends to zero, suggesting that a very small value of  $\epsilon$  should be used.

Moreover, we see from Fig. 3 that the sensitivities initially increase as  $\epsilon$  is decreased and then saturates at some level, implying that as long as we choose  $\epsilon < 10^{-1}$ , the sensitivities will be more or less the same. However, as seen in Fig. 4, the statistical error in the estimate of the sensitivities behaves asymptotically as  $\epsilon^{-0.65}$  and the number of trajectories required to assert that the statistical error is below a given tolerance varies as  $M \propto \epsilon^{-1.3}$ . As we shall see below, the time-discretization



**Fig. 4.** Plot of the dependence on  $\epsilon$  of the relative statistical error of the sensitivities based on  $M = 10^6$  trajectories and  $N = 10$  time steps. Legend: Relative statistical error (solid thin, the lower batch corresponds to  $\nu \in \{2, \dots, 10\}$  and the upper batch to  $\nu \in \{11, \dots, 21\}$ ); reference line with slope  $-0.65$  (solid thick).

**Table 2**

Simulation results for the swaption price in the LIBOR market model example. The upper half of the table corresponds to  $\epsilon = 10^{-1}$  and the lower half to  $\epsilon = 10^{-2}$ .

$N$	$\bar{u}_p^{\epsilon, \Delta, M}$	$\bar{E}_d^{\Delta, M}$	$\bar{E}_{d,s}^{\Delta, M}$	$\bar{E}_s^{\Delta, M}$
10	0.12772	$1.79 \cdot 10^{-4}$	$2.27 \cdot 10^{-8}$	$3.77 \cdot 10^{-5}$
20	0.12750	$8.96 \cdot 10^{-5}$	$1.12 \cdot 10^{-8}$	$3.74 \cdot 10^{-5}$
40	0.12744	$4.49 \cdot 10^{-5}$	$5.53 \cdot 10^{-9}$	$3.72 \cdot 10^{-5}$
80	0.12737	$2.25 \cdot 10^{-5}$	$2.77 \cdot 10^{-9}$	$3.71 \cdot 10^{-5}$
10	0.09845	$1.86 \cdot 10^{-4}$	$8.54 \cdot 10^{-8}$	$3.76 \cdot 10^{-5}$
20	0.09824	$9.32 \cdot 10^{-5}$	$4.27 \cdot 10^{-8}$	$3.75 \cdot 10^{-5}$
40	0.09816	$4.65 \cdot 10^{-5}$	$2.12 \cdot 10^{-8}$	$3.72 \cdot 10^{-5}$
80	0.09811	$2.33 \cdot 10^{-5}$	$1.07 \cdot 10^{-8}$	$3.72 \cdot 10^{-5}$

error for the sensitivities also increases as  $\epsilon$  is decreased and consequently, as far as the sensitivities are concerned, we should not choose  $\epsilon$  too small. Based on this discussion, we have chosen to work with the two cases  $\epsilon = 10^{-1}$  and  $\epsilon = 10^{-2}$  in the numerical simulations below.

With  $\epsilon$  fixed, we first consider the calculation of  $u_p^\epsilon$  and the corresponding time-discretization error and statistical error. Table 2 displays the results of a simulation with  $M = 1.25 \cdot 10^7$  trajectories. As expected the time-discretization errors are of order  $\mathcal{O}(\Delta_N^*)$ . Note also the fourfold increase in statistical discretization error as  $\epsilon$  is decreased from  $10^{-1}$  to  $10^{-2}$ . This is due to the fact that the derivatives  $\partial_\alpha g_p^\epsilon(x)$ , for  $x \in (-\epsilon, \epsilon)$  and  $|\alpha| \geq 2$ , diverge as  $\epsilon \rightarrow 0$ . Extrapolating the values of  $\bar{u}_p^{\epsilon, \Delta, M}$  for the two choices of  $\epsilon$ , we conclude that the true values of  $\bar{u}_p^\epsilon$  are approximately 0.12730 and 0.09803, respectively, and hence we can also conclude that the estimated time-discretization errors,  $\bar{E}_d^{\Delta, M}$ , certainly are of the correct order.

Next, we consider the estimates of the sensitivities. Fig. 5 displays the sensitivities  $(\partial_{\theta_{\sigma(\nu)}} u_p^\epsilon)(0, x, (0, 0))$  for different choices of  $\nu$  and we see that the sensitivities increase rapidly for  $1 \leq \nu < p$  but are fairly constant for  $p \leq \nu \leq n$ . To ensure that the sensitivity estimate in (6.52) gives the correct value, we have performed a finite difference approximation of the sensitivity with respect to  $\tilde{\sigma}^{(21)}$  for  $\epsilon = 10^{-2}$ . In particular, using a finite difference approximation based on  $\theta_{\sigma(21)} = 0$ ,  $\theta_{\sigma(21)} = 0.05$ ,  $M = 1.2 \cdot 10^8$  trajectories and  $N = 10$  time steps, we obtained the approximation

$$(\partial_{\theta_{\sigma(\nu)}} u_p^\epsilon)(0, x, (0, 0)) \approx 0.00196 \pm 0.00048. \tag{6.59}$$

The corresponding estimate based on (6.52) is  $0.001932 \pm 0.000002$  which certainly is of the same order. Note here that, as there is no known way of quantifying the error due to the choice  $\theta_{\sigma(21)} = 0.05$  in the finite difference approximation, the margin of error in (6.59) is only based on the statistical error.

Next we consider the estimates of the time-discretization and statistical time-discretization errors for the sensitivities and, for brevity, we shall only consider the sensitivity corresponding to  $\nu = 21$ . Fig. 6 shows, for  $\epsilon = 10^{-1}$  and  $\epsilon = 10^{-2}$  respectively, the dependence of the terms including  $\bar{A}_{F_{\sigma,k}}^\Delta$ ,  $\bar{B}_{F_{\sigma,k}}^\Delta$  and  $\bar{C}_{F_{\sigma,k,h}}^\Delta$  on the number of time steps. The first two terms are of order  $\mathcal{O}(\Delta_N^*)$  whereas the last term turns out to be of order  $\mathcal{O}((\Delta_N^*)^2)$  and can thus be omitted in this setting. Note also that the error terms containing  $\bar{A}_{F_{\sigma,k}}^\Delta$  and  $\bar{C}_{F_{\sigma,k,h}}^\Delta$  increase with  $\epsilon$  but that the term containing  $\bar{B}_{F_{\sigma,k}}^\Delta$  appears to be independent of  $\epsilon$ . Table 3 displays the results of a simulation with  $M = 2 \cdot 10^6$  trajectories. Note that since the estimates

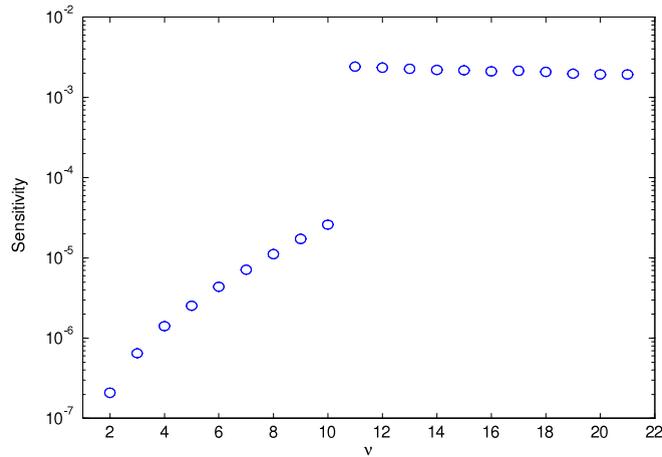


Fig. 5. Plot of the sensitivities as a function of  $\nu$  based on  $M = 2 \cdot 10^6$  trajectories and  $N = 10$  time steps.

Table 3

Simulation results for the sensitivities in the LIBOR market model example. The upper half of the table corresponds to  $\epsilon = 10^{-1}$  and the lower half to  $\epsilon = 10^{-2}$ .

$N$	$\partial_{\theta_{\sigma(2)}} u_p^{\epsilon, \Delta, M}$	$\bar{E}_{\sigma, d}^{\Delta, M}$	$\bar{E}_{\sigma, d, s}^{\Delta, M}$	$\bar{E}_{\sigma, s}^{\Delta, M}$
10	$1.796 \cdot 10^{-3}$	$5.18 \cdot 10^{-5}$	$5.57 \cdot 10^{-8}$	$6.94 \cdot 10^{-7}$
20	$1.788 \cdot 10^{-3}$	$2.39 \cdot 10^{-5}$	$2.65 \cdot 10^{-8}$	$5.85 \cdot 10^{-7}$
40	$1.783 \cdot 10^{-3}$	$1.11 \cdot 10^{-5}$	$1.16 \cdot 10^{-8}$	$6.64 \cdot 10^{-7}$
10	$1.928 \cdot 10^{-3}$	$1.00 \cdot 10^{-3}$	$2.68 \cdot 10^{-6}$	$1.79 \cdot 10^{-6}$
20	$1.927 \cdot 10^{-3}$	$4.18 \cdot 10^{-4}$	$1.18 \cdot 10^{-6}$	$1.74 \cdot 10^{-6}$
40	$1.920 \cdot 10^{-3}$	$1.65 \cdot 10^{-4}$	$5.02 \cdot 10^{-7}$	$1.65 \cdot 10^{-6}$

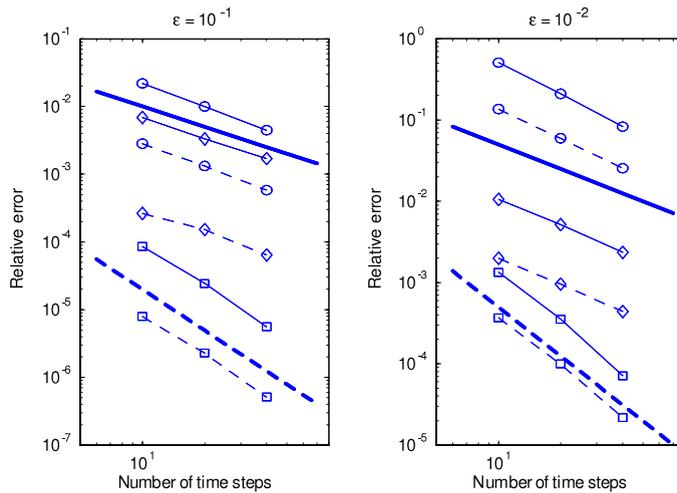


Fig. 6. Plot of the three terms of the time-discretization and statistical time-discretization error for the sensitivities as a function of the number of time steps for  $M = 200$  trajectories. The left plot corresponds to  $\epsilon = 10^{-1}$  and the right plot to  $\epsilon = 10^{-2}$ . Legend: Time-discretization error (solid thin); statistical time-discretization error (dash thin); terms containing  $\bar{A}_{F_{\sigma, k}}^{\Delta}$  (circles); terms containing  $\bar{B}_{F_{\sigma, k}}^{\Delta}$  (diamonds); terms containing  $\bar{C}_{F_{\sigma, k, h}}^{\Delta}$  (squares); reference line with slope  $-1$  (solid thick); reference line with slope  $-2$  (dash thick).

of the time-discretization error turn to infinity as  $\epsilon \rightarrow 0$ , the term  $\bar{E}_{\sigma, d}^{\Delta, M}$  will increase as  $\epsilon$  is decreased. Nevertheless, the time-discretization error is bounded by  $\bar{E}_{\sigma, d}^{\Delta, M}$ .

7. Summary and discussion

Any numerical algorithm (Monte Carlo algorithm) for stochastic differential equations produces a time-discretization error and a statistical error in the process of pricing financial derivatives and calculating the associated ‘Greeks’. In this

article we have shown how a posteriori error estimates and adaptive methods for stochastic differential equations can be used to control both these errors in the context of pricing and hedging of financial derivatives. In particular, we have derived expansions, with leading order terms which are computable in a posteriori form, of the time-discretization errors for the price and the associated ‘Greeks’. These expansions allow the user to simultaneously first control the time-discretization errors in an adaptive fashion, when calculating the price, sensitivities and hedging parameters with respect to a large number of parameters, and then subsequently to ensure that the total errors are, with prescribed probability, within tolerance. Furthermore, we have demonstrated the methodology outlined through two numerical examples.

One point left open in the bulk of the article is a discussion of the importance of the ellipticity condition in (1.2). This condition is used to ensure the appropriate elliptic regularity theory for the operator  $\partial_t + L$ . However, there are many important classes of systems and operators which do not satisfy this ellipticity condition. One important class of operators, relevant in the context of mathematical finance, is the class of second order differential operators of Kolmogorov type of the form  $\partial_t + L$ , where

$$L = \sum_{i,j=1}^m a_{ij}(t, x) \partial_{ij} + \sum_{i=1}^m b_i(t, x) \partial_i + \sum_{i,j=1}^N b_{ij} x_i \partial_j, \quad (7.1)$$

$(t, x) \in \mathbb{R}^{N+1}$ ,  $m$  is a positive integer satisfying  $m < N$ , the functions  $\{a_{ij}(\cdot, \cdot)\}$  and  $\{b_i(\cdot, \cdot)\}$  are continuous and bounded and  $B = \{b_{ij}\}$  is a matrix containing constant real numbers. As  $m < N$  these operators cannot be uniformly elliptic–parabolic but this is compensated for by assuming appropriate regularity on the coefficients and by assuming that the operator

$$\partial_t + K = \partial_t + \sum_{i=1}^m \partial_{ii} + \sum_{i,j=1}^N b_{ij} x_i \partial_j \quad (7.2)$$

is hypoelliptic, i.e., every distributional solution of  $(\partial_t + K)u = f$  is, whenever  $f$  is infinitely smooth, an infinitely smooth solution. Let

$$Y = \sum_{i,j=1}^N b_{ij} x_i \partial_j + \partial_t, \quad (7.3)$$

and let  $\text{Lie}(Y, \partial_1, \dots, \partial_m)$  denote the Lie algebra generated by the vector fields  $Y, \partial_1, \dots, \partial_m$ . Then it is well known that the above assumption of hypoellipticity of  $\partial_t + K$  can be stated in terms of the well-known Hörmander condition [27]:

$$\text{rank Lie}(Y, \partial_1, \dots, \partial_m) = N + 1. \quad (7.4)$$

We comment that, assuming that the coefficients of the operator  $\partial_t + L$  belong to  $C_b^\infty(\mathbb{R}_+ \times \mathbb{R}^N)$ , hypoellipticity of  $\partial_t + L$  is sufficient for the methodology outlined in this article. As a consequence, the methodology is also applicable to operators of Kolmogorov type. For applications where these operators occur we refer to [28–32] and the references in these articles.

As mentioned earlier, the methodology outlined in this article can be used in a variety of situations. In a future article we will investigate several extensions of the example in Section 6, including an extension to the setting of SABR/LIBOR market models [33].

## Acknowledgements

Nyström was partially supported by grant VR-70629701 from the Swedish research council VR. Önskog was financially supported by the Swedish Defense Research Agency, FOI.

## References

- [1] Paul Glasserman, Monte Carlo Methods in Financial Engineering, Springer-Verlag, 2003.
- [2] Nan Chen, Paul Glasserman, Malliavin Greeks without Malliavin calculus, Stochastic Processes and their Applications 117 (2007) 1689–1723.
- [3] Paul Glasserman, Mike Giles, Smoking adjoints; fast Monte Carlo Greeks, Risk Magazine 19 (2006) 88–92.
- [4] Hans-Peter Bermin, Arturo Kohatsu-Higa, Miquel Montero, Local Vega index and variance reduction methods, Mathematical Finance 13 (2003) 85–97.
- [5] Guillaume Bernis, Emmanuel Gobet, Arturo Kohatsu-Higa, Monte Carlo evaluation of Greeks for multidimensional barrier and lookback options, Mathematical Finance 13 (2003) 99–113.
- [6] Bruno Bouchard, Ivar Ekeland, Nizar Touzi, On the Malliavin approach to Monte Carlo approximation of conditional expectations, Finance and Stochastics 8 (1) (2004) 45–71.
- [7] Eric Fournié, Jean-Michel Lasry, Jérôme Lebouchoux, Pierre-Louis Lions, Applications of Malliavin calculus to Monte-Carlo methods in finance. II, Finance and Stochastics 5 (2) (2001) 201–236.
- [8] Eric Fournié, Jean-Michel Lasry, Jérôme Lebouchoux, Pierre-Louis Lions, Nizar Touzi, Applications of Malliavin calculus to Monte Carlo methods in finance, Finance and Stochastics 3 (4) (1999) 391–412.
- [9] Peter W. Glynn, Stochastic approximation for Monte Carlo optimization, in: James R. Wilson, James O. Henriksen and Stephen D. Roberts (Eds.) Proceedings of the 1986 Winter Simulation Conference, 1986, pp. 356–365.
- [10] Peter W. Glynn, Likelihood ratio gradient estimation: an overview, in: Arne Thesen, Hank Grant, W. David Kelton (Eds.) Proceedings of the 1987 Winter Simulation Conference, 1987, pp. 366–375.
- [11] Emmanuel Gobet, Arturo Kohatsu-Higa, Computation of Greeks for barrier and lookback options using Malliavin calculus, Electronic Communications in Probability 8 (2003) 51–62.

- [12] Emmanuel Gobet, Rémi Munos, Sensitivity analysis using Itô–Malliavin calculus and martingales and applications to stochastic optimal control, *SIAM Journal on Control and Optimization* 43 (5) (2005) 1676–1713.
- [13] Arturo Kohatsu-Higa, Miquel Montero, Malliavin calculus in finance, in: *Handbook of Computational and Numerical Methods in Finance*, Birkhäuser, 2004, pp. 111–174.
- [14] Jichuan Yang, Harold J. Kushner, A Monte Carlo method for the sensitivity analysis and parametric optimization of nonlinear stochastic systems, *SIAM Journal on Control and Optimization* 29 (1991) 1216–1249.
- [15] Anders Szepessy, Raúl Tempone, Georgios E. Zouraris, Adaptive weak approximation of stochastic differential equations, *Communications of Pure and Applied Mathematics* 54 (2001) 1169–1214.
- [16] Thomas Önskog, *The Skorohod Problem and Weak Approximation of Stochastic Differential Equations*, Ph.D. Thesis, Umeå University, 2009.
- [17] Avner Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, 1964.
- [18] Ioannis Karatzas, Steven E. Shreve, *Brownian Motion and Stochastic Calculus*, second ed., Springer, 1991.
- [19] David Nualart, *The Malliavin Calculus and Related Topics*, Springer-Verlag, 1995.
- [20] Richard Durrett, *Probability: Theory and Examples*, second ed., Wadsworth Pub. Co, Pacific Grove, CA, 1995.
- [21] Alan Brace, Dariusz Gatarek, Marek Musiela, The market model of interest rate dynamics, *Mathematical Finance* 7 (1997) 127–155.
- [22] Farshid Jamshidian, LIBOR and swap market models and measures, *Finance and Stochastics* 1 (1997) 293–330.
- [23] Kristian R. Miltersen, Klaus Sandmann, Dieter Sondermann, Closed-form solutions for term structure derivatives with lognormal interest rate, *Journal of Finance* 52 (1997) 409–430.
- [24] Damiano Brigo, Fabio Mercurio, *Interest Rate Models – Theory and Practice*, Springer Verlag, 2006.
- [25] John Schoenmakers, *Robust Libor Modelling and Pricing of Derivative Products*, Chapman & Hall, Boca Raton, 2005.
- [26] Riccardo Rebonato, *Volatility and Correlation*, John Wiley & Sons, 1999.
- [27] Lars Hörmander, Hypoelliptic second order differential equations, *Acta Mathematica* 119 (1967) 147–171.
- [28] Marco Di Francesco, Andrea Pascucci, On the complete model with stochastic volatility of Hobson and Rogers, *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences* 460 (2051) (2004) 3327–3338.
- [29] Marco Di Francesco, Andrea Pascucci, Sergio Polidoro, The obstacle problem for a class of hypoelliptic ultraparabolic equations, *Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences* 464 (2008) 155–176.
- [30] Marie Frentz, Kaj Nyström, Andrea Pascucci, Sergio Polidoro, Optimal regularity in the obstacle problem for Kolmogorov operators related to American Asian options, *Mathematische Annalen* 347 (4) (2010) 805–838.
- [31] Ermanno Lanconelli, Andrea Pascucci, Sergio Polidoro, Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance, in: *Nonlinear Problems in Mathematical Physics and Related Topics, II*, in: *International Mathematics Series (New York)*, vol. 2, Kluwer/Plenum, New York, 2002, pp. 243–265.
- [32] Andrea Pascucci, Free boundary and optimal stopping problems for American Asian options, *Finance and Stochastics* 12 (1) (2008) 21–41.
- [33] Riccardo Rebonato, Kenneth McKay, Richard White, *The SABR/LIBOR Market Model. Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, John Wiley & Sons, 2009.