Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Monte Carlo Greeks for financial products via approximative Greenian Kernels

Jörg Kampen, [1] Anastasia Kolodko, [2] and John Schoenmakers [3]

submitted: 8th February 2007

[1] Weierstrass Institute
for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117
Berlin, Germany
E-Mail: kampen@wias-berlin.de

[2] Weierstrass Institute
for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117
Berlin, Germany
E-Mail: kolodko@wias-berlin.de

[3] Weierstrass Institute
for Applied Analysis and Stochastics,
Mohrenstr. 39, 10117
Berlin, Germany
E-Mail: schoenma@wias-berlin.de

No. 1208
Berlin 2007

2000 *Mathematics Subject Classification.* 60H10, 62G07, 65C05.

Key words and phrases. American options, Sensitivities, Monte-Carlo methods, WKB expansions.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we introduce efficient Monte Carlo estimators for the valuation of high-dimensional derivatives and their sensitivities ("Greeks"). These estimators are based on an analytical, usually approximative representation of the underlying density. We study approximative densities obtained by the WKB method. The results are applied in the context of a Libor market model.

1 Introduction

Valuation methods for high-dimensional derivative products are typically based on Monte Carlo simulation of the underlying process. The dynamics of the underlyings are usually given via a (jump-)diffusion SDE. In case of a diffusion SDE, the underlying process may be simulated using an Euler scheme or a (weak) second order scheme e.g. see Kloeden & Platen (1992) or Milstein & Tretyakov (2004). For simulation of jump-diffusions see e.g. Cont & Tankov (2003), and Glasserman & Merener (2003) for simulation of (Libor) interest rate models with jumps. For the evaluation of option sensitivities, Greeks in financial terminology, several works follow a pathwise approach, see Glasserman & Zhao (1999), Piterbarg (2004), Milstein & Schoenmakers (2002). Chen & Glasserman (2006) provide estimators which are connected with Malliavin methods.

In general the cost of computing prices and sensitivities for high-dimensional derivatives can be considerably reduced if one has a procedure for simulating the underlying process directly at the first exercise date. In the ideal case, the density of the underlying process at a fixed point in time is known explictly and an efficient method to sample from it is available. In practice, however, usually non of this is true. In the context of a the Libor interest rate model, Kurbanmuradov, Sabelfeld & Schoenmakers (2002) considered lognormal approximations for the transition density. Among other approaches for speeding up the simulation of a Libor model we mention drift approximations by Hunter, Jäckel, & Joshi (2001), and also Pelsser, Pietersz & van Regenmortel (2004). However, accurate enough lognormal approximations or drift approximations for the underlying process do not always exist (for certain jump-diffusions for example). In this paper we therefore choose for a more general approach and consider the following objectives in order to tackle these issues.

- Construction of a "good" analytical approximation for the density of the underlying process by using (convergent) WKB[1] methods;
- Developing efficient probabilistic representations for the product price and price sensitivities, based on an analytical approximation of the underlying density (e.g. a

[1] The historical origin of the name is the work of **W**entzel, **K**ramers, **B**rioullin in the context of semi-classical solutions of the Schrödinger equation. The meaning of WKB has broadened since; nowadays, it refers to analytic expansions of exponential form.

WKB approximation) and a possibly rougher approximative standard density (e.g. a lognormal density) which is basically used as an importance sampler.

The structure of the paper is as follows. In Section 2 we set up the model class for which we exemplify our methods and specify the financial products (including Bermudan callables) for which prices and sensitivities are to be determined. In Section 3 we introduce probabilistic representations for integral functionals of kernel type and their derivatives. As a particular result we prove that the corresponding estimator for the derivatives has non-exploding variance for sharply peaked kernels in contrast to some existing weighted Monte Carlo schemes. This estimator thus allows for efficient Monte Carlo estimation of option sensitivities, in particular with respect to underlyings (Deltas), even in situations where the densities are sharply peaked (for instance when volatilities are small). The general probabilistic representations introduced in Section 3 are applied to the computation of Deltas for Bermudan callable products in Section 4. Section 5 deals with the WKB-theory of densities of diffusion equations (densities of processes which have continuous paths). A remark about extensions to analytic expansions of densities of Feller processes is included. These extensions will be considered in detail in Kampen (2007). In Section 5.1 we summarize some results concerning pointwise valid WKB-representations of densities obtained in Kampen (2006). Since in practice only finitely many terms of a WKB expansion can be computed, it will be necessary to use a truncated form of the WKB-representation for actual computations. In Section 5.2. we analyze the effect of this truncation error on approximations of solutions of Cauchy problems and their derivatives.The case of non-autonomous diffusion models is discussed in Section 5.3. The results of Sections 2-5 are applied in Section 6 to the Libor market model. In Section 6.1. we compute explicitly the first three coefficients of the WKB representation of the Libor model density. In Section 6.2 we compute prices and Deltas in a case study of European swaptions.

2 Basic setup

Let (X, B) be an underlying asset process in $\mathbb{R}^n_+ \times \mathbb{R}_+$ ($\mathbb{R}_+ := \{x : x > 0\}$) on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[t_0,T]}, P)$, consisting of n risky assets $X = (X^1, ..., X^n)$, and a numeraire B. We assume that the filtration $(\mathcal{F}_t)$ satisfies the usual conditions and that the system (t, X_t) is Markovian with respect to this filtration. Moreover we assume that X has an absolute continuous transition kernel with density $p(t, x, s, y)$, which has derivatives of any order in $0 \le t < s$, $x, y \in \mathbb{R}^n_+$. Further we assume that B_t, $t > 0$, is adapted to $(X_s,\ 0 \le s \le t)$ and is of finite variation. A popular framework for the system (X, B) is, for instance, the class of jump-diffusions (e.g. Cont & Tankov (2003)). For simplicity however we mainly consider ordinary diffusions in the present article, but, note that the main results generally extend to jump processes as well (see Remark 11 for example). We thus consider a system (X, B) given by its dynamics

$$\frac{dX^i}{X^i} = r(t, X)dt + \sum_{j=1}^{n} \sigma^{ij}(t, X)dW^j, \frac{dB}{B} = r(t, X)dt, \quad 1 \le i, j \le n, \tag{2.1}$$

in the (risk-neutral) measure P. In (2.1) $W = (W^1, ..., W^n)^\top$ is an adapted n-dimensional standard Wiener process, where as usual $(\mathcal{F}_t)$ is the P-augmentation of the filtration generated by W. W.l.o.g. we take for Ω the space of all continuous functions $\omega : [0, \infty) \to \mathbb{R}^n$.

Hence a generic $\omega \in \Omega$ under the measure P is a trajectory of the Wiener process It is assumed that the interest rate $r(t, X)$ and the matrix $\sigma(t,x) = \left(\sigma^{ij}(t,x)\right), t \in [t_0, T], x \in \mathbb{R}^n_+$ are such that for all $x_0 \in \mathbb{R}^n_+$, $b_0 > 0$, there exists a unique solution $t \to (X_t, B_t) \in \mathbb{R}^n_+ \times \mathbb{R}_+$ of (2.1) for $t_0 \le t \le T$ satisfying $(X_{t_0}, B_{t_0}) = (x_0, b_0) =: (X^{t_0,x_0}_{t_0}, B^{t_0,x_0,b_0}_{t_0})$, such that all X^i/B are (true) martingales on $[t_0, T]$ under the risk-neutral measure P. Thus, since the number of Brownian motions equals the number of tradables, the price system (X, B) constitutes a complete market (e.g. Karatzas & Shreve (1998)). It is further required that the Markov process X has a transition density $p(t, x, s, y)$ which is differentiable with respect to $x, y \in \mathbb{R}^n_+$, $s, t \in [t_0, T]$, $t > s$, up to any order. To meet all these requirements, it is sufficient to assume that the functions $r(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ are bounded and have bounded derivatives up to any order, and that the volatility matrix $\sigma(t, x)$ is regular with

$$0 < \lambda_1 \le \left| \left(\sigma\sigma^\top\right)(t,x) \right| \le \lambda_2 \tag{2.2}$$

for all (t, x), $t \in [t_0, T]$, $x \in \mathbb{R}^n_+$, and some $0 < \lambda_1 < \lambda_2$ (see for example Bally & Talay (1996)).

Let us take (w.l.o.g.) $b_0 = 1$ and consider contingent claims with pay-off function of the form $f\ (X_\tau)B_\tau$ at some $(\mathcal{F}_\cdot)$-stopping time τ. By completeness such claims are uniquely priced at time t_0 by

$$v(t_0, x_0) = E\ f(X^{t_0,x_0}_\tau)$$

(e.g. Duffie (2001)). For deterministic τ, say $\tau \equiv T$, we have a European claim, and for $t_0 \le t \le T$ its value process can be represented by

$$v_t := v(t, X_t, B_t) := B_t E^{\mathcal{F}_t} f(X_T) = e^{\int_{t_0}^t r(s,X_s)ds} E^{\mathcal{F}_t} f(X_T).$$

Hence the discounted price process $u_t := v_t/B_t$ depends on X only, i.e.,

$$u_t := u(t, X_t) := E^{\mathcal{F}_t} f(X_T) = \int p(t, X_t, T, y) f(y) dy,$$

where

$$u(t,x) = \int p(t, x, T, y) f(y) dy \tag{2.3}$$

is the unique solution of the Cauchy problem

$$\frac{\partial u}{\partial t} + \frac{1}{2} \sum_{i,j=1}^n x^i x^j \left(\sigma\sigma^\top\right)^{ij}(t,x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1}^n x^i r(t,x) \frac{\partial u}{\partial x^i} = 0, \tag{2.4}$$

$$u(T, x) = f(x).$$

The density kernel $p(\cdot, \cdot, T, y)$ is the unique (weak) solution of (2.4) with $p(T, x, T, y) = \delta(x - y)$, where δ is the Dirac-delta function in Schwarz distribution sense.

Of particular importance are Bermudan callable contracts. A Bermudan contract starting at t_0, is specified by a set of exercise dates $\{t_1, t_2, ..., t_\mathcal{J}\}$, where $t_0 < t_1 < ... < t_\mathcal{J} < T$, and corresponding (discounted) pay-off functions $f_i(x)$, $1 \le i \le \mathcal{J}$. According to the contract, the holder has the right to call (once) a cash-flow $f_i(X^{t_0,x_0}_{t_i}) B^{t_0,x_0,1}_{t_i}$ at an exercise date t_i

of his choice. It is well known that the fair value of this contract at time t, $t_0 \le t \le T$, assuming that no exercise took place before t, is given by

$$v(t,x,b) := bu(t,x) := \sup_{\tau \in \mathcal{T}_{i,\mathcal{J}}} bEf(X_\tau^{t,x}) = bEf(X_{\tau_*^{t,x}}^{t,x}), \quad t_{i-1} < t \le t_i, \tag{2.5}$$

where $x = X_t^{t_0,x_0}$, $b = B_t^{t_0,x_0,1}$, $\mathcal{T}_{i,\mathcal{J}}$ the set of stopping times τ taking values in $\{t_i, t_{i+1}, ..., t_{\mathcal{J}}\}$, and $\tau_*^{t,x}$ is an optimal stopping time.

3 Probabilistic representations and their estimators

In this section we consider for a given smooth function $u : \mathbb{R}_+^n \to \mathbb{R}_+$ and a smooth kernel function $p : \mathbb{R}_+^n \times \mathbb{R}_+^n \to \mathbb{R}_+$ (which may or may not be a transition kernel), probabilistic representations for the integral

$$\begin{aligned} I(x) &:= \int p(x,y)u(y)dy, \quad \text{and its gradient} \\ \frac{\partial I}{\partial x}(x) &= \int \frac{\partial}{\partial x} p(x,y)u(y)dy, \quad \text{with} \\ \frac{\partial}{\partial x} &:= \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right). \end{aligned}$$

Let ζ be some random variable with density ϕ on $\mathbb{R}_+^n$, $\phi > 0$. Then, obviously,

$$I(x) = E\ p(x,\zeta) \frac{u(\zeta)}{\phi(\zeta)} \tag{3.6}$$

is a probabilistic representation for (3.6) which may be estimated by the unbiased Monte Carlo estimator

$$\widehat{I}(x) := \frac{1}{M} \sum_{m=1}^{M} p(x, {}_m\zeta) \frac{u({}_m\zeta)}{\phi({}_m\zeta)}, \tag{3.7}$$

where for $m = 1, ..., M$, ${}_m\zeta$ are i.i.d. samples from a distribution with density ϕ. By taking gradients in (3.6) we readily obtain the probabilistic representation

$$\frac{\partial I}{\partial x}(x) = E\ \frac{\partial}{\partial x} p(x,\zeta) \frac{u(\zeta)}{\phi(\zeta)}, \tag{3.8}$$

with corresponding estimator,

$$\widehat{\frac{\partial I}{\partial x}}(x) := \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial x} p(x, {}_m\zeta) \frac{u({}_m\zeta)}{\phi({}_m\zeta)}. \tag{3.9}$$

While as a rule (3.7) is an effective estimator for $I(x)$ for a proper choice of ϕ, unfortunately the gradient estimator (3.9) has a serious drawback: If the kernel $p(x,\cdot)$ is sharply peaked (nearly proportional to a 'delta-function'), its variance may be extremely high. This fact is demonstrated by the following stylistic example of a multi-asset model, which is in order of magnitude realistic though.

Example 1. Consider for fixed $x_0 \in \mathbb{R}^n_+$, parameters $s > 0$, and $\sigma > 0$, the n-dimensional lognormal density

$$p(s, \sigma; x_0, y) := \frac{1}{(2\pi\sigma^2 s)^{n/2}} \prod_{i=1}^{n} \frac{\exp\left[-\frac{1}{2\sigma^2 s} \ln^2 \frac{y^i}{x_0^i}\right]}{y^i}. \tag{3.10}$$

In (3.10) $p(s, \sigma; x_0, \cdot)$ is the density of the random variable

$$(x_0^1 e^{\sigma\sqrt{s}\xi^1}, ..., x_0^n e^{\sigma\sqrt{s}\xi^n}),$$

where ξ^i , $i = 1, ..., d$, are i.i.d. standard normal random variables. Thus, for small s and σ, $p(s, \sigma; x_0, \cdot)$ is peaked ('delta-shaped') around x_0. Let us now take $\phi(\cdot) := p(s, \sigma; x_0, \cdot)$ in (3.6) and (3.8), respectively, and $u \equiv ||x_0||$ (a constant of order x_0 in magnitude). Clearly, estimator (3.7) equals $||x_0||$ almost surely and so has zero variance. However, estimator (3.9) is not deterministic and we have

$$\begin{aligned}
\widehat{\frac{\partial I}{\partial x^j}}(x_0) &:= \frac{1}{M} \sum_{m=1}^{M} \frac{||x_0||}{p(s, \sigma; x_{0,m}\,\zeta)} \frac{\partial}{\partial x^j} p(s, \sigma; x_{0,m}\,\zeta) \\
&= \frac{||x_0||}{M} \sum_{m=1}^{M} \frac{\partial}{\partial x^j} \ln p(s, \sigma; x_{0,m}\,\zeta) \\
&= \frac{||x_0||}{M} \sum_{m=1}^{M} \frac{\ln \frac{{}_m\zeta^j}{x_0^j}}{\sigma^2 s x_0^j} = \frac{||x_0||}{M} \sum_{m=1}^{M} \frac{{}_m\xi^1}{\sigma\sqrt{s} x_0^j}.
\end{aligned}$$

Hence, $E\left[\widehat{\frac{\partial I}{\partial x^j}}(x_0)\right] = 0$ as should be, but,

$$\mathrm{Var}\left[\widehat{\frac{\partial I}{\partial x^j}}(x_0)\right] = \frac{||x_0/x_0^j||^2}{M} \frac{1}{\sigma^2 s} \tag{3.11}$$

which explodes when $\sigma^2 s$ goes to zero!

Remark 2. In Fries & Kampen (2006) estimators (3.7) and (3.9) are used for computing prices and sensitivities of European Libor options, respectively. In their numerical examples they used 50% (rather high) volatility in order to amplify Monte Carlo errors. While, indeed, a larger volatility generally gives rise to a large Monte Carlor error of (3.7), Example 1 shows that the opposite is true for estimator (3.9). For example, 50% volatility in combination with 0.5 yr. maturity corresponds to a (just moderate) variance factor $1/\left(\sigma^2 s\right) = 8.0$ in (3.11), while a more usual Libor volatility, e.g. 14%, and 0.5 y maturity would give a factor 102.0(!).

In the present paper we propose sensitivity estimators which are efficient on a broad time and volatility scale. As a result, the next theorem provides a tool for constructing sensitivity (gradient) estimators with non-exploding variance.

Theorem 3. *Let λ be a reference density on $\mathbb{R}^n$ with $\lambda(z) \neq 0$ for all z (for example, the standard normal density). Let ξ be an $\mathbb{R}^n$-valued random variable with density λ and $g : \mathbb{R}^n_+ \times \mathbb{R}^n \to \mathbb{R}^n_+$ be a smooth map with $|\partial g(x, z)/\partial z| \neq 0$, such that for each $x \in \mathbb{R}^n_+$ the* *random variable $\zeta^x := g(x,\xi)$ has a density $\phi(x,\cdot)$ on $\mathbb{R}^n_+$. Then, we have the probabilistic representation*

$$\frac{\partial I}{\partial x}(x) = E\,\frac{\partial}{\partial x}\frac{p(x,\zeta^x)u(\zeta^x)}{\phi(x,\zeta^x)} = E\,\frac{\partial}{\partial x}\frac{p(x,g(x,\xi))u(g(x,\xi))}{\phi(x,g(x,\xi))}, \tag{3.12}$$

with corresponding Monte Carlo estimator

$$\widehat{\frac{\partial I}{\partial x}}(x) = \frac{1}{M}\sum_{m=1}^{M}\frac{\partial}{\partial x}\frac{p(x,g(x,_m\xi))u(g(x,_m\xi))}{\phi(x,g(x,_m\xi))}. \tag{3.13}$$

Let $|\cdot|$ denote either a vector norm or a compatible matrix norm. Then it holds

$$E\left|\frac{\partial}{\partial x}\frac{p(x,g(x,\xi))u(g(x,\xi))}{\phi(x,g(x,\xi))}\right|^2 \le 2M_4^2\left(M_2^2M_3^2 + 4M_1^2M_5^2 + 4M_1^2M_3^2M_6^2\right), \tag{3.14}$$

hence the second moments of the Monte Carlo samplers for the components of $\partial I/\partial x$ are bounded by the right-hand-side of (3.14), if for fixed $x \in \mathbb{R}^n_+$, there are $\alpha_1,...,\alpha_6 > 1$ with

$$\frac{1}{\alpha_4}+\frac{1}{\alpha_1}+\frac{1}{\alpha_5}=1, \quad \frac{1}{\alpha_4}+\frac{1}{\alpha_2}+\frac{1}{\alpha_3}=1, \quad \frac{1}{\alpha_4}+\frac{1}{\alpha_1}+\frac{1}{\alpha_6}+\frac{1}{\alpha_3}=1,$$

such that,

$$\begin{aligned}
E\,u^{2\alpha_1}(g(x,\xi)) &= \int u^{2\alpha_1}(y)\phi(x,y)dy \le M_1^{2\alpha_1},\\
E\left|\frac{\partial u}{\partial y}(g(x,\xi))\right|^{2\alpha_2} &= \int\left|\frac{\partial u}{\partial y}(y)\right|^{2\alpha_2}\phi(x,y)dy \le M_2^{2\alpha_2},\\
E\left|\frac{\partial g}{\partial x}(x,\xi)\right|^{2\alpha_3} &= \int\left|\frac{\partial g}{\partial x}(x,z)\right|^{2\alpha_3}\lambda(z)dz \le M_3^{2\alpha_3},\\
E\left(\frac{p(x,g(x,\xi))}{\phi(x,g(x,\xi))}\right)^{2\alpha_4} &= \int\left(\frac{p(x,y)}{\phi(x,y)}\right)^{2\alpha_4}\phi(x,y)dy \le M_4^{2\alpha_4},
\end{aligned}$$

$$E\left|\frac{p_x(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_x(x,g(x,\xi))}{\phi(x,g(x,\xi))}\right|^{2\alpha_5} = \int\left|\frac{p_x(x,y)}{p(x,y)} - \frac{\phi_x(x,y)}{\phi(x,y)}\right|^{2\alpha_5}\phi(x,y)dy \le M_5^{2\alpha_5},$$

and

$$E\left|\frac{p_y(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_y(x,g(x,\xi))}{\phi(x,g(x,\xi))}\right|^{2\alpha_6} = \int\left|\frac{p_y(x,y)}{p(x,y)} - \frac{\phi_y(x,y)}{\phi(x,y)}\right|^{2\alpha_6}\phi(x,y)dy \le M_6^{2\alpha_6},$$

with shorthands $p_x := \partial p/\partial x$, etc.

Proof. For any bounded measurable $\psi : \mathbb{R}^n_+ \to \mathbb{R}$, we have

$$\begin{aligned}\int \psi(g(x,z))\lambda(z)dz = E\,\psi(g(x,\xi)) = E\,\psi(\zeta^x) = \int \psi(y)\phi(x,y)dy\\ = \int \psi(g(x,z))\phi(x,g(x,z)) \left|\frac{\partial g(x,z)}{\partial z}\right| dz.\end{aligned}$$

Therefore, the densities ϕ and g are connected via the relationship

$$\phi(x,g(x,z)) \left|\frac{\partial g(x,z)}{\partial z}\right| = \lambda(z). \tag{3.15}$$

By taking the derivative outside the expectation in the right-hand-side of (3.12) and using (3.15) we get (3.12),

$$\begin{aligned}\frac{\partial}{\partial x} E\,\frac{p(x,g(x,\xi))u(g(x,\xi))}{\phi(x,g(x,\xi))} &= \frac{\partial}{\partial x}\int \frac{p(x,g(x,z))u(g(x,z))}{\phi(x,g(x,z))}\lambda(z)dz\\ &= \frac{\partial}{\partial x}\int p(x,g(x,z))u(g(x,z))\left|\frac{\partial g(x,z)}{\partial z}\right| dz\\ &= \frac{\partial}{\partial x}\int p(x,y)u(y)dy = \frac{\partial I}{\partial x}(x).\end{aligned}$$

To prove the moment estimation (3.14), we observe that

$$\begin{aligned}&E\left|\frac{\partial}{\partial x}\frac{p(x,\zeta^x)u(\zeta^x)}{\phi(x,\zeta^x)}\right|^2 = E\left|\frac{\partial}{\partial x}\frac{p(x,g(x,\xi))u(g(x,\xi))}{\phi(x,g(x,\xi))}\right|^2 =\\ &= E\left|u(g(x,\xi))\frac{\partial}{\partial x}\frac{p(x,g(x,\xi))}{\phi(x,g(x,\xi))} + \frac{p(x,g(x,\xi))}{\phi(x,g(x,\xi))}\frac{\partial u}{\partial y}(g(x,\xi))\frac{\partial g}{\partial x}(x,\xi)\right|^2\\ &\le 2E\,\frac{p^2(x,g(x,\xi))}{\phi^2(x,g(x,\xi))}\left|\frac{\partial u}{\partial y}(g(x,\xi))\right|^2\left|\frac{\partial g}{\partial x}(x,\xi)\right|^2\\ &\quad + 2E\,u^2(g(x,\xi))\left|\frac{\partial}{\partial x}\frac{p(x,g(x,\xi))}{\phi(x,g(x,\xi))}\right|^2\\ &=: 2(I) + 2(II).\end{aligned}$$

Then by Hölders inequality,

$$\begin{aligned}(I) &= E\,\frac{p^2(x,g(x,\xi))}{\phi^2(x,g(x,\xi))}\left|\frac{\partial u}{\partial y}(g(x,\xi))\right|^2\left|\frac{\partial g}{\partial x}(x,\xi)\right|^2\\ &\le \sqrt[\alpha_2]{E\left|\frac{\partial u}{\partial y}(g(x,\xi))\right|^{2\alpha_2}}\sqrt[\alpha_3]{E\left|\frac{\partial g}{\partial x}(x,\xi)\right|^{2\alpha_3}}\sqrt[\alpha_4]{E\frac{p^{2\alpha_4}(x,g(x,\xi))}{\phi^{2\alpha_4}(x,g(x,\xi))}}\\ &\le M_2^2 M_3^2 M_4^2.\end{aligned}$$

For (II) we have

$$
\begin{aligned}
(II) &= E\, u^2(g(x,\xi)) \left| \frac{\partial}{\partial x} \frac{p(x,g(x,\xi))}{\phi(x,g(x,\xi))} \right|^2 \\
&= E\, u^2(g(x,\xi)) \frac{p^2(x,g(x,\xi))}{\phi^2(x,g(x,\xi))} \left| \frac{p_x(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_x(x,g(x,\xi))}{\phi(x,g(x,\xi))} \right. \\
&+ \left. \left(\frac{p_y(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_y(x,g(x,\xi))}{\phi(x,g(x,\xi))} \right) \frac{\partial g}{\partial x}(x,\xi) \right|^2 \\
&\leq 2E\, u^2(g(x,\xi)) \frac{p^2(x,g(x,\xi))}{\phi^2(x,g(x,\xi))} \left| \frac{p_x(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_x(x,g(x,\xi))}{\phi(x,g(x,\xi))} \right|^2 \\
&+ 2E\, u^2(g(x,\xi)) \frac{p^2(x,g(x,\xi))}{\phi^2(x,g(x,\xi))} \left| \frac{p_y(x,g(x,\xi))}{p(x,g(x,\xi))} - \frac{\phi_y(x,g(x,\xi))}{\phi(x,g(x,\xi))} \right|^2 \left| \frac{\partial g}{\partial x}(x,\xi) \right|^2,
\end{aligned}
$$

and then again by Hölders inequality,

$$
(II) \leq 2M_1^2 M_4^2 M_5^2 + 2M_1^2 M_3^2 M_4^2 M_6^2.
$$

$\square$

Remark 4. If one takes $\phi(x,y) \equiv \phi(x_0,y)$ estimator (3.13) collapses to (3.9). The delicate bound in Theorem 3 is M_5. Indeed, in Example 1 where $\phi(x,y) \equiv p(x_0,y)$ in fact, M_5 cannot taken to be small when $\sigma^2 s$ is small, i.e. when p is highly peaked around x. In contrast, if for fixed x, $\phi(x,\cdot)$ is approximately proportional to $p(x,\cdot)$ and $\partial \ln \phi(x,\cdot)/\partial x$ $\approx \partial \ln p(x,\cdot)/\partial x$ (both with respect to the weight function $\phi(x,\cdot)$), a small M_5 may exists. Note that for $\phi(\cdot,\cdot)$ exactly proportional to $p(\cdot,\cdot)$, we may take $M_5 = 0$.

Remark 5. It can be shown that Theorem 3 can be extended to probabilistic representations and corresponding estimators for higher order derivatives,

$$
\frac{\partial I}{\partial x^\beta}(x) = E \frac{\partial}{\partial x^\beta} \frac{p(x,\zeta^x)u(\zeta^x)}{\phi(x,\zeta^x)} = E \frac{\partial}{\partial x^\beta} \frac{p(x,g(x,\xi))u(g(x,\xi))}{\phi(x,g(x,\xi))},
$$

with corresponding Monte Carlo estimator

$$
\widehat{\frac{\partial I}{\partial x^\beta}}(x) = \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial x^\beta} \frac{p(x,g(x,_m\xi))u(g(x,_m\xi))}{\phi(x,g(x,_m\xi))}, \tag{3.16}
$$

where $\beta := (\beta_1,\ldots,\beta_n)$, $\beta_i \in \{0,1,2,\ldots\}$ is a multi-index with (formally) $\partial x^\beta = \partial x_1^{\beta_1} \partial x_2^{\beta_2} \cdots \partial x_n^{\beta_n}$. Loosely speaking, the variance of the higher order derivative estimator (3.16) can be bounded from above by an expression like (3.14) involving (i) sufficiently high moments of the derivatives,

$$
y \to \frac{\partial u}{\partial y^\gamma}, \quad \text{and} \quad z \to \frac{\partial g(x,z)}{\partial z^\gamma}
$$

for fixed x, $\gamma \leq \beta$ (component wise), with respect to weight functions $y \to \phi(x,y)$ and $z \to \lambda(z)$, respectively, and, (ii) for fixed x, $L^q(\mathbb{R}_+^n, \phi(x,y)dy)$-norms of

$$
y \to \frac{\partial}{\partial x^\gamma}\left(\frac{\phi(x,y)}{p(x,y)}\right), \quad \text{and} \quad y \to \frac{\partial}{\partial y^\gamma}\left(\frac{\phi(x,y)}{p(x,y)}\right), \quad \gamma \leq \beta,
$$

for q large enough.

Remark 6. In financial applications $I(x)$ is usually the price of a derivative contract considered in dependence of the argument x which may stand for the underlying process or some parameter (vector) which affects the dynamics of the underlying process (e.g. volatilities).

4 Sensitivities for Bermudan options

Theorem 3 may be applied in general for computing sensitivities ("Greeks") of derivative products. For estimator (3.9) the danger of exploding variance is typically the largest when derivatives of prices with respect to underlyings (Deltas, Gammas) are considered. We therefore consider in this section only (first order) derivatives with respect to the underlying process, hence Deltas.

Let $\tau : \Omega \to \mathbb{R}_+$ be a given stopping time with respect to the filtration $(\mathcal{F}_\cdot)$, and define $\tau^{s,x}(\omega) := s + \tau(X^{s,x}_{s+\cdot}(\omega))$ (recall that Ω is the space of continuous trajectories $[0,\infty) \to \mathbb{R}^d$). We now consider the Bermudan contract introduced in Section 2. For fixed t, t^+, $t_0 \le t \le t^+ \le t_1$, $x \in \mathbb{R}^n_+$, we have $\tau^{t,x}_* = \tau^{t^+, X^{t,x}_{t^+}}_*$ since $\tau^{t,x}_* \ge t_1$, and we thus may write

$$
\begin{aligned}
u(t,x) &:= Ef(X^{t,x}_{\tau^{t,x}_*}) = EE^{\mathcal{F}_{t^+}} f(X^{t^+, X^{t,x}_{t^+}}_{\tau_*^{t^+, X^{t,x}_{t^+}}}) \\
&= \int p(t,x,t^+,y) dz Ef(X^{t^+,y}_{\tau^{t^+,z}_*}) \\
&= \int p(t,x,t^+,y) u(t^+,y) dy,
\end{aligned}
$$

by the Chapman-Kolmogorov equation.

For each t, t^+ as above, let $\phi(t,x,t^+,y)$, $g(t,x,t^+,y)$, and reference density $\lambda(t,t^+,z)$ be as in Theorem 3. We then have the probabilistic representation

$$
u(t,x) = E\, \frac{p(t,x,t^+, g(t,x,t^+,\xi))}{\phi(t,x,t^+, g(t,x,t^+,\xi))} f(X^{t^+, g(t,x,t^+,\xi)}_{\tau_*^{t^+, g(t,x,t^+,\xi)}}), \tag{4.17}
$$

with Monte Carlo estimator

$$
\widehat{u}(t,x) := \frac{1}{M} \sum_{m=1}^{M} \frac{p(t,x,t^+, g(t,x,t^+, {}_m\xi))}{\phi(t,x,t^+, g(t,x,t^+, {}_m\xi))} f(X^{t^+, g(t,x,t^+,\, {}_m\xi)}_{\tau_*^{t^+, g(t,x,t^+,\, {}_m\xi)}}), \tag{4.18}
$$

and for the gradients (Deltas) we have the probabilistic representation

$$
\Delta_i := \frac{\partial u}{\partial x^i}(t,x) = E\, \frac{\partial}{\partial x^i} \left(\frac{p(t,x,t^+, g(t,x,t^+,\xi))}{\phi(t,x,t^+, g(t,x,t^+,\xi))} f(X^{t^+, g(t,x,t^+,\xi)}_{\tau_*^{t^+, g(t,x,t^+,\xi)}}) \right) \tag{4.19}
$$

with Monte Carlo estimator

$$
\widehat{\Delta}_i := \frac{1}{M} \sum_{m=1}^{M} \frac{\partial}{\partial x^i} \left(\frac{p(t,x,t^+, g(t,x,t^+, {}_m\xi))}{\phi(t,x,t^+, g(t,x,t^+, {}_m\xi))} f(X^{t^+, g(t,x,t^+,\, {}_m\xi)}_{\tau_*^{t^+, g(t,x,t^+,\, {}_m\xi)}}) \right), \tag{4.20}
$$

where ${}_m\xi$, $m = 1, ..., M$, are i.i.d. samples from the reference density λ. Indeed, by precondtioning on $\mathcal{F}_{t^+}$ and then taking expectations we see that (4.18) and (4.20) are unbiased

Monte Carlo estimators for the price (4.17) and 'deltas' (4.19), respectively. Moreover, if ϕ is close to p in the sense of Theorem 3, it is not difficult to see that also gradient estimator (4.20) has non-exploding variance when $t^+ \downarrow t$.

Estimators (4.18) and (4.20) are useful if one has an analytic approximation $\widehat{p}(t,x,t^+,y)$ of the density $p(t,x,t^+,y)$ and known densities $\phi(x,\cdot)$ for $x \in \mathbb{R}^n_+$. The approximation $\widehat{p}$ may be obtained by some specific method, for example by a WKB expansion as presented in Section 5, or some lognormal approximation as proposed in Kurbanmuradov, Sabelfeld & Schoenmakers (2002) for the Libor market model. Of course the density ϕ has to be chosen with some care. If it is possible to sample directly from $\widehat{p}$ (e.g. in case of a log-normal approximation) we may take $\phi = \widehat{p}$. If not, (e.g. in the case of a WKB expansion) one may take for ϕ a (not necessarily very accurate) lognormal approximation of the density p.

A canonical lognormal approximation for $p(t,x,t^+,z)$ is obtained by freezing X in the coefficients of (2.1) at the initial time. We thus yield

$$X^{(g)t,x;i}_{t^+} := x^i \exp\left(-\frac{1}{2}\sum_{j=1}^n \int_t^{t^+} (\sigma^{ij})^2(s,x)ds + \int_t^{t^+} r(s,x)ds \right.$$
$$\left. + \sum_{j=1}^n \int_t^{t^+} \sigma^{ij}(s,x)dW^j_s \right) =: x_i \exp(\xi_i). \tag{4.21}$$

Here, $(\xi_i)_{i=1}^n$ is a gaussian random vector with

$$E\xi_i = -\frac{1}{2}\sum_{j=1}^n \int_t^{t^+} (\sigma^{ij})^2(s,x)ds + \int_t^{t^+} r(s,x)ds =: \mu^{i;t,t^+,x}, \quad 1 \le i \le n,$$

and

$$Cov(\xi_i,\xi_j) = \sum_{l=1}^n \int_t^{t^+} \sigma^{il}(s,x)\sigma^{jl}(s,x)\,ds =: \sigma^{ij;t,t^+,x}, \quad 1 \le i,j \le n.$$

Clearly, the density ϕ is then given by

$$\phi(t,x,t^+,y) := \frac{\psi_{\mu^{t,t^+,x},\sigma^{t,t^+,x}}(\ln\frac{y^1}{x^1}, \ln\frac{y^2}{x^2}, ..., \ln\frac{y^n}{x^n})}{y^1 y^2 \cdots y^n}, \tag{4.22}$$

$y^i > 0$, $1 \le i \le n$, with $\psi_{\mu^{t,t^+,x},\sigma^{t,t^+,x}}$ being the density of the n-dimensional normal distribution $\mathcal{N}_n(\mu^{t,t^+,x},\sigma^{t,t^+,x})$ with $\mu^{t,t^+,x} := (\mu^{i;t,t^+,x})_{1\le i\le n}$ and $\sigma^{t,t^+,x} := (\sigma^{ij;t,t^+,x})_{1\le i,j\le n}$.

For practical applications it is most useful to discretize estimator (4.20) to

$$\widehat{\Delta}^h_i := \frac{1}{M}\sum_{m=1}^M \frac{1}{2h}\left(\frac{p(t,x+\mathfrak{h}_i,t^+,g(t,x+\mathfrak{h}_i,t^+,{}_m\xi))}{\phi(t,x+\mathfrak{h}_i,t^+,g(t,x+\mathfrak{h}_i,t^+,{}_m\xi))} f(X^{t^+,g(t,x+\mathfrak{h}_i,t^+,\ {}_m\xi)}_{\tau_*^{t^+,g(t,x+\mathfrak{h}_i,t^+,\ {}_m\xi)}}) \right.$$
$$\left. - \frac{p(t,x-\mathfrak{h}_i,t^+,g(t,x-\mathfrak{h}_i,t^+,{}_m\xi))}{\phi(t,x-\mathfrak{h}_i,t^+,g(t,x-\mathfrak{h}_i,t^+,{}_m\xi))} f(X^{t^+,g(t,x-\mathfrak{h}_i,t^+,\ {}_m\xi)}_{\tau_*^{t^+,g(t,x-\mathfrak{h}_i,t^+,\ {}_m\xi)}}) \right), \tag{4.23}$$

where $\mathfrak{h}_i := h(\delta_{i1},\ldots,\delta_{in})$, for small enough $h > 0$. As an alternative, it is also possible to push the derivatives in (4.20) through. Since the set of exercise dates is discrete, $\tau_*^{t^+,y}(\omega)$

may be considered locally constant in y on a fixed ω. Thus, by differentiating (4.20) pathwise we obtain

$$\widehat{\Delta}_i :== \frac{1}{M}\sum_{m=1}^{M} f(X^{t^+,g(t,x,t^+,\ _m\xi)}_{\tau_*^{t^+,g(t,x,t^+,\ _m\xi)}})\frac{\partial}{\partial x^i}\left(\frac{p(t,x,t^+,g(t,x,t^+,_m\xi))}{\phi(t,x,t^+,g(t,x,t^+,_m\xi))}\right) \qquad (4.24)$$

$$+\frac{1}{M}\sum_{m=1}^{M}\frac{p(t,x,t^+,g(t,x,t^+,_m\xi))}{\phi(t,x,t^+,g(t,x,t^+,_m\xi))}\frac{\partial f}{\partial z}(X^{t^+,g(t,x,t^+,\ _m\xi)}_{\tau_*^{t^+,g(t,x,t^+,\ _m\xi)}})\cdot$$

$$\cdot\,\partial_y X^{t^+,y}_{\tau_*^{t^+,g(t,x,t^+,\ _m\xi)}}(g(t,x,t^+,\ _m\xi))\frac{\partial g(t,x,t^+,\ _m\xi)}{\partial x^i},$$

where

$$\frac{\partial}{\partial x^i}\frac{p(t,x,t^+,y)}{\phi(t,x,t^+,y)},\qquad \frac{\partial g(t,x,t^+,\ _m\xi)}{\partial x^i},\qquad \frac{\partial f}{\partial z}$$

can in principle be expressed analytically, and the vector process

$$\partial_y X_s^{t^+,y}(\cdot) := \frac{\partial X_s^{t^+,y}}{\partial y}(\cdot),\qquad s\geq t^+,$$

can in principle be simulated via a variational system of SDEs (e.g. see Protter (1990), Milstein & Schoenmakers (2002), Giles & Glasserman (2006)).

In this paper we will prefere the discretized version (4.23) of (4.20) for our applications. The algorithm is as follows. We first choose a $h>0$, and sample $_m\xi$ for $m=1,\ldots,M$ from the reference (usually normal) density. Next we simulate for each m a pair of trajectories $_mX^{\pm}$, which start in $_mg^{\pm} := g(t,x\pm h,t^+,\ _m\xi)$ at t^+, and end at the optimal stopping times $_m\tau_*^{\pm} := \tau_*^{t^+,_mg^{\pm}}$. Of course the optimal exercise dates $_m\tau_*^{\pm}$ are generally unknown in practice, but we assume that we have good approximations $_m\tau^{\pm}$ at hand, which are constructed via some well known procedure. For example, in a pre-computation we may construct an exercise boundary via the regression method of Longstaff & Schwartz (2001) (Tsitsiklis & Van Roy (2001)), or as an alternative, we may use the policy iteration method of Kolodko & Schoenmakers (2006), see also Bender & Schoenmakers (2006). For each m we compute also the values $_mp^{\pm} := p(t,x\pm h,t^+,\ _mg^{\pm})$ and $_m\phi^{\pm} := \phi(t,x\pm h,t^+,\ _mg^{\pm})$, and finally compute the estimate

$$\widehat{\Delta}_i^h := \frac{1}{M}\sum_{m=1}^{M}\frac{1}{2h}\left(\frac{_mp^+}{_m\phi^+}f(_mX^+) - \frac{_mp^-}{_m\phi^-}f(_mX^-)\right). \qquad (4.25)$$

Remark 7. The presented estimators are particularly effective for European products and for longer dated Bermudans where $t\ll t^+\leq t_1$.

Remark 8. In the previous sections vector and matrix components are denoted by superscripts, so that time parameters of processes can be denoted by subscripts. In the next sections we deflect from this convention and use subscripts for vector and matrix components.

5 WKB approximations for Greenian kernels

5.1 Recap of WKB theory

We summarize some results concerning WKB-expansions of parabolic equations (cf. Kampen (2006) for details). Let us consider the parabolic diffusion operator

$$\frac{\partial u}{\partial t} + Lu \equiv \frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial u}{\partial x_i}. \tag{5.26}$$

For simplicity of notation and without loss of generality it is assumed that the diffusion coefficients a_{ij} and the first order coefficients b_i in (5.26) depend on the spatial variable x only. In the following let $\delta t = T - t$, and let the functions

$$(x,y) \to d(x,y) \geq 0, \quad (x,y) \to c_k(x,y), \; k \geq 0,$$

be defined on $\mathbb{R}^n \times \mathbb{R}^n$, with d^2 and c_k, $k \geq 0$, being smooth. Then a set of (simplified) conditions sufficient for pointwise valid WKB-representations of the form

$$p(t,x,T,y) = \frac{1}{\sqrt{2\pi\delta t}^{n}} \exp\left(-\frac{d^2(x,y)}{2\delta t} + \sum_{k=0}^{\infty} c_k(x,y)\delta t^k \right), \tag{5.27}$$

for the solution $(t,x) \to p(t,x,T,y)$ of the final value problem

$$\begin{aligned} \frac{\partial p}{\partial t} + Lp &= 0, \quad \text{with final value} \\ p(T,x,T,y) &= \delta(x-y), \quad y \in \mathbb{R}^n \quad \text{fixed}, \end{aligned} \tag{5.28}$$

is given by

(A) The operator L is uniformly elliptic in $\mathbb{R}^n$, i.e. as in (2.2) the matrix norm of $(a_{ij}(x))$ is bounded below and above by $0 < \lambda < \Lambda < \infty$ uniformly in x,

(B) the smooth functions $x \to a_{ij}(x)$ and $x \to b_i(x)$ and all their derivatives are bounded.

For more subtle (and partially weaker conditions) we refer to Kampen (2006). If we add the uniform boundedness condition

(C) there exists a constant c such that for each multiindex α and for all $1 \leq i,j,k \leq n$,

$$\left|\frac{\partial a_{jk}}{\partial x^\alpha}\right|, \; \left|\frac{\partial b_i}{\partial x^\alpha}\right| \leq c \exp\left(c|x|^2\right), \tag{5.29}$$

then the Taylor expansions of the functions d and c_k around $y \in \mathbb{R}^n$ are equal to d and $c_k, k \geq 0$ globally. I.e. we have the power series representations

$$d^2(x,y) = \sum_\alpha d_\alpha(y)\delta x^\alpha \tag{5.30}$$

$$c_k(x,y) = \sum_\alpha c_{k,\alpha}(y)\delta x^\alpha, \qquad k \geq 0. \tag{5.31}$$

Note that (C) is implied by the stronger condition that all derivatives in (5.29) have a uniform bound. Summing up we have the following theorem:

Theorem 9. *If the hypotheses (A),(B) are satisfied, then the fundamental solution p has the representation*

$$p(\delta t, x, y) = \frac{1}{\sqrt{2\pi\delta t}^n} \exp\left(-\frac{d^2(x,y)}{2\delta t} + \sum_{k\geq 0} c_k(x,y)\delta t^k\right), \tag{5.32}$$

where d and c_k are smooth functions, which are unique global solutions of the first order differential equations (5.33),(5.34), *and* (5.36) *below. Especially,*

$$(\delta t, x, y) \to \delta t \ln p(\delta t, x, y) = -\frac{n}{2}\delta t \ln 2\pi\delta t - \frac{d^2}{2} + \sum_{k\geq 0} c_k(x,y)\delta t^{k+1}$$

is a smooth function which converges to $-\frac{d^2}{2}$ as $\delta t \searrow 0$, where d is the Riemannian distance induced by the line element $ds^2 = \sum_{ij} a_{ij}^{-1} dx_i dx_j$, where with a slight abuse of notation (a_{ij}^{-1}) denotes the matrix inverse of (a_{ij}). If the hypotheses (A),(B) and (C) are satisfied, then in addition the functions d, $c_k, k \geq 0$ equal their Taylor expansion around y globally, i.e. we have (5.30)-(5.31).

The recursion formulas for d and $c_k,\ k \geq 0$ are obtained by plugging the Ansatz (5.27) into the parabolic equation (5.28), and ordering terms with respect to the monoms $\delta t^i = (T-t)^i$ for $i \geq -2$. By collecting terms of order δt^{-2} we obtain

$$d^2 = \frac{1}{4}\sum_{ij} d^2_{x_i} a_{ij} d^2_{x_j}, \tag{5.33}$$

where $d^2_{x_k}$ denotes the derivative of the function d^2 with respect to the variable x_k, with the boundary condition $d(x,y) = 0$ for $x = y$. Collecting terms of order δt^{-1} yields

$$-\frac{n}{2} + \frac{1}{2}Ld^2 + \frac{1}{2}\sum_i \left(\sum_j \left(a_{ij}(x) + a_{ji}(x)\right)\frac{d^2_{x_j}}{2}\right)\frac{\partial c_0}{\partial x_i}(x,y) = 0, \tag{5.34}$$

where the boundary condition

$$c_0(y,y) = -\frac{1}{2}\ln\sqrt{\det\left(a^{ij}(y)\right)} \tag{5.35}$$

determines c_0 uniquely for each $y \in \mathbb{R}^n$. Finally, for $k+1 \geq 1$ we obtain

$$\begin{aligned} &(k+1)c_{k+1}(x,y) + \tfrac{1}{2}\textstyle\sum_{ij} a_{ij}(x)\left(\frac{d^2_{x_i}}{2}\frac{\partial c_{k+1}}{\partial x_j} + \frac{d^2_{x_j}}{2}\frac{\partial c_{k+1}}{\partial x_i}\right) \\ &= \tfrac{1}{2}\textstyle\sum_{ij} a_{ij}(x)\sum_{l=0}^k \frac{\partial c_l}{\partial x_i}\frac{\partial c_{k-l}}{\partial x_j} + \frac{1}{2}\sum_{ij} a_{ij}(x)\frac{\partial^2 c_k}{\partial x_i \partial x_j} + \sum_i b_i(x)\frac{\partial c_k}{\partial x_i}, \end{aligned} \tag{5.36}$$

with boundary conditions

$$c_{k+1}(x,y) = R_k(y,y) \text{ if } \quad x = y, \tag{5.37}$$

R_k being the right side of (5.36). For some classical models in finance a global transformation of the diffusion operator to the Laplace operator is possible (at the price of more complicated first order terms however). We observe this in the case of a Libor market model (Section 6). By a straightforward derivation we get the next proposition.

Proposition 10. *There is a global coordinate transformation for the operator (5.26) such that the second order part of the transformed operator equals the Laplacian, iff* $a_{ij} = (\sigma\sigma^\top)_{ij}$ *for a (square) matrix function* σ *which satisfies*

$$\sum_{l=1}^{n} \frac{\partial \sigma_{ik}(x)}{\partial x_l} \sigma_{lj}(x) = \sum_{l=1}^{n} \frac{\partial \sigma_{ij}(x)}{\partial x_l} \sigma_{lk}(x), \quad x \in \mathbb{R}^n. \tag{5.38}$$

The latter fact is also observed and proved in Ait-Sahalia (2006). If the condition of Proposition 10 is satisfied, then coordinate transformation leads to second order coefficients of the form $a_{ij} \equiv \delta_{ij}$, so that the solution of (5.33) becomes

$$d^2(x,y) = \sum_i (x_i - y_i)^2.$$

If conditions (A), (B), (C), and (5.38) hold, then in the transformed coordinates, explicit formulas for the coefficient functions $c_k, k \geq 0$ can be computed via the formulas

$$c_0(x,y) = \sum_i (y_i - x_i) \int_0^1 b_i(y + s(x-y))ds, \tag{5.39}$$

and

$$c_{k+1}(x,y) = \int_0^1 R_k(y + s(x-y), y) s^k ds, \tag{5.40}$$

with R_k being the right-hand-side of (5.36) where $a_{ij} = \delta_{ij}$. Similar formulas are obtained in Ait-Sahalia (2006). In Kampen (2006) it is shown in addition how the coefficients c_k can be computed explicitly in terms of power series approximations of the diffusion coefficients a_{ij} and b_i. However, in high dimensional models such as the Libor market model direct computation of the coefficients c_k seems more feasible as it turns out that the computation up to the coefficient c_1 is sufficient for our purposes. We conclude this Section with a final remark concerning possible generalizations to Feller processes.

Remark 11. Pointwise converging analytic expansion can be obtained for a large class of densities p_F of Feller processes which are fundamental solutions of the equation

$$\frac{\partial u}{\partial \delta t} = \frac{1}{2} \sum_{ij} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i(x) \frac{\partial u}{\partial x_i} + I\left[u\right],$$

where a_{ij} and b_i satisfy conditions (A), (B), and (C), and

$$I\left[f\right] \equiv \int_{\mathbb{R}^n} \left\{ f(x+y) - f(x) - 1_{[-1,1]^n}(y) \frac{\partial f}{\partial x} y \right\} \nu(x,dy)$$

with a jump measure $\nu(x,dy)$. If the jump measure $\nu(x,.)$ is a Radon measure on $\mathbb{R}^n \setminus \{0\}$, and

$$\int_{\mathbb{R}^n \setminus \{0\}} z^\alpha \nu(x,dz) =: \epsilon_\alpha(x)$$

holds with functions $\epsilon_\alpha \in C^\infty(\mathbb{R}^n)$ where

$$\frac{\sup_{x \in \mathbb{R}^n}(\epsilon_\alpha(x))}{\alpha!} \downarrow 0 \text{ as } \alpha \uparrow \infty,$$

then there exists a pointwise converging representation of the form

$$p_F(t,x,T,y) = p_{\mathrm{WKB}}(t,x,T,y)\left(\sum_{l\geq 0} r_l(x,y)\delta t^l\right)$$

where p_{WKB} is the WKB-expansion (5.27) of the diffusion equation without integral term and r_l are coefficient functions which satisfy certain first order equations similar to those of the WKB-coefficients c_k (cf. Kampen (2007)).

5.2 Error estimates

We now study the approximation error of a truncated WKB expansion (and its derivatives), which is essential for the convergence of the Monte Carlo schemes. In this respect we will show how the derivatives (up to second order) of the product value fuction with respect to the underlyings computed by means of a truncated WKB-expansion converge in supremum norm and Hölder norms. Let us consider a WKB-approximation of the fundamental solution p of the form

$$p_l(t,x,T,y) = \frac{1}{\sqrt{2\pi\delta t}^n}\exp\left(-\frac{d^2(x,y)}{2\delta t} + \sum_{k=0}^{l} c_k(x,y)\delta t^k\right), \tag{5.41}$$

i.e. we assume that the coefficients d^2 and $c_k,\ 0\leq k\leq l$ have been computed up to order l. We use the following a priori estimate: Let us denote the domain of the Cauchy problem by $D=(0,T)\times\mathbb{R}^n$. For integers $n\geq 0$ and real numbers $\delta\in(0,1)$ let $C^{m+\delta/2,n+\delta}(D)$ be the space of m (n) times differentiable functions such that the mth (nth) derivative with respect to time (space) is Hölder continuous with exponent $\frac{\delta}{2}$ (δ). Furthermore, $|.|_{m+\delta/2,n+\delta}$ denote the natural norms associated with these function spaces. Then a consequence of Safanov's theorem (cf. Krylov (1996)) is

Theorem 12. *Assume that* $(A),(B),$ *and* (C) *are satisfied and let* $g\in C^{2+\delta}(\mathbb{R}^n)$ *and* $f\in C^{\delta/2,\delta}(D)$. *If*

$$c\leq -\lambda \ \textit{ for some }\ \lambda>0, \tag{5.42}$$

then the Cauchy problem

$$\begin{cases} \frac{\partial w}{\partial t} + \frac{1}{2}\sum_{ij} a_{ij}(x)\frac{\partial^2 w}{\partial x_i\partial x_j} + \sum_i b_i(x)\frac{\partial w}{\partial x_i} + c(x)w = f(t,x) \ \textit{in } D \\[2ex] w(T,x) = g(x) \ \textit{for } x\in\mathbb{R}^n \end{cases} \tag{5.43}$$

has a unique solution w, *and there exists a constant* c *depending only on* δ, n λ,Λ *and* $K=\max\{|a|_\delta,|b|_\delta,|c|_\delta\}$ *such that*

$$|w|_{1+\delta/2,2+\delta}\leq c\left[|f|_{\delta/2,\delta} + |g|_{2+\delta}\right]. \tag{5.44}$$

In order to analyze the truncation error of the Cauchy problem with data g we consider he function

$$u^{\Delta}(t,x) = u(t,x) - u_l(t,x),$$

where

$$u(t,x)=\int_{\mathbb{R}^n} g(y)p(t,x,T,y)dy,$$

and

$$u_l(t,x)=\int_{\mathbb{R}^n} g(y)p_l(t,x,T,y)dy.$$

Observe that

$$\frac{\partial u^\Delta}{\partial t}+Lu^\Delta=-\frac{\partial u_l}{\partial t}-Lu_l=:f_{u_l}(t,x)$$

Now assume that $g\in C_0^\delta(\mathbb{R}^n)$, i.e. g has compact support. Then we apply the estimate (5.44) to the function

$$w(t,x)=e^{-rt}u^\Delta$$

which solves (5.42) for a constant $r>0$. Then we get

$$|u^\Delta|_{1+\delta/2,2+\delta}\leq c|f_{u_l}|_{\delta/2,\delta}. \tag{5.45}$$

In order to compute the term on the right side of (5.45) we first plug (5.41) into the left-hand side of (5.28) the parabolic equation satisfied by the exact fundamental solution p. We get up to higher order terms in δt (recall that $\delta t=T-t$)

$$\begin{aligned}&\tfrac{\partial p_l}{\partial t}+\tfrac{1}{2}\textstyle\sum_{ij}a_{ij}\tfrac{\partial^2 p_l}{\partial x_i\partial x_j}+\sum_i b_i\tfrac{\partial p}{\partial x_i}\\ &=\left[\textstyle\sum_i b_i c_{l,x_i}\delta t^l-\sum_{p=l}^{2l}\left(\sum_{k+m=p}c_{k,x_i}c_{m,x_j}\right)\delta t^p\right]p_l.\end{aligned}$$

Since the coefficients a_{ij} and b_i are bounded and applying a priori estmates for p_l we get

Theorem 13. *Assume that conditions (A), (B), and (C) hold and that $g\in C_0^\delta(\mathbb{R}^n)$. Then*

$$|u(t,x,y)-u_l(t,x,y)|_{1+\delta/2,2+\delta}=O(\delta t^{l-\frac{\delta}{2}}).$$

It is well known that $g\in C_0^\delta(\mathbb{R}^n)$ is not essentially restrictive for practical purposes. Note that the estimate above is in a very strong norm which ensues pointwise convergence in the Hölder sense up to the second derivatives of the value function. The result can easily extended to the case where the functions $d^2, c_k, \leq l$ are known only in terms of their Taylor representation up to some order. The case where c_k are computed up to $k=1$ is the first case where the truncation error for first and second derivatives converges to zero (in supremum norm with order $O(\delta t)$ and in Hölder- extension of supremum norm with order $O(\delta t)^{1-\frac{\delta}{2}}$). This implies that our Monte Carlo computation scheme for the Greeks converges.

5.3 Extension to diffusions with time-dependent coefficients

The results of Section 5.1 and 5.2. extend to time-dependent diffusions without great difficulties. However, some remarks are in order as the numerical example of the Libor market model below belongs to that class of diffusions. First, for a diffusion equation of the form

$$\tfrac{\partial u}{\partial t}+L_t u\equiv\tfrac{\partial u}{\partial t}+\tfrac{1}{2}\textstyle\sum_{i,j}a_{ij}(t,x)\tfrac{\partial^2 u}{\partial x_i\partial x_j}+\sum_i b_i(t,x)\tfrac{\partial u}{\partial x_i}=0$$

Theorem 9 and Proposition 10 hold with time-dependent coefficient functions, and the corresponding WKB expansion is of the form

$$p(t,x,T,y)=\frac{1}{\sqrt{2\pi\delta t}^{n}}\exp\left(-\frac{d^2(t,x,T,y)}{2\delta t}+\sum_{k=0}^{\infty}c_k(t,x,y)\delta t^k\right). \tag{5.46}$$

Formally, recursion equations for the WKB-coefficient functions c_k can be obtained in a similar way as in the time-independent case. Application of a convergence theorem in (Kampen, 2006) leads to a pointwise valid representation (5.46). In case a coordinate transformation in the sense of Proposition 10 is possible, we obtain the final value problem

$$\begin{aligned}\frac{\partial p}{\partial t}+\Delta p+\sum_i \mu_i(t,x)\frac{\partial p}{\partial x_i} &= 0, \quad \text{with final value}\\ p(T,x,T,y) &= \delta(x-y),\end{aligned}$$

for the WKB kernel $(t,x)\to p(t,x,T,y)$ in the transformed coordinates. In particular, the function $d^2(x,y)=\sum_i(x_i-y_i)^2$ does not depend on the time parameter anymore, and the time-dependent WKB-coefficient functions c_k are recursively defined by the first order partial differential equations

$$\sum_i (x_i-y_i)\left(\mu_i(t,x)+\frac{\partial c_0}{\partial x_i}\right)(t,x,y)=0, \tag{5.47}$$

for $k=0$, and

$$\begin{aligned}&(k+1)c_{k+1}(t,x,y)+\textstyle\sum_i(x_i-y_i)\frac{\partial c_{k+1}}{\partial x_i}(t,x,y)\\ &=R_k(t,x,y)\equiv \frac{\partial c_k}{\partial t}+\frac{1}{2}\textstyle\sum_i\sum_{l=0}^{k}\frac{\partial c_l}{\partial x_i}\frac{\partial c_{k-l}}{\partial x_i}+\frac{1}{2}\sum_i\frac{\partial^2 c_k}{\partial x_i^2}+\sum_i\mu_i(t,x)\frac{\partial c_k}{\partial x_i},\end{aligned} \tag{5.48}$$

for $k+1\geq 1$, which have to be solved with the boundary conditions $c_0(t,x,y)=0$ if $x=y$, and $c_{k+1}(t,x,y)=R_k(t,y,y)$ if $x=y$. Analogue to the time-inhomogenous case we compute

$$\begin{aligned}c_{k+1}(t,x,y) &= \int_0^1 R_k(t,y+s(x-y),y)s^k ds = R_k(t,x,y)\frac{1}{k+1}\\ &\quad -\int_0^1\sum_i(x_i-y_i)\partial_i R_k(t,y+s(x-y),y)\frac{s^{k+1}}{k+1}ds,\end{aligned}$$

and obtain

$$\begin{aligned}(k+1)c_{k+1}(t,x,y) &= R_k(t,x,y)-\\ &\quad \int_0^1\sum_i(x_i-y_i)\partial_i R_k(t,y+s(x-y),y)s^{k+1}ds\\ &= R_k(t,x,y)-\sum_i(x_i-y_i)\frac{\partial c_{k+1}}{\partial x_i}.\end{aligned}$$

As we see, the solution is formally the same as in the time-homogeneous case.

6 Applications to the Libor market model

We consider a Libor market model with respect to a tenor structure $0 < T_1 \ldots < T_{n+1}$ in the terminal measure P_{n+1} (induced by the terminal zero coupon bond $B_{n+1}(t)$). The dynamics of the forward Libors $L_i(t)$, defined in the interval $[0, T_i]$ for $1 \le i \le n$, are governed by the following system of SDE's (e.g., see Jamshidian (1997)),

$$dL_i = -\sum_{j=i+1}^{n} \frac{\delta_j L_i L_j \gamma_i^\top \gamma_j}{1 + \delta_j L_j} dt + L_i \gamma_i^\top dW_{n+1} =: \mu_i(t, L) + L_i \gamma_i^\top dW_{n+1}, \tag{6.49}$$

where $\delta_i = T_{i+1} - T_i$ are day count fractions and $t \to \gamma_i(t) = (\gamma_{i,1}(t), \ldots, \gamma_{i,d}(t))$, $(\gamma_i^\top \gamma_j)_{i,j=1}^n =: \rho$ are deterministic volatility vector functions defined in $[0, T_i]$. We denote the matrix with rows $\gamma_i^\top$ by Γ and assume that Γ is invertible. In (6.49), $(W_{n+1}(t) \mid 0 \le t \le T_n)$ is a standard d-dimensional Wiener process under the measure P_{n+1} with d, $1 \le d \le n$, being the number of driving factors. In what follows we consider the full-factor Libor model with $d = n$ in the time interval $[0,\, T_1)$.

6.1 WKB approximations for the Libor kernel

To approximate the transition density $p^L(s, u, t, v)$ of Libor process for $0 < s < t$, we transform (6.49) to an equation of form

$$dY_i = \mu_i^Y(t, Y)dt + dW_{n+1}^i, \quad 1 \le i \le n. \tag{6.50}$$

For the transition density $p^Y(s, x, t, y)$, we can compute c_k, $k = 0, 1, \ldots$ in (5.46) via (5.47)-(5.48). After that we find $p^L(s, u, t, v)$ by a density transformation formula. In order to find μ^Y in (6.50), we first transform (6.49) to $K_i = \log L_i$, $1 \le i \le n$,

$$dK_i = \frac{1}{L_i} dL_i - \frac{1}{2L_i^2} d\langle L_i \rangle = \left(-\frac{\gamma_i^\top \gamma_i}{2} + \mu_i(t, e^{L_1}, \ldots, e^{L_n}) \right) dt + \gamma_i^\top dW_{n+1}.$$

Then, by the transformation $Y = \Gamma^{-1} K$ we get

$$\mu_i^Y(t, Y) = (MY)_i + V_i + \sum_{j=1}^{n} \Gamma_{ij}^{-1} \mu_j(t, e^{(\Gamma Y)_1}, \ldots, e^{(\Gamma Y)_n}),$$

where

$$M_{ij} = \sum_{l=1}^{n} \frac{d\Gamma_{il}^{-1}}{dt} \Gamma_{lj}, \quad V_i = -\sum_{j=1}^{n} \Gamma_{ij}^{-1} \frac{|\gamma_j|^2}{2}.$$

In our case WKB coefficients have a rather complicated form. However, according to our experience, the convergence of the WKB expansion is typically very fast. In our example (see Section 6.2) a WKB series with c_0 and c_1 only provides a very good appoximation for the transition density. We now give formulas for c_0 and its derivatives $\frac{\partial c_0}{\partial x_p}$, $\frac{\partial^2 c_0}{\partial x_p^2}$, $1 \le p \le n$, for further use in our numerical simulation (to compute c_1, we integrate $R_0(s, x, y)$ in (5.48) numerically by the trapezoidal rule). Using the notation

$$F_l(s, x, y) := \frac{1}{(\Gamma(x - y))_l} \ln \frac{1 + \delta_l e^{(\Gamma x)_l}}{1 + \delta_l e^{(\Gamma y)_l}}, \quad 1 \le l \le n.$$

we may write,

$$
\begin{aligned}
c_0(s,x,y) &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}(y_i-x_i)(y_j+x_j) + \sum_{i=1}^{n} V_i(y_i-x_i) \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{n} \Gamma_{ij}^{-1}(y_i-x_i) \sum_{l=j+1}^{n} \rho_{jl} F_l(s,x,y),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial c_0}{\partial x_p}(s,x,y) &= \frac{1}{2}\sum_{i=1}^{n}(M_{ip}(y_i-x_i) - M_{pi}(y_i+x_i)) - V_p + \\
&\quad \sum_{j=1}^{n} \Gamma_{pj}^{-1} \sum_{l=j+1}^{n} \rho_{jl} F_l(s,x,y) - \\
&\quad \sum_{i=1}^{n}\sum_{j=1}^{n} \Gamma_{ij}^{-1}(y_i-x_i) \sum_{l=j+1}^{n} \rho_{jl} \frac{\partial F_l(s,x,y)}{\partial x_p},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 c_0}{\partial x_p^2}(s,x,y) &= -M_{pp} + 2\sum_{j=1}^{n} \Gamma_{pj}^{-1} \sum_{l=j+1}^{n} \rho_{jl} \frac{\partial F_l(s,x,y)}{\partial x_p} - \\
&\quad \sum_{i=1}^{n}\sum_{j=1}^{n} \Gamma_{ij}^{-1}(y_i-x_i) \sum_{l=j+1}^{n} \rho_{jl} \frac{\partial^2 F_l(s,x,y)}{\partial x_p^2},
\end{aligned}
$$

where

$$
\frac{\partial F_l(s,x,y)}{\partial x_p} = \frac{\Gamma_{lp}(s)}{(\Gamma(x-y))_l}\left(\frac{\delta_l e^{(\Gamma x)_l}}{1+\delta_l e^{(\Gamma x)_l}} - F_l(s,x,y),\right)
$$

and

$$
\begin{aligned}
\frac{\partial^2 F_l(s,x,y)}{\partial x_p^2} &= \frac{2\Gamma_{lp}^2}{(\Gamma(x-y))_l^2}\left(F_l(s,x,y) - \frac{\delta_l e^{(\Gamma x)_l}}{1+\delta_l e^{(\Gamma x)_l}}\right) \\
&\quad + \frac{\Gamma_{lp}^2}{(\Gamma(x-y))_l}\frac{\delta_l e^{(\Gamma x)_l}}{(1+\delta_l e^{(\Gamma x)_l})^2}.
\end{aligned}
$$

The function $F(s,x,y)$ is analytical in the whole domain. In particluar, there is no singularity when $(\Gamma x)_l = (\Gamma y)_l$, $1 \leq l \leq p$. By L'Hospital's Rule we get,

$$
\begin{aligned}
\lim_{(\Gamma x)_l \to (\Gamma y)_l} F_l(s,x,y) &= \frac{\delta_l e^{(\Gamma y)_l}}{1+\delta_l e^{(\Gamma y)_l}}, \\
\lim_{(\Gamma x)_l \to (\Gamma y)_l} \frac{\partial F_l(s,x,y)}{\partial x_p} &= \frac{\Gamma_{lp}\delta_l e^{(\Gamma y)_l}}{2(1+\delta_l e^{(\Gamma y)_l})^2}, \\
\lim_{(\Gamma x)_l \to (\Gamma y)_l} \frac{\partial^2 F_l(s,x,y)}{\partial x_p^2} &= \frac{\Gamma_{lp}^2\delta_l e^{(\Gamma y)_l}(1-\delta_l e^{(\Gamma y)_l})}{3(1+\delta_l e^{(\Gamma y)_l})^3}.
\end{aligned}
$$

We finally obtain $p^L(s,u,t,v)$ by density transformation formula,

$$
p^L(s,u,t,v) = p^Y(s, S_s^{-1}(u), t, S_t^{-1}(v)) \left|\frac{\partial S_t^{-1}(v)}{\partial v}\right|
$$

with

$$S_t^{-1}(v) := \Gamma^{-1}(t)(\log v_1, \ldots, \log v_n)^\top.$$

In our numerical example (Section 6.2) we assume for simplicity that the matrix Γ is upper triangular and does not depend on t. In this case we have,

$$\begin{aligned} p^L(s,u,t,v) \quad &= \quad \frac{1}{\sqrt{2\pi(t-s)}^n} \prod_{i=1}^n \frac{\Gamma_{ii}^{-1}}{v_i} \exp\left(-\frac{\left(\Gamma^{-1}(\log \frac{v_1}{u_1}, \ldots, \log \frac{v_n}{u_n})^\top\right)^2}{2(t-s)} \right. \\ &\qquad \left. + \sum_{k=0}^{\infty} c_k(s, S^{-1}(u), S^{-1}(v))(t-s)^k \right). \end{aligned}$$

6.2 Case study: European swaptions

Estimators (3.7) and (3.13) will be tested for pricing European swaptions and Deltas in a Libor market model for different maturities T_1. A (payer) swaption contract with maturity T_i and strike θ with principal \$1 gives the right to contract at T_i for paying a fixed coupon θ and receiving floating Libor at the settlement dates $T_{i+1}, \ldots, T_n$. So, in (3.7) and (3.13) we take

$$u(L) = \frac{B_{n+1}(0)}{B_{n+1}(T_1)} \left(\sum_{j=i}^n B_{j+1}(T_1)\left(\delta_j L_j(T_1) - \theta\right) \right)^+. \tag{6.51}$$

and then the discretized analogue of (3.13) reads,

$$\begin{aligned} \widehat{\frac{\partial I}{\partial x_i}}^{(h)}(x) \quad := \quad & \frac{1}{M} \sum_{m=1}^M \frac{1}{2h} \left(\frac{p(x+\mathfrak{h}_i, g(x+\mathfrak{h}_i, {}_m\xi))u(g(x+\mathfrak{h}_i, {}_m\xi))}{\phi(x+\mathfrak{h}_i, g(x+\mathfrak{h}_i, {}_m\xi))} - \right. \\ & \left. \frac{p(x-\mathfrak{h}_i, g(x-\mathfrak{h}_i, {}_m\xi))u(g(x-\mathfrak{h}_i, {}_m\xi))}{\phi(x-\mathfrak{h}_i, g(x-\mathfrak{h}_i, {}_m\xi))} \right), \quad 1 \le i \le n, \end{aligned} \tag{6.52}$$

where $\mathfrak{h}_i = h(\delta_{i1}, \ldots, \delta_{in})$.

For our experiments we take $\delta_i \equiv 0.5$ for $i \ge 1$, flat 3.5% initial Libor curve, and constant volatility loadings

$$\gamma_i(t) \equiv 0.2 e_i,$$

in the Libor market model (6.49), where e_i are n-dimensional unit vectors decomposing an input correlation matrix ρ,

$$\rho_{ij} = \exp\left[\frac{|j-i|}{n-1} \ln \rho_\infty\right], \quad 1 \le i, j \le n. \tag{6.53}$$

with $n > 2$ and $\rho_\infty = 0.3$ (for more general correlation structures we refer to Schoenmakers (2005)). We consider at-the-money ($\theta = 3.5\%$) swaption over a period $[T_1, T_{19}]$.

In our case, the Libor transitional kernel $p^L(s,x,t,y)$ has a pronounced 'delta-shaped' form. See Fig. 1 for cross-sections $\alpha \to p^L(0, L(0), t, \alpha L(0))$ of its WKB approximations p_0^L and p_1^L, with c_0 only, and with c_0 and c_1, respectively. Because of the 'delta-shape', it is very important to find a proper density φ in (3.7) and (6.52) to make the estimators

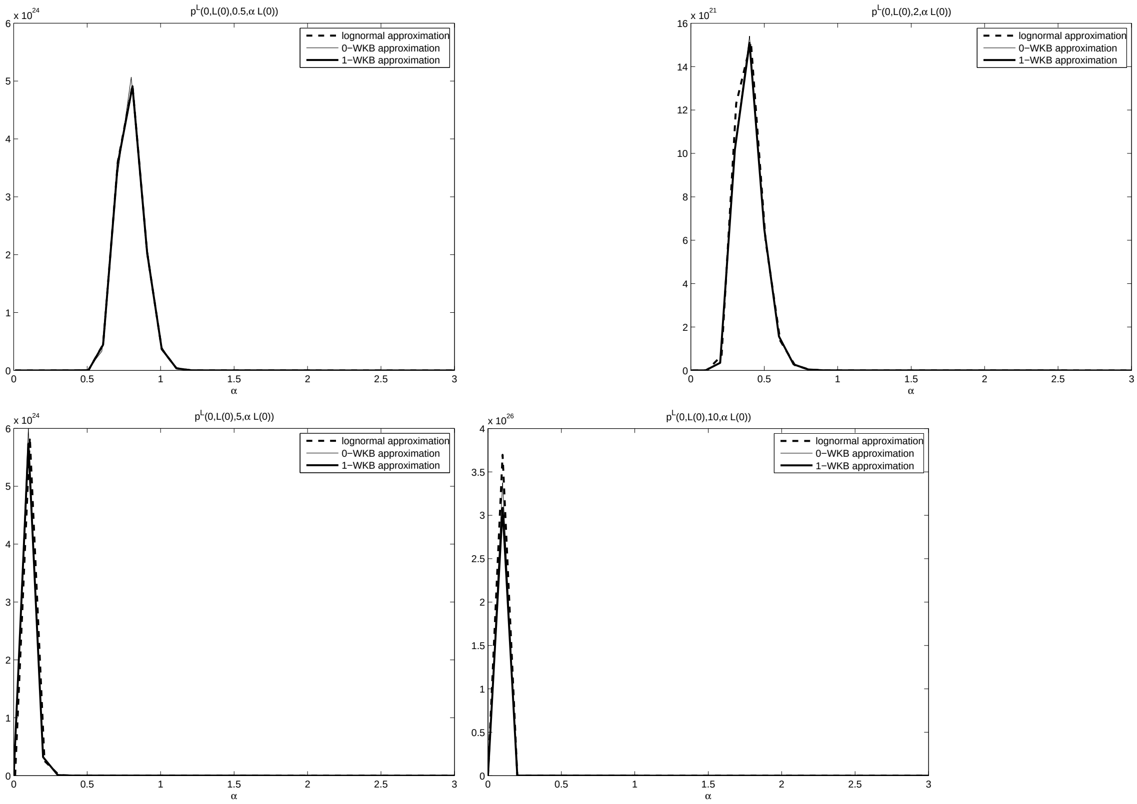

Figure 1: WKB approximations of the Libor transitional density (cross-section) $p^L(0, L(0), t, \alpha L(0))$ for different t (thin line for p_0^L, bold line for p_1^L) and lognormal approximation p_{ln}^L(dashed line).

efficient. Here we take for φ a canonical lognormal approximation of the transition kernel $p_{ln}^L(s, x, t, y)$ defined by (4.22),

$$\begin{aligned} p_{ln}^L(s, u, t, v) &= \frac{1}{\sqrt{2\pi(t-s)}^n} \prod_{i=1}^n \frac{\Gamma_{ii}^{-1}}{v_i} \times \\ &\quad \exp\left(-\frac{(\Gamma^{-1}((\log \frac{v_1}{u_1} \ldots \log \frac{v_n}{u_n}) - \mu^{ln}(s, t, x))^T)^2}{2(t-s)} \right) \end{aligned}$$

with

$$\mu_i^{ln}(s, t, x) = (t-s) \left(\frac{|\gamma_i|^2}{2} - \sum_{j=i+1}^n \frac{|\gamma_i||\gamma_j|\rho_{ij}\delta_j x_j}{1+\delta_j x_j} \right), \quad 1 \le i \le n.$$

The figure shows, that $p_{ln}^L(0, L(0), t, \alpha L(0))$ is rather close to WKB approximation. However, simulating Libors from $p_{ln}^L(0, L(0), t, \cdot)$ provides rather crude estimations for European swaptions and Deltas. In this example the bias is about 5% for prices of European swaptions and 3% for the Deltas, see Table 1 and Table 2.

Now we consider estimators (3.7) and (6.52) with payoff (6.51) for $x = L(0)$, where

$$\begin{aligned} &\varphi(x, \cdot) = p_{ln}^L(0, x, T_1, \cdot), \\ &p(x, \cdot) = p_0^L(0, x, T_1, \cdot) \quad \text{and} \quad p(x, \cdot) = p_1^L(0, x, T_1, \cdot). \end{aligned}$$

Let us denote them by $\widehat{I}_0$ and $\widehat{I}_1$, respectively. The estimates due to (3.7) and (6.52) are compared with 'exact' values of European swaptions and Delta. For the 'exact' values, we simulate M Libor trajectories of (6.49) by a log-Euler scheme with very small time steps, $\Delta t = \delta_i/10$, and take

$$\begin{aligned} \widehat{I}_{ex} &= \frac{1}{M}\sum_{m=1}^{M} u({}_mL_{T_1}^{0,x}), \\ \widehat{\frac{\partial I_{ex}}{\partial x_i}}^{(h)} &= \frac{1}{M}\sum_{m=1}^{M} \frac{u({}_mL_{T_1}^{0,x+h}) - u({}_mL_{T_1}^{0,x-h})}{2h}, \quad 1 \le i \le n. \end{aligned} \tag{6.54}$$

Analogue to (6.54), we compute $\widehat{I}_{ln}$ and $\frac{\partial \widehat{I}_{ln}}{\partial x_i}$ due to a lognormal approximation of the Libor model with transition kernel $p_{ln}^L(0, x, T_1, \cdot)$.

In Table 1 and Table 2, we give time 0 values of European swaptions and Deltas, computed via estimators $\widehat{I}_{ex}$, $\widehat{I}_{ln}$, $\widehat{I}_0$, $\widehat{I}_1$ and $\widehat{\frac{\partial I_{ex}}{\partial x_1}}^{(h)}$, $\widehat{\frac{\partial I_{ln}}{\partial x_1}}^{(h)}$, $\widehat{\frac{\partial I_0}{\partial x_1}}^{(h)}$, $\widehat{\frac{\partial I_1}{\partial x_1}}^{(h)}$, respectively, for different maturities T_1. To compute the values in the tables, we take $h = 3.5 \times 10^{-5}$, M equal to 3×10^5 and 2×10^5, respectively, to keep standard deviations within 0.5% relative to the values. As we see, the WKB approximation with only two coefficients, c_0 and c_1, provides a very close estimate of the European swaptions and Deltas, also for large maturities. The distance between $\widehat{I}_{ex}$ and $\widehat{I}_1$ and $\widehat{\frac{\partial I_{ex}}{\partial x_1}}$ and $\widehat{\frac{\partial I_1}{\partial x_1}}$ is smaller than 0.5% relative to the value.

Table 1. (values in basis points)

T_1	$\widehat{I}_{ex}$ (SD)	$\widehat{I}_{ln}$ (SD)	$\widehat{I}_0$ (SD)	$\widehat{I}_1$ (SD)
0.5	129.6(0.4)	128.9(0.4)	129.1(0.4)	128.4(0.4)
1.0	179.1(0.5)	179.4(0.5)	180.6(0.6)	178.7(0.5)
2.0	243.8(0.8)	246.0(0.8)	251.4(0.8)	245.1(0.8)
5.0	351.2(1.3)	357.8(1.3)	376.3(1.4)	349.4(1.3)
10.0	430.3(2.0)	453.3(2.2)	499.4(2.1)	430.6(1.8)

Table 2. (values in basis points)

T_1	$\widehat{\frac{\partial I_{ex}}{\partial x_1}}^{(h)}$ (SD)	$\widehat{\frac{\partial I_{ln}}{\partial x_1}}^{(h)}$ (SD)	$\widehat{\frac{\partial I_0}{\partial x_1}}^{(h)}$ (SD)	$\widehat{\frac{\partial I_1}{\partial x_1}}^{(h)}$ (SD)
0.5	2475.3(5.6)	2470.3(6.0)	2485.4(6.0)	2470.5(6.0)
1.0	2450.6(6.2)	2451.7(6.2)	2480.0(6.6)	2450.1(6.1)
2.0	2401.4(6.4)	2405.2(6.4)	2460.3(6.6)	2400.4(6.4)
5.0	2257.2(7.1)	2261.2(7.2)	2386.7(7.4)	2239.1(6.9)
10.0	2017.9(8.3)	2077.3(8.8)	2299.8(9.0)	2010.2(7.7)

References

Ait-Sahalia, Y., Closed-form Likelyhood expansions for multivariate diffusions (Preprint, 2006)

Bally, V. and Talay, D. (1996) The law of the Euler scheme for stochastic differential equations I: convergence rate of the density. *Monte Carlo Methods Appl.*, **2**, 93–128.

Bender, C., Schoenmakers, J. (2006) An iterative algorithm for multiple stopping: Convergence and stability. *Advances in Appl. Prob.*, Volume 38, Number 3 (2006), pp. 729-749

Chen, N. and Glasserman, P.: Malliavin Greeks without Malliavin Calculus. Working paper (2006).

Cont, R. and Tankov, P. (2003). *Financial Modelling with Jump processes,* Chapman & Hall.

Duffie, D., *Dynamic Asset Pricing Theory.* Princeton: Princeton University Press 2001

Fries, C., Kampen, J.: Proxy Simulation Schemes for generic robust Monte Carlo sensitivities, process oriented importance sampling and high accuracy drift approximation (with applications to the LIBOR market model), *Journal of Computational Finance,* Vol. 10, Nr. 2, 97-128, 2007.

Giles, M., and Glasserman, P. Smoking Adjoints: Fast Monte Carlo Greeks. *Risk*, vol. 19, 88-92, 2006.

Glasserman, P. and Merener, N.: Numerical Solution of Jump-Diffusion LIBOR Market Models. *Finance and Stochastics*, 7, 1-27, 2003.

Glasserman, P. and Zhao, X.: Fast Greeks by Simulation in Forward Libor Models. *Journal of Computational Finance,* Vol. 3, 1, 5-39, 1999.

Hunter, C., Jäckel, P. & Joshi, M.: Getting the Drift. *Risk* 14, 81-84 (July, 2001)

Jamshidian, F.: LIBOR and swap market models and measures. Finance and Stochastics, 1, 293-330 (1997)

Kampen, J.: *The WKB-Expansion of the fundamental solution of linear parabolic equations and its applications* . book manuscript, submitted to Memoirs of the American Mathematical Society, (downloadable at http://papers.ssrn.com/abstract=914122 (Soc. Sc. Res. Net), 2006.

Kampen, J.: Pointwise Convergent analytic expansions for densities of Feller processes (in preparation).

Karatzas, I., Shreve, S.E.: *Methods of Mathematical Finance.* Springer-Verlag Berlin Heidelberg (1998)

Kloeden, P.E., Platen, E., *Numerical solution of stochastic differential equations.* Springer Verlag Berlin, 1992.

Kolodko, A., Schoenmakers, J. Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics*, **10**(1) (2006), 27-49.

Krylov, N.V., *Lectures on Elliptic and Parabolic Equations in Hölder Spaces,* Graduate Studies in Mathematics, Vol. 12, American Mathematical Society, 1996.

Kurbanmuradov, O., Sabelfeld, K., and Schoenmakers, J. Lognormal approximations to libor market models. *Journal of Computational Finance*, Vol. 6, 1, 69-100, 2002.

Longstaff, F. A., Schwartz, R. S. Valuing American options by simulation: A simple least-square approach. *Review of Financial Studies*, **14** (2001), 113–147

Milstein, G., and Schoenmakers, J. Monte Carlo construction of hedging strategies against multi-asset European claims. *Stochastics Stochastics Rep.*, 73(1-2):125–157, 2002.

Milstein, G.N., Tretyakov, M.V., *Stochastic Numerics for Mathematical Physics*, Springer Verlag Berlin, 2004.

Pelsser, A., Pietersz, R., and van Regenmortel, M.: Fast Drift-Approximated Pricing in the BGM model. *Journal of Computational Finance,* Vol. 8(1) (2004), 93-124

Piterbarg, V.V.: Risk sensitivities of Bermuda swaptions. *International Journal of Theoretical and Applied Finance,* **7**, No. 4, (2004) 465-509.

Protter, P. *Stochastic Integration and Differential Equations.* Springer 1990.

J. Schoenmakers. *Robust Libor Modelling and Pricing of Derivative Products.* Financial Mathematics. Chapman & Hall/CRC, 2005.

Tsitsiklis, J. N. and Van Roy, B. (2001) Regression methods for pricing complex American-style options. *IEEE Trans. on Neural Networks*, **12** (4), 694–703.