Fast drift approximated pricing in the BGM model[1]

Raoul Pietersz[2], Antoon Pelsser[3], Marcel van Regenmortel[4]

First version: 18 March 2002, this version: 20 November 2002

Abstract. This paper presents a method for fast drift approximated pricing in the BGM model (Brace, Gątarek and Musiela, 1997). It is a significant addition to the predictor-corrector drift approximation method introduced by Hunter, Jäckel and Joshi (HJJ, 2001). HJJ use the drift approximation only to speed up their Monte Carlo by reducing it to single time-step simulation. We show that much more efficient numerical methods (e.g. finite differences) may be used at the cost of a minor additional assumption, *separability.* We also present a new drift approximation and propose a method to measure the accuracy of a drift approximation. This measure shows that our drift approximation is more accurate than the one of HJJ. We compare fast drift approximated pricing with Monte Carlo simulation for a Bermudan swaption, reporting a computational speed increase of a factor 10.

Key words: BGM model, predictor-corrector, drift approximation, Markov processes, separability, Feynman-Kac, Bermudan swaption

JEL Classification: G13

---

[1] We are grateful for the comments of Tomas Björk, Dick Boswinkel, Marije Elkenbracht, Lane Hughston, Martin Martens, Ton Vorst, Lars Wulf, and seminar participants at ABN AMRO Bank, ECFR and Mathematics in Finance 2002 South Africa. We thank Glyn Baker for the use of his hopscotch implementation. The first author is grateful for the financial support of the Erasmus Center of Financial Research.

[2] Erasmus Research Institute of Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (email: pietersz@few.eur.nl, tel: +31 10 4088932, fax: +31 10 4089136) and Market Risk – Modelling and Product Analysis (HQ1059), ABN AMRO Bank, P.O. Box 283, 1000 EA Amsterdam, The Netherlands

[3] Econometric Institute, Faculty of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (email: pelsser@few.eur.nl, tel: +31 10 4081270, fax: +31 10 4089169) and Nationale-Nederlanden, P.O. Box 796, 3000 AT Rotterdam, The Netherlands

[4] Product Development Group (HQ7000), ABN AMRO Bank, P.O. Box 283, 1000 EA Amsterdam, The Netherlands (email: marcel.van.regenmortel@nl.abnamro.com, tel: +31 20 3836065, fax: +31 20 3834827)

1 Introduction

The BGM framework, developed by Brace, Gątarek, Musiela [BGM97], Miltersen, Sandmann, Sondermann [MSS97] and Jamshidian [Jam96], [Jam97], is now one of the most popular models for pricing interest rate derivatives. Within the BGM framework, almost all prices are computed using Monte Carlo (MC) simulation. An advantage of MC is its applicability to almost any product. However, MC has the drawback of being rather slow computationally. In an attempt to limit MC computational time, Hunter, Jäckel and Joshi [HJJ01], [HJJ01a], Jäckel [Jäc02] (section 12.5) and Kurbanmuradov, Sabelfeld and Schoenmakers [KSS99], [KSS02] introduced predictor-corrector drift approximations. These reduce the MC to single time-step simulation.

This paper presents a significant addition to the predictor-corrector drift approximation method. We show that much more efficient numerical methods (either numerical integration or finite differences) may be used at the cost of a minor additional assumption, *separability*. The latter is a nonrestrictive requirement on the form of the volatility function. The drift approximation together with separability renders the state of the BGM model completely determined by a low-dimensional Markov process. This enables efficient implementation.

We give an example of the fast drift approximate pricing framework for Bermudan swaptions. A comparison is made with prices obtained by least-squares Monte Carlo simulation in the BGM model. This includes the use of the Longstaff and Schwartz [LoS01] method.

The computational speed increase by use of finite differences for BGM drift approximations is the main result. This paper also contains two other sub results. The first sub result is a method to measure the accuracy of a drift approximation. This is the timing inconsistency test as outlined in section 9. The second sub result is a new drift approximation using a Brownian bridge as introduced in section 3. We show that this approximation outperforms the HJJ one in the timing inconsistency test.

The outline of this paper is as follows. First, some basic notation and the most important formulas for the BGM model are stated. Second, the predictor-corrector drift approximation based on the Brownian bridge is briefly summarized. Third, the fast drift approximation pricing framework is introduced. Fourth and fifth, two fast computational techniques are discussed, numerical integration and finite differences, respectively. Sixth, the proposed framework is worked out for the one-factor case. Seventh, an example is given for the pricing of Bermudan swaptions, both for a one- and two-factor model. Eighth, a test is developed to assess the quality of a drift approximation. Ninth, conclusions are made.

2 BGM – Notation

In this section our notation of the BGM model is introduced.

Consider an arbitrage-free and complete BGM model $\mathcal{M}$[5]. Such a model $\mathcal{M}$ features N forward rates L_i, $i = 1, \ldots, N$, where forward i accrues from time T_i to time T_{i+1}, $0 < T_1 < \cdots < T_{N+1}$. Define the accrual factor δ_i to be $T_{i+1} - T_i$. Denote by $B_i(t)$ the time-t price of a discount bond expiring at time T_i. Bond prices and forward rates are linked by the relation below,

$$1 + \delta_i L_i(t) = \frac{B_i(t)}{B_{i+1}(t)}.$$

Each forward rate is driven by a d-dimensional Brownian motion $\mathbf{W}$ as follows, (d is the number of stochastic factors of the BGM model.)

$$\frac{\mathrm{d}L_i(t)}{L_i(t)} = \mu_i(t)\mathrm{d}t + \sigma_i(t) \cdot \mathrm{d}\mathbf{W}(t). \tag{1}$$

Here σ_i is the d-dimensional volatility vector. μ_i is the drift term; its form will in general depend on the choice of probability measure. Throughout this paper, we use the numeraire probability measure associated with the bond maturing at time T_{N+1}, the so called *terminal measure*. There is a specific reason why we use the terminal measure, this is explained in remark 5 of section 4. For the terminal measure, the drift term will have the following form, for $i < N$,

$$\mu_i(t, L_{i+1}, \ldots, L_N) = -\sum_{j=i+1}^{N} \frac{\delta_j L_j \sigma_j(t) \cdot \sigma_i(t)}{1 + \delta_j L_j}. \tag{2}$$

The drift term is zero for $i = N$. This simply expresses the well-known fact that a forward rate is a martingale under its associated forward measure.

For the remainder of this paper it will be useful to have stochastic differential equation (SDE) (1) in integrated form. A simple application of the Itô formula yields the following (formal) solution.

$$L_i(t) = L_i(0) e^{\int_0^t \left\{ \mu_i\left(s, L_{i+1}(s), \ldots, L_N(s)\right) - \frac{1}{2}\|\sigma_i(s)\|^2 \right\} \mathrm{d}s + \int_0^t \sigma_i(s) \cdot \mathrm{d}\mathbf{W}^{N+1}(s)}. \tag{3}$$

Lastly, we introduce the notion of 'all available forward rates at a given point in time'. Define $i(t)$ to be the smallest integer i such that $t \leq T_i$.

[5] A construction of such a model may be found in, e.g., [MuR97], [Pel00] or [BrM01].

Define $\mathbf{L}$ to consist of all forward rates that have not yet expired at time t, i.e.,

$$\mathbf{L}(t) = \big(\ L_{i(t)}(t), \ldots, L_N(t)\ \big).$$

3 Predictor-corrector drift approximation

This section discusses the predictor-corrector drift approximations introduced by Hunter, Jäckel and Joshi, [HJJ01], and Kurbanmuradov, Sabelfeld and Schoenmakers, [KSS99]. These authors introduce drift approximations for Monte Carlo simulation of SDE (1) that allows single time steps of up to twenty years, whereas normally, due to the strong state dependency of the drift term, small time steps are required. Our main result is that drift approximations plus separability yield low-dimensional tractability of the BGM model. To this end, it does not really matter what drift approximation is used. In the following, we describe a novel drift approximation based on the Brownian bridge.

An estimate is made of the forward rates $\mathbf{L}(t^*)$ given the values $\{\int_0^{t^*} \sigma_i(s) \cdot \mathrm{d}\mathbf{W}(s)\ ;\ i = 1, \ldots, N\}$ without further knowledge of the path traversed by $\mathbf{W}$. The method works recursively from forward rate N down to the first available forward rate. Forward rate N can be determined exactly, because it is fully determined by the value of $\int_0^{t^*} \sigma_N(s) \cdot \mathrm{d}\mathbf{W}(s)$, see equation (3). Assume that forward rates $j = i+1, \ldots, N$ have been estimated by $L_j^{\mathrm{D-A}}(t^*)$, respectively. We will make an estimate of forward rate i. By inspection of equation (3) it follows that this requires us to estimate the following integrals containing unknown functions $L_j(\cdot)$ with known begin and end points,

$$\int_0^{t^*} -\frac{\delta_j L_j(s)\sigma_j(s) \cdot \sigma_i(s)}{1 + \delta_j L_j(s)}\mathrm{d}s, \quad j = i+1, \ldots, N. \tag{4}$$

Here the value of $L_j(0)$ is given by today's market data and the value of $L_j(t^*)$ has been previously estimated by $L_j^{\mathrm{D-A}}(t^*)$ by recursion. Ignoring the drift term in L_j, the following approximating process L'_j is obtained,

$$\frac{\mathrm{d}L'_j(t)}{L'_j(t)} = \sigma_i(t) \cdot \mathrm{d}\mathbf{W}(t), \quad L'_j(0) = L_j(0), \quad L'_j(t^*) = L_j^{\mathrm{D-A}}(t^*). \tag{5}$$

L'_j follows a generalized geometric Brownian bridge. We will replace any term $L_j(s)$ occurring in the integral (4) with the mean $m_j(s)$ of $L'_j(s)$. See Figure 1 for an illustration. It may be shown (see Appendix A) that the

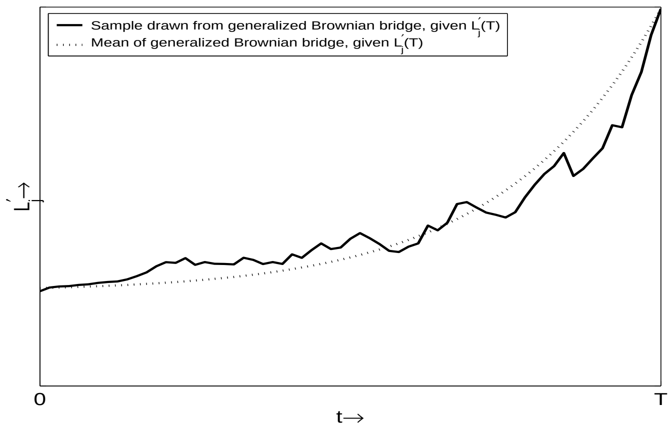

Figure 1: Proposed drift approximation; interpolation of L_j'.

mean m_j of L_j' at time t is equal to

$$\begin{aligned} m_j(t) &:= \mathbb{E}\big[\, L_j'(t) \,\big| L_j'(t^*) = L_j^{\mathrm{D-A}}(t^*) \,\big] \\ &= L_j(0)\Big(\frac{L_j^{\mathrm{D-A}}(t^*)}{L_j(0)} \Big)^{\frac{\Sigma_j^2(t)}{\Sigma_j^2(t^*)}} \exp\Big\{ \frac{1}{2}\frac{\Sigma_j^2(t)}{\Sigma_j^2(t^*)}\big(\Sigma_j^2(t^*) - \Sigma_j^2(t) \big) \Big\}, \end{aligned}$$

where

$$\Sigma_j^2(t) = \int_0^t \|\sigma_j(s)\|^2 \mathrm{d}s.$$

In this way, the integral equations (4) may be approximated by

$$\int_0^{t^*} -\frac{\delta_j m_j(s)\sigma_j(s)\cdot\sigma_i(s)}{1+\delta_j m_j(s)}\mathrm{d}s, \quad j = i+1,\ldots,N. \tag{6}$$

The sum $\sum_{j=i+1}^N$ of the expression above will be denoted by

$$\mathrm{Drift} - \mathrm{Approximation}_i\big(\, t^*, L_{i+1}^{\mathrm{D-A}}(t^*),\ldots, L_N^{\mathrm{D-A}}(t^*) \,\big).$$

Substituting this into equation (3) yields an estimate $L_i^{\mathrm{D-A}}(t^*)$ of $L_i(t^*)$.

4 Drift approximation pricing framework

The key idea in the development of the drift approximation pricing framework is to derive a sufficient condition on the volatility structure that permits the dynamics of the drift approximated forward rates process to be represented by a low-dimensional Markov process. This condition, deemed *separability*, is defined below.

Definition 1 (Separability) *A collection of instantaneous volatility functions* $\sigma_i : [0, T_i] \to \mathbb{R}^d$, $i = 1, \ldots, N$, *is called* separable *if there exists a vector valued function* $\sigma : [0, T] \to \mathbb{R}^d$ *and vectors* $\mathbf{v}_i \in \mathbb{R}^d$, $i = 1, \ldots, N$, *such that*

$$\sigma_i(t) = \mathbf{v}_i \sigma(t) \tag{7}$$

(no vector product; entry-by-entry multiplication) for $0 \le t \le T_i$, $i = 1, \ldots, N$.

Separability seems to appear quite regularly in the context of requiring a process to be Markov. We mention three examples. First, we mention Ritchken and Sankarasubramanian (RS, [RiS95], Proposition 2.1). Working in the HJM model (Heath, Jarrow and Morton, [HJM92]), RS show that separability is a necessary and sufficient condition on the volatility structure such that the dynamics of the term structure may be represented by a two-dimensional Markov process. Second, we mention the Wiener chaos expansion framework of Hughston and Rafailidis [HuR02]. In this framework any interest rate model is completely characterized by its so called *Wiener chaos expansion*. The n^{th} chaos expansion is represented by a function $\phi_n : \mathbb{R}_+^n \to \mathbb{R}$ satisfying certain integrability conditions. If all ϕ_n are separable, then the resulting interest rate model turns out to be Markovian. Third, we mention the finite dimensional Markovian realizations for stochastic volatility forward rate models. See Björk, Landén and Svensson [BLS02]. Here a necessary condition for a stochastic volatility model to have a finite dimensional Markovian realization is the following. The drift term and each component of the volatility term in the Stratonovich representation of the short rate SDE should be a sum of functions that are separable in time to expiry and the stochastic volatility driver.

We give an important example of a separable volatility function in the case of a one-factor model ($d = 1$).

Example 2 (*Mean reversion, De Jong, Driessen, Pelsser (2002)*) Following [JDP02], the instantaneous volatility may be specified as

$$\sigma_i(t) = \gamma_i e^{-\kappa(T_i - t)}. \tag{8}$$

The constant κ is usually referred to as the *mean reversion parameter*. It may be readily seen that the above specified instantaneous volatility structure is separable.

The following proposition shows that drift approximation plus separability yields low-dimensional representability.

Proposition 3 *Suppose $\mathcal{M}$ is a d-factor BGM model, for which the instantaneous volatility structure is separable. Then the drift approximated forward rates process may be represented by a d-dimensional Markovian process.*

Proof: Define the Markov process $\mathbf{X} : [0,T] \to \mathbb{R}^d$ by

$$\mathbf{X}(t) = \int_0^t \sigma(s)\mathrm{d}\mathbf{W}^{N+1}(s),$$

(entry-by-entry multiplication) where σ is as in Definition 1. Then the drift approximated process $\mathbf{L}^{\mathrm{D-A}} : [0,T] \to (0,\infty)^{n-i(t)+1}$ at time t is the unique solution to the set of $n - i(t) + 1$ non-linear equations

(9)
$$L_i^{\mathrm{D-A}}(t) = L_i(0)e^{\text{Drift-Approximation}_i(\ t, L_{i+1}^{\mathrm{D-A}}(t), \ldots, L_N^{\mathrm{D-A}}(t)\ ) - \int_0^t \frac{1}{2}\|\sigma_i(s)\|^2 \mathrm{d}s + \mathbf{v}_i \cdot \mathbf{X}(t)},$$

$i = i(t), \ldots, N$. The claim readily follows, bar two clarifying remarks:

First, the set of equations (9) has a unique solution. This is readily seen to hold as follows, using an inductive kind of argument. The last forward rate L_N is determined by the Markov factor $\mathbf{X}$ only. Forward rate i in turn is determined by the forward rates $j = i+1, \ldots, N$.

Second, the third term in the exponent of equation (9) is exactly equal to the stochastic part occurring in the BGM SDE (1), in virtue of the separability of the volatility structure:

$$\begin{aligned}\int_0^t \sigma_i(s) \cdot \mathrm{d}\mathbf{W}(s) &= \int_0^t \big(\ \mathbf{v}_i\ \sigma(s)\ \big) \cdot \mathrm{d}\mathbf{W}^{N+1}(s) \\ &= \mathbf{v}_i \cdot \mathbf{X}(t),\end{aligned}$$

where the notation of Definition 1 has been used. □

Remark 4 The vector of (drift approximated) forwards may be considered (if separability holds) as a time-dependent function of the Markov process $\mathbf{X}$, i.e.,

$$\mathbf{L}^{\mathrm{D-A}}(t) = \mathbf{f}(t, \mathbf{X}(t)),$$

for some function $\mathbf{f}$. Hunt, Kennedy and Pelsser ([HKP00], Theorem 1) showed that this is impossible to achieve for the true BGM forward rates themselves, in case of $\mathbf{X}$ being one-dimensional and under some technical restrictions. So in a way, drift approximations plus separability brings one as close as possible to low-dimensional representability in BGM.

Another essential building block for the fast drift approximated pricing framework is use of the terminal measure. This is explained in the following remark.

Remark 5 (*Choice of numeraire*) For the workings of the fast drift approximated pricing algorithm it is essential that the terminal measure is used. This is explained as follows. As proven in proposition 3, the time-t drift approximated forward rates are fully determined by $\mathbf{X}(t)$. This result holds for any measure/numeraire choice. However, for the terminal numeraire, the numeraire value at time t is fully determined by the forward rate values at time t, but this does not hold in case of for example the spot numeraire. Namely, the latter is generally determined by bond values observed at earlier times. The spot numeraire B_0 rolls its holdings over by the spot LIBOR account. Its time T_i-value is

$$B_0(T_i) = \frac{1}{\prod_{j=1}^{i} B_j(T_{j-1})}, \quad T_0 := 0.$$

Put in another way, the spot numeraire value is path dependent whereas the terminal numeraire value is not. For pricing (corollary 7 and equation (10)) it is essential that the numeraire value is known given the value of $\mathbf{X}(t)$. Therefore the fast drift approximated framework requires the use of the terminal numeraire.

This brings us in a position to define the fast drift approximated pricing framework.

Definition 6 (Framework for fast drift approximated pricing in BGM) *The framework for fast drift approximated pricing is based on the arbitrage-free pricing formula, where*

(i) the forward rates process is replaced by its drift approximated equivalent,

(ii) the volatility structure is assumed to be separable,

(iii) the terminal measure is used.

Because of the drift approximation, the framework will *not* be arbitrage-free; this is investigated in section 9.

The arbitrage-free pricing formula may be evaluated through integration or finite differences. The following two sections provide details.

5 Numerical integration

In the interest rate market the following payoff structure is quite common: A fixing is made at time T_i and the contract pays at time T_{i+1}. If a payment of $Y(T_i)$ is made at time T_{i+1}, then this is equivalent to receiving $Y(T_i)B_{i+1}(T_i)$ at time T_i.

As a Corollary, the following is obtained.

Corollary 7 *The drift approximation framework pricing formula for a T_i fixing / T_{i+1} payoff pair is given by*

$$\begin{aligned} V(0) &= B_{N+1}(0)\mathbb{E}^{N+1}\Big[\, \frac{Y(T_i)B_{i+1}(T_i)}{B_{N+1}(T_i)} \,\Big] \\ &= B_{N+1}(0)\mathbb{E}^{N+1}\Big[\, Y(T_i) \prod_{j=i+1}^{N} \big(\, 1+\delta_j L_j^{\mathrm{D-A}}(T_i) \,\big) \,\Big], \end{aligned}$$

where the drift approximated forward rates should be considered as a function of $\mathbf{X}$, i.e., $\mathbf{L}^{\mathrm{D-A}}(t) = \mathbf{L}^{\mathrm{D-A}}(t, \mathbf{X}(t))$.

Note that $\mathbf{X}(t)$ is normally distributed,

$$\mathbf{X}(t) \sim \mathcal{N}\big(\, \mathbf{0}, (\Sigma_{jj'}^2)_{j,j'=1}^d \,\big), \;\; \Sigma_{jj}^2 = \int_0^t \|\sigma_j(s)\|^2 \mathrm{d}s, \;\; \Sigma_{jj'}^2 = 0 \; (j \neq j'),$$

such that the expectation occurring in Corollary 7 reduces to the numerical evaluation of a (multi-dimensional) integral, whenever the time-T_i payoff is determined by the values of the forward rates at time T_i and not at earlier times. The numerical integration technique may be applied to caps, floors and European swaptions.

6 Finite differences

If an asset pays, at time T, an amount $Y(T)$ determined solely by the time-T values of the forward rates, then the drift approximated BGM price $V(0)$ is given by

$$V(0) = B_{N+1}(0)\mathbb{E}^{N+1}\big[\frac{Y(T)}{B_{N+1}(T)}\big],$$

where each of the stochastic variables should be considered as a function of $\mathbf{L}^{\mathrm{D-A}}(T, \mathbf{X}(T))$. Here $\mathbf{X}$ satisfies the SDE

$$\mathrm{d}\mathbf{X}(t) = \sigma(t)\mathrm{d}\mathbf{W}^{N+1}(t).$$

Define the relative price Π at time t to be

$$\Pi(t, \mathbf{X}) = \mathbb{E}^{N+1}_{\mathbf{X}(t)=\mathbf{x}}\big[\frac{Y(T)}{B_{N+1}(T)}\big].$$

Using the Feynman-Kac equation the following partial differential equation (PDE) is obtained.

$$\frac{\partial \Pi}{\partial t} + \frac{1}{2}\sum_{j=1}^{d} \|\sigma_j(t)\|^2 \frac{\partial^2 \Pi}{\partial X_j^2} = 0, \tag{10}$$

$$\Pi(T, \mathbf{X}) = \frac{Y(\ \mathbf{L}^{\mathrm{D-A}}(T, \mathbf{X})\ )}{B_{N+1}(\ \mathbf{L}^{\mathrm{D-A}}(T, \mathbf{X})\ )}.$$

This introduces the possibility of valuing for example Bermudan swaptions, trigger swaps and discrete barrier caps.

7 Example: One-factor drift approximated BGM framework

This section illustrates the framework for fast drift approximated pricing in BGM by setting it up in the special case of a one-factor model with a volatility structure as in Example 2. This structure may be written as follows,

$$\sigma_i(t) = \tilde{\gamma}_i e^{\kappa t},$$

for certain constants $\tilde{\gamma}_i$. The corresponding Markov factor X is then defined as

$$X(t) = \int_0^t e^{\kappa s} \mathrm{d}W(s).$$

$X(t)$ is normally distributed, $X(t) \sim \mathcal{N}(0, \Sigma^2(t))$, where

$$\Sigma^2(t) = \int_0^t e^{2\kappa s} \mathrm{d}s = \begin{cases} \frac{e^{2\kappa t} - 1}{2\kappa}, & \kappa \neq 0, \\ t, & \kappa = 0. \end{cases}$$

The drift approximated forward rates process is then given by

$$L_i^{\mathrm{D-A}}(t) = L_i(0) e^{\text{Drift-Approximation}_i(\ t, L_{i+1}^{\mathrm{D-A}}(t), \ldots, L_N^{\mathrm{D-A}}(t)\ ) - \frac{1}{2}\tilde{\gamma}_i^2 \Sigma^2(t) + \tilde{\gamma}_i X(t)}.$$

Prices may now be computed either by numerical integration or finite differences. In case of numerical integration, simply integrate over the normally distributed random variable $X(t)$. In case of finite differences, Feynman-Kac yields the following PDE for the price relative to the terminal bond.

$$\frac{\partial \Pi}{\partial t} + \frac{1}{2} e^{2\kappa t} \frac{\partial^2 \Pi}{\partial X^2} = 0.$$

A simple numerical example. We will evolve 5 annual ($\delta_i = 1$) forward rates over a one year period. Forward rate i accrues from year i till year $i+1$, $i = 1, \ldots, 5$. Take $L_i(0) = 7\%$, $\tilde{\gamma}_i = 25\%$, $\kappa = 15\%$, then $\Sigma^2(1) \approx 1.166196$. Suppose that after one year, the process X jumps to 1, so $X(1) = 1$. All computations are displayed in Table 1. Column (II) is determined by equation (2). To evaluate the effect of the Brownian bridge drift approximation

Table 1: A simple numerical example.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
i	$L_i(0)$	$\mu_i(0)$	$-\frac{1}{2}\tilde{\gamma}_i^2\Sigma^2(1)$	$\tilde{\gamma}_i X(1)$	drift frozen $L_i(1)$	equation $(6)_i$	predicted-corrected $L_i(1)$
5	7.00%	0.00000	-0.03644	0.25	8.67%	-0.00569	8.67%
4	7.00%	-0.00409	-0.03644	0.25	8.63%	-0.00567	8.62%
3	7.00%	-0.00818	-0.03644	0.25	8.60%	-0.00564	8.57%
2	7.00%	-0.01227	-0.03644	0.25	8.56%	-0.00562	8.53%
1	7.00%	-0.01636	-0.03644	0.25	8.53%		8.47%

over the most simple drift approximation, the 'drift frozen' forward rates (where the drift is evaluated at time zero) have been displayed in column (V), using the equation (V) = (I) exp((II) + (III) + (IV)). Then, we start with computing the predicted-corrected forward rate 5 and work back till forward rate 1. Forward rate 5 is easily computed; there are no drift terms involved. To compute the drift term integral at time 1 for forward rate 4, we compute the drift term integral of equation (6) for forward rate 5. The result is displayed in column (VI). This we may then use to compute the predicted-corrected forward rate 4 (see column (VII)), where we use the equation $(\mathrm{VII})_i = (\mathrm{I})\exp(\ \{\sum_{j=i+1}^{N}(\mathrm{VI})_j\} + (\mathrm{III}) + (\mathrm{IV})\ )$. $(\mathrm{VI})_1$ is not reported in the table because it is not used in the calculation. In general we would have to update column (VI) each time we drift-approximate a forward rate due to the presence of $\sigma_i(\cdot)$ in equation (6). In the simple setup of this example however $\sigma_i(\cdot)$ is identical for all forward rates i. Continuing, we compute the drift for forward rate 3 using only the predicted-corrected forward rates 4 and 5. And so on till all forward rates have been computed.

8 Example: Bermudan swaption

As an example of the drift approximation pricing framework, an analysis is made for Bermudan swaptions in comparison with a BGM model combined with the least-squares Monte Carlo method introduced by Longstaff and Schwartz [LoS01]. The one-factor set-up introduced in the previous section was used with zero mean reversion.

Callable Bermudan and European payer swaptions were priced in a one-factor BGM model, for various tenors and non-call periods. The zero rates were taken to be flat at 5%, the volatility of the forwards flat at 15%. The Bermudans were priced on a grid, the Europeans through numerical integration. The drift approximation PDE was solved using an explicit finite difference scheme. The explanatory variable in the least-squares Monte Carlo was taken to be the underlying swap NPV. This was regressed onto a constant and a linear term. These two basis functions yield sufficiently accurate results, because the value of a Bermudan swaption increases almost linearly with the value of the underlying swap.

Problems may possibly occur for American style derivatives in the drift approximated framework. Since the framework is not arbitrage-free, spurious early exercise may take place to collect the arbitrage opportunity. The effects of these phenomena have been analyzed by comparing the exercise boundaries of Longstaff Schwartz[6] and drift approximated BGM. In both models, the exercise rule turned out to be of the following form: Exercise whenever the underlying swap NPV S is larger than a certain value S^*, which is then defined to be the *exercise boundary*.

For a full deal description, see Table 2. Results have been summarized in Table 3. Computational times may be found in Table 4. Exercise boundaries for the 8 year deal are displayed in Figure 2, including confidence bounds on the LS boundaries[7]. We looked at exercise boundaries for other deals as well and these revealed similar pictures. The results show that the drift approximation BGM pricing framework indeed prices the Bermudan swaptions close to Longstaff Schwartz (LS). In all cases, the difference is within twice the simulation standard error. Moreover, the computational time involved is a factor 10 less. Note that the exercise boundary is calculated

[6]In case of Longstaff Schwartz, the future discounted cash flows are regressed against the underlying swap NPV with a constant and linear term, say with coefficients a and b. So the option is exercised whenever

$$S > a + bS \;\;\Leftrightarrow\;\; S > \frac{a}{1-b} =: S^*,$$

where it is assumed $b < 1$, which turns out to hold in practice. Hence the exercise boundary S^* may be computed from the regression coefficients by the above formula.

[7]The empirical covariance matrix of the regression-estimated coefficients a and b may be used to obtain the empirical variance of S^*. Denote random errors in a and b by ϵ_a and ϵ_b, respectively. Assuming these errors are relatively small, a Taylor expansion yields (ignoring second order terms)

$$S^* \approx \frac{a}{1-b}\Big(\,1 + \frac{\epsilon_a}{a} + \frac{\epsilon_b}{1-b}\,\Big).$$

We thus obtain the empirical variance of S^* (as well as its standard error). Assuming S^* is normally distributed, then a 95% confidence interval is given by plus and minus twice the standard error.

Table 2: Specification of the Bermudan swaption comparison deal.

Callable Bermudan swaption	
Market data	
Zero rates	Flat @ 5%
Volatility	Flat @ 15%
Product specification	
Tenor	Variable (2-8 Y)
Non-call period	Variable
Call dates	Semi-Annual
Pay / Receive	Pay fixed
Fixed leg properties	
Frequency	Semi-Annual
Date roll	None
Day count	Half year = 0.5
Fixed rate	5.06978% (ATM)
Floating leg properties	
Frequency	Semi-Annual
Date roll	None
Day count	Half year = 0.5
Margin	0%
Numerics	
Simulation paths	10,000
Finite difference scheme	Explicit
Longstaff Schwartz	
Explanatory variable	Swap NPV
Basis function type	Monomials
No. basis functions	2 (Constant and linear)

Table 3: Results of the Bermudan swaption comparison deal. The notation XNCY denotes a X year underlying swap with a non-call period of Y years. In case of a European swaption, it means that the swaption is exercisable exactly after Y years. All prices and standard errors are in basis points.

	Bermudan			**European**		
	Drift Approx BGM	Longstaff Schwartz	Stnd Err	Drift Approx BGM	Monte Carlo BGM	Stnd Err
2NC1	29.40	28.85	0.42	27.36	26.88	0.43
3NC1	64.33	62.78	0.83	53.78	52.92	0.83
4NC1	101.66	101.51	1.29	78.04	78.77	1.24
4NC3	44.09	43.59	0.70	42.93	42.55	0.71
5NC1	141.22	137.95	1.68	100.85	99.31	1.55
5NC3	89.25	86.75	1.34	83.08	80.83	1.36
6NC1	182.16	179.48	2.22	122.27	123.36	1.92
6NC3	134.88	136.43	2.01	120.60	123.06	2.03
6NC5	50.93	50.79	0.86	50.07	50.09	0.87
7NC1	224.40	221.38	2.61	142.93	140.66	2.19
7NC3	181.20	177.11	2.53	156.15	153.71	2.53
7NC5	101.84	100.59	1.64	97.28	96.57	1.65
8NC1	266.63	266.35	3.15	159.38	161.00	2.50
8NC3	226.55	226.94	3.14	185.20	190.98	3.08
8NC5	151.23	151.13	2.38	137.73	140.95	2.41
8NC7	54.20	53.70	0.96	52.38	53.12	0.96

Table 4: Computational times for the Bermudan swaption comparison deal for a computer with a 700 MHz processor. The notation XNCY denotes a X year underlying swap with a non-call period of Y years. In the drift approximation framework Bermudans are priced on a grid and Europeans are priced through numerical integration. All computational times are denoted in seconds.

	Bermudan		**European**	
	Drift Approx BGM	Longstaff Schwartz	Drift Approx BGM	Monte Carlo BGM
2NC1	0.4	3.0	0.0	1.9
3NC1	0.4	6.6	0.1	3.7
4NC1	0.7	11.1	0.2	6.1
4NC3	0.2	4.5	0.1	3.4
5NC1	1.4	17.3	0.6	9.1
5NC3	0.3	9.0	0.1	6.2
6NC1	2.4	24.5	0.6	12.8
6NC3	0.7	14.6	0.2	9.8
6NC5	0.2	5.8	0.0	4.8
7NC1	4.0	33.1	0.8	16.8
7NC3	1.4	21.2	0.4	13.5
7NC5	0.3	11.4	0.2	8.6
8NC1	5.6	45.9	1.2	23.9
8NC3	2.2	30.2	0.6	18.8
8NC5	0.6	18.4	0.2	13.5
8NC7	0.1	7.4	0.0	7.8

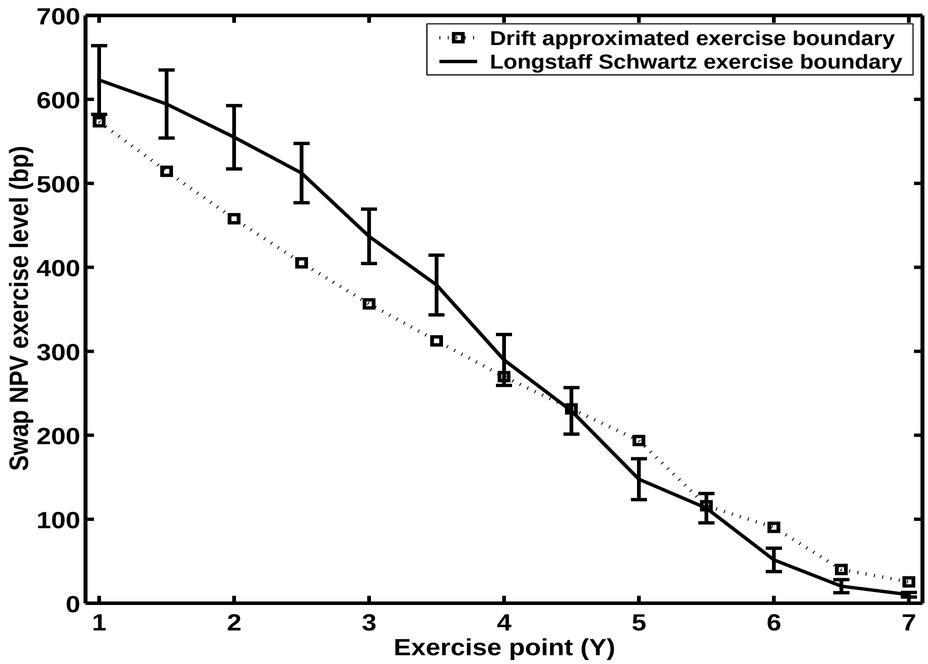

Figure 2: Exercise boundaries for 8 year deal.

Table 5: BGM pricing simulation re-run for 500,000 paths using pre-computed exercise boundaries. The standard errors for both prices were virtually the same in all cases, therefore only a single standard error is reported. All prices and standard errors are in basis points.

BGM simulation price			
	LS pre-computed exercise boundaries	D-A pre-computed exercise boundaries	Standard error
2NC1	28.63	28.62	0.06
3NC1	62.80	62.77	0.12
4NC1	99.51	99.58	0.18
5NC1	138.38	138.55	0.24
6NC1	178.08	179.41	0.30
7NC1	221.51	222.49	0.36
8NC1	263.05	265.27	0.42

slightly differently by the LS and drift approximated (D-A) approach. To determine which approach computed the best exercise boundaries, the BGM pricing simulation was re-run for 500,000 paths using the pre-computed exercise boundaries. Results may be found in Table 5. The results show that the drift approximated exercise boundaries are not worse than their Longstaff Schwartz counterparts and even slightly better[8]. Hence there is no problem with the spurious early exercise opportunities connected with the absence of no arbitrage in the fast drift approximation framework. The non-arbitrage-free issue is investigated further in the next section. This section ends with results for a 2-factor model.

2-Factor Model. We consider a 2-factor model with the same setup as above with the exception of the volatility structure, which we now take as

$$\frac{\mathrm{d}L_i(t)}{L_i(t)} = v_{i,1} dW_1^{i+1}(t) + v_{i,2} dW_2^{i+1}(t).$$

Here $|\mathbf{v}_i| = 15\%$. For a model with forward expiry structure $T_1 < \cdots < T_N$ we take the $\mathbf{v}_i \in \mathbb{R}^2$ to be

$$\mathbf{v}_i = (15\%) \left(a_i, \sqrt{1 - a_i^2} \right), \quad a_i = \frac{T_i - T_1}{T_N - T_1}.$$

This instantaneous volatility structure is purely hypothetical. It has the property that correlation steadily drops between more separated forward rates. To solve the 2-dimensional PDE version of equation (10) we used the hopscotch method, see paragraph 48.5 of [Wil98]. Results for the 2-factor model have been displayed in Table 6. In a 2-factor model (with de-correlation) the exercise decision not longer depends only on the underlying swap NPV but also on all forward swap rates. Therefore we see that prices agree more when LS regresses on all forward swap rates available in the model. The computational time of the fast drift approximated pricing 2D-grid was on average only a fourth of the Monte Carlo computational time.

9 Drift approximation accuracy test

Besides the approximation of the drift, the framework (Definition 6) contains a timing inconsistency. The inconsistency is best described by example. See Figure 3. Suppose that the underlying Markov process $\mathbf{X}$ jumps to $\mathbf{X}(2)$, say, in two years. Consider computing the value of the forwards at

[8] This does not necessarily mean that the D-A framework outperforms LS, because we only regress on the underlying swap NPV. LS may possibly yield better exercise boundaries when it is regressed onto more explanatory variables.

Table 6: 2-Factor model comparison. 50,000 paths were used for the LS simulation. 'Swap NPV only' or 'All forward rates' denote that LS regressed on only the swap NPV or on all forward swap rates, respectively. All prices and standard errors are in basis points.

	Fast Drift Approximation	LS Swap NPV only	LS All forward rates	LS Standard error
2NC1	25.45	23.27	24.64	0.2
3NC1	59.22	55.79	58.08	0.3
4NC1	94.67	89.54	93.00	0.5
5NC1	132.35	124.79	129.42	0.7
6NC1	171.41	162.89	169.76	0.9
7NC1	212.15	202.97	210.89	1.1
8NC1	252.49	242.59	251.88	1.3
9NC1	292.62	283.89	294.68	1.5

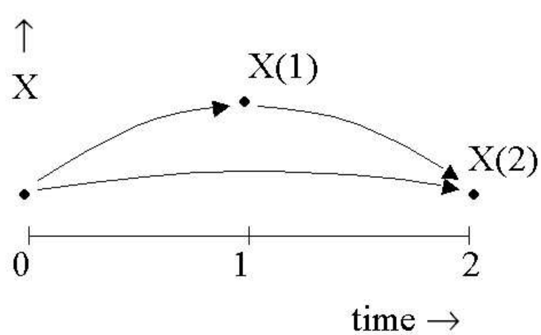

Figure 3: Timing inconsistency in framework for fast drift approximated BGM.

year 2. We could jump immediately to year 2 and calculate the forwards there. Alternatively, we could consider first calculating the forwards at time 1 (under assumption that $\mathbf{X}$ jumps to some value $\mathbf{X}(1)$) and from this point calculate the forwards at time 2 (assuming that $\mathbf{X}$ then jumps to the very same $\mathbf{X}(2)$). In general, the so computed forwards at time 2 will be different.

In a way, 'any low-dimensional approximation of BGM will exhibit this timing inconsistency'. Consider the following. Given the value of $\mathbf{X}(t)$, we cannot determine all time-t forward rates. We do know however the value of $L_N(t)$, because L_N has zero drift under the terminal measure $N+1$. The value of any other forward rate $L_i(t)$ does not solely depend on the value of $\mathbf{X}(t)$, but is dependent of the whole path that $\mathbf{X}$ traversed on the interval $[0,t]$. The framework for fast drift approximated pricing simply makes a guess as to what could be the most likely value for $L_i(t)$ given the value of $\mathbf{X}(t)$. If we start from a different initial model state (for example, if we start from the state determined by $\mathbf{X}(1)$) then almost surely our guess to the most likely value of $L_i(t)$ will be different. In this way, it is not really fair to consider this timing inconsistency, but we will nonetheless investigate it. In the following, a test will be proposed to evaluate the size of the inconsistency error.

Drift approximation accuracy test based on no-arbitrage. The accuracy test is described as follows. Consider some time T at which forwards $i,\dots,N$ have not yet expired. The framework for fast drift approximated pricing yields time-T forward rates as a function of $\mathbf{X}(T)$. Under the assumption of

(i) the model state being determined by the Markov process $\mathbf{X}$, and

(ii) the framework being arbitrage free,

the fundamental arbitrage-free pricing formula will yield values of forward rates at time $t<T$ *as a function of* $\mathbf{X}(t)$ given by the following formula[9].

$$\begin{aligned} L_i^{\mathrm{A-F}}(t,\mathbf{x}) &= \frac{1}{\delta_i}\left\{ \frac{B_i^{\mathrm{A-F}}(t)/B_{N+1}^{\mathrm{A-F}}(t)}{B_{i+1}^{\mathrm{A-F}}(t)/B_{N+1}^{\mathrm{A-F}}(t)} - 1 \right\} \\ &= \frac{1}{\delta_i}\left\{ \frac{\mathbb{E}^{N+1}\left[\frac{B_i^{\mathrm{D-A}}(T)}{B_{N+1}^{\mathrm{D-A}}(T)} \mid \mathbf{X}(t)=\mathbf{x} \right]}{\mathbb{E}^{N+1}\left[\frac{B_{i+1}^{\mathrm{D-A}}(T)}{B_{N+1}^{\mathrm{D-A}}(T)} \mid \mathbf{X}(t)=\mathbf{x} \right]} - 1 \right\} \end{aligned} \tag{11}$$

where each of the above stated T-random variables should be evaluated at $(T,\mathbf{X}(T))$. The second equality follows from $B_i^{\mathrm{A-F}}/B_{N+1}^{\mathrm{A-F}}$ being a martingale

[9]Here the notation 'A-F' is used for 'arbitrage-free' and 'D-A' is used for 'drift approximated'.

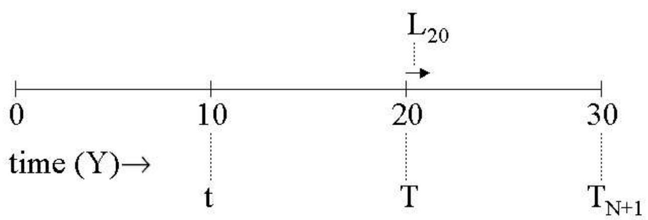

Figure 4: Inconsistency test setup.

Table 7: Quality of drift approximations: Comparison of $L_{20}^{\mathrm{A-F}}(10)$ and $L_{20}^{\mathrm{D-A}}(10)$ under different $X(10)$ moves for the volatility/mean reversion scenario 15%/10%. SD denotes the standard deviation of $X(10)$. All variables below are evaluated at time t=10.

Brownian Bridge				**HJJ**			
$X(10)$	$L_{20}^{\mathrm{A-F}}$	$L_{20}^{\mathrm{D-A}}$	$L_{20}^{\mathrm{D-A}} - L_{20}^{\mathrm{A-F}}$ (bp)	$X(10)$	$L_{20}^{\mathrm{A-F}}$	$L_{20}^{\mathrm{D-A}}$	$L_{20}^{\mathrm{D-A}} - L_{20}^{\mathrm{A-F}}$ (bp)
−SD	3.75%	3.81%	5.11	−SD	3.74 %	3.81 %	7.17
−SD/2	4.23%	4.27%	4.03	−SD/2	4.19 %	4.27 %	7.94
0	4.77%	4.79%	2.37	0	4.70 %	4.79 %	8.81
+SD/2	5.38%	5.38%	-0.05	+SD/2	5.28 %	5.38 %	9.79
+SD	6.07%	6.03%	-3.47	+SD	5.92 %	6.03 %	10.91

by assumption of no arbitrage. The so obtained ‘arbitrage-free’ forward rates $L_i^{\mathrm{A-F}}(t, \mathbf{x})$ may then be compared with forward rates $L_i^{\mathrm{D-A}}(t, \mathbf{x})$ obtained by drift approximation.

Numerical results for drift approximation test. The inconsistency test was performed under the following setup. Ten annual forward rates were considered where forward rate i accrued from year i to $i+1$, for $i = 20, \ldots, 29$. Under the notation of the previous section, t was taken to be 10 years, T was taken to be 20 years and T_{N+1} was taken to be 30 years. See also Figure 4. $L_i(0)$ was taken to be 5% and mean reversion κ was varied at 0%, 5% and 10%. The $\tilde{\gamma}_i$ were chosen such that the corresponding caplet volatility was equal to some general volatility level v, which was varied at 10%, 15% and 20%. Let SD denote the standard deviation of $X(10)$. $X(10)$

Table 8: Quality of drift approximations: Maximum of $|L_{20}^{\mathrm{A-F}}(10)-L_{20}^{\mathrm{D-A}}(10)|$ over $X(10)$ moves $0, \pm\mathrm{SD}/2, \pm\mathrm{SD}$ for different volatility/mean reversion scenarios. SD denotes the standard deviation of $X(10)$. Differences are denoted in basis points.

Brownian Bridge			
Mean reversion	Volatility level v 10%	15%	20%
0%	2.97	9.34	28.73
5%	2.56	8.21	19.46
10%	1.46	5.11	12.56

HJJ			
Mean reversion	Volatility level v 10%	15%	20%
0%	2.86	8.60	37.45
5%	2.32	12.29	53.85
10%	1.69	10.91	44.59

moves were considered for $0, \pm\mathrm{SD}/2, \pm\mathrm{SD}$. For the volatility/mean reversion scenario 15%/10% the results may be found in Table 7. The comparison is only reported for L_{20} because this forward rate contains the most drift terms, and therefore its corresponding error is the largest amongst $i = 20, \ldots, 29$. Note that the error for L_{29} is always zero as it is fully determined by X. In Table 8 the maximum error (over the five considered $X(10)$ moves) between $L_{20}^{\mathrm{A-F}}(10)$ and $L_{20}^{\mathrm{D-A}}(10)$ is reported.

The test was performed for both the Brownian bridge drift approximation and the HJJ drift approximation, as outlined in [HJJ01]. The results show that the Brownian bridge outperforms HJJ in the timing inconsistency test.

The inconsistency test results show that for less volatile market scenarios, the drift approximation performs very accurately with errors only up to a few basis points. For more volatile market scenarios the approximation becomes worse. But for realistic yield curve and forward volatility scenarios there are no problems with respect to pricing, see section 8. The approximation worsening for more volatile scenarios is what may be expected from the nature of the drift approximations; as the 'model dimensions' increase, the drift approximation will break up. With *model dimensions* we mean either volatility level, tenor of deal, difference between forward index i and N or time zero forward rates etc. Care should be taken in the application of the drift approximated framework for BGM that the market scenario does not violate the realm where the drift approximation is reasonably valid.

10 Conclusions

We have introduced a fast approximate pricing framework as an addition to the predictor-corrector drift approximation introduced by Hunter, Jäckel and Joshi [HJJ01]. HJJ use the drift approximation only to speed up their Monte Carlo by reducing it to single time-step simulation. We have shown that, at a slight cost, instead much faster computational methods may be used, such as numerical integration or finite differences. The additional cost is a nonrestrictive assumption, namely separability of the volatility function. The proposed drift approximation framework was applied to the pricing of Bermudan swaptions. It yielded very accurate prices at much lower computation times.

References

[BGM97] Brace, A., Gątarek, D., Musiela, M.: The Market Model of Interest Rate Dynamics. *Mathematical Finance* **7** (2) 127-155 (1997)

[BLS02] Björk, T., Landén, C., Svensson, L.: Finite Dimensional Markovian Realizations for Stochastic Volatility Forward Rate Models. *Working paper* (2002)

[BrM01] Brigo, D., Mercurio, F.: Interest Rate Models. Theory and Practice. *Berlin: Springer-Verlag* (2001)

[HJJ01] Hunter, C.J., Jäckel, P., Joshi, M.S.: Drift Approximations in a Forward-Rate-Based LIBOR Market Model. *Working Paper* (2001) *Downloadable from:* `www.rebonato.com/MarketModelPredictorCorrector.pdf`

[HJJ01a] Hunter, C.J., Jäckel, P., Joshi, M.S.: Getting the drift. *Risk Magazine* (July 2001)

[HJM92] Heath, D., Jarrow, R., Morton, A.: Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica* **60** 77-105 (1992)

[HKP00] Hunt, P., Kennedy, J., Pelsser, A.A.J.: Markov-functional Interest Rate Models. *Finance and Stochastics* **4** 391-408 (2000)

[HuR02] Hughston, L.P., Rafailidis, A.: A Chaotic Approach to Interest Rate Modelling. *London: King's College London Working Paper* (2002)

[Jäc02] Jäckel, P.: Monte Carlo Methods in Finance. *Chichester: John Wiley & Sons* (2002)

[Jam96] Jamshidian, F.: LIBOR and Swap Market Models and Measures. *London: Sakura Global Capital Working Paper* (1996)

[Jam97] Jamshidian, F.: LIBOR and Swap Market Models and Measures. *Finance and Stochastics* **1** 293-330 (1997)

[JDP02] De Jong, F., Driessen, J., Pelsser, A.: Libor Market Models versus Swap Market Models for Pricing of Interest Rate Derivatives: An Empirical Analysis. *European Finance Review* **5** (3) 201-237 (2002)

[KaS91] Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. 2^{nd} edition. New York: Springer (1991)

[KSS99] Kurbanmuradov, O., Sabelfeld, K., Schoenmakers, J.: Lognormal Random Field Approximations to LIBOR Market Models. *Berlin: Weierstrass Institute Working Paper* (1999)
Downloadable from:
`www.wias-berlin.de/publications/preprints/481/document.pdf`

[KSS02] Kurbanmuradov, O., Sabelfeld, K., Schoenmakers, J.: Lognormal Approximations to Libor Market Models. *Journal of Computational Finance* **6** (1) (2002)

[LoS01] Longstaff, F.A., Schwartz, E.S.: Valuing American Options by Simulation: A Simple Least-squares Approach. *The Review of Financial Studies* **14** (1) 113-147 (2001)

[MSS97] Miltersen, K.R., Sandmann, K., Sondermann, D.: Closed Form Solutions for Term Structure Derivatives with Log-normal Interest Rates. *Journal of Finance* **52** (1) 409-430 (1997)

[MuR97] Musiela, M., Rutkowski, M.: Continuous-time Term Structure Models: Forward Measure Approach. *Finance and Stochastics* **1** 261-292 (1997)

[Pel00] Pelsser, A.A.J.: Efficient Methods for Valuing Interest Rate Derivatives. *Berlin: Springer-Verlag* (2000)

[RiS95] Ritchken, P., Sankarasubramanian, L.: Volatility Structures of the Forward Rates and the Dynamics of the Term Structure. *Mathematical Finance* **5** (1) 55-72 (1995)

[Wil98] Wilmott, P.: Derivatives: The Theory and Practice of Financial Engineering. *Chichester: John Wiley & Sons* (1998)

Appendix A: Mean of generalized geometric Brownian bridge

In this section, the time-t mean of the process L'_j defined in equation (5) is determined. Equivalently, we may determine the time-t mean of the process Y, given by

$$\frac{\mathrm{d}Y(t)}{Y(t)} = \sigma(t) \cdot \mathrm{d}\mathbf{W}(t), \;\; Y(0) = y_0, \;\; Y(t^*) = y^*.$$

(Compare with equation (5).) The solution of Y (unconditional of time-t^*) is given by

$$Y(t) = y_0 e^{X(t) - \frac{1}{2}\Sigma^2(t)},$$

where

$$X(t) := \int_0^t \sigma(s) \cdot \mathrm{d}\mathbf{W}(s), \;\; \Sigma^2(t) := \int_0^t \|\sigma(s)\|^2 \mathrm{d}s.$$

Note that

$$\{ \; \omega \in \Omega \; ; \; Y(t^*) = y^* \; \} = \{ \; \omega \in \Omega \; ; \; X(t^*) = \log(y^*/y_0) + \frac{1}{2}\Sigma^2(t^*) =: x^* \}.$$

According to the martingale time change theorem, for example Theorem 4.6 of [KaS91], we have that $X(\tau(\cdot))$ is a Brownian motion, where the time change τ is defined by

$$\tau(t) = \inf\{s \geq 0; \; \Sigma^2(t) > s\}.$$

Working in the time-changed time coordinates, $X(\cdot)|X(\tau^*) = x^*$ will be a standard Brownian bridge, and so, according to section 5.6.B of [KaS91],

$$X(\tau)\big|X(\tau^*) = x^* \quad \sim \quad \mathcal{N}\big(\; \frac{\tau}{\tau^*}x^*, \tau - \frac{\tau^2}{\tau^*} \; \big).$$

Back in the original time coordinates, this translates to

$$X(t)\big|X(t^*) = x^* \quad \sim \quad \mathcal{N}\big(\; \frac{\Sigma^2(t)}{\Sigma^2(t^*)}x^*, \Sigma^2(t) - \frac{(\Sigma^2(t))^2}{\Sigma^2(t^*)} \; \big).$$

With this, we may evaluate the mean of $Y(t)|Y(t^*) = y^*$ to be

$$\mathbb{E}\big[\; Y(t) \; \big| \; Y(t^*) = y^* \; \big] = y_0 \Big(\; \frac{y^*}{y_0} \; \Big)^{\frac{\Sigma^2(t)}{\Sigma^2(T)}} \exp\Big\{ \; \frac{1}{2}\frac{\Sigma^2(t)}{\Sigma^2(T)} \big(\; \Sigma^2(T) - \Sigma^2(t) \; \big) \; \Big\},$$

where the following simple rule has been used, $\mathbb{E}[e^Z] = e^{\beta + \tau^2/2}$ whenever Z is normally distributed, $Z \sim \mathcal{N}(\beta, \tau^2)$.