Bias-Free Joint Simulation of Multi-Factor Short Rate Models and Discount Factor

Marcio Ferrão Lopes

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

July 24, 2018

MPhil in Mathematical Finance,
University of Cape Town.

University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy to the University of Cape Town. It has not been submitted before for any degree or examination to any other university.

Marcio Lopes

July 24, 2018

Abstract

This dissertation explores the use of single- and multi-factor Gaussian short rate models for the valuation of interest rate sensitive European options. Specifically, the focus is on deriving the joint distribution of the short rate and the discount factor, so that an exact and unbiased simulation scheme can be derived for risk-neutral valuation. We see that the derivation of the joint distribution remains tractable when working with the class of Gaussian short rate models.

The dissertation compares three joint and exact simulation schemes for the short rate and the discount factor in the single-factor case; and two schemes in the multi-factor case. We price European floor options and European swaptions using a two-factor Gaussian short rate model and explore the use of variance reduction techniques. We compare the exact and unbiased schemes to other solutions available in the literature: simulating the short rate under the forward measure and approximating the discount factor using quadrature.

Acknowledgements

I would like to thank my supervisors, Tom McWalter and Jörg Kienitz, for taking time out of their busy schedules to support my research. Thank you for being available, even at unusual hours.

I would also like to thank Mario Giuricich for playing an instrumental mentorship role in my academic career. Thank you for your selfless guidance and support.

I would also like to acknowledge David Taylor for his support throughout the year.

Lastly, I would like to thank my friends and family, who supported and fuelled me throughout my studies.

Contents

List of Figures

Chapter 1

Introduction

A central problem in mathematical finance is risk-neutral option valuation. To solve this problem, we need to estimate both the option payoff and numéraire. For interest rate sensitive payoffs, the numéraire, or the discount factor, will be correlated to the short rate. In order to honour this correlation, these two quantities need to be jointly simulated to ensure an unbiased price estimate. This dissertation presents exact and bias-free joint simulation schemes for multi-factor short rate models and the discount factor, which can be used for risk-neutral valuation of interest rate sensitive instruments.

Chapter 1 begins by examining interest rate modelling paradigms, with a primary focus on short rate models. The appeal of multi-factor models is then presented, and we include an overture to forward rate models. Chapter 1 then clearly details the problem and contextualizes the need for bias-free simulation.

Chapter 2 addresses joint and bias-free simulation methods for single-factor short rate models, while Chapter 3 extends these to multi-factor models. Chapter 4 concludes the dissertation with the Monte Carlo pricing of European options using the bias-free simulation methods presented, as well as alternative approaches suggested by the literature.

1.1 Interest Rate Modelling Paradigms

1.1.1 Endogenous Short-Rate Term Structure Models

In 1977, Vašíček developed an explicit characterization of the term structure (TS) of interest rates in an efficient market. Vašíček (1977) proposed that the instantaneous spot rate (r) follows the so-called Ornstein-Uhlenbeck process under the risk-neutral measure $\mathbb{Q}$

$$dr(t) = \kappa\left(\vartheta - r(t)\right)dt + \sigma dW(t), \quad r(0) = r_0. \tag{1.1}$$

Under this model, the process converges to its long-term mean ϑ with velocity κ, while its variance does not explode. For $\kappa > 0$, the Ornstein-Uhlenbeck process possesses a stationary terminal distribution. In addition, the stochastic differential equation (SDE) is linear and can be solved explicitly which leads to closed-form solutions for bond pricing purposes; a useful feature for calibration and risk-management.

The model is also a Markov process with normally distributed increments. This implies a Gaussian distribution for the short rate; which in turn implies that the short rate can assume negative values with positive probability. For this reason, the model fell out of favour for not being compatible with market-implied distributions at the time.

Later, Cox, Ingersoll and Ross (1985) specified a diffusion process model which retained the mean-reversion properties of the Vašíček model, but which constrained the short rate to the positive domain for $(\kappa, \vartheta, \sigma)$ within a suitable region known as the Feller condition. The short rate is assumed to follow the SDE

$$dr(t) = \kappa\left(\vartheta - r(t)\right)dt + \sigma\sqrt{r(t)}dW(t), \quad r(0) = r_0, \tag{1.2}$$

$$2\kappa\theta > \sigma^2. \tag{1.3}$$

This characterized the instantaneous short rate by a non-central chi-squared distribution. The model maintains analytical tractability but is less tractable than the Vašíček (1977) model. In particular, this means that bias-free joint simulation of the short rate and the discount factor becomes intractable (Glasserman, 2003, pg. 129). In addition, whereas European bond-option prices under Gaussian models feature the instantaneous rate only implicitly through the bond price, under the Cox, Ingersoll and Ross model (and all other square-root models) $r(t)$ appears explicitly (Brigo and Mercurio, 2006). This may be undesirable for products that do not directly depend on r (Brigo and Mercurio, 2006, pg. 82).

1.1.2 Exogenous Short-Rate Term Structure Models

A feature common to both the Vašíček (1977) and Cox, Ingersoll and Ross (1985) models is that they are *endogenous* in the sense that at $t = 0$, the interest rate curve is an output of the model. Therefore, if we wish to incorporate the market-observable initial zero curve, we would need to optimise for $(\kappa, \vartheta, \sigma)$ such that the $t = 0$ model curve best matches the market curve. This is not ideal since (a) three parameters may not satisfactorily reproduce a given term structure, and (b) certain shapes, such as an inverted curve, can never be reproduced (Brigo and Mercurio, 2006).

To remedy this, Hull and White (1990), extended these models with the inclusion of deterministic time-dependent parameters $\kappa(t)$, $\vartheta(t)$ and $\sigma(t)$. Using these so-called *exogenous* TS models allowed the observable TS of interest rates and TS of volatilities (at $t = 0$) to be specified as an input, and so allowed for better calibration to option data.

In the literature, this extension is often parametrized as

$$dr(t) = \kappa(t)\left(\vartheta(t) - r(t)\right)dt + \sigma(t)dW_t, \quad r(0) = r_0, \tag{1.4}$$

however, the parametrisation,

$$dr(t) = \left(\theta(t) - \alpha(t)r(t)\right)dt + \sigma(t)dW_t, \quad r(0) = r_0, \tag{1.5}$$

is also popular.

When fitting the TS of interest rates under this model, the perfect fitting of volatility can be 'dangerous' and must be carefully dealt with (Brigo and Mercurio, 2006, pg. 73). Brigo and Mercurio give two reasons for this: (a) not all volatilities quoted are significant (liquidity issues may make these quotes insignificant or unreliable), and (b) the volatility structures implied by (1.4) are unlikely to be realistic or comply with typical market shapes. The latter point was in fact asserted by Hull and White (1995) themselves, and is the subject of their Hull and White (1994) work, where $\vartheta(t)$ was assumed to be a deterministic and time-dependent function, while κ and σ were kept constant. The SDE for such a model is given by

$$dr(t) = \kappa\left(\vartheta(t) - r(t)\right)dt + \sigma dW_t, \quad r(0) = r_0. \tag{1.6}$$

We refer to Hull and White's extension of the Vašíček model (1.4), where $(\kappa, \vartheta, \sigma)$ are all time-dependent parameters, the extended Vašíček model, and the extension (1.6), where only ϑ (or θ) is time-dependent, the classical Hull-White model.

In 1991, Black and Karasinski proposed that the natural log of the short rate follows an Ornstein-Uhlenbeck process with time-dependent parameters ($\kappa(t)$, $\vartheta(t)$, $\sigma(t)$). This ensured that interest rates remained positive, and was seen as a natural choice given that market formulas for caps and swaptions were based on the assumption of log-normal rates (Brigo and Mercurio, 2006, pg. 83). The model is specified with SDE

$$d\ln r(t) = \kappa(t)\left(\vartheta(t) - \ln r(t)\right)dt + \sigma(t)dW(t), \quad r(0) = r_0. \tag{1.7}$$

Under this model, the short rate follows a log-normal distribution and this means that analytical tractability is lost altogether.

Since the 2008/09 financial crisis, negative interest rates appeared as central banks struggled to stimulate economies and control inflation. In particular, central banks in Europe and Japan lowered interest rates into negative territory, and this transmitted directly to other instruments, such as government bonds (Ueno, 2017). As of December 2017, negative-yielding sovereign debt stands at an estimated \$9.7 trillion (Fitch Ratings, 2017), and at the time of writing CHF (Swiss Franc) trades at a negative overnight rate. This has helped lead to a resurgence of Gaussian short rate (GSR) models.

1.1.3 Multi-Factor Models

For all of the 'single-factor' short rate models presented in Section 1.1.1 and Section 1.1.2, a single source of noise drives the entire evolution of the yield curve through the basic bond price relation

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)\middle|\mathcal{F}_t\right], \tag{1.8}$$

where $P(t,T)$ denotes the price of a zero-coupon bond at time t and with payout of 1 at maturity, T. $\mathbb{Q}$ denotes the risk-neutral measure; and F_t is the filtration which defines the market at time t.

A leading drawback of this approach is that it implies a perfect correlation between LIBOR rates. Additionally, single-factor models may not sufficiently capture all of the observed variability in market yield curves. This is the motivation for multi-factor models, which attempt to more realistically capture market yield curve dynamics by using multiple stochastic factors to drive the short rate process.

In single-factor models, instantaneous shocks to forward rates are perfectly correlated because of the single Brownian motion driving factor — a shock to the yield curve at time t is transmitted equally across all maturities on the yield curve (Brigo and Mercurio, 2006). Multi-factor models, on the other hand, allow for the decorrelation of the yield curve, which is in line with what is observed in reality, where multiple market players operate at different segments of the yield curve.

Single-factor models still prove useful when derivative payoffs depend only on a single interest rate and not on the correlation across different rates. Furthermore, single-factor models may be useful for risk-management purposes, where the rates that affect the payoff are close enough that a perfect correlation approximation is acceptable (Brigo and Mercurio, 2006). In general, however, multi-factor models are more useful when correlation plays an important role, when higher precision is needed, or for valuing exotic options (e.g. barrier and Bermudan options).

To study the correlation term structures in single- and multi-factor models, we examine a class of short rate models known as affine term structure (ATS) models, although the results hold for all single- and multi-factor models in general.

Definition 1.1 (Affine term structure). A short rate model is said to possess affine term structure if bond prices can be written as

$$P(t,T) = \exp(A(t,T) - B(t,T)r(t)), \tag{1.9}$$

where $A(t,T)$, $B(t,T)$ are (sufficiently regular) deterministic functions.

For a d-factor affine term structure model, the definition requires use of a d-dimensional row-vector valued function $B(t,T)^\top$, so that bond prices are linear functions of the state variables $x(t)$[1].

Under an ATS model, the continuously compounded spot rate $R(t,T)$ is itself an affine transform of $r(t)$

$$R(t,T) = -\frac{A(t,T)}{T-t} + \frac{B(t,T)}{T-t}r(t) =: a(t,T) + b(t,T)r(t), \tag{1.10}$$

so that the correlation between rates $R(t,T_1)$ and $R(t,T_2)$ ($T_1 < T_2$) is 1. On the other hand, for a two-factor model we have

$$\begin{aligned} R(t,T) &= -\frac{A(t,T)}{T-t} + \frac{B_1(t,T)}{T-t}x_1(t) + \frac{B_2(t,T)}{T-t}x_2(t) \\ &=: a(t,T) + b_1(t,T)x_1(t) + b_2(t,T)x_2(t). \end{aligned} \tag{1.11}$$

In this case, we have that the correlation between rates $R(t,T_1)$ and $R(t,T_2)$ is not necessarily 1, due to the dependency on the correlation between the state variables $x(t)$

$$\begin{aligned} &\text{Corr}\left(R(t,T_1),\, R(t,T_2)\right) \\ &\quad = \text{Corr}\left(b_1(t,T_1)x_1(t) + b_2(t,T_1)x_2(t),\, b_1(t,T_2)x_1(t) + b_2(t,T_2)x_2(t)\right). \end{aligned} \tag{1.12}$$

Figure 1.1 plots the forward rate correlation term structures for a two-factor GSR model, where $\rho(t,T_1,T_2)$ denotes the time t instantaneous correlation between $f(t,T_1)$, $f(t,T_2)$. It is assumed that the correlation structure is *time-stationary*, in the sense that $\rho(t,T_1,T_2)$ does not depend explicitly on t, but only the length of the intervals $T_1 - t$ and $T_2 - t$. Such an assumption is generally strongly preferred for practical purposes (Andersen and Piterbarg, 2010).

A natural question to ask, then, is how many factors are appropriate? Empirical studies by Jamshidian and Zhu (1996) show that, when one analyses the variability of interest rates using principal components, one component explains 68%–76%, two components 85%–90%, and three components capture 93%–94% of the variability in the market yield curve under the real-world measure. This suggests that

[1] For a single-factor model, the scalar $r(t)$ is the only state variable.

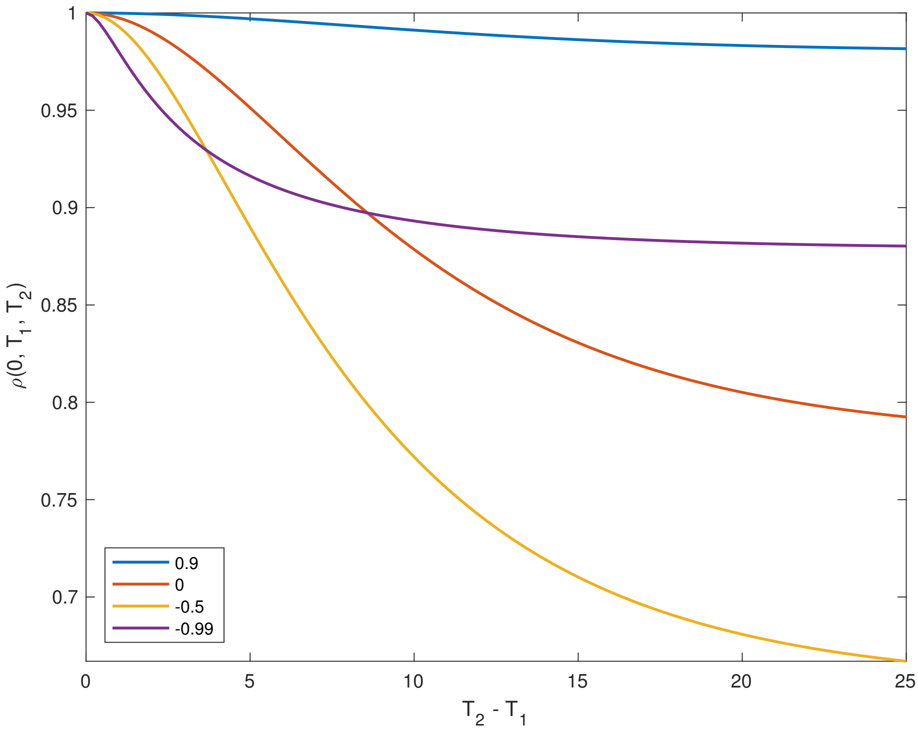


Fig. 1.1: Example correlation term structures for a two-factor GSR model for various correlations between state variables $x_1(t)$ and $x_2(t)$. Parameters used: $t = 0$, $T_1 = 0.1$, $\kappa_1 = 0.1$, $\kappa_2 = 0.25$, $\sigma_1 = 0.025$, $\sigma_2 = 0.02$.

a two- or three-factor model maintains a sufficient balance between accuracy and model simplicity.

An example of a multi-factor model we work with is Brigo and Mercurio's (2006) so-called two-additive-factor Gaussian model (G2++) (Section 3.1), which assumes that the instantaneous short rate has dynamics under $\mathbb{Q}$

$$r(t) = x_1(t) + x_2(t) + \varphi(t), \quad r(0) = r_0, \tag{1.13}$$

where $x_1(t)$ and $x_2(t)$ satisfy

$$\begin{aligned} dx_1(t) &= -\kappa_1 x_1(t)dt + \sigma_1 dW_1(t), \quad x_1(0) = 0, \\ dx_2(t) &= -\kappa_2 x_2(t)dt + \sigma_2 dW_2(t), \quad x_2(0) = 0, \end{aligned} \tag{1.14}$$

and where $dW_1(t)dW_2(t) = \rho dt$. The function φ is deterministic and selected such that the model fits the initial TS of interest rates.

1.1.4 Heath-Jarrow-Morton Framework

In Section 1.1.3 we saw that it can be unreasonable to assume that the short rate is the only explanatory variable for the evolution of the yield curve. While multi-factor models overcome some of these disadvantages, they are also more difficult to interpret; especially when the volatilities for each of the factors are calibrated against the results of a principal component analysis (Ouwehand, 2017).

Under the Heath, Jarrow and Morton (1992) (HJM) approach, the dynamics for an entire family of infinitely many forward rates are specified as (infinitely many) SDEs. That is, for each maturity $T \geq 0$, assume that the forward rate f has $\mathbb{Q}$-dynamics

$$\begin{aligned} df(t,T) &= \alpha_f(t,T)\,dt + \sigma_f(t,T)^\top\,dW_t, \quad t \leq T, \\ f(0,T) &= f^M(0,T), \end{aligned} \tag{1.15}$$

where W_t is an adapted d-dimensional $\mathbb{Q}$-Brownian motion. It is assumed that $\sigma_f(t,T)^\top$ is a d-dimensional adapted and sufficiently regular[2] stochastic process. It is also useful to note that for every maturity T, each SDE has an initial condition $f(0,T) = f^M(0,T)$ so that the entire yield curve is automatically fitted to the initial term structure of interest rates.

The resulting class of models is quite broad, and it is for this reason that it is referred to as a framework as opposed to a particular model. In particular, for a specific choice of $\sigma_f(t,T)$ the resulting HJM model coincides with some of the short rate models already examined in Section 1.1.2.

While we have infinitely many SDEs, we have assumed only finitely many sources of information in the economy, so as to not make this sheer-dimensionality unmanageable.

Given $\alpha_f(t,T)$ and $\sigma_f(t,T)$, we can solve the SDE in (1.15) from which we have the entire TS at all times and maturities; and therefore we have the entire TS of bond prices via $P(t,T) = \exp(-\int_t^T f(t,s)ds)$. However, the dynamics in (1.15) are not necessarily arbitrage-free. Heath, Jarrow and Morton (1992) proved that in order for a unique equivalent martingale measure to exist, $\alpha_f(t,T)$ cannot be specified arbitrarily, but must rather be determined by a suitable transform of $\sigma_f(t,T)$.

Proposition 1.2 (HJM Drift Condition). *Under the risk-neutral measure $\mathbb{Q}$, the process $f(t,T)$ is governed by SDE*

$$df(t,T) = \sigma_f(t,T)^\top \int_t^T \sigma_f(t,u)\,du\,dt + \sigma_f(t,T)^\top\,dW_t, \quad t \leq T. \tag{1.16}$$

[2] i.e. bounded and jointly measurable.

This is labelled Lemma 4.4.1 in Andersen and Piterbarg (2010), and the proof is therein. Proposition 1.2 states that once the volatility structure $\sigma_f(t,T)$ is specified, the drift of the forward rates is automatically specified under $\mathbb{Q}$, and this is often considered as the main insight of Heath, Jarrow and Morton. As a final remark, note that while under the HJM framework forward rates $f(0,T) = f^M(0,T)$ are specified exogenously, $\sigma_f(t,T)$ is free to be set from empirical studies or from market price calibration (Andersen and Piterbarg, 2010, pg. 183).

Markovianity

In contrast to short rate models, we can easily let $\sigma_f(t,T)$ (and $\alpha_f(t,T)$) depend on past history, and hence only a restricted class of volatilities implies a Markov short rate process. To see this, Andersen and Piterbarg (2010, pg. 184) examine the path-dependent term $D(t) = \int_0^t \sigma_f(u,t)^\top dW(u)$ in

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma_f(u,t)^\top \int_u^t \sigma_f(u,s)\,ds\,du + \int_0^t \sigma_f(u,t)^\top dW(u), \quad (1.17)$$

and note that unless the bracketed term in

$$D(T) = D(t) + \int_t^T \sigma_f(u,T)^\top dW(u) + \left\{ \int_0^t \sigma_f(u,T)^\top dW(u) - D(t) \right\} \quad (1.18)$$

is deterministic, $\mathbb{E}[D(T)\,|\,D(t)] \neq \mathbb{E}[D(T)\,|\,\mathcal{F}_t]$.

Therefore, in general, computationally expensive methods such as a non-recombining lattice may be required to price interest rate sensitive instruments (Brigo and Mercurio, 2006, pg. 184).

Lemma 1.3. *Consider the special case where $\sigma_f(t,T) = \xi(t)\varphi(T)$, where $\xi : \mathbb{R} \to \mathbb{R}^{d\times d}$ can take any sign and $\varphi : \mathbb{R} \to \mathbb{R}^{d\times 1}$. Then the short rate satisfies an SDE of the type*

$$dr(t) = [\theta(t) - \alpha(t)r(t)]\;dt + \sigma_r(t)^\top dW(u), \quad (1.19)$$

and is therefore Markov.

Proof. For $\sigma_f(t,T) = \xi(t)\varphi(T)$, (1.17) reduces to

$$r(t) = f(0,t) + \varphi(t)^\top \int_0^t \xi(u)^\top \xi(u) \int_u^t \varphi(s)\,ds\,du + \varphi(t)^\top \int_0^t \xi(u)^\top dW(u). \quad (1.20)$$

Next, by defining $\gamma(t) := f(0,t) + \varphi(t)^\top \int_0^t \xi(u)^\top \xi(u) \int_u^t \varphi(s)\,ds\,du$, the SDE for $r(t)$ is given by

$$dr(t) = \left[\gamma'(t) + \varphi'(t)^\top \int_0^t \xi(u)^\top dW(u)\right]\;dt + \varphi(t)^\top \xi(t)^\top dW(u). \quad (1.21)$$

Lastly, by using (1.20), (1.21) can be rewritten in the form of an Itô diffusion process as

$$dr(t) = [\theta(t) - \alpha(t)r(t)]\ dt + \sigma_r(t)^\top dW(u), \tag{1.22}$$

where

$$\begin{aligned}\theta(t) &= \gamma'(t) - \frac{\varphi'(t)^\top \gamma(t)}{\varphi(t)^\top},\\ \alpha(t) &= -\frac{\varphi'(t)^\top}{\varphi(t)^\top},\\ \sigma_r(t)^\top &= \xi(t)\varphi(t).\end{aligned} \tag{1.23}$$

1.2 Bias-Free Joint Simulation of Short Rate and Numéraire

Definition 1.4 (Point estimator). A point estimator is a rule for calculating an estimate of a given quantity based on observed data. It is a function that maps the sample space to a set of sample estimates, and is usually denoted by the symbol $\hat{\theta}$.

Note that an estimator is itself a random variable. An example of an estimator is the sample mean $\bar{x}$, an estimator of the population mean μ of some random variable X, which has distribution $\bar{x} \sim N(\mu, \sigma^2)$. An estimator is distinguished from an *estimate*, which is a realisation of the estimator, i.e. $\hat{\theta}(X = x)$.

Definition 1.5 (Bias). The bias of an estimator is defined as $\mathbb{B}[\hat{\theta}] = \mathbb{E}[\hat{\theta} - \theta] = \mathbb{E}[\hat{\theta}] - \theta$. It is the difference between the mean of the estimator and the true value of the parameter being estimated, θ. An estimator is said to be *unbiased* if $\mathbb{E}[\hat{\theta}] = \theta$.

If samples are repeated many times, and the same estimation procedure is applied each time, then an unbiased estimator will average out to correct answer in the long run (Stewart and Thiart, 2005). Keeping with our example of the sample mean $\bar{x} \sim N(\mu, \sigma^2)$, it is clear that $\bar{x}$ is an unbiased estimator of the population mean μ. In contrast, it can be shown that the estimator for population variance defined as $s_n^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2$ is biased. This is why the usual estimate for population variance is $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2$. The replacement of $\frac{1}{n}$ by $\frac{1}{n-1}$ is known as Bessel's correction in the statistical literature (Reichmann, 1961).

The problem we would now like to solve is to price an interest rate sensitive payoff X_T. To do so, it remains to evaluate

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(s)ds\right) X_T \,\middle|\, \mathcal{F}_t\right]. \tag{1.24}$$

Given an exact representation of the short rate, we can evaluate the payoff X_T. However, it still remains to specify the numéraire. Since the numéraire identically depends on the evolution of the short rate, the numéraire will be correlated with the payoff X_T, and so these quantities need to be jointly simulated.

A simple approach is to approximate the numéraire via quadrature. Specifically, approximate

$$\int_t^T r(s)ds \approx \sum_{i=1}^{N} r(t_i)\Delta_i, \tag{1.25}$$

so that

$$V_t \approx \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\sum_{i=1}^{N} r(t_i)\Delta_i\right) X_T \,\middle|\, \mathcal{F}_t\right], \tag{1.26}$$

where $\Delta_i = t_i - t_{i-1}$.

However, as noted by Fries (2016), such an approximation may generate bias especially if the model uses large simulation time-steps. In order for such an approximation to be effective, we need to increase the number of points in the discretised time schedule. However, this comes at the expense of computational time. By deriving the joint distribution of the short rate and the numéraire, or some constituent of the numéraire, we are able to avoid this approximation and 'long-step' the short rate, and the numéraire, to the required terminal time. Fries (2016) writes that is of interest in the calculation of xVAs, where there are many risk factors and where computational resources need to be saved.

An alternative solution to the above problem is to use a change of numéraire technique

$$V_t = P(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[X_T \,|\, \mathcal{F}_t\right]. \tag{1.27}$$

Under the T-forward measure, with $P(t,T)$ as the numéraire, X_T is a martingale and therefore all that remains to price the option is simulation of the short rate. The problem with such an approach is that it cannot be used for all types of options (including Bermudan swaptions, ratchet options etc.).

It may even be that this change of numéraire technique holds some clues in the later derivation of the joint distribution. Since

$$\mathbb{C}\mathrm{ov}\left[X,\, Y\right] = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right], \tag{1.28}$$

we have that

$$\mathbb{C}\mathrm{ov}_t\left[r(T),\, \exp\left(-\int_t^T r(s)ds\right)\right] = P(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[r(T)\,|\,\mathcal{F}_t\right] - P(t,T)\mathbb{E}^{\mathbb{Q}}\left[r(T)\,|\,\mathcal{F}_t\right], \tag{1.29}$$

where the subscript t on the $\mathbb{C}\mathrm{ov}$ operator indicates that the covariance is conditional on $\mathcal{F}_t$. We can then compute $\mathbb{E}^{\mathbb{Q}}\left[r(T)\,|\,\mathcal{F}_t\right]$ by solving the SDE for r, and with an appropriate Girsanov transformation we can compute $\mathbb{E}^{\mathbb{Q}^T}\left[r(T)\,|\,\mathcal{F}_t\right]$.

Despite these constraints, Brigo and Mercurio (2006) as well as Andersen and Piterbarg (2010) state that using either technique is often convenient in practice. In fact, Brigo and Mercurio (2006) do not derive the joint distribution of the short rate and the numéraire, which we shall present in this dissertation. Deriving the joint distribution is important for Bermudan options and correlation sensitive products, like LIBOR in arrears.

Chapter 2

Bias-Free Single-Factor Simulation

This chapter focuses on generating joint and bias-free realizations of the short rate and discount factor under the Hull and White single-factor model in (1.6). This chapter will examine three different methods for such simulation: the 'standard' method by Brigo and Mercurio (2006) solves the SDE for r and simulates from r directly; the method by Andersen and Piterbarg (2010) derives the distributional properties of a transform of r, namely $x(t) = r(t) - f(0,t)$; while the method by Fries (2016) uses a deterministic shift representation of the short rate model to simulate r and the numéraire indirectly from two underlying processes. We then develop analytical bond pricing formulae under each of the methods.

2.1 Hull and White (1990)

The Hull and White (1990) model with SDE

$$dr(t) = \kappa \left(\vartheta(t) - r(t)\right) dt + \sigma dW_t, \quad r(0) = r_0, \tag{2.1}$$

can be solved by a simple application of Itô's lemma. Integrating for $t_1 < t_2$, we get

$$r(t_2) = e^{-\kappa(t_2-t_1)} r(t_1) + \kappa \int_{t_1}^{t_2} e^{-\kappa(t_2-u)} \vartheta(u) du + \sigma \int_{t_1}^{t_2} e^{-\kappa(t_2-u)} dW_u. \tag{2.2}$$

If $\vartheta(t)$ is chosen such that it fits the initial TS of interest rates, i.e. letting

$$\vartheta(t) := \frac{1}{\kappa} \frac{\partial f^M(0,t)}{\partial T} + f^M(0,t) + \frac{\sigma^2}{2\kappa^2}\left(1 - e^{-2\kappa t}\right), \tag{2.3}$$

where $f^M(0,t)$ denotes the market-implied instantaneous forward rate, then (2.2) further simplifies to

$$\begin{aligned} r(t_2) &= e^{-\kappa(t_2-t_1)} r(t_1) + \alpha(t) - \alpha(s) e^{-\kappa(t_2-t_1)} + \sigma \int_{t_1}^{t_2} e^{-\kappa(t_2-u)} dW_u, \\ \alpha(t) &:= f^M(0,t) + \frac{\sigma^2}{2\kappa^2}\left(1 - e^{-2\kappa t}\right)^2. \end{aligned} \tag{2.4}$$

Corollary 2.1. *Define a process x by*

$$dx(t) = -\kappa x(t)dt + \sigma dW_t. \tag{2.5}$$

Then $r(t) = x(t) + \alpha(t)$, where $\alpha(t)$ is defined as in (2.4).

The proof follows by solving (2.5) and comparing to (2.4). This will be a useful reflection point when defining the analogous multi-factor model in Section 3.1.

Continuing, we see from (2.4) that given $r(t_1)$, $r(t_2)$ is Gaussian with mean

$$\begin{aligned}\mathbb{E}\left[r(t_2)\,|\,\mathcal{F}_{t_1}\right] &= e^{-\kappa(t_2-t_1)}r(t_1) + \kappa\int_{t_1}^{t_2} e^{-\kappa(t_2-u)}\vartheta(u)du \\ &= e^{-\kappa(t_2-s)}r(s) + \alpha(t_2) - \alpha(s)e^{-\kappa(t_2-s)},\end{aligned} \tag{2.6}$$

and variance

$$\begin{aligned}\mathbb{V}\text{ar}\left[r(t_2)\,|\,\mathcal{F}_{t_1}\right] &= \sigma^2\int_{s}^{t_2} e^{-2\kappa(t_2-u)}du \\ &= \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t_2-t_1)}\right).\end{aligned} \tag{2.7}$$

Corollary 2.2. *The classical Hull-White model (2.1) permits negative interest rates with risk-neutral probability*

$$\mathbb{Q}(r(t) < 0) = \Phi\left(-\frac{\alpha(t)}{\sqrt{\frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t_2-t_1)}\right)}}\right), \tag{2.8}$$

where Φ is cumulative distribution function (CDF) of the standard normal distribution.

Therefore, having the moments of r, we can simulate at times $0 = t_0 < t_1 < \ldots < t_n$ by

$$r(t_{i+1}) = \mathbb{E}\left[r(t_{i+1})\,|\,r(t_i)\right] + \sqrt{\mathbb{V}\text{ar}\left[r(t_{i+1})\,|\,r(t_i)\right]}Z_i \tag{2.9}$$

for independent $Z_1, Z_2, \ldots, Z_n \sim N(0,1)$.

Having generated a sample path for r, we are now in a position to compute bond prices, or the numéraire, by integrating $\int_{t_1}^{t_2} r(u)du$ numerically. However, as noted in Section 1.2, in doing so we introduce a discretisation error, or bias, which is dependent on the number of dates in the schedule. Instead, we can use the fact that $\int_{t_1}^{t_2} r(u)du$ is itself a Gaussian process to simulate the numéraire directly, and without bias.

Using an application of Fubini's theorem, we have that $\int_{t_1}^{t_2} r(u)du$ has mean

$$\begin{aligned}\mathbb{E}\left[\int_t^T r(u)du \,\middle|\, \mathcal{F}_t\right] &= \int_t^T \mathbb{E}\left[r(u)\,|\,\mathcal{F}_t\right]du \\ &= \frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right)r(t) + \int_t^T \left(1-e^{-\kappa(T-s)}\right)\vartheta(s)ds \\ &= A(t,T)\left[r(t)-\alpha(t)\right] + \ln\frac{P^M(0,t)}{P^M(0,T)} + \frac{\sigma^2}{2\kappa^2}\left[\beta(T)-\beta(t)\right],\end{aligned} \tag{2.10}$$

and variance

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\int_t^T r(u)du \,\middle|\, \mathcal{F}_t\right] &= \mathbb{C}\mathrm{ov}_t\left[\int_t^T r(u)du,\ \int_t^T r(s)ds\right] \\ &= 2\int_t^T\int_t^u \mathbb{C}\mathrm{ov}_t\left[r(u), r(s)\right]ds\,du \\ &= 2\int_t^T\int_t^u\int_t^s \sigma^2 e^{-\kappa(u-v)}e^{-\kappa(s-v)}dv\,ds\,du \\ &= \frac{\sigma^2}{\kappa^2}\left[(T-t) - A(t,T) - \frac{\kappa}{2}A^2(t,T)\right],\end{aligned} \tag{2.11}$$

where

$$A(t,T) = \frac{1}{\kappa}\left(1-e^{-\kappa(T-t)}\right), \qquad \beta(t) = t - \frac{1}{2\kappa}e^{-2\kappa t} - \frac{2}{\kappa}e^{-2\kappa t}. \tag{2.12}$$

We now have the marginal distributions of $r(t_2)$ and $\int_{t_1}^{t_2} r(u)du$ conditional on $\mathcal{F}_{t_1}$, under the risk-neutral measure $\mathbb{Q}$. However, it remains to specify the covariance structure in order to jointly simulate the pair $(r(t_2), \int_{t_1}^{t_2} r(u)du)$ conditional on $\mathcal{F}_{t_1}$. For ease of notation, define $Y(t) := \int_0^t r(u)du$, and note that

$$\mathbb{E}\left[Y(t_2)\,|\,\mathcal{F}_{t_1}\right] = Y(t_1) + \mathbb{E}\left[\int_{t_1}^{t_2} r(u)du\,|\,\mathcal{F}_{t_1}\right], \tag{2.13}$$

and

$$\mathbb{V}\mathrm{ar}\left[Y(t_2)\,|\,\mathcal{F}_{t_1}\right] = \mathbb{V}\mathrm{ar}\left[\int_{t_1}^{t_2} r(u)du\,|\,\mathcal{F}_{t_1}\right]. \tag{2.14}$$

The covariance, conditional on $\mathcal{F}_{t_1}$ is then given by

$$\begin{aligned}\mathbb{C}\mathrm{ov}_{t_1}\left[Y(t_2),\ r(t_2)\right] &= \int_{t_1}^{t_2} \mathbb{C}\mathrm{ov}_{t_1}\left[r(u),\ r(t_2)\right]du \\ &= \int_{t_1}^{t_2}\int_{t_1}^{s} \sigma^2 e^{-\kappa(t_2-v)}e^{-\kappa(u-v)}dv\,du \\ &= \frac{\sigma^2}{2}A^2(t_1,t_2),\end{aligned} \tag{2.15}$$

so that the correlation is

$$\rho_{rY}(t_1,t_2) = \frac{\mathbb{C}\mathrm{ov}_{t_1}\left[Y(t_2),\, r(t_2)\right]}{\sqrt{\mathbb{V}\mathrm{ar}\left[Y(t_2)\,|\,\mathcal{F}_{t_1}\right]}\sqrt{\mathbb{V}\mathrm{ar}\left[r(t_2)\,|\,\mathcal{F}_{t_1}\right]}}. \tag{2.16}$$

Having now derived the joint distribution of $(r(t),\, Y(t))$, we can simulate a realisation of the pair at times $0 = t_0 < t_1 < \ldots < t_n$ by

$$\begin{aligned} r(t_{i+1}) &= \mathbb{E}\left[r(t_{i+1})\,|\,\mathcal{F}_i\right] + \sqrt{\mathbb{V}\mathrm{ar}\left[r(t_{i+1})\,|\,\mathcal{F}_i\right]} Z_1(i) \\ Y(t_{i+1}) &= \mathbb{E}\left[Y(t_{i+1})\,|\,\mathcal{F}_i\right] + \sqrt{\mathbb{V}\mathrm{ar}\left[Y(t_{i+1})\,|\,\mathcal{F}_i\right]} \times \\ &\quad \left[\rho_{rY}(t_i,t_{i+1}) Z_1(i) + \sqrt{1-\rho_{rY}^2(t_i,t_{i+1})} Z_2(i)\right], \end{aligned} \tag{2.17}$$

for $(Z_1, Z_2) \sim N_2(\mathbf{0},\mathbf{1})$ independent bivariate normal vectors.

2.2 Andersen and Piterbarg (2010)

Andersen and Piterbarg parameterise their one-factor GSR model by the SDE

$$dr(t) = \kappa(t)\left(\vartheta(t) - r(t)\right) dt + \sigma(t) dW(t), \quad r(0) = r_0. \tag{2.18}$$

However, for this model to match the initial yield curve, we require

$$\vartheta(t) := \frac{1}{\kappa(t)} \frac{\partial f^M(0,t)}{\partial t} + f^M(0,t) + \frac{1}{\kappa(t)} \int_0^t \exp\left(-2\int_u^t \kappa(s) ds\right) \sigma(u)^2 du, \tag{2.19}$$

which involves a $\partial f^M(0,t)/\partial t$ term. This can be problematic when the initial forward curve is not smooth, which can commonly arise if the curve is bootstrapped. For this reason, Andersen and Piterbarg transform $r(t)$ to $x(t) := r(t) - f(0,t)$, and go on to derive the dynamics of $x(t)$, as well as the bond reconstitution formulae in terms of $x(t)$.

Proposition 2.3. *Define $x(t) := r(t) - f(0,t)$, then for the model (2.18),*

$$dx(t) = \left(y(t) - \kappa(t)x(t)\right) dt + \sigma(t) dW(t), \quad x(0) = 0, \tag{2.20}$$

where

$$y(t) = \int_0^t \exp\left(-2\int_u^t \kappa(s) ds\right) \sigma(u)^2 du. \tag{2.21}$$

Proof. From Lemma 1.3, recognize that the dynamics from (2.18) must originate from an HJM model (1.16) with $\sigma_f(t,T) = \xi(t)\varphi(T)$ — an HJM model that admits a Markovian short rate r. With knowledge of this, the result follows by making the substitutions $\sigma_f(t,T) = \sigma(t)\exp(-\int_t^T \kappa(u)du)$ and $x(t) = r(t) - f(0,t)$ to (1.19).

Proposition 2.4. *The price at time t of a zero-coupon bond maturing at time T with unit face value is*

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left(-x(t)G(t,T) - \frac{1}{2}y(t)G^2(t,T)\right), \tag{2.22}$$

where

$$G(t,T) = \int_t^T \exp\left(-\int_t^u \kappa(s)\right) ds\, du. \tag{2.23}$$

Proof. Andersen and Piterbarg (2010) derive the bond reconstitution formula by inserting the expression obtained for $f(t,T)$, where $\sigma_f(t,T) = \xi(t)\varphi(T)$ in (1.16), into $P(t,T) = \exp(-\int_t^T f(t,u)du)$ and integrating. Alternatively, by recognizing that (2.20) is affine in $x(t)$, we can solve the term structure PDEs to solve for functions $A(t,T)$ and $B(t,T)$ in $P(t,T) = \exp(A(t,T) - B(t,T)x(t))$.

The SDE in (2.20) can be solved and can be simulated exactly by discretising the time-interval so that for $t_1 < t_2$

$$\begin{aligned} x(t_2) = \exp&\left(-\int_{t_1}^{t_2} \kappa(u)du\right) x(t_1) + \int_{t_1}^{t_2} \exp\left(-\int_s^{t_2} \kappa(u)du\right) y(s)ds \\ &+ \int_{t_1}^{t_2} \exp\left(-\int_s^{t_2} \kappa(u)du\right) \sigma(s)dW(s). \end{aligned} \tag{2.24}$$

This is again Gaussian and can be simulated bias-free as in (2.9) with

$$\begin{aligned} \mathbb{E}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right] &= \exp\left(-\int_{t_1}^{t_2} \kappa(u)du\right) x(t_1) + \int_{t_1}^{t_2} \exp\left(-\int_s^{t_2} \kappa(u)du\right) y(s)ds, \\ \mathbb{V}\mathrm{ar}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right] &= \int_{t_1}^{t_2} \left(\exp\left(-\int_s^{t_2} \kappa(u)du\right) \sigma(s)\right)^2 ds. \end{aligned} \tag{2.25}$$

Next, by defining $I(t) := -\int_0^t x(u)du$, and finding the joint distribution of $x(t_2)$ and $I(t_2)$ conditional on $\mathcal{F}_{t_1}$, we can simulate the discount factor jointly and bias free. $I(t)$ has moments

$$\begin{aligned} \mathbb{E}\left[I(t_2)\,|\,\mathcal{F}_{t_1}\right] &= -\int_0^{t_2} \mathbb{E}\left[x(u)\,|\,\mathcal{F}_{t_1}\right] du \\ &= -I(t_1) - \int_{t_1}^{t_2} \mathbb{E}\left[x(u)\,|\,\mathcal{F}_{t_1}\right] du \\ &= I(t_1) - x(t_1)G(t_1,t_2) - \int_{t_1}^{t_2}\int_{t_1}^{u} e^{-\int_s^u \kappa(v)dv} y(s)ds\, du, \end{aligned} \tag{2.26}$$

and

$$\begin{aligned} \mathbb{V}\mathrm{ar}\left[I(t_2)\,|\,\mathcal{F}_{t_1}\right] &= 2\int_{t_1}^{t_2}\int_{t_1}^{u} \mathbb{C}\mathrm{ov}_{t_1}\left[x(u),\, x(s)\right] ds\, du \\ &= 2\int_{t_1}^{t_2}\int_{t_1}^{u}\int_{t_1}^{s} \sigma^2(v) e^{-\int_v^u \kappa(u)du} e^{-\int_v^s \kappa(u)du} dv\, ds\, du. \end{aligned} \tag{2.27}$$

These results follow from an application of Fubini's theorem, and thereafter algebraic manipulations for Gaussian random variables, including use of results similar to (2.13) and (2.14).

Lastly, the covariance between $I(t_2)$ and $x(t_2)$ conditional on $\mathcal{F}_{t_1}$ is

$$\begin{aligned}\mathbb{C}\text{ov}_{t_1}\left[I(t_2),\, x(t_2)\right] &= -\int_{t_1}^{t_2} \mathbb{C}\text{ov}_{t_1}\left[x(u),\, x(t_2)\right] du \\ &= -\int_{t_1}^{t_2}\int_{t_1}^{s} \sigma(v)^2 e^{-\int_v^{t_2}\kappa(s)ds} e^{-\int_v^{u}\kappa(s)ds} dv\, du,\end{aligned} \tag{2.28}$$

so that the correlation is

$$\rho_{xI}(t_1, t_2) = \frac{\mathbb{C}\text{ov}_{t_1}\left[I(t_2),\, x(t_2)\right]}{\sqrt{\mathbb{V}\text{ar}\left[I(t_2)\,|\,\mathcal{F}_{t_1}\right]}\sqrt{\mathbb{V}\text{ar}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right]}}. \tag{2.29}$$

The derivation is similar to the workings leading up to (2.16), except that now κ and σ are time varying functions.

Having now derived the joint distribution of $(x(t),\, I(t))$, we can simulate a realisation of the pair at times $0 = t_0 < t_1 < \ldots < t_n$ by

$$\begin{aligned}x(t_{i+1}) &= \mathbb{E}\left[x(t_{i+1})\,|\,\mathcal{F}_{t_i}\right] + \sqrt{\mathbb{V}\text{ar}\left[x(t_{i+1})\,|\,\mathcal{F}_{t_i}\right]} Z_1(i) \\ I(t_{i+1}) &= \mathbb{E}\left[I(t_{i+1})\,|\,\mathcal{F}_{t_i}\right] + \sqrt{\mathbb{V}\text{ar}\left[I(t_{i+1})\,|\,\mathcal{F}_{t_i}\right]}\times \\ &\quad \left[\rho_{xI}(t_i, t_{i+1})Z_1(i) + \sqrt{1-\rho_{xI}^2(t_i, t_{i+1})}Z_2(i)\right],\end{aligned} \tag{2.30}$$

for $(Z_1, Z_2) \sim N_2(\mathbf{0}, \mathbf{1})$ independent bivariate normal vectors.

Using the joint distribution of $(x(t),\, I(t))$, we can price a derivative security with payoff X_T, where X_T depends explicitly on $r(t) : 0 \le t \le T$, and not the transform $x(t)$, using the following result

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[X_T e^{-\int_t^T r(u)du}\,\middle|\,\mathcal{F}_t\right] \tag{2.31}$$

$$= P(t,T)\mathbb{E}^{\mathbb{Q}}\left[X_T^x e^{-\int_t^T x(u)du}\,\middle|\,\mathcal{F}_t\right], \tag{2.32}$$

which follows from the way in which we defined $x(t) := r(t) - f(0,t)$. The payoff function X_T^x now depends explicitly on $x(t) : 0 \le t \le T$ instead of $r(t)$, because we choose to simulate x instead of r. In practice, it is easy to compute X_T given $r(t) = x(t) + f(0,t)$.

2.3 Fries (2016)

Fries (2016) parametrises the extended Vašíček model with SDE

$$dr(t) = (\theta(t) - \alpha(t)r(t))\, dt + \sigma(t)dW(t), \quad r(0) = r_0. \tag{2.33}$$

By defining

$$M(t) := \exp\left(\int_0^t \alpha(s)ds\right), \quad M(t,T) := \frac{M(t)}{M(T)} = \exp\left(-\int_t^T \alpha(s)ds\right), \tag{2.34}$$

Fries (2016) shows that by solving $d(M(t)r(t)) = M(t)\theta(t)dt + \sigma(t)dW(t)$,

$$r(t) = r(t_i) - \tilde{a}(t_i,t)r(t_i)(t-t_i) + \int_{t_i}^t M(s,t)\theta(s)ds + \int_{t_i}^t M(s,t)\sigma(s)dW(s), \tag{2.35}$$

where

$$\tilde{a}(t_i,t) = \frac{1-M(t_i,t)}{t-t_i} = \frac{1-\exp(-\int_{t_i}^t \alpha(s)ds)}{t-t_i}. \tag{2.36}$$

This is Gaussian, and so an exact Euler step for the short rate r is given by

$$r(t_{i+1}) = \mathbb{E}\left[r(t_{i+1}) \,|\, r(t_i)\right] + \sqrt{\mathbb{V}\text{ar}\left[r(t_{i+1}) \,|\, r(t_i)\right]} Z_i \tag{2.37}$$

with

$$\begin{aligned} \mathbb{E}\left[r(t_{i+1}) \,|\, r(t_i)\right] &= r(t_i) - \tilde{a}(t_i,t_{i+1})r(t_i)(t_{i+1}-t_i) + \int_{t_i}^{t_{i+1}} M(s,t_{i+1})\theta(s)ds, \\ \mathbb{V}\text{ar}\left[r(t_{i+1}) \,|\, r(t_i)\right] &= \int_{t_i}^{t_{i+1}} \left(M(s,t_{i+1})\sigma(s)\right)^2 ds, \end{aligned} \tag{2.38}$$

for $i = 0, \ldots, N-1$ independent standard normal variates Z.

Alternatively, by using a so-called deterministic shift representation attributed to Brigo and Mercurio (2001), an exact simulation scheme for the short rate is given as

$$r(t_{i+1}) = x(t_i) + f(0,t_i) + \frac{1}{2}\frac{dV(t_i)}{dt}, \tag{2.39}$$

where

$$\begin{aligned} x(t_{i+1}) &= x(t_i) - \tilde{a}(t_i,t_{i+1})x(t_i)(t_{i+1}-t_i) + \int_{t_i}^{t_{i+1}} M(s,t_{i+1})\sigma(s)dW(s), \\ x(t_0) &= 0. \end{aligned} \tag{2.40}$$

This differs from before, in that now we do not simulate the short rate directly, but rather simulate the underlying stochastic process x, and thereafter compute the short rate as a deterministic shift of x using (2.39). For the numéraire $N(t)$, by integrating (2.35) with respect to t from t_i to T, we get for the diffusion part

$$\begin{aligned} \int_{t_i}^T \int_{t_i}^t M(s,t)\sigma(s)dW(s)\,dt &= \int_{t_i}^T \int_s^t M(s,t)\sigma(s)dt\,dW(s) \\ &= \int_{t_i}^T B(s,T)\sigma(s)dW(s), \end{aligned} \tag{2.41}$$

so that the correlation between $r(t_{i+1}) - r(t_i)$ and $\int_{t_i}^{t_{i+1}} r(s)ds$ is

$$\rho(t_i, t_{i+1}) = \frac{\int_{t_i}^{t_{i+1}} M(s, t_{i+1}) B(s, t_{i+1}) \sigma^2(s) ds}{\sqrt{\int_{t_i}^{t_{i+1}} \left(M(s, t_{i+1})\sigma(s)\right)^2 ds} \sqrt{\int_{t_i}^{t_{i+1}} \left(B(s, t_{i+1})\sigma(s)\right)^2 ds}}, \tag{2.42}$$

where

$$B(t, T) = \int_t^T M(t, s) ds. \tag{2.43}$$

To complete the derivation of an exact joint-simulation scheme for $(r(t),\ N(t))$, we need to get the drift part of $\int_{t_i}^{t_{i+1}} r(s)ds$ by integrating (2.35) with respect to t, and then making the appropriate substitutions to get $N(t_i) = \exp(\int_0^{t_i} r(s)ds)$. Thereafter, Fries (2016) again cites the earlier work from Brigo and Mercurio (2001) on the deterministic shift representation of short rate models to specify an exact sampling scheme for the short rate and numéraire as

$$\begin{aligned} r(t_{i+1}) &= x(t_i) + f(0, t_i) + \frac{1}{2}\frac{dV(t_i)}{dt}, \\ N(t_i) &= \exp\left(y(t_i) + \frac{1}{2}V(t_i)\right) / P(t_i), \end{aligned} \tag{2.44}$$

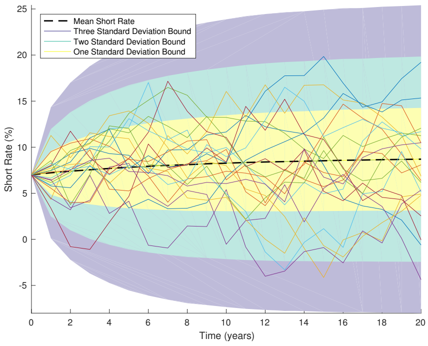
where

$$\begin{aligned} x(t_{i+1}) &= x(t_i) - \tilde{a}(t_i, t_{i+1}) x(t_i) \Delta t_i + \sigma_i \sqrt{\Delta t_i} Z_1(i), \\ y(t_{i+1}) &= y(t_i) + \tilde{b}(t_i, t_{i+1}) x(t_i) \Delta t_i + \gamma_i \sqrt{\Delta t_i} \left(\rho_i Z_1(i) + \sqrt{1 - \rho_i^2} Z_2(i)\right), \end{aligned} \tag{2.45}$$

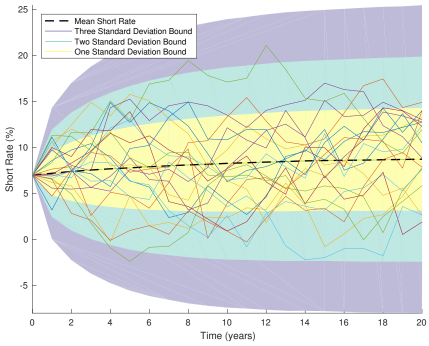
for $(Z_1, Z_2) \sim N_2(\mathbf{0}, \mathbf{1})$ independent bivariate normal vectors and with initial values $(x(t_0),\ y(t_0)) = \mathbf{0}$. For brevity, functions $\tilde{b}$, V, ρ_i, σ_i and γ_i are left to Appendix A.1, although, for example, σ_i and γ_i have been implicitly defined in the denominator of (2.42).

Figure 2.1 plots 20 sample paths for classical Hull-White model (2.1) with $r_0 = 0.07$, $\kappa = 0.1$, $\vartheta = 0.09$, $\sigma = 0.025$ for the methods of Section 2.1, Section 2.2, and Section 2.3[1]. Since the short rate is Gaussian, the one, two, and three standard deviation bounds correspond to 68%, 95%, and 99.7% confidence intervals respectively.

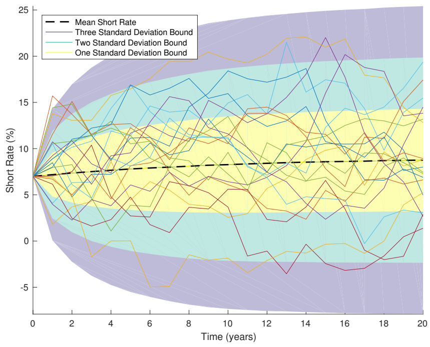
[1] For the parametrisation used by Fries in (2.33), the corresponding parameters are $r_0 = 0.07$, $\alpha = 0.1$, $\theta = 0.009$, $\sigma = 0.025$.



(a) Hull and White (1990)



(b) Andersen and Piterbarg (2010)



(c) Fries (2016)

Fig. 2.1: 20 sample paths for classical Hull-White model (2.1) under all three simulation methods, including the mean short rate and standard deviation bounds. Parameters used: $r_0 = 0.07$, $\kappa = 0.1$ ($\alpha = 0.1$), $\vartheta = 0.09$ ($\theta = 0.009$), $\sigma = 0.025$.

Chapter 3

Extension to Multi-Factor Models

In this chapter we present the multi-factor extensions for the methods presented in Section 2.1 and Section 2.2. Specifically, with ease of exposition in mind, a two-factor extension will be presented. As highlighted in Section 1.1.3, a two-factor model captures between 85%–90% of variability in the market yield curve, and hence will be considered sufficiently parsimonious. In Brigo and Mercurio (2006) a two-factor model is presented, and in Andersen and Piterbarg (2010) the two-factor model is considered the 'important case'. This chapter will, however, make clear how an extension to a three- or higher dimensioned multi-factor model proceeds.

3.1 Brigo and Mercurio (2006) G2++ model

Brigo and Mercurio's (2006) so-called two-additive-factor Gaussian model (G2++) assumes that the instantaneous short rate has dynamics under $\mathbb{Q}$

$$r(t) = x_1(t) + x_2(t) + \varphi(t), \quad r(0) = r_0, \tag{3.1}$$

where $x_1(t)$ and $x_2(t)$ satisfy

$$\begin{aligned} dx_1(t) &= -\kappa_1 x_1(t)dt + \sigma_1 dW_1(t), \quad x_1(0) = 0, \\ dx_2(t) &= -\kappa_2 x_2(t)dt + \sigma_2 dW_2(t), \quad x_2(0) = 0, \end{aligned} \tag{3.2}$$

and where $dW_1(t)dW_2(t) = \rho dt$. The function φ is deterministic and well-defined on a given time interval $[0, T^*]$, where T^* is typically 10, 30 or 50 years (Brigo and Mercurio, 2006). Since $r(0) = 0$, $x(0) = 0$, and $y(0) = 0$, we must have that $\varphi(0) = r_0$.

Proposition 3.1. *The model (3.1) fits the initial term structure of interest rates if and only if, for each $T \in [0, T^*]$*

$$\begin{aligned} \varphi(T) = f(0,T) &+ \frac{\sigma_1^2}{2\kappa_1^2}\left(1 - e^{-\kappa_1 T}\right)^2 + \frac{\sigma_2^2}{2\kappa_2^2}\left(1 - e^{-\kappa_2 T}\right)^2 \\ &+ \rho\frac{\sigma_1\sigma_2}{\kappa_1\kappa_2}\left(1 - e^{-\kappa_1 T}\right)\left(1 - e^{-\kappa_2 T}\right). \end{aligned} \tag{3.3}$$

This is labelled Corollary 4.2.1 in Brigo and Mercurio (2006), and the proof is therein.

From Proposition 3.1, notice that by dropping the stochastic process $x_2(t)$, the model (3.1) reduces to $r(t) = x_1(t) + \varphi(t)$, where $\varphi(t) = \alpha(t)$ from Corollary 2.1. Indeed, (3.1) is the two-factor analogue of the method described in Section 2.1.

To derive a sampling scheme for the G2++ model of (3.1), Brigo and Mercurio (2006) begin by integrating (3.2) for $t_1 < t_2$, to get

$$\begin{aligned} r(t_2) = e^{-\kappa_1(t_2-t_1)}x_1(t_1) + e^{-\kappa_2(t-t_1)}x_2(t_1) \\ + \sigma_1 \int_{t_1}^{t_2} e^{-\kappa_1(t_2-u)} dW_1(u) + \sigma_2 \int_{t_1}^{t_2} e^{-\kappa_2(t_2-u)} dW_2(u) + \varphi(t_2). \end{aligned} \tag{3.4}$$

This is Gaussian with

$$\begin{aligned} \mathbb{E}\left[r(t_2) \,|\, \mathcal{F}_{t_1}\right] &= e^{-\kappa_1(t_2-t_1)}x_1(t_1) + e^{-\kappa_2(t_2-t_1)}x_2(t_1) + \varphi(t_2), \\ \mathbb{V}\text{ar}\left[r(t_2) \,|\, \mathcal{F}_{t_1}\right] &= \frac{\sigma_1^2}{2\kappa_1}\left[1 - e^{-2\kappa_1(t_2-t_1)}\right] + \frac{\sigma_2^2}{2\kappa_2}\left[1 - e^{-2\kappa_2(t_2-t_1)}\right] \\ &\quad + 2\rho\frac{\sigma_1\sigma_2}{\kappa_1+\kappa_2}\left[1 - e^{-(\kappa_1+\kappa_2)(t_2-t_1)}\right]. \end{aligned} \tag{3.5}$$

This is comparable to (2.6) and (2.7), but differs due to the addition of the second correlated Brownian motion, $W_2(t)$. In particular, this correlation structure is visible as the third term in the conditional variance of r. For an n dimensional Gn++ model, this changes to

$$\begin{aligned} \mathbb{E}\left[r(t_2) \,|\, \mathcal{F}_{t_1}\right] &= \sum_{i=1}^{n} e^{-\kappa_i(t_2-t_1)}x_i(t_1) + \varphi(t_2), \\ \mathbb{V}\text{ar}\left[r(t_2) \,|\, \mathcal{F}_{t_1}\right] &= \sum_{i=1}^{n} \frac{\sigma_i^2}{2\kappa_i}\left[1 - e^{-2\kappa_i(t_2-t_1)}\right] + 2\sum_{1\le i<j}^{n} \rho_{ij}\frac{\sigma_i\sigma_j}{\kappa_i+\kappa_j}\left[1 - e^{-(\kappa_i+\kappa_j)(t_2-t_1)}\right]. \end{aligned} \tag{3.6}$$

where $dW_i(t)dW_j(t) = \rho_{ij}dt$.

For simulation of the numéraire, (Brigo and Mercurio, 2006, pg. 145) show that the random variable $Y(t,T) = \int_t^T [x_1(u) + x_2(u)]du$ conditional on $\mathcal{F}_t$ is normally distributed with mean

$$\begin{aligned} \mathbb{E}\left[Y(t_1,t_2) \,|\, \mathcal{F}_{t_1}\right] &= \int_{t_1}^{t_2} \mathbb{E}\left[x_1(u) + x_2(u) \,|\, \mathcal{F}_{t_1}\right] du \\ &= \frac{1 - e^{-\kappa_1(T-t_1)}}{\kappa_1}x(t_1) + \frac{1 - e^{-\kappa_2(T-t_1)}}{\kappa_2}x_2(t_1) \\ &=: M(t_1,t_2), \end{aligned} \tag{3.7}$$

and variance

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}\left[Y(t_1,t_2)\,|\,\mathcal{F}_{t_1}\right] &= \mathbb{C}\mathrm{ov}_{t_1}\left[\int_{t_1}^{t_2} x_1(u)+x_2(u)\,du,\ \int_{t_1}^{t_2} x_1(s)+x_2(s)\,ds\right] \\
&= \frac{\sigma_1^2}{\kappa_1^2}\left[(t_2-t_1)-A_1(t_1,t_2)-\frac{\kappa_1}{2}A_1^2(t_1,t_2)\right] \\
&\quad+\frac{\sigma_2^2}{\kappa_2^2}\left[(t_2-t_1)-A_2(t_1,t_2)-\frac{\kappa_2}{2}A_2^2(t_1,t_2)\right] \\
&\quad+2\rho\frac{\sigma_1\sigma_2}{\kappa_1\kappa_2}\Bigg[(t_2-t_1)-A_1(t_1,t_2)-A_2(t_1,t_2) \\
&\qquad\qquad\qquad\qquad-\frac{e^{-(\kappa_1+\kappa_2)(t_2-t_1)}-1}{\kappa_1+\kappa_2}\Bigg] \\
&=: V(t_1,t_2),
\end{aligned}
\tag{3.8}
$$

where

$$
A_i(t_1,t_2)=\frac{1}{\kappa_i}\left(1-e^{-\kappa_i(t_2-t_1)}\right). \tag{3.9}
$$

The mean and variance of $Y(t,T)$ is similarly extended for a Gn++ model as in (3.6).

Corollary 3.2. *The price at time t of a zero-coupon bond maturing at time T with unit face value is*

$$
P(t,T)=\exp\left(-\int_t^T\varphi(u)du-M(t,T)+\frac{1}{2}V(t,T)\right). \tag{3.10}
$$

Proof. Since $Y(t,T)$ is normally distributed with mean $M(t,T)$ and variance $V(t,T)$, and since

$$
\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)\middle|\,\mathcal{F}_t\right]=\exp\left(-\int_t^T\varphi(u)du\right)\mathbb{E}^{\mathbb{Q}}\left[\exp(-Y(t,T))\,|\,\mathcal{F}_t\right],
$$

the result follows by using the moment generating function for normal random variables.

In the single-factor case, we only needed to derive the correlation between the short rate and the discount factor. Now we have that the Brownian motions driving $x_1(t)$ and $x_2(t)$ are correlated, with $dW_1(t)dW_2(t)=\rho dt$, and so $I(t,T)$ is correlated to both $x_1(t)$ and $x_2(t)$. Hence, in order to jointly simulate the short rate and the discount factor bias-free, we now need to find the cross-correlation matrix for $(x_1(t),x_2(t),Y(t,T))$.

Notice that the variance expression in (3.5) is simply the sum of $\mathbb{V}\mathrm{ar}\left[x_1(t_2)\,|\,\mathcal{F}_{t_1}\right]$, $\mathbb{V}\mathrm{ar}\left[x_2(t_2)\,|\,\mathcal{F}_{t_1}\right]$ and $2\times\mathbb{C}\mathrm{ov}_{t_1}\left[x_1(t_2),\ x_2(t_2)\right]$. From (3.8) we have $\mathbb{V}\mathrm{ar}\left[Y(t_1,t_2)\,|\,\mathcal{F}_{t_1}\right]$, and so all that is missing to populate the cross-covariance (and hence cross-correlation) matrix is to compute $\mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, Y(t_1,t_2)]$ and $\mathbb{C}\mathrm{ov}_{t_1}[x_2(t_2),\, Y(t_1,t_2)]$. $\mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, Y(t_1,t_2)]$ is computed as follows

$$\begin{aligned}\mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, Y(t_1,t_2)] &= \int_{t_1}^{t_2} \mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, x_2(u)] + \mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, x_1(u)]\; du \\ &= \int_{t_1}^{t_2}\int_{t_1}^{u} \rho\sigma_1\sigma_2 e^{-\kappa_1(t_2-s)}e^{-\kappa_2(u-s)} ds\, du \\ &\quad + \int_{t_1}^{t_2}\int_{t_1}^{u} \sigma_1^2 e^{-\kappa_1(t_2-s)}e^{-\kappa_1(u-s)} ds\, du \\ &= \rho\frac{\sigma_1\sigma_2}{\kappa_1+\kappa_2}\left[A_2(t_1,t_2) + \frac{1}{\kappa_2}\left(e^{-(\kappa_1+\kappa_2)(t_2-t_1)} - e^{-\kappa_1(t_2-t_1)}\right)\right] \\ &\quad + \frac{\sigma_1^2}{2\kappa_1}\left[A_1(t_1,t_2) + \frac{1}{\kappa_1}\left(e^{-2\kappa_1(t_2-t_1)} - e^{-\kappa_1(t_2-t_1)}\right)\right].\end{aligned} \tag{3.11}$$

The first line follows since the covariance operator is linear. Thereafter straightforward calculations for Gaussian random variables apply. $\mathbb{C}\mathrm{ov}_{t_1}[x_2(t_2),\, Y(t_1,t_2)]$ is computed similarly.

We therefore have all the necessary ingredients to populate the 3×3 instantaneous cross-correlation matrix. For $(x_1(t), x_2(t), Y(t,T))$, the upper diagonal entries of the symmetric matrix Σ are

$$\begin{bmatrix} 1 & \dfrac{\mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, x_2(t_2)]}{\sqrt{\mathbb{V}\mathrm{ar}[x_1(t_2)\,|\,\mathcal{F}_{t_1}]}\sqrt{\mathbb{V}\mathrm{ar}[x_2(t_2)\,|\,\mathcal{F}_{t_1}]}} & \dfrac{\mathbb{C}\mathrm{ov}_{t_1}[x_1(t_2),\, Y(t_1,t_2)]}{\sqrt{\mathbb{V}\mathrm{ar}[x_1(t_2)\,|\,\mathcal{F}_{t_1}]}\sqrt{\mathbb{V}\mathrm{ar}[Y(t_1,t_2)\,|\,\mathcal{F}_{t_1}]}} \\ & 1 & \dfrac{\mathbb{C}\mathrm{ov}_{t_1}[x_2(t_2),\, Y(t_1,t_2)]}{\sqrt{\mathbb{V}\mathrm{ar}[x_2(t_2)\,|\,\mathcal{F}_{t_1}]}\sqrt{\mathbb{V}\mathrm{ar}[Y(t_1,t_2)\,|\,\mathcal{F}_{t_1}]}} \\ & & 1 \end{bmatrix}. \tag{3.12}$$

For ease of representation, the lower diagonal entries are omitted. However, the matrix is symmetric and hence these entries correspond to the opposite upper entries. For a Gn++ model, the matrix Σ has dimensions $(n+1)\times(n+1)$.

Having derived the correlation structure of $(x_1(t), x_2(t), Y(t,T))$, we can now jointly simulate a sample realisation of the short rate $r(t)$ and $Y(t,T)$ at times $0 = t_0 < t_1 < \ldots < t_n$. Let L be the lower triangular Cholesky decomposition of the correlation matrix, Σ, and generate a random normal $n\times 3$ matrix Z. Then $X = LZ$ is a $n \times 3$ matrix of correlated normal random variables which can be used for updates of $(x_1(t), x_2(t), Y(t,T))$. Advancement of the schedule proceeds as follows

$$\begin{aligned} x_1(t_{i+1}) &= \mathbb{E}[x_1(t_{i+1})\,|\,\mathcal{F}_{t_i}] + \sqrt{\mathbb{V}\mathrm{ar}[x_1(t_{i+1})\,|\,\mathcal{F}_{t_i}]}X(i,1) \\ x_2(t_{i+1}) &= \mathbb{E}[x_2(t_{i+1})\,|\,\mathcal{F}_{t_i}] + \sqrt{\mathbb{V}\mathrm{ar}[x_2(t_{i+1})\,|\,\mathcal{F}_{t_i}]}X(i,2) \\ Y(t_{i+1}) &= \mathbb{E}[Y(t_{i+1})\,|\,\mathcal{F}_{t_i}] + \sqrt{\mathbb{V}\mathrm{ar}[Y(t_{i+1})\,|\,\mathcal{F}_{t_i}]}X(i,3). \end{aligned} \tag{3.13}$$

To compute the short rate $r(t)$ from updates of $x_1(t)$ and $x_2(t)$, we use the relationship in (3.1). However, it is often of interest to work with the processes $x_1(t)$ and $x_2(t)$ explicitly, since, for example, bond prices (see (3.2)) and therefore other option prices of interest depend explicitly on $x_1(t)$ and $x_2(t)$. This is after all, by design of multi-factor models: option prices are not characterised by merely one stochastic process.

Brigo and Mercurio (2006) do not go so far as to derive this correlation structure, choosing instead to derive the distributional properties of $r(t)$ under the T-forward measure $\mathbb{Q}^T$, or using quadrature as the situation dictates.

3.2 Andersen and Piterbarg (2010)

The multi-factor Gaussian model can be developed in two different ways: the 'classical' way; 'from the bottom up', which Andersen and Piterbarg (2010) comment as involving 'laborious details', or via the 'modern' way; within an HJM setting from a separability condition. In this section we will develop the multi-factor model via the HJM setting, as we did for the single-factor model in Section 2.2.

Remark 3.3. The single-factor model in Section 2.1, as well as the multi-factor model in Section 3.1, are developed in what Andersen and Piterbarg would call the 'classical' way. In choosing to develop Andersen and Piterbarg's method via an HJM framework, this dissertation is able to showcase both approaches.

We begin by recalling a result from Section 1.1.4. Assume that the forward rate volatility is separable, as in Lemma 1.3. That is, assume $\sigma_f(t,T) = \xi(t)\varphi(T)$. Then

$$f(t,T) = f(0,T) + \varphi(T)^\top \int_0^t \xi(u)^\top \xi(u) \int_u^T \varphi(s)\, ds\, du + \varphi(T)^\top \int_0^t \xi(u)^\top dW(u). \tag{3.14}$$

Next, define $z(t) = \int_0^t \xi(u)^\top dW(u)$, and notice that $z(t)$ is a d-dimensional random vector satisfying

$$dz(t) = \xi(t)^\top dW(t), \quad z(0) = 0. \tag{3.15}$$

This demonstrates that the forward curve can be reconstructed from d Gaussian martingale variables. However, Andersen and Piterbarg assert that the choice of d state variables is not unique, and in fact a martingale representation may have numerical disadvantages since components in $\xi(t)$ grow exponentially over time. A common approach is to therefore shift the variables to have a mean reverting drift (Andersen and Piterbarg, 2010, pg. 481).

Andersen and Piterbarg (2010) go on to demonstrate one particular construction of a multi-factor model, which is summarised here. Begin with the following definitions

$$\Phi(t) = \mathrm{diag}(\varphi(t)) = \begin{bmatrix} \varphi_1(t) & 0 & \ddots & 0 \\ 0 & \varphi_2(t) & \ddots & \ddots \\ \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & 0 & \varphi_d(t) \end{bmatrix}, \tag{3.16}$$

$$\kappa(t) = -\frac{d\Phi(t)}{dt}\Phi(t)^{-1} = \mathrm{diag}((\kappa_1(t), \kappa_2(t))^\top). \tag{3.17}$$

Here $\kappa(t)$ is a $d \times d$ dimensional diagonal matrix, and we assume that $\Phi(t)$ is invertible (i.e. that $\varphi_i(t) \neq 0, i = 1, \ldots d$).

In addition, set

$$x(t) = \Phi(t) \int_0^t \xi(s)^\top \xi(s) \int_s^t \varphi(u) du\, ds + \Phi(t) \int_0^t \xi(u)^\top dW(u), \tag{3.18}$$

$$y(t) = \Phi(t) \left(\int_0^t \xi(s)^\top \xi(s) ds \right) \Phi(t). \tag{3.19}$$

Here $x(t)$ is a d-dimensional random vector, and $y(t)$ is a deterministic $d \times d$ symmetric matrix.

Proposition 3.4. *Let the forward rate volatility be separable (as in Lemma 1.3). Then, for $\mathbf{1} = (1, 1, \ldots, 1)^\top \in \mathbb{R}^d$, and for $\kappa(t)$, $x(t)$, $y(t)$ defined as in (3.17), (3.18), and (3.19),*

$$dx(t) = (y(t)\mathbf{1} - \kappa(t)x(t))\, dt + \sigma_x(t)^\top dW(t), \quad \sigma_x(t) = \xi(t)\Phi(t), \tag{3.20}$$

and, with $M(t,T) := \Phi(T)\Phi(t)^{-1}\mathbf{1}$,

$$f(t,T) = f(0,T) + M(t,T)^\top \left(x(t) + y(t) \int_t^T M(t,u) du \right). \tag{3.21}$$

In particular, for $T = t$, we have

$$r(t) = f(t,t) = f(0,t) + \mathbf{1}^\top x(t) = f(0,t) + \sum_{i=1}^d x_i(t). \tag{3.22}$$

This is labelled Proposition 12.1.2 in Andersen and Piterbarg (2010) and the proof is therein. As a sketch, (3.20) is proved by applying the Leibniz integration rule to $x(t)$. (3.21) follows by recognizing that $\varphi(T)^\top = \mathbf{1}^\top \Phi(T)$ in (3.14), and then by an exercise in algebraic manipulation and in applying the definition of $M(t,T)$.

Proposition 3.4 therefore defines a multi-factor Gaussian model that is consistent with the workings developed for the single-factor model in Section 2.2. For the important case of $d = 2$, we have that $r(t) = f(0,t) + x_1(t) + x_2(t)$, in contrast to the single-factor case where $r(t) = f(0,t) + x(t)$. In what follows, we will demonstrate the two-factor case in an effort to make the theory more practical.

First, set

$$\xi(t) = \begin{bmatrix} \sigma_{11}(t)e^{\int_0^t \kappa_1(u)du} & 0 \\ \sigma_{21}(t)e^{\int_0^t \kappa_1(u)du} & \sigma_{22}(t)e^{\int_0^t \kappa_2(u)du} \end{bmatrix}, \quad \varphi(t) = \begin{bmatrix} e^{-\int_0^t \kappa_1(u)du} \\ e^{-\int_0^t \kappa_2(u)du} \end{bmatrix}. \tag{3.23}$$

$\xi(t)$ is assumed, without loss of generality, to be lower diagonal. We then have that $x(t) = (x_1(t), x_2(t))^\top$ satisfies (3.20), with

$$\sigma_x(t) = \xi(t)\Phi(t) = \begin{bmatrix} \sigma_{11}(t) & 0 \\ \sigma_{21}(t) & \sigma_{22}(t) \end{bmatrix}, \quad x(0) = 0. \tag{3.24}$$

Since $x(t)$ has instantaneous variance-covariance matrix

$$\sigma_x(t)^\top \sigma_x(t) = \begin{bmatrix} \sigma_{11}(t)^2 + \sigma_{21}(t)^2 & \sigma_{21}(t)\sigma_{22}(t) \\ \sigma_{22}(t)\sigma_{21}(t) & \sigma_{22}(t)^2 \end{bmatrix}, \tag{3.25}$$

it is easy to see that the instantaneous correlation between $x_1(t)$ and $x_2(t)$ is

$$\rho_x(t) = \frac{\sigma_{21}(t)\sigma_{22}(t)}{\sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2}\sqrt{\sigma_{22}(t)^2}}. \tag{3.26}$$

In other words, we can reparametrize (3.20) as

$$dx(t) = (y(t)\mathbf{1} - \kappa(t)x(t))\,dt + \sigma_x^*(t)dW^*(t), \tag{3.27}$$

where now $dW_1^*(t)dW_2^*(t) = \rho_x(t)dt$ and $\sigma_x^*(t)$ is diagonal with non-negative entries

$$\sigma_x^*(t) = \begin{bmatrix} \sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2} & 0 \\ 0 & \sqrt{\sigma_{22}(t)^2} \end{bmatrix} := \begin{bmatrix} \sigma_1(t) & 0 \\ 0 & \sigma_2(t) \end{bmatrix}. \tag{3.28}$$

Andersen and Piterbarg (2010) consider the specification of the two-factor model in (3.27) more intuitive. This specification is also more consistent with the presentation of Brigo and Mercurio's G2++ model in Section 3.1, and is in some senses more amiable when it comes to deriving the variance from first principles.

Proposition 3.5. *The price at time t of a zero-coupon bond maturing at time T with unit face value is*

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left(-G(t,T)^\top x(t) - \frac{1}{2}G(t,T)^\top y(t)G(t,T)\right), \tag{3.29}$$

where

$$G(t,T) = \int_t^T M(t,u)du, \quad M(t,T) = \left(e^{-\int_t^T \kappa_1(u)du},\, e^{-\int_t^T \kappa_2(u)du}\right)^\top. \tag{3.30}$$

Proof. The result follows after an exercise in integration and algebraic manipulations when substituting the expression for $f(t,T)$ from (3.21) into $P(t,T) = \exp(-\int_t^T f(t,u)du)$.

By integrating (3.20) for $t_1 < t_2$, we get that

$$x(t_2) = e^{-\int_{t_1}^{t_2} \kappa(u)du} x(t_1) + \int_{t_1}^{t_2} e^{-\int_s^{t_2} \kappa(u)du} y(s)\mathbf{1}\, ds \\ + \int_{t_1}^{t_2} e^{-\int_s^{t_2} \kappa(u)du} \sigma_x(s)^\top dW(s), \tag{3.31}$$

so that $x(t_2)$ conditional on $\mathcal{F}_{t_1}$ is d-dimensional Gaussian with mean

$$\mathbb{E}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right] = e^{-\int_{t_1}^{t_2} \kappa(u)du} x(t_1) + \int_{t_1}^{t_2} e^{-\int_s^{t_2} \kappa(u)du} y(s)\mathbf{1}\, ds, \tag{3.32}$$

and variance-covariance matrix

$$\mathbb{V}\text{ar}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right] = \int_{t_1}^{t_2} e^{-\int_s^{t_2} \kappa(u)du} \sigma_x(s)\sigma_x(s)^\top e^{-\int_s^{t_2} \kappa(u)^\top du} ds. \tag{3.33}$$

Next, to compute

$$V_t = \mathbb{E}^{\mathbb{Q}}\left[X_T e^{-\int_t^T r(u)du} \,\middle|\, \mathcal{F}_t\right] = P(t,T)\mathbb{E}^{\mathbb{Q}}\left[X_T^x e^{-\int_t^T \mathbf{1}^\top x(u)du} \,\middle|\, \mathcal{F}_t\right], \tag{3.34}$$

it remains to simulate the quantity $-\int_t^T \mathbf{1}^\top x(u)du$. Define $I(T) = -\int_0^T \mathbf{1}^\top x(u)du$, then, as in Section 2.2, we need to derive the joint distribution of $(x(t_2), I(t_2))$ conditional on $\mathcal{F}_{t_1}$ for bias-free joint simulation of the short rate and discount factor. In this instance, however, $x(t)$ is a 2-dimensional column vector, and so what we are after is a 3-dimensional covariance matrix as in (3.12).

As in the single-factor case ((2.26) and (2.27)), $I(T)$ is Gaussian with moments

$$\mathbb{E}\left[I(t_2)\,|\,\mathcal{F}_{t_i}\right] = -\int_0^{t_2} \mathbb{E}\left[\mathbf{1}^\top x(u)\,|\,\mathcal{F}_{t_1}\right] du \\ = I(t_1) - G(t_1,t_2)^\top x(t_1) - \int_{t_1}^{t_2}\int_{t_1}^{u} \mathbf{1}^\top e^{-\int_s^u \kappa(v)dv} y(s)\mathbf{1}\, ds\, du, \tag{3.35}$$

$$\mathbb{V}\text{ar}\left[I(t_2)\,|\,\mathcal{F}_{t_1}\right] = 2\int_{t_1}^{t_2}\int_{t_1}^{u} \mathbb{C}\text{ov}_{t_1}\left[\mathbf{1}^\top x(u), \mathbf{1}^\top x(s)\right] ds\, du \\ = 2\int_{t_1}^{t_2}\int_{t_1}^{u}\int_{t_1}^{s} \mathbf{1}^\top e^{-\int_v^u \kappa(u)du} \sigma_x(v)\sigma_x(v)^\top e^{-\int_v^s \kappa(u)^\top du} \mathbf{1}\, dv\, ds\, du. \tag{3.36}$$

The covariance between $I(t_2)$ and the vector $x(t_2)$ conditional on $\mathcal{F}_{t_1}$ is

$$\begin{aligned}\mathbb{C}\mathrm{ov}_{t_i}\left[I(t_{i+1}),\, x(t_{i+1})\right] &= -\int_{t_i}^{t_{i+1}} \mathbb{C}\mathrm{ov}_{t_i}\left[\mathbf{1}^\top x(u), x(t_{i+1})\right] du \\ &= -\int_{t_i}^{t_{i+1}} \int_{t_i}^{s} \mathbf{1}^\top e^{-\int_v^{t_{i+1}} \kappa(s)ds} \sigma_x(v)\sigma_x(v)^\top e^{-\int_v^u \kappa(s)^\top ds} dv\, du,\end{aligned} \tag{3.37}$$

so that the correlation is computed as

$$\rho_{xI}(t_1, t_2) = \frac{\mathbb{C}\mathrm{ov}_{t_1}\left[I(t_2),\, x(t_2)\right]}{\sqrt{\mathbb{V}\mathrm{ar}\left[I(t_2)\,|\,\mathcal{F}_{t_1}\right]}\sqrt{\mathbb{V}\mathrm{ar}\left[x(t_2)\,|\,\mathcal{F}_{t_1}\right]}}. \tag{3.38}$$

(3.38) is a 2-dimensional cross-correlation vector with entries for the correlation between $x_1(t_2)$ and $I(t_2)$, and $x_2(t_2)$ and $I(t_2)$. Finally, using (3.38), as well as the covariance matrix in (3.33), we can construct a 3×3 correlation matrix for $(x_1(t),\, x_2(t),\, I(t))$ as in (3.12) for the G2++ model. Using this, Monte Carlo simulation proceeds in the same way as in (3.13).

Chapter 4

Monte Carlo Option Pricing

As an implementation exercise of the multi-factor models presented in Section 3.1 and Section 3.2, this chapter is concerned with the Monte Carlo pricing of European caplets and floorlets (Section 4.1), caps and floors (Section 4.2), and swaptions (Section 4.3). Since we have closed-form solutions for these instruments at our disposal, we will then be able to examine the extent of bias in option pricing where the discount factor is simulated via quadrature. The chapter concludes with a comparison of the bias-free methods of Brigo and Mercurio (2006) and Andersen and Piterbarg (2010), as well as simulation via a change of numéraire.

In all the simulations that follow, we will make the following assumptions:

Assumption 4.1. Assume a two-factor Gaussian short rate model with mean reversion parameters $\kappa_1 = 0.1$, $\kappa_2 = 0.25$; volatilities $\sigma_1 = 0.025$, $\sigma_2 = 0.02$, and correlation $\rho = -0.6$.[1]

Assumption 4.2. Assume the initial term structure of interest rates is governed by a Vašíček (1977) short rate process (as in (1.1)) with mean reversion rate $\kappa = 0.1$, mean reversion level $\vartheta = 0.09$, volatility $\sigma = 0.04$, and initial short rate $r_0 = 0.07$.

4.1 Caplets and Floorlets (Zero-Coupon Bond Options)

This section will begin by illustrating that caplets and floorlets are equivalent to certain European zero-coupon bond (ZCB) options, for which we have closed-form pricing solutions. Subsequently, in Section 4.2, we will be able to price a cap (floor), which is simply a strip of caplets (floorlets). We will see that, where possible, using a change of numéraire for option valuation results in significantly lower Monte Carlo error. Then, using the closed-form prices, we will compare the schemes of Brigo and Mercurio (2006) and Andersen and Piterbarg (2010) to procedures which approximate the numéraire via simple and trapezoidal quadrature.

[1] ρ_x in Andersen and Piterbarg (2010)

A caplet (floorlet) is a call (put) option on a floating rate, and protects the holder against rising (falling) interest rates over a period from T_1 to T_2. It is specified by a floating rate, a fixed strike rate K, and a notional amount F. We will assume zero spread, and work with in-arrears caplets (floorlets). A caplet has payoff, at time T_2,

$$F\left[L(T_1,T_2)-K\right]^+\Delta_1, \tag{4.1}$$

where Δ_1 denotes the year fraction applicable over the period T_1 to T_2. $L(T_1,T_2)$ denotes the simple floating rate observable at time T_1, which applies over the period T_1 to T_2.

Similarly, a floorlet has payoff, at T_2,

$$F\left[K-L(T_1,T_2)\right]^+\Delta_1. \tag{4.2}$$

Proposition 4.3. *The price, at time t, of an in-arrears caplet (floorlet) for period T_1 to T_2, with nominal value F, and strike K, is equivalent to the price of a European put (call) option on a zero-coupon bond $P(t,T_2)$, with expiry T_1, and with face value*

$$F' = F(1+K\Delta_n), \tag{4.3}$$

and strike

$$K' = \frac{1}{1+K\Delta_n}. \tag{4.4}$$

More succinctly,

$$\boldsymbol{Cpl}(t,T_1,T_2,F,K) = \boldsymbol{ZBP}(t,T_1,T_2,F',K'), \tag{4.5}$$

and

$$\boldsymbol{Fll}(t,T_1,T_2,F,K) = \boldsymbol{ZBC}(t,T_1,T_2,F',K'). \tag{4.6}$$

Proof. Consider the caplet cashflow in (4.1), and note that this is equivalent to a cashflow of

$$\frac{F\left[L(T_1,T_2)-K\right]^+\Delta_n}{1+L(T_1,T_2)\Delta_n}, \tag{4.7}$$

at time T_1, which, on rearrangement is equivalent to

$$F(1+K\Delta_1)\left(\frac{1}{1+K\Delta_1}-\frac{1}{1+L(T_1,T_2)\Delta_1}\right)^+. \tag{4.8}$$

Lastly, recognize (4.8) as the payoff of $F(1+K\Delta_1)$-many put options on $P(t,T_2)$, with strike $\frac{1}{1+K\Delta_n}$ and expiry T_1. The proof follows similarly for floorlets.

The analytical prices of European ZCB call and put options, ZBC and ZBP respectively, for a two-factor linear additive GSR model are left to Appendix A.2.

Figure 4.1 illustrates the Monte Carlo price estimates of a floorlet option as a function of sample size, as well as a three standard deviation bound for these estimates around the closed-form price. The option parameters used were $T_1 = 1$,

$T_2 = 2$, $F = 10$, $K = 0.25$. The Monte Carlo prices are computed using both the Brigo and Mercurio (2006) 'standard simulation' method from Section 3.1 and the Andersen and Piterbarg (2010) method from Section 3.2. In addition, we plot the Monte Carlo option price using a change of numéraire to the measure associated with the ZCB maturing at T_1.

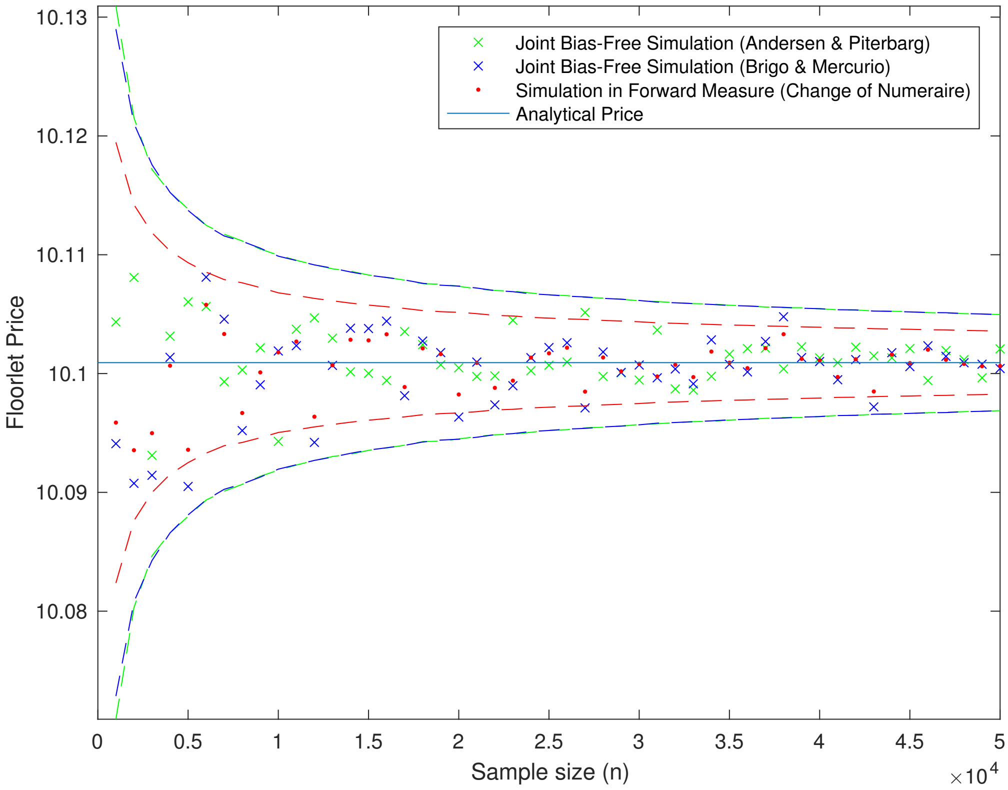


Fig. 4.1: Comparison of floorlet option price as a function of sample size using the methods of Section 3.1 and Section 3.2, as well as via a change of numéraire. Option parameters: $T_1 = 1, T_2 = 2, F = 10, K = 0.25$.

From Figure 4.1, it is immediately clear that the Monte Carlo three standard deviation bounds for the simulation in the T-forward measure, $\mathbb{Q}^T$, are narrower than for both of the joint and bias-free simulation methods considered. This follows intuition, since under the forward measure we only need to simulate two processes, x_1 and x_2, whereas in the joint and bias-free simulation methods, we need to simulate the processes x_1, x_2, *and* the discount factor, or more specifically the integral of $x_1 + x_2$[2]. As opposed to using the observable bond price today as a discount fac-

[2] $Y(t,T) = \int_t^T [x(u) + y(u)]du$ in the case of Brigo and Mercurio (2006) and $I(T) = -\int_0^T [x_1(u) +$

tor, the need to simulate this additional quantity introduces additional stochastic variability — and hence the wider bounds.

Also, notice that the error bounds for both joint and bias-free methods overlap. This follows since the variance functions for the state variables x_1 and x_2 are identical under both methods, as well as the variance functions for quantities $Y(t, T)$ and $I(T)$ (conditional on $\mathcal{F}_t$).

On closer inspection of Figure 4.1, observe that the estimates for the change of numéraire technique appear related to the estimates of the Brigo and Mercurio (2006) method. This is no accident, and arises because the former estimates were derived by using the method of Brigo and Mercurio (2006). That is, using the same short rate realizations[3] used in the bias-free method of Brigo and Mercurio (2006). The differences arise, because, in the latter method, the discount factor is stochastic. The use of the same short rate realisations are deliberate, in an attempt to highlight the effect of the stochastic discount factor on the price estimates.

For the option parameters in Figure 4.1, the three standard deviation bounds using a change of numéraire are approximately 65% of the size of the bounds using the joint and bias-free simulation methods. From the Central Limit Theorem, we know that Monte Carlo integration converges with $\sqrt{n}$. This means that we can achieve the same error bound using $0.65^2 \approx 42\%$ of the sample. In other words, simulation via a change of numéraire is, for the parameters used in Figure 4.1, approximately 2.3 times more efficient when measured by sample size.

The emphasis on the parameters used in Figure 4.1 for the percentages quoted above is due to the fact that the error bounds change as the option parameters change. For example, with $T_1 = 10$, $T_2 = 11$ (i.e. when there is greater uncertainty, or variance, on the discount factor), the size of the error bounds on the change of numéraire simulation are approximately 15% of the size of the bounds when using the joint and bias-free simulation methods. This translates to a mere $0.15^2 \approx 2\%$ of the sample size on the joint and bias-free methods for the same error bound — 46 times more efficient.

Figure 4.2 plots the floorlet option price as in Figure 4.1, but where the Monte Carlo sample size for the simulation via a change in numéraire is reduced to be 42% of the sample size used in Figure 4.1. The upper x-axis reflects this change. As postulated above, we see that the three standard deviation error bounds coincide for all three simulation methods.

Next, we consider comparing the joint and bias-free methods to simulations where the numéraire is approximated using quadrature. As we have seen above,

$x_2(u)]du$ in the case of Andersen and Piterbarg (2010).

[3] More specifically, the same realisations of x_1 and x_2.

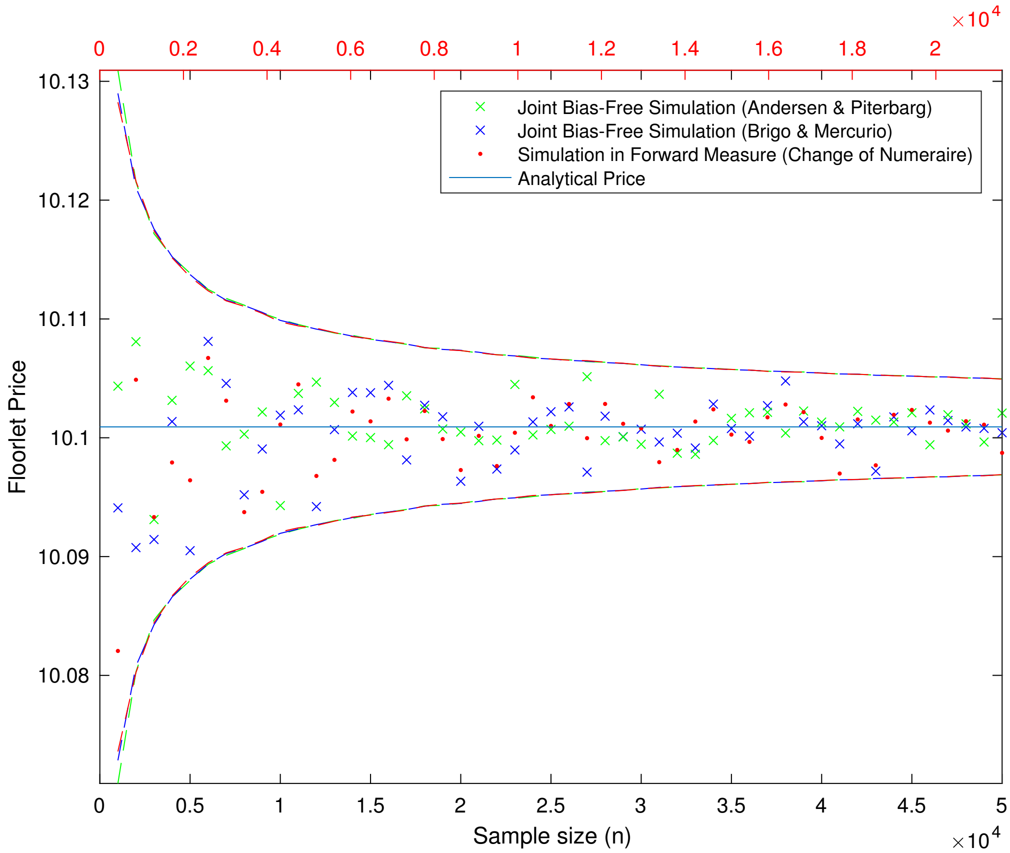


Fig. 4.2: Comparison of floorlet option price as a function of sample size using the methods of Section 3.1 and Section 3.2, as well as via a change of numéraire with reduced sample size. Option parameters: $T_1 = 1$, $T_2 = 2$, $F = 10$, $K = 0.25$.

both joint and bias-free methods have the same error bounds, and so for easier comprehension, we will only consider comparing quadrature approximations against the method of Andersen and Piterbarg (2010). We consider simple quadrature, which, in the case of Andersen and Piterbarg's work approximates $I(T)$ on the schedule $\{t_i\}_{i=0}^{N}$ as

$$I(T) \approx -\mathbf{1}^\top \sum_{i=1}^{N} x(t_i)\Delta_i, \tag{4.9}$$

and, the more accurate trapezoidal quadrature

$$I(T) \approx -\mathbf{1}^\top \sum_{i=1}^{N} \left(x(t_{i-1}) + x(t_i)\right) \frac{\Delta_i}{2}, \tag{4.10}$$

where $\Delta_i = t_i - t_{i-1}$.

In previous simulations (methods of Brigo and Mercurio (2006) and Andersen and Piterbarg (2010) in Figure 4.1 and Figure 4.2), it was sufficient (and efficient) to 'long-step' the short rate simulation to time T_1: the floorlet (or caplet) option payoff depends only on the realised short rate at time T_1, and not on the path leading up to T_1. However, for quadrature to be viable, we cannot long-step the short rate to time T_1, since this will make the approximation of $I(T_1)$ inaccurate. We therefore need to simulate an entire path up to T_1, with more points in the schedule leading to improved accuracy. Of course, this comes at the expense of computation time.

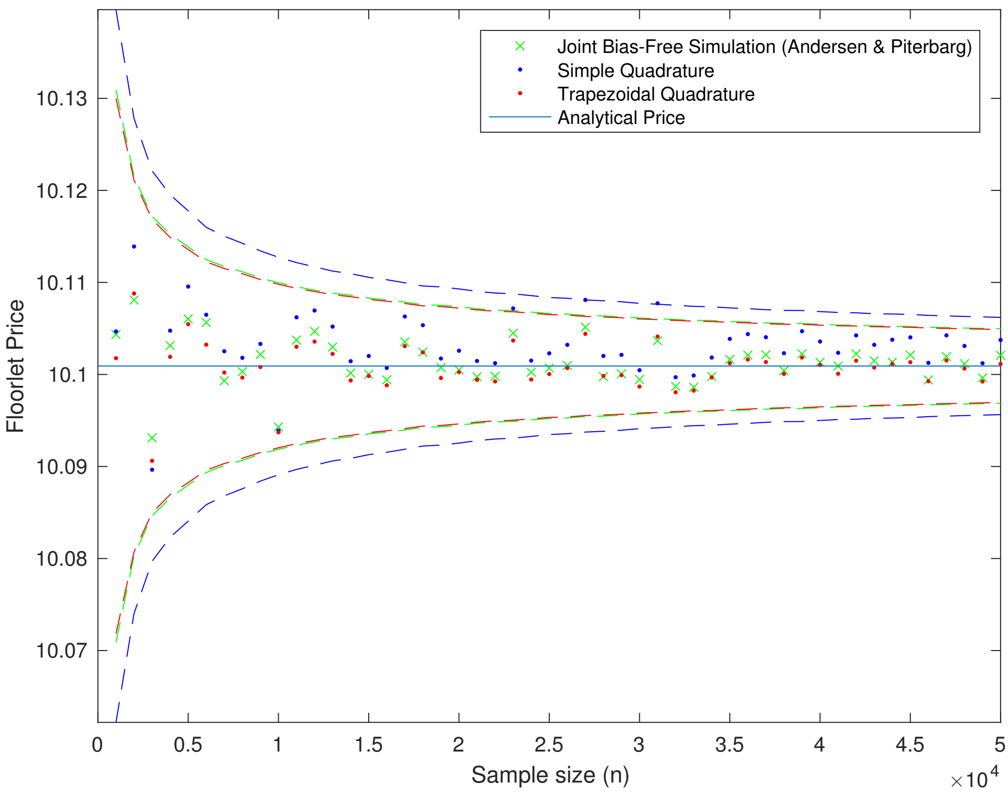


Fig. 4.3: Comparison of floorlet option price as a function of sample size using exact simulation (Section 3.2) as well as simple right-hand sum and trapezoidal quadrature with a single time-step. Option parameters: $T_1 = 1$, $T_2 = 2$, $F = 10$, $K = 0.25$.

Figure 4.3 compares the Monte Carlo price estimates for the exact simulation as well as the two quadrature approximations where the short rate was 'long-stepped' to time T_1[4]. While there is still an element of bias in the trapezoidal approximation,

[4] i.e. $N = 1$ in (4.9) and (4.10).

it follows the bias-free simulation fairly closely, even with a large time-step. On the other hand, the simple quadrature performs poorly, with 44/50 Monte Carlo estimates above the analytical solution. If the approximation were unbiased, we would expect that an estimate would over/under-estimate the true price with 50% probability. Under this hypothesis, the observed statistic, 44/50, would occur with virtually 0% probability[5]. The trapezoidal approximation has no noticeable impact on computation time when compared to simple quadrature.

This bias can be controlled by inserting extra dates in the schedule. Andersen and Piterbarg (2010) write that it is often more convenient to change measure or to compute $I(T)$ by numerical integration as in (4.9). The former technique can not be used for all types of options (including Bermudan swaptions, ratchet options etc.), and the latter becomes more and more inefficient as option maturity increases. This therefore highlights an advantage of using a bias-free method.

4.2 Caps and Floors

An interest rate cap (floor) protects the holder against rising (falling) interest rates. It consists of a strip of caplets (floorlets), where each caplet (floorlet) is a call (put) option on a floating rate. A cap, or floor, is specified by a tenor structure $\mathcal{T} = \{T_0, T_1, \dots, T_N\}$, a floating rate, a fixed strike rate K, and a notional amount F. The tenor structure $\mathcal{T}$ is the set of all payment dates augmented with the first reset date. Again, we assume zero spread, and work with in-arrears caps and floors. As stated by Fries (2016), this latter assumption is an important test for the correct conditional numéraire sampling.

A cap consists of N caplets, where the n^{th} caplet c_n has payoff

$$c_n(T_n) = F\left[L(T_{n-1}, T_n) - K\right]^+ \Delta_n \quad \text{at time } T_n, \tag{4.11}$$

for $n > 0$, and where Δ_n denotes the year fraction applicable over the period T_{n-1} to T_n. $L(T_{n-1}, T_n)$ denotes the simple floating rate observable at time T_n and which applies over the period T_{n-1} to T_n.

Similarly, an interest rate floor consists of N floorlets, with each floorlet f_n paying

$$f_n(T_n) = F\left[K - L(T_{n-1}, T_n)\right]^+ \Delta_n \quad \text{at time } T_n. \tag{4.12}$$

Pricing such an instrument is simple once we know how to price caplets and floorlets. Let $\Delta = \{\Delta_1, \dots, \Delta_N\}$ be the set of year fractions corresponding to $\mathcal{T}$. Then, following from Proposition 4.3, we have that the time t price of a cap, and

[5] $p < 0.00000002$.

floor, is given by

$$\mathbf{Cap}(t,\mathcal{T},\Delta,F,K)=\sum_{i=1}^{n}\mathbf{Cpl}(t,T_i,T_{i-1},F,K)=\sum_{i=1}^{n}\mathbf{ZBP}(t,T_i,T_{i-1},F_i',K_i'), \quad (4.13)$$

and

$$\mathbf{Flr}(t,\mathcal{T},\Delta,F,K)=\sum_{i=1}^{n}\mathbf{Fll}(t,T_i,T_{i-1},F,K)=\sum_{i=1}^{n}\mathbf{ZBC}(t,T_i,T_{i-1},F_i',K_i'). \quad (4.14)$$

The use of subscript i on F_i' and K_i' is due to the fact that the notional (4.3) and strike (4.4) on the corresponding call or put option depends on the year fraction Δ_i.

Figure 4.4 plots the difference between the Monte Carlo estimate and the analytical solution for a floor option with an increasing number of floorlet options. The first data point on the x-axis corresponds to a floor option with $\mathcal{T}=\{0,1\}$[6], while the second data point corresponds to a floor option with $\mathcal{T}=\{0,1,2\}$, so that the payment dates are at time 1 and time 2, and so forth. The remaining option parameters were set at $F=10$, $K=0.25$, and $\Delta_i=1$ is assumed constant.

Figure 4.4 plots the Monte Carlo price deviation for the bias-free methods of Brigo and Mercurio (2006) and Andersen and Piterbarg (2010). Importantly, notice that 0 deviation is contained within the error bounds of both bias-free methods, as one would expect from a bias-free simulation. In addition, and this time to demonstrate the approximations relating to Brigo and Mercurio's method, we consider simple and trapezoidal approximations of $Y(t,T)$ with ten equidistant points in a one year period. As expected, the bias from a simple approximation increases as the number of floorlets increase. This follows since the error in the approximated discount factor increases over longer periods. Trapezoidal quadrature performs remarkably well, even when the number of points in the schedule is reduced. The computation time difference between the two approximation schemes is negligible, and so the trapezoidal approximation is always preferable.

The deviation from the analytical price is computed as the Monte Carlo estimate less the analytical price. Therefore, the increasing trend in the simple approximation means that it is over-estimating the actual price. The reason it over-estimates the price is because the instantaneous forward curve implied by Assumption 4.2 is downward sloping, and hence the simple (right-hand) quadrature approximation overstates the discount factor, and hence over-states the price estimate.

[6] This is simply a floorlet option with reset date 0 and payment date at time 1.

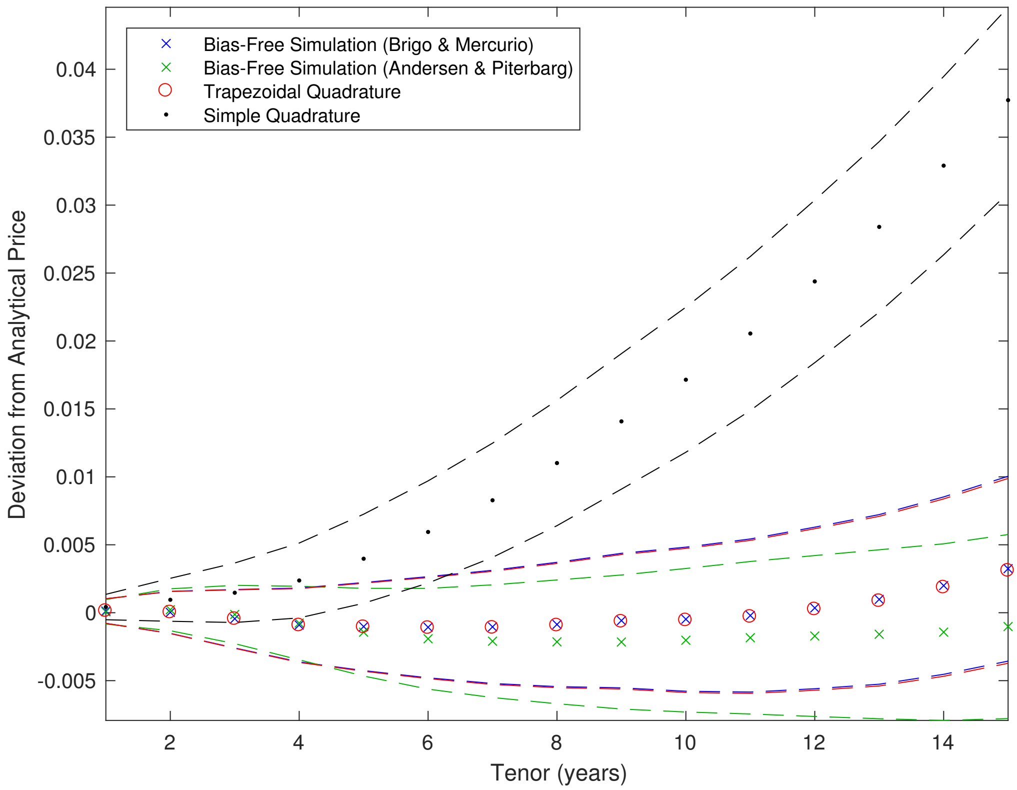


Fig. 4.4: Deviation of Monte Carlo floor option price from analytical solution, using 1,000,000 samples, as the number of floorlet options in the floor option increase. Compares the bias-free schemes to simple and trapezoidal quadrature applied to the scheme of Brigo and Mercurio (2006) with 10 time-points in a year. Option parameters: $\mathcal{T} = \{0, 1, 2, \dots 15\}$, $F = 10$, $K = 0.25$.

4.3 Swaptions

A payer (receiver) swaption (short for "swap option") is the option to enter a pay-fixed (pay-floating) swap at some future date T at a strike rate K. If the tenor structure of the swap is $\mathcal{T} = \{T = T_0, T_1, T_2, \dots, T_N\}$ and the notional is F, then the swap gives the holder the right (but no obligation) to receive payments

$$F\omega\left[L(T_{n-1}, T_n) - K\right]\Delta_n \quad \text{at time } T_n, \tag{4.15}$$

for $n > 0$, and where $\omega = 1$ $(\omega = -1)$ for a payer (receiver) swaption.

Assuming zero spread, entering into a swap at T entitles the holder to receive

$$F\omega\left[L(T_{n-1}, T_n) - S_T\right]\Delta_n \quad \text{at time } T_n, \tag{4.16}$$

for $n > 0$, and where S_T is the fair swap rate, so that the swap has zero initial value. The holder of a payer (receiver) swaption will therefore exercise if $K \leq S_T$ ($K \geq S_T$), and discard otherwise, so that the net payoff at maturity T is

$$F\omega \left[S_T - K\right]^+ \sum_{n=1}^{N} P(T, T_n)\Delta_n. \tag{4.17}$$

Figure 4.5 plots the Monte Carlo price estimates of a payer swaption as a function of sample size using the bias-free method of Section 3.1, as well as simple and trapezoidal quadrature on $Y(t, T)$. The error bounds for simple and trapezoidal quadrature approximations overlap with the error bounds using bias-free simulation. In addition, we plot the Monte Carlo swaption price estimates using a floorlet option with parameters $T_1 = 10$, $T_2 = 11$, $F = 10$, $K = 0.25$ as a control variate. The analytical price of a European swaption for a two-factor linear additive GSR model can be found in Appendix A.2.

In Section 4.1, we saw that simulation under the forward measure resulted in significantly narrower Monte Carlo error bounds for floorlet prices. Specifically, with parameters $T_1 = 10$, $T_2 = 11$, $F = 10$, $K = 0.25$, the error bounds were approximately 15% of the size of the bounds when using the joint and bias-free simulation methods. It is for this reason that we use the floorlet option, simulated under the forward measure, as a control variate for pricing swaptions. Figure 4.5 illustrates that by using the added information of the control variates, we are able to significantly reduce the size of error bounds. In Figure 4.5, using control variates reduces the Monte Carlo error bound by 44%.

As a comparison, Figure A.1 uses the same floorlet option as a control variate, but where the short rate and discount factor used to price the floorlet option are simulated under the risk-neutral measure using Brigo and Mercurio's method. As one would expect, the error bounds using the control variates are narrower than without, but are not as pronounced as in Figure 4.5. When simulating the floorlet option price under $\mathbb{Q}$, the error bounds using control variates are reduced by 16%.

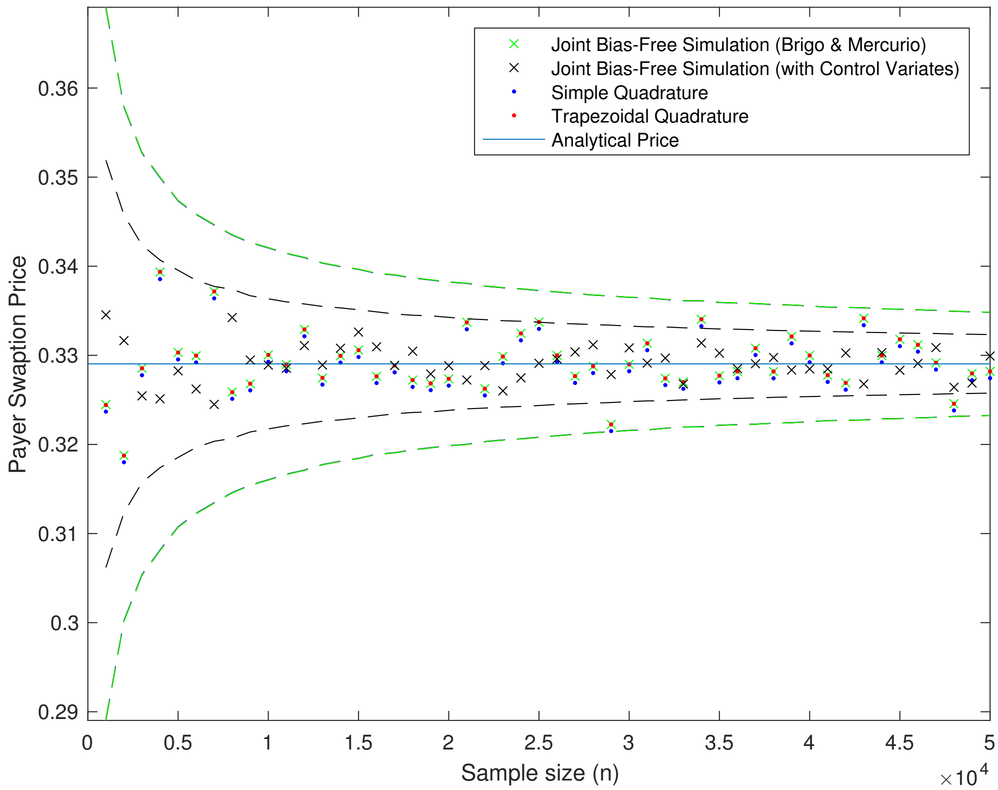


Fig. 4.5: Comparison of payer swaption price as a function of sample size using exact simulation (Section 3.1) as well as simple and trapezoidal quadrature with a ten time-points in a one-year period. Option parameters: $T = 10$, $\mathcal{T} = \{10, 11, 12, \ldots, 19\}$, $F = 10$, $K = 0.05$.

4.4 Conclusions

In Section 1.1 we examined short rate modelling paradigms over time, and saw that single-factor models imply that the evolution of a single state variable is responsible for all interest rate sensitive option prices, and that shocks across yield curves were perfectly correlated across all maturities. In addition, we saw that single-factor models may fail to capture a significant portion of market-implied interest rate variability. We used multi-factor models to address some concerns.

We saw that GSR models were amiable to closed-form solutions for certain options. The analytical tractability of GSR models allowed for the joint and bias-free simulation of the short rate and the discount factor, a central problem in risk-neutral option pricing (Section 4.1).

In Chapter 2, we derived exact and efficient bias-free sampling schemes for the short rate and the discount factor using three different approaches. The 'standard' approach (Section 2.1), involves simulation from the short rate directly, whereas the method by Andersen and Piterbarg (Section 2.2) requires simulation from some transform of the short rate. The third method (Section 2.3) uses a deterministic shift representation of the short rate, and so simulates from an underlying stochastic process. The methods produce the same error bounds.

In Chapter 3, we extend the single-factor methods of Section 2.1 and Section 2.2. Andersen and Piterbarg (2010) capture their single (Section 2.2) and multi-factor (Section 3.2) model in one formulation: the multi-factor representation, in matrix form, reduces simply and elegantly to the single-factor case. In contrast, Brigo and Mercurio (2006) do not present their G2++ model (Section 3.1) in matrix form, and so add terms to the single-factor case. While more elegant, the Andersen and Piterbarg (2010) approach requires a firm grasp of linear algebra and matrix calculus, and some results in the multi-factor case can be elusive if one does not clearly understand the matrices involved. In contrast, the Brigo and Mercurio (2006) approach lends itself to more intuitive and familiar calculations, especially when compared to single-factor models.

In addition, the presentation by Brigo and Mercurio (2006) is more restrictive when compared to the presentation by Andersen and Piterbarg (2010), where parameters are not necessarily assumed constant. However, certain practical considerations may restrict the use of time-varying parameters. For example, the assumption of time-stationarity (Section 1.1.3) places certain restrictions on these parameters (Andersen and Piterbarg, 2010, pg. 493), as well as the added (and often significant) computational considerations that may arise when numerically evaluating iterated integrals.

In Chapter 4, we conclude by implementing the multi-factor methods of Section 3.1 and Section 3.2 to price European options. We compare the bias-free simulation methods to the techniques used by Brigo and Mercurio (2006) and Andersen and Piterbarg (2010), noting their limitations. We showed we are able to recover the values of European option prices independently of the time discretisation schedule using an exact, joint and bias-free implementation of a two-dimensional GSR.

Bibliography

Andersen, L. and Piterbarg, V. (2010). *Interest Rate Modeling – Volume II: Term Structure Models*, 1st edn, Atlantic Financial Press.

Black, F. and Karasinski, P. (1991). Bond and option pricing when short rates are lognormal, *Financial Analysts Journal* **47**(4): 52–59.

Brigo, D. and Mercurio, F. (2001). A deterministic–shift extension of analytically–tractable and time–homogeneous short–rate models, *Finance and Stochastics* **5**(3): 369–387.

Brigo, D. and Mercurio, F. (2006). *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*, Springer Finance, Springer Berlin Heidelberg.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica: Journal of the Econometric Society* pp. 385–407.

Fitch Ratings (2017). Fitch: $9.7t of neg yielding debt despite monetary normalization. Available at: https://www.fitchratings.com/site/pr/10014210.

Fries, C. P. (2016). A short note on the exact stochastic simulation scheme of the Hull-White model and its implementation. Available at SSRN: https://ssrn.com/abstract=2737091.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*, Vol. 53, 1 edn, Springer-Verlag New York.

Heath, D., Jarrow, R. and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica* **60**(1): 77–105.

Hull, J. C. and White, A. D. (1995). A note on the models of Hull and White for pricing options on the term structure: response, *The Journal of Fixed Income* **5**(2): 97–102.

Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities, *The Review of Financial Studies* **3**(4): 573–592.

Hull, J. and White, A. (1994). Branching out, *Risk* **7**(7): 34–37.

Jamshidian, F. and Zhu, Y. (1996). Scenario simulation: Theory and methodology, *Finance and stochastics* **1**(1): 43–67.

Ouwehand, P. (2017). *Forward Rate Models*, African Institute of Financial Markets and Risk Management (AIFMRM). Lecture notes for Stochastic Calculus for Finance II (unpublished).

Reichmann, W. (1961). *Use and abuse of statistics*, 1 edn, Chapman and Hall.

Stewart, T. and Thiart, C. (2005). *Topics in Probability Distribution Theory*, Department of Statistical Sciences, University of Cape Town. Lecture notes for Statistical Theory and Inference (unpublished).

Ueno, Y. (2017). Term structure models with negative interest rates, *Technical report*, Institute for Monetary and Economic Studies, Bank of Japan.

Vašíček, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2): 177 – 188.

Appendix A

Appendix

A.1 Fries (2016) Defined Functions

The following function definitions relate to the work in Section 2.3.

$$M(t) := \exp\left(\int_0^t \alpha(s)ds\right), \tag{A.1}$$

$$M(t,T) := \frac{M(t)}{M(T)} = \exp\left(-\int_t^T \alpha(s)ds\right), \tag{A.2}$$

$$V(t,T) := \int_t^T \left(\int_s^T \frac{1}{M(u)} du M(s)\sigma(s)\right)^2 ds, \tag{A.3}$$

$$B(t,T) := \int_t^T M(t,s)ds, \tag{A.4}$$

$$\tilde{a}(t_i,t_{i+1}) := \frac{1-M(t_i,t_{i+1})}{t_{i+1}-t_i} = \frac{1-\exp(-\int_{t_i}^{t_{i+1}} \alpha(s)ds)}{t_{i+1}-t_i}, \tag{A.5}$$

$$\tilde{b}(t_i,t_{i+1}) := \frac{B(t_i,t_{i+1})}{t_{i+1}-t_i}, \tag{A.6}$$

$$\sigma_i(t_i,t_{i+1}) := \sqrt{\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} (M(s,t_{i+1})\sigma(s))^2 ds}, \tag{A.7}$$

$$\gamma_i(t_i,t_{i+1}) := \sqrt{\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} (B(s,t_{i+1})\sigma(s))^2 ds}, \tag{A.8}$$

$$\rho_i(t_i,t_{i+1}) := \frac{1}{\sigma_i\gamma_i}\frac{1}{t_{i+1}-t_i}\int_{t_i}^{t_{i+1}} M(s,t_{i+1})B(s,t_{i+1})\sigma^2(s)ds. \tag{A.9}$$

The shorthand notation $\sigma_i = \sigma_i(t_i,t_{i+1})$, $\gamma_i = \gamma_i(t_i,t_{i+1})$, and $\rho_i = \rho_i(t_i,t_{i+1})$ is also used.

These definitions relate to the extended Vašíček model with SDE

$$dr(t) = (\theta(t) - \alpha(t)r(t))\,dt + \sigma(t)dW(t), \quad r(0) = r_0. \tag{A.10}$$

A.2 Analytical Prices for European Options under Two-Factor Additive Linear Gaussian Model

The following lemma is labelled Lemma 4.2.2 in Brigo and Mercurio (2006) and the proof is therein.

Lemma A.1. *The processes $x_1(t)$ and $x_2(t)$ from (3.2) evolve under the forward measure $\mathbb{Q}^T$ according to*

$$\begin{aligned} dx_1(t) &= \left[\kappa_1 x_1(t) - \frac{\sigma_1^2}{\kappa_1}(1-e^{-\kappa_1(T-t)}) - \rho\frac{\sigma_1\sigma_2}{\kappa_2}(1-e^{-\kappa_2(T-t)})\right]dt + \sigma_1 dW_1^T(t), \\ dx_2(t) &= \left[\kappa_2 x_2(t) - \frac{\sigma_2^2}{\kappa_2}(1-e^{-\kappa_2(T-t)}) - \rho\frac{\sigma_1\sigma_2}{\kappa_1}(1-e^{-\kappa_1(T-t)})\right]dt + \sigma_2 dW_2^T(t), \end{aligned} \tag{A.11}$$

where W_1^T and W_2^T are two correlated Brownian motions under $\mathbb{Q}^T$ with $dW_1^T(t)dW_2^T(t) = \rho dt$. Moreover, the explicit solutions to (A.11) are, for $t_1 \leq t_2 \leq T$,

$$\begin{aligned} x_1(t_2) &= x_1(t_1)e^{-\kappa_1(t_2-t_1)} - M_1^T(t_1,t_2) + \sigma_1\int_{t_1}^{t_2} e^{-\kappa_1(t_2-u)}dW_1^T(u), \\ x_2(t_2) &= x_2(t_1)e^{-\kappa_2(t_2-t_1)} - M_2^T(t_1,t_2) + \sigma_2\int_{t_1}^{t_2} e^{-\kappa_2(t_2-u)}dW_2^T(u), \end{aligned} \tag{A.12}$$

where

$$\begin{aligned} M_1^T(s,t) = &\left(\frac{\sigma_1^2}{\kappa_1^2} + \rho\frac{\sigma_1\sigma_2}{\kappa_1\kappa_2}\right)\left[1-e^{-\kappa_1(t-s)}\right] - \frac{\sigma_1^2}{2\kappa_1^2}\left[e^{-\kappa_1(T-t)} - e^{-\kappa_1(T+t-2s)}\right] \\ &- \frac{\rho\sigma_1\sigma_2}{\kappa_2(\kappa_1+\kappa_2)}\left[e^{-\kappa_2(T-t)} - e^{-\kappa_2 T - \kappa_1 t + (\kappa_1+\kappa_2)s}\right], \end{aligned} \tag{A.13}$$

$$\begin{aligned} M_2^T(s,t) = &\left(\frac{\sigma_2^2}{\kappa_2^2} + \rho\frac{\sigma_1\sigma_2}{\kappa_1\kappa_2}\right)\left[1-e^{-\kappa_2(t-s)}\right] - \frac{\kappa_1^2}{2\kappa_2^2}\left[e^{-\kappa_2(T-t)} - e^{-\kappa_2(T+t-2s)}\right] \\ &- \frac{\rho\sigma_1\sigma_2}{\kappa_1(\kappa_1+\kappa_2)}\left[e^{-\kappa_1(T-t)} - e^{-\kappa_1 T - \kappa_2 t + (\kappa_1+\kappa_2)s}\right]. \end{aligned} \tag{A.14}$$

A.2.1 Zero-Coupon Bond Options

The following theorem is labelled Corollary 4.2.2 in Brigo and Mercurio (2006) and the proof is therein.

Theorem A.2. *The price at time t of a European call option with maturity T and strike K, written on a zero-coupon bond with face value N and maturity S is given by*

$$\begin{aligned} \boldsymbol{ZBC}(t,T,S,N,K) = NP(t,S)\Phi&\left(\frac{\ln\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)} + \frac{1}{2}\Sigma(t,T,S)\right) \\ - P(t,T)K\Phi&\left(\frac{\ln\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)} - \frac{1}{2}\Sigma(t,T,S)\right), \end{aligned} \tag{A.15}$$

where

$$\Sigma(t,T,S)^2 = \frac{\sigma_1^2}{2\kappa_1^3}\left[1-e^{-\kappa_1(S-T)}\right]^2\left[1-e^{-2\kappa_1(T-t)}\right]$$
$$+\frac{\sigma_2^2}{2\kappa_2^3}\left[1-e^{-\kappa_2(S-T)}\right]^2\left[1-e^{-2\kappa_2(T-t)}\right]$$
$$+2\frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2(\kappa_1+\kappa_2)}\left[1-e^{-\kappa_1(S-T)}\right]\left[1-e^{-\kappa_2(S-T)}\right]\left[1-e^{-(\kappa_1+\kappa_2)(T-t)}\right].$$

Analogously, the price at time t of the corresponding put option is

$$\begin{aligned}\boldsymbol{ZBP}(t,T,S,N,K) = &-NP(t,S)\Phi\left(\frac{\ln\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)}-\frac{1}{2}\Sigma(t,T,S)\right)\\ &+P(t,T)K\Phi\left(\frac{\ln\frac{NP(t,S)}{KP(t,T)}}{\Sigma(t,T,S)}+\frac{1}{2}\Sigma(t,T,S)\right).\end{aligned}\tag{A.16}$$

A.2.2 Swaptions

The following theorem is labelled Theorem 4.2.3 in Brigo and Mercurio (2006) and the proof is therein.

Theorem A.3. *The arbitrage-free price at time t = 0 of the above European swaption is given by numerically computing the following one-dimensional integral:*

$$\boldsymbol{ES}(0,T,\mathcal{T},N,X,\omega) =$$
$$N\omega P(0,T)\int_{-\infty}^{+\infty}\frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x\sqrt{2\pi}}\left[\Phi(-\omega h_1(x))-\sum_{i=1}^{n}\lambda_i(x)e^{\gamma_i(x)}\Phi(-wh_2(x))\right]dx,$$

where ω = 1 (ω = -1) for a payer (receiver) swaption,

$$h_1(x) := \frac{\bar{y}-\mu_y}{\sigma_y\sqrt{1-\rho_{xy}^2}}-\frac{\rho_{xy}(x-\mu_x)}{\sigma_x\sqrt{1-\rho_{xy}^2}},\tag{A.17}$$

$$h_2(x) := h_1(x)+A_2(T,T_i)\sigma_y\sqrt{1-\rho_{xy}^2},\tag{A.18}$$

$$B(t,T) := \frac{P(0,T)}{P(0,t)}\exp\left(\frac{1}{2}[V(t,T)-V(0,T)+V(0,t)]\right)\tag{A.19}$$

$$\lambda_i(x) := c_i B(T,t_i)e^{-A_1(T,t_i)x},\tag{A.20}$$

$$\gamma_i(x) := -A_2(T,t_i)\left[\mu_y-\frac{1}{2}(1-\rho_{xy}^2)\sigma_y^2A_2(T,t_i)+\rho_{xy}\sigma_y\frac{x-\mu_x}{\sigma_x}\right],\tag{A.21}$$

$\bar{y} = \bar{y}(x)$ is the unique solution of the following equation

$$\sum_{i=1}^{n}c_iB(T,t_i)e^{-A_1(T,t_i)x-A_2(T,t_i)\bar{y}} = 1,\tag{A.22}$$

and

$$\mu_x := -M_1^T(0,T), \tag{A.23}$$
$$\mu_y := -M_2^T(0,T), \tag{A.24}$$
$$\sigma_x := \sigma_1 \sqrt{\frac{1-e^{-2\kappa_1 T}}{2\kappa_1}}, \tag{A.25}$$
$$\sigma_y := \sigma_2 \sqrt{\frac{1-e^{-2\kappa_2 T}}{2\kappa_2}}, \tag{A.26}$$
$$\rho_{xy} := \frac{\rho\sigma_1\sigma_2}{(\kappa_1+\kappa_2)\sigma_x\sigma_y}\left[1-e^{-(\kappa_1+\kappa_2)T}\right]. \tag{A.27}$$

$V(t,T)$, $A_i(t,T)$, and $P(t,T)$ are as defined in (3.8), (3.9), and (3.10) respectively.

A.3 Additional Plots

Figure A.1 plots the Monte Carlo price estimates of a payer swaption as a function of sample size using the bias-free method of Section 3.1, as well as simple and trapezoidal quadrature on $Y(t,T)$. In addition, Figure A.1 illustrates the Monte Carlo swaption price estimates using a floorlet option with parameters $T_1 = 10$, $T_2 = 11$, $F = 10$, $K = 0.25$ as a control variate.

Figure A.1 uses Monte Carlo price estimates of a floorlet option simulated under $\mathbb{Q}$, where the short rate and discount factor and jointly simulated, bias-free. Figure A.1 illustrates that by using the added information of the control variates, we are able to significantly reduce the size of error bounds. In Figure A.1, using control variates reduces the Monte Carlo error bound by 16%.

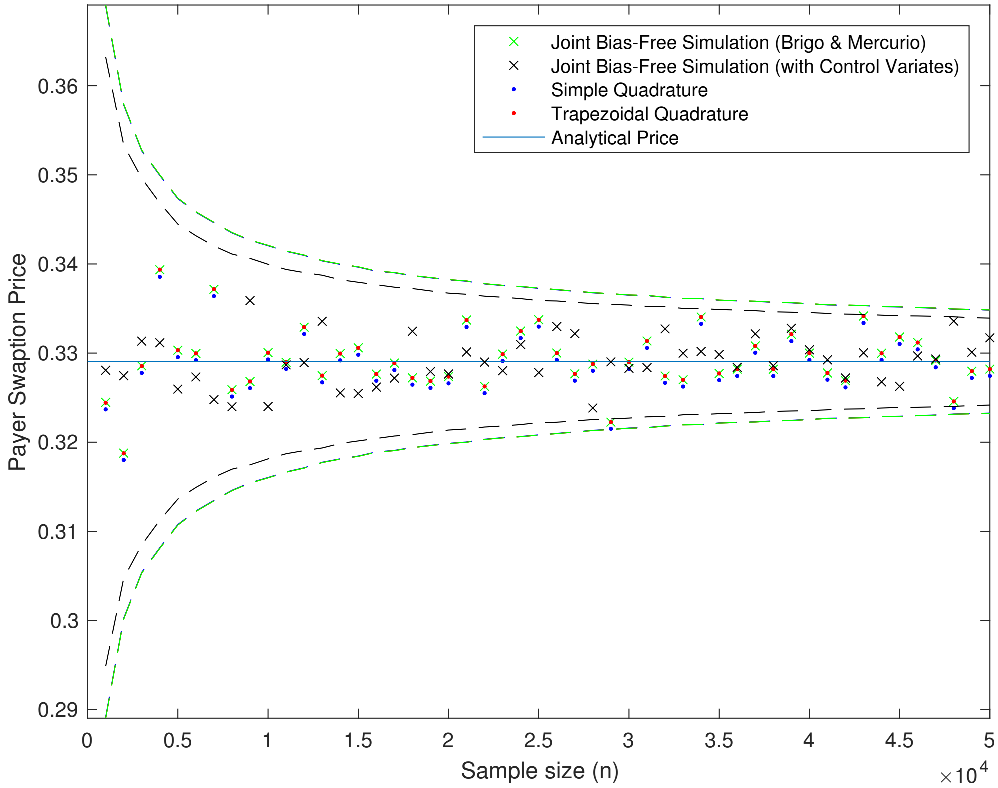


Fig. A.1: Comparison of payer swaption price as a function of sample size using exact simulation (Section 3.1) as well as simple and trapezoidal quadrature with a ten time-points in a one-year period. Option parameters: $T = 10$, $\mathcal{T} = \{10, 11, 12, \ldots, 19\}$, $F = 10$, $K = 0.05$.