Trolle-Schwartz HJM Interest Rate Model

Gareth William Schumann

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.

November 10, 2016

MPhil in Mathematical Finance,
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 in the University of Cape Town. It has not been submitted before for any degree or examination in any other University.

Signature removed

November 10, 2016

Acknowledgments

I wish to extend my deepest gratitude to my supervisors, Tom McWalter and Jörg Kienitz. Their guidance and insight have been invaluable. To my classmates, who brightened up every moment of this journey that we took together.

Finally, I wish to thank my family and those closest to me for providing the much needed support and encouragement when it mattered most.

Abstract

The Trolle and Schwartz (2009) interest rate model prices interest rate derivatives in a generalised stochastic volatility framework. It is a reformulation of the multi-factor Heath, Jarrow and Morton (1992) framework with stochastic volatility terms presented in an analogous fashion to the seminal Heston (1993) model. The Trolle and Schwartz (2009) model provides semi-analytical pricing formulas for zero-coupon bonds and zero-coupon bond options. These formulas are extended to price interest rate caplets, and therefore caps, as well as swaptions. These formulas are described as semi-analytical because of the use of numerical methods as well as their dependency on unobserved state variables. These state variables are estimated by applying an extended Kalman filter on a dataset of interest rates and interest rate derivative prices. Although Trolle and Schwartz (2009) confirm the accuracy of their model when testing against empirical prices, they do not provide an analysis of the consistency between the semi-analytical formulas and Monte Carlo pricing. Presenting this test for consistency seeks to confirm the validity of these pricing formulas. The aim of this dissertation is to implement the Trolle and Schwartz (2009) model and discuss the performance of the semi-analytical pricing formulas against a Monte Carlo simulation. Emphasis will be placed firstly on reviewing the derivations outlined in Trolle and Schwartz (2009) and secondly, building a Monte Carlo framework capable of comparing prices with the semi-analytical pricing formulas. Simulated data will be considered for the purpose of confirming that the estimation of the state vector is sufficiently accurate. Thereafter, an analysis on an empirical dataset can determine whether the results hold across different sets of data.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Interest rate models vary significantly in their characterisation of term structure dynamics as well as their methods for pricing interest rate derivatives. An important consideration is how to define the dynamics of interest rate volatility. Modelling this volatility stochastically is a desirable feature as it better reflects the stochastic nature of volatility in option pricing as well as addressing the observed price skews found in the market (Brigo and Mercurio, 2007).

Trolle and Schwartz (2009) introduce stochastic volatility into the multi-factor Heath, Jarrow and Morton (1992) (hereafter HJM) framework for modelling the instantaneous forward rate. The stochastic volatility terms are represented by unobserved volatility state variables in an analogous fashion to the Heston (1993) stochastic volatility model. These volatility state variables are correlated with the instantaneous forward rate which provides a mechanism to match price skew in interest rate derivative markets. The volatility state variables join with additional term structure specific state variables to form the full state vector for the model.

These state variables are incorporated in semi-analytical pricing formulas for bonds and interest rate derivatives. The formulas are described as semi-analytical as they require numerical methods and are conditional on the unobserved state vector. The state vector process is specified as a finite dimensional affine diffusion, placing this model in the Duffie and Kan (1996) class of affine dynamic term structure models. The affine structure proves useful when estimating the unobserved state variables. This structure holds under both the risk-neutral, $\mathbb{Q}$, and real-world, $\mathbb{P}$, measures by incorporating an appropriate choice for the market-price of risk.

Semi-analytical pricing formulas are derived for zero-coupon bonds and zero-coupon bond options and then extended to pricing formulas for interest rate caplets, caps and swaptions. This dissertation will explore both the origins and performance of these formulas. The former will involve an investigation into the original results presented by Trolle and Schwartz (2009) and the latter, a test for consistency with Monte Carlo pricing.

Features of the Trolle and Schwartz (2009) model:

There are prominent features of the Trolle and Schwartz (2009) interest rate model (hereafter referred to as the TS model) which incorporate stylised facts of interest rates, of which modelling interest rate volatility as a stochastic process is a fundamental component. Heston (1993) notably constructed the first stochastic volatility interest rate model by modelling volatility of bond prices as an additional state variable that follows a mean-reverting, square root process. Many models have since incorporated stochastic volatility, however there are notably few instances of this being applied for the HJM framework.

As a reformulation of the HJM framework, the TS model inherits the useful feature of fitting the initial yield curve, however this gives the framework a high-dimensional structure. In order to lower the dimensions, Cheyette (1992) introduces a Markovian form for the forward rate volatility that reduces the model dimensions. Chiarella and Kwon (2000) then integrate stochastic volatility into the forward rate volatilities to form a stochastic volatility HJM framework. The TS model however, specifies deterministic forward rate volatilities and incorporates stochastic volatility as additional state variables. Stochastic volatility models have since been built around the market models of Brace, Gatarek and Musiela (1997) and Jamshidian (1997). It is worthwhile noting that the HJM framework has favourable simulation features compared to these market models because it requires fewer Brownian motions whilst market models require a one for each forward rate on the curve.

Trolle and Schwartz (2009) compare their model to the market models of Han (2007) and Jarrow, Li and Zhao (2007), as both utilise interest rate derivative prices in their state vector estimation. Han (2007) uses swaption, Jarrow, Li and Zhao (2007) use cap price skews while Trolle and Schwartz (2009) use both swaption price data and cap price skew data. Both Han (2007) and Jarrow, Li and Zhao (2007) do not incorporate correlations between interest rates and interest rate volatility. Studies on short-rate dynamics by Andersen and Lund (1997) and Ball and Torous (1999) conclude that a non-zero correlation exists between interest rates and their volatility while Casassus, Collin-Dufresne and Goldstein (2005) maintain that permitting non-zero correlations is essential in fitting cap skew data. Han (2007) instead focuses on swaption pricing without skews, while Jarrow, Li and Zhao (2007) use a jump process to fit observed pricing skews in interest rate cap prices. Trolle and Schwartz (2009) suggest that extending these competing market models to non-zero correlations disturbs the volatility dynamics under the forward measure which results in intractable features. Jarrow, Li and Zhao (2007) in fact recognise a strong negative correlation in their own model between the predominant volatility

state variable and interest rates, despite specifying a zero correlation constraint.

These competing market models all imply that, given the state variables, forward LIBOR and swap rates are log-normally distributed under the appropriate forward measure. In the TS model, LIBOR and swap rates are approximately normally distributed under the appropriate forward measure, which implies a possibility of negative forward rates. This has in fact become a desirable feature of interest rate models, given the recent history of observed negative interest rates in certain developed markets, most notably in the EUR, CHF and JPY interest rate markets.

Another practical feature required of interest rate models is to facilitate shocks to the term structures. Trolle and Schwartz (2009) specify their forward rate volatilities using a Markovian formulation that allows for hump-shaped shocks to the term structure which ensures path-dependency of the forward curve and proves essential when fitting implied cap price skews in their sample. The choice for the forward rate volatility links the TS model to other variations of stochastic volatility interest rate models. The stochastic Hull and White (1990) term structure model developed by Casassus, Collin-Dufresne and Goldstein (2005) as well as a stochastic volatility Ho and Lee (1986) model can both be recovered from the TS model by changing the specification of the forward rate volatility.

The TS model also has the notable feature of including unspanned stochastic volatility factors. These are additional Brownian motions that drive interest rate volatility, and thus interest rate derivative prices, without necessarily affecting the term structure. Including these factors within a dynamic term structure model improves performance in pricing interest rate derivatives, whilst leaving the term structure unaffected. This was first identified by Collin-Dufresne and Goldstein (2002), who note that these factors arise naturally within the HJM framework.

There are several studies supporting existence of unspanned stochastic volatility factors. Collin-Dufresne and Goldstein (2002) maintain that not all interest rate derivatives can be hedged with bonds alone and attribute this to unspanned stochastic volatility factors. This result is supported by the difficulties Li and Zhao (2006) encountered in hedging volatility dependent cap straddles with bonds alone. Furthermore, Heidari and Wu (2003) found that a three-factor term structure model explained only 60% of changes to swaption volatilities but including additional unspanned factors improved this to 97%. Trolle and Schwartz (2009) confirm that multiple unspanned factors are required to accurately represent interest rate volatility. Andersen and Benzoni (2010) suggest that the existence of these factors prevents bond markets from spanning interest rate derivatives markets which implies that volatility risk can not be hedged fully by a position in the underlying bond market.

Requirements for Monte Carlo simulation:

In order to test the consistency of the semi-analytical pricing formulas, a scheme is required for Monte Carlo simulation. Although citing results linked to a Monte Carlo formulation, Trolle and Schwartz (2009) do not provide a simulation scheme as such. Thus, in the absence of a an exact simulation scheme, a model specific simulation setup must be constructed. To simulate the dynamics of the instantaneous forward rate and the volatility state variables a discrete simulation structure is developed using simulation techniques consistent with the HJM framework and Heston (1993) model.

To determine the discrete simulation scheme for the instantaneous forward rate in the multi-factor HJM framework, Fan, Gupta and Ritchken (2003) and Glasserman (2003) both express convenient formulations of the discretised drift term which ensure no-arbitrage. Glasserman (2003) provides steps to determine the drift term by ensuring that for the discrete simulation discounted bond prices remain martingales under $\mathbb{Q}$. Fan, Gupta and Ritchken (2003), on the other hand, apply a slightly different simplification to discretise the forward rate volatilities and incorporate the HJM no-arbitrage restriction. Unlike Glasserman (2003) however, they do not provide steps for the construction of their simulation.

Determining the discrete simulation of the volatility state variables involves reviewing methods used for simulating the Heston (1993) equity model. When considering a standard Euler scheme, Lord, Koekkoek and Dijk (2010) address the possibility of negative volatility by proposing a full truncation scheme which imposes a max function on the volatility at each time step. This can be extended to a Milstein scheme, although even in this setting the volatility is not guaranteed to remain non-negative. However, full truncation is effective in preventing the process from prolonged periods of negative values. Kahl and Jäckel (2006) suggest an implicit discrete Milstein scheme combined with a slightly more sophisticated scheme for the stock price. Andersen (2007), however, criticises this method for producing biased results and for its lack of robustness. Instead, Andersen (2007) advocates an alternative quadratic-exponential (QE) scheme which applies moment matching methods to simulate the volatility from the exact density and infers a correlation for the stock price through a carefully chosen discrete scheme; this produces notably smaller biases. Broadie and Kaya (2006) develop an exact simulation scheme by using the true volatility transition density of a non-central chi-squared, although Andersen (2007) states that the computational inefficiency of the scheme makes it impractical. A detailed analysis of the implementation of these techniques is presented by Rouah (2013). In constructing a Monte Carlo simulation for the TS model, we propose a full truncated Milstein approach in conjunction with the insights pro-

posed by Glasserman (2003) as well as a modified version of the QE scheme in the context of the instantaneous forward rate rather than a stock price process.

The simulation structure depends on the values of the state vector at the initial time of simulation. These state variables are however constructs of the TS model and are thus unobservable from market data. To illustrate the full implementation of the TS model, it is therefore necessary to employ "Kalman filtering" to estimate the progression of the latent state vector through time with the use of observable data-points. Duffee and Stanton (2012) suggest that in the estimation of term structure models, a quasi-maximum likelihood (QML) estimation in combination with the Kalman filter is preferable for combined estimation of latent state variables and model parameters. Trolle and Schwartz (2009) apply QML in accordance with an extended Kalman filter (EKF), which alters the Kalman filter as a result of the influence of the non-linear semi-analytical formulas.

For the purposes of this dissertation it is important to note that estimation of the model parameters will not be required. The only requirement is consistency between the parameters used for the semi-analytical pricing and those used for Monte Carlo simulation. Therefore, a full re-estimation of the parameter set through an application of QML estimation is not performed. The parameter set will instead be adopted from the results of the original estimation by Trolle and Schwartz (2009), so that only the extended Kalman filter is required for estimating the latent state variables. This allows the focus to remain on estimating the state variables rather than efficiency of the filtering process, which would be the case when estimating the parameter set.

Implementation of this model will involve simulating a path of the state vector and applying the model pricing formula to simulate a dataset of interest rates and interest rate derivatives and then filtering on this dataset. The accuracy of the extended Kalman filter can then be determined by comparing the estimated state vector to the simulated state vector. In order to verify the results from the simulated dataset, an analysis will then also be conducted using a historical dataset, similar to the dataset of Trolle and Schwartz (2009) in their analysis. If indeed the semi-analytical pricing formulas appear consistent with the Monte Carlo simulation, it will support the use of these formulas to price bonds and interest rate derivatives.

In Chapter 2, the model specifications and formulation of the semi-analytical pricing formulas, as outlined by Trolle and Schwartz (2009), are introduced. In Chapter 3, the implementation of the extended Kalman Filter to this model is reviewed and the outline from Trolle and Schwartz (2009) is explored in further detail. In Chapter 4, the results of the Monte Carlo pricing comparison are analysed. The dissertation concludes with Chapter 5 .

Chapter 2

Model specifications

These specifications closely follow those by Trolle and Schwartz (2009) in defining their model dynamics and deriving the semi-analytical pricing formulas.

2.1 Dynamics under risk-neutral measure $\mathbb{Q}$

Consider the time t instantaneous forward rate, $f(t,T)$, for instantaneous borrowing at time T, with $t \leq T$, and volatility state variables $v_i(t)$, where $i = 1, \ldots, N$, that inform the diffusion of $f(t,T)$. Then, the model dynamics specified by Trolle and Schwartz (2009) under $\mathbb{Q}$ are

$$df(t,T) = \mu_f(t,T)\,dt + \sum_{i=1}^{N} \sigma_{f,i}(t,T)\sqrt{v_i(t)}\,dW_i^{\mathbb{Q}}(t), \tag{2.1}$$

$$d\,v_i(t) = \kappa_i(\theta_i - v_i(t))\,dt + \sigma_i\sqrt{v_i(t)}\left(\rho_i\,dW_i^{\mathbb{Q}}(t) + \sqrt{1-\rho_i^2}\,dZ_i^{\mathbb{Q}}(t)\right), \tag{2.2}$$

where $W_i^{\mathbb{Q}}(t)$ and $Z_i^{\mathbb{Q}}(t)$ are independent standard Brownian motions under $\mathbb{Q}$. The $v_i(t)$ follow a mean-reverting, square-root process with speed of mean-reversion, κ_i, long-run mean, θ_i, and volatility σ_i.

In the case where $i = N$, the dynamics exhibit $N \times 2$ factors, or Brownian motions, in total. The $W_i^{\mathbb{Q}}(t)$ represent the term structure specific factors and the $Z_i^{\mathbb{Q}}(t)$ represent the unspanned stochastic volatility factors, which arise naturally in this multi-factor HJM framework as noted by Collin-Dufresne and Goldstein (2002). The unspanned nature of these factors will become fully apparent upon closer inspection of the factors that drive the zero-coupon bond curve.

The value of the correlations, ρ_i, between the forward rate and the volatility state variables can reduce the number of unspanned factors by setting $\rho_i = \pm 1$. Upon validating the model structure, Trolle and Schwartz (2009) find that having fewer unspanned factors than term structure factors reduces the performance of the model significantly. Setting $\rho_i = 0$ for all i eliminates all correlations, which

is a common assumption amongst market models, and radically reduces the ability of the model to fit cap pricing skews. This correlation structure resembles the dynamics of the Heston (1993) model in a multi-factor context.

An important feature highlighted in the seminal work of Heath, Jarrow and Morton (1992) is the no-arbitrage condition on the drift of $df(t,T)$. Applying the no-arbitrage requirement in this setting fully determines the drift term $\mu_f(t,T)$ as

$$\mu_f(t,T) = \sum_{i=1}^{N} v_i(t)\sigma_{f,i}(t,T) \int_t^T \sigma_{f,i}(t,u)\, du. \tag{2.3}$$

In order to generalise the dynamics further and maintain a Markovian structure, Trolle and Schwartz (2009) define the forward rate volatilities, $\sigma_{f,i}(t,T)$, as

$$\sigma_{f,i}(t,T) = (\alpha_{0,i} + \alpha_{1,i}(T-t))e^{-\gamma_i(T-t)}. \tag{2.4}$$

This ensures that the process is Markov and has finite dimensions while maintaining time-homogeneity of the forward rate volatilities. Trolle and Schwartz (2009) describe this volatility structure as essential for accurate pricing of long-term derivatives. It also allows for a hump-shaped forward rate volatility structure for $\sigma_{f,i}(t,T)$, much like the formulation in Cheyette (1992). In fact, the TS model can alternatively be described as a stochastic volatility Cheyette (1992) model.

Simpler stochastic interest rate models are embedded within the model dynamics. Setting $N = 1$ and $\alpha_{1,1} = 0$ produces the stochastic Hull and White (1990) model as developed by Casassus, Collin-Dufresne and Goldstein (2005) and additionally setting $\gamma_i = 0$ recovers the stochastic Ho and Lee (1986) model.

2.2 Monte Carlo simulation

A discrete Monte Carlo simulation for the TS model is neither produced explicitly by Trolle and Schwartz (2009) nor by any subsequent literature. Simulation schemes are developed here that combine techniques used in the HJM framework and Heston (1993) model. Two alternative methods are proposed.

2.2.1 Method 1: Full truncated Milstein scheme

In the first approach, the volatility state variables are simulated using a Milstein scheme coupled with the full truncation method of Lord, Koekkoek and Dijk (2010). This seeks to avoid negative volatility values in the discretisation with the use of a max function. The instantaneous forward rate is simulated using a discrete time Euler scheme, with the drift term determined using the analysis of Glasserman (2003).

Let the discretised forward rate, $\hat{f}(t_k, t_j)$, and volatility state variables, $\hat{v}_i(t_k)$, be defined for time t_k as

$$\begin{aligned}\hat{f}(t_k, t_j) &= \hat{f}(t_{k-1}, t_j) + \hat{\mu}_f(t_{k-1}, t_j)\Delta t_k \\ &\quad + \sum_{i=1}^{N} \hat{\sigma}_{f,i}(t_{k-1}, t_j)\sqrt{\hat{v}_i(t_{k-1})^+}\sqrt{\Delta t_k}X_{i,k}\,, \qquad (2.5)\\ \hat{v}_i(t_k) &= \hat{v}_i(t_{k-1}) + \kappa_i(\theta_i - \hat{v}_i(t_{k-1})^+)\Delta t_k \\ &\quad + \sigma_i\sqrt{\hat{v}_i(t_{k-1})^+}\left(\sqrt{\Delta t_k}Y_{i,k}\right) + \frac{1}{4}\sigma_i^2\left((\sqrt{\Delta t_k}Y_{i,k})^2 - 1\right)\,, \qquad (2.6)\end{aligned}$$

where $\Delta t_k = (t_k - t_{k-1})$ and $X_{i,k}, Y_{i,k} \sim \mathcal{N}(0,1)$ with correlation ρ_i between each $X_{i,k}$ and $Y_{i,k}$. This correlation is introduced using a Cholesky decomposition for the normal random variables. The drift term, $\hat{\mu}_f(t_{k-1}, t_j)$, is specified as

$$\hat{\mu}_f(t_{k-1}, t_j) = \frac{\sum_{i=1}^{N}\hat{v}_i(t_{k-1})(S_{i,j}^2 - S_{i,j-1}^2)}{2\Delta t_k}, \qquad (2.7)$$

where

$$S_{i,n} = \sum_{l=k}^{n}\hat{\sigma}_{f,i}(t_{k-1}, t_l)\Delta t_k. \qquad (2.8)$$

The derivation of this discretisation is outlined in Appendix A.

2.2.2 Method 2: Adapted QE scheme

The second approach follows Andersen (2007) and the quadratic exponential (QE) method for simulating $\hat{v}_i(t_k)$. This scheme is developed exclusively for square-root processes, such as in (2.2), and considers the exact non-central chi-squared density of $v_i(t_k)$ using properties of similar densities. Full description of the scheme can be found in Andersen (2007). For implementation in the context of the TS model a brief outline is presented for the expressions required for simulation. The QE method follows the results of Broadie and Kaya (2006), which specify the exact transition density of $v_i(t_k)$ as a scaled non-central chi-squared random variable, conditional on $v_i(t_{k-1})$.

Knowing the exact form of the exact distribution, Andersen (2007) notes that $v_i(t_{k-1})$ is directly proportional to the non-centrality parameter of the non-central chi-squared variable, $v_i(t_k)$. Therefore, for the discretised process $\hat{v}_i(t_k)$, when $\hat{v}_i(t_{k-1})$ is large, then $\hat{v}_i(t_k)$ and can be simulated as

$$\hat{v}_i(t_k) = a_i(b_i + X_{i,k})^2, \qquad (2.9)$$

where $X_{i,k}$ is a standard normal random variable and a_i and b_i are constants which depend on $\hat{v}_i(t_{k-1})$. For low values of $\hat{v}_i(t_{k-1})$, the value for $\hat{v}_i(t_k)$ is given by

$$\hat{v}_i(t_k) = L^{-1}(U_{i,k}), \tag{2.10}$$

where $U_{i,k}$ is a uniform random variable and

$$L^{-1}(u) = \begin{cases} 0 & \text{if} \quad 0 \leq u \leq p_i \\ \beta_i^{-1}\log\left(\frac{1-p_i}{1-u}\right) & \text{if} \quad p_i < u \leq 1, \end{cases} \tag{2.11}$$

where the constants p_i and β_i are dependent on $\hat{v}_i(t_{k-1})$. Andersen (2007) employs moment-matching techniques to define a_i, b_i, p_i and β_i as

$$a_i = \frac{m_i}{1+b_i^2}, \tag{2.12}$$

$$b_i^2 = 2\psi_i^{-1} - 1 + 2\sqrt{2\psi_i^{-1}} \cdot \sqrt{2\psi_i^{-1} - 1}, \tag{2.13}$$

$$p_i = \frac{\psi_i - 1}{\psi_i + 1}, \tag{2.14}$$

$$\beta_i = \frac{2}{m_i(1+\psi_i)}, \tag{2.15}$$

with additional constants m_i, s_i and ψ_i defined as

$$m_i = \theta_i + (\hat{v}_i(t_{k-1}) - \theta_i)e^{-\kappa_i \Delta t_k}, \tag{2.16}$$

$$s_i^2 = \frac{\hat{v}_i(t_{k-1})\sigma_i^2 e^{-\kappa_i \Delta t_k}}{\kappa_i}(1 - e^{-\kappa_i \Delta t_k}) + \frac{\theta_i \sigma_i^2}{2\kappa_i}(1 - e^{-\kappa_i \Delta t_k})^2, \tag{2.17}$$

$$\psi_i = \frac{s^2}{m^2}. \tag{2.18}$$

For a choice of $\psi_c \in [1, 2]$, if $\psi_i \leq \psi_c$ then (2.9) applies and if $\psi_i > \psi_c$, then (2.10) applies. According to Andersen (2007) the choice of ψ_c is arbitrary, hence we set $\psi_c = 1.5$. To address the absence of the correlations, ρ_i, an Euler scheme is adopted for $\hat{f}(t_k, t_j)$, similar to (2.5) but without full truncation

$$\begin{aligned} \hat{f}(t_k, t_j) =& \hat{f}(t_{k-1}, t_j) + \hat{\mu}_f(t_{k-1}, t_j)\Delta t_k \\ &+ \sum_{i=1}^{N} \hat{\sigma}_{f,i}(t_{k-1}, t_j)\sqrt{\hat{v}_i(t_{k-1})}\sqrt{\Delta t_k} Y_{i,k}, \end{aligned} \tag{2.19}$$

where $\Delta t_k = (t_k - t_{k-1})$ and $Y_{i,k}$ is a standard normal random variable and has correlation ρ_i with $X_{i,k}$ in (2.9). Again, the analysis of Glasserman (2003) is used to determine the discretised drift term.

Andersen (2007) noted that this approach does experience "leaking correlations" because (2.9) and (2.10) are non-linear. However, Andersen (2007) states that in practice the implied correlations are consistently close to ρ_i. The adapted discretisation and correlation relationship is discussed further in Appendix A with reference to Andersen (2007).

2.3 Semi-analytical pricing formulas

2.3.1 Pricing formula for zero-coupon bonds

Trolle and Schwartz (2009) show that the model specifications for $f(t,T)$ imply that the time t semi-analytical zero-coupon bond price, for maturity T, is

$$\begin{aligned} P(t,T) =& \exp\left[-\int_t^T f(t,u)du\right] \\ =& \frac{P(0,T)}{P(0,t)} \exp\left[\sum_{i=1}^{N} B_{x_i}(T-t)x_i(t) + \sum_{i=1}^{N}\sum_{j=1}^{6} B_{\phi_{j,i}}(T-t)\phi_{j,i}(t)\right], \end{aligned} \tag{2.20}$$

where

$$B_{x_i}(\tau) = \frac{\alpha_{1,i}}{\gamma_i}\left(\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right)(e^{-\gamma_i\tau} - 1) + \tau e^{-\gamma_i\tau}\right), \tag{2.21}$$

$$B_{\phi_{1,i}}(\tau) = \frac{\alpha_{1,i}}{\gamma_i}(e^{-\gamma_i\tau} - 1), \tag{2.22}$$

$$B_{\phi_{2,i}}(\tau) = \left(\frac{\alpha_{1,i}}{\gamma_i}\right)^2\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right)\left(\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right)(e^{-\gamma_i\tau} - 1) + \tau e^{-\gamma_i\tau}\right), \tag{2.23}$$

$$\begin{aligned} B_{\phi_{3,i}}(\tau) = -\frac{\alpha_{1,i}}{\gamma_i^2}\Bigg(&\left(\frac{\alpha_{1,i}}{2\gamma_i^2} + \frac{\alpha_{0,i}}{\gamma_i} + \frac{\alpha_{0,i}^2}{2\alpha_{1,i}}\right)(e^{-2\gamma_i\tau} - 1) \\ &+ \left(\frac{\alpha_{1,i}}{\gamma_i} + \alpha_{0,i}\right)\tau e^{-2\gamma_i\tau} + \frac{\alpha_{1,i}}{2}\tau^2 e^{-2\gamma_i\tau}\Bigg), \end{aligned} \tag{2.24}$$

$$B_{\phi_{4,i}}(\tau) = \left(\frac{\alpha_{1,i}}{\gamma_i}\right)^2\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right)(e^{-\gamma_i\tau} - 1), \tag{2.25}$$

$$B_{\phi_{5,i}}(\tau) = -\frac{\alpha_{1,i}}{\gamma_i^2}\left(\left(\frac{\alpha_{1,i}}{\gamma_i} + \alpha_{0,i}\right)(e^{-2\gamma_i\tau} - 1) + \alpha_{1,i}\tau e^{-2\gamma_i\tau}\right), \tag{2.26}$$

$$B_{\phi_{6,i}}(\tau) = -\frac{1}{2}\left(\frac{\alpha_{1,i}}{\gamma_i}\right)^2(e^{-2\gamma_i\tau} - 1), \tag{2.27}$$

and the dynamics of the additional state variables are defined as

$$dx_i(t) = -\gamma_i x_i(t)\,dt + \sqrt{v_i(t)}\,dW_i^{\mathbb{Q}}(t), \tag{2.28}$$

$$d\phi_{1,i}(t) = (x_i(t) - \gamma_i\phi_{1,i}(t))\,dt, \tag{2.29}$$

$$d\phi_{2,i}(t) = (v_i(t) - \gamma_i\phi_{2,i}(t))\,dt, \tag{2.30}$$

$$d\phi_{3,i}(t) = (x_i(t) - 2\gamma_i\phi_{3,i}(t))\,dt, \tag{2.31}$$

$$d\phi_{4,i}(t) = (\phi_{2,i}(t) - \gamma_i\phi_{4,i}(t))\,dt, \tag{2.32}$$

$$d\phi_{5,i}(t) = (\phi_{3,i}(t) - 2\gamma_i\phi_{5,i}(t))\,dt, \tag{2.33}$$

$$d\phi_{6,i}(t) = (2\phi_{5,i}(t) - 2\gamma_i\phi_{6,i}(t))\,dt, \tag{2.34}$$

with initial conditions $x_i(0) = \phi_{1,i}(0) = \ldots = \phi_{6,i}(0) = 0$, to fit the initial term structure. Further expansion of this exact result is found in Appendix B.

The $x_i(t)$ state variables are driven by the $W_i^{\mathbb{Q}}(t)$ Brownian motions and are the only stochastic variables that affect the zero-coupon bond curve. As a result, $W_i^{\mathbb{Q}}(t)$ refers to term structure specific factors. The $Z_i^{\mathbb{Q}}(t)$ do not affect the term structure directly, although they continue to drive $v_i(t)$. As a result, these factors, or Brownian motions, are considered unspanned stochastic volatility factors.

This model is initially time-inhomogeneous as it fits the initial term structure by construction. In their estimation procedure, Trolle and Schwartz (2009) estimate a time-homogeneous model by removing the dependence on the initial term structure and replacing $\frac{P(0,T)}{P(0,t)}$ in (2.20) with $\exp(-\varphi(T-t))$, where φ represents a forward rate for infinite maturity.

2.3.2 Pricing formula for zero-coupon bond options

In order to price options on zero-coupon bonds the methods of Duffie, Pan and Singleton (2000) and Collin-Dufresne and Goldstein (2003) are employed. Application in the case of the TS model requires the introduction of the transform

$$\Psi(u,t,T_0,T_1) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_0} r_s\,ds\right)\exp\Big(u\log\big(P(T_0,T_1)\big)\Big)\right]. \tag{2.35}$$

This expression is familiar in the context of risk-neutral pricing for zero-coupon bond options and, according to Collin-Dufresne and Goldstein (2003), has an exponentially affine solution. The solution is determined as

$$\begin{aligned}\Psi(u,t,T_0,T_1) = \exp\Bigg[&M(T_0-t) + \sum_{i=1}^{N} N_i(T_0-t)v_i(t) + u\log\big(P(t,T_1)\big)\\ &+ (1-u)\log\big(P(t,T_0)\big)\Bigg],\end{aligned} \tag{2.36}$$

where $M(\tau)$ and $N(\tau)$ are the solutions to the ordinary differential equations

$$\frac{\partial M(\tau)}{\partial \tau} = \sum_{i=1}^{N} N_i(\tau)\kappa_i\theta_i, \tag{2.37}$$

$$\begin{aligned}\frac{\partial N_i(\tau)}{\partial \tau} = {}& N_i(\tau)\big(-\kappa_i + \sigma_i\rho_i\big(uB_{x_i}(T_1-T_0+\tau) + (1-u)B_{x_i}(\tau)\big)\big)\\ &+ \frac{1}{2}N_i(\tau)^2\sigma_i^2 + \frac{1}{2}(u^2-u)B_{x_i}(T_1-T_0+\tau)^2\\ &+ \frac{1}{2}((1-u)^2-(1-u))B_{x_i}(\tau)^2\\ &+ u(1-u)B_{x_i}(T_1-T_0+\tau)B_{x_i}(\tau),\end{aligned} \tag{2.38}$$

with boundary conditions $M(0)=0$ and $N_i(0)=0$. Further details of this derivation are contained in Appendix B, which provides a breakdown of this exact result.

It is important to note here that when calculating derivative prices as part of the Monte Carlo simulation, $P(t,T_0)$ and $P(t,T_1)$ in (2.36) will be calculated using the bond pricing formula in (2.20). When calculating derivative prices for the estimation procedure, these will be computed using the observed zero-coupon bond curve. This is explored in more detail in Chapter 3.

Trolle and Schwartz (2009) then price a put option on a zero-coupon bond using the result that the price of a time t put option on a zero-coupon bond struck at K with option expiry at T_0 and maturity of the bond at T_1, with $T_0 < T1$, can be expressed as

$$\mathcal{P}(t,T_0,T_1,K) = KG_{0,1}\big(t,T_0,t_1,\log(K)\big) - G_{1,1}\big(t,T_0,T_1,\log(K)\big), \tag{2.39}$$

where $G_{a,b}(t,T_0,T_1,y)$ is defined as

$$G_{a,b}(t,T_0,T_1,y) = \frac{\Psi(a,t,T_0,T_1)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathcal{I}m[\Psi(a+\mathrm{i}ub,t,T_0,T_1)e^{-\mathrm{i}uy}]}{u}\,du. \tag{2.40}$$

The above result for the price of a zero-coupon bond option is derived in Trolle and Schwartz (2009) using the inverse Fourier transform approach of Duffie, Pan and Singleton (2000) and Collin-Dufresne and Goldstein (2003). An alternative approach is considered in Appendix B using a change of measure approach. These semi-analytical pricing formulas, (2.20) and (2.39), provide solutions for pricing LIBOR and swap rates as well as caplets/caps and swaptions.

2.3.3 Numerical methods in pricing formulas

The pricing formula for zero-coupon bonds requires straight forward calculations for any maturity given the parameters and the latent term structure state variables $x_i(t)$, $\phi_{1,i}(t),\ldots,\phi_{6,i}(t)$ for $i = 1,\ldots,N$. This formula is semi-analytical only because it is conditional on the estimated state variables.

Evaluating the formula for zero-coupon bond options requires the use of numerical methods. This formula is described as semi-analytical because of its dependency on the estimated state vector as well as the numerical methods required. There are two instances requiring numerical techniques: Solving the ordinary differential equations (ODEs) in (2.37) and (2.38) and evaluating the integral in (2.40). The solutions to the ODEs $M(\tau)$ and $N_i(\tau)$ are required to evaluate (2.36) and are solved using the standard 4th-order Runge-Kutta algorithm built into Matlab's *ode45* functionality for solving differential equations. The algorithm is simple and efficient and solves first for $N_i(\tau)$ across a range for u before applying a trapezoidal rule to solve for $M(\tau)$.

Gauss-Legendre quadrature is used to calculate the integral in (2.40). This method of quadrature calculates a weighted sum of n values, each evaluated using the integrated expression at the i-th root, for $i = 1, \ldots, n$, of an nth-degree Legendre polynomial. Although the integral under consideration has integration bounds from zero to infinity, Trolle and Schwartz (2009) note that the expression vanishes rapidly to zero well before u approaches 2000. Trolle and Schwartz (2009) use 20 points in the interval [0,1000] interval and 20 further points in the interval [1000, 8000]. See Appendix B for an illustration of the convergence of the integral.

The expression in (2.40) is required for any derivatives pricing, however the range of integration and chosen quadrature points for u can remain static throughout. This means the ODEs for $M(\tau)$ and $N_i(\tau)$ can be calculated in a single Runge-Kutta algorithm for multiple values of u for any derivative, as long as the quadrature bounds are sufficiently large. This use of predetermined quadrature points for Gauss-Legendre quadrature makes having a large number of integration points less computationally intensive. The larger number of integration points is required for pricing accuracy and although it is possible to reduce the number of points, as in Trolle and Schwartz (2009), the focus here is on the accuracy of the pricing formulas so a larger number of points is maintained.

2.4 Interest rate pricing formulas

The ability to price both zero-coupon bonds and zero-coupon bond options allows for the pricing of underlying interest rates and vanilla interest rate derivatives using the pricing formulas below.

LIBOR and swap rates

The time t value of a simple LIBOR rate for terminal time T_1, with $t < T_1$, is

$$L(t, T_1) = \frac{1 - P(t, T_1)}{(T_1 - t)P(t, T_1)}. \tag{2.41}$$

The time t value of a swap rate for the period t to T_n, with equal fixed payment legs υ, is defined as

$$S(t, T_n) = \frac{1 - P(t, T_n)}{\upsilon \sum_{j=1}^{n} P(t, T_j)}. \tag{2.42}$$

The time t value of a forward swap rate for the period starting at T_m to T_n, with $T_m < T_n$, is similarly

$$S(t, T_m, T_n) = \frac{P(t, T_m) - P(t, T_n)}{\upsilon \sum_{j=m+1}^{n} P(t, T_j)}, \tag{2.43}$$

where setting $T_m = t$ corresponds to the time t swap rate.

Cap price

The time t cap price is the sum of caplet prices for reset dates $T_1, ..., T_{n-1}$ and corresponding payment dates $T_2, ..., T_n$, where the first caplet from t to T_1 is not included. It can be shown that by manipulating the terminal caplet payoff

$$\text{Cpl}(T_j, T_j, K) = \upsilon \left(L(T_j - \upsilon, T_j) - K\right)^+, \tag{2.44}$$

for payment date T_j, payment leg υ and strike K, each caplet can be calculated exactly as a scaled put option on a zero-coupon bond. The time t caplet price can then be expressed as

$$\text{Cpl}(t, T_j, K) = (1 + \upsilon K)\mathcal{P}\left(t, T_j - \upsilon, T_j, \frac{1}{1 + \upsilon K}\right), \tag{2.45}$$

so that the time t price of an interest rate cap maturing at time T_n is

$$\text{Cap}(t, T_n, K) = \sum_{j=2}^{n} \text{Cpl}(t, T_j, K). \tag{2.46}$$

This pricing technique requires only the pricing formula for zero-coupon bond options in (2.39). For a cap to be priced at-the-money-forward (ATMF) the strike price is set at $\bar{K} = S(t, T_1, T_n)$, which is the fair forward swap rate for initial payment T_1 with terminal payment at T_n.

Swaption price

A payer swaption can be shown to be priced as a scaled put option on a coupon-bearing bond by manipulating the terminal payoff

$$\text{Swpn}(T_m, T_m, T_n, K) = \upsilon \sum_{i=m+1}^{n} P(T_m, T_i) \left(S(T_m, T_n) - K\right)^+, \tag{2.47}$$

with option maturity T_m, swap dates $T_{m+1}, \ldots, T_n$, swap maturity of T_n and strike K. The time t price of the corresponding coupon-bearing bond with cashflows $Y(T_j)$ for $j = m + 1, \ldots, n$ is

$$P^c(t) = \sum_{j=m+1}^{n} Y(T_j) P(t, T_j), \tag{2.48}$$

with $Y(T_j) = K\upsilon$ for $j = m + 1, ..., n - 1$ and $Y(T_n) = 1 + K\upsilon$.

As there is no pricing formula for a coupon-bearing bond in Trolle and Schwartz (2009), the swaption price is approximated using the stochastic duration method developed by Wei (1997) and Munk (1999). They approximate an option on a

coupon-bearing bond with an option on a zero-coupon bond that has the same stochastic duration as the coupon bond. They define the time t stochastic duration, $D(t)$, of a coupon bond as the maturity of a zero-coupon bond with equivalent relative volatility. Trolle and Schwartz (2009) deduce that considering the volatility of the bond price dynamics under the risk-neutral measure, $D(t)$ is the solution to

$$\sum_{i=1}^{N} v_i(t) B_{x_i}\big(D(t)\big)^2 = \sum_{i=1}^{N} \left(\sum_{j=m+1}^{n} w_j B_{x_i}(T_j - t) \right)^2, \tag{2.49}$$

where $w_j = \frac{Y(T_j)P(t,T_j)}{P^c(t)}$. As a result of the analysis of Munk (1999), $D(t)$ is known to have a unique solution if $B_{x_i}(\tau)$ is strictly decreasing, which Trolle and Schwartz (2009) confirm in the case of their parameter estimates.

Therefore, the time t price of a payer swaption is a scaled option on a zero-coupon bond of the form

$$\text{Swpn}(t, T_m, T_n, K) = \xi \mathcal{P}(t, T_m, t + D(t), \xi^{-1}), \tag{2.50}$$

with the scaling factor $\xi = \frac{P^c(t)}{P(t,t+D(t))}$.

2.5 Dynamics under real-world measure, $\mathbb{P}$

Trolle and Schwartz (2009) also determine the dynamics of the state variables under the real-world measure $\mathbb{P}$ following a change of measure. These dynamics form an important part of the filtering and estimation process, which is discussed in Chapter 3. They apply an "extended affine market price of risk" condition recommended by Cheridito, Filipović and Kimmel (2007). This is a significantly robust specification under which the state vector retains an affine structure under the change of measure. Under $\mathbb{P}$, the state variables dynamics can be specified as

$$dx_i(t) = \left(\eta_i^{\mathbb{P}} + \kappa_{x,i}^{\mathbb{P}} x_i(t) + \kappa_{xv,i}^{\mathbb{P}} v_i(t) \right) dt + \sqrt{v_i(t)}\, dW_i^{\mathbb{P}}(t), \tag{2.51}$$

$$dv_i(t) = \kappa_i^{\mathbb{P}} \left(\theta_i^{\mathbb{P}} - v_i(t) \right) dt + \sigma_i \sqrt{v_i(t)} \left(\rho_i dW_i^P(t) + \sqrt{1-\rho_i} dZ_i^{\mathbb{P}}(t) \right), \tag{2.52}$$

while the dynamics of the state variables $\phi_{1,i}(t), ..., \phi_{6,N}(t)$ are unaffected as they do not contain any stochastic elements. Trolle and Schwartz (2009) note that in order for this market price of risk condition to hold, $v_i(t)$ should not reach zero under either $\mathbb{Q}$ or $\mathbb{P}$, which gives rise to the following Feller conditions

$$2\kappa_i \theta_i \geq \sigma_i^2, \tag{2.53}$$

$$2\kappa_i^{\mathbb{P}} \theta_i^{\mathbb{P}} \geq \sigma_i^2. \tag{2.54}$$

This ensures that the market prices of risk remain finite, although in practice this is often not a necessity and is instead a beneficial feature for numerical tractability.

2.6 Approximate pricing using implied volatilities

This section introduces a further semi-analytical pricing method for the TS model that approximates implied normal swaption and caplet volatilities. It is based on the result that the forward swap rate is approximately normally distributed under the forward swap measure. This is in contrast to popular market models, where forward LIBOR rates, as in Brace, Gatarek and Musiela (1997), or forward swap rates, as in Jamshidian (1997), are log-normally distributed.

To determine the distribution of the forward swap rate, Trolle and Schwartz (2009) analyse the dynamics under the forward swap measure where forward swap rates are martingales. Considering (2.43) and applying Ito's Lemma before switching to the forward swap measure produces the forward swap rate dynamics

$$dS(u,T_m,T_n) = \sum_{i=1}^{N}\left(\sum_{j=m}^{n}\zeta_j(u)B_{x_i}(T_j-u)\right)\sqrt{v_i(u)}\,dW_i^{\mathbb{Q}^{T_m,T_n}}(u), \tag{2.55}$$

where the time-u dynamics of $v_i(u)$ are

$$\begin{aligned} dv_i(u) = &\left(\kappa_i(\theta_i - v_i(u)) + v_i(u)\sigma_i\rho_i\upsilon\sum_{j=m+1}^{n}\epsilon_j(u)B_{x_i}(T_j-u)\right)du \\ &+ \sigma_i\sqrt{v_i(u)}\left(\rho_i\,dW_i^{\mathbb{Q}^{T_m,T_n}}(u) + \sqrt{1-\rho_i^2}\,dZ_i^{\mathbb{Q}^{T_m,T_n}}(u)\right), \end{aligned} \tag{2.56}$$

where the terms required in these dynamics are

$$\epsilon_j(u) = \frac{P(u,T_j)}{PVBP(u)}, \tag{2.57}$$

$$\zeta_m(u) = \frac{P(u,T_m)}{PVBP(u)}, \tag{2.58}$$

$$\zeta_j(u) = -\upsilon S(u,T_m,T_n)\frac{P(u,T_j)}{PVBP(u)} \quad \text{for } j = m+1,...,n-1, \tag{2.59}$$

$$\zeta_n(u) = -(1+\upsilon S(u,T_m,Tn)\frac{P(u,T_j)}{PVBP(u)}, \tag{2.60}$$

and

$$PVBP(u) = \upsilon\sum_{j=m+1}^{n}P(u,T_j). \tag{2.61}$$

Initially, forward swap rates are not normally distributed due to stochastic elements $\zeta_j(u)$ in the diffusion term. The stochastic volatility also becomes non-affine, due to stochastic elements $\epsilon_j(u)$ in its drift term. To address this, these stochastic terms which are martingales under the forward swap measure are approximated by taking their time t expectations, which is just their time t values.

These approximations then imply that forward swap rates are approximately normally distributed, conditional on $v_i(t)$ as well as $v_i(t)$ retaining its affine diffusion. A second approximation is made by taking the time t expected value of $v_i(u)$ so that the volatility of the forward swap rate is now no longer conditionally normally distributed. These two approximations lead to

$$S(T_m, T_n) \sim \mathcal{N}(S(t, T_m, T_n), \sigma_N^2(t, T_m, T_n)(T_m - t)). \tag{2.62}$$

The normal implied volatility term is then defined as

$$\begin{aligned}&\sigma_N(t, T_m, T_n)\\ &= \left(\frac{1}{T_m - t} \int_t^{T_m} \sum_{i=1}^{N} \left(\sum_{j=m}^{n} \zeta_j(t) B_{x_i}(T_j - u) \right)^2 E_t^{\mathbb{Q}^{T_m, T_n}}[v_i(u)]\, du \right)^{\frac{1}{2}}.\end{aligned} \tag{2.63}$$

In the implementation of this approximate pricing formula, the calculation of the expected value of $v_i(u)$ in equation(2.63) is evaluated in the form of a Monte Carlo simulation. The dynamics of $v_i(u)$ under the forward swap measure evolve according to equation (2.56) and an Euler scheme is implemented for these purposes. Gauss-Legendre quadrature is used again to evaluate the integral in (2.63).

Trolle and Schwartz (2009) state that this pricing method makes for efficiency in the early stages of filtering and parameter estimation. This is done by using a third and final approximation to transform the normal implied volatility term to a log-normal implied volatility according to

$$\sigma_{LN}(t, T_m, T_n) = \frac{\sigma_N(t, T_m, T_n)}{S(t, T_m, T_n)}. \tag{2.64}$$

There are alternate approximations to this relationship that are not discussed here which may yield different approximation results. Trolle and Schwartz (2009) state that according to unreported Monte Carlo analysis, this pricing approximation produces "reasonably accurate" results for ATMF swaption pricing. As these results are unreported, it will be a worthwhile exercise to include this pricing formula in our Monte Carlo analysis.

As caplets are special cases of swaptions, where the underlying swap is just a single resetting period, this approximate pricing can be implemented for caps. By using this approximation formula for all caplets associated with a particular cap tenor, this produces an alternative pricing method for interest rate caps. In order to distinguish this alternative semi-analytical pricing method from the previous pricing formulas, these formulas will be referred to as the volatility approximation pricing formulas, while the previous pricing formulas will continue to be referred to as the semi-analytical pricing formulas.

Chapter 3

Model estimation

The semi-analytical pricing formulas as well as the volatility approximation formula can be fully evaluated given both the the latent state variables and the model parameters under $\mathbb{Q}$. To estimate the unobserved state variables and the model parameters, Trolle and Schwartz (2009) apply a Kalman filter in conjunction with maximum likelihood estimation. The filter predicts a path for the latent state vector given a set of data points and the maximum likelihood estimation finds the most likely parameter values by maximising a likelihood function. For the purposes of this analysis, if the parameter set is consistent and reasonable for both semi-analytical pricing and Monte Carlo simulation then any inconsistencies observed should not be as a result of the parameters. Therefore parameter estimation will not be required, only estimation of the latent state variables.

3.1 Data required for estimation

The initial set of data analysed will come from a simulated path of the state vector which then implies a set of interest rates and interest rate derivative prices. This will assess the validity of the implemented Kalman filter by comparing the estimated state vector with the simulated paths. A secondary analysis will involve using a subset of the historical data used by Trolle and Schwartz (2009) which contains US market rates and prices from Bloomberg.

The simulated dataset will be constructed using a discrete time Euler simulation of the state vector under the real-world measure $\mathbb{P}$. This simulated state vector can then imply LIBOR and swap rates as well as at-the-money-forward (ATMF) swaption and cap prices using the semi-analytical pricing formulas and the given parameter set. The simulated observations are generated for 101 weekly intervals with 50 time steps between each of the weekly observation dates. The size of the dataset remains computational practical while ensuring sufficiently accurate estimation of the state vector.

	$N=1$	$N=3$		
	$i=1$	$i=1$	$i=2$	$i=3$
κ_i	0.0553	0.5509	1.0187	0.1330
σ_i	0.3325	1.0497	1.4274	0.5157
$\alpha_{0,i}$	0.0045	0.0000	0.0020	-0.0097
$\alpha_{1,i}$	0.0131	0.0046	0.0265	0.0323
γ_i	0.3341	0.1777	1.1623	0.8282
ρ_i	0.4615	0.3270	0.2268	0.1777
$\kappa^P_{x,i}$	0.9767	0.7677	0.5650	0.8739
$\kappa^P_{xv,i}$	3.4479	0.0988	1.7115	1.6425
η^P_i	1.1964	-1.1288	0.8528	1.0453
κ^P_i	2.1476	2.3698	3.1794	1.7372
θ^P_i	0.7542	2.1070	0.7875	0.6330
ψ	0.0832		0.0680	
σ_{rates}	0.0054		0.0004	
σ_{deriv}	0.0003		0.0126	

Tab. 3.1: Model parameters as per original TS Model implementation.

The parameter set shown in Table 3.1 follows from the original parameter estimation by Trolle and Schwartz (2009), for $N = 1$ and $N = 3$ estimated using interest rate and swap rate data as well as both swaption prices and cap pricing skews. This parameters set is sufficient to determine the dynamics of the state vector through time under $\mathbb{P}$, as defined by equations (2.51) and (2.52).

The only adjustment from the original parameter set is a smaller σ_{deriv}, the standard deviation of the measurement errors for the derivative prices. This improves performance of the filter to accurately estimate the volatility state variables using fewer interest rate derivative prices. This improves computational efficiency as fewer derivative prices are required for estimation while maintaining accuracy in recovering the simulated state variables.

The interest rates considered include LIBOR rates calculated for 3, 6 and 9 month tenors. The swap rates calculated are the 1-15 yearly rates with six months fixed payment legs, requiring $\upsilon = 0.5$ in (2.42). The ATMF cap prices considered are yearly caps ranging from 1-15 year tenors. The ATMF swaption prices considered are expressed in terms of an $A \times B$ swaption, where the option maturity is A and the tenor of the underlying swap is B. The US market convention of three month payment legs between reset and payment dates is applied for all interest rate derivatives, so $\upsilon = 0.25$ in equations (2.45) and (2.47).

3.2 Estimation using extended Kalman filter (EKF)

As mentioned previously, the real-world dynamics of the state vector follow an affine diffusion process set out in (2.51) and (2.52) as well as (2.29) - (2.34). The standard Kalman filter is governed by two equations which represent the progression of the state vector through time and the relation of the state vector to the observed estimation inputs.

The transition equation identifies the discrete evolution of the state variables through time as

$$X_t = \Phi(X_{t-1}) + w_t, \quad w_t \quad \text{iid.}, \quad \mathbb{E}[w_t] = 0, \quad \text{Var}[w_t] = Q(X_{t-1}). \tag{3.1}$$

The function $\Phi(X_{t-1})$ is determined by the real-world continuous time dynamics of the state variables under $\mathbb{P}$. The variance of the disturbance term is described by the function $Q(X_{t-1})$. For a chosen value of N, the state vector X_t is constructed as

$$X_t = (x_1(1), ..., x_N(t), \phi_{1,1}(t), ..., \phi_{6,N}(t), v_1(t), ..., v_N(t))', \tag{3.2}$$

As X_t, which is an $(N \times 8)$-element column vector, follows an affine diffusion process the specifications for $\Phi(X_t)$ and $Q(X_t)$ can be determined in closed form. Following the results described by Fisher and Gilles (1996) for general estimation of exponential-affine term structure models, the functional forms above can be determined as

$$\Phi(X_t) = \Phi_0 + \Phi_X X_t, \tag{3.3}$$

$$Q(X_t) = \sum_{i=1}^{N} Q_{v,i} v_i(t). \tag{3.4}$$

The expressions for Φ_0, Φ_X and $Q_{v,i}$ are found using the continuous-time dynamics of the state vector, $dX_t = \mu_X(X_t)\,dt + \sigma_X^T(X_t)\,dW_t^{\mathbb{P}}$. Although these terms are referred to by Trolle and Schwartz (2009) there are no expressions revealed. The final closed form solutions used in this implementation can be found in Appendix C.

The vector of disturbance, w_t is iid. however, contrary to the standard Kalman filter setup, it is not multivariate normally distributed. This is because the continuous time dynamics of the state vector include the state variables $v_i(t)$ in the diffusion and thus the volatility $\sigma_X^T(X_t)$ is not deterministic. Trolle and Schwartz (2009) do however make the assumption that the disturbance vector w_t is normally distributed, which then lends itself to the method of quasi-maximum likelihood estimation. Assuming w_t is normally distributed, the transition equation can be adjusted from (3.1) and takes the form

$$X_t = \Phi_0 + \Phi_X X_{t-1} + w_t, \quad w_t \sim \text{iid.} \quad \mathcal{N}(0, Q(X_{t-1})). \tag{3.5}$$

The second equation defining the Kalman filter is the measurement equation, which describes the relationship between the state variables and the observed inputs used in the estimation. It is given by

$$y_t = h(X_t) + u_t, \qquad u_t \sim \text{iid.} \quad \mathcal{N}(0, \mathcal{S}). \tag{3.6}$$

In this case, y_t is a size M column vector of input prices which in this case are interest rates and interest rate derivatives. The interest rates are the simulated LIBOR and swap rates and the derivative prices are the ATMF swaption and cap prices. The function $h(X_t)$ implements the semi-analytical pricing formulas to match the elements of y_t and u_t is a column vector of measurement errors which is normally distributed with $\mathcal{S}$ as its covariance matrix. The choice of specification for $\mathcal{S}$ follows the standard Kalman filtering approach of assuming uncorrelated measurement errors and a separate variance for all interest rates and derivatives. Thus, $\mathcal{S}$ is a diagonal matrix with a single variance for all interest rate measurement, σ_{rates}, and another variance for all the observed derivative prices, σ_{deriv}.

The pricing function $h(X_t)$ calculates interest rates using (2.20), (2.41) and (2.42). These equations are non-linearly related to the state variables $x_1(1), ..., x_N(t), \phi_{1,1}(t), \ldots, \phi_{6,N}(t)$. Interest rate derivatives are priced using (2.36) and (2.39) applied to caps and swaptions. These equations are non-linearly related to only the volatility state variables $v_1(t), ..., v_N(t)$, not the term structure state variables. This follows from the use of the calibrated zero-coupon bond curve applied to (2.36).

This implies that $h(X_t)$ is therefore non-linear with respect to the state vector. In order to apply Kalman filtering, $h(X_t)$ must be linear and thus the extended Kalman Filter (EKF) applies. This modifies the measurement equation, with the use of

$$H_t' = \left.\frac{\partial h(X_t)}{\partial X_t'}\right|_{X_t = \hat{X}_{t|t-1}}, \tag{3.7}$$

which is the Jacobian matrix of the pricing function $h(X_t)$ evaluated at $\hat{X}_{t|t-1} = \mathbb{E}_{t-1}[X_t]$, the expectation of X_t given only X_{t-1}. Although Trolle and Schwartz (2009) outline the EKF procedure in the context of the TS model they do not provide insight into the exact specifications required for the modified measurement equation and H_t. For further detail of the modified measurement equation and H_t used here in this implementation see Appendix C.

The recursions of the EKF are defined using $P_{t|t-1}$, the covariance matrix of the estimation error given P_{t-1} and are specified by the prediction equation

$$\hat{X}_{t|t-1} = \Phi_0 + \Phi_X \hat{X}_{t-1} \tag{3.8}$$

$$P_{t|t-1} = \Phi_X P_{t-1} \Phi_X' + Q(\hat{X}_{t-1}), \tag{3.9}$$

and the update equation, which incorporates the information from y_t into the estimation,

$$\hat{X}_t = \hat{X}_{t|t-1} + P_{t|t-1} H_t' F_t^{-1} \epsilon_t, \tag{3.10}$$

$$P_t = P_{t|t-1} - P_{t|t-1} H_t' F_t^{-1} H_t P_{t|t-1}, \tag{3.11}$$

where

$$\epsilon_t = y_t - h(\hat{X}_{t|t-1}), \tag{3.12}$$

$$F_t = H_t P_{t|t-1} H_t' + S. \tag{3.13}$$

Here, define $\hat{X}_t = E_t^{\mathbb{P}}[X_t]$ as the expectation of X_t after including y_t, while P_t is the covariance matrix of the estimation error after including y_t. The covariance P_t is initialised as an $(N \times 8) \times (N \times 8)$ diagonal matrix, usually with a very small positive value across the whole diagonal.

Quasi-Maximum likelihood (QML)

As this implementation does not require parameter estimation, the QML method applied in Trolle and Schwartz (2009) is shown here for completeness.

As mentioned previously, the disturbance vector w_t is not multivariate normally distributed. Making the assumption that the disturbance vector is normally distributed makes this a quasi-maximum likelihood (QML) estimation. With this assumption, the estimation maximises the resultant log-likelihood function

$$\log L = -\frac{1}{2}\log(2\pi) \sum_{t=1}^{T} N_t - \frac{1}{2} \sum_{t=1}^{T} \log|F_t| - \frac{1}{2} \sum_{t=1}^{T} \epsilon_t' F_t^{-1} \epsilon_t, \tag{3.14}$$

for T observation dates, where each N_t is the length of y_t on each date. Duffee and Stanton (2012) consider a number of estimation techniques in the context of term structure modelling and conclude that QML combined with Kalman filtering has better convergence properties in finite samples and is preferable in terms of computational efficiency.

Trolle and Schwartz (2009) mention that QML estimation is not consistent in the context of an extended Kalman filter because of the linearisation and the dependence of Q on the $v_i(t)$ state variable at each time recursion, see (3.4). A study on the small-sample bias properties of the filter is addressed by Trolle and Schwartz (2009) and they find no significant bias present in this setup.

3.3 State variable estimation results

Figure 3.1 illustrates the estimation of the volatility state variables. Estimation was implemented in both cases for the $N = 3$ simulated dataset of LIBOR rates, swap rates and interest rate derivatives.

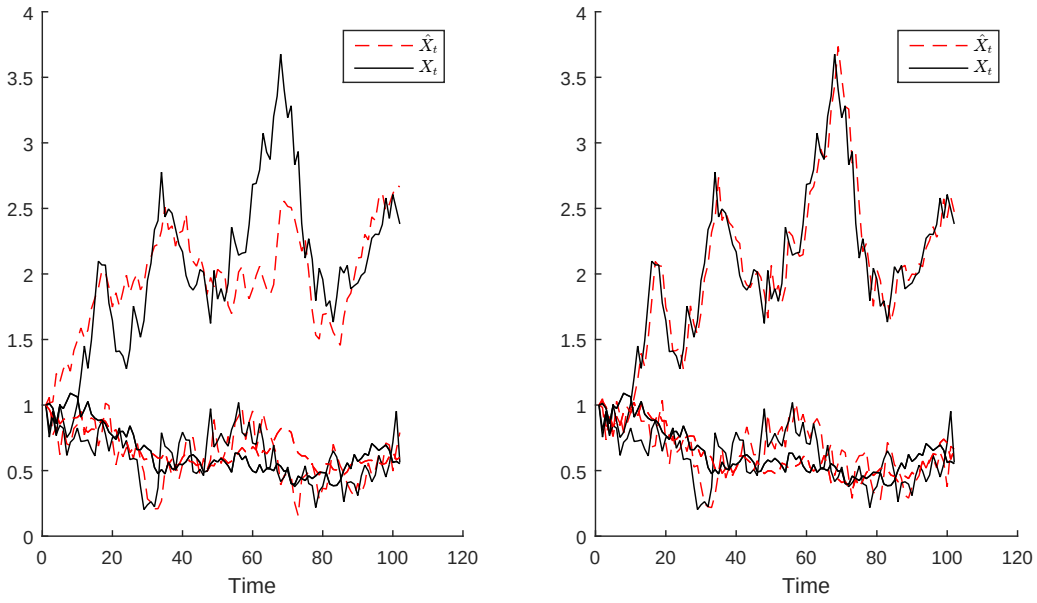

Fig. 3.1: Simulated volatility state variables $v_i(t)$ for $N = 3$ factor model.

The notation X_t illustrates the simulated state variables while $\hat{X}_t$ illustrates the estimated state variables. The left-hand plot shows the results of estimating the state vector using three short-term swaptions, with maturity and swap tenor combinations of $3M \times 1Y$, $6M \times 2Y$ and $1Y \times 3Y$. Adjusting this to three longer dated tenor and maturity combinations for the inputted swaptions is shown in the right-hand plot and notably improves estimation, using $3M \times 1Y$, $3Y \times 5Y$ and $5Y \times 10Y$ swaptions. This fit proved sufficiently accurate in estimating the volatility state variables and therefore no cap prices were used, as swaption prices proved computationally superior without surrendering any accuracy in the estimation.

In their implementation, Trolle and Schwartz (2009) use a much larger set of swaption data and cap skew data, which proves sufficient for accurate parameter estimation. As our implementation avoids all parameter estimation, the only requirement is recovering the latent state variables with a suitable degree of accuracy. Figure 3.1 provides confirmation of the success of the estimation procedure for the volatility state variables. For the full estimated state vector for $N = 3$ (including the term structure state variables) see Appendix C.

Consideration must be made regarding the choice of N in the TS model. In practice, the focus is on reproducing $N = 1, 2$ and 3. Results from principal component analysis reveal that three term structure factors should be sufficient to model the term structure (Litterman and Scheinkman, 1991). In the TS model this aligns with setting $N = 3$, which provides three term structure factors and three additional unspanned stochastic volatility factors. This requires a total number of $N \times 2$ factors. As N increases, the model complexity increases significantly, since the total number of state variables in the state vector is an $N \times 8$ system. For the purposes of this analysis, first $N = 1$ will be considered and thereafter the complexity will be increased by setting $N = 3$.

As mentioned in Chapter 2, a calibrated zero-coupon bond curve will be used to price bonds in (2.36) during estimation with the Kalman filter. Trolle and Schwartz (2009) state that this allows for a cleaner estimation of the volatility state variables because the derivative pricing is no longer affected by the term structure state variables, $x_i(t), \phi_{1,i}(t), ..., \phi_{6,i}(t)$. Therefore, a poor fit to the term structure should not impact on the estimation of the $v_i(t)$ state variables.

As an alternative to bootstrapping the curve according to a market convention, the method adopted by Trolle and Schwartz (2009) is to fit a parametric form for the yield curve, using a parsimonious Nelson and Siegel (1987) calibration. This calibrated yield curve implies a zero-coupon bond curve from interest rate data and is then considered as the observed zero-coupon bond curve. Calibrating to the simulated yield curve out to fifteen years, this parametric model specifies the yield at time t for maturity at time T as

$$y(t,T) = \beta_0 + (\beta_1 + \beta_2)\left(\frac{\tau_1}{T-t}\right)\left(1 - e^{-\frac{T-t}{\tau_1}}\right) - \beta_2 \exp\left(-\frac{T-t}{\tau_1}\right). \tag{3.15}$$

At each point in time, the curve parameters $\beta_0, \beta_1, \beta_2$ and τ_1 are used to fit the observed term structure by minimising the average squared percentage error between rates implied from this parametric curve and the rates from the simulated dataset. Based on the calibrated zero-coupon bond curve, derivative prices are computed from the observed lognormal implied volatilities using Black (1976) pricing formulas for swaptions and caps.

Once analysis has been performed on the simulated dataset, a historical dataset is constructed as a subset of the original dataset used by Trolle and Schwartz (2009). Their estimation dataset consists of United States based interest rates and interest rate derivative prices from 21$^{\text{st}}$ August 1998 to 26$^{\text{th}}$ January 2005, comprising 360 weekly observations of mid-price quotes. The historical dataset inherited for this study are the final 101 weekly observations for interest rates, swap rates, swaption and cap prices.

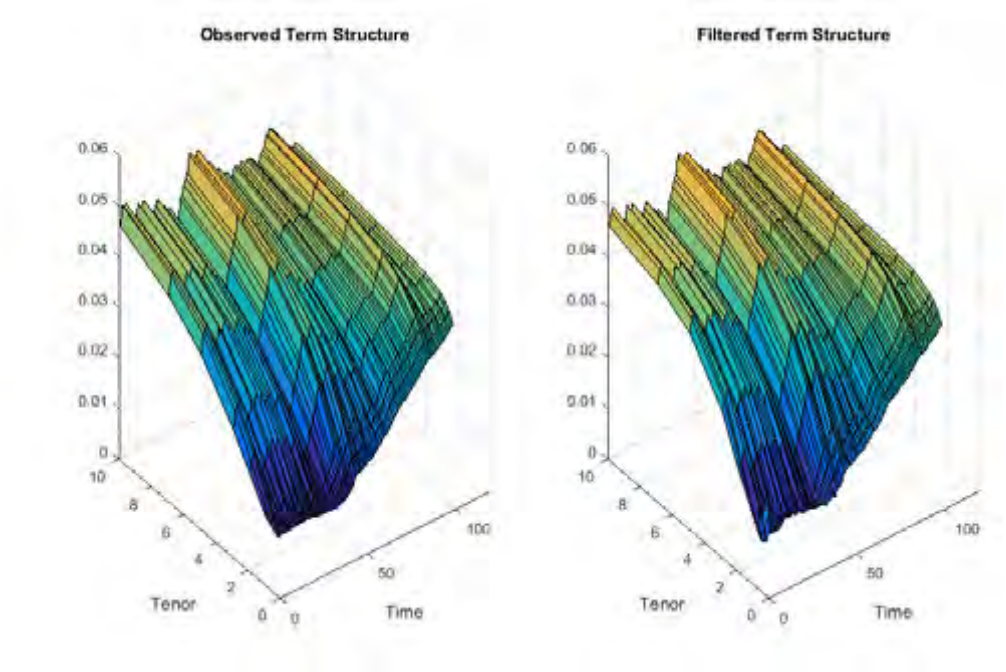

Fig. 3.2: Observed and filtered term structure for empirical dataset, $N = 3$.

Figure 3.2 illustrates the observed term structure for the final 101 observation dates of the Trolle and Schwartz (2009) dataset. The figure on the right is the estimated term structure from the Kalman filter for $N = 3$. The fit to the observed rates appears accurate, which supports the success of the state vector estimation relative to observed market data.

LIBOR Rates		**Swap Rates**		**Swaptions**	
$L(0, 0.25)$	10.35	$S(0, 1)$	-9.68	$3M \times 1Y$	7.58
$L(0, 0.5)$	-2.79	$S(0, 5)$	10.25	$3Y \times 5Y$	-0.25
$L(0, 0.75)$	-7.83	$S(0, 15)$	-9.04	$5Y \times 10Y$	0.02

Tab. 3.2: Mean pricing errors for input instruments using estimated $\hat{X}_t$.

Table 3.2 shows the mean pricing errors for interest rates, as the difference between fitted and observed rates in basis points, and the mean pricing error for swaptions, as the difference between fitted and observed prices as a percentage of the observed prices. Trolle and Schwartz (2009) do achieve superior mean pricing errors in their estimation. This could be attributed to the larger size of their dataset as well as incorporating more derivative prices into their estimation. However, the mean pricing errors in our estimation still seem small enough as to suggest that the estimated state variables successfully recover the observed market data over time.

Chapter 4

Monte Carlo results

The Monte Carlo analysis is presented in two parts; firstly by comparing the results of the different Monte Carlo schemes to compare how each scheme performs and secondly by interpreting the results of the semi-analytical and volatility approximation formulas for pricing bonds, caps and swaptions.

For the purposes of this section, we take the features of the TS model as given and simply consider variations in the model's implementation. The variations considered are

1. Choice of value for N: consider $N = 1$ and $N = 3$.
2. Source of state vector estimation: Basic, simulated and historical.
3. Monte Carlo scheme: full truncation Euler, Milstein or adapted QE.

The state vector is derived from three different sources: the basic state vectors generated without any estimation procedure; the simulated state vectors estimated from the simulated dataset and the historical state vectors estimated using the historical dataset. The schemes employed will include the newly developed full truncation Milstein scheme using (2.5) and (2.6), and the adapted QE scheme using (2.9), (2.10) and (2.19). In addition, a full truncation Euler scheme will be employed for a comparative result, where the dynamics resemble those of the Milstein scheme but without the last term in (2.6). An example of a particular model variation might combine a basic state vector with a $N = 1$ model and the full truncation Euler scheme.

The semi-analytical pricing formulas (2.20), (2.46) and (2.50) are analysed for consistency with the Monte Carlo schemes by comparing the formula results with the Monte Carlo error bounds. A three standard error bound around the Monte Carlo price will be used for general comparison. The discrete Monte Carlo simulations should provide suitable prices as more sample paths are introduced and the time steps become smaller. The tested instruments include zero-coupon bonds, zero-coupon bond options and ATMF swaptions and caps.

4.1 Analysis of TS Monte Carlo schemes

To set up the model for pricing we need only the model parameters in Table 3.1 and the estimated state vector. The EKF estimation described in Chapter 3 provides the method for retrieving the state vector and the initial time-t_0 state vectors used in this Monte Carlo analysis are shown below.

	$N=1$				$N=3$		
	Basic	Simulated	Historical		Basic	Simulated	Historical
$\hat{x}_1(t_0)$	0	22.839642	-1.325143	$\hat{x}_1(t_0)$	0	1.938370	-2.177431
$\hat{\phi}_{1,1}(t_0)$	0	14.117994	-3.575916	$\hat{x}_2(t_0)$	0	8.760007	0.273416
$\hat{\phi}_{2,1}(t_0)$	0	1.126711	1.739020	$\hat{x}_3(t_0)$	0	8.813625	-0.176194
$\hat{\phi}_{3,1}(t_0)$	0	0.844322	1.260232	$\hat{\phi}_{1,1}(t_0)$	0	1.853301	-2.499880
$\hat{\phi}_{4,1}(t_0)$	0	1.032915	1.764088	$\hat{\phi}_{1,2}(t_0)$	0	4.090686	-0.735175
$\hat{\phi}_{5,1}(t_0)$	0	0.684835	1.151111	$\hat{\phi}_{1,3}(t_0)$	0	4.359235	-0.189152
$\hat{\phi}_{6,1}(t_0)$	0	0.816412	1.432192	$\hat{\phi}_{2,1}(t_0)$	0	3.393962	1.768812
$\hat{v}_1(t_0)$	1	0.759818	0.804407	$\hat{\phi}_{2,2}(t_0)$	0	0.450992	0.157256
				$\hat{\phi}_{2,3}(t_0)$	0	0.552160	1.032486
				$\hat{\phi}_{3,1}(t_0)$	0	2.944533	1.429598
				$\hat{\phi}_{3,2}(t_0)$	0	0.232744	0.074260
				$\hat{\phi}_{3,3}(t_0)$	0	0.320707	0.547722
				$\hat{\phi}_{4,1}(t_0)$	0	2.790457	2.157690
				$\hat{\phi}_{4,2}(t_0)$	0	0.318436	0.110386
				$\hat{\phi}_{4,3}(t_0)$	0	0.435114	0.923794
				$\hat{\phi}_{5,1}(t_0)$	0	2.287188	1.682523
				$\hat{\phi}_{5,2}(t_0)$	0	0.105238	0.042562
				$\hat{\phi}_{5,3}(t_0)$	0	0.172332	0.352524
				$\hat{\phi}_{6,1}(t_0)$	0	2.518307	2.333830
				$\hat{\phi}_{6,2}(t_0)$	0	0.086694	0.035133
				$\hat{\phi}_{6,3}(t_0)$	0	0.173249	0.375247
				$\hat{v}_1(t_0)$	1	2.473681	0.378429
				$\hat{v}_2(t_0)$	1	0.663182	0.146628
				$\hat{v}_3(t_0)$	1	0.661061	0.623278

Tab. 4.1: State vector $\hat{X}(t_0)$ used for $N=1$ and $N=3$ Monte Carlo analysis.

The basic state vectors are determined by the initial conditions for the term structure state variables and long run means of the volatility state variables. The term structure state variables, $\hat{x}_i(t_0), \hat{\phi}_{1,i}(t_0), ..., \hat{\phi}_{6,i}(t_0)$, determine the initial bond curve, $P(t_0, t_j)$, which is used to find $\hat{f}(t_0, t_j)$ in (2.5) and (2.19).

	$N=1$			$N=3$		
	Euler	Milstein	QE	Euler	Milstein	QE
Bonds						
$P(t,1)$	0.920168	0.920168	0.920166	0.934255	0.934267	0.934264
	(0.00007)	*(0.00007)*	*(0.00007)*	*(0.00007)*	*(0.00007)*	*(0.00007)*
$P(t,5)$	0.659658	0.659722	0.659680	0.711702	0.711804	0.711789
	(0.00106)	*(0.00106)*	*(0.00107)*	*(0.00092)*	*(0.00092)*	*(0.00092)*
$P(t,10)$	0.435095	0.435233	0.435038	0.506805	0.506399	0.506479
	(0.00185)	*(0.00184)*	*(0.00185)*	*(0.00164)*	*(0.00164)*	*(0.00163)*
Caps						
$1Y$	0.001669	0.001661	0.001666	0.001699	0.001705	0.001721
	(0.00005)	*(0.00005)*	*(0.00005)*	*(0.00005)*	*(0.00005)*	*(0.00005)*
$5Y$	0.028858	0.028673	0.028856	0.027255	0.027292	0.027315
	(0.00082)	*(0.00082)*	*(0.00081)*	*(0.00077)*	*(0.00078)*	*(0.00077)*
$10Y$	0.06525	0.065331	0.065931	0.061605	0.061604	0.062322
	(0.00182)	*(0.00182)*	*(0.00180)*	*(0.00171)*	*(0.00171)*	*(0.00170)*
Swaptions						
$3M \times 1Y$	0.001896	0.001892	0.001889	0.001891	0.001879	0.001893
	(0.00005)	*(0.00005)*	*(0.00005)*	*(0.00005)*	*(0.00005)*	*(0.00005)*
$2Y \times 3Y$	0.018903	0.019070	0.019036	0.015496	0.015334	0.015447
	(0.00053)	*(0.00053)*	*(0.00053)*	*(0.00044)*	*(0.00043)*	*(0.00043)*
$5Y \times 5Y$	0.029919	0.030168	0.030178	0.025889	0.025703	0.025956
	(0.00080)	*(0.00080)*	*(0.00079)*	*(0.00071)*	*(0.00071)*	*(0.00070)*

Tab. 4.2: Monte Carlo prices and error bounds for variations of basic model.

Note: Monte Carlo schemes use 100,000 sample paths and 1M time steps. Size of three standard error bounds around Monte Carlo prices are in brackets. Caps and swaptions prices are for ATMF options.

Table 4.2 shows prices and convergence levels across the three Monte Carlo schemes using the basic state vectors in Table 4.1. The standard error bounds indicate convergence properties of the Monte Carlo schemes. Simulation results for bond prices appears consistent across all schemes. A marginal improvement in convergence in the QE prices for caps and swaptions is in line with intentions of the scheme to improve simulation for $v_i(t)$ and thus improve option pricing. Convergence of bond and option prices for longer maturities improve when setting $N = 3$. This is also expected, as setting $N = 3$ introduces more factors and improves simulation performance.

As these simulations are discrete there might be bias which depends on the size of the discrete time steps. To test the robustness of these simulations we can decrease the size of the time steps and look for stability of prices. Separately we also increase the number of sample paths used to test convergence.

	6M	3M	1M	2W
$P(t,10)$	0.506935	0.506505	0.506502	0.506376
	(0.00233)	*(0.00232)*	*(0.00232)*	*(0.00230)*
$10Y$	/	0.062168	0.062046	0.061888
		(0.00243)	*(0.00243)*	*(0.00242)*
$5Y \times 5Y$	/	0.025790	0.025869	0.025813
		(0.00100)	*(0.00101)*	*(0.00100)*

Tab. 4.3: Effect of decreasing size of time steps for discretised Monte Carlo

Note: Prices reflect the basis state vector, $N = 3$ and an Euler scheme with 50000 sample paths. Results are similar for both Milstein and QE implementations.

Table 4.3 reveals marginal differences in convergence attributed to decreasing the size of the time steps. Using smaller time steps attempts to improve the accuracy of the simulation and since the prices appear fairly stable it suggests that the bias in the discrete Monte Carlo simulation might be small. Time steps of 1M and 2W are used for the remainder of the analysis to produce consistent and accurate pricing. Using smaller time steps is even more important when setting $N = 3$.

	10000	50000	100000	200000
$P(t,10)$	0.506708	0.506481	0.506582	0.506733
	(0.00519)	*(0.00234)*	*(0.00163)*	*(0.00116)*
$10Y$	0.062587	0.062088	0.062331	0.062315
	(0.00544)	*(0.00242)*	*(0.00172)*	*(0.00121)*
$5Y \times 5Y$	0.026218	0.025821	0.025855	0.025878
	(0.00228)	*(0.00100)*	*(0.00071)*	*(0.00050)*

Tab. 4.4: Convergence for increase in sample paths of Monte Carlo simulation

Note: Prices reflect the basis, $N = 3$, Euler scheme with 50000 sample paths. Results are consistent across both the Milstein and QE implementations.

As expected, Table 4.4 illustrates a noticeable decrease in error bounds as number of sample paths increase. The smaller error bounds provide credibility when comparing with the tested pricing formulas. In comparing prices for the remainder of this analysis either 100,000 or 50,000 sample paths will be used.

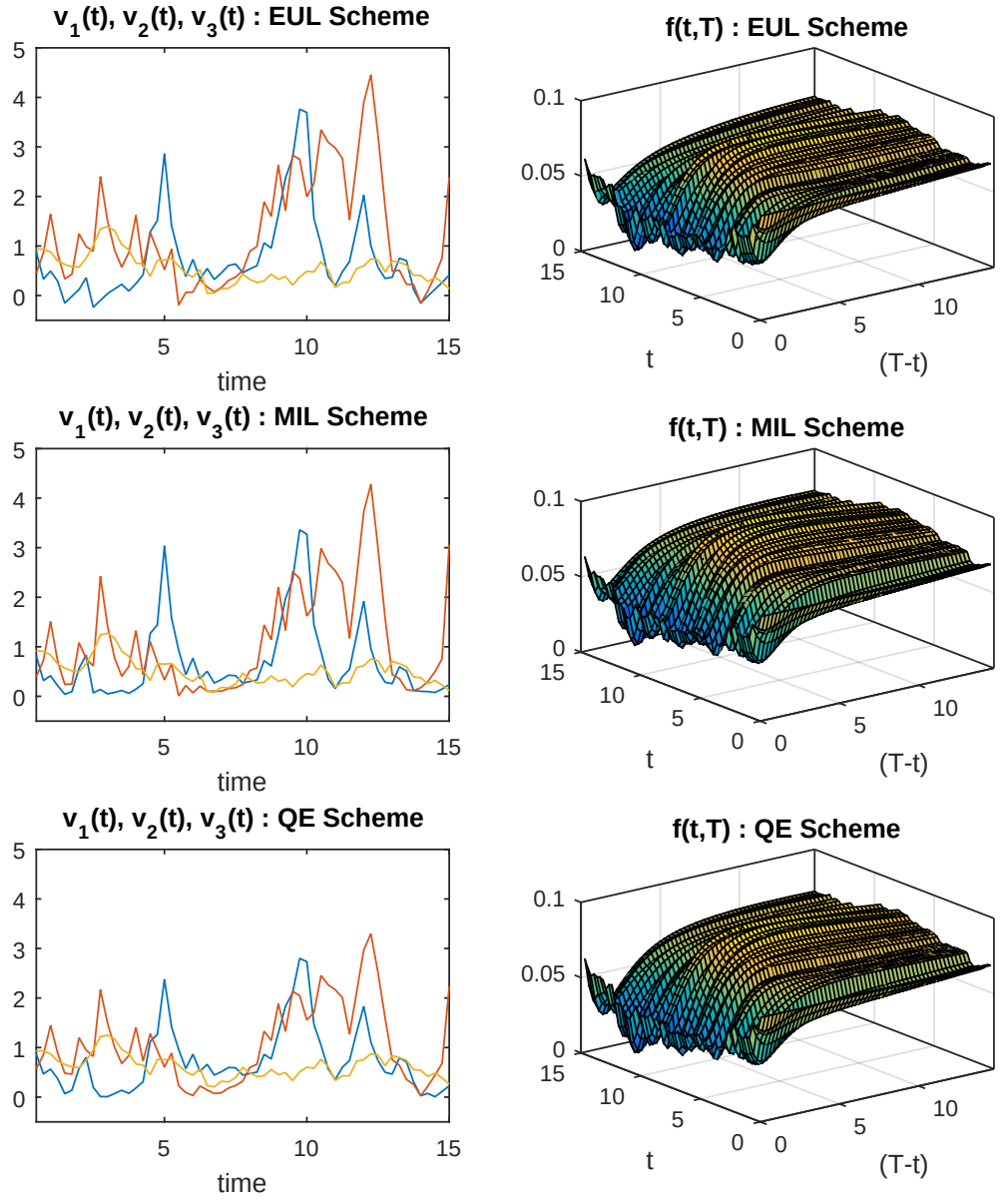

Fig. 4.1: Euler, Milstein and QE scheme simulation comparison for $N = 3$ basic state vector.

Figure 4.1 shows the results of a single simulation for $\hat{v}_i(t)$ and $\hat{f}(t,T)$ under all three schemes, for the same set of random variables. These diagrams illustrate similar properties, with the QE scheme exhibiting narrower bounds for the simulated volatility state variables. The Euler and Milstein simulations are similar, which is to be expected, and both experience negative volatility. The Milstein scheme does however seem to have shorter periods of sustained negative $\hat{v}_i(t)$.

4.2 Analysis of TS pricing formulas

4.2.1 Semi-analytical pricing formula

	Simulated		Historical	
	$N = 1$	$N = 3$	$N = 1$	$N = 3$
Bonds				
$P(t, 1)$	0.638715	0.717961	0.968484	0.961509
$P(t, 5)$	0.085961	0.313909	0.803662	0.809937
$P(t, 10)$	0.019415	0.192943	0.568556	0.641927
Caps				
$1Y$	0.006423	0.002486	0.001559	0.001171
$5Y$	0.028354	0.092566	0.034343	0.024452
$10Y$	0.049241	0.170776	0.088305*	0.062960
Swaptions				
$3M \times 1Y$	0.001293	0.001395	0.001757	0.001137
$2Y \times 3Y$	0.004376***	0.008967	0.019634	0.013545
$5Y \times 5Y$	0.002229***	0.012584	0.033465***	0.027123

Tab. 4.5: Pricing results for semi-analytical for bonds and ATMF derivatives

* Exceeds three standard error bound for QE scheme.

** Exceeds three standard error bound for Milstein scheme.

*** Exceeds three standard error bound for QE and Milstein schemes.

The semi-analytical pricing formulas in Table 4.5 are compared to Monte Carlo results using 100,000 sample paths and 1M time steps. For consistency, the bond prices in (2.36) are determined by the semi-analytical formulas (2.20). Prices are compared to Milstein and QE error bounds. Firstly, zero-coupon bond pricing appears consistent across all model variations. Secondly, the form of the estimated state variable appears to influence results. Comparing caps and swaptions across simulated and historical model reveals mispricing in the simulated but not the historical variations. Thirdly, semi-analytical option pricing is more consistent with the Monte Carlo for $N = 3$ than $N = 1$. The $10Y$ cap in the $N = 1$ historical model is inconsistent with the QE scheme as well as inconsistency seeing in swaption pricing for both simulated and historical.

This is likely a deficiency in the $N = 1$ Monte Carlo simulations, due to the insufficient number of factors in the simulation. The $N = 3$ models allow for refined derivative pricing because the additional factors improve the fit to these non-linear derivatives. See Appendix D for full Monte Carlo pricing results.

The flexible features of the semi-analytical pricing formula appear in both the quadrature method used in evaluating (2.40) and the number of time steps for solving the ODEs in equations (2.37) and (2.38).

	Original	500 [0; 500]	2000 [0; 4000]	8000 [0; 4000]	4000 [0; 4000]
$1Y \times 2Y$ ZCB Option	0.002169	0.002054	0.002169	0.002169	0.002169
$2Y \times 3Y$ ZCB Option	0.013935	0.013936	0.013936	0.013936	0.013936
$5Y \times 5Y$ ZCB Option	0.031233	0.031428	0.031428	0.031428	0.031428
$1Y$ Cap	0.001168	0.000491	0.001171	0.001171	0.001171
$7Y$ Cap	0.039342	0.037058	0.039349	0.039349	0.039349
$15Y$ Cap	0.101256	0.098994	0.101260	0.101260	0.101260
$3M \times 1Y$ Swaption	0.001137	0.000964	0.001137	0.001137	0.001137
$4Y \times 4Y$ Swaption	0.021495	0.021496	0.021496	0.021496	0.021496
$5Y \times 10Y$ Swaption	0.041746	0.041480	0.041480	0.041480	0.041480

Tab. 4.6: Results of quadrature analysis for semi-analytical pricing formulas

The original quadrature method of Trolle and Schwartz (2009) using 20 points across [0; 1000] and another 20 across [1000; 8000] performs favourably in comparison to greater numbers of points over larger intervals. A conservative choice is made to use 4000 quadrature points over [0; 4000] in further pricing. This becomes more important for longer dated derivatives.

	10	100	1000	300
$1Y \times 2Y$ ZCB Option	0.002169	0.002169	0.002169	0.002169
$2Y \times 3Y$ ZCB Option	0.013930	0.013936	0.013936	0.013936
$5Y \times 5Y$ ZCB Option	0.031391	0.031428	0.031428	0.031428
$1Y$ Cap	0.001171	0.001171	0.001171	0.001171
$7Y$ Cap	0.039335	0.039349	0.039349	0.039349
$15Y$ Cap	0.101128	0.101259	0.101260	0.101260
$3M \times 1Y$ Swaption	0.001137	0.001137	0.001137	0.001137
$4Y \times 4Y$ Swaption	0.021472	0.021496	0.021496	0.021496
$5Y \times 10Y$ Swaption	0.041413	0.041479	0.041480	0.041480

Tab. 4.7: Results of ODE analysis for semi-analytical pricing formulas

Table 4.7 shows the effect of changing the number of time steps used to solve ODEs (2.37) and (2.38). The choice of 300 remains on the cautious side but does provide consistency in pricing for multiple tenors.

4.2.2 Volatility approximation pricing formula

	Simulated		Historical	
	$N = 1$	$N = 3$	$N = 1$	$N = 3$
Caps				
$1Y$	0.006428	0.00250	0.001556	0.001173
$5Y$	0.028364	0.09258	0.033705***	0.024621
$10Y$	0.049240	0.17103***	0.087360*	0.062945
Swaptions				
$3M \times 1Y$	0.001299	0.001412	0.001758	0.001129
$2Y \times 3Y$	0.004559	0.008973	0.019794	0.013722**
$5Y \times 5Y$	0.002398**	0.012808**	0.034754**	0.027788**

Tab. 4.8: Pricing results for volatility approximation for ATMF derivatives

* Exceeds three standard error bound for QE scheme.

** Exceeds three standard error bound for Milstein scheme.

*** Exceeds three standard error bound for QE and Milstein schemes.

The volatility approximation formula for swaption pricing in equation (2.63) presents an alternative pricing formula for swaptions which is also used to price caplets and therefore caps. This pricing method is more efficient than the semi-analytical pricing formulas, a result supported by Trolle and Schwartz (2009). However, they do state that this formula is less accurate than the stochastic duration approach used in the semi-analytical pricing formulas.

Table 4.8 explores the accuracy of the approximation formula. These results appear to confirm certain results shown in the semi-analytical pricing section. There is difficulty in consistently pricing longer dated options, while shorter dated options are priced quite consistently. Swaption pricing is consistent with the QE scheme for all model variations but fails for the Milstein scheme. This points to inconsistency between the Milstein and QE schemes. It is not yet clear whether this is a result of the "leaking correlation" of the QE scheme or possibly a bias inherent in the Milstein scheme. The inconsistency of the QE results with the Milstein results will become clearer when we take a closer look into the pricing results for caps and swaptions.

This formula seems inconsistent for calculation of ATMF cap prices. This is not surprising because the method requires the use of the volatility approximation formula for each individual caplet. This can cause even slight errors in caplet pricing to accumulate into noticeable errors in cap pricing. See Appendix D for full Monte Carlo pricing results.

The two features that require robustness testing for the volatility approximation pricing formula are the quadrature method used to calculate the the integral in (2.63) and calculating the expectation of $v_i(t)$ under the forward swap rate measure.

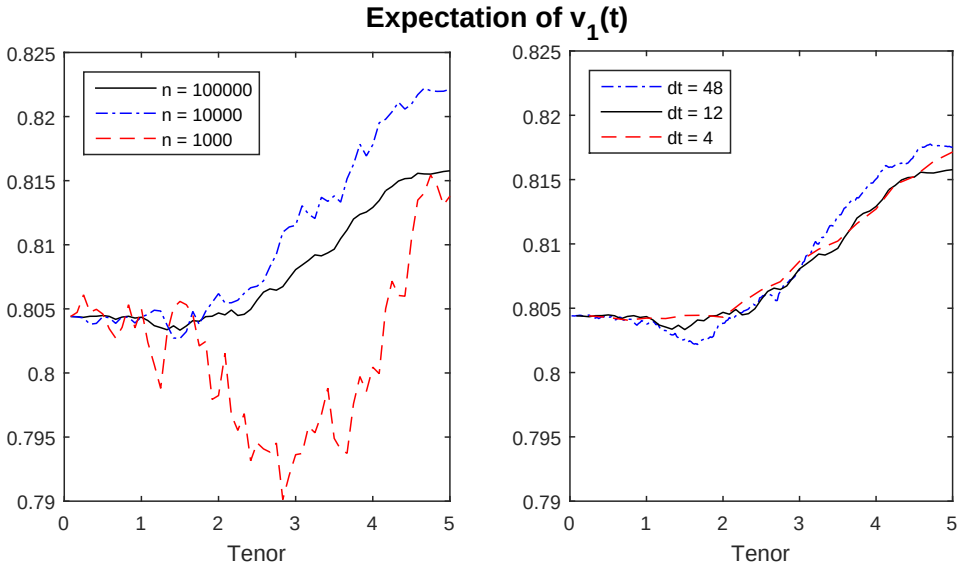

Fig. 4.2: Expectation of $v_1(t)$ in volatility approximation formula for 5-year tenor, $N = 1$ historical state vector.

Figure 4.2 illustrates both the effects of the number of sample paths used in the Monte Carlo simulation to evaluate the expectation of $v_i(t)$ in (2.63) as well as the number of quadrature points per year used to evaluate the integral in (2.63). The expectation is evaluated using a simple Euler scheme on the dynamics of $v_i(t)$ in equation (2.56) in a Monte Carlo simulation. The first plot illustrates the use of this Monte Carlo simulation to estimate the value of the expectation of $v_i(t)$ using increased sample paths. There is clearly a large discrepancy between the number of sample paths used. However, this turns out to only have marginal effects on the pricing, as even using 1000 sample paths yields fairly consistent pricing when compared to pricing using 100,000 sample paths. This will be illustrated further in the sections below.

The more important consideration is the number of the quadrature points used. This has been seen to cause deviations in pricing, however these discrepancies are still relatively small when comparing to the Monte Carlo prices. This gives the impression that this pricing formula is quite robust and computationally efficient although this approximation does appear less consistent with Monte Carlo pricing than the semi-analytical pricing formulas for interest rate derivatives.

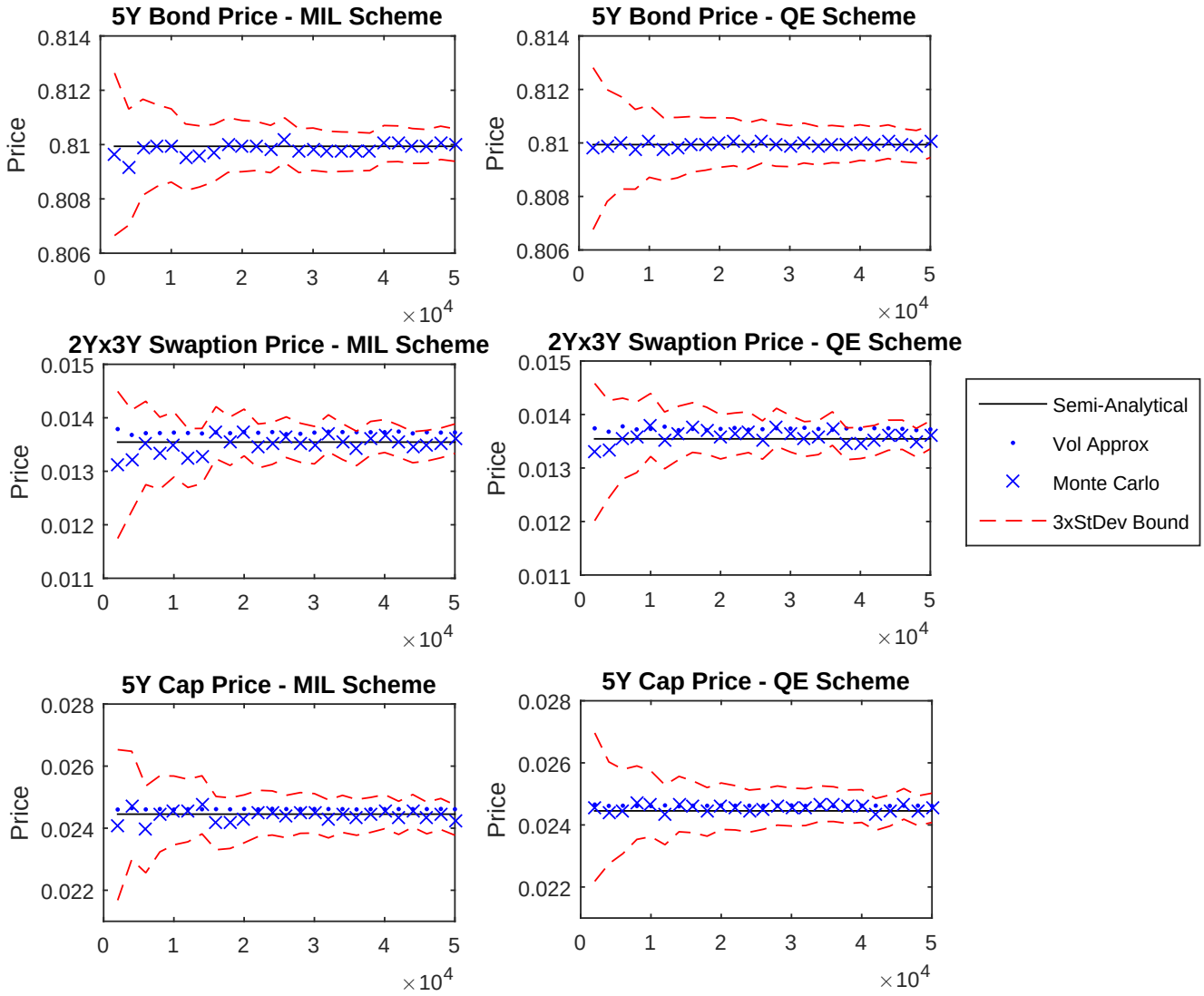

Fig. 4.3: Monte Carlo simulation for $N = 3$ historical state vector and 2W discretised time steps.

Figure 4.3 provides comfort in the accuracy of both the semi-analytical and volatility approximation formulas and the convergence of both the Milstein and QE schemes. Similar results can be shown for different option tenors. A point to note is the slight divergence of the volatility approximation formula in swaption pricing. This trend does illustrate some of the concerns of inconsistency in the volatility approximation formula. It is also worth noting that the volatility approximation formula is using the same number of sample paths as the comparative Monte Carlo. The pricing remains very stable, even as the number of sample paths used in evaluating the expectation increases.

Results thus far suggest that a better method is required for comparison of the semi-analytical and the volatility approximation formulas. An R^2 statistic is calculated to measure the goodness-of-fit between the semi-analytical and Monte Carlo prices. The R^2 statistic measures the propensity for two datasets to move in the same direction, where the value is between zero and one and one indicates a perfect fit.

		N=1		N=3	
		S-A	Vol Approx	S-A	Vol Approx
Caps:	SIM	0.999972	0.999971	0.999985	0.999954
	HIST	0.998972	0.998846	0.998522	0.998695
Swaptions:	SIM	0.985082	0.989729	0.991923	0.991247
	HIST	0.995791	0.994628	0.996053	0.992390

Tab. 4.9: Goodness of fit analysis for pricing formulas using R^2 statistic

Note: Fit compared for 1000 sample paths for simulated Milstein models across 42 Swaption prices and 30 cap prices. Calculation of R^2 statistic for both semi-analytical and volatility approximation formulas

Table 4.9 shows results for a very low number of sample paths because a larger number of paths resulted in R^2 statistics very close to 1 and thus provided less insight into pricing trends. This result supports both pricing methods overall for their use in the TS framework. What this table does show is that the semi-analytical pricing formulas appear to outperform the volatility approximation formulas for all but the $N = 1$ simulated model. This supports the results found by Trolle and Schwartz (2009) that the semi-analytical formulas are more accurate than the volatility approximation formulas.

Option Pricing Analysis

It is important to analyse the best technique to price the interest rate options using the pricing formula and identify where the pricing formulas fail. It is useful to test the pricing formulas across a great number of option tenors and compare the results with the Monte Carlo. At the heart of the semi-analytical formulas is the pricing formula for zero-coupon bond options. This pricing formula is tested here in order to consider the error that might filter through to the dependent option pricing formulas.

So far we have seen that derivative pricing performs better when setting $N = 3$, both for the semi-analytical and volatility approximation pricing formulas. This is encouraging as it is preferable to pricing in an $N = 3$ context in practice.

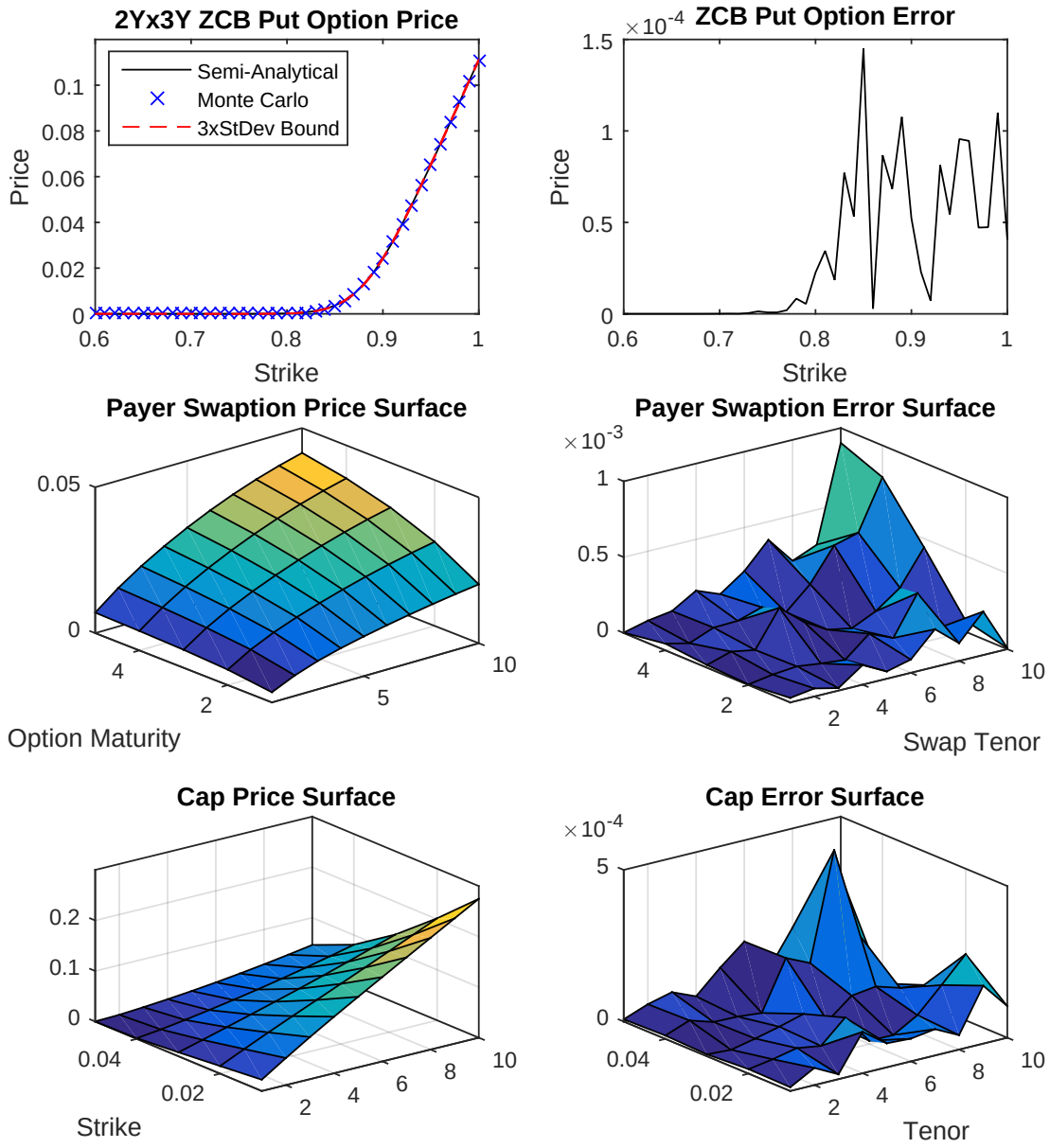

Fig. 4.4: Derivative pricing errors as per semi-analytical and Milstein Monte Carlo simulation.

The absolute pricing error shown in Figure 4.4 for zero-coupon option pricing shows larger errors for pricing in-the money (ITM) and ATMF options. The swaption error surface highlights longer dated swap tenors as the most inconsistent and the cap error surface fails for longer dated out-the-money (OTM) options.

As the larger error in the ZCB option pricing are seen for ATMF levels, or rather where the pricing curve is at its highest point of convexity, the fact that ATMF

interest rate options are still pricing accurately supports the performance of the semi-analytical pricing formulas. The larger error seen for longer dated swaptions is unsurprising as the approximation formula seeks to replace the coupon bond underlying the swaption with a zero coupon bond. The longer the tenor of the coupon bond being approximated the worse the stochastic duration approximation appears to be. For caps, the error in pricing longer dated OTM options may in fact not be due to the semi-analytical pricing formulas failing, but the Milstein scheme unsuccessfully capturing the correlation dynamics. This would result in poorer distributional tails and a reduced ability to price OTM options.

Caplet pricing:

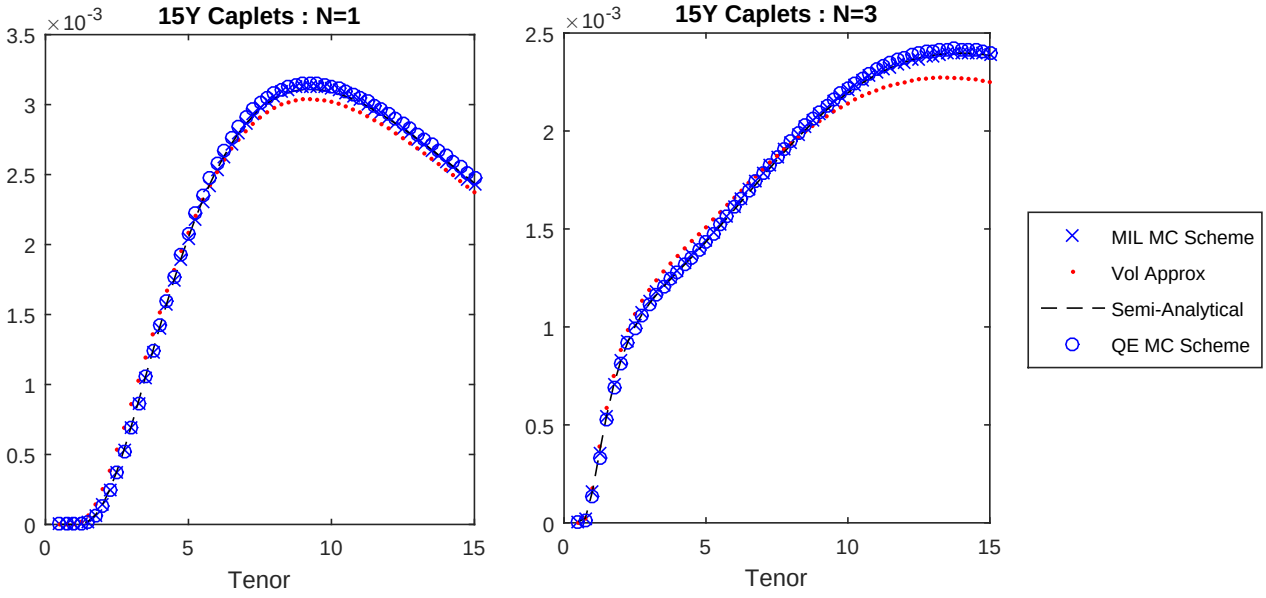

Fig. 4.5: Caplet pricing comparison for $N = 1$ and $N = 3$ with historical state vector.

Caplet pricing reveals vital features of both pricing formulas and the Monte Carlo schemes. It is clear that the lognormal approximation is not the ideal method for pricing longer dated options in the $N = 1$ or $N = 3$ model variations. What is slightly less obvious from Figure 4.5 is the slight deviation of the QE scheme in overestimating caplet prices. This slight discrepancy might link back to the pitfalls warned of by Andersen (2007) for incorrectly accounting for the correlation and subsequent drift adjustment, see Appendix A. Otherwise, pricing appears consistent across the Milstein and semi-analytical pricing formulas. These results were obtained using 100,000 sample paths with a 1M time step.

Swaption pricing

	Milstein	Semi-Analytic	Vol Approx
$1Y \times 9Y$	0.019102	0.019129	0.019576
	[0.018824; 0.019381]	$D(t) = 8.120157$	
$5Y \times 5Y$	0.027202	0.027123	0.027828
	[0.026820; 0.027584]	$D(t) = 9.429723$	
$9Y \times 1Y$	0.007492	0.007522	0.007577
	[0.007391; 0.007593]	$D(t) = 9.981236$	

Tab. 4.10: Swaption comparison for semi-analytical and volatility approximation

Note: Prices compared to 100,000 Milstein scheme for $N = 3$ with historical state vector and 2W time steps.

In this analysis there is a clearer picture of the comparison between the semi-analytical and volatility approximation pricing results for ATMF swaptions. From this analysis it can be shown that the semi-analytic pricing for the $N = 3$ historical state vector model appears consistent across variations for swap tenors and option maturities. The same cannot be said for the volatility approximation price. Inconsistency occurs for both options where the underlying swap tenor is large. This could reflect that the formula fails when the approximations made for the components of the underlying swap under the forward swap measure are inaccurate and as a result the true dynamics are no longer being accurately accounted for.

Where the volatility approximation formula does succeed is for pricing options with shorter dated underlying maturities. It would appear that the volatility approximation formula can be used for swaptions with shorter dated underlying swap tenors. This is useful as the volatility approximation pricing formula is more efficient for pricing swaptions.

Chapter 5

Conclusion

Analysis of the semi-analytical pricing formulas derived in Trolle and Schwartz (2009) requires knowledge of latent state variables inherent in this modelling framework. As a result, the use and extent of the extended Kalman filter is crucial in determining an accurate estimation of the state vector. Following the success of the estimation of the state variables, the semi-analytical pricing formulas appear consistent for parts of both the $N = 1$ and $N = 3$ settings.

The accuracy of the pricing formulas for zero-coupon bonds confirms that the Monte Carlo simulation is performing adequately, for both the adapted Milstein and QE approaches. The slight deviation in the drift of the QE scheme suggests that the Milstein scheme might be more reliable for simulation. The semi-analytical derivative pricing formulas are accurate for pricing of ATMF caps and swaptions. Cap pricing appears consistent across different model variations, which is in line with their derivation from an exact solution for zero-coupon bond options. Swaption pricing proves slightly less accurate, although it is notable that the pricing is better in the $N = 3$ environment compared to the $N = 1$ environment.

Application of the volatility approximation pricing formula for interest rate derivatives appears to show accuracy for a limited set of shorter term derivatives. This method is computationally more efficient than the stochastic duration approximation for swaption pricing and could be used as an alternative pricing method for swaptions with shorter term underlying swap tenors.

This dissertation provides discrete Monte Carlo schemes which may prove useful for extending the analysis to price more complicated interest rate derivatives. Many results are shown here in detail, to assist in replicating the model implementation, that are not shown in great detail by Trolle and Schwartz (2009). An extension of this analysis could involve improving methods for implementing the volatility approximation formula and addressing the bias seen over longer tenors. It may also be worthwhile to develop a consistent QE scheme, better suited to handle the specific correlation structure of this model.

Bibliography

Andersen, L. B. (2007). Efficient Simulation of the Heston Stochastic Volatility Model, *Available at SSRN 946405* .

Andersen, T. G. and Benzoni, L. (2010). Do Bonds Span Volatility Risk in the US Treasury Market? A Specification Test for Affine Term Structure Models, *The Journal of Finance* **65**(2): 603–653.

Andersen, T. G. and Lund, J. (1997). Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate, *Journal of econometrics* **77**(2): 343–377.

Ball, C. A. and Torous, W. N. (1999). The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence, *The Journal of Finance* **54**(6): 2339–2359.

Black, F. (1976). The Pricing of Commodity Contracts, *Journal of financial economics* **3**(1): 167–179.

Brace, A., Gatarek, D. and Musiela, M. (1997). The Market Model of Interest Rate Dynamics, *Mathematical finance* **7**(2): 127–155.

Brigo, D. and Mercurio, F. (2007). *Interest Rate Models-Theory and Practice: with smile, inflation and credit*, Springer Science & Business Media.

Broadie, M. and Kaya, Ö. (2006). Exact Simulation of Stochastic Volatility and other Affine Jump Diffusion Processes, *Operations Research* **54**(2): 217–231.

Casassus, J., Collin-Dufresne, P. and Goldstein, B. (2005). Unspanned Stochastic Volatility and Fixed Income Derivatives Pricing, *Journal of Banking & Finance* **29**(11): 2723–2749.

Cheridito, P., Filipović, D. and Kimmel, R. L. (2007). Market Price of Risk Specifications for Affine Models: Theory and Evidence, *Journal of Financial Economics* **83**(1): 123–170.

Cheyette, O. (1992). Markov representation of the heath-jarrow-morton model, *Available at SSRN 6073* .

Chiarella, C. and Kwon, O. K. (2000). A complete markovian stochastic volatility model in the hjm framework, *Asia-Pacific Financial Markets* **7**(4): 293–304.

Collin-Dufresne, P. and Goldstein, R. S. (2002). Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility, *The Journal of Finance* **57**(4): 1685–1730.

Collin-Dufresne, P. and Goldstein, R. S. (2003). Generalizing the Affine Framework to HJM and Random Field Models, *Available at SSRN 410421* .

Duffee, G. R. and Stanton, R. H. (2012). Estimation of Dynamic Term Structure Models, *The Quarterly Journal of Finance* **2**(02): 1250008.

Duffie, D. and Kan, R. (1996). A Yield-Factor Model of Interest Rates, *Mathematical finance* **6**: 379–406.

Duffie, D., Pan, J. and Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* **68**(6): 13431376.

Fan, R., Gupta, A. and Ritchken, P. (2003). Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets, *Journal of Finance* pp. 2219–2248.

Fisher, M. and Gilles, C. (1996). Estimating Exponential-Affine Models of the Term Structure, *Unpublished working paper, Federal Reserve Board, Washington, DC* .

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*, Vol. 53, Springer Science & Business Media.

Han, B. (2007). Stochastic Volatilities and Correlations of Bond Yields, *The Journal of Finance* **62**(3): 1491–1524.

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.

Heidari, M. and Wu, L. (2003). Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?, *Journal of Fixed Income* **13**(1): 75–86.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of financial studies* **6**(2): 327–343.

Ho, T. S. and Lee, S.-B. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance* pp. 1011–1029.

Hull, J. and White, A. (1990). Pricing Interest-Rate-Derivative Securities, *Review of financial studies* **3**(4): 573–592.

Jamshidian, F. (1997). LIBOR and Swap Market Models and Measures, *Finance and Stochastics* **1**(4): 293–330.

Jarrow, R., Li, H. and Zhao, F. (2007). Interest Rate Caps Smile too! But can the LIBOR Market Models Capture the Smile?, *The Journal of Finance* **62**(1): 345–382.

Kahl, C. and Jäckel, P. (2006). Fast Strong Approximation Monte Carlo Schemes for Stochastic Volatility Models, *Quantitative Finance* **6**(6): 513–536.

Li, H. and Zhao, F. (2006). Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives, *The Journal of Finance* **61**(1): 341–378.

Litterman, R. B. and Scheinkman, J. (1991). Common Factors Affecting Bond Returns, *The Journal of Fixed Income* **1**(1): 54–61.

Lord, R., Koekkoek, R. and Dijk, D. V. (2010). A Comparison of Biased Simulation Schemes for Stochastic Volatility Models, *Quantitative Finance* **10**(2): 177–194.

Munk, C. (1999). Stochastic Duration and Fast Coupon Bond Option Pricing in Multi-Factor Models, *Review of Derivatives Research* **3**(2): 157–181.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves, *Journal of business* pp. 473–489.

Rouah, F. D. (2013). *The Heston Model and Its Extensions in Matlab and C#*, John Wiley & Sons.

Trolle, A. B. and Schwartz, E. S. (2009). A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives, *Review of Financial Studies* **22**(5): 2007–2057.

Wei, J. Z. (1997). A Simple Approach to Bond Option Pricing, *Journal of Futures Markets* **17**(2): 131–160.

Appendix A

Discrete Monte Carlo formulation

Method 1: Specify $\hat{\mu}_f(t,T)$ for simulating $\hat{f}(t,T)$

The discrete-time dynamics of $f(t,T)$ and $v_i(t)$ can be specified by simple Euler and Milstein discretisations respectively. Adopting the full truncation scheme of Lord, Koekkoek and Dijk (2010) for $\hat{v}_i(t_k)$, the discretised forward rate is

$$\hat{f}(t_k,t_j) = \hat{f}(t_{k-1},t_j) + \hat{\mu}_f(t_{k-1},t_j)\,\Delta t_k + \sum_{i=1}^{N} \hat{\sigma}_{f,i}(t_{k-1},t_j)\sqrt{\hat{v}_i(t_{k-1})^+}\sqrt{\Delta t_k}\,X_{i,k},$$

where *iid.* $X_{i,k} \sim N(0,1)$, $\Delta t_k = (t_k - t_{k-1})$, full truncation applies where $\hat{v}_i(t_{k-1})^+ = \max(\hat{v}_i(t_{k-1},0)$ addresses negative volatility, while $\hat{\sigma}_{f,i}(t_{k-1},t_j) = (\alpha_{0,i} + \alpha_{1,i}(t_j - t_{k-1}))e^{-\gamma_i(t_j - t_{k-1})}$. What remains is to determine $\hat{\mu}_f(t_{k-1},t_j)$.

The form of $\hat{\mu}_f(t_{k-1},t_j)$ can be determined following analysis by Glasserman (2003), which ensures that under the discretised model, discounted bond prices are still martingales. Let us consider a zero-coupon bond and the numeraire, the money market account

$$\hat{B}(t_k,t_j) = \exp\left(-\sum_{l=k}^{j-1} \hat{f}(t_k,t_l)(t_{l+1}-t_l)\right), \tag{A.1}$$

$$\hat{\beta}_k = \exp\left(-\sum_{l=0}^{k-1} \hat{f}(t_l,t_l)(t_{l+1}-t_l)\right), \tag{A.2}$$

where

$$\mathbb{E}^{\mathbb{Q}}_{t_{k-1}}\left[\hat{B}(t_k,t_j)\right] = \frac{\hat{\beta}_k}{\hat{\beta}_{k-1}}\hat{B}(t_{k-1},t_j)$$

$$\mathbb{E}^{\mathbb{Q}}_{t_{k-1}}\left[\exp\left(-\sum_{l=k}^{j-1} \hat{f}(t_k,t_l)(t_{l+1}-t_l)\right)\right] = \exp\left(-\sum_{l=k}^{j-1} \hat{f}(t_{k-1},t_l)(t_{l+1}-t_l)\right).$$

Substituting (A.1) from above result in

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}_{t_{k-1}}\left[\exp\left(-\sum_{l=k}^{j-1}\sum_{i=1}^{N}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\sqrt{\hat{v}_i(t_{k-1})}\sqrt{(t_k-t_{k-1})}X_{i,k}\right)\right]\\
&\quad=\exp\left(\sum_{l=k}^{j-1}\hat{\mu}_f(t_{k-1},t_l)(t_k-t_{k-1})(t_{l+1}-t_l)\right)\\
&\mathbb{E}^{\mathbb{Q}}_{t_{k-1}}\left[\exp\left(-\sum_{i=1}^{N}\sum_{l=k}^{j-1}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\sqrt{\hat{v}_i(t_{k-1})}\sqrt{(t_k-t_{k-1})}X_{i,k}\right)\right]\\
&\quad=\exp\left(\sum_{l=k}^{j-1}\hat{\mu}_f(t_{k-1},t_l)(t_k-t_{k-1})(t_{l+1}-t_l)\right)\\
&\mathbb{E}^{\mathbb{Q}}_{t_{k-1}}\left[\exp\left(-\sum_{i=1}^{N}\sqrt{\hat{v}_i(t_{k-1})}\sqrt{(t_k-t_{k-1})}X_{i,k}\sum_{l=k}^{j-1}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\right)\right]\\
&\quad=\exp\left(\sum_{l=k}^{j-1}\hat{\mu}_f(t_{k-1},t_l)(t_k-t_{k-1})(t_{l+1}-t_l)\right)\\
&\qquad\exp\left(\frac{1}{2}\sum_{i=1}^{N}\hat{v}_i(t_{k-1})(t_k-t_{k-1})\left(\sum_{l=k}^{j-1}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\right)^2\right)\\
&\quad=\exp\left(\sum_{l=k}^{j-1}\hat{\mu}_f(t_{k-1},t_j)(t_k-t_{k-1})(t_{l+1}-t_l)\right),
\end{aligned}\tag{A.3}$$

where the last step follows from the characteristic function for the sum of the independent, normally distributed $X_{i,k}$. Therefore, for (A.3) to hold

$$\frac{1}{2}\sum_{i=1}^{N}\hat{v}_i(t_{k-1})\left(\sum_{l=k}^{j-1}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\right)^2=\sum_{l=k}^{j-1}\hat{\mu}_f(t_{k-1},t_l)(t_{l+1}-t_l).\tag{A.4}$$

This is only true if

$$\begin{aligned}
\hat{\mu}_f(t_{k-1},t_j)(t_{l+1}-t_l)&=\frac{1}{2}\sum_{i=1}^{N}\hat{v}_i(t_{k-1})\left(\sum_{l=k}^{j}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\right)^2\\
&\quad-\frac{1}{2}\sum_{i=1}^{N}\hat{v}_i(t_{k-1})\left(\sum_{l=k}^{j-1}\hat{\sigma}_{f,i}(t_{k-1},t_l)(t_{l+1}-t_l)\right)^2\\
\hat{\mu}_f(t_{k-1},t_j)&=\frac{\sum_{i=1}^{N}\hat{v}_i(t_{k-1})(S_{i,j}^2-S_{i,j-1}^2)}{2(t_{l+1}-t_l)},
\end{aligned}\tag{A.5}$$

where $\Delta t_k=(t_k-t_{k-1})=(t_l-t_{l-1})$ and

$$S_{i,n}=\sum_{l=k}^{n}\hat{\sigma}_{f,i}(t_{k-1},t_l)\Delta t_k.$$

Method 2: Adapted $\hat{f}(t,T)$ for QE algorithm

The analysis by Andersen (2007) reveals inconsistency in the Euler discretisation described in (2.19). For the Heston (1993) model's stock price process Andersen (2007) abandons the standard Euler discretisation because of the non-linear relationship between the standard normal $X_{i,k}$ and $\hat{v}_i(t_k)$. Andersen (2007) instead revisits the exact solution of the original SDE as part of the QE algorithm and uses a Cholesky decomposition to invoke this lost correlation. This approach is not replicable in the context of the TS model, as the diffusion term for $\hat{f}(t,T)$ includes the forward rate volatilities $\sigma_{f,i}(t,T)$, which prevents direct substitution of the exact solution for $v_i(t)$ because the exact solution does not contain forward rate volatilities.

Andersen (2007) considers the formulation in (2.19) as "leaking correlation" and thus not the ideal simulation approach. However, results in the study do show that this may have less of an impact in practice when comparing resulting implied correlations. Andersen (2007) also states that even when revisiting the exact solutions a martingale correction is required to correctly adjust the drift term in simulations.

To complete this simple approach to the QE scheme, the dynamics for $f(t,T)$ are restated below for an Euler scheme in line with what Andersen (2007) considered the naive approach for simulation

$$df(t,T) = \mu_f(t,T)\,dt + \sum_{i=1}^{N} \sigma_{f,i}(t,T)\sqrt{v_i(t)}\,dW_i^{\mathbb{Q}}(t),$$

where $W_i^{\mathbb{Q}}(t)$ are *iid.* standard Brownian motions under $\mathbb{Q}$. It follows that the exact solution for $f(t,T)$ is

$$f(t_k,t_j) = f(t_{k-1},t_j) + \int_{t_{k-1}}^{t_k} \mu_f(s,t_j)ds + \sum_{i=1}^{N} \int_{t_{k-1}}^{t_k} \sigma_{f,i}(s,t_j)\sqrt{v_i(s)}\,dW_i^{\mathbb{Q}}(s).$$

The drift component can then be approximated with an Euler-like discretisation. Following from the definitions for $\hat{\mu}_f(t_{k-1},t_j)$ in (A.5) the discretised drift term is defined as

$$\int_{t_{k-1}}^{t_k} \mu_f(s,t_j)ds \approx \hat{\mu}_f(t_{k-1},t_j)\Delta t_k. \tag{A.6}$$

Since $W_i^{Q}(t)$ are independent of $v_i(t)$, conditional on knowing $v_i(t_{k-1})$, the Ito integrals in (A.6) are normally distributed with mean zero and variance of

$$\int_{t_{k-1}}^{t_k} \sigma_{f,i}^2(s,t_j)v_i(s)ds \approx \hat{\sigma}_{f,i}^2(t_{k-1},t_j)\hat{v}_i(t_{k-1}),\Delta t_k \tag{A.7}$$

where $\hat{\sigma}_{f,i}(t_{k-1},t_k)$ follows from the results from Appendix A above. Using correlated normally distributed random variables, the approximation formula is defined as in (2.19).

Appendix B

Derivation of semi-analytical pricing formulas

The following derivations are expansions of already proven results shown by Trolle and Schwartz (2009). The original proofs are expanded to illustrate their origins and identify any appropriate approximations required for pricing.

Semi-analytical price for $P(t,T)$

Applying the Markovian representation for forward rate volatilities in (2.4), (2.3) becomes

$$\begin{aligned}
\mu_f(t,T) &= \sum_{i=1}^{N} v_i(t)\left(\alpha_{0,i}+\alpha_{i,1}(T-t)\right)e^{-\gamma_i(T-t)}\int_t^T\left(\alpha_{0,i}+\alpha_{i,1}(u-t)\right)e^{-\gamma_i(u-t)}du \\
&= \sum_{i=1}^{N} v_i(t)\left(\alpha_{0,i}+\alpha_{i,1}(T-t)\right)e^{-\gamma_i(T-t)} \\
&\quad \left[\frac{\alpha_{0,i}}{\gamma_i}\left(1-e^{-\gamma_i(T-t)}\right)-(T-t)\frac{\alpha_{1,i}}{\gamma_i}e^{-\gamma_i(T-t)}+\frac{\alpha_{1,i}}{\gamma_i^2}\left(1-e^{-\gamma_i(T-t)}\right)\right] \\
&= \sum_{i=1}^{N} v_i(t)\left[\frac{\alpha_{0,i}\alpha_{1,i}}{\gamma_i}\left(\frac{1}{\gamma_i}+\frac{\alpha_{0,i}}{\alpha_{1,i}}\right)\left(e^{-\gamma_i(T-t)}-e^{-2\gamma_i(T-t)}\right)\right. \\
&\quad -(T-t)\frac{\alpha_{0,i}\alpha_{1,i}}{\gamma_i}e^{-2\gamma_i(T-t)}+\frac{\alpha_{1,i}^2}{\gamma_i}\left(\frac{1}{\gamma_i}+\frac{\alpha_{0,i}}{\alpha_{1,i}}\right) \\
&\quad \left.(T-t)\left(e^{-\gamma_i(T-t)}-e^{-2\gamma_i(T-t)}\right)-(T-t)^2\frac{\alpha_{1,i}^2}{\gamma_i}e^{-2\gamma_i(T-t)}\right]. \qquad \text{(B.1)}
\end{aligned}$$

Therefore, the solution, $f(t,T)$, at time-t for instantaneous borrowing at time-T is

$$f(t,T) = f(0,T)+\int_0^t \mu_f(s,T)ds+\sum_{i=1}^{N}\int_0^t \sigma_{f,i}(s,T)\sqrt{v_i(s)}\,dW_i^{\mathbb{Q}}(s),$$

which can be evaluated as

$$
\begin{aligned}
f(t,T) = f(0,T) + \int_0^t \sum_{i=1}^N v_i(s) \Bigg[& \frac{\alpha_{0,i}\alpha_{1,i}}{\gamma_i}\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right)\left(e^{-\gamma_i(T-s)} - e^{-2\gamma_i(T-s)}\right) \\
& - (T-s)\frac{\alpha_{0,i}\alpha_{1,i}}{\gamma_i}e^{-2\gamma_i(T-s)} + \frac{\alpha_{1,i}^2}{\gamma_i}\left(\frac{1}{\gamma_i} + \frac{\alpha_{0,i}}{\alpha_{1,i}}\right) \times \\
& (T-s)\left(e^{-\gamma_i(T-s)} - e^{-2\gamma_i(T-s)}\right) - (T-s)^2\frac{\alpha_{1,i}^2}{\gamma_i}e^{-2\gamma_i(T-s)}\Bigg] ds \\
& + \sum_{i=1}^N \int_0^t \left(\alpha_{0,i} + \alpha_{i,1}(T-s)\right) e^{-\gamma_i(T-s)}\sqrt{v_i(s)}\, dW_i^{\mathbb{Q}}(s) \\
= \sum_{i=1}^N \Bigg(& \mathcal{B}_{\phi_{2,i}}(T-t)\phi_{2,i}(t) + \mathcal{B}_{\phi_{3,i}}(T-t)\phi_{3,i}(t) \\
& + \mathcal{B}_{\phi_{4,i}}(T-t)\phi_{4,i}(t) + \mathcal{B}_{\phi_{5,i}}(T-t)\phi_{5,i}(t) + \mathcal{B}_{\phi_{6,i}}(T-t)\phi_{6,i}(t)\Bigg) \\
& + \sum_{i=1}^N \Big(\mathcal{B}_{x_i}(T-t)x_i(t) + \mathcal{B}_{\phi_{1,i}}(T-t)\phi_{1,i}(t)\Big) \\
= f(0,T) + \sum_{i=1}^N & \mathcal{B}_{x_i}(T-t)x_i(t) + \sum_{i=1}^N\sum_{j=1}^6 \mathcal{B}_{\phi_{j,i}}(T-t)\phi_{j,i}(t). \qquad \text{(B.2)}
\end{aligned}
$$

Appropriate simplification shows $\mathcal{B}_{x_i}(T-t)$ and $\mathcal{B}_{\phi_{j,i}}(T-t)$ as constants, where $B_{x_i}(T-t)$ and $B_{\phi_{j,i}}(T-t)$ in (2.20) follow from simple integration of $\mathcal{B}_{x_i}(T-t)$ and $\mathcal{B}_{\phi_{j,i}}(T-t)$. The state variables take the form

$$x_i(t) = \int_0^t \sqrt{v_i(t)}e^{-\gamma_i(t-s)}\, dW_i^{\mathbb{Q}}(s), \qquad \text{(B.3)}$$

$$\phi_{1,i}(t) = \int_0^t \sqrt{v_i(s)}(t-s)e^{-\gamma_i(t-s)}\, dW_i^{\mathbb{Q}}(s), \qquad \text{(B.4)}$$

$$\phi_{2,i}(t) = \int_0^t v_i(s)e^{-\gamma_i(t-s)}ds, \qquad \text{(B.5)}$$

$$\phi_{3,i}(t) = \int_0^t v_i(s)e^{-2\gamma_i(t-s)}ds, \qquad \text{(B.6)}$$

$$\phi_{4,i}(t) = \int_0^t v_i(s)(t-s)e^{-\gamma_i(t-s)}ds, \qquad \text{(B.7)}$$

$$\phi_{5,i}(t) = \int_0^t v_i(s)(t-s)e^{-2\gamma_i(t-s)}ds, \qquad \text{(B.8)}$$

$$\phi_{6,i}(t) = \int_0^t v_i(s)(t-s)^2e^{-2\gamma_i(t-s)}ds. \qquad \text{(B.9)}$$

As an illustrative example, the dynamics of $\phi_{2,i}(t)$ as expressed in (2.30) is found by changing (B.5) to

$$\phi_{2,i}(t) = e^{-\gamma_i t}\int_0^t v_i(s)e^{\gamma_i s}ds. \qquad \text{(B.10)}$$

Ito's Lemma and the fundamental theorem of calculus then implies

$$\begin{aligned} d\phi_{2,i}(t) &= \left[e^{-\gamma_i t} v_i(t) e^{\gamma_i t} - \gamma_i e^{-\gamma_i t} \int_0^t v_i(s) e^{\gamma_i s} ds\right] dt \\ &= \left[v_i(t) - \gamma_i \int_0^t v_i(s) e^{-\gamma_i (t-s)} ds\right] dt \\ &= \left[v_i(t) - \gamma_i \phi_{2,i}(t)\right] dt. \end{aligned}$$

Solution to transform $\Psi(u, t, T_0, T_1)$

The expressions for (2.36), (2.37) and (2.38), follow from the methods of Duffie, Pan and Singleton (2000) and Collin-Dufresne and Goldstein (2003). First note that

$$\begin{aligned} \Psi(u, T_0, T_0, T_1) &= \mathbb{E}_{T_0}^{\mathbb{Q}} \left[\exp\left(-\int_{T_0}^{T_0} r_s ds\right) \exp\left(u \log(P(T_0, T_1))\right)\right] \\ &= \exp\left(u \log(P(T_0, T_1))\right), \end{aligned} \tag{B.11}$$

which implies that

$$\exp\left(-\int_0^t r_s ds\right) \Psi(u, t, T_0, T_1) = \mathbb{E}_t^{\mathbb{Q}} \left[\exp\left(-\int_0^{T_0} r_s ds\right) \Psi(u, T_0, T_0, T_1)\right]. \tag{B.12}$$

Therefore, an expression must be found for $\Psi(u, t, T_0, T_1)$, such that the process

$$\eta(t) = \exp\left(-\int_0^t r_s ds\right) \Psi(u, t, T_0, T_1), \tag{B.13}$$

is a martingale under $\mathbb{Q}$. To ensure that this is the case, we need to show that:

$$\mathbb{E}_t^{\mathbb{Q}} \left[\frac{d\eta(t)}{\eta(t)}\right] = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{d\Psi(u, t, T_0, T_1)}{\Psi(u, t, T_0, T_1)} - r dt\right] = 0.$$

This is shown by a careful choice of $\Psi(u, t, T_0, T_1)$, and specifying $\frac{\partial M(T_0 - t)}{\partial t}$ and $\frac{\partial N_i(T_0 - t)}{\partial t}$ to set the drift, minus the risk-free rate, to zero. Collin-Dufresne and Goldstein (2003) specify the solution of $\Psi(u, t, T_0, T_1)$ as seen in (2.36). Application of Ito's Lemma shows

$$\begin{aligned} \frac{d\Psi(u, t, T_0, T_1)}{\Psi(u, t, T_0, T_1)} &= \left(-\frac{\partial M(T_0 - t)}{\partial t} - \sum_{i=1}^N \frac{\partial N_i(T_0 - t)}{\partial t} v_i(t)\right) dt + \sum_{i=1}^N N_i(T_0 - t) dv_i(t) \\ &+ u \frac{dP(t, T_1)}{P(t, T_1)} + (1 - u) \frac{dP(t, T_0)}{P(t, T_0)} + u \frac{dP(t, T_1)}{P(t, T_1)} \sum_{i=1}^N N_i(T_0 - t) dv_i(t) \\ &+ (1 - u) \frac{dP(t, T_0)}{P(t, T_0)} \sum_{i=1}^N N_i(T_0 - t) dv_i(t) \\ &+ u(1 - u) \frac{dP(t, T_0) dP(t, T_1)}{P(t, T_0) P(t, T_1)} + \frac{1}{2} \sum_{i=1}^N N_i(T_0 - t)^2 (dv_i(t))^2 \\ &+ \frac{1}{2}(u^2 - u) \frac{(dP(t, T_1))^2}{P(t, T_1)^2} + \frac{1}{2}((1 - u)^2 - (1 - u)) \frac{(dP(t, T_0))^2}{P(t, T_0)^2}. \end{aligned}$$

Trolle and Schwartz (2009) state the dynamics of $P(t,T)$ from (2.20) under $\mathbb{Q}$ as

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + \sum_{i=1}^{N} B_{x_i}(T-t)\sqrt{v_i(t)}dW_i^{\mathbb{Q}}(t). \tag{B.14}$$

Using the dynamics of $v_i(t), P(t,T_0)$ and $P(t,T_1)$ under $\mathbb{Q}$, it follows that

$$\begin{aligned}
\frac{d\Psi(u,t,T_0,T_1)}{\Psi(u,t,T_0,T_1)} = \Bigg(& -\frac{\partial M(T_0-t)}{\partial t} - \sum_{i=1}^{N} \frac{\partial N_i(T_0-t)}{\partial t} v_i(t) \\
& + \sum_{i=1}^{N} N_i(T_0-t)\kappa_i(\theta_i - v_i(t)) + ur_t + (1-u)r_t \\
& + u\sum_{i=1}^{N} N_i(T_0-t)\sigma_i\rho_i v_i(t) B_{x_i}(T_1-t) \\
& + (1-u)\sum_{i=1}^{N} N_i(T_0-t)\sigma_i\rho_i v_i(t) B_{x_i}(T_0-t) \\
& + u(1-u)\sum_{i=1}^{N} B_{x_i}(T_1-t)B_{x_i}(T_0-t)v_i(t) \\
& + \frac{1}{2}\sum_{i=1}^{N} N_i(T_0-t)^2\sigma_i^2 v_i(t) + \frac{1}{2}(u^2-u)\sum_{i=1}^{N} B_{x_i}(T_1-t)^2 v_i(t) \\
& + \frac{1}{2}((1-u)^2-(1-u))\sum_{i=1}^{N} B_{x_i}(T_1-t)^2 v_i(t)\Bigg)dt + \sum_{i=1}^{N}(\dots)dW_i^{Q}(t).
\end{aligned}$$

After subtracting the risk-free rate, the drift term is set to zero by choosing $\partial M(T_0-t)/\partial t$ to cancel terms without $v_i(t)$ and $\partial N_i(T_0-t)/\partial t$ for terms with $v_i(t)$ so that

$$\frac{\partial M(T_0-t)}{\partial t} = \sum_{i=1}^{N} N_i(T_0-t)\kappa_i\theta_i, \tag{B.15}$$

$$\begin{aligned}
\frac{\partial N_i(T_0-t)}{\partial t} = & N_i(T_0-t)(-\kappa_i + \sigma_i\rho_i(uB_{x_i}(T_1-t) + (1-u)B_{x_i}(T_0-t))) \\
& + \frac{1}{2}N_i(T_0-t)^2\sigma_i^2 + \frac{1}{2}(u^2-u)B_{x_i}(T_1-t)^2 \\
& + \frac{1}{2}((1-u)^2-(1-u))B_{x_i}(T_0-t)^2 \\
& + u(1-u)B_{x_i}(T_1-t)B_{x_i}(T_0-t),
\end{aligned} \tag{B.16}$$

with boundary conditions $M(0) = 0$ and $N_i(0) = 0$, to satisfy $\Psi(u,T_0,T_0,T_1) = \exp(u\log(P(T_0,T_1))$. All the terms cancel and thus $\eta(t)$ is a martingale as required. Previously, the ODEs were in terms of $d\tau$, which is just the time reversed equivalent to the above result.

Semi-analytical price for $\mathcal{P}(t, T_0, T_1, K)$

Let us consider the numeraire as the bank account, $A_t = \exp\left(\int_0^t r_s ds\right)$, then

$$\frac{1}{A_t}\Psi(u,t,T_0,T_1) = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{1}{A_{T_0}}\exp\left(u\log(P(T_0,T_1))\right)\right]. \tag{B.17}$$

The Radon-Nikodym derivative for the $\mathbb{Q}^{T_0}$ measure is

$$\frac{d\mathbb{Q}^{T_0}}{d\mathbb{Q}} = \frac{P(T_0,T_0)/P(0,T_0)}{A_{T_0}/A_0}. \tag{B.18}$$

Therefore,

$$\begin{aligned}
\frac{1}{A_t}\Psi(u,t,T_0,T_1) &= \mathbb{E}_t^{\mathbb{Q}^{T_0}}\left[\frac{1}{A_{T_0}}\exp\left(u\log(P(T_0,T_1)\right)\frac{A_{T_0}/A_t}{P(T_0,T_0)/P(t,T_0)}\right] \\
\Psi(u,t,T_0,T_1) &= P(t,T_0)\,\mathbb{E}_t^{Q^{T_0}}\left[\exp\left(u\log(P(T_0,T_1)\right)\right] \\
\Rightarrow \Psi(iu,t,T_0,T_1) &= P(t,T_0)\,\varphi_{T_0}(u),
\end{aligned} \tag{B.19}$$

where $\varphi_{T_0}(u)$ is the characteristic function of $\log(P(t,T_0))$. A well-known result for the price of a put option on a zero-coupon bond with a change of numeriare is

$$\begin{aligned}
\mathcal{P}(t,T_0,T_1,K) &= \mathbb{E}_t^{\mathbb{Q}^{T_0}}\left[\exp\left(-\int_t^{T_0} r_s ds\right)(K - P(T_0,T_1))\mathbf{1}_{P(T_0,T_1)\leq K}\right] \\
&= KP(t,T_0)\mathbb{Q}^{T_0}(P(T_0,T_1)\leq K) - P(t,T_1)\mathbb{Q}^{T_1}(P(T_0,T_1)\leq K).
\end{aligned} \tag{B.20}$$

Applying the Gil-Pelaez theorem, where $k = \log(K)$ and using the fact that $\mathcal{R}e(z) = \mathcal{I}m(iu)$

$$\begin{aligned}
\mathbb{Q}^{T_0}(\log(P(T_0,T_1))\leq k) &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{R}e\left[\frac{\varphi_{T_0}(u)e^{-iuk}}{iu}\right]du \\
&= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{I}m\left[\frac{\Psi(iu,t,T_0,T_1)e^{-iuk}}{P(t,T_0)u}\right]du.
\end{aligned} \tag{B.21}$$

Considering $P(t,T_0)$, it can be seen that

$$P(t,T_0) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_0} r_s ds\right)\right] = \Psi(0,t,T_0,T_1). \tag{B.22}$$

Therefore,

$$\begin{aligned}
P(t,T_0)\mathbb{Q}^{T_0}(\log(P(T_0,T_1))\leq k) &= \frac{P(t,T_0)}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{I}m\left[\frac{\Psi(iu,t,T_0,T_1)e^{-iuk}}{u}\right]du \\
&= \frac{\Psi(0,t,T_0,T_1)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathcal{I}m\left[\Psi(iu,t,T_0,T_1)e^{-iuk}\right]}{u}du \\
&= G_{0,1}(t,T_0,T_1,k),
\end{aligned} \tag{B.23}$$

where $G_{a,b}(t,T_0,T_1,k)$ is defined by (2.40). Similarly, it can be shown that

$$P(t,T_1)\mathbb{Q}^{T_0}(\log(P(T_0,T_1)) \leq k) = G_{1,1}(t,T_0,T_1,k). \tag{B.24}$$

Consider the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}^{T_1}}{d\mathbb{Q}} = \frac{P(T_0,T_1)/P(t,T_1)}{A_{T_0}/A_t}, \tag{B.25}$$

so that

$$\begin{aligned}
\frac{1}{A_t}\Psi(u,t,T_0,T_1) &= \mathbb{E}_t^{\mathbb{Q}^{T_1}}\left[\frac{1}{A_{T_0}}\exp\left(u\log(P(T_0,T_1))\,\frac{A_{T_1}/A_t}{P(T_0,T_1)/P(t,T_1)}\right]\right. \\
\Psi(u,t,T_0,T_1) &= P(t,T_1)\,\mathbb{E}_t^{Q^{T_1}}\left[\exp\left((u-1)\log(P(T_0,T_1))\right]\right. \\
\Rightarrow \Psi(1+iu,t,T_0,T_1) &= P(t,T_1)\,\varphi_{T_1}(u),
\end{aligned} \tag{B.26}$$

where $\varphi_{T_1}(u)$ is the characteristic function of $\log(P(t,T_1))$. Again, by Gil-Pelaez

$$\begin{aligned}
\mathbb{Q}^{T_1}(\log(P(T_0,T_1)) \leq k) &= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{R}e\left[\frac{\varphi_{T_1}(u)e^{-iuk}}{iu}\right]du \\
&= \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{I}m\left[\frac{\Psi(1+iu,t,T_0,T_1)e^{-iuk}}{P(t,T_1)u}\right]du
\end{aligned} \tag{B.27}$$

Considering $P(t,T_1)$, it can be shown that

$$\begin{aligned}
P(t,T_1) &= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_1} r_s ds\right)\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_0} r_s ds\right)\exp\left(-\int_{T_0}^{T_1} r_s ds\right)\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_0} r_s ds\right)P(T_0,T_1)\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^{T_0} r_s ds\right)\exp\left(\log(P(T_0,T_1))\right)\right] \\
&= \Psi(1,t,T_0,T_1),
\end{aligned} \tag{B.28}$$

which implies that

$$\begin{aligned}
&P(t,T_1)\,\mathbb{Q}^{T_1}(\log(P(T_0,T_1)) \leq k) \\
&\qquad = \frac{P(t,T_1)}{2} - \frac{1}{\pi}\int_0^\infty \mathcal{I}m\left[\frac{\Psi(1+iu,t,T_0,T_1)e^{-iuk}}{u}\right]du \\
&\qquad = \frac{\Psi(1,t,T_0,T_1)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathcal{I}m\left[\Psi(1+iu,t,T_0,T_1)e^{-iuk}\right]}{u}du \\
&\qquad = G_{1,1}(t,T_0,T_1,k),
\end{aligned} \tag{B.29}$$

which completes (2.39).

Gauss-Legendre integral quadrature

Figure B.1 below illustrates the expression for the integral in equation (2.40) when $a = 0$ and $b = 1$ for varying levels of strike, K. This seeks to show both the oscillating nature of the expression as well as the rapid convergence seen for this integral to support the range of integration applied for the Gauss-Legendre quadrature used by Trolle and Schwartz (2009).

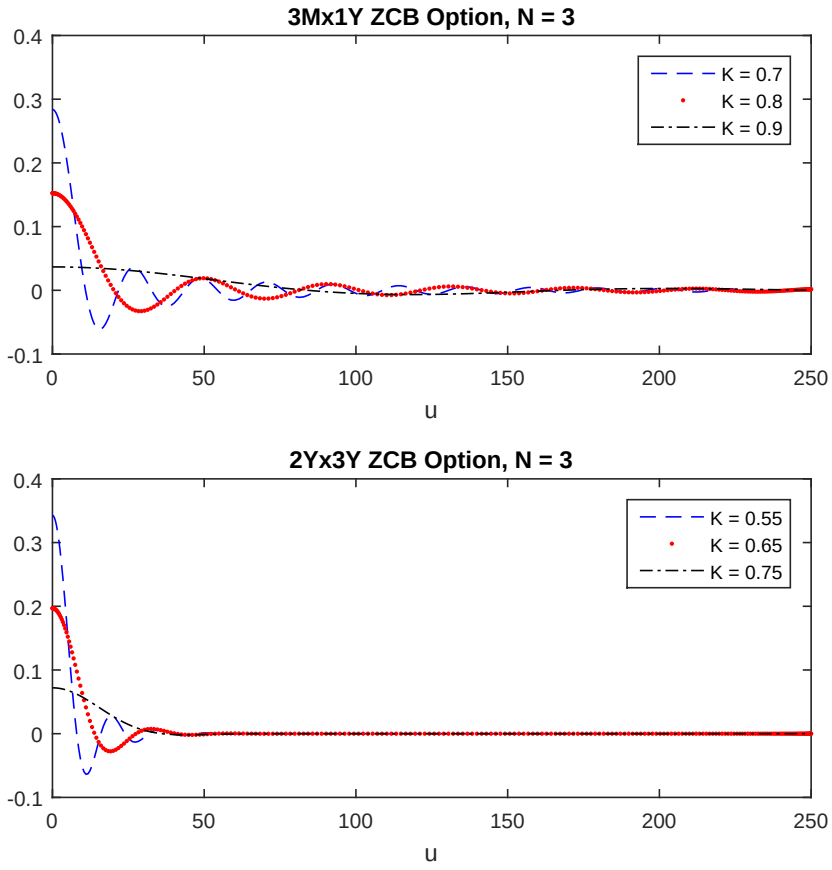

Fig. B.1: Integral for ZCB Option semi-analytical pricing formula.

It would seem that oscillation is more prevalent in options with shorter tenors and that the convergence is clear in both instances, supporting the quadrature approach described in Chapter 2.

Appendix C

Extended Kalman filter specifications

Expansion of elements for transition equation

The closed-form solutions for Φ_0, Φ_X and $Q_{v,i}$ are determined by the continuous-time dynamics of the state vector, X_t under $\mathbb{P}$. These are implied by (2.51), (2.52), and (2.29) - (2.34), which define the state vector dynamics, $dX_t = \mu_X(X_t)\,dt + \sigma_X(X_t)^\top\, dW_t^{\mathbb{P}}$. Since X_t has an affine structure, the dynamics in matrix form are

$$\mu_X(X_t) = b_0 + BX_t, \quad \sigma_x(X_t)^\top \sigma_x(X_t) = \sum_{i=1}^{N} G_i v_i(t), \tag{C.1}$$

where b_0 and B are expressed using the state vector dynamics under $\mathbb{P}$ so that

$$b_0 = \begin{pmatrix} \eta_1^P \\ \vdots \\ \eta_N^P \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \vdots \\ \kappa_1^P\theta_1^P \\ \vdots \\ \kappa_N^P\theta_N^P \end{pmatrix}, \quad B = \begin{bmatrix} \kappa_{x,1}^P & 0 & 0 & 0 & 0 & 0 & 0 & \kappa_{x,1}^P & \\ & \ddots & & & & & & & \ddots \\ 1 & -\gamma_1 & 0 & 0 & 0 & 0 & 0 & 0 & \\ & \ddots & \ddots & & & & & & \\ 0 & 0 & -\gamma_1 & 0 & 0 & 0 & 0 & 1 & \\ & & & \ddots & & & & & \ddots \\ 0 & 0 & 0 & -2\gamma_1 & 0 & 0 & 0 & 1 & \\ & & & & \ddots & & & & \ddots \\ 0 & 0 & 1 & 0 & -\gamma_1 & 0 & 0 & 0 & \\ & & & \ddots & & \ddots & & & \\ 0 & 0 & 0 & 1 & 0 & -2\gamma_1 & 0 & 0 & \\ & & & & \ddots & & \ddots & & \\ 0 & 0 & 0 & 0 & 0 & 2 & -2\gamma_1 & 0 & \\ & & & & & & \ddots & \ddots & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\kappa_1^P & \\ & & & & & & & & \ddots \end{bmatrix}. \tag{C.2}$$

The expression for G_i for $i = 1, .., N$ then takes the form of

$$G_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_i\rho_i & \\ & \ddots & & & & & & & \ddots \\ & & & - & & & & & \\ & & & - & & & & & \\ & & & - & & & & & \\ & & & - & & & & & \\ & & & - & & & & & \\ & & & - & & & & & \\ \sigma_i\rho_i & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_i^2 & \\ & \ddots & & & & & & & \ddots \end{bmatrix}. \tag{C.3}$$

Hence, b_0 is an $[(N \times 8) \times 1]$ column vector, while B and Gi are $[(N \times 8) \times (N \times 8)]$ square matrices. Then define

$$\phi(t) = \exp\{B\, t\}, \tag{C.4}$$

$$\mathcal{D}(t) = \int_0^t \phi(s)ds, \tag{C.5}$$

to fully specify the closed-form expressions for Φ_0 and Φ_X as

$$\Phi_0 = \mathcal{D}(t_{i+1} - t_i)b_0, \tag{C.6}$$

$$\Phi_X = \phi(t_{i+1} - t_i). \tag{C.7}$$

Then, $\Phi(X_{t_i})$ is equal to the expected value of $X_{t_{i+1}}$ given X_{t_i}

$$\begin{aligned} \Phi(X_{t_i}) &= \mathbb{E}_t^{\mathcal{P}}[X_{t+1}] \\ &= \hat{X}(t_i, t_{i+1}) \\ &= \Phi_0 + \Phi_X X_{t_i} \\ &= \mathcal{D}(t_{i+1} - t_i)b_0 + \phi(t_{i+1} - t_i)X_{t_i}. \end{aligned} \tag{C.8}$$

The solution for the variance, $Q(X_t)$, is then also fully specified as

$$\begin{aligned} Q(X_t) &= \sum_{i=1}^{N} Q_{v,i} v_i(t) \\ &= \int_{t_i}^{t_{i+1}} \phi(t_{i+1} - s) \left(\sum_{i=1}^{N} G_i \hat{v}_i(t_i, s) \right) \phi(t_{i+1} - s)^\top ds, \end{aligned} \tag{C.9}$$

where $\hat{v}_i(t_i, s)$ are the elements of $\hat{X}(t_i, s)$ pertaining to $v_i(t)$.

Expression for H_t required in linearised measurement equation for EKF

Consider the case where the function h in the measurement equation is non-linear. Linearising h results in the linearised measurement equation

$$y_t = H_t' X_t + (h(\hat{X}_{t|t-1}) - H_t' \hat{X}_{t|t-1}) + u_t, \qquad u_t \sim \mathcal{N}(0, \mathcal{S}), \tag{C.10}$$

which requires the use of the EKF. Here, the Jacobian of h is specified as H_t where

$$H_t = \frac{\delta h(X_t)'}{\delta X_t}\bigg|_{X_t=\hat{X}_{t|t-1}}.$$

As before, $\hat{X}_{t|t-1} = \mathbb{E}^{\mathbb{P}}_{t-1}[X_t]$ is the prediction of X_t excluding y_t. If M is the length of y_t and $h(X_t)$, then H_t is a $[M \times (N \times 8)]$ two-dimensional matrix.

The partial derivatives for the interest rate pricing formulas (2.41) and (2.42) have closed-form expressions with respect to the term structure state variables $x_i(t), \phi_{1,i}(t), \ldots, \phi_{6,i}(t)$. Both pricing formulas remain independent of each $v_i(t)$. The elements of the Jacobian, H_t for the pricing formulas for time-t LIBOR rates are

$$\frac{\delta L(t,T_1)}{\delta x_i(t)} = -\frac{B_{x_i}(T_1-t)}{(T_1-t)P(t,T_1)}, \tag{C.11}$$

$$\frac{\delta L(t,T_1)}{\delta \phi_{j,i}(t)} = -\frac{B_{\phi_{j,i}}(T_1-t)}{(T_1-t)P(t,T_1)}, \tag{C.12}$$

and for the pricing formulas for time-t swap rates are

$$\frac{\delta S(t,T_n)}{\delta x_i(t)} = \frac{P(t,T_n)(P_{B_{x_i}}(t,T_n) - B_{x_i}(T_n-t)P^s(t,T_n)) - P_{B_{x_i}}(t,T_n)}{\upsilon(P^s(t,T_n))^2}, \tag{C.13}$$

$$\frac{\delta S(t,T_n)}{\delta \phi_{k,i}(t)} = \frac{P(t,T_n)(P_{B_{\phi_{k,i}}}(t,T_n) - B_{\phi_{k,i}(t)}(T_n-t)P^s(t,T_n)) - P_{B_{\phi_{k,i}}}(t,T_n)}{\upsilon(P^s(t,T_n))^2}, \tag{C.14}$$

where $P^s(t,T_n) = \sum_{j=2}^{n} P(t,T_j)$, $P_{B_{x_i}}(t,T_n) = \sum_{j=2}^{n} P(t,T_j)B_{x_i}(T_j-t)$ and $P_{B_{\phi_{k,i}}}(t,T_n) = \sum_{j=2}^{n} P(t,T_j)B_{\phi_{k,i}}(T_j-t)$.

Examining the pricing formulas (2.46) and (2.50) reveals no closed-form partial derivatives, thus the expression must be computed numerically. An important note is that derivatives are priced using the fitted zero-coupon bond curve and are thus independent of $x_i(t), \phi_{1,i}(t), ..., \phi_{6,i}(t)$. A forward differencing scheme was chosen for small perturbations of each $v_i(t)$. The elements of the Jacobian H_t for the caps and swaptions pricing formulas are simply, for $\Delta v = 0.001$,

$$\frac{\delta \,\text{Cap}(t,T_n,K,v_i(t))}{\delta v_i(t)} \approx \Big(\text{Cap}(t,T_n,K,v_i(t)+\Delta v) - \text{Cap}(t,T_n,K,v_i(t))\Big)/\,\Delta v, \tag{C.15}$$

$$\frac{\delta \,\text{Swpn}(t,T_m,T_n,K,v_i(t))}{\delta v_i(t)} \approx \Big(\text{Swpn}(t,T_m,T_n,K,v_i(t)+\Delta v) - \text{Swpn}(t,T_m,T_n,K,v_i(t))\Big)/\,\Delta v. \tag{C.16}$$

Estimated state vectors for simulated and empirical data

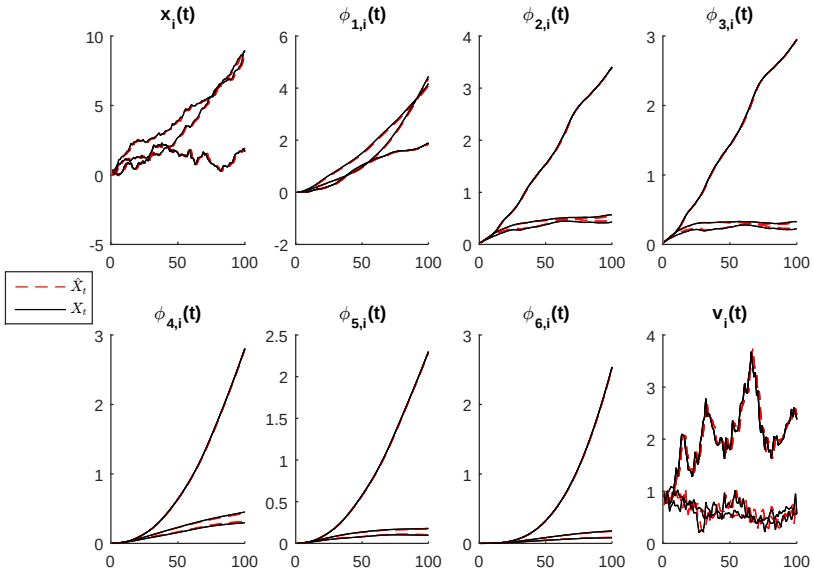

Fig. C.1: Simulated (X_t) and estimated ($\hat{X}_t$) state vector, simulated dataset, $N = 3$

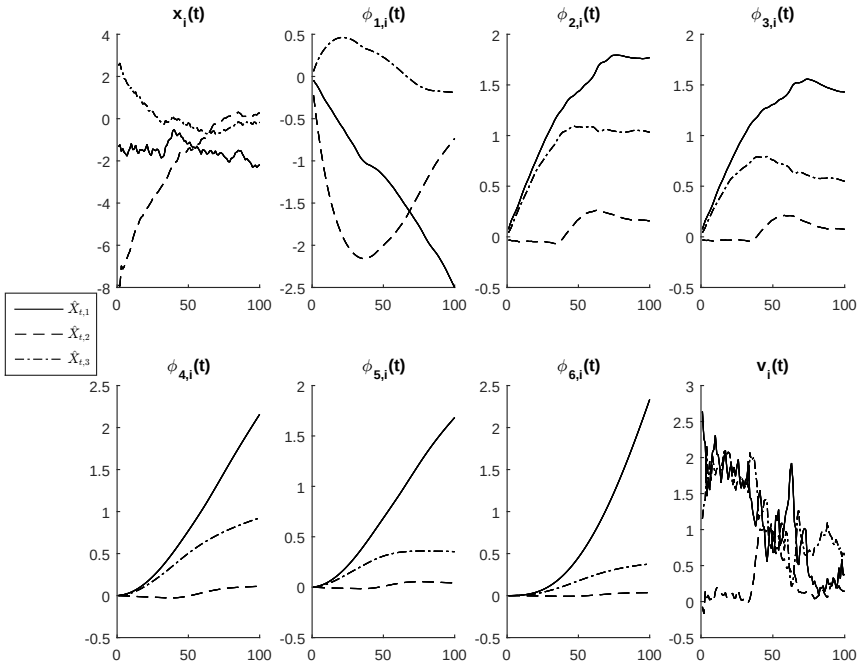

Fig. C.2: Estimated ($\hat{X}_t$) state vector, empirical dataset for $i = 1, 2, 3$

Appendix D

Monte Carlo comparative prices

This shows the full set of results used to compare the Monte Carlo simulation to the semi-analytical and volatility approximation prices as seen in Chapter 4. These prices and error bounds are a result of using 100,000 sample paths with 1M time steps for $N = 1$ and 2W time steps for $N = 3$ in Monte Carlo simulations using the Euler, Milstein and QE schemes.

	Simulated		Historical	
	$N = 1$	$N = 3$	$N = 1$	$N = 3$
Bonds				
$P(t, 1)$	0.638715	0.717962	0.968483	0.961509
	[0.638693; 0.638736]	[0.717938; 0.717986]	[0.968449; 0.968517]	[0.961493; 0.961525]
$P(t, 5)$	0.085969	0.313879	0.803634	0.809922
	[0.085908; 0.08603]	[0.313666; 0.314092]	[0.803046; 0.804222]	[0.809498; 0.810347]
$P(t, 10)$	0.019421	0.192935	0.568447	0.642177
	[0.019384; 0.019457]	[0.192578; 0.193291]	[0.567346; 0.569548]	[0.641277; 0.643077]
Caps				
$1Y$	0.006423	0.002485	0.001566	0.001160
	[0.006403; 0.006444]	[0.002464; 0.002505]	[0.001544; 0.001588]	[0.001145; 0.001176]
$5Y$	0.028357	0.092566	0.034300	0.024453
	[0.028219; 0.028495]	[0.09244; 0.092693]	[0.03387; 0.034731]	[0.024106; 0.024800]
$10Y$	0.049254	0.170801	0.088262	0.063042
	[0.049095; 0.049413]	[0.170598; 0.171003]	[0.087239; 0.089285]	[0.062173; 0.063910]
Swaption				
$3M \times 1Y$	0.001288	0.001404	0.001763	0.001132
	[0.001269; 0.001306]	[0.001384; 0.001424]	[0.001738; 0.001787]	[0.001116; 0.001148]
$2Y \times 3Y$	0.004516	0.008871	0.019677	0.013548
	[0.004451; 0.00458]	[0.008747; 0.008996]	[0.019398; 0.019956]	[0.013355; 0.013741]
$5Y \times 5Y$	0.002346	0.012590	0.033954	0.027154
	[0.002314; 0.002379]	[0.012421; 0.012759]	[0.033494; 0.034414]	[0.026774; 0.027534]

Tab. D.1: Full Monte Carlo pricing results for Euler scheme

	Simulated		Historical	
	$N = 1$	$N = 3$	$N = 1$	$N = 3$
Bonds				
$P(t, 1)$	0.638714	0.717957	0.968485	0.961509
	[0.638693; 0.638736]	[0.717933; 0.717981]	[0.968451; 0.968519]	[0.96149; 0.961528]
$P(t, 5)$	0.085953	0.313906	0.803653	0.809958
	[0.085892; 0.086013]	[0.313692; 0.314119]	[0.803067; 0.804239]	[0.80953; 0.810386]
$P(t, 10)$	0.019415	0.192914	0.568457	0.642015
	[0.019379; 0.019452]	[0.192561; 0.193268]	[0.567359; 0.569555]	[0.641113; 0.642917]
Caps				
$1Y$	0.006424	0.002482	0.001561	0.00116
	[0.006404; 0.006444]	[0.002461; 0.002503]	[0.001539; 0.001583]	[0.001144; 0.001176]
$5Y$	0.028369	0.092569	0.034311	0.024511
	[0.02823; 0.028508]	[0.092443; 0.092694]	[0.033881; 0.03474]	[0.024163; 0.02486]
$10Y$	0.049222	0.170789	0.088243	0.062798
	[0.049064; 0.049381]	[0.170586; 0.170992]	[0.087221; 0.089265]	[0.061929; 0.063667]
Swaption				
$3M \times 1Y$	0.001298	0.0014	0.001757	0.001123
	[0.00128; 0.001316]	[0.00138; 0.00142]	[0.001732; 0.001782]	[0.001107; 0.001139]
$2Y \times 3Y$	0.004511	0.008909	0.019746	0.013505
	[0.004447; 0.004576]	[0.008784; 0.009034]	[0.019466; 0.020027]	[0.013312; 0.013698]
$5Y \times 5Y$	0.00235	0.012545	0.034016	0.027076
	[0.002318; 0.002382]	[0.012376; 0.012714]	[0.033554; 0.034478]	[0.026694; 0.027458]

Tab. D.2: Full Monte Carlo pricing results for Milstein scheme

	Simulated		Historical	
	$N = 1$	$N = 3$	$N = 1$	$N = 3$
Bonds				
$P(t, 1)$	0.638715	0.717962	0.968483	0.961512
	[0.638693; 0.638737]	[0.717938; 0.717986]	[0.968449; 0.968517]	[0.961493; 0.961531]
$P(t, 5)$	0.085962	0.313884	0.803666	0.809919
	[0.085901; 0.086023]	[0.313671; 0.314097]	[0.803082; 0.80425]	[0.809495; 0.810343]
$P(t, 10)$	0.01942	0.192944	0.568574	0.6417
	[0.019384; 0.019457]	[0.192584; 0.193303]	[0.567474; 0.569675]	[0.640799; 0.642601]
Caps				
$1Y$	0.006426	0.002482	0.00157	0.001161
	[0.006406; 0.006447]	[0.002462 ;0.002503]	[0.001548; 0.001592]	[0.001145; 0.001176]
$5Y$	0.02837	0.092582	0.034642	0.024555
	[0.028231; 0.028508]	[0.092456; 0.092708]	[0.034216; 0.035068]	[0.024211; 0.024899]
$10Y$	0.04924	0.170791	0.089501	0.063445
	[0.049082; 0.049399]	[0.17059; 0.170993]	[0.088486; 0.090515]	[0.062581; 0.064308]
Swaption				
$3M \times 1Y$	0.001292	0.001409	0.001767	0.001127
	[0.001274; 0.00131]	[0.001389; 0.001429]	[0.001742; 0.001791]	[0.001111; 0.001143]
$2Y \times 3Y$	0.004516	0.008942	0.019771	0.013602
	[0.004453; 0.00458]	[0.008819; 0.009066]	[0.019495; 0.020048]	[0.01341; 0.013793]
$5Y \times 5Y$	0.002373	0.012687	0.034441	0.027483
	[0.002341; 0.002405]	[0.012519; 0.012854]	[0.033984; 0.034899]	[0.027105; 0.027861]

Tab. D.3: Full Monte Carlo pricing results for QE scheme