

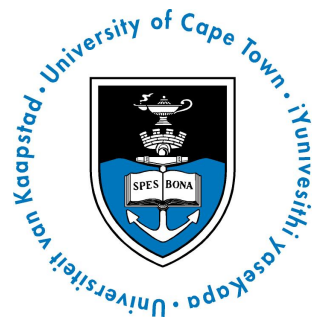
# Linear-Rational Term Structure Models With Flexible Level-Dependent Volatility

Adrian Schwellnus

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

July 22, 2018

*MPhil in Mathematical Finance,  
University of Cape Town.*



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

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

# Declaration

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

---

Adrian Schwellnus

July 22, 2018

# Abstract

The Linear-Rational Framework for the modelling of interest rates is a framework which allows for the addition of spanned and unspanned factors, while maintaining a lower bound on rates and tractable valuation of interest rate derivatives, particularly swaptions. The advantages of having all these properties are significant. This dissertation presents the Linear-Rational Framework, and specializes the factor process to a class of diffusion models which allows for the degree of state dependence of volatility to be estimated. This dissertation then finds that the estimated state dependent volatility structure is significantly different to that of typical models, where it is set *a priori*. The effect the added degree of freedom has on the model implied swaption skew is then analysed.

# Acknowledgements

I am deeply grateful to my supervisor, Mr Alex Backwell, for not only setting such an engaging dissertation topic which has allowed me to learn so much, but for also being a constant source of guidance and assistance. I would also like extend a sincere thank you to the anonymous individual who was kind enough to provide data for use in this dissertation. Lastly, I would like to thank my family for being a constant source of encouragement throughout my university years.

# Contents

<b>1. Introduction</b>	1
<b>2. Linear-Rational Term Structure Models</b>	3
2.1 Framework Overview	3
2.2 Framework Specification	4
2.3 Unspanned Factors	5
2.4 CEV Process	7
<b>3. Parameter Estimation</b>	9
3.1 Unscented Kalman Filter	9
3.1.1 Algorithm	10
3.2 Parameter Recovery	12
3.3 Empirical Results	13
3.3.1 Discussion	14
<b>4. Swaptions</b>	20
4.1 Swaption Valuation	20
4.2 Swaption IV Skew	23
<b>5. Conclusions</b>	26
<b>Bibliography</b>	27
<b>A. Chapter 2</b>	28
A.1 Radon-Nikodým Process	28
A.2 Derivations for Section 2.1	28
A.2.1 Conditional Expectation of the Factor Process	28
A.2.2 ZCB Value Formula	29
A.2.3 Short Rate	29
<b>B. Chapter 4</b>	30

# List of Figures

3.1	Simulated vs Filtered Factor Processes . . . . .	13
3.2	Filtered Factor Processes - CEV Parameters . . . . .	15
3.3	Filtered Swap Rate Error - CEV Parameters . . . . .	16
3.4	Time Series of Data . . . . .	17
3.5	$X^\alpha$ for CEV and Sqrt Process . . . . .	18
3.6	$R_{\text{fair}}$ vs Factor Values . . . . .	19
4.1	Crude Monte-Carlo vs Fourier Value . . . . .	23
4.2	CEV vs Sqrt Black 76 Implied Vol for Multiple Strikes . . . . .	24

# List of Tables

3.1	Parameter Recovery . . . . .	12
3.2	Parameter Estimates . . . . .	14



## Chapter 1

# Introduction

The Linear-Rational Framework for the modelling of interest rates, developed by [Filipović, Larsson and Trolle \(2017\)](#), is a particularly powerful class of term structure models. The framework's ability to respect a lower bound on rates is of major importance, given the low interest rate environment currently being experienced in certain markets. The framework allows for the addition of multiple spanned factors, as well as multiple unspanned factors (factors which can not be hedged using zero-coupon bonds). Despite this significant freedom, the exact valuation of interest rate derivatives remains tractable. In particular, under certain conditions, exact pseudo-analytical solutions for swaption values exist. This can be contrasted with the approximations that are necessary in the well known Affine Term Structure Models, such as in [Schrager and Pelsser \(2006\)](#).

A significant portion of [Filipović \*et al.\* \(2017\)](#) is a specification, and empirical analysis, for the factor process following square-root, CIR-type, dynamics. In this specification the volatility of the factor process is imposed to be a scaled square root of the position of the current state. That is the degree of level-dependence of volatility, which is how the volatility of the factor process changes as its position changes, is set *a priori*. This dissertation adds a degree of freedom to the parameter set, which allows for estimation of the degree of level-dependence of volatility. The specification used, which allows for the extra degree of freedom, is one where the factor process follows CEV-type dynamics.

An empirical study is then undertaken to estimate model parameters by attempting to directly reprice daily ZAR swap rates, over the period 27/09/2012-30/03/2015. An Unscented Kalman Filter is used during this procedure. Special reference is made to the estimation of drift parameters. In particular, the estimated drift parameters of the factor process represent the most likely cross-sectionally fitting parameters, and not necessarily to match the observed time series factor process drift dynamics. A significant finding within the empirical study is that when allowing for the extra degree of freedom in the CEV-type specification outlined

above, then the degree of level-dependence of volatility is substantially different to that of if it was set *a priori*, using a square-root factor process. Specifically, the square-root factor process underestimates the contribution of the level-dependence on the volatility.

The degree of level-dependence of volatility is connected to the skew in swaptions, and this relationship is then examined. In particular, it is confirmed that when using the CEV parameters to value swaptions it results in a steeper skew, when compared to the square-root process parameters.

The structure of the dissertation is as follows. Chapter 2 introduces the Linear-Rational Framework, and shows some important results thereof, such as the zero-coupon bond valuation formula. It then briefly introduces the idea of unspanned factors, despite the fact that the topic is not developed beyond this section, as they are an important feature of Linear-Rational Term Structure Models. The CEV specification for the factor process, which is widely used within this dissertation, is then introduced. The empirical study is undertaken in Chapter 3, where model parameters are estimated. Chapter 4 begins with the swaption valuation methodology, followed by presenting the model implied swaption implied volatility skew.

## Chapter 2

# Linear-Rational Term Structure Models

This section aims to give a brief and general overview of Linear-Rational Term Structure Models, as well as motivate some of the advantages of these models. It then specializes the framework to a class of diffusion models, which will be studied in greater detail during the course of the dissertation.

### 2.1 Framework Overview

The fundamental theorem of asset pricing ([Harrison and Pliska, 1981](#)) states that the absence of arbitrage is equivalent to the existence of a risk-neutral measure. Following the approach taken in [Constantinides \(1992\)](#), the existence of a risk-neutral measure is equivalent to the existence of a positive adapted process  $\{\zeta_t\}$ , known as the state price density<sup>1</sup>, such that the time- $t$  value  $V(t, T)$ , of a time- $T$  cash flow  $C_T$ , is given by

$$V(t, T) = \frac{1}{\zeta_t} \mathbb{E}_t[\zeta_T C_T]. \quad (2.1)$$

This is assuming a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ , where the expectation is an  $\mathcal{F}_t$ -conditional expectation under  $\mathbb{P}$ , which is considered the real-world measure. It is worth emphasising that the state price density has both a value deflating *and* risk adjustment role. This is due to the fact that  $\{\zeta_t\}$  involves a Radon-Nikodým process, and as a result of this the value expression in Equation (2.1) does not require explicit use of a risk-neutral,  $\mathbb{Q}$ , measure. This discussion is developed slightly further in Appendix A.1.

---

<sup>1</sup> Otherwise known as a stochastic discount factor, or pricing kernel. For further background on state price density asset pricing, the reader is referred to [Flesaker and Hughston \(1996\)](#), [Rogers \(1997\)](#) and [Macrina \(2014\)](#).

The approach taken in [Filipović \*et al.\* \(2017\)](#) is to specify the state price density as a linear function of a multivariate factor process. This results in bond values, and the short rate, becoming linear-rational functions (that is, ratios of linear functions) of the current state. Therefore it has been conveniently named the “Linear-Rational” framework. Linear-Rational Term Structure Models are immediately arbitrage free by specification, which is due to the if and only if nature of the fundamental theorem of asset pricing. That is, if we specify a form for the state price density, then the model is immediately arbitrage free.

## 2.2 Framework Specification

In this framework it is specified that the multivariate factor process,  $\{X_t\}$ , whose state space is some subset  $E \subset \mathbb{R}^d$ , has the dynamics

$$dX_t = \kappa(\theta - X_t)dt + dM_t, \quad (2.2)$$

where  $\kappa \in \mathbb{R}^{d \times d}$ ,  $\theta \in \mathbb{R}^d$ , and  $\{M_t\}$  is a  $d$ -dimensional martingale. It is clear that drift term of this process results in a mean reverting interpretation. The state price density is specified to be

$$\zeta_t = e^{-\beta t}(\phi + \psi^\top X_t), \quad (2.3)$$

for some  $\theta \in \mathbb{R}$ ,  $\psi \in \mathbb{R}^d$ , such that  $\phi + \psi^\top x > 0$  for all  $x \in E$ , and for some  $\beta \in \mathbb{R}$ . The linearity of the drift in Equation (2.2) allows for the time- $t$  conditional expectation of  $X_T$  to be solved, and the calculation in the one factor case is explicitly shown in Appendix A.2.1. One gets

$$\mathbb{E}_t[X_T] = \theta + e^{-\kappa(T-t)}(X_t - \theta), \quad (2.4)$$

for  $t \leq T$ . In order to find the zero-coupon bond (ZCB) values under this framework, one sets the time- $T$  cash flow  $C_T = 1$  in the fundamental valuation Equation (2.1). The ZCB values are given by  $P(t, T) = F(T - t, X_t)$ , as is explicitly shown in Appendix A.2.2, where  $F(\tau, x)$  is

$$F(\tau, x) = e^{-\beta\tau} \frac{\phi + \psi^\top \theta + \psi^\top e^{-\tau\kappa}(x - \theta)}{\phi + \psi^\top x}. \quad (2.5)$$

The short rate is then obtained through the formula:  $r_t = -\frac{\partial}{\partial T} \log P(t, T)|_{T=t}$ , and the result is shown in Appendix A.2.3 to be

$$r_t = \beta - \frac{\psi^\top \kappa(\theta - X_t)}{\phi + \psi^\top X_t}. \quad (2.6)$$

This gives a good intuition for the role of the  $\beta$  parameter. If  $\frac{\psi^\top \kappa(\theta - X_t)}{\phi + \psi^\top X_t}$  is bounded on the state space  $E$ , then a large enough  $\beta$  can be chosen to ensure a chosen lower bound on the short rate is preserved. A natural choice is the smallest  $\beta$  that results in a non-negative short rate. Defining

$$\beta^* = \sup_{x \in E} \frac{\psi^\top \kappa(\theta - X_t)}{\phi + \psi^\top X_t},$$

$$\beta_* = \inf_{x \in E} \frac{\psi^\top \kappa(\theta - X_t)}{\phi + \psi^\top X_t},$$

it is natural to set  $\beta = \beta^*$ . It then follows that the short rate satisfies  $r_t \in [0, \beta^* - \beta_*]$ .

## 2.3 Unspanned Factors

The idea of unspanned factors is based on the fact that there may be directions along which the factor process,  $X_t$ , can move that results in no changes to ZCB values, but do effect the values of interest rate derivatives. The existence of such directions is shown in [Collin-Dufresne and Goldstein \(2002\)](#). These directions are known as the “unspanned directions” of the factor process. While this dissertation will not develop the idea of unspanned factors any further beyond this section, it is worth including a discussion on them, given that the ease at which unspanned factors are added being a major benefit of the framework. To introduce these factors, the definition of the kernel of a function is used:  $\ker f(x) = \{\zeta \in \mathbb{R}^d : \nabla f(x)^\top \zeta = 0\}$ . Then the *term structure kernel*, denoted by  $\mathcal{U}$ , is given by

$$\mathcal{U} = \bigcap_{\tau \geq 0} \ker F(\tau, x).$$

Intuitively,  $\mathcal{U}$  is the space of all directions of  $\zeta$  along which the ZCB value formula,  $F(T - t, X_t)$ , does not change for all time. On those directions along which  $F(T - t, X_t)$  does not change, it is clear that the location of  $X_t$  cannot be recovered just by knowledge of the value of  $F(T - t, X_t)$ . In other words, knowledge of the term structure doesn't given sufficient information about the location of  $X_t$  along directions in the kernel,  $\mathcal{U}$ . This leads to the statement that the term structure kernel is unspanned by the term structure. A major implication of the fact that ZCB values do not change when moving along unspanned directions, is that movements along unspanned directions can not be hedged by ZCBs. This implies that interest rate derivatives are not redundant securities, as they could be required to hedge the unspanned directions.

As will soon be evident, it is helpful to transform the state space so that the unspanned directions correspond to the last components of the state vector. Let

$S$  be an invertible matrix on  $\mathbb{R}^d$ . The transformed factor process  $\hat{X}_t = SX_t$  then satisfies

$$d\hat{X}_t = \hat{\kappa}(\hat{\theta} - \hat{X}_t)dt + d\hat{M}_t,$$

where

$$\hat{\kappa} = S\kappa S^{-1}, \hat{\theta} = S\theta, \hat{M} = SM_t.$$

Defining

$$\hat{\phi} = \phi, \hat{\psi} = S^{-\top}\psi,$$

then the state price density given in Equation (2.3) can be written in terms of the transformed factor process

$$\zeta_t = e^{-\beta t}(\hat{\phi} + \hat{\psi}^\top \hat{X}_t).$$

The idea is that the transformation,  $S$ , maps the term structure kernel, which has dimension  $n$ , onto the last  $n$  components of the transformed state space. Letting  $\mathcal{U}$  be the directions along which the term structure remains unchanged (the term structure kernel), then

$$S(\mathcal{U}) = \{0\} \times \mathbb{R}^n, \quad (2.7)$$

where as previously mentioned  $n = \dim \mathcal{U}$ , which implies that  $m = d - n$  is the dimension of the spanned directions. This decomposition leads to the result that the factor process becomes  $\hat{X}_t = (Z_t, U_t)$ , where the  $Z_t$  factors effect the term structure and is  $m$ -dimensional, while the  $U_t$  factors do not and is  $n$ -dimensional.

Assuming that the technical conditions required for Equation (2.7) hold, and writing  $Sx = (z, u)$ , the ZCB value formula becomes

$$\hat{F}(\tau, z) = e^{-\beta\tau} \frac{\hat{\phi} + \hat{\psi}_z^\top \hat{\theta}_z + \hat{\psi}_z^\top e^{-\tau \hat{\kappa}_{zz}} (z - \hat{\theta}_z)}{\hat{\phi} + \hat{\psi}_z^\top z}, \quad (2.8)$$

where  $\hat{\kappa}_{zz}$ , is the  $m \times m$  upper left block of the transformed  $\kappa$  parameter,  $\hat{\theta}_z$  is the first  $m$  elements of  $\hat{\theta}$ , and similarly for  $\hat{\psi}_z$ . The new formula has no effect on the value of ZCBs and  $\hat{F}(\tau, z) = F(\tau, x)$ , as is required for consistency.

Bond values are now given by  $P(t, T) = \hat{F}(T - t, Z_t)$ , which clearly gives the interpretation that the  $U_t$  components are the unspanned factors, as the  $U_t$  term plays no part in the value of a ZCB. That is, the  $U_t$  components of the factor process can not be hedged using ZCBs, and a snapshot of the term structure at any time does not give information about the factors  $U_t$ . However, the  $Z_t$  factors directly effect the term structure, and can be retrieved from knowledge of the term structure, provided that a few technical conditions hold. That is, the  $Z_t$  components of the factor process can be hedged using ZCBs. The  $Z_t$  factors are then referred to as term structure factors.

## 2.4 CEV Process

This dissertation specialises the Linear-Rational framework to a class of diffusion models. In particular, it will examine the case that the factor process, Equation (2.2), follows constant elasticity of variance (CEV) type dynamics, which has a state space  $E = \mathbb{R}_+^d$ , and is of the form (Cox, 1975)

$$dX_t = \kappa(\theta - X_t)dt + \text{Diag}(\sigma_1 X_{1t}^\alpha, \dots, \sigma_d X_{dt}^\alpha)dB_t, \quad (2.9)$$

with  $\kappa, \theta, \sigma > 0$ . In the case that  $0 < \alpha < 1$ , then the factor process will have increasing volatility as its value decreases. If  $\alpha > 1$  then the factor process will have increasing volatility as its value increases, and the local martingale component of the factor process,  $\sigma_i X_{it}^\alpha dB_t$ , will not be a true martingale (Mijatović and Urusov, 2012). This results in the conditional expectation and bond pricing equations; Equations (2.4) and (2.5), not holding. For this reason, only values of  $\alpha \leq 1$  are examined.

This process is a generalisation of the Linear-Rational Square Root (LRSQ, or just sqrt) process introduced, and empirically examined, in Filipović *et al.* (2017). An LRSQ process can be recovered by setting  $\alpha = \frac{1}{2}$ , which will occasionally be done for comparison, and to motivate the tractability of swaption valuation in Linear-Rational Term Structure Models. It is worth emphasising that introducing a process such as this is only possible because of the generality of the framework where, as is stipulated in Equation (2.2), it is possible to use any martingale in the factor process, while still making use of the results from Section 2.2.

A reflective boundary at  $X_t = 0$  is used throughout this dissertation. The reason for this is twofold. The first, and most important, is that as is outlined and developed in Andersen and Piterbarg (2007), this guarantees uniqueness of the solution to this SDE for  $0 < \alpha < \frac{1}{2}$ , as zero is an attainable boundary in this case. The second reason is to guarantee that no numerical errors from the discretization of Equation (2.9) results in negative values for the process. If  $\alpha = \frac{1}{2}$  then the Yamada conditions guarantee uniqueness of the solution to the SDE, if the well known Feller conditions do not hold, as zero is an attainable boundary in this case. If  $\alpha > \frac{1}{2}$  then zero is an unattainable boundary (Andersen and Piterbarg, 2007).

In our specification  $\kappa$  is a diagonal matrix by construction. This guarantees uniqueness in the multi-dimensional specification, as each dimension is examined individually as a single dimensional SDE. As the state space admits a canonical representation (Filipović, 2009), we set  $\phi = 1$  and  $\psi = 1$ . Intuitively, this removes extra degrees of freedom in the parameter set of the model. Recall from Equation (2.6),  $\beta$  is a parameter which is chosen to be large enough to guarantee non-negative

---

rates. With the above specification, this value is precisely  $\beta = \mathbf{1}^\top \kappa \theta$ . Estimating the remaining model parameters is attempted using a Kalman Filter.



## Chapter 3

# Parameter Estimation

In order to estimate parameters, we attempt to directly reprice market observed swap rates. The advantage of using swap rates directly is that no bootstrapping scheme is required to perform the parameter estimation, as would be the case if one uses ZCBs.

From the perspective of the fixed-rate payer, a vanilla, uncollateralized<sup>1</sup>, fixed-for-floating swap, with payment dates:  $T_0 < \dots < T_n$ , constant  $\tau = T_i - T_{i-1}$ , and a fixed rate of  $R$ , has value at time  $t \leq T_0$  being the difference between the floating and fixed legs;

$$V_t^{swap} = P(t, T_0) - P(t, T_n) - \tau R \sum_{i=1}^n P(t, T_i). \quad (3.1)$$

The fixed rate,  $R$ , is typically chosen fairly to give the swap an initial value of zero, in which case

$$R_{\text{fair}} = \frac{P(t, T_0) - P(t, T_n)}{\tau \sum_{i=1}^n P(t, T_i)}. \quad (3.2)$$

Given the non-linear relationship between the current state  $X_t$ , and the fair swap rate,  $R_{\text{fair}}$ , an Unscented Kalman Filter is used to estimate parameters.

### 3.1 Unscented Kalman Filter

A Kalman filter is a tool which is used to estimate, and give a distribution for, noisily measured variables. This is achieved by predicting the next state of a variable, the so called “hidden state”, based on information in the previous state, and

---

<sup>1</sup> In practice, swaps are typically collateralized. This requires the valuation of a swap to take the LIBOR-OIS spread into account. Extending the Linear-Rational Framework to do so is discussed in [Filipović \*et al.\* \(2017\)](#), and they do achieve a mildly better fit to data when using the extended framework. It is however stated that simulations show the two model specifications have a very similar behaviour for IRS rates and swaption IVs. This implies that their results, which assume uncollateralized trades, are valid.

then updating the estimate based on observed information, the so called “measurements”. The algorithm is then applied recursively.

An Unscented Kalman Filter is a generalization of the classical Kalman Filter, which can be used when the prediction and measurement functions are non-linear. It makes use of an Unscented Transform, which approximates Gaussian random variables by advancing so called “sigma points” through the non-linear function. The Unscented Kalman Filter approximates means and covariances as a linear combination of the sigma points of the Unscented Transform. In the context of this dissertation, the measurements are swap rates, and the hidden state is the factor process. In order to make a prediction of the factor process, an Euler-Maruyama discretization of Equation (2.9) is used. The measurement function is a function which takes as input the state,  $X_t$ , and the output is a swap rate. For a given set of model parameters, a likelihood function value can be computed. The aim is then to discover a set of model parameters which maximize this likelihood.

### 3.1.1 Algorithm

We now present the Non-linear Unscented Kalman Filter Algorithm (Särkkä, 2010). In the algorithm,  $L$  is the dimension of the hidden state vector (the number of factors),  $\gamma$  and  $\epsilon$  reflect the spread of the sigma points, and  $\chi$  can be used to incorporate prior knowledge of the distribution of the hidden state. There is then a scaling parameter,  $\lambda = \gamma^2(L + \epsilon) - L$ , as well as weightings for the sigma points

$$W_1^m = \frac{\lambda}{L + \lambda},$$

$$W_1^c = \frac{\lambda}{L + \lambda} + (1 - \gamma^2 + \chi),$$

$$W_i^m = W_i^c = \frac{1}{2(L + \lambda)}, \quad i = 2, \dots, 2L + 1.$$

The algorithm is presented in Algorithm 1.

**Algorithm 1** Unscented Kalman Filter Algorithm

1. Form the matrix of sigma points from the time  $k - 1$  mean.

$$\mathbf{S}_{k-1} = [\mathbf{m}_{k-1} \dots \mathbf{m}_{k-1}] + \sqrt{L + \lambda} [0 \quad \sqrt{\mathbf{P}_{k-1}} \quad -\sqrt{\mathbf{P}_{k-1}}].$$

2. Advance the sigma points by one time step.

$$\hat{\mathbf{S}}_{k,i} = f(\mathbf{S}_{k-1,i}), \quad i = 1, \dots, 2L + 1,$$

where  $f(\mathbf{S}_{k-1,i})$  is the expected value of the Euler-Maruyama discretization of Equation (2.9).

3. Compute the predicted mean and covariance respectively.

$$\hat{\mathbf{m}}_k = \sum_i W_i^m \hat{\mathbf{S}}_{k,i},$$

$$\hat{\mathbf{P}}_k = \sum_i W_i^c (\hat{\mathbf{S}}_{k,i} - \hat{\mathbf{m}}) (\hat{\mathbf{S}}_{k,i} - \hat{\mathbf{m}})^\top + \mathbf{Q}_{k-1}(m_{k-1}),$$

where  $\mathbf{Q}_{k-1}$  is the variance of the Euler-Maruyama discretization of Equation (2.9).

4. Now form a matrix of sigma points.

$$\bar{\mathbf{S}}_k = [\hat{\mathbf{m}}_k \dots \hat{\mathbf{m}}_k] + \sqrt{L + \lambda} [0 \quad \sqrt{\hat{\mathbf{P}}_k} \quad -\sqrt{\hat{\mathbf{P}}_k}].$$

5. Compute measurements of the sigma points.

$$\hat{\mathbf{Y}}_{k,i} = h(\bar{\mathbf{S}}_{k,i}), \quad i = 1, \dots, 2L + 1,$$

where  $h(\bar{\mathbf{S}}_{k,i})$  is a function from the current state to a swap rate, given by Equation (3.2).

6. Compute the measured mean, variance and covariances respectively.

$$\boldsymbol{\mu}_k = \sum_i W_i^m \hat{\mathbf{Y}}_{k,i},$$

$$\mathbf{V}_k = \sum_i W_i^c (\hat{\mathbf{Y}}_{k,i} - \boldsymbol{\mu}_k) (\hat{\mathbf{Y}}_{k,i} - \boldsymbol{\mu}_k)^\top + \mathbf{R}_k,$$

$$\mathbf{C}_k = \sum_i W_i^c (\bar{\mathbf{S}}_k - \hat{\mathbf{m}}_k) (\hat{\mathbf{Y}}_{k,i} - \boldsymbol{\mu}_k)^\top,$$

where  $\mathbf{R}_k$  is a diagonal matrix with identical elements on the diagonal, to reflect the assumption of measurement variances being i.i.d. The diagonal element is named  $\sigma_{rates}^2$ , and is estimated. The standard deviation,  $\sigma_{rates}$ , is reported.

7. Compute the Kalman gain, the filtered mean and the measurement conditional covariance respectively.

$$\mathbf{K}_k = \mathbf{C}_k \mathbf{V}_k^{-1},$$

$$\mathbf{m}_k = \hat{\mathbf{m}}_k + \mathbf{K}_k [\mathbf{Y}_k - \boldsymbol{\mu}_k],$$

$$\mathbf{P}_k = \hat{\mathbf{P}}_k - \mathbf{K}_k \mathbf{V}_k \mathbf{K}_k^\top,$$

where  $\mathbf{Y}_k$  is the observed measurement at step  $k$ .

This algorithm is then applied iteratively. A likelihood value is then calculated from quantities calculated in Algorithm 1 (Lütkepohl, 2005).

$$\log(\text{Likelihood}) = -0.5 \sum_{k=1}^N [\mathbf{n} \times \log(2\pi) + \log(|\mathbf{V}_k|) + (\mathbf{Y}_k - \boldsymbol{\mu}_k)^\top \mathbf{V}_k^{-1} (\mathbf{Y}_k - \boldsymbol{\mu}_k)]$$

where  $N$  is the total number of time points, and  $\mathbf{n}$  is the number of measurements.

The Likelihood function value is maximised numerically.

## 3.2 Parameter Recovery

In order to illustrate that the Kalman Filter implementation and resultant likelihood is working correctly, it is a helpful exercise to simulate a dataset of measurements (swap rates in this case), and attempt to recover the known underlying parameters. A simulated data set of 500 points, each with four swap rates of maturities 2, 5, 10 and 15 years, is used. The results from recovering simulated parameters, for a two-factor model, are presented Table 3.1.

**Tab. 3.1:** Parameter Recovery

Parameter	Known	Estimation
$\alpha$	0.5	0.4831
$\kappa_{11}$	0.2	0.1971
$\kappa_{22}$	0.1	0.1033
$\theta_1$	0.4	0.3976
$\theta_2$	0.3	0.3010
$\sigma_1$	0.2	0.1947
$\sigma_2$	0.3	0.3072
$\sigma_{rates}$	$10^{-4}$	0.0001

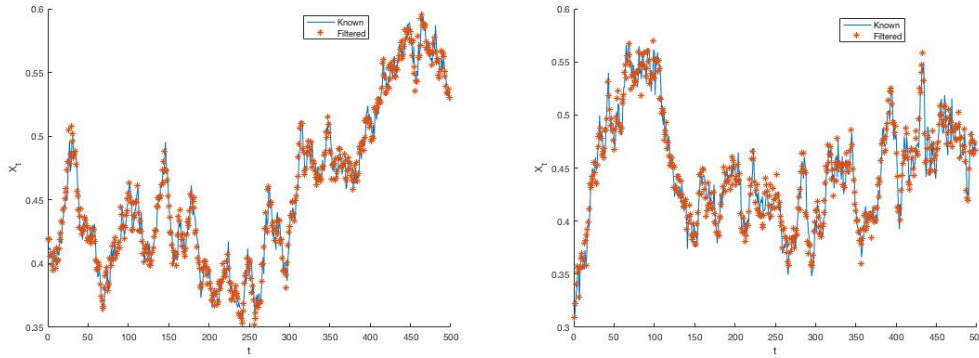
All the parameters were recovered to within a reasonable accuracy. There is, however, a point of clarity with regards to the estimation of the drift parameters,  $\kappa$  and  $\theta$ . It was noted in Section 2.1 that  $\{\zeta_t\}$  contains a Radon-Nikodým process, so there is an implicit measure change in the machinery of  $\{\zeta_t\}$ . That is, the parameter estimation is done over the risk-neutral,  $\mathbb{Q}$ , measure cross-sectionally to value ZCBs, and over the real-world,  $\mathbb{P}$ , measure to match the time series factor process dynamics. As is discussed in Duffee and Stanton (2012), the likelihood value is insensitive to the real world drift parameters, which implies the most likely parameters are given by the best fitting risk-neutral, cross-sectional, ZCB valuation parameters. Recall from Equation (2.5) that the  $\kappa$  and  $\theta$  parameters are prevalent in the valuation of ZCBs. Therefore, the estimated parameters do not necessarily match the real-world observed factor process dynamics. In the case for simulated data, the same parameters are used for the, cross-sectional, ZCB valuation and for the real-world drift dynamics, so the real-world drift parameter recovery falsely appears successful.

It is for the above reason that Filipović *et al.* (2017) finds the specification which is used in this dissertation too restrictive. They address it by introducing an extended state price density specification. In the extension they develop the state price density under some auxiliary measure  $\mathbb{A} \equiv \mathbb{P}$ , and introduce a variable,  $\delta_t$ ,

which allows for freedom to express the market price of risk as  $\lambda_t^{\mathbb{P}} = \lambda_t^{\mathbb{A}} + \delta_t$ . The exogenous choice of  $\delta_t$  allows for the introduction of factors which effect the future distribution of bond values under  $\mathbb{P}$ , but not under  $\mathbb{Q}$ .

As a further check for consistency the known, simulated, factor processes are compared with the filtered factor processes in Figure 3.1.

**Fig. 3.1:** Simulated vs Filtered Factor Processes



Once the optimal parameters have been found, then one can use the Kalman Filter to infer the location of the underlying states from the noisily measured variables, which are swap rates.

### 3.3 Empirical Results

A South African market data set is used in the empirical analysis. In this data set there are 626 daily fair ZAR swap rates, over the period 27/09/2012-30/03/2015, for multiple maturities. In our empirical analysis we use 2Y, 5Y, 10Y and 20Y swaps. Estimation is done for a two-factor model. Principle components analysis shows that two principle components account for 99.81% of swap rate variation, suggesting that a two-factor model can account for nearly all of the variation observed in the data set.

Following on from [Filipović \*et al.\* \(2017\)](#),  $\alpha$  is initially fixed to be 0.5, and then the remaining parameters are estimated. The full parametrization from Section 2.4, which also requires estimation of  $\alpha$ , is then considered. The results from estimating parameters for the two specifications, from the market swap data, are presented in Table 3.2.

**Tab. 3.2:** Parameter Estimates

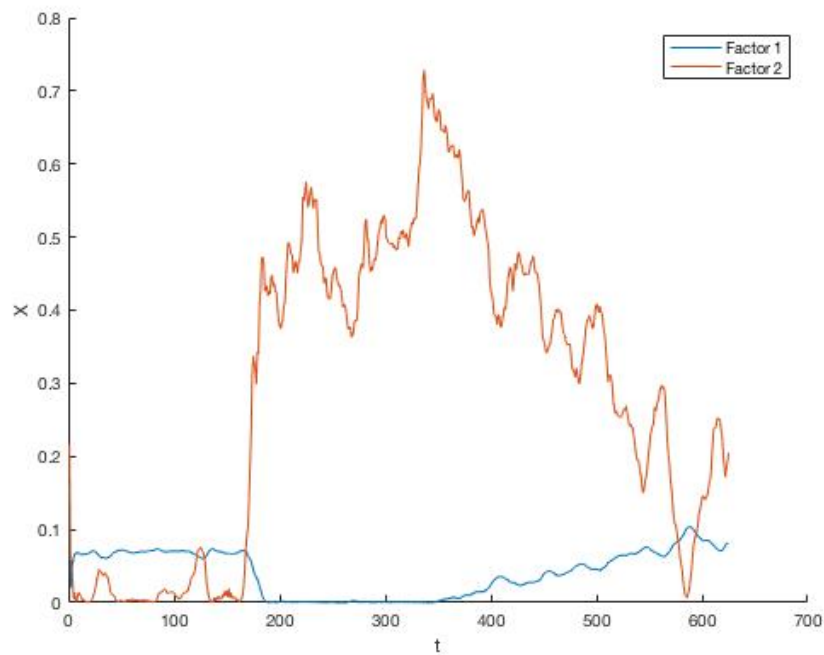
<u>Sqrt process parameter</u>	<u>Estimation</u>
$\alpha$	Set to be = 0.5
$\kappa_{11}$	0.5183
$\kappa_{22}$	0.0378
$\theta_1$	0.0928
$\theta_2$	1.5869
$\sigma_1$	0.1653
$\sigma_2$	0.6425
$\sigma_{rates}$	0.0013
log(likel)	12684.76
<u>CEV process parameter</u>	<u>Estimation</u>
$\alpha$	0.3215
$\kappa_{11}$	0.5116
$\kappa_{22}$	0.0380
$\theta_1$	0.0943
$\theta_2$	1.5734
$\sigma_1$	0.1003
$\sigma_2$	0.3940
$\sigma_{rates}$	0.0013
log(likel)	12726.37

### 3.3.1 Discussion

It is noted, and will be shortly discussed, that when allowing for  $\alpha$  to vary from the fixed sqrt specification of 0.5, then the estimate changes significantly. Similarly, the estimated  $\sigma$  parameters are substantially different between the two specifications. This can be contrasted with the estimated strength of mean reversion,  $\kappa$ , and the level of mean reversion,  $\theta$ , parameter estimates, which have the appearance of being stable. This is due to the fact that they are similar in value between the two specifications. Given the fact that the martingale term in the factor process dynamics, Equation (2.2), plays no part in the valuation of ZCBs, as is seen by Equation (2.5), then this is a good check for consistency. That is, the change of the martingale term between the two specifications does not effect swap rates, which are made up of ZCB values as seen in Equation (3.2). Recall that the  $\theta$  and  $\kappa$  estimates represent the most likely risk-neutral, cross-sectional, estimates. This is due to the likelihood value being insensitive to the real-world drift parameters (Duffee

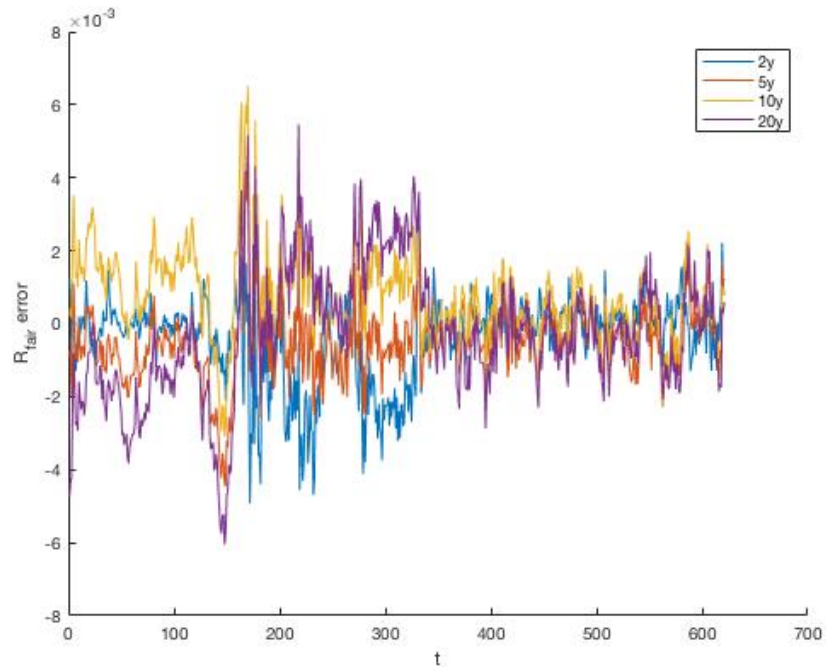
and Stanton, 2012). Therefore  $\theta$  and  $\kappa$  are estimated to give the most likely term structure fit, which remains the same between the two specifications. The filtered factor processes, using CEV parameters, are presented in Figure 3.2. This is done in light of the discussion undertaken in Section 3.2.

**Fig. 3.2:** Filtered Factor Processes - CEV Parameters



The effect of the implicit measure change in  $\{\zeta_t\}$  can be observed in the drift parameter estimates. The clearest is for the  $\theta_2$  estimate, which is approximately 1.58. It is observed that factor 2 remains on the interval  $[0, 1]$ , which is not consistent with a real world level of mean reversion of 1.58.

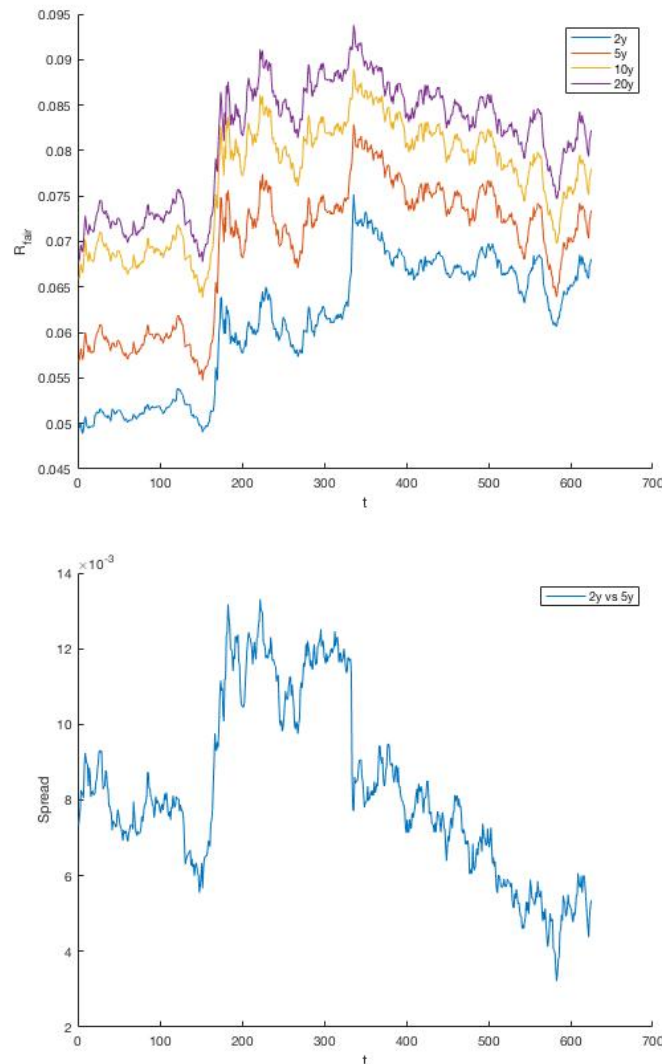
Despite not using the extended state price density specification of Filipović *et al.* (2017), a good fit to swap data is achieved. The time series filtered swap rate error is presented in Figure 3.3.

**Fig. 3.3:** Filtered Swap Rate Error - CEV Parameters

The middle period of greater error came during a time of significant moves in the market, and in particular, a steepening in the front end of the yield curve and an increase in rates. To illustrate this, the time series swap data, as well as the spread between 5y and 2y swaps, is presented in Figure 3.4.



Fig. 3.4: Time Series of Data



A significant result is that when allowing the  $\alpha$  parameter to vary from a fixed level of 0.5, then the value is substantially different to 0.5. There is an accompanying increase in the value of the likelihood of the parameters as is expected, given the fact that a sqrt process is a subset of a CEV process. There is however not an accompanying increase in cross-sectional fit, as the measurement standard deviation,  $\sigma_{rates}$ , is stable. The stable measurement standard deviation,  $\sigma_{rates}$ , is consistent with the  $\kappa$  and  $\theta$  parameters being stable, as follows from the above discussion around the term structure fit. This implies that the accompanying improved fit to the time series swap rate volatility dynamics is significant, as this is where the increased likelihood originates. The extra degree of freedom, in allowing  $\alpha$  to vary,

allows the level-dependence of the volatility to be estimated instead of imposed. In Figure 3.5, the difference in volatility level dependence is illustrated by plotting the state value,  $X$ , versus  $X^\alpha$ , for a sqrt process ( $\alpha = 0.5$ ), and the estimated CEV parameter, ( $\alpha = 0.3215$ ). The factor values are plotted on the interval  $[0, 1]$ , which is consistent with what is observed in Figure 3.2.

**Fig. 3.5:**  $X^\alpha$  for CEV and Sqrt Process

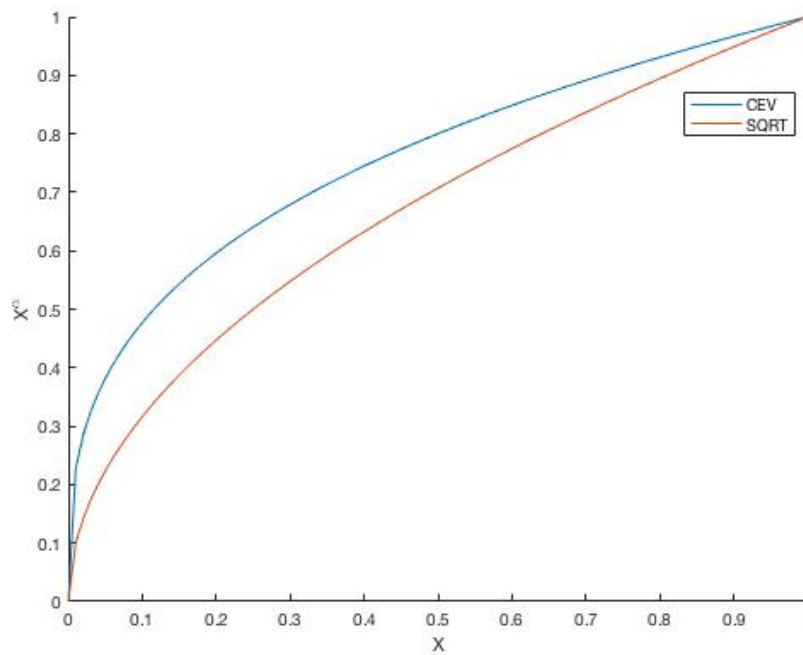
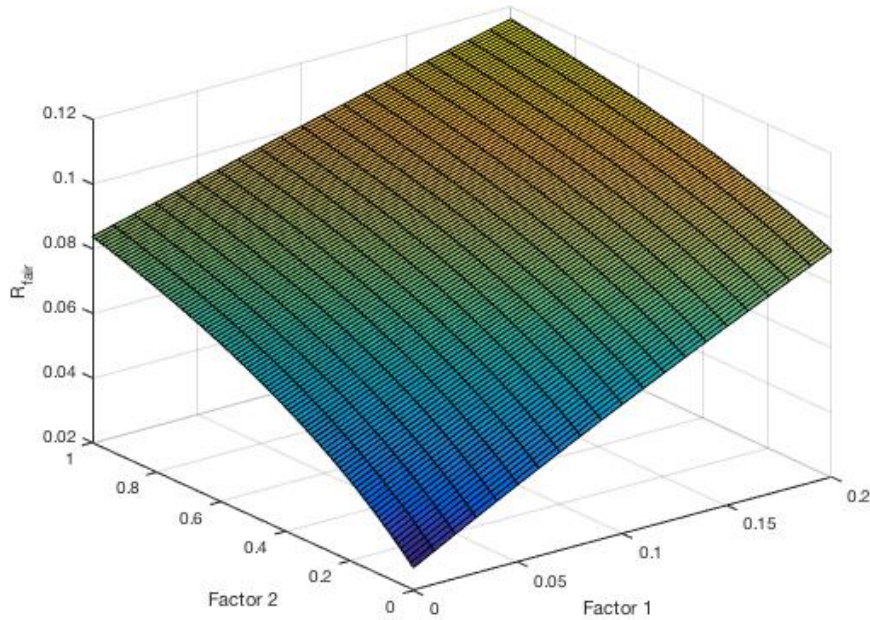


Figure 3.5 indicates that the sqrt process underestimates the contribution of the degree of state dependence on the volatility of the process. This observation is reflected in the parameter estimates for  $\sigma$ , where it is noticed that to account for the underestimated level dependent volatility contribution, the  $\sigma$  parameters are greater for the sqrt process than for the CEV process. What is not obvious, is the relationship between the value of the factor process and the swap rate. In Figure 3.6 we present the fair swap rate,  $R_{\text{fair}}$ , using the CEV parameters, versus the factor process values. The first factor process is plotted to be on the interval  $[0, 0.2]$ , and the second factor process being on the interval  $[0, 1]$ . These intervals are consistent with what is observed in Figure 3.2.

Fig. 3.6:  $R_{\text{fair}}$  vs Factor Values

It is clear that as the values of the factor processes decrease, then it results in a decreased fair swap rate. It has also been seen, in Figure 3.5, that decreasing the values of the factor process results in increased level dependent volatility when using estimated CEV parameters, if compared to sqrt parameters. This implies that decreased swap rates are accompanied with increased level dependent volatility when using estimated CEV parameters, if compared to sqrt parameters. An empirical study done in [Filipović \*et al.\* \(2017\)](#) where changes in swap implied volatility were regressed on changes in the swap rate level, conditional on the swap rate being within certain ranges, was undertaken. The study found that changes in swap rates had more statistical significance when conditioning on being in a lower interest rate regime. When using a CEV factor process, then the model parameters can be estimated to accurately reflect the relationship outlined in the [Filipović \*et al.\* \(2017\)](#) empirical study. This can be contrasted with the case where the level dependence of volatility is imposed as an exogenous choice, such as for a sqrt process, which as has been noted would have resulted in an underestimation of the contribution to the level dependence of volatility for the data set used in this dissertation.

The level dependence of volatility is connected to the skew in swaptions, and this is developed in Chapter 4.

## Chapter 4

# Swaptions

A major advantage of the Linear-Rational framework is that despite the ability to add multiple spanned and unspanned factors, exact swaption valuation remains tractable. This can be contrasted with the popular Affine Term Structure Models, where approximations are often required. An example would be [Schrager and Pelsser \(2006\)](#).

### 4.1 Swaption Valuation

A European payer swaption expiring at  $T_0$ , has a value at  $T_0$  of

$$\begin{aligned} C_{T_0} &= (V_{T_0}^{swap})^+ \\ &= \left( \sum_{i=0}^n c_i P(T_0, T_i) \right)^+ \\ &= \frac{1}{\zeta_{T_0}} \left( \sum_{i=0}^n c_i \mathbb{E}_{T_0}[\zeta_{T_i}] \right)^+ \\ &= \frac{(v_{swap}(X_{T_0}))^+}{\zeta_{T_0}}. \end{aligned} \tag{4.1}$$

The coefficients,  $c_i$ , in the second equation are read from Equation (3.1). The third equation follows directly from the ZCB value using Equation (2.1). The conditional expectations,  $\mathbb{E}_{T_0}[\zeta_{T_i}]$ , can be recovered from Equations (2.3) and (2.4), as was done in Appendix A.2.2, in line  $\star$ . The expression for  $v_{swap}$  is given by

$$v_{swap}(x) = \sum_{i=0}^n c_i e^{-\beta T_i} (\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T_i - T_0)} (x - \theta)). \tag{4.2}$$

For the time  $t \leq T_0$  value of this derivative, the fundamental valuation formula, Equation (2.1), is used

$$\begin{aligned} V_t^{swaption} &= \frac{1}{\zeta_t} \mathbb{E}_t[\zeta_{T_0} C_{T_0}] \\ &= \frac{1}{\zeta_t} \mathbb{E}_t[v_{swap}(X_{t_0})^+] \end{aligned} \quad (4.3)$$

The  $d$ -dimensional integral in Equation (4.3) can be reduced to a 1-dimensional integral by making use of Fourier methods. Theorem 4 of Filipović *et al.* (2017) gives the tractable form

$$V_t^{swaption} = \frac{1}{\pi \zeta_t} \int_0^\infty \operatorname{Re} \left[ \frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda \quad (4.4)$$

for  $\mu > 0$ , where

$$\hat{q}(z) = \mathbb{E}_t[\exp(zv_{swap}(X_{T_0}))]. \quad (4.5)$$

This theorem uses a result from Fourier analysis to express the positive part of a function as a Fourier integral, and makes use of Fubini's theorem to change the order of Fourier integration and expectation. The use of Fubini's theorem places technical conditions on  $\mu$ .

A useful result is that  $\hat{q}(z)$  in Equation (4.5) can be evaluated by using Equation (4.6), which known as the Exponential-Affine Transform Formula, if the underlying dynamics are affine (Filipović, 2009). That is, when the drift vector and diffusion matrices can be written in the form  $u(x) = u + \sum_i x_i \mu_i$ ,  $m(x) = m + \sum_i x_i \Omega_i$  respectively. This is only the case when  $\alpha = 0$  or  $\alpha = \frac{1}{2}$ , so is not applicable in our full specification.

$$\mathbb{E}_t[\exp(u + v^\top X_{T_0})] = \exp(\Phi(T_0) + \Psi(T_0)^\top X_t). \quad (4.6)$$

Within Equation (4.6),  $u$  and  $v$  are read off using Equation (4.2), while  $\Phi$  and  $\Psi$  satisfy a system of Riccati equations, displayed in Equation (4.7).

$$\frac{\partial \Phi}{\partial t} = u^\top \Psi + \frac{1}{2} \Psi^\top m \Psi, \quad (4.7)$$

$$\frac{\partial \Psi_i}{\partial t} = \mu_i^\top \Psi + \frac{1}{2} \Psi_i^\top \Omega_i \Psi_i,$$

with  $\Phi(0) = u$ , and  $\Psi(0) = v$ . While the intention was to keep the expressions outlined in Equations (4.6) and (4.7) general, the equations for the sqrt case ( $\alpha = \frac{1}{2}$ ) are explicitly shown in the Appendix B. In order to value a swaption the factor process value has to be inverted from knowledge of the current term structure, which is not

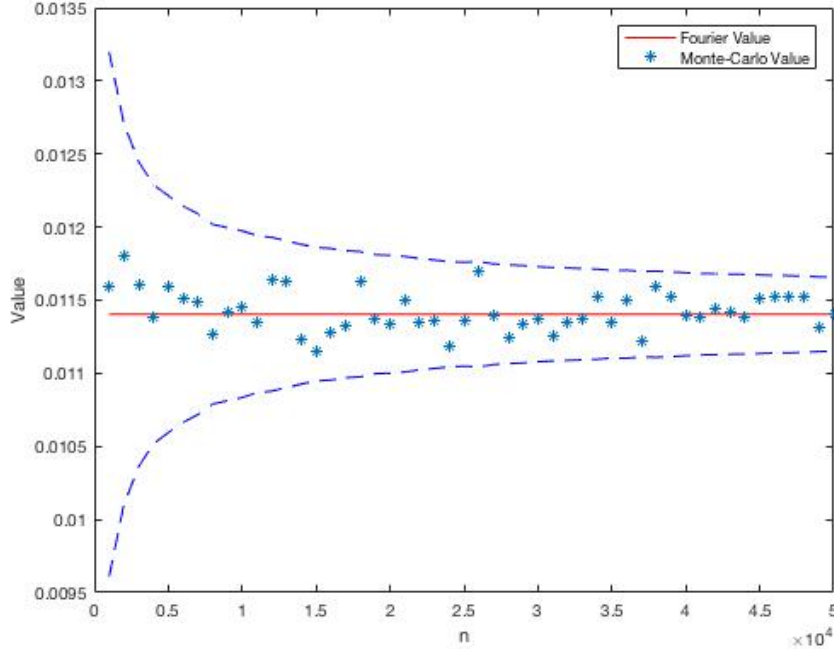
assumed to be an exact fit. It is possible to extend the swaption valuation methodology to have a perfect fit to the initial term structure. This extension is developed in the online appendix of [Filipović \*et al.\* \(2017\)](#).

In order to take into the account results of the parameter estimates from Chapter 3, and their effect on the value of swaptions, then we can not restrict ourselves to affine processes. We therefore also make use of crude Monte-Carlo methods to value a swaption. To verify consistency between the Monte-Carlo and Fourier swaption values we use the following parameters, which specifically sets  $\alpha = 0.5$ .

<u>Parameter</u>	<u>Value</u>
$\alpha$	0.5
$\kappa_{11}$	0.1
$\kappa_{22}$	0.2
$\theta_1$	0.2
$\theta_2$	0.8
$\sigma_1$	0.2
$\sigma_2$	0.3

In Figure 4.1 we see the crude Monte-Carlo 2Y2Y ATM payer swaption value with a three standard-deviation envelope, for  $n$  independent samples, as well as the value achieved by using Fourier methods. The initial state is set to be  $X_0 = [0.5, 0.5]^T$ .

Fig. 4.1: Crude Monte-Carlo vs Fourier Value



The algorithm for the Monte-Carlo value of a swaption struck at rate  $R$ , for maturity  $T_0$ , is presented in Algorithm 2.

---

**Algorithm 2** Monte-Carlo Algorithm
 

---

1. Compute the initial value of the state price density:  $\zeta_0 = \phi + \psi^\top X_0$ .
  2. Advance  $X_t$  using the Euler-Maruyama discretization of Equation (2.9), for  $n$  independent Monte-Carlo simulations, up until the maturity of the swaption,  $T_0$ .
  3. Compute the pay-off of the swaption for each simulation,  $\text{payoff} = (R_{fair_{T_0}} - R)^+ \tau \sum_{i=1}^k P(T_0, T_i)$ .
  4. Compute the  $T_0$  value of each simulated state price density  $\zeta_{T_0} = \phi + \psi^\top X_0$ .
  5. Compute the value of the swaption according to the Monte-Carlo approximation of the fundamental valuation equation, Equation (2.1),  $V_0^{swaption} = \frac{\frac{1}{n} \sum_{i=1}^n \zeta_{T_0, i} \text{payoff}_i}{\zeta_0}$ .
- 

## 4.2 Swaption IV Skew

There are two ways to determine parameters that achieve the correct volatility properties in a model. The first is a cross-sectional approach, and that is to look at a skew in liquidly traded options and calibrate parameters to agree with those

implied volatilities. This is only possible if there are liquidly traded options which are struck away from the money. This is not the case for ZAR swaptions, and therefore this is not possible. The other approach is to estimate parameters that provide the most likely time-series fit to the volatility dynamics, and this is the approach taken in Chapter 3. We therefore observe a skew for ZAR swaptions, which has been implied by the model estimated parameters. The implied vol presented is the Black-76 Implied Volatility (IV). The Black-76 Formula is as follows

$$V_0^{Black76} = \sum_{i=1}^n P(0, T_i) \tau(F_{T_0} \Phi(d_1) - K \Phi(d_2)), \text{ with}$$

$$d_1 = \frac{\log(F_{T_0}/K) + \frac{1}{2}\sigma^2 T_0}{\sigma\sqrt{T_0}}, \quad d_2 = d_1 - \sigma\sqrt{T_0}.$$

When a premium has been calculated ( $\text{premium}_{MC}$ ) using Algorithm 2, then the Black-76 IV is found by numerically solving for a root to the equation  $f(\sigma) = \text{premium}_{MC} - V_0^{Black76}(\sigma)$ .

When using the parameters from Chapter 3, Section 3.3, and a current state of  $X = [0.1, 0.1]^T$  the model implied IV skew for 2Y2Y payer swaptions is visible in Figure 4.2. There are  $n = 200000$  Monte-Carlo simulations used for the valuation of each strike.

**Fig. 4.2:** CEV vs Sqrt Black-76 Implied Vol for Multiple Strikes

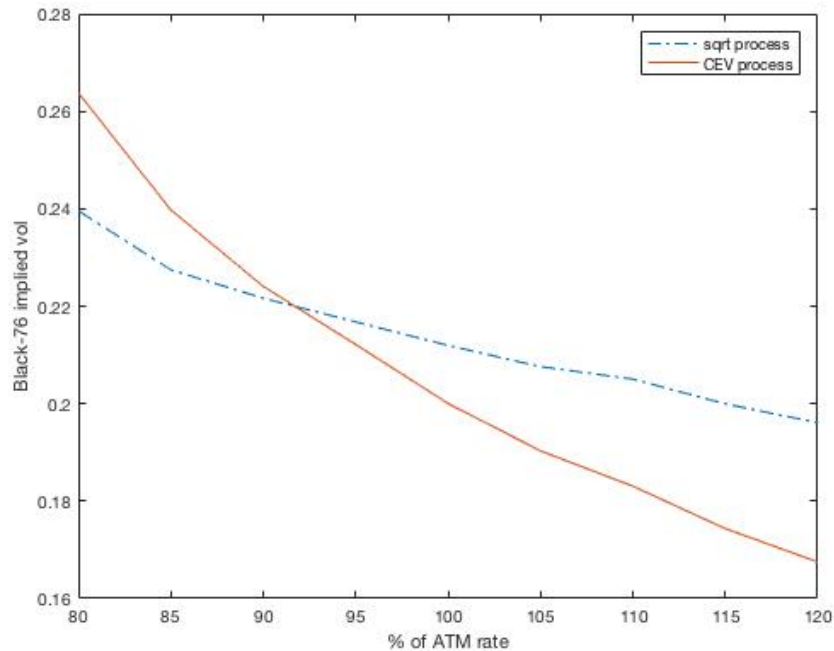


Figure 4.2 should be interpreted in light of the discussion in Chapter 3 and recalling Figure 3.5. That is, the CEV case is expected to provide a more accurate rep-



resentation of the swaption IV skew, as the state dependent volatility is no longer imposed, but is estimated to give the most likely fit to data. In the case that swap rates are low, then when using the CEV parameters there is a higher state dependent volatility, when compared to the sqrt parameters. Therefore, a CEV process will be more volatile for lower rates. This implies a higher IV for deep in-the-money payer swaptions when using CEV parameters, despite the estimated CEV  $\sigma$  parameters being smaller, as this valuation includes lower swap rates. In the case that rates are high, then the CEV and sqrt parameters have comparatively similar state dependent volatilities (Figures 3.5 and 3.6), but the  $\sigma$  parameters estimated for the sqrt process are much higher than that of the CEV process, and therefore the sqrt process will be more volatile for higher rates. This implies that the IV for the sqrt process is expected to be higher for deep out-the-money payer swaptions, as this valuation only takes into account higher swap rates. Therefore based on the parameters found in Chapter 3, one would expect a steeper skew when using the CEV parameters, as is observed in Figure 4.2.

## Chapter 5

# Conclusions

This dissertation introduced Linear-Rational Term Structure Models, and highlighted their significant benefits, which include the following: a lower bound on rates, the ability to add multiple spanned and unspanned factors, and the ability to maintain tractable valuation of interest rate derivatives. The factor process was specialized to follow a CEV-type process, which allows for the degree of state dependence of volatility to be estimated. The freedom to do so is significant as empirical evidence suggests that changes in swap rates have more statistical significance to changes in swap implied volatility, when conditioning on being in a lower interest rate regime. A CEV-type process allows for this relationship to be accurately reflected. When estimating parameters for a two-factor CEV process, on swap market data, a good fit was achieved. This despite the fact that there were significant market moves during the data period. The estimated parameters suggest that the degree of state dependent of volatility is significantly different to that of models where the state dependence is set *a priori*, such as in a square root process. In particular, it finds that the state dependent volatility is underestimated when using a square root process. The effect this has on the model-implied swaption was discussed, and it confirmed that when using the estimated CEV process parameters the swaption IV skew was steeper, when compared to a using square-root process parameters.

# Bibliography

- Andersen, L. and Piterbarg, V. (2007). Moment Explosions in Stochastic Volatility Models, *Finance and Stochastics* **11**(1): 29–50.
- Collin-Dufresne, P. and Goldstein, R. (2002). Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility, *The Journal of Finance* **57**(4): 1685–1730.
- Constantinides, G. (1992). A Theory of the Nominal Term Structure of Interest Rates, *Review of Financial Studies* **5**(4): 531–552.
- Cox, J. (1975). Notes on Option Pricing I: Constant Elasticity of Diffusions. Stanford University.
- Duffee, G. and Stanton, R. (2012). Estimation of Dynamic Term Structure Models, *The Quarterly Journal of Finance* **2**(2): 111–135.
- Filipović, D. (2009). *Term Structure Models: A Graduate Course*, Springer-Finance.
- Filipović, D., Larsson, M. and Trolle, A. (2017). Linear-Rational Term Structure Models, *The Journal of Finance* **72**(2): 655–704.
- Flesaker, B. and Hughston, L. (1996). *Risk*, **1**: 96–99.
- Harrison, J. and Pliska, R. (1981). Martingales and stochastic integrals in the theory of continuous trading, *Journal of Economic Theory* **11**(3): 215–260.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag.
- Macrina, A. (2014). Heat Kernel Models for Asset Pricing, *International Journal of Theoretical and Applied Finance* **17**(7): 1–34.
- Mijatović, A. and Urusov, M. (2012). On The Martingale Property of Certain Local Martingales, *Probability Theory and Related Fields* **152**(1-2): 1–30.
- Rogers, L. (1997). The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates, *Mathematical Finance* **7**(2): 157–176.
- Särkkä, S. (2010). Lecture 5 - Unscented Kalman Filter, and General Gaussian Filtering. Aalto University.
- Schrager, D. and Pelsser, A. (2006). Pricing Swaptions and Coupon Bond Options in Affine Term Structure Models, *Mathematical Finance* **16**(4): 673–694.

## Appendix A

# Chapter 2

### A.1 Radon-Nikodým Process

In order to value a cash flow using no arbitrage principles, then the measure under which the  $\mathcal{F}_t$ -conditional expectation is taken must result in the deflated cash flow being a martingale. Let this measure be  $\mathbb{Q}$ ,  $\{V_t\}$  be the value process of an asset, and  $\{A_t\}$  be a suitable numéraire, then for the no arbitrage condition to hold

$$\frac{V_t}{A_t} = \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T}{A_T} \mid \mathcal{F}_t\right].$$

Letting  $Z_t = \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t}$ , with some measure  $\mathbb{P}$  which is equivalent to  $\mathbb{Q}$ , then

$$V_t \frac{Z_t}{A_t} = \mathbb{E}^{\mathbb{P}}\left[V_T \frac{Z_T}{A_T} \mid \mathcal{F}_t\right].$$

Letting  $\zeta_t = \frac{Z_t}{A_t}$ , then

$$V_t = \frac{1}{\zeta_t} \mathbb{E}^{\mathbb{P}}\left[V_T \zeta_T \mid \mathcal{F}_t\right].$$

Therefore, the time- $t$  value,  $V_t$ , of a time- $T$  cash flow,  $C_T$ , is given by

$$V_t = \frac{1}{\zeta_t} \mathbb{E}^{\mathbb{P}}\left[C_T \zeta_T \mid \mathcal{F}_t\right].$$

### A.2 Derivations for Section 2.1

#### A.2.1 Conditional Expectation of the Factor Process

Rearranging a 1-factor case of Equation (2.2) as

$$dX_t + \kappa X_t dt = \kappa \theta dt + dM_t.$$

Multiplying through by an integrating factor,  $e^{\kappa(t-t_0)}$ , and grouping terms

$$d(X_t e^{\kappa(t-t_0)}) = e^{\kappa(t-t_0)} \kappa \theta dt + e^{\kappa(t-t_0)} dM_t.$$

Integrating and the rearranging to get

$$X_t = \theta + e^{-\kappa(t-t_0)}(X_{t_0} - \theta) + e^{-\kappa(t-t_0)} \int_{t_0}^t e^{\kappa(s-t_0)} dM_s.$$

Taking the time- $t_0$  conditional expectation

$$\mathbb{E}_{t_0}[X_t] = \theta + e^{-\kappa(t-t_0)}(X_{t_0} - \theta).$$

### A.2.2 ZCB Value Formula

Using the fundamental valuation equation, Equation (2.1), and setting  $C_T = 1$  to get the value of a ZCB,  $P(t, T)$ .

$$P(t, T) = \frac{1}{\zeta_t} \mathbb{E}_t[\zeta_T].$$

Taking conditional expectations of the state price density, given by Equation (2.3), as is required in the ZCB value formula above

$$\mathbb{E}_t[\zeta_T] = e^{-\beta T} (\phi + \psi^\top \mathbb{E}_t[X_T]).$$

Substituting in for the conditional expectation of  $X_T$ , given by Equation (2.4)

$$\mathbb{E}_t[\zeta_T] = e^{-\beta T} (\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T-t)} (X_t - \theta)). \quad \star$$

Substituting this into the ZCB value formula

$$P(t, T) = \frac{1}{e^{-\beta t} (\phi + \psi^\top X_t)} e^{-\beta T} (\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T-t)} (X_t - \theta)).$$

Which simplifies to

$$P(t, T) = e^{-\beta(T-t)} \frac{\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T-t)} (X_t - \theta)}{\phi + \psi^\top X_t}.$$

### A.2.3 Short Rate

The short rate is given by

$$r_t = -\frac{\partial}{\partial T} \log P(t, T) \Big|_{T=t}.$$

This can be written in a more convenient form as

$$r_t = -\frac{1}{P(t, T)} \partial_T P(t, T) \Big|_{T=t}.$$

Substituting in to get

$$r_t = e^{\beta(T-t)} \frac{\phi + \psi^\top X_t}{\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T-t)} (X_t - \theta)} \Big|_{T=t} \times \left[ \beta e^{-\beta(T-t)} \frac{\phi + \psi^\top \theta + \psi^\top e^{-\kappa(T-t)} (X_t - \theta)}{\phi + \psi^\top X_t} + e^{-\beta(T-t)} \frac{\psi^\top \kappa e^{-\kappa(T-t)} (X_t - \theta)}{\phi + \psi^\top X_t} \right] \Big|_{T=t}.$$

Evaluating at  $T = t$ , and the performing cancellations, to get the result

$$r_t = \beta - \frac{\psi^\top \kappa (\theta - X_t)}{\phi + \psi^\top X_t}$$

## Appendix B

# Chapter 4

The expression for  $u$  and  $v$  are given, as well as the system of Riccati equations in matrix form. This is for the the case that  $\alpha = 0.5$ .

$$u = \sum_{i=0}^n c_i e^{-\beta T_i} (\phi + \psi^\top \theta - \psi^\top e^{-\kappa(T_i - T_0)} \theta),$$

$$v = \sum_{i=0}^n c_i e^{-\beta T_i} \psi^\top e^{-\kappa(T_i - T_0)}.$$

$$\frac{\partial \Phi}{\partial t} = (\kappa \theta)^\top \Psi,$$

$$\frac{\partial \Psi}{\partial t} = -\kappa \Psi + \frac{1}{2} \text{diag}(\sigma^2) \text{diag}(\Psi) \Psi, \text{ where } \sigma^2 \text{ is a vector of squared } \sigma \text{ values.}$$