Model Fitting of a Two-Factor Arbitrage-Free Model for the Term Structure of Interest Rates using Markov Chain Monte Carlo: Part 1: Theory and Methodology

Charnchai Leuwattanachotinan[1], Andrew J.G. Cairns[2], and George Streftaris[3]

April, 2012

Abstract

In this paper we use Markov chain Monte Carlo (MCMC) simulation to calibrate a two-factor arbitrage-free model for the term structure of interest rates, proposed by Cairns (2004) based on the positive-interest framework (Flesaker and Hughston, 1996). The model is a time-homogeneous model driven by latent state variables which follow a two-dimensional Ornstein-Uhlenbeck process. The standard Metropolis-Hastings (MH) algorithm was first employed for estimating both model parameters and latent variables using simulated data in order to validate the algorithm and ensure that it can result in reasonable and reliable estimates. According to the results, it turns out that the chains of the estimation tend to converge slowly. Therefore, we carry out some improvements by using the adaptive MH algorithm associated with a blocking strategy and reparameterising the log posterior of Cairns bond prices.

Keywords: positive interest; term structure model; Ornstein-Uhlenbeck; Bayesian inference; MCMC; Metropolis-Hastings.

[1]Charnchai Leuwattanachotinan: Erasmus School of Economics, Erasmus University Rotterdam, the Netherlands. E: Leuwattanachotinan@ese.eur.nl.

[2]Andrew J.G. Cairns: Maxwell Institute for Mathematical Sciences, and Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS, UK. E: A.Cairns@ma.hw.ac.uk.

[3]George Streftaris: Maxwell Institute for Mathematical Sciences, and Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS, UK. E: G.Streftaris@ma.hw.ac.uk.

1 Introduction

Interest rate plays a substantial role in several kinds of investment. It lends itself to a form of security and also the underlying of other securities such as derivatives. In light of asset and liability management, a change in interest rates affects the valuation on both sides of a bank's balance sheet. From an economic point of view, the interest rate influences decision making for investors and is a key indicator for the economy that determines the levels of investment, saving and consumption. By definition, interest rate can simply be thought of as the cost of borrowing from one to another, but its modelling is far more complicated than this simple definition suggests.

Typically, interest rates are considered for a wide range of maturities as the yield curve (i.e. the term structure). In order to develop a model for it, we first need to understand the behaviour of interest rates. With reference to Cairns (2004), some desirable characteristics for a term structure model are as follows:

- All interest rates should be positive and can remain values close to zero.
- The model should be arbitrage-free and framed in continuous-time in order to be able to use for derivative pricing and hedging.
- The model should have a mean-reverting process reflecting that in reality interest rates will not be completely allowed to move freely, but once they reach extreme levels, they will be pulled back to some long-term rates (e.g. an intervention by the Central banks).
- The model should be able to produce yield curves similar to what we can observe in historical data.

In recent decades, a considerable number of arbitrage-free models have been proposed for the term structure of interest rates. At the beginning, much attention was drawn to one-factor models for the short-term rates (i.e. risk-free rates) such that their dynamics are characterised by stochastic processes. Unfortunately, from empirical research (e.g. Litterman et al., 1991), it is suggested that one-factor models are unlikely to sufficiently capture the dynamics of real market data. Accordingly, multifactor arbitrage-free models have been developed thereafter. Despite an increase of the number of factors and their rigorous frameworks, several multifactor models are yet required to impose some restrictions in order to guarantee interest rates being positive. In effect, this makes those models less flexible and more difficult to implement. Nevertheless, a new framework introduced by Fleasaker and Hughton (1996) (the positive-interest framework) allows us to develop a multifactor model that can ensure the positivity of interest rates in a natural way.

In aspect of model implementation, multifactor models are often incorporated with unobservable state variables so that advanced and modern statistical techniques are required for estimation. The main methodologies that frequently appear in literature related to term structure modelling are maximum likelihood (ML), general method of moments (GMM) and efficient method of moments (EMM) which all follow the frequentist statistical approach. For the Bayesian approach, Markov chain Monte Carlo (MCMC) simulation is the prevailing method for estimation.

In an early stage, likelihood-based estimation played a key role in statistical inference and modelling. It was widely used in many applications, including the estimation of term structure models. For instance, Pearson and Sun (1994) and Nowman (1997) used the ML method to estimate a two-factor CIR model and a one-factor CKLS model respectively. After a while, much of the attention moved to the general method of moments when it was first formalised by Hansen (1982). GMM generalises the standard method of moments (MM) in the sense that the number of moment functions can be greater than the number of parameters being estimated. More precisely, the moment functions are initially defined and then we solve an optimisation problem of equating the sample average of the moment functions to zero. In case that the number of moment functions is just equal to the number of parameters, the exact solution thus can be achieved. Examples of applications to term structure modelling are those by Longstaff and Schwartz (1992) and Chan et al. (1992), where they applied the GMM methodology to estimate the CIR and CKLS models respectively.

For the models incorporated with unobserved or latent variables, the GMM method may not be applicable if the moment functions cannot be numerically evaluated. Furthermore, in a presence of the latent variables, the complete likelihood function may also be hard to obtain and hence the ML method is unlikely to be feasible. Under these circumstances, one may use the simulated method of moments (SMM). The technical properties of SMM methodology can be found in Duffie and Singleton (1993). The efficient method of moments, described by Gallant and Tauchen (1996), is a kind of SMM. Since introduced, it has appeared in several well-known literatures of calibrating term structure models such as stochastic volatility models (Andersen and Lund, 1997), affine term structure models (Dai and Singleton, 2000) and quadratic term structure models (Ahn et al., 2002). Despite its popularity, one prevailing shortcoming of EMM is computationally expensive comparing to GMM and ML.

In recent years, Bayesian estimation has also increasingly been paid attention due to the development of the MCMC simulation. In the past, the Bayesian approach was less preferable because of the difficulty of implementation, particularly for high-dimensional problems. The existence of the MCMC methodology allows us to tackle such problems in more flexible and feasible ways. One distinct advantage of MCMC is that we can obtain information about parameter uncertainty directly from the simulation output. Specifically, MCMC avoids relying on an asymptotic approximation as do GMM and EMM. With application to term structure modelling, Eraker(2001) applied MCMC to fit a two-factor model with a latent stochastic volatility component using weekly US Treasury data from Jannuary 1954 to May 1997. Hu (2005) estimated multifactor affine models using MCMC but those models are not in a general form since Wiener processes are assumed to be uncorrelated ($\rho = 0$). Moreover, the chains of several parameters converge rather poorly and need to be improved. Pooter et al. (2007) employed a Bayesian approach to estimate term structure models incorporating macroeconomic variables but some results appeared to have a substantial problem of convergence since some parameters did not converge at all. Other examples of implementing term structure models using MCMC can also be found in Bester (2004), and Lamoureux and Witte (2002).

In this paper, we calibrate a specific two-factor arbitrage-free term structure model developed by Cairns (2004) using simulated data with Markov chain Monte Carlo being the central methodology for our estimation. A new family of the Cairns models is based on the positive-interest framework that can be used in long-term risk management. In the Cairns model, we are required to estimate both model parameters and time-varying latent state variables, driven by a two-dimensional Ornstein-Uhlenbeck process. Numerical integration methods are also used for computing the bond prices since their closed-form solution cannot be achieved analytically. Here, theoretical bond prices are numerically computed using the Trapezoidal rule since it is most convenient for the programming and, more importantly, we found that by this simple method there is no significant difference to the prices compared to using more complicated techniques such as the adaptive Simpson quadrature.

Consequently, the full joint posterior density of the Cairns bond price is derived and hence the latent variables and model parameters are estimated where the Metropolis-Hastings (MH) algorithms play a key role for our estimation. We initially employ the standard MH which can give us fairly acceptable results but it is evident that the chains converge rather slowly. Accordingly, we then use the adaptive MH with a blocking strategy and repameterise the bond posterior distribution for improving the chain convergence.

1.1 Outline of the paper

The remainder of this paper is as follows. In Section 2, we introduce our estimation framework and the Cairns term structure model. Section 3 outlines the dataset that we will be used to calibrate the Cairns model. In Section 4, we describe the MCMC algorithms which are the core methodology for our estimation. In Section 5 and 6, we discuss and analyse the estimation results of using the standadard and the adaptive MH algorithms with some improvement respectively. Section 7 concludes.

2 Estimation Framework

Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space. We first set up an estimation framework by assuming interest rates in the market follow the Cairns model such that the observations

$$P(t, \tau_{tj}) = C(\tau_{tj}; X(t), \theta) + \varepsilon(t, j), \tag{1}$$

where $\varepsilon(t, j) \sim$ i.i.d. $N(0, \sigma_\varepsilon^2)$ for some constant σ_ε, $P(t, \tau_{tj})$ is a zero-coupon bond price at time t for a bond that pays 1 at time $t + \tau_{tj}$ of the maturities τ_{tj}, for $j = 1, 2, \ldots, N_t$, and $C(\tau_{tj}; X(t), \theta)$ is the theoretical bond price by the two-factor Cairns model, i.e.

$$C(\tau, x, \theta) = \frac{\int_\tau^\infty H(u, x)du}{\int_0^\infty H(u, x)du}, \tag{2}$$

where

$$H(u, x) = \exp\Big[-\beta u + \sum_{i=1}^{2} \sigma_i x_i \, e^{-\alpha_i u} - \frac{1}{2} \sum_{i,j=1}^{2} \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j} e^{-(\alpha_i+\alpha_j)u} \Big], \tag{3}$$

$\theta = (\beta, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho, \gamma_1, \gamma_2)'$ is the model parameter vector and $X(t) = (X_1(t), X_2(t))'$, for $t = 1, \ldots, M$, are the latent variables which follow

$$dX_i(t) = \alpha_i(\gamma_i - X_i(t))dt + \sum_{j=1}^{2} \sigma_{ij} dW_j(t), \tag{4}$$

where $W_1(t)$ and $W_2(t)$ are two independent Wiener processes with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ under the real world probability P.

With this setting, we have $P(t, \tau_{tj}) \sim N(C(\tau_{tj}; X(t), \theta), \sigma_\varepsilon^2)$.

3 Data

In an initial stage, we consider the estimation of the latent variables and model parameters with a simulated dataset which will allow us to check the accuracy of the algorithm with reference to the true values for $X(t) = (X_1(t), X_2(t))'$ and θ. To simulate bond prices according to (1), we define 20 constant maturities ($N_t = 20$, for all t): 0.25, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 12.5, 15.0, 17.5, 20.0, 22.5, 25.0, 27.5, 30.0 years, and then generate unit normal random variables $\varepsilon(t, j)$, for $j = 1, 2, \ldots, 20$ with the mean and covariance matrix:

$$\mu = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_\varepsilon & 0 & \cdots & 0 \\ 0 & \sigma_\varepsilon & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon \end{bmatrix}, \text{ where } \sigma_\varepsilon = 0.001.$$

Clearly, we assume here that the bond prices of each time t are independent with a fixed normal randomness. Additionally, the other model parameter values are chosen with respect to Cairns (2004) as

$$\beta = 0.04, \alpha_1 = 0.6, \alpha_2 = 0.06, \sigma_1 = 0.6, \sigma_2 = 0.4, \rho = -0.5,$$

whereas the latent variables can be simulated from the exact solution of a two-dimensional Ornstein-Uhlenbeck process using the following proposition.

Proposition 1 *[The Exact Solution of a Two-Dimensional Ornstein-Uhlenbeck Process] Suppose that $X(t) = (X_1(t), X_2(t))'$ follows a two-dimensional Ornstein-Uhlenbeck process such that*

$$d\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}\begin{pmatrix} \gamma_1 - X_1(t) \\ \gamma_2 - X_2(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} d\begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix}, \quad or$$

$$dX_i(t) = \alpha_i(\gamma_i - X_i(t))dt + \sum_{j=1}^{2} \sigma_{ij} dW_j(t),$$

where $X_1(0) = \hat{x}_1, X_2(0) = \hat{x}_2$, $W_1(t)$ and $W_2(t)$ are two independent Wiener processes. Then, the exact solution of $X(t)$ can be found and achieved at

$$X_i(t) = \gamma_i + (X_i(0) - \gamma_i)e^{-\alpha_i t} + \sum_{j=1}^{2} \sigma_{ij} \int_0^t e^{-\alpha_i(t-s)} dW_j(s). \tag{5}$$

Hence, $(X_1(t), X_2(t))'$ is bivariate normal with

$$\begin{aligned} \mathbb{E}(X_i(t)) &= \gamma_i + (X_i(0) - \gamma_i)e^{-\alpha_i t}, \\ Var(X_i(t)) &= \frac{(\sigma_{i1}^2 + \sigma_{i2}^2)}{2\alpha_i}(1 - e^{-2\alpha_i t}), \\ Cov(X_1(t), X_2(t)) &= \frac{\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}}{\alpha_1 + \alpha_2}(1 - e^{-(\alpha_1+\alpha_2)t}). \end{aligned}$$

The latent variables $X(t)$ under the real world probability measure P can be simulated from the exact solution in Proposition 1. Given $\gamma_1 = \gamma_2 = 0$ and the instantaneous correlation matrix

$$\rho = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = AA^T, \quad \text{where} \quad A = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix},$$

it follows that

$$X_1(t) = X_1(0)e^{-\alpha_1 t} + \int_0^t e^{-\alpha_1(t-s)} dW_1(s),$$

$$X_2(t) = X_2(0)e^{-\alpha_2 t} + \rho \int_o^t e^{-\alpha_2(t-s)} dW_1(s) + \sqrt{1-\rho^2} \int_0^t e^{-\alpha_2(t-s)} dW_2(s). \quad (6)$$

Specifically,

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} \sim N_2 \left(\begin{pmatrix} X_1(0)e^{-\alpha_1 t} \\ X_2(0)e^{-\alpha_2 t} \end{pmatrix}, \begin{pmatrix} \frac{1}{2\alpha_1}(1-e^{-2\alpha_1 t}) & \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}) \\ \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)t}) & \frac{1}{2\alpha_2}(1-e^{-2\alpha_2 t}) \end{pmatrix} \right).$$

Likewise, from time t_k to t_{k+1}, where $\Delta t = t_{k+1} - t_k$, we will have

$$\begin{pmatrix} X_1(t_{k+1}) \\ X_2(t_{k+1}) \end{pmatrix} \sim N_2 \left(\begin{pmatrix} X_1(t_k)e^{-\alpha_1 \Delta t} \\ X_2(t_k)e^{-\alpha_2 \Delta t} \end{pmatrix}, \Sigma_{\Delta t} \right),$$

where

$$\Sigma_{\Delta t} = \begin{pmatrix} \frac{1}{2\alpha_1}(1-e^{-2\alpha_1 \Delta t}) & \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t}) \\ \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t}) & \frac{1}{2\alpha_2}(1-e^{-2\alpha_2 \Delta t}) \end{pmatrix}.$$

Figure 1 shows the simulation results of $X_1(t)$ and $X_2(t)$ from $t = 1, 2, ..., M = 1,000$ with time step $\Delta t = 1/12$, given the initial and parameter values: $X_1(1) = 2, X_2(1) = 3, \alpha_1 = 0.6, \alpha_2 = 0.06, \rho = -0.5$.

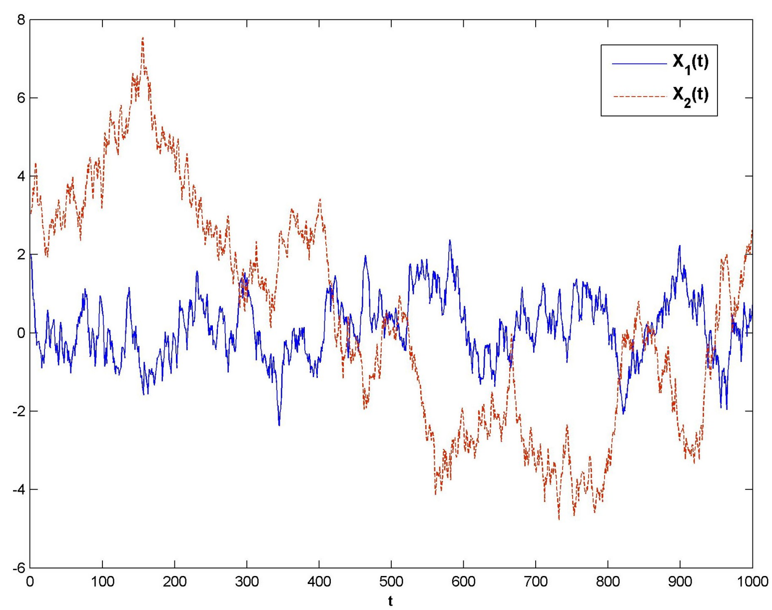

Figure 1: The simulated $X_1(t)$(solid) and $X_2(t)$(dotted) from the exact solution for $t = 1, ..., 1000, \Delta t = 1/12$ with $\beta = 0.04, \alpha_1 = 0.6, \alpha_2 = 0.06, \sigma_1 = 0.6, \sigma_2 = 0.4, \rho = -0.5, \gamma_1 = 0$ and $\gamma_2 = 0$.

4 Estimation Method

In this section, we discuss main MCMC algorithms which are based on a Bayesian approach and will be our core methodology for estimation. Given observed data $y = (y_1, \ldots, y_N)$, an unknown parameter vector θ of the underlying model in the Bayesian paradigm is treated as a random variable with some prior beliefs. This is contrary to the classical approach that θ is supposed to be a fixed quantity. The heart of the Bayesian approach is Bayes theorem. Here we assume that θ is continuous, and initially the joint distribution of y and θ can be written as

$$f(y, \theta) = f(y|\theta)f(\theta) = f(\theta|y)f(y).$$

Hence, it follows that the posterior distribution of θ conditional on y is

$$\begin{aligned} f(\theta|y) &= \frac{f(y|\theta)f(\theta)}{f(y)} = \frac{f(y|\theta)f(\theta)}{\int f(y|\theta)f(\theta)d\theta} \\ f(\theta|y) &\propto f(y|\theta)f(\theta), \end{aligned} \tag{7}$$

where $f(y|\theta)$ is the likelihood of the data, $f(\theta)$ is the prior density and $\int f(y|\theta)f(\theta)d\theta$ is the normalising constant satisfying $\int f(\theta|y)d\theta = 1$.

4.1 Joint Posterior Distribution of the Cairns Bond Price

Initially, we will look at a mean-reverting bivariate vector autoregressive VAR(1) model since it can be thought of as a model for the latent variables in discrete-time version. The derived likelihood will be part of the full posterior distribution of the Cairns bond price which is required for the MCMC simulation.

Proposition 2 *[Likelihood of the Bivariate Normal VAR(1) Model] Suppose that* $X(t) = (X_1(t),\ X_2(t))'$ *follows*

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \begin{pmatrix} X_1(t-1) - \mu_1 \\ X_2(t-1) - \mu_2 \end{pmatrix} + \begin{pmatrix} Z_1(t) \\ Z_2(t) \end{pmatrix}, \quad or \tag{8}$$

$$X(t) = \mu + K(X(t-1) - \mu) + Z(t),$$

where $Z(t) \sim N_2(0, \Sigma)$ *and* $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$ *is a covariance matrix.*

Then, the likelihood of $\boldsymbol{X} = \{X(1), X(2), \ldots, X(M)\}$*, given the parameter vector* $\theta = (\mu, K, \Sigma)$*, is*

$$
\begin{aligned}
f(\boldsymbol{X}|\mu, K, \Sigma) &= L_1(X(1)|\theta) \cdot L_2(X(2), \ldots, X(M)|\theta, X(1)) \\
&= (2\pi)^{-1}|\Omega_x|^{-1/2} \exp\left\{-\frac{1}{2}(X(1)-\mu)'\Omega_x^{-1}(X(1)-\mu)\right\} \cdot \\
&\quad (2\pi)^{-(M-1)}|\Sigma|^{-\frac{1}{2}(M-1)} \exp\left\{-\frac{1}{2}\sum_{t=2}^{M} \hat{Z}(t)'\Sigma^{-1}\hat{Z}(t)\right\}, \qquad (9)
\end{aligned}
$$

where $\hat{Z}(t) = X(t) - \mu - K(X(t-1) - \mu)$ *and*

$$
\Omega_x = \begin{pmatrix} \frac{\sigma_{11}}{1-k_1^2} & \frac{\sigma_{12}}{1-k_1k_2} \\ \frac{\sigma_{12}}{1-k_1k_2} & \frac{\sigma_{22}}{1-k_2^2} \end{pmatrix}.
$$

Providing the framework in (1), the full joint posterior distribution of the Cairns bond price can be written as

$$
\begin{aligned}
f(\Theta|\mathbb{P}) \propto f(\mathbb{P}, \Theta) &= f(\mathbb{P}|\Theta)f(\Theta) \\
&= \prod_{t=1}^{M} f_1(\mathbb{P}|X(t), \theta) \\
&\quad \times \prod_{t=2}^{M} f_{2c}(X(t)|X(t-1), \theta_2) \times f_{2u}(X(1)|\theta) \times f_0(\theta), \quad (10)
\end{aligned}
$$

where $\mathbb{P}$ is all bond price data, f_1, f_{2u} and f_{2c} are the normal density functions, f_0 is the prior density function and $\Theta = (\mathbf{X}, \theta)$, where $\mathbf{X} = \{X(1), X(2), ..., X(M)\}, \theta = \theta_1 \cup \theta_2$, where $\theta_1 = \{\beta, \sigma_1, \sigma_2, \sigma_\varepsilon\}, \theta_2 = \{\alpha_1, \alpha_2, \rho, \gamma_1, \gamma_2\}$.

We partition the parameter vector θ into θ_1 and θ_2 to make clear which group of parameters is in the dynamic of the latent variables. In addition, it can be observed that the full posterior distribution (10) consists of four main components:

- the likelihood of the pricing data; f_1 (measurement equation),
- the conditional likelihood of the latent variables $X(t)$; f_{2c} (transition equation),
- the unconditional likelihood of $X(1)$; $f_{2u}(X(1)|\theta)$, and
- the prior density of of model parameters $f_0(\theta)$.

Since the likelihood f_{2c} and $f_{2u}(X(1)|\theta)$ are given in Proposition 2, the log posterior eventually is

$$
F(\Theta) = \log f(\Theta|\mathbb{P}) = k + \log f(\mathbb{P}, \Theta), \qquad (11)
$$

where k is a constant and

$$\begin{aligned}
\log f(\mathbb{P},\Theta) &= \sum_{t=1}^{M}\log f_1(\mathbb{P}|X(t),\theta) \\
&\quad +\sum_{t=2}^{M}\log f_{2c}(X(t)|X(t-1),\theta_2)+\log f_{2u}(X(1)|\theta)+\log f_0(\theta) \\
&= \sum_{t=1}^{M}\left\{-\frac{N_t}{2}\log(2\pi\sigma_\varepsilon^2)-\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{N_t}\left(P(t,\tau_{tj})-C(\tau_{tj};X(t),\theta)\right)^2\right\} \\
&\quad -(M-1)\log(2\pi)-\frac{(M-1)}{2}\log|\Sigma|-\frac{1}{2}\sum_{t=2}^{M}\hat{Z}(t)'\Sigma^{-1}\hat{Z}(t) \\
&\quad -\log(2\pi)-\frac{1}{2}\log|\Omega_x|-\frac{1}{2}(X(1)-\gamma)'\Omega_x^{-1}(X(1)-\gamma)+\log f_0(\theta)
\end{aligned}$$

$$\begin{aligned}
\log f(\mathbb{P},\Theta) &= -\frac{MN_t}{2}\log(2\pi\sigma_\varepsilon^2)-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{M}\sum_{j=1}^{N_t}\left(P(t,\tau_{tj})-C(\tau_{tj};X(t),\theta)\right)^2 \\
&\quad -M\log(2\pi)-\frac{1}{2}\log|\Omega_x|-\frac{(M-1)}{2}\log|\Sigma| \\
&\quad -\frac{1}{2}(X(1)-\gamma)'\Omega_x^{-1}(X(1)-\gamma)-\frac{1}{2}\sum_{t=2}^{M}\hat{Z}(t)'\Sigma^{-1}\hat{Z}(t)+\log f_0(\theta),
\end{aligned}$$

where $f_0(\theta)$ is the prior, $C(\tau_{tj};X(t),\theta)$ is theoretical price, $\mathbb{P}$ is all bond price data and $P(t,\tau_{tj})$ is observed price at time t for the maturity τ_{tj} such that

$$P(t,\tau_{tj})\sim N(C(\tau_{tj};X(t),\theta),\sigma_\varepsilon^2),$$

$$\hat{Z}(t)=X(t)-\gamma-K(X(t-1)-\gamma),$$

$$\Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}=\begin{pmatrix}\frac{1}{2\alpha_1}(1-e^{-2\alpha_1\Delta t}) & \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t})\\ \frac{\rho}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t}) & \frac{1}{2\alpha_2}(1-e^{-2\alpha_2\Delta t})\end{pmatrix},$$

$$\Omega_x=\begin{pmatrix}\frac{\sigma_{11}}{1-k_1^2} & \frac{\sigma_{12}}{1-k_1k_2}\\ \frac{\sigma_{21}}{1-k_1k_2} & \frac{\sigma_{22}}{1-k_2^2}\end{pmatrix},K=\begin{pmatrix}e^{-\alpha_1\Delta t} & 0\\ 0 & e^{-\alpha_2\Delta t}\end{pmatrix}.$$

4.2 Markov Chain Monte Carlo (MCMC)

According to (7), one can see that the posterior distribution for a complex model frequently cannot be analytically obtained in a closed-form and hence this was a crucial shortfall for implementing the Bayesian estimation in the past. Fortunately, the development of MCMC methods in recent decades has enabled us to deal with such problems by achieving a posterior distribution by simulation. Here we describe the Metropolis-Hastings (MH) algorithms that will be used for calibrating the Cairns model.

The MH algorithm is a popular MCMC updating scheme since it can eliminate the difficulty of drawing a sample from the full posterior distribution. Generally, the algorithm comprises two main steps. First, a candidate point is drawn from an arbitrary proposal distribution. Second, the candidate point is then used for calculating an acceptance probability in order to decide a movement of the chain. If the candidate point is rejected, the chain remains the previous value, otherwise it moves to the next state. An appropriate choice of the proposal distributions is therefore crucial for succeeding in implementing the MH algorithm.

With respect to the estimation framework (1), suppose that $\mathbb{P}$ represents all bond price data which are generated from the Cairns term structure model with an unknown parameter and latent variable vector $\Theta = (\Theta_1, \ldots, \Theta_d)$. Then, we know from (11) that the posterior distribution of each element of Θ, denoted as Θ_i, is achieved at

$$g(\Theta_i|\mathbb{P}, \Theta_{-i}) \quad \propto \quad \exp(\log f(\mathbb{P}, \Theta_i|\Theta_{-i})), \tag{12}$$

where Θ_{-i} is a vector of all model parameters and latent variables excluding Θ_i. Remark that the joint density $\log f(\mathbb{P}, \Theta)$ includes $f_0(\theta)$, the prior of model parameters, in which we let

$$f_0(\theta) = f_0(\beta)f_0(\alpha_1)f_0(\alpha_2)f_0(\sigma_1)f_0(\sigma_2)f_0(\rho)f_0(\gamma_1)f_0(\gamma_2)$$

such that

$$\begin{aligned} f_0(\beta), f_0(\alpha_1), f_0(\alpha_2), f_0(\sigma_1), f_0(\sigma_2), &\sim \Gamma(0.01, 100), \\ f_0(\rho) &\sim U[-1, 1], \\ f_0(\gamma_1), f_0(\gamma_2) &\sim N(0, 1.0 \times 10^5). \end{aligned} \tag{13}$$

In this case, a gamma prior is assigned to the non-negative parameters and a uniform prior with values in the range $[-1, 1]$ to the correlation parameter ρ. Note that, for simplicity, σ_z does not appear here since we fix it as a constant for this time. Further, a candidate point $y^*_{\Theta_i}$, for each Θ_i, will be drawn from the normal proposal distribution q_{Θ_i}, with mean depending on its previous value with some constant variance. The detailed procedure of two main MH algorithms for simulating the model parameters and latent variables from the Cairns bond posterior in (11) are outlined as follows.

Algorithm 1. **[Standard Metropolis-Hastings Algorithm]**

1. Initialise the chain at $j = 1$ and start the iteration at $j = 2$ and set $i = 1$.
2. Generate a candidate point $y^*_{\Theta_i}$ from the proposal
$$q_{\Theta_i} \sim N(\Theta_i(j-1), vol^2_{\Theta_i}), \tag{14}$$
where $vol^2_{\Theta_i}$ is some constant volatility.
3. If
$$\begin{cases} y^*_{\Theta_i} < 0, \text{ for } \Theta_i = \beta, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \\ |y^*_{\Theta_i}| > 1, \text{ for } \Theta_i = \rho, \end{cases}$$
set $\Theta_i(j) = \Theta_i(j-1)$, and then go to step 5, otherwise go to step 4.
4. Then,
   - Generate U from $\sim U(0,1)$.
   - Compute the acceptance probability $\eta_{\Theta_i}(\Theta_i(j-1), y^*_{\Theta_i}) =$
$$\min\left\{1, \frac{g(y^*_{\Theta_i}|\mathbb{P}, \Theta_{-i}) \cdot q_{\Theta_i}(\Theta_i(j-1)|y^*_{\Theta_i})}{g(\Theta_i(j-1)|\mathbb{P}, \Theta_{-i}) \cdot q_{\Theta_i}(y^*_{\Theta_i}|\Theta_i(j-1))}\right\}. \tag{15}$$
   - If $U \leq \eta_{\Theta_i}(\Theta_i(j-1), y^*_{\Theta_i})$, then $\Theta_i(j) = y^*_{\Theta_i}$; otherwise $\Theta_i(j) = \Theta_i(j-1)$.
5. Set $i = i + 1$ and repeat step 2 to 4 until $i = d$ (the last element of Θ). For the same iteration j, the recent value of the Θ_i, which is already updated, will be used, rather than its value from the previous step $j - 1$.
6. Set $j = j+1, i = 1$ and repeat step 2 to 5 until the last iteration (convergence).

Algorithm 2. **[Adaptive Metropolis-Hastings Algorithm]**

Referring to Haario et al. (2005), another idea for improving the proposal distribution is to use the empirical variance computed from its previous sample path (recent n_1 values), after using a constant variance for some initial iteration n_0. More precisely, the proposal variance is given by

$$vol^2_{\Theta_i}(j) = \begin{cases} Var_{n_0} & j \leq n_0 \\ k \cdot Var(\Theta_i(j-n_1), \ldots, \Theta_i(j-1)) & j > n_0, \end{cases} \tag{16}$$

where Var_{n_0} is some constant value, $Var(\cdot)$ is the sample variance of values in the argument and k is a scaling number. For updating several parameters, the covariance matrix can be used in order to draw a set of candidate points with correlation, i.e.

$$Cov_{\Theta_i}(j) = \begin{cases} Cov_{n_0} & j \leq n_0 \\ k \cdot Cov(\Theta_i(j-n_1), \ldots, \Theta_i(j-1)) & j > n_0, \end{cases} \tag{17}$$

where Cov_{n_0} is an initial covariance matrix, $Cov(\cdot)$ is the sample covariance of a series of values in the argument and k is a scaling number.

5 Estimation Results using Standard MH Algorithm

In this section, we are implementing the standard MH algorithm (Algorithm 1) to the two-factor Cairns term structure model in which the sampling is facilitated by the normal proposal distribution with a "constant" variance using 100 months of the simulated data (with the latent variables from time $t = 1$ to 100 in Figure 1 and model parameter values as specified in Section 3). With this dataset, 200 latent variables and 8 model parameters are being estimated. Regarding the number of maturities of the bond prices, we use 20 constant maturities ($N_t = 20$, for all t): 0.25, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 12.5, 15.0, 17.5, 20.0, 22.5, 25.0, 27.5, 30.0 years. We simulate each parameter for 15,000 iterations starting from the true values in order to shorten the runtime and particularly concentrate on the convergence assessment. We now demonstrate the results in a series of figures, and comment as follows:

- Figure 2 illustrates the sample paths of model parameters, where γ_1 and γ_2 are fixed to be zero and each parameter is updated individually. From the figure, we can observe that all chains encompass reasonably well the true values but with different velocities of convergence. Compared with the others, σ_1 tends to converge slowest whereas α_1 is found to be most stable and has long excursions in our simulation. Despite poor convergence of σ_1, the means of all parameters are clearly close to the true values (within one standard deviation) as can be seen from the summary statistics in Table 1. Table 1 also shows the constant normal proposal standard deviations used for the simulation in which these values are discovered to be suitable for this dataset after extensive tuning up. It should be mentioned that for the model parameters we noticed that the constant proposal standard deviations should be as close to the posterior standard deviations as possible, otherwise the chains will easily drift away since these parameters are too sensitive to be sampled from the proposal with high and low standard deviations. For the MH acceptance rates, we obtain the rates varying from 8.0% to 22.0%.

- Figure 3 shows the selected sample paths of latent variables $X_1(t)$ and $X_2(t)$, for $t = 1, 20, 40, 60, 80$ and 100, where they are updated as a pair for each t in which we use the constant normal proposal standard deviations 0.055 and 0.04 for all t since it is not practical to tune up the proposal distribution for each t individually. Unlike the model parameters, these values are approximately 1.5 to 2.5 times higher than the resulting posterior standard deviations of all $X_1(t)$ and $X_2(t)$ (the posterior standard deviations of $X_1(t)$ range from 0.020 to 0.036 and of $X_2(t)$ from 0.013 to 0.019). According to the MH acceptance rates, the pairs of $X_1(t)$ and $X_2(t)$ are accepted with the rates between 5.12% to 10.03% which is rather low. It is noticed that the latent variables seem to be least sensitive quantities in the Cairns term structure model and hence they are easier controlled than the model parameters. As to the results, the chains strongly converge for almost all t except the unconditional latent variable $X_1(1)$ that is found to be most negatively correlated with σ_1.

- Figure 4 provides the 95% credible interval constructed from the sample paths of latent variables. As can be observed, all include the true values fairly well.

- In Figure 5, the scatter plots for the model parameters are shown and we also consider their cross-correlations in Table 2 in order to further our analysis in the interactions among all unknown quantities in the Cairns model. Among the model parameters, it turns out that there exist strong positive correlations between α_1 and σ_1 and between α_2 and σ_2, while the former is in the lesser degree. Furthermore, β is strongly negatively correlated with a pair of (α_2, σ_2), whereas ρ is least correlated to all other parameters.

  Between the model parameters and latent variables, it can be found that α_2 and σ_2 have moderate negative correlation with almost all $X_2(t)$. While α_1 is hardly correlated to $X_1(t)$, σ_1 is found to be most correlated to $X_1(t)$ (can be either positive or negative, with very strong negative correlation with the first $X_1(1), X_1(2), \ldots, X_1(6)$). Furthermore, ρ has consistently positive correlation (around 0.3 to 0.5) to all $X_1(t)$ but almost zero correlation with all $X_2(t)$.

  Among all the latent variables $X_1(t)$ and $X_2(t)$ (not shown in the figure), there is no evidence of high correlations except in the group of $X_1(t)$ for $t = 1, 2, \ldots, 6$, in which they are highly positively correlated (around 0.48 to 0.83).

- The components of the log-likelihood are also monitored during the simulation as illustrated in Figure 6. As can be seen, although the log-likelihood of pricing data constitutes of the largest part of the total log-likelihood, it is most stable and hence we may infer that the variations influencing the overall MH sampling are actually from the log-likelihood of latent variables and priors. The total log-likelihood of all components has a very similar picture to the log-likelihood of pricing data but has not been shown here.

We first summarise for this section that tuning for the suitable constant variances of the proposal distribution plays a substantial role in order to obtain good MH convergence. Due to high interaction complexity among the unknown quantities in this model, too small and too large variances of only one parameter can easily make the chain either drift away or diverge. Once one parameter, for example α_2, starts shifting away from the true values to some extent, so do the others (to which α_2 is highly correlated). The total variation of the simulation is hard to be controlled by using the constant proposal variance and much relies on the variation of the log-likelihood of latent variables. In general, the parameters and latent variables in the same term of the function $H(u, x)$ in (3) tend to be correlated to one another in some way.

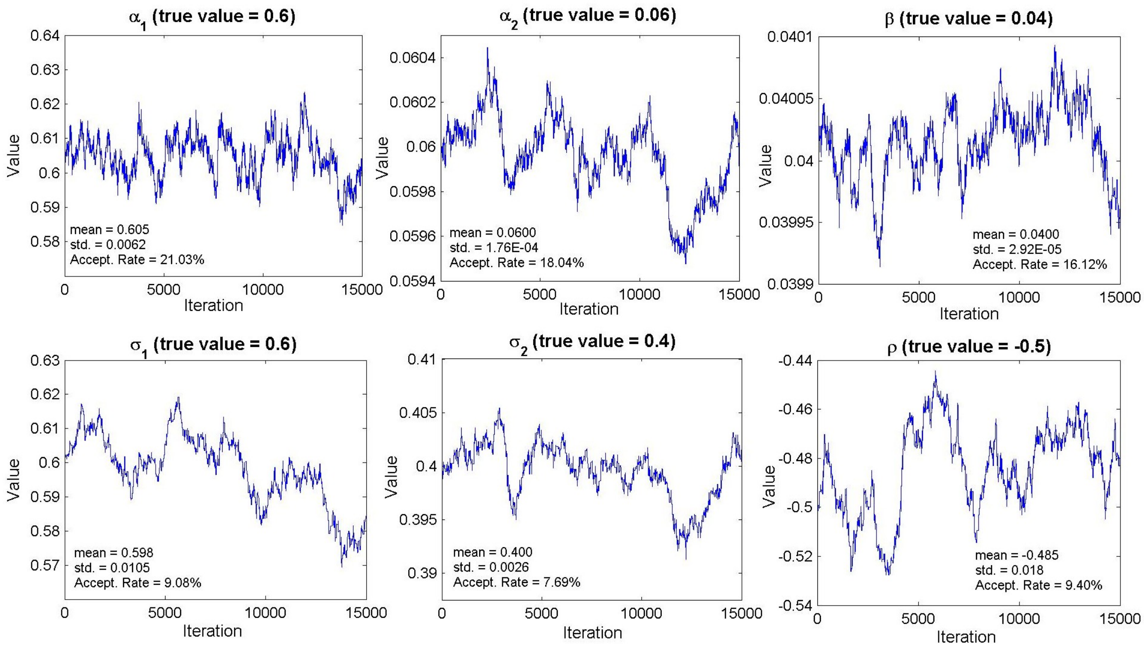

Figure 2: Sample paths of model parameters of the two-factor Cairns term structure model using the standard MH algorithm with constant normal proposal variance.

	(True value)	Mean	Std.	95% Credible Interval	Proposal std.	Acceptance rate
β	(0.04)	0.0400	0.00003	(0.03995, 0.04007)	0.00003	16.12%
α_1	(0.6)	0.605	0.0062	(0.5921, 0.6158)	0.0060	21.03%
α_2	(0.06)	0.0600	0.00018	(0.05956, 0.06026)	0.00010	18.04%
σ_1	(0.6)	0.598	0.0104	(0.5756, 0.6150)	0.0100	9.08%
σ_2	(0.4)	0.400	0.0026	(0.3935, 0.4038)	0.0050	7.69%
ρ	(-0.5)	-0.485	0.0180	(-0.5218, -0.4555)	0.0200	9.40%

Table 1: Summary statistics of parameter posterior estimates of the two-factor Cairns term structure model using the standard MH algorithm with constant normal proposal variance (15,000 iterations).

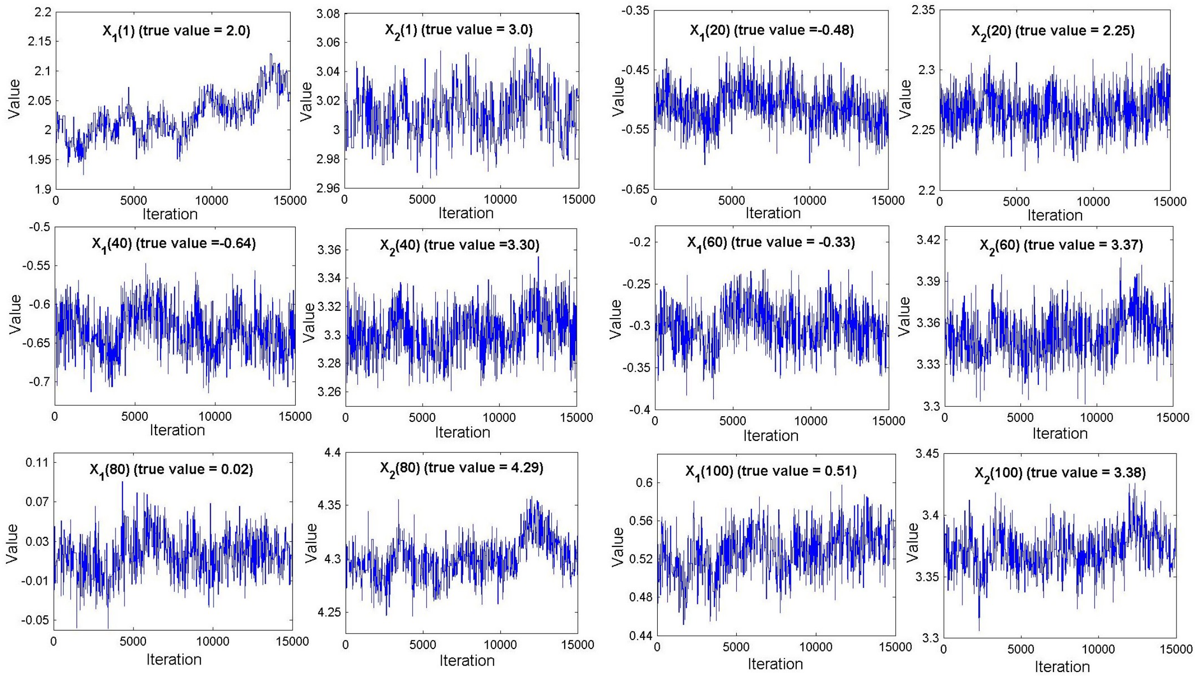

Figure 3: Sample paths of latent variables (for $t = 1, 20, 40, 60, 80$ and 100) of the two-factor Cairns term structure model using the standard MH algorithm with constant normal proposal variance.

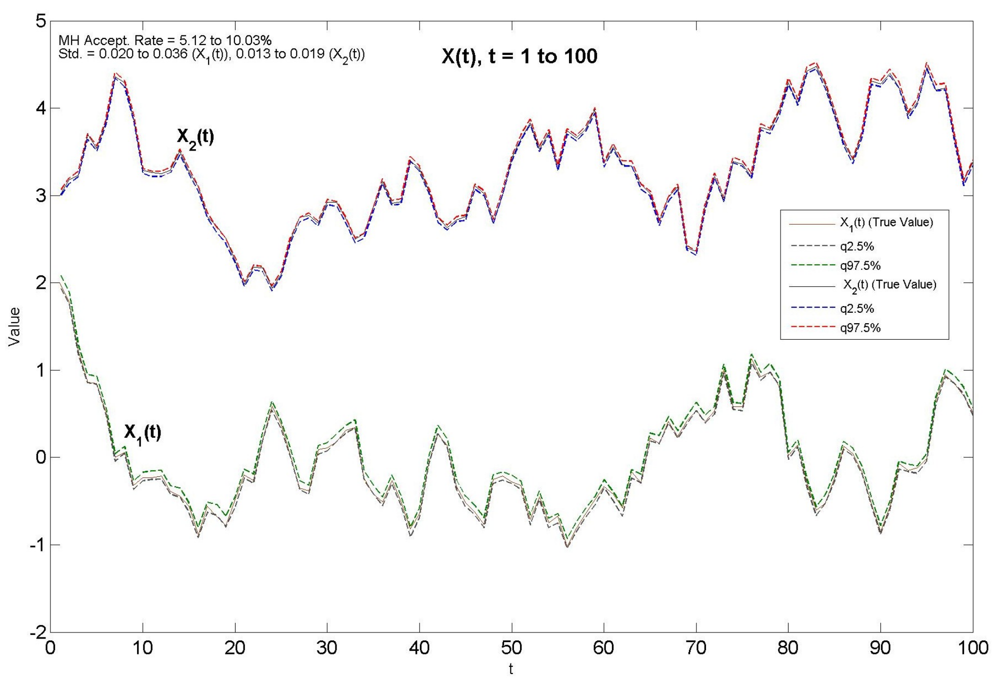

Figure 4: Plots of 95% credible interval constructed from the sample paths with the true values of $X_1(t)$ and $X_2(t)$ for $t = 1, \ldots, 100$, of the two-factor Cairns term structure model using the standard MH algorithm with constant normal proposal variance.

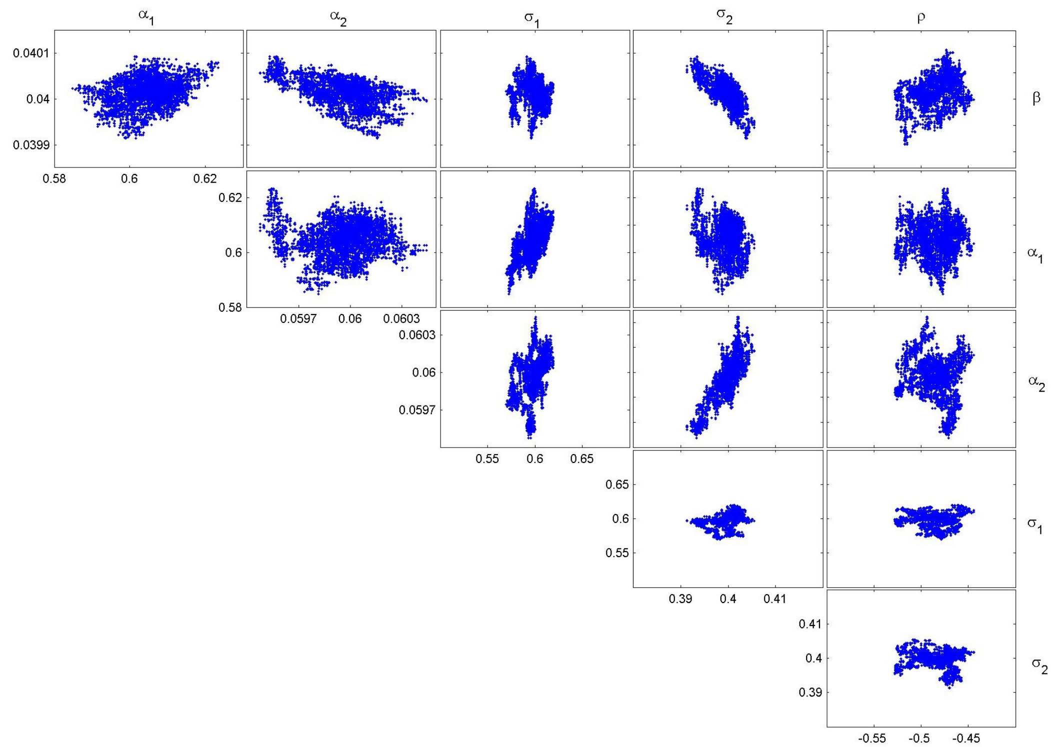

Figure 5: Scatter plots of model parameters of the two-factor Cairns term structure model using the standard MH algorithm with constant normal proposal variance.

	β	α_1	α_2	σ_1	σ_2	ρ	$X_1(t)$	$X_2(t)$
β	1.00	0.28	-0.54	-0.08	-0.71	0.37	0.00 to 0.30	-0.37 to 0.35
α_1		1.00	0.01	0.57	-0.19	0.02	-0.42 to 0.22	-0.42 to 0.32
α_2			1.00	0.39	0.85	-0.16	-0.38 to 0.21	-0.68 to 0.02
σ_1				1.00	0.31	-0.02	-0.86 to 0.52	-0.86 to 0.02
σ_2					1.00	-0.18	-0.30 to 0.23	-0.68 to 0.17
ρ						1.00	0.29 to 0.55	-0.22 to 0.31

Table 2: Correlation matrix of model parameters and latent variables of the two-factor Cairns term structure model using the MH algorithm with constant normal proposal variance.

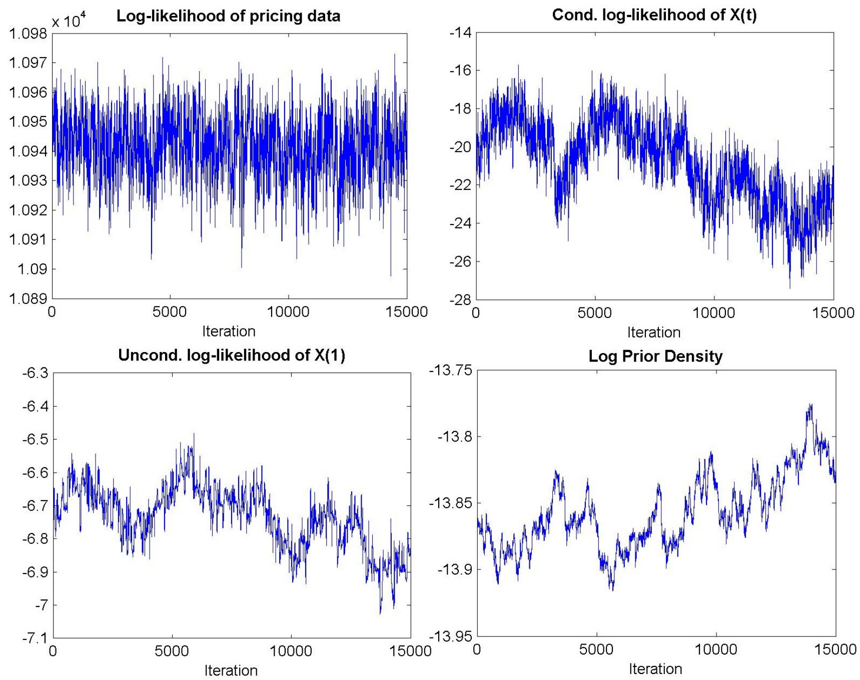

Figure 6: Log posterior components (referring to equation 5.9): log-likelihood of pricing data (top, left), conditional and unconditional log-likelihood of latent variables (top, right and bottom, left) and log prior density (bottom, right) using the MH algorithm with constant normal proposal variance.

6 Estimation Results using Adaptive MH Algorithm with Some Improvement

Although the chains according to the results in the previous section tend to converge slowly, they result in reasonable inference for all the model parameters and latent variables. Nonetheless, some improvements are yet required in order to implement the MH algorithm with more complex data. Having the standard MH algorithm with constant normal proposal variance (Algorithm 1) as a base case, we here consider to improve the proposal distribution using the adaptive Metropolis-Hastings algorithm (Algorithm 2) associated with a blocking strategy, reparameterising the log posterior of Cairns bond prices and re-evaluating the prior distributions. All of these will be first described below, before we move to look at the results.

6.1 Reparameterising

From Table 2, there is evidence of strong correlation between some parameters, particularly σ_1 and σ_2 where they are highly correlated to $X_1(t)$ and $X_2(t)$ for all t

respectively. In the bond price formula, σ_1 and σ_2 are actually the local volatilities of all the latent variables $X_1(t)$ and $X_2(t)$. Thus, here we attempt to eliminate such correlations by re-parameterisation.

Let $Y(t) = (Y_1(t), Y_2(t))'$, where $Y_1(t) = \sigma_1 X_1(t)$ and $Y_2(t) = \sigma_2 X_2(t)$ and $\gamma_y = (\gamma_{y_1}, \gamma_{y_2})'$, where $\gamma_{y_1} = \sigma_1\gamma_1$ and $\gamma_{y_2} = \sigma_2\gamma_2$. Then, the log posterior in (11) can be re-written as

$$\begin{aligned}
\log f(\Theta|\mathbb{P}) \propto & -\frac{MN_t}{2}\log(2\pi\sigma_\varepsilon^2) - \frac{1}{2\sigma_\varepsilon^2}\sum_{t=1}^{M}\sum_{j=1}^{N_t}\left(P(t,\tau_{tj}) - C_Y(\tau_{tj}; Y(t), \theta)\right)^2 \\
& -M\log(2\pi) - \frac{1}{2}\log|\Omega_y| - \frac{(M-1)}{2}\log|\Sigma_Y| \\
& -\frac{1}{2}(Y(1)-\gamma_y)'\Omega_y^{-1}(Y(1)-\gamma_y) \\
& -\frac{1}{2}\sum_{t=2}^{M}\hat{Z}_Y(t)'\Sigma^{-1}\hat{Z}_Y(t) + \log f_0(\theta), \qquad (18)
\end{aligned}$$

where

$$C_Y(\tau, y, \theta) = \frac{\int_\tau^\infty H(u,y)du}{\int_0^\infty H(u,y)du},$$

$$H(u,y) = \exp\left[-\beta u + y_1\, e^{-\alpha_1 u} + y_2\, e^{-\alpha_2 u} - \frac{1}{2}\sum_{i,j=1}^{2}\frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i+\alpha_j}e^{-(\alpha_i+\alpha_j)u}\right]$$

$$\hat{Z}_Y(t) = Y(t) - \gamma_y - K(Y(t-1) - \gamma_y),$$

$$\Sigma_Y = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \frac{\sigma_1^2}{2\alpha_1}(1-e^{-2\alpha_1\Delta t}) & \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t}) \\ \frac{\rho\sigma_1\sigma_2}{\alpha_1+\alpha_2}(1-e^{-(\alpha_1+\alpha_2)\Delta t}) & \frac{\sigma_2^2}{2\alpha_2}(1-e^{-2\alpha_2\Delta t}) \end{pmatrix},$$

$$\Omega_y = \begin{pmatrix} \frac{\sigma_{11}}{1-k_1^2} & \frac{\sigma_{12}}{1-k_1k_2} \\ \frac{\sigma_{21}}{1-k_1k_2} & \frac{\sigma_{22}}{1-k_2^2} \end{pmatrix}, K = \begin{pmatrix} e^{-\alpha_1\Delta t} & 0 \\ 0 & e^{-\alpha_2\Delta t} \end{pmatrix}.$$

6.2 Adaptive MH Algorithm and Blocking Strategy

Previously, we made an observation that the normal proposal distribution with constant variance may not facilitate the MH algorithm for the model parameters well. Moreover, it did spend long time to explore suitable variance values. As already described in an earlier section, one way to improve this is to use the adaptive Metropolis-Hastings algorithm (Algorithm 2) in which it will be here associated with a blocking strategy. First of all, based on some evidence of the correlation structure, we group the model parameters and latent variables as follows.

I. α_1, σ_1 and ρ.

II. α_2, σ_2 and β.

III. γ_1 and γ_2.

IV. $X_1(t)$ and $X_2(t)$ for each t.

In each group, the parameters or latent variables will be updated together. For our simulation, candidate points for the parameters in group I and II will be sampled from a multivariate normal distribution where means depend on their previous values with a covariance matrix computed from most recent previous 200 values of sample paths of the parameters in each group (an arbitrary fixed covariance matrix will be used for the first 200 iterations). For example, at 601st iteration, the covariance matrix of the proposal distribution is of sample paths from 401st to 600th iteration, at 602nd iteration from 402nd to 601st iteration and so on. Similarly, candidate points for the parameters and latent variables in group III and IV, will be generated by the same way as the first two groups except we will impose zero correlation to the proposal covariance matrix at all time.

6.3 Re-evaluating the Priors

Finally, we allow priors of the parameters $\beta, \alpha_1, \alpha_2, \sigma_1$ and σ_2 to be slightly more informative. We set their means with reference to the posterior means of the earlier simulations and reduce the coefficient of variation to around 5.3 (previously a $\Gamma(0.01, 100)$ prior, which provides a mean of 1.0 and coefficient of variation of 10, was assigned to all of these parameters). Specifically, the priors of all the model parameters now become

$$\begin{aligned}
f_0(\beta) &\sim \Gamma(0.036, 1.13), \\
f_0(\alpha_1), f_0(\sigma_1) &\sim \Gamma(0.036, 16.67), \\
f_0(\alpha_2) &\sim \Gamma(0.036, 1.67), \\
f_0(\sigma_2) &\sim \Gamma(0.036, 11.25), \\
f_0(\rho) &\sim U[-1, 1], \\
f_0(\gamma_1), f_0(\gamma_2) &\sim N(0, 1.0 \times 10^5).
\end{aligned} \tag{19}$$

6.4 Estimation Results with Improvements

To begin with, we provide a narrative summary of the effects for each improvement from several simulations compared with the base case as follows.

- We first started using adaptive proposal distributions for each parameter individually, but it was not possible to notice any distinct difference according to the results.

- Then, the reparameterisation for the latent variables $X_1(t)$ and $X_2(t)$ (i.e. define $Y_1(t) = \sigma_1 X_1(t), Y_2 = \sigma_2 X_2(t)$) was therefore considered. In effect, it turned out that the convergence of σ_1 was clearly improved but σ_2 still converged slower than expected.
- Next, we re-evaluated the priors for $\beta, \alpha_1, \alpha_2, \sigma_1, \sigma_2$. That is, their prior means were shifted from 1.0 to the values with respect to their posterior means with variances of 10.0. However, any difference was hard to notice. Consequently, we also decreased the coefficient variations of the priors to around 5.3 but once again distinct improvement could not be observed.
- Eventually, we attempted to improve the convergence by incorporating the blocking strategy. In effect, we found that the convergence of σ_2 was evidently better.

With the evidences mentioned above, we therefore exclude the re-evaluation for the priors as an improvement. In the following results, we run MCMC simulation for 20,000 iterations for each chain using the adaptive MH algorithm with the reparameterised log posterior in (18) and a blocking strategy where γ_{y_1} and γ_{y_2} are now unrestricted and then estimated.

- Figure 7 shows the resulting sample paths of model parameters. We can easily see that the convergence of the parameters σ_1, α_2 and β are all clearly improved (compared with Figure 2). Parameter α_1 still converges well same as the previous result, while ρ is not significantly different. Furthermore, it can also be noticed that the posterior standard deviations (Table 3) are relatively higher (about twice than before). In addition, γ_{y_1} and γ_{y_2} undoubtedly get stationary although the range of γ_{y_2} is rather wide.
- Figure 8 demonstrates the sample paths of latent variables $Y_1(t)$ and $Y_2(t)$, for $t = 1, 20, 40, 60, 80,$ and 100. As can be noticed, the convergence of $Y_1(1)$ compared with $X_1(1)$ in Figure 3 is evidently better, whereas the remaining still converge well.
- In Figure 9, plots of the 95% credible interval constructed from the sample paths of $Y_1(t)$ and $Y_2(t)$ for all t are illustrated. Comparing to those of $X_1(t)$ and $X_2(t)$ in Figure 4, these intervals are generally wider which infers higher posterior standard deviations.
- Figure 10 provides scatter plots of the model parameters. We can observe that overall the correlation structure of the model parameters is much better than those obtained with the standard MH algorithm (Figure 5). However, we observe strongly positive correlations among all parameters in group II $(\alpha_2, \sigma_2, \beta)$. In the previous result (Table 2), although β was found to be strongly negatively correlated to α_2 and σ_2, we found that this was not always the case. Specifically, since the chains converged rather slowly, their estimated

correlations are much less reliable than those after the reparameterisation in which we can easily see that the chains generally converge much faster.

Between the model parameters and latent variables, it turns out that ρ is most likely to correlate with $Y_1(t)$ for all t. Among the latent variables themselves, there is no evidence of any strong correlation among them. Moreover, the strong negative correlations previously found for the first $X_1(1), X_1(2), \ldots, X_1(6)$ now disappear.

- Additionally, we also compare the log-likelihood components monitored during the simulation as shown in Figure 11. While the log-likelihood of pricing data roughly remains unchanged, the variations of other components are much more stable. This corresponds to the better convergence achieved for both model parameters and latent variables.

To this end, we conclude for this section that the achieved results using the adaptive MH algorithm with a blocking strategy and reparameterising the log posterior distribution were substantially improved from those with the standard MH algorithm in terms of both convergence and correlation structure. Furthermore, we observed that the chains of almost all parameters and latent variables can achieve stationarity much easier than using the standard MH algorithm. Although parameter ρ is still rather sensitive to the proposal variance and hard to converge, this is not a surprising result since ρ appears in the minor term in the bond price posterior which tends to be most difficult to be estimated accurately.

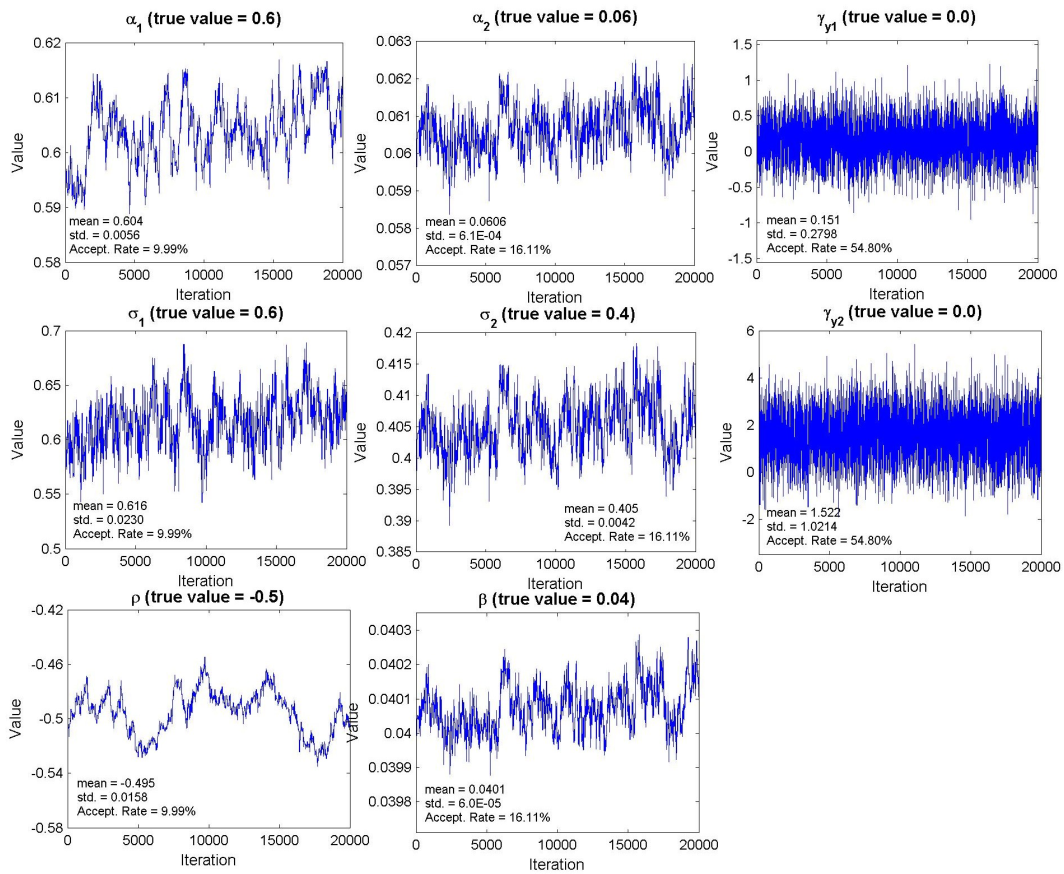

Figure 7: Sample paths of model parameters of the two-factor Cairns term structure model using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

	(True value)	Mean	Std.	95% Credible Interval	Acceptance rate	Scaling of the proposal std.
α_1	(0.6)	0.604	0.0056	(0.5924, 0.6143)	9.99%	2.0
σ_1	(0.6)	0.616	0.0230	(0.5688, 0.6616)	9.99%	2.0
ρ	(-0.5)	-0.495	0.0158	(-0.5250, -0.4673)	9.99%	2.0
α_2	(0.06)	0.0606	0.00061	(0.05958, 0.06190)	16.11%	1.4
σ_2	(0.4)	0.405	0.0042	(0.3977, 0.4138)	16.11%	1.4
β	(0.04)	0.0401	0.00006	(0.03996, 0.04021)	16.11%	1.4
γ_{y_1}	(0.0)	0.151	0.2798	(-0.3976, 0.7117)	54.80%	1.0
γ_{y_2}	(0.0)	1.522	1.0214	(-0.5068, 3.4789)	54.80%	1.0

Table 3: Summary statistics of parameter posterior estimates of the two-factor Cairns term structure model using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy (20,000 iterations).

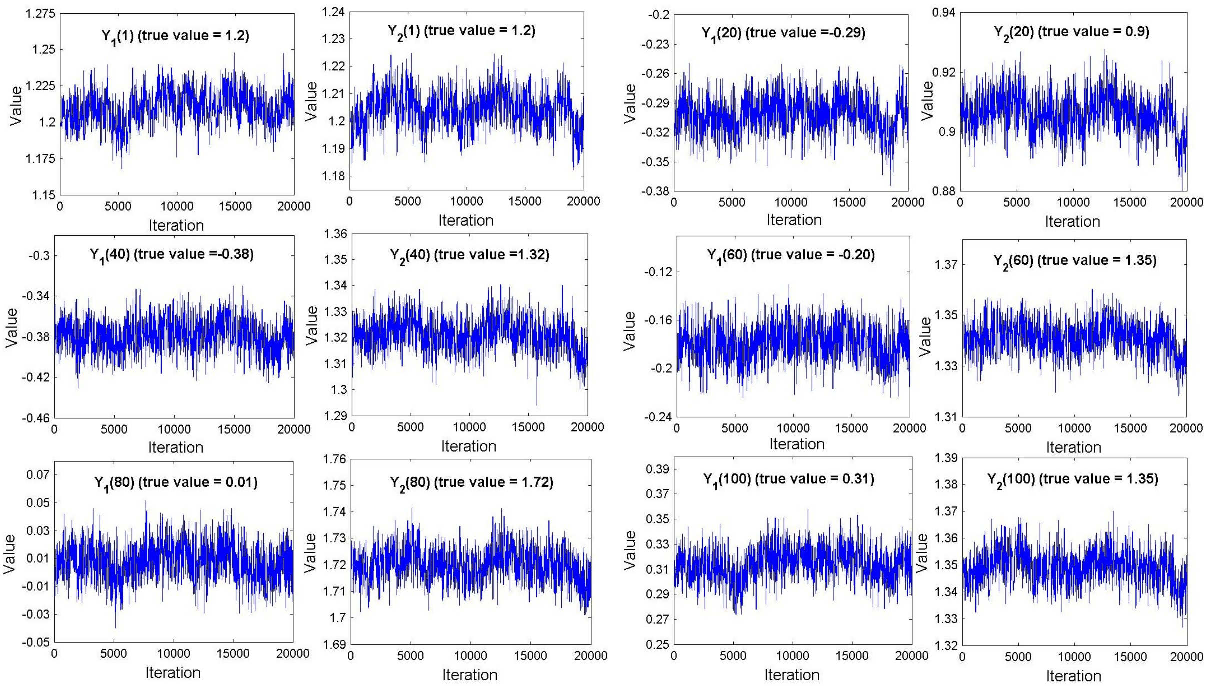

Figure 8: Sample paths of latent variables (for $t = 1, 20, 40, 60, 80$ and 100) of the two-factor Cairns term structure model using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

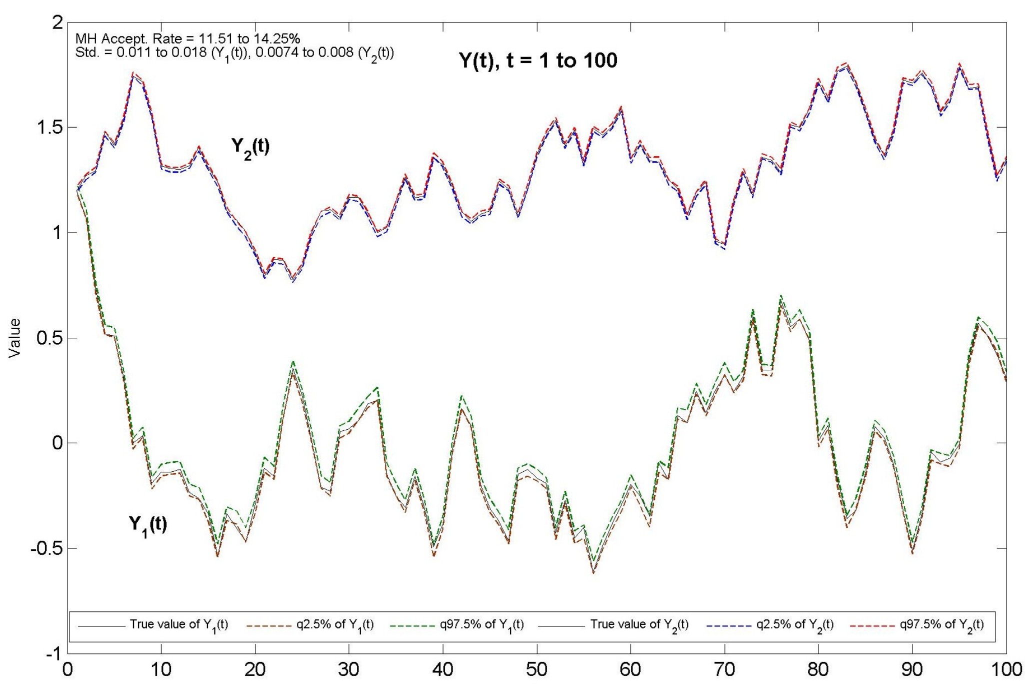

Figure 9: Plots of 95% credible interval constructed from the sample paths with the true values of $Y_1(t)$ and $Y_2(t)$ for $t = 1, \ldots, 100$, of the two-factor Cairns term structure model using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

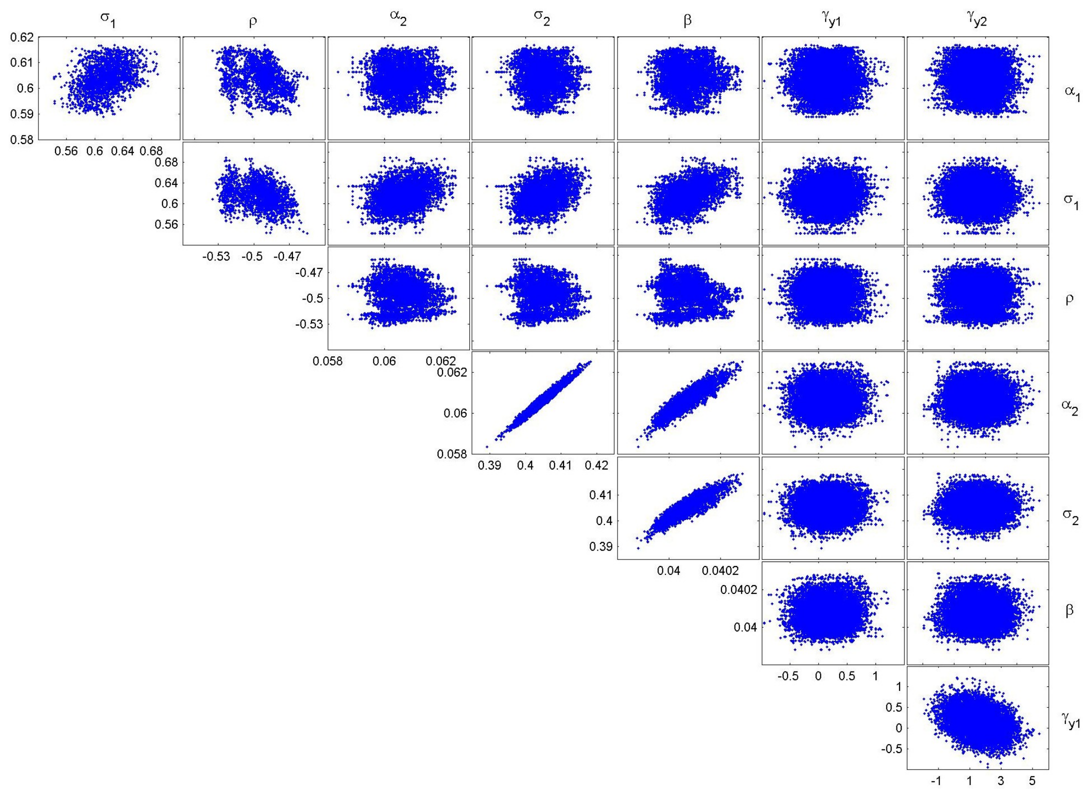

Figure 10: Scatter plots of model parameters of the two-factor Cairns term structure model using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

	α_1	σ_1	ρ	α_2	σ_2	β	γ_{y_1}	γ_{y_2}	$Y_1(t)$	$Y_2(t)$
α_1	1.00	0.37	-0.27	0.00	-0.02	0.03	0.01	-0.03	-0.23 to 0.20	-0.20 to 0.29
σ_1		1.00	-0.34	0.29	0.35	0.43	0.05	-0.01	-0.10 to 0.29	-0.22 to 0.29
ρ			1.00	-0.02	-0.05	-0.03	-0.01	0.02	0.24 to 0.44	-0.10 to 0.32
α_2				1.00	0.98	0.91	0.06	0.03	-0.11 to 0.45	-0.40 to 0.30
σ_2					1.00	0.90	0.06	0.03	-0.09 to 0.46	-0.35 to 0.31
β						1.00	0.07	0.03	-0.06 to 0.49	-0.52 to 0.37
γ_{y_1}							1.00	-0.32	-0.03 to 0.07	-0.07 to 0.045
γ_{y_2}								1.00	-0.04 to 0.05	-0.04 to 0.04

Table 4: Correlation matrix of the simulation using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

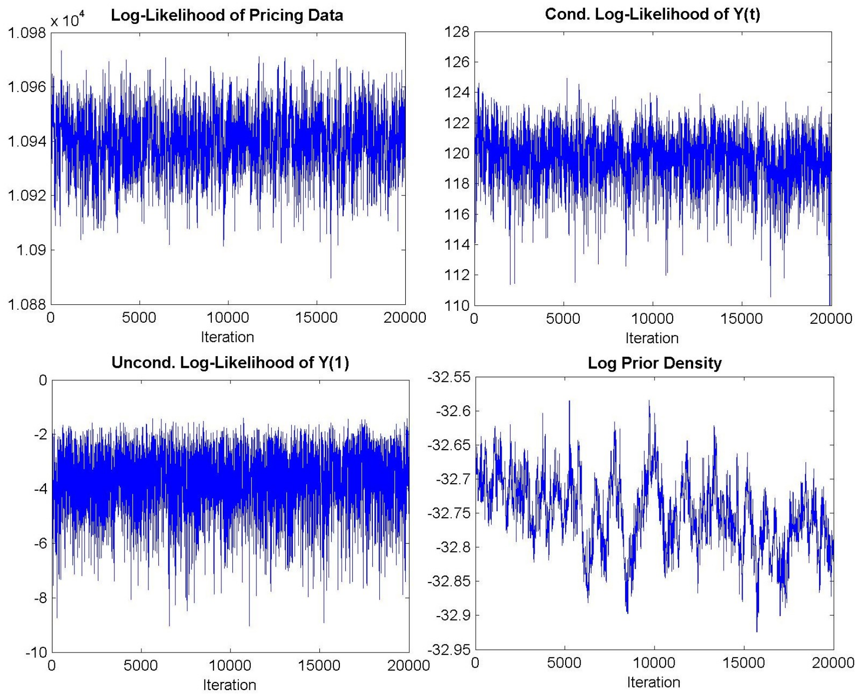

Figure 11: Log posterior components (referring to equation 5.9): log-likelihood of pricing data (top, left), conditional and unconditional log-likelihood of latent variables (top, right and bottom, left) and log prior density (bottom, right) using the adaptive MH algorithm with the reparameterised posterior and a blocking strategy.

7 Conclusions

In this paper, we have developed MCMC algorithms (specifically, the Metropolis-Hastings algorithms) to estimate the two-factor Cairns term structure model using simulated data. The main contribution is therefore the development and use of MCMC simulation for estimating both the driving latent variables and model parameters of the Cairns arbitrage-free model which has a non-linear bond price formula. The existence of latent variables is a common issue that causes difficulty to the estimation of many continuous-time term structure models, and here it can be effectively dealt with by using MCMC methodology under a Bayesian approach. According to the results, the standard MH algorithm gives rise to reasonable estimation but the chains of model parameters converge rather slowly. Nevertheless, by using the adaptive MH algorithm associated with a blocking strategy and reparameterising the log posterior distribution, the resulting estimates are significantly improved. We can conclude that our algorithm based on the MCMC framework estimates the term structure model very well and it seems efficient enough to deal with real market data which is more complex than simulation data.

References

Ahn, A.-H., Dittmar, R.F., and Gallant, A.R. (2002). Quadratic term structure models: theory and evidence. *Review of Financial Studies*, 15: 243-288.

Andersen, T.G., and Lund, J. (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, 77: 343-377.

Bester, A.C. (2004). Random Fields and Affine Models for Interest Rates: An Empirical Comparison. Working Paper, Duke University.

Cairns, A.J.G. (2004). A Family of term-structure models for long-term risk management and derivative pricing. *Mathematical Finance*, 14: 415-444.

Chan, K.C., Karolyi, G.A., Longstaff, F.A., and Sanders, A.B. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47: 1209-1227.

Dai, Q., and Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55: 1943-1978.

Duffie, D., and Singleton, K.J. (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica*, 61: 929-952.

Eraker, B. (2001). MCMC Analysis of Diffusion Models with Application to Finance. *Journal of Business* & *Economic Statistics*, **19**, 177-191.

Fleasaker, B., and Hughston, L. (1996). Positive interest. *Risk*, 9(1): 46-49.

Gallant, A.R., and Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12: 657-681.

Haario, H., Laine M., Mira, A., and Saksman E. (2005). DRAM - efficient adaptive MCMC. *Statistics and Computing*, 16: 339-354.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50: 1029-1055.

Hu, H. (2005). Markov Chain Monte Carlo Estimation of Multi-Factor Affine Term-Structure Models. Ph.D. Dissertation, University of California Los Angeles.

Lamoureux, C.G., and Witte, H.D. (2002). Empirical Analysis of the Yield Curve: The Information in the Data Viewed through the Window of Cox, Ingersoll, and Ross. *Journal of Finance*, 57: 1479-1520.

Litterman, R., and Scheinkman, J.A. (1991). Common factors affecting bond returns. *Journal of Fixed Income*, 1: 54-61.

Longstaff, F.A., and Schwartz, E.S. (1992). Interest rate volatility and the term structure: a two-factor general equilibrium model. *Journal of Finance*, 47: 1259-1282.

Nowman, K.B. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance*, 52: 1695-1706.

Pearson, N., and Sun, T.-S. (1994). An empirical examination of the Cox, Ingersoll and Ross model of the term structure of interest rates using the method of maximum likelihood. *Journal of Finance*, 54: 929-959.

Pooter, M., Ravazzolo, F., and Dijk, D. (2007). Predicting the Term Structure of Interest Rates: Incorporating parameter uncertainty, model uncertainty and macroeconomic information. Tinbergen Institute Disccusion Paper, TI2007-028/4.