Classification of Two- and Three-Factor Time-Homogeneous Separable LMMs

Jaka Gogala*[1] and Joanne E. Kennedy†[1]

[1]Department of Statistics, University of Warwick
Coventry, CV4 7AL, United Kingodm

23rd January 2017

Abstract

The flexibility of parameterisations of the LIBOR market model comes at a cost, namely the LIBOR market model is high-dimensional, which makes it cumbersome to use when pricing derivatives with early exercise features. One way to overcome this issue for short and medium term time-horizons is by imposing the separability condition on the volatility functions and approximating the model using a single time-step approximation.

In this paper we examine the flexibility of separable LIBOR market models under the relaxed assumption that the driving Brownian motions can be correlated. In particular, we are interested in how the separability condition interacts with time-homogeneity – a desirable property of a LIBOR market model. We show that the two concepts can be related using a Levi-Civitá equation and provide a characterization of two- and three-factor separable and time-homogeneous LIBOR market models and show that they are of practical interest. The results presented in this paper are also applicable to local-volatility LIBOR market models. These separable volatility structures can be used for the driver of a two- or three-dimensional Markov-functional model - in which case no (single time step) approximation is needed and the resultant model is both time-homogeneous and arbitrage-free.

Keywords: Levi-Civitá equation, LIBOR market model, Markov-functional models, separability, time-homogeneity.

1 Introduction

The *LIBOR Market Models* (LMMs) are one of the most popular classes of terms structure models. One of the reasons for their popularity can be attributed to the flexibility of their parameterisations. However, this flexibility comes with a major drawback, the Markovian dimension of a LMM is equal to the number of forward rates in the model. This makes them particularly cumbersome to use for pricing of derivatives with early exercise features.

To overcome the issue of high-dimensionality Pietersz et al. (2004) proposed the *separability* constraint on the volatility structure of the LMM and proved that

*jaka.gogala@gmail.com (corresponding author)
†j.e.kennedy@warwick.ac.uk

a separable LMM has an approximation with Markovian dimension equal to the number of Brownian motions driving the model dynamics. This process came with two drawbacks. Firstly, it greatly restricted the class of available parameterisations. In particular, it was noted in Joshi (2011) that the separability condition is too restrictive to use when the instantaneous volatilities are time-homogeneous. Secondly, the approximation obtained is not arbitrage free and is only useful for time horizons up to 15 years.

In this paper we mainly address the first issue. In particular, we characterise two- and three-factor separable parameterisations of the LMM when components of the Brownian motion driving the model's dynamics are allowed to be correlated. We then analyse the obtained parameterisations and show that they are of practical interest.

We briefly comment on the second issue by pointing out the relationship between the separable LMMs to the Markov-functional models (MFMs) (Hunt et al., 2000). In particular, the characterised parameterisations can be used to define two- and three-dimensional MFMs that can be implemented efficiently and are arbitrage-free. Furthermore, we note that the ideas presented here can be extended to a more general class of local-volatility LMMs (Andersen and Andreasen, 2000).

The remainder of the paper is structured as follows. In Section 2 we introduce the basic concepts of LMMs. The separability condition is discussed and generalised in Section 3. In Section 4 we characterise the two- and three-factor separable LMM with time-homogeneous instantaneous volatilities. In Section 5 we discuss the models obtained from a practical point of view. Section 6 concludes.

2 LIBOR Market Models

Throughout the paper we will assume we are working on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ supporting a Brownian motion and satisfying the usual conditions. We will be interested in a single currency economy consisting of zero-coupon bonds (ZCBs) maturing on dates $T_1 < \ldots < T_{n+1}$ and will denote the time $t \leq T_i, i = 1, \ldots, n+1$, price of a T_i-maturity ZCB by D_{t,T_i}. We will model the prices of ZCBs indirectly via the forward LIBORs $L^i, i = 1, \ldots, n$, defined by

$$L^i_t = \frac{D_{t,T_i} - D_{t,T_{i+1}}}{\alpha_i D_{t,T_{i+1}}}, \quad t \leq T_i, \ i = 1, \ldots, n, \tag{2.1}$$

where α_i is the accrual factor associated with the period $[T_i, T_{i+1}]$.

Over the last few years, since the financial crisis of 2008, the interest-rate markets have evolved and it is now no longer sufficient to assume that equation (2.1) provides an accurate representation of the connection between discount factors and LIBORs. We now live in a 'multi-curve' world where discounting is usually driven by overnight index swaps (OIS) (since collateral deposits usually receive interest based on overnight rates) and LIBORs are correctly treated as separate. In the context of a term-structure model there are various levels of sophistication one could adopt in generalising (2.1). The simplest, which is sufficient for most applications in practice, would be to take

$$L^i_t - \frac{D_{t,T_i} - D_{t,T_{i+1}}}{\alpha_i D_{t,T_{i+1}}} = s^i_t, \quad t \leq T_i, \ i = 1, \ldots, n, \tag{2.2}$$

where, for each i, $s^i_t = s^i$ is some constant. More generally one could model s^i_t as a stochastic process.

For the purposes of this paper we will stick with the definition (2.1), for ease of exposition, but remark that extending our results to (2.2) when s_t^i is non-stochastic is straightforward.

Amongst the most popular models of the described economy is the *LIBOR Market Model* (LMM). It was developed in the 1990s by Miltersen et al. (1997), Brace et al. (1997), Musiela and Rutkowski (1997) and Jamshidian (1997). The basic idea behind the LMM is that the process $(L_t^i)_{t\in[0,T_i]}$ is a log-normal martingale under the T_{i+1}-forward measure associated with the T_{i+1}-maturity ZCB as the numeraire. In particular the prices of caplets on each of the forward LIBORs are given by the Black (1976) formula. To fully specify a LMM we need to specify the joint dynamics of the forward LIBORs under a common equivalent martingale measure (EMM).

A d-factor LMM under the T_{n+1}-forward measure, usually referred to as the terminal measure, is given by a system of SDEs

$$dL_t^i = L_t^i\langle\tilde{\sigma}^i(t), d\tilde{W}_t\rangle - L_t^i\sum_{j=i+1}^{n}\frac{\alpha_j L_t^j\langle\tilde{\sigma}^i(t),\tilde{\sigma}^j(t)\rangle}{1+\alpha_j L_t^j}dt, \quad t\le T_i,\ i=1,\dots,n, \tag{2.3}$$

where $\tilde{W}$ is a standard d-dimensional standard Brownian motion under the measure $\mathbb{F}^{n+1}$ and $\tilde{\sigma}^i : [0,T_i]\to\mathbb{R}^d, i=1,\dots,n$, are bounded measurable functions and $\langle x,y\rangle$ denotes the inner product of vectors. One can show that under these conditions the system of SDEs (2.3) admits a strictly positive strong solution when the initial forward LIBORs $L_0^i, i=1,\dots,n$, are strictly positive (see Section 14.2 in Andersen and Piterbarg (2010) for more details).

The specification of a LMM as in (2.3) is particularly useful from a computational perspective. For example it allows for a straight-forward implementation via Monte Carlo methods. However, it offers little intuition about the model's dynamics. It is therefore often useful to introduce *instantaneous volatility* and *instantaneous correlation* functions. The instantaneous volatility functions are given by

$$\sigma^{inst,i}(t) = \sqrt{\langle\tilde{\sigma}^i(t),\tilde{\sigma}^i(t)\rangle}, \quad t\le T_i,\ i=1,\dots,n, \tag{2.4}$$

and the instantaneous correlation functions are given by

$$\rho_{i,j}^{inst}(t) = \frac{\langle\tilde{\sigma}^i(t),\tilde{\sigma}^j(t)\rangle}{\sigma^{inst,i}(t)\sigma^{inst,j}(t)}, \quad t\le T_i\wedge T_j,\ i,j=1,\dots,n. \tag{2.5}$$

It is easy to see that

$$d(\log L_t^i)d(\log L_t^j) = \rho_{i,j}^{inst}(t)\sigma^{inst,i}(t)\sigma^{inst,j}(t)dt, \quad t\le T_i\wedge T_j,\ i,j=1,\dots,n, \tag{2.6}$$

and one can show that the instantaneous volatility and correlation functions uniquely determine a LMM. Furthermore, the time $t\le T_i$ implied volatility of a caplet written on $L_{T_i}^i$ is a deterministic function given by

$$\sigma^{impl,i}(t) = \frac{1}{\sqrt{T_i-t}}\left(\int_t^{T_i}\sigma^{inst,i}(s)^2ds\right)^{\frac{1}{2}}, \quad t\le T_i,\ i=1,\dots,n. \tag{2.7}$$

It is often convenient to fix a calendar time t and consider the time t implied volatilities as a function of the maturity of the caplet, i.e.

$$T_i \mapsto \sigma_t^{impl,i}, \quad T_i > t. \tag{2.8}$$

We will refer to such function as the time t *term structure of volatilities* or simply term structure of volatilities when t is clear from the context.

Observe that by specifying the instantaneous volatility functions one implicitly specifies the evolution of the term structure of volatilities over time. In practice one often does not have a particular view on the dynamics of volatility surface and is faced with two natural choices. Either he chooses the implied volatilities to be constant functions of time (i.e. depend only on the maturity of the caplet) or that the implied volatilities are a function of the time to maturity (i.e. depend on the difference $T_i - t$) (see Section 6.2 in Rebonato (2002)). In this paper we will focus on the latter choice. It is easy to see that the implied volatility of a caplet will depend on the time to maturity if the instantaneous volatility functions satisfy the *time-homogeneity* condition

$$\sigma^{inst,i}(t) = \sigma^{inst}(T_i - t), \quad t \leq T_i, \ i = 1, \dots, n, \tag{2.9}$$

where $\sigma^{inst} : [0, T_n] \to \mathbb{R}_+$ is some bounded measurable function. In particular, σ^{inst} is often taken to be of the form

$$\sigma^{inst}(T_i - t) = \big(a + b(T_i - t)\big) \exp\big(-c(T_i - t)\big) + d. \tag{2.10}$$

This parameterisation was proposed by Rebonato (1999) and remains a popular choice amongst practitioners.

Let us now turn our attention back to the specification of the LMM. Recall that we assumed that the d-dimensional Brownian motion $\tilde{W}$ has independent components. While this assumption is in general non-restrictive, it turns out to be beneficial to relax it when there are additional constraints associated with functions $\tilde{\sigma}^i, i = 1, \dots, n$.

Suppose that $\rho : [0, T_n] \to \mathbb{R}^{d \times d}$ is a continuous matrix valued function such that $\rho(t)$ is a full rank correlation matrix for $t \leq T_n$. Then there exists a continuous matrix valued function $R : [0, T_n] \to \mathbb{R}^{d \times d}$ such that $R(t)$ is positive definite and $R(t)R(t) = \rho(t)$ for $t \leq T_n$. Then we can define a d-dimensional Brownian motion W (with correlated components) by

$$W_t = \int_0^t R(s) d\tilde{W}_s, \quad t \leq T_n, \tag{2.11}$$

and clearly $dW_t^T dW_t = R(t)R(t)dt = \rho(t)dt$. Now we can define functions $\sigma^i : [0, T_i] \to \mathbb{R}^d$ by

$$\sigma^i(t) = R(t)^{-1} \tilde{\sigma}^i(t), \quad t \leq T_i, \ i = 1, \dots, n. \tag{2.12}$$

Observe that $\langle \tilde{\sigma}^i(t), \tilde{W}_t \rangle = \langle \sigma^i(t), dW_t \rangle$ and $\langle \tilde{\sigma}^i(t), \tilde{\sigma}^j(t) \rangle = \langle \sigma^i(t), \rho(t)\sigma^j(t) \rangle$. Then if $(L^1, \dots, L^n)$ is a strong solution to the system of SDEs (2.3) it is also a strong solution to

$$dL_t^i = L_t^i \langle \sigma^i(t), dW_t \rangle - L_t^i \sum_{j=i+1}^{n} \frac{\alpha_j L_t^j \langle \sigma^i(t), \rho(t)\sigma^j(t) \rangle}{1 + \alpha_j L_t^j} dt, \quad t \leq T_i, \ i = 1, \dots, n. \tag{2.13}$$

We will refer to the collection of functions $\{\sigma^i\}_{i=1}^n$ in (2.13) as the volatility structure and will say that a LMM $(L^i)_{i=1}^n$ is parametrised by the pair $(\{\sigma^i\}_{i=1}^n, \rho)$. We can express the instantaneous volatility and correlation functions in terms of functions $\sigma^1, \dots, \sigma^n$ and ρ as

$$\sigma^{inst,i}(t) = \sqrt{\langle \sigma^i(t), \rho(t)\sigma^i(t) \rangle}, \quad t \leq T_i, \ i = 1, \dots, n, \tag{2.14}$$

and

$$\rho_{i,j}^{inst}(t) = \frac{\langle \sigma^i(t), \rho(t)\sigma^j(t) \rangle}{\sigma^{inst,i}(t)\sigma^{inst,j}(t)}, \quad t \leq T_i \wedge T_j, \ i, j = 1, \dots, n. \tag{2.15}$$

Remark 2.1. *Note that we could start by specifying a LMM as in* (2.13)*. This would allow for* $\rho : [0, T_n] \to [-1, 1]^{d\times d}$ *to be any correlation matrix valued function. In particular, if* $\rho(t)$ *is of rank* $d' < d$ *for* $t \le T_n$*, we get a* d *factor parameterisation of a* d' *factor LMM. This may seem suboptimal for implementation purposes, however as we will later observe this is not necessarily the case.*

Let us conclude this section by briefly discussing the implementation of the LMM. It turns out that one of the biggest challenges when implementing the LMM is the state dependent drift occurring in the SDEs for the forward LIBORs (see equations (2.3) and (2.13)). In particular this ensures that the LMM is Markovian in dimension n regardless of the dimension of the Brownian motion driving the dynamics. Furthermore, there are no closed form solutions for the joint distribution of the LIBORs at any date $t > 0$. Therefore, in order to implement the LMM it is necessary to approximate it. This is usually done in the log-space since

$$d \log L_t^i = \langle \sigma^i(t), dW_t\rangle - \left(\frac{1}{2}\sigma^{inst,i}(t)^2 + \sum_{j=i+1}^{n} \frac{\alpha_j L_t^j \langle \sigma^i(t), \rho(t)\sigma^j(t)\rangle}{1+\alpha_j L_t^j}\right)dt \tag{2.16}$$

and the distribution of $\int_{t_1}^{t_2} \langle \sigma^i(t), dW_t\rangle$ is known explicitly.

In this paper we will focus on the approximation in which the forward LIBORs are evolved from time 0 to time t in a single time-step. An early description of this method can be found in Hunter et al. (2001), however we will closely follow the approach and notation in Pietersz et al. (2004). Let us denote by Z a vector valued process, where the ith component, $i = 1, \ldots, n$, Z_i is given by

$$Z_i(t) = \int_0^t \langle \sigma^i(t), dW_t\rangle, \quad t \le T_i. \tag{2.17}$$

We say that $(L^{STSA,i})_{i=1}^n$ is a *single time-step approximation* of $(L^i)_{i=1}^n$ if

$$\log L_t^{STSA,i} = \log L_0^i + Z_i(t) + \mu^i(t, Z(t)), \quad t \le T_i, i = 1, \ldots, n, \tag{2.18}$$

where μ^i is defined by the drift approximation used (e.g. Euler, Brownian bridge, see Joshi and Stacey (2008)). Note that the drift approximation implicitly depends on the the initial term structure. Furthermore, observe that the process Z is in general an n-dimensional Markov process.

Remark 2.2. *Observe that the process* Z_i *is only well defined for* $t \le T_i$*, hence the drift approximation* μ^j *at time* $t \le T_j$ *may only depend on the* i*th component of vector* Z *if* $t \le T_i$*. However, this does not cause problems since the drift part of* $\log L^j$ *only depends on state of the* $L^{j+1}, \ldots, L^n$*.*

Remark 2.3. *Instead of approximating the LMM under the terminal measure, we could have used any* T_i *forward measure or the spot measure.*

The single-time step approximation is a powerful computational tool, however it does come with one major drawback. Like most approximations of the LMM it is not arbitrage free. In particular, the quality of approximation decreases with time and is typically only useful for time-horizons up to 10–15 years. Beyond that the arbitrage in the approximation becomes noticeable and care must be taken when using it for longer time horizons. This is typically less of a problem for the schemes that use many time steps to evolve the forward LIBORs in time. Nevertheless, the single time-step approximation is a useful method for short- and medium-term time horizons and its true power will be demonstrated in Section 4.

3 Separability

We have noted in previous section that a d-factor LMM is Markovian in dimension n. Therefore, one typically needs to implement it by using Monte Carlo methods, which are particularly cumbersome to use when pricing derivatives with early exercise features such as Bermudan swaptions. However, it was first shown by Pietersz et al. (2004) that a single-time step approximation of a d-factor LMM is Markovian in dimension d if we impose the *separability* condition on the volatility structure.[1]

Definition 3.1. *A volatility structure $\{\sigma^i : [0, T_i] \to \mathbb{R}^d\}_{i=1}^n$ is separable if there exist a function $\sigma : [0, T_n] \to \mathbb{R}^d$ and vectors $v^1, \ldots, v^n \in \mathbb{R}^d$ such that*

$$\sigma^i(t) = v^i * \sigma(t), \quad t \le T_i, \; i = 1, \ldots, n, \tag{3.1}$$

where operator $*$ denotes entry-by-entry multiplication of vectors.

We say that a d-factor LMM is separable if it can be parametrised by $(\{\sigma^i\}_{i=1}^n, \rho)$ where the volatility structure $\{\sigma^i\}_{i=1}^n$ is separable.

Definition 3.1 generalises the one given in Pietersz et al. (2004). In particular, it allows for the parameterisation of an LMM to be driven by a Brownian motion with correlated components. In fact Definition 3.1 is equivalent to the 'matrix separability' as defined in Denson and Joshi (2009) and the earlier two-factor extension by Piterbarg (2004) (see Appendix A). We chose to work with the above definition as it is more natural for the problem we consider in the next section when we consider the time-homogeneous separable LMMs.

Proposition 3.2. *Suppose forward LIBORs $(L^i)_{i=1}^n$ are given by a d-factor separable LMM and let $(L^{STSA,i})_{i=1}^n$ be a single-time step approximation to $(L^i)_{i=1}^n$. Then there exists a d-dimensional Markov process $x = (x_t)_{t \in [0,T_n]}$ and functions f^i : $[0, T_i] \times \mathbb{R}^d \to \mathbb{R}^+, i = 1, \ldots, n$, such that*

$$L_t^{STSA,i} = f^i(t, x_t), \quad t \le T_i, \; i = 1, \ldots, n. \tag{3.2}$$

Proof. Since $(L^i)_{i=1}^n$ are given by a separable d-factor LMM, there exists a parameterisation $(\{\sigma^i\}_{i=1}^n, \rho)$ such that the volatility structure $\{\sigma^i\}_{i=1}^n$ is separable, i.e. there exists function $\sigma : [0, T_n] \to \mathbb{R}^d$ and vectors $v^1, \ldots, v^n \in \mathbb{R}^d$ satisfying (3.1).

Let W be the d-dimensional Brownian motion, such that $dW_t dW_t^T = \rho(t)$, driving the dynamics of the LMM (under the terminal measure) and define the vector valued process $Z = (Z_i)_{i=1}^n$ as in (2.17). Now define a d-dimensional Markov process $x = (x_t)_{t \in [0,T_n]}$ by

$$x_t = \int_0^t \sigma(s) * dW_s, \quad t, \le T_n, \tag{3.3}$$

and observe that

$$Z_i(t) = \langle v^i, x_t \rangle, \quad t \le T_i, \; i = 1, \ldots, n. \tag{3.4}$$

In particular $Z(t) = v x_t$, where $v = [v^1, \ldots, v^n]^T$. Then any single time-step approximation $(L^{STSA,i})_{i=1}^n$ of $(L^i)_{i=1}^n$ is of the form

$$\log L_t^{STSA,i} = \log L_0^i + \langle v^i, x_t \rangle + \mu^i(t, v x_t), \quad t \le T_i, \; i = 1, \ldots, n, \tag{3.5}$$

[1]While separability has been used before to reduce the dimension of an interest rate model, for example Carverhill (1994) used it in the context of Heath et al. (1992) framework, its application to LMMs was introduced by Pietersz et al. (2004).

where μ^i depends on the drift approximation used. In particular there exist functions $f^i : [0, T_i] \times \mathbb{R}^d \to \mathbb{R}_+, i = 1, \ldots, n$, such that

$$L_t^{STSA,i} = f^i(t, x_t), \quad t \leq T_i, \; i = 1, \ldots, n. \tag{3.6}$$

□

The Proposition 3.2 is in fact independent of the equivalent martingale measure used to specify the model and the single time-step approximation. It was originally argued by Pietersz et al. (2004) that if one is to implement the single time-step approximation on a grid the terminal measure needs to be used to avoid the path dependence of the numeraire. However, one can easily implement the single time-step approximation under the spot measure associated with the rolling bank account numeraire by using same ideas as in the implementation of a Markov-functional model under the spot measure (Fries and Rott, 2004).

Since a single time-step approximation of a separable LMM can significantly reduce the computational effort needed for valuation of callable derivatives it is a natural question to ask how flexible are the separable LMMs. We will address this question in Section 3.

4 Time-Homogeneous Separable LMMs

We have pointed out in Section 2 that time-homogeneity of instantaneous volatilities is usually a desirable property of a LMM. In this section we will be interested which time-homogeneous instantaneous volatility functions can be obtained in a d-factor LMM when we also impose the separability condition on the volatility structure. In particular we will be interested in solutions of the system of functional equations

$$\sigma^{inst}(T_i - t)^2 = \langle v^i * \sigma(t), \rho(t)(v^i * \sigma(t))\rangle, \quad t \leq T_i, \; i = 1, \ldots, n. \tag{4.1}$$

Note that the system (4.1) implicitly depends on the choice of reset dates $T_1, \ldots, T_n$. It is therefore reasonable to only search for the solutions that continuously depend on the reset dates. This can be simply achieved by searching for the solutions of the functional equation

$$\sigma^{inst}(T - t)^2 = \langle v(T) * \sigma(t), \rho_t(v(T) * \sigma(t))\rangle, \quad t \leq T. \tag{4.2}$$

where we require $v : [0, \infty) \to \mathbb{R}^d$ to be a continuous function.

We will first consider one-factor volatility structures. This problem has already been examined in Joshi (2011), however it is instructional to study it first as it points out some of the important aspects of the problem that will be encountered later. In the one-factor case equation (4.2) can be simply rewritten as

$$\sigma^{inst}(T - t)^2 = v(T)^2 \sigma(t)^2, \quad t \leq T. \tag{4.3}$$

Note that if $\sigma^{inst}(x) = 0$ for some $x \geq 0$, then $\sigma^{inst} \equiv 0$ and either $v \equiv 0$ or $\sigma \equiv 0$ (or both). Clearly, such solution is not of interest and we can therefore assume without loss of generality that $\sigma^{inst}(x) \neq 0, x \geq 0$.

Next we define functions f, g, h, by $f(x) = \sigma^{inst}(x)^2$, $g(y) = \sigma(-y)^2$, and $h(x) = v(x)^2$, where $x \geq 0$ and $-x \leq y \leq 0$. Then we can rewrite (4.3) as

$$f(x + y) = h(x)g(y), \quad x \geq 0, -x \leq y \leq 0. \tag{4.4}$$

Equation (4.4) is commonly known as the Pexider equation. It can be shown that under the assumption that f is a continuous function[2] the general solution to the Pexider equation is of the form $f(x) = ab\exp(cx)$, $g(y) = a\exp(cy)$ and $h(x) = b\exp(cx)$, where $a, b, c \in \mathbb{R}$ (see Section 3.1 in Aczél (1966)).

Recall that $f(x) = \sigma^{inst}(x)^2 > 0$ and hence we are only interested in positive solutions to the Pexider equation and we need to restrict the parameters to $a, b > 0$. Furthermore, each solution to f, g, h, can be mapped to four solutions of equation (4.3):

1. $\sigma(t) = \sqrt{a}\exp(-\frac{1}{2}ct)$ and $v(T) = \sqrt{b}\exp(\frac{1}{2}cT)$;
2. $\sigma(t) = -\sqrt{a}\exp(-\frac{1}{2}ct)$ and $v(T) = \sqrt{b}\exp(\frac{1}{2}cT)$;
3. $\sigma(t) = \sqrt{a}\exp(-\frac{1}{2}ct)$ and $v(T) = -\sqrt{b}\exp(\frac{1}{2}cT)$;
4. $\sigma(t) = -\sqrt{a}\exp(-\frac{1}{2}ct)$ and $v(T) = -\sqrt{b}\exp(\frac{1}{2}cT)$.

and in all cases $\sigma^{inst}(T-t) = \sqrt{ab}\exp(\frac{1}{2}c(T-t))$. Now recall that σ and v affect the dynamics of the LMM through their product. Furthermore, the sign of the product $v(T)\sigma(t)$ can be absorbed into the Brownian motion driving the dynamics. Therefore, all four solutions lead to the same LMM and we can without loss of generality assume that one of the parameters a and b is equal to one.

Therefore a one-factor time-homogeneous and separable LMM can be parametrised as

$$\sigma(t) = \alpha\exp(\beta t), \tag{4.5}$$
$$v(T) = \exp(-\beta T), \tag{4.6}$$
$$\sigma^{inst} = \alpha\exp(-\beta(T-t)), \tag{4.7}$$

where $\alpha > 0$ and $\beta \in \mathbb{R}$.

As mentioned earlier the one-factor time-homogeneous separable LMMs was already characterised in Joshi (2011). However, there are two important observations we can make from our thought process. Firstly, although we imposed the continuity condition on function f this turned out not to be a restriction since a solution to the Pexider equation is either smooth or nowhere-continuous. Secondly, any solution to the Pexider equation corresponded to four solutions of (4.3) which all lead to the same LMM. We will see that above observations also hold in a d-factor setting where (4.2) can be transformed to a Levi-Civitá equation

$$f(x+y) = \sum_{i=1}^{k} g_i(x)h_i(y), \tag{4.8}$$

where $k = \frac{1}{2}(d^2 + d)$.

It can be shown that if $f, g_i, h_i, i = 1, \ldots, k$ is a continuous solution to (4.8) then $f, g_i, h_i \in \mathcal{C}^\infty$ and f is of the form

$$f(x) = \sum_i P_i(x)\exp(\lambda_i x), \tag{4.9}$$

where P_i is a polynomial of degree $k_i - 1$, such that $\sum_i k_i = k$, and $\lambda_i \in \mathbb{C}$ (See Section 4.2 in Aczél (1966)).

4.1 Two Factor Case

In the two factor case (4.2) can be rewritten as

$$\begin{aligned}\sigma^{inst}(T-t)^2 = {} & v_1(T)^2\sigma_1(t)^2 + v_2(T)^2\sigma_2(t)^2 \\ & + 2v_1(T)v_2(T)\rho_{1,2}(t)\sigma_1(t)\sigma_2(t).\end{aligned} \tag{4.10}$$

[2] It is enough to assume that f is continuous at a single point.

To simplify the analysis of (4.10) we introduce functions

$$f(x) = \sigma^{inst}(x)^2, \tag{4.11}$$

$$g_i(x) = \sigma_i(x)^2, \quad i = 1, 2, \tag{4.12}$$

$$g_3(x) = 2\rho_{1,2}(x)\sigma_1(x)\sigma_2(x), \tag{4.13}$$

$$h_i(x) = v_i(x)^2, \quad i = 1, 2, \tag{4.14}$$

$$h_3(x) = v_1(x)v_2(x). \tag{4.15}$$

We can then rewrite (4.10) as

$$f(T-t) = \sum_{i=1}^{3} g_i(t)h_i(T). \tag{4.16}$$

Note that equation (4.16) can be easily transformed to the form of equation (4.8) by the following change of coordinates

$$(x(T,t), y(T,t)) = (T, -t). \tag{4.17}$$

Therefore, if we assume that f, g_i, h_i are continuous functions, f is of the form as in equation (4.9).

Theorem 4.1. *Let $v, \sigma : \mathbb{R}_+ \to \mathbb{R}^2$ and $\rho_{1,2} : \mathbb{R}_+ \to [-1, 1]$ be continuous functions such that equation* (4.10) *holds for some function $\sigma^{inst} : \mathbb{R}_+ \to \mathbb{R}_+$.*

Then v, σ and $\rho_{1,2}$ are parametrised up to the uniqueness of σ^{inst} by one of the following parameterisations

2.1. $\alpha_1, \alpha_2 \geq 0, \beta_1,\ \beta_2 \in \mathbb{R}$ *and* $\gamma \in [-1, 1]$

$$v(T) = \begin{bmatrix} \exp(-\beta_1 T) \\ \exp(-\beta_2 T) \end{bmatrix}, \tag{4.18}$$

$$\sigma(t) = \begin{bmatrix} \alpha_1 \exp(\beta_1 t) \\ \alpha_2 \exp(\beta_2 t) \end{bmatrix}, \tag{4.19}$$

$$\rho_{1,2}(t) = \gamma; \tag{4.20}$$

2.2. $\alpha > 0, \beta \in \mathbb{R}, \gamma \geq 0$ *and* $\lambda \in \mathbb{R}$

$$v(T) = \begin{bmatrix} T\exp(-\lambda T) \\ \exp(-\lambda T) \end{bmatrix}, \tag{4.21}$$

$$\sigma(t) = \begin{bmatrix} \alpha \exp(\lambda t) \\ \alpha\sqrt{(t+\beta)^2 + \gamma}\exp(\lambda t) \end{bmatrix}, \tag{4.22}$$

$$\rho_{1,2}(t) = -\frac{t+\beta}{\sqrt{(t+\beta)^2+\gamma}}; \tag{4.23}$$

2.3. $\alpha, \beta, \theta, \lambda \in \mathbb{R},\ \gamma \geq \sqrt{\alpha^2+\beta^2}$

$$v(T) = \begin{bmatrix} \operatorname{sgn}(\cos\frac{\theta T}{2} + \sin\frac{\theta T}{2})\sqrt{1+\sin(\theta T)}\exp(-\lambda T) \\ \operatorname{sgn}(\cos\frac{\theta T}{2} - \sin\frac{\theta T}{2})\sqrt{1-\sin(\theta T)}\exp(-\lambda T) \end{bmatrix}, \tag{4.24}$$

(a) If $\alpha^2 + \beta^2 > \gamma^2$

$$\sigma(t) = \begin{bmatrix} \sqrt{\gamma + \alpha\cos(\theta t) + \beta\sin(\theta t)}\exp(\lambda t) \\ \sqrt{\gamma - \alpha\cos(\theta t) - \beta\sin(\theta t)}\exp(\lambda t) \end{bmatrix}, \tag{4.25}$$

$$\rho_{1,2} = \frac{\beta\cos(\theta t) - \alpha\sin(\theta t)}{\sqrt{\gamma^2 - (\alpha\cos(\theta t) + \beta\sin(\theta t))^2}}; \tag{4.26}$$

(b) If $\alpha^2 + \beta^2 = \gamma^2$

$$\sigma(t) = \begin{bmatrix} \operatorname{sgn}(\cos\frac{\theta t-\phi}{2})\sqrt{\alpha\cos(\theta t)+\beta\sin(\theta t)+\gamma}\exp(\lambda t) \\ -\operatorname{sgn}(\sin\frac{\theta t-\phi}{2})\sqrt{-\alpha\cos(\theta t)-\beta\sin(\theta t)+\gamma}\exp(\lambda t) \end{bmatrix}, \tag{4.27}$$

$$\rho_{1,2} = 1, \tag{4.28}$$

where

$$\phi = \begin{cases} \arccos\frac{\alpha}{\gamma}; & \beta \geq 0 \\ -\arccos\frac{\alpha}{\gamma}; & \beta < 0 \end{cases}. \tag{4.29}$$

The proof can be found in Appendix B.

We analyse the parameterisations obtained in Theorem 4.1 in Section 5. At this point let us just mention that one of them can capture the 'hump' and the long term level of volatility simultaneously. For this reason we now consider the three factor case.

4.2 Three Factor Case

In the three factor case (4.2) can be rewritten as

$$\begin{aligned} \sigma^{inst}(T-t)^2 = {} & v_1(T)^2\sigma_1(t)^2 + v_2(T)^2\sigma_2(t)^2 + v_3(T)^2\sigma_3(t)^2 \\ & + 2v_1(T)v_2(T)\rho_{1,2}(t)\sigma_1(t)\sigma_2(t) \\ & + 2v_1(T)v_3(T)\rho_{1,3}(t)\sigma_1(t)\sigma_3(t) \\ & + 2v_2(T)v_3(T)\rho_{2,3}(t)\sigma_2(t)\sigma_3(t). \end{aligned} \tag{4.30}$$

We can now proceed similarly as in the two-factor case and we define functions $f, g_i, h_i, i = 1, \ldots, 6$ by

$$f(x) = \sigma^{inst}(x)^2, \tag{4.31}$$
$$g_i(x) = \sigma_i(x)^2, \quad i = 1, 2, 3 \tag{4.32}$$
$$g_4(x) = 2\rho_{1,2}(x)\sigma_1(x)\sigma_2(x), \tag{4.33}$$
$$g_5(x) = 2\rho_{1,3}(x)\sigma_1(x)\sigma_3(x), \tag{4.34}$$
$$g_6(x) = 2\rho_{2,3}(x)\sigma_2(x)\sigma_3(x), \tag{4.35}$$
$$h_i(x) = v_i(x)^2, \quad i = 1, 2, 3, \tag{4.36}$$
$$h_4(x) = v_1(x)v_2(x), \tag{4.37}$$
$$h_5(x) = v_1(x)v_3(x), \tag{4.38}$$
$$h_6(x) = v_2(x)v_3(x). \tag{4.39}$$

We can then rewrite (4.30) as

$$f(T-t) = \sum_{i=1}^{6} g_i(t)h_i(T). \tag{4.40}$$

Again we obtain an equation that can be easily transformed to equation (4.8) by the change of coordinates $(x(T,t), y(T,t)) = (T, -t)$. If we assume that σ, v and ρ are continuous functions then so are $g_i, h_i, i = 1, \ldots, 6$, and function f has to be of the form as in equation (4.9). In the three-factor case we will only be interested in solutions where the coefficients λ_i in (4.9) are real numbers.

Theorem 4.2. *Let $\sigma^{inst} : \mathbb{R}_+ \to \mathbb{R}_+$, $v, \sigma : \mathbb{R}_+ \to \mathbb{R}^2$ and $\rho_{1,2}, \rho_{1,3}, \rho_{2,3} : \mathbb{R}_+ \to [-1,1]$ be continuous functions. Furthermore, assume that matrix*

$$\rho(t) = \begin{bmatrix} 1 & \rho_{1,2}(t) & \rho_{1,3}(t) \\ \rho_{1,2}(t) & 1 & \rho_{2,3}(t) \\ \rho_{1,3}(t) & \rho_{2,3}(t) & 1 \end{bmatrix} \tag{4.41}$$

is a correlation matrix for $t \geq 0$.

Then the following parameterisations are solutions to equation (4.30):

3.1. $\alpha_1, \alpha_2, \alpha_3 \geq 0, \beta_1, \beta_2, \beta_3 \in \mathbb{R}$ *and* $\gamma \in [-1,1]^{3\times 3}$ *a correlation matrix*

$$v(T) = \begin{bmatrix} \exp(-\beta_1 T) \\ \exp(-\beta_2 T) \\ \exp(-\beta_3 T) \end{bmatrix}, \tag{4.42}$$

$$\sigma(t) = \begin{bmatrix} \alpha_1 \exp(\beta_1 t) \\ \alpha_3 \exp(\beta_2 t) \\ \alpha_2 \exp(\beta_3 t) \end{bmatrix}, \tag{4.43}$$

$$\rho(t) = \gamma; \tag{4.44}$$

3.2. $\alpha > 0$, $\gamma, \delta \geq 0$, $\beta, \lambda, \mu \in \mathbb{R}$, $\eta \in [-1,1]$ *and* $\varepsilon \in [\delta - \sqrt{\beta\eta^{-2} - \beta}, \delta + \sqrt{\beta\eta^{-2} - \beta}]$

$$v(T) = \begin{bmatrix} T\exp(-\lambda T) \\ \exp(-\lambda T) \\ \exp(-\mu T) \end{bmatrix}, \tag{4.45}$$

$$\sigma(t) = \begin{bmatrix} \alpha \exp(\lambda t) \\ \alpha\sqrt{(t+\beta)^2 + \gamma}\exp(\lambda t) \\ \delta \exp(\mu t) \end{bmatrix} \tag{4.46}$$

and ρ defined by

$$\rho_{1,2}(t) = -\frac{t+\beta}{\sqrt{(t+\beta)^2+\gamma}}, \tag{4.47}$$

$$\rho_{1,3}(t) = \eta, \tag{4.48}$$

$$\rho_{2,3}(t) = -\eta\frac{t+\varepsilon}{\sqrt{(t+\beta)^2+\gamma}}. \tag{4.49}$$

3.3. $\alpha, \gamma, \delta \geq 0$, $\beta, \lambda \in \mathbb{R}$

$$v(T) = \begin{bmatrix} T^2\exp(-\lambda T) \\ T\exp(-\lambda T) \\ \exp(-\lambda T) \end{bmatrix}, \tag{4.50}$$

$$\sigma(t) = \begin{bmatrix} \alpha \exp(\lambda t) \\ \alpha\sqrt{4(t+\beta)^2 + \gamma}\exp(\lambda t) \\ \alpha\sqrt{(t+\beta)^4 + \gamma(t+\beta)^2 + \delta}\exp(\lambda t) \end{bmatrix} \tag{4.51}$$

$$\tag{4.52}$$

and ρ defined by

$$\rho_{1,2}(t) = -\frac{2(t+\beta)}{\sqrt{4(t+\beta)^2+\gamma}}, \tag{4.53}$$

$$\rho_{1,3}(t) = \frac{(t+\beta)^2}{\sqrt{(t+\beta)^4+\gamma(t+\beta)^2+\delta}}, \tag{4.54}$$

$$\rho_{2,3}(t) = -\frac{2(t+\beta)^2+\gamma(t+\beta)}{\sqrt{(4(t+\beta)^2+\gamma)((t+\beta)^4+\gamma(t+\beta)^2+\delta)}}. \tag{4.55}$$

The proof of the Theorem 4.2, can be simply done by verifying that parameterisations presented are valid ($\rho(t)$ needs to be a correlation matrix) and satisfy the time-homogeneity condition.

Remark 4.3. *Theorem 4.2 does not classify all 3-factor separable time-homogeneous parameterisations of the LMM, in particular restrictions on the Parameterisation 3.3 could be relaxed. However, one can show that Parameterisations 3.1 and 3.2 cannot be generalised. Furthermore, it characterises all parameterisations where* $(\sigma^{inst})^2$ *captures the long term level of volatility and is a sum of exponential polynomials with real coefficients.*

5 Analysis

Recall that a separable LMM is given by vectors $v_1, \dots, v_n$, vector valued function σ and correlation matrix valued function ρ. However, to analyse the dynamics of a LMM it is more intuitive to think in terms of instantaneous volatility and correlation functions. These can be expressed in terms of $v_i, i = 1, \dots, n$, σ and ρ by combining equations (2.14), (2.15) and (3.1) as

$$\sigma^{inst,i}(t) = \sqrt{\langle v^i * \sigma(t), \rho(t)(v^i * \sigma(t))\rangle}, \tag{5.1}$$

$$\rho^{inst}_{i,j}(t) = \frac{\langle v^i * \sigma(t), \rho(t)(v^j * \sigma(t))\rangle}{\sigma^{inst,i}(t)\sigma^{inst,j}(t)}. \tag{5.2}$$

Recall that we have imposed the time-homogeneity condition on the instantaneous volatility functions explicitly in Theorems 4.1 and 4.2. However, it turns out that the parameterisations characterised in the theorems result in instantaneous correlation functions $\rho^{inst}_{i,j}, i, j = 1, \dots, n$, that also depend on the maturities T_i, T_j and the calendar time t only through the times to maturity $T_i - t$ and $T_j - t$. Moreover, the parameterisations obtained in Theorems 4.1 and 4.2 are independent of the choice of the setting dates $T_1, \dots, T_n$. Therefore we can think of instantaneous volatilities and correlations for the purposes of this section as functions $\sigma^{inst} : \mathbb{R}_+ \to \mathbb{R}_+$ and $\rho^{inst} : \mathbb{R}^2_+ \to [-1, 1]$, whose arguments are times to maturity.

In the two-factor model we get the following parameterisations of the instantaneous volatility and correlation:

2.1. $\alpha_1, \alpha_2 \geq 0$, $\beta_1, \beta_2 \in \mathbb{R}$ and $\gamma \in [-1, 1]$

$$\begin{aligned}\sigma^{inst}(x)^2 = {} & \alpha_1^2 \exp(-2\beta_1 x) + \alpha_2^2 \exp(-2\beta_2 x) \\ & + 2\alpha_1\alpha_2\gamma \exp(-(\beta_1 + \beta_2)x),\end{aligned} \tag{5.3}$$

$$\begin{aligned}\rho^{inst}(x_1, x_2) = {} & \Big(\alpha_1^2 \exp(-\beta_1(x_1 + x_2)) + \alpha_2^2 \exp(-\beta_2(x_1 - x_2)) \\ & + \alpha_1\alpha_2\gamma \exp(-\beta_1 x_1 - \beta_2 x_2) \\ & + \alpha_1\alpha_2\gamma \exp(-\beta_2 x_1 - \beta_1 x_2)\Big) \\ & /(\sigma^{inst}(x_1)\sigma^{inst}(x_2));\end{aligned} \tag{5.4}$$

$$\tag{5.5}$$

2.2. $\alpha > 0$, $\beta, \lambda \in \mathbb{R}$ and $\gamma \geq 0$

$$\sigma^{inst}(x)^2 = \alpha^2((x - \beta)^2 + \gamma)\exp(-2\lambda x), \tag{5.6}$$

$$\tag{5.7}$$

$$\rho^{inst}(x_1, x_2) = \frac{(x_1 - \beta)(x_2 - \beta) + \gamma}{\sqrt{((x_1 - \beta)^2 + \gamma)((x_2 - \beta)^2 + \gamma)}}. \tag{5.8}$$

2.3. $\alpha, \beta, \theta, \lambda \in \mathbb{R}$, $\gamma \geq \sqrt{\alpha^2 + \beta^2}$

$$\sigma^{inst}(x)^2 = 2\big(\alpha\cos(\theta x) + \beta\sin(\theta x) + \gamma\big)\exp(-2\lambda x) \tag{5.9}$$

$$\rho^{inst}(x_1, x_2) = \frac{\alpha\sin\big(\frac{\theta}{2}(x_1+x_2)\big) + \beta\cos\big(\frac{\theta}{2}(x_1+x_2)\big) + \gamma\cos\big(\frac{\theta}{2}(x_1-x_2)\big)}{\sqrt{(\alpha\cos(\theta x_1) + \beta\sin(\theta x_1) + \gamma)(\alpha\cos(\theta x_2) + \beta\sin(\theta x_2) + \gamma)}} \tag{5.10}$$

In the three-factor case we get the following parameterisations:

3.1. $\alpha_1, \alpha_2, \alpha_3 \geq 0$, $\beta_1, \beta_2, \beta_3 \in \mathbb{R}$ and $\gamma_{1,2}\gamma_{1,3}, \gamma_{2,3} \in [-1, 1]$

$$\sigma^{inst}(x)^2 = \sum_{i=1}^{3}\sum_{j=1}^{3} \alpha_i\alpha_j\gamma_{i,j}\exp(-(\beta_i + \beta_j)x), \tag{5.11}$$

$$\rho^{inst}(x_1, x_2) = \frac{\sum_{i=1}^{3}\sum_{j=1}^{3} \alpha_i\alpha_j\gamma_{i,j}\exp(-\beta_i x_1 - \beta_j x_2)}{\sigma^{inst}(x_1)\sigma^{inst}(x_2)}, \tag{5.12}$$

where $\gamma_{i,j} := \gamma_{j,i}$ and $\gamma_{i,i} := 1$ and $\Gamma = (\gamma_{i,j})_{i,j=1}^{3}$ is a correlation matrix;

3.2. $\alpha > 0$, $\gamma, \delta \geq 0$, $\beta, \lambda, \mu \in \mathbb{R}$, $\eta \in [-1, 1]$ and $\varepsilon \in [\delta - \sqrt{\beta\eta^{-2} - \beta}, \delta + \sqrt{\beta\eta^{-2} - \beta}]$

$$\begin{aligned}\sigma^{inst}(x)^2 = {} & \alpha^2((x-\beta)^2 + \gamma)\exp(-2\lambda x) \\ & + 2\alpha\delta\eta(x-\varepsilon)\exp(-(\lambda+\mu)x) \\ & + \delta^2\exp(-2\mu x)\end{aligned} \tag{5.13}$$

$$\begin{aligned}\rho^{inst}(x_1, x_2) = {} & \Big(\alpha^2((x_1-\beta)(x_2-\beta) + \gamma)\exp(-2\lambda(x_1+x_2)) \\ & + \alpha\delta\eta(x_1-\varepsilon)\exp(-\lambda x_1 - \mu x_2) \\ & + \alpha\delta\eta(x_2-\varepsilon)\exp(-\lambda x_2 - \mu x_1) \\ & + \delta^2\exp(-2\mu(x_1+x_2))\Big)\Big/\big(\sigma^{inst}(x_1)\sigma^{inst}(x_2)\big);\end{aligned} \tag{5.14}$$

3.3. $\alpha, \gamma, \delta \geq 0$, $\beta, \lambda \in \mathbb{R}$

$$\sigma^{inst}(x)^2 = \alpha^2((x-\beta)^4 + \gamma(x-\beta)^2 + \delta)\exp(-2\lambda x) \tag{5.15}$$

$$\begin{aligned}\rho^{inst}(x_1, x_2) = {} & \Big((x_1-\beta)^2(x_2-\beta)^2 + \gamma(x_1-\beta)(x_2-\beta) + \delta\Big) \\ & \Big/\Big(\sqrt{((x_1-\beta)^4 + \gamma(x_1-\beta)^2 + \delta)((x_2-\beta)^4 + \gamma(x_2-\beta)^2 + \delta)}\Big).\end{aligned} \tag{5.16}$$

Note that Parameterisation 2.1. can be seen as a special case of Parameterisation 3.1. by setting $\alpha_3 = 0$ and $\gamma_{1,2} = \gamma$ and that Parameterisation 2.2. can be seen as a special case of Parameterisation 3.2. by setting $\delta = 0$.

In the rest of the section we analyse the obtained instantaneous volatility by relating them to the implied volatilities which can be observed on the market. Then we consider the implied volatilities and we conclude by pointing out some practical implications of using two- and three-factor separable and time-homogeneous LMMs.

5.1 Instantaneous Volatiltiy

We have noted in Section 2 that time-homogeneity of instantaneous volatilities is a desirable property of LMM. This motivated us to characterise the two- and three-factor time-homogeneous and separable LMMs. Next we analyse the flexibility of the obtained instantaneous volatility functions.

In practice the instantaneous volatilities of forward rates cannot be observed directly but we can observe the term-structure of volatiltiy for a finite set of different times to maturity. Section 6.3 in Rebonato (2002) contains an analysis of historical data on term-structure of volatility. In particular, he points out that the term-structure remains relatively stable over time and at each date has one of the following shapes

- *Hump shape*: the term structure of volatilities first increases with time to maturity up to some time T' and after T' decreases as time to maturity increases;
- *Monotonically decreasing*: the term structure monotonically decreases with time to maturity.

Furthermore, he observes that the implied volatilities do not decrease to zero as the time to maturity increases but approach some non-negative constant, which we will call the *long-term level of volatility*.

Under the assumption that the instantaneous volatilities are time-homogeneous, i.e. there exists a function σ^{inst} such that condition (2.9) holds, then it is easy to observe:

- If σ^{inst} is hump shaped then the term structure of volatilities is hump shaped;
- If σ^{inst} is monotonically decreasing then the term structure of volatilities is monotonically decreasing.

Moreover, if $\lim_{x\to\infty} \sigma^{inst}(x) = 0$ then

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \sigma^{inst}(x)^2 dx = 0. \tag{5.17}$$

In particular, if σ^{inst} is a decreasing function on an interval (a, ∞) for some $a \geq 0$ then the implied volatilities will converge to some non-zero long term level if and only if $\lim_{x\to\infty} \sigma^{inst}(x) \neq 0$.

Therefore, a good parameterisation of a time-homogeneous instantaneous volatility function will converge to a positive constant as time to maturity increases and will be able to represent both hump-shaped and monotonically decreasing instantaneous volatilities.

Two Factors

We begin by analysing the instantaneous volatility functions we can obtain in the two-factor case and which are given in equations (5.3) and (5.6).

Parameterisation 2.1 The instantaneous volatility function for the Parameterisation 2.1 is given by the parameters $\alpha_1, \alpha_2 \geq 0$, $\beta_1, \beta_2 \in \mathbb{R}$, $\gamma \in [-1, 1]$ and equation (5.3). For the purpose of this discussion we will assume that $\alpha_1, \alpha_2 > 0$ and $\beta_1 \neq \beta_2$ as the instantaneous volatility function otherwise reduces to a single exponential. Furthermore we will assume that $0 \leq \beta_1 < \beta_2$ to ensure that the instantaneous volatility function is bounded on $\mathbb{R}_+$. Figure 1 shows plots of the instantaneous volatility function for various choices of parameter values.

Clearly this parameterisation can capture the long-term level of volatility when $\beta_1 = 0$ in this case $\lim_{x\to\infty} \sigma^{inst}(x) = \alpha_1$. Moreover, when $\gamma \in [0, 1]$ the function σ^{inst} is strictly decreasing. On the other hand if $\gamma \in [-1, 0)$ the instantaneous volatility function has a local minimum at $x' = \frac{1}{\beta_2} \log \frac{\alpha_1}{-\alpha_2\gamma}$. When $x' \leq 0$ the instantaneous volatility function is strictly increasing (on $\mathbb{R}_+$) and when $x' > 0$ the instantaneous

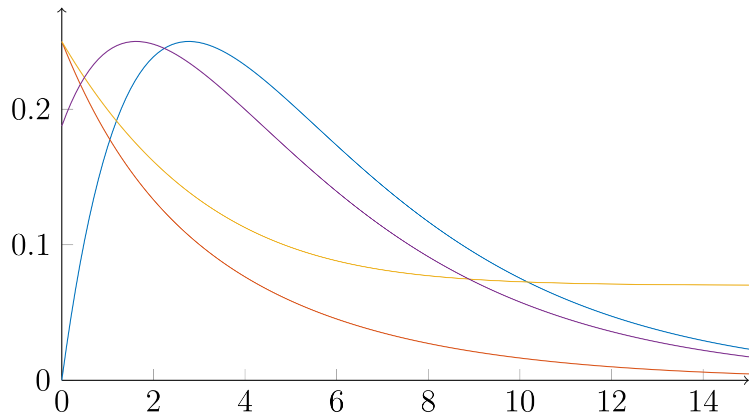

Figure 1: Plots of instantaneous volatility as a function of time to maturity corresponding to Parameterisation 2.1 (equation (5.3)) for various different choices of parameter values.

volatility function is strictly decreasing on $[0, x')$ and strictly increasing on (x', ∞). In particular when $\beta_1 = 0$ the instantaneous volatility function cannot capture the hump, but it can capture the monotonically decreasing instantaneous volatilities and the long-term level of volatility.

Let us now consider the case when $\beta_1 > 0$. In this case it is obvious that $\lim_{x\to\infty} \sigma^{inst} = 0$ and the instantaneous volatility cannot capture the long-term level of volatility. Furthermore, when $\gamma \geq 0$ it is easy to observe that the instantaneous volatility function is strictly decreasing. One can show that σ^{inst} has two local extrema x_1' and x_2' (on $\mathbb{R}$) if and only if

$$\gamma < -2\frac{\sqrt{\beta_1\beta_2}}{\beta_1 + \beta_2}. \tag{5.18}$$

In particular when $\gamma = -1$ the local extrema occur at

$$x_1' = \frac{1}{\beta_2 - \beta_1} \log \frac{\alpha_2}{\alpha_1}, \qquad x_2' = \frac{1}{\beta_2 - \beta_1} \log \frac{\alpha_2\beta_2}{\alpha_1\beta_1}. \tag{5.19}$$

Since $\beta_1 < \beta_2$ it follows $x_1' < x_2'$ and the local minimum is attained at x_1' and the local maximum is attained at x_2'. Note that when $\alpha_1 \geq \alpha_2$ then $x_1' \leq 0$ and σ^{inst} is strictly increasing on $(0, x_2')$ and strictly decreasing towards zero on (x_2, ∞) and is therefore hump shaped.

To summarise, the instantaneous volatility function given by Parameterisation 2.1 cannot capture the hump and the long-term level simultaneously. However, it can capture monotonically decreasing volatilities together with the long-term level of volatility.

Parameterisation 2.2 Next we analyse the instantaneous volatility function corresponding to Parameterisation 2.2 given in equation (5.6). Figure 2 shows plots of the instantaneous volatility function for various choices of parameter values.

First observe that σ^{inst} will be bounded (on $\mathbb{R}_+$) if and only if $\lambda > 0$, which we will assume throughout the analysis. In this case it is clear that $\lim_{x\to\infty} \sigma^{inst}(x) = 0$ and the instantaneous volatility function cannot capture the long-term level of volatility.

Secondly note that the parameter α is a scale parameter and does not affect the shape of the instantaneous volatility function, which is affected only by the parameters β, γ and λ. Parameter λ controls the speed of decay of instantaneous

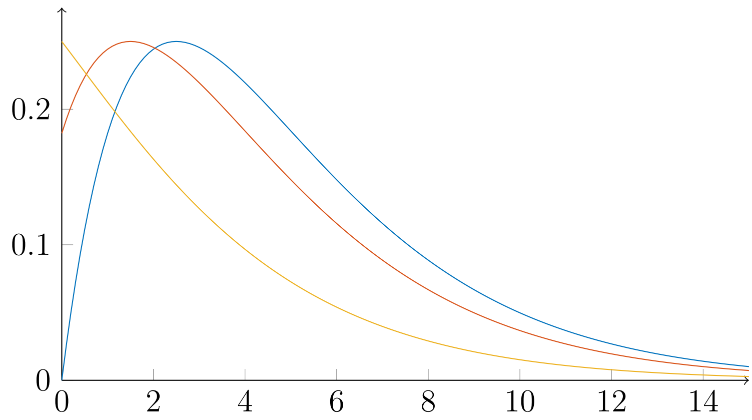

Figure 2: Plots of instantaneous volatility as a function of time to maturity corresponding to Parameterisation 2.2 (equation (5.6)) for various different choices of parameter values.

volatility function and one can think of β and γ as a shfit along x and y axis respectively. Note however that the shift will be non-linear and affected by the decay, i.e. the effect of varying β and γ on the instantaneous volatility will decrease as time to maturity increases.

It is then easy to observe that σ^{inst} has local extrema (on $\mathbb{R}$) if and only if

$$\gamma < \frac{1}{4\lambda^2}, \tag{5.20}$$

which is in practice a relatively mild constraint. The local extrema are then attained at

$$x_1' = \beta + \frac{1 - \sqrt{1 - 4\gamma\lambda^2}}{2\lambda}, \qquad x_2' = \beta + \frac{1 + \sqrt{1 - 4\gamma\lambda^2}}{2\lambda}. \tag{5.21}$$

In particular, x_1' is a local minimum and x_2' is a local maximum.[3] Note that $x_1' < x_2'$ and that changing the parameter β will shift the location of the local extrema, which is in line with the intuitive interpretation of the parameter β. When $x_1' \leq 0 < x_2'$ the instantaneous volatility function is strictly increasing on $(0, x_2')$ and strictly decreasing on (x_2', ∞) and can therefore capture the hump. Furthermore, when $x_2' \leq 0$ the instantaneous volatility function is strictly decreasing on R_+. Note that in both cases $\beta < 0$.

To summarise, Parameterisation 2.2 can represent both monotonically decreasing and hump shaped volatilities. However it cannot capture the long-term level of volatility.

Three Factors

We have seen that the two-factor parameterisations cannot capture the hump and the long-term level of volatility simultaneously. We will show that introducing the third factor leads to significantly more flexible instantaneous volatility parameterisations, given by equations (5.11), (5.13) and (5.15), that can capture the hump and the long-term level of volatility simultaneously.

Parameterisation 3.1. First we consider the instantaneous volatility function given by equation (5.11). Figure 3 shows plots of the volatility function for various choices of parameter values.

[3]When $\gamma = \frac{1}{4\lambda^2}$ then $x_1' = x_2'$ is a saddle point.

Note that, by setting $\alpha_3 = 0$ the instantaneous volatility function reduces to the one we get in Parameterisation 2.1. Therefore we can assume that $\alpha_1, \alpha_2, \alpha_3 > 0$. Furthermore, in order for the instantaneous volatility function to be bounded we will additionally require $\beta_1, \beta_2, \beta_3 \geq 0$.

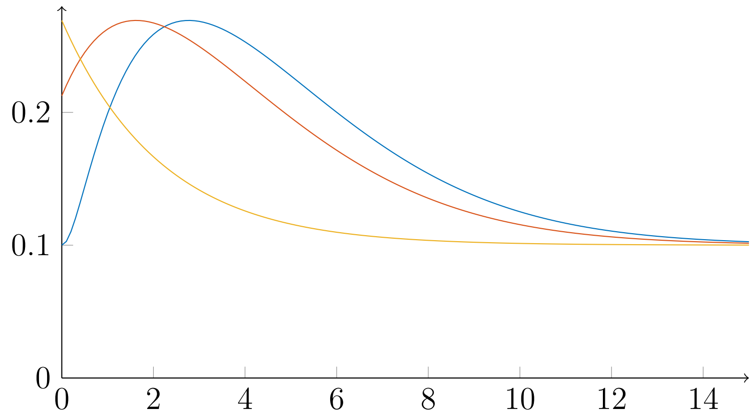

Figure 3: Plots of instantaneous volatility as a function of time to maturity corresponding to Parameterisation 3.1 (equation (5.11)) for various different choices of parameter values.

Recall that the main weakness of the Parameterisation 2.1 is its inability to capture the hump and the long-term level of volatility simultaneously. We will therefore only concentrate on the case when $\beta_3 = 0$ and $\beta_1 \neq \beta_2$. In this case we can interpret the parameter α_3 as the long-term level of volatility.

For the Parameterisation 3.1 to be valid, the matrix value function $\rho(t)$ describing the time t correlation structure of the Brownian motion driving the model needs to be a correlation matrix. In the case of Parameterisation 3.1 ρ is given by

$$\rho(t) = \begin{bmatrix} 1 & \gamma_{1,2} & \gamma_{1,3} \\ \gamma_{1,2} & 1 & \gamma_{2,3} \\ \gamma_{1,3} & \gamma_{3,3} & 1 \end{bmatrix} \tag{5.22}$$

and is a correlation matrix if and only if $\gamma_{1,2}, \gamma_{1,3}, \gamma_{2,3} \in [-1, 1]$ and

$$\det \rho(t) = 1 - (\gamma_{1,2}^2 + \gamma_{1,3}^2 + \gamma_{2,3}^2) + 2\gamma_{1,2}\gamma_{1,3}\gamma_{2,3} \geq 0. \tag{5.23}$$

When the third factor is independent of the first two (i.e. $\gamma_{1,3} = \gamma_{2,3} = 0$), equation (5.23) is satisfied for any $\gamma_{1,2} \in [-1, 1]$ and σ^{inst} has local extrema (on $\mathbb{R}$) if and only if

$$\gamma_{1,2} < -2\frac{\sqrt{\beta_1 \beta_2}}{\beta_1 + \beta_2}. \tag{5.24}$$

Note, that this is essentially the same condition as in the Parameterisation 2.1. Moreover, it is easy to verify that the local extrema are attained at the same points as for the Parameterisation 2.1.

When the third factor is correlated with the first two, one cannot in general explicitly find the local extrema, due to the first derivative being highly non-linear. However, allowing the third factor to be correlated with the first two clearly introduces additional flexibility to the instantaneous volatility parameterisation. In particular, this flexibility is necessary when the implied volatilities of caplets with short times to maturity are below the long-term level of volatility.

To summarise, Parameterisation 3.1 can capture both the hump and monotonically decreasing volatilities while it also captures the long term level of volatility. Its main downside is that it becomes less intuitive (but remains analytically tractable) when the factor representing the long-term level of volatility is correlated with the other two factors.

Parameterisation 3.2. The instantaneous volatility Parameterisation 3.2 given by equation (5.13) is perhaps the most interesting parameterisation we can achieve in a three-factor separable and time-homogeneous model. Figure 4 shows the plots of the volatility function for various choices of parameter values.

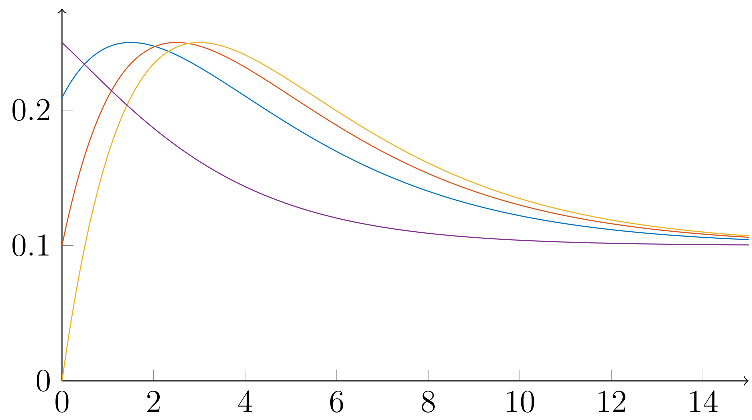

Figure 4: Plots of instantaneous volatility as a function of time to maturity corresponding to Parameterisation 3.2 (equation (5.13)) for various different choices of parameter values.

Note that setting the parameter $\delta = 0$ reduces the instantaneous volatility function to the one obtained in Parameterisation 2.1. In particular, we noted that the main drawback of Parameterisation 2.1 is its inability to capture the long term level of volatility.

Parameterisation 3.2 can capture the long-term level of volatility simply by setting $\mu = 0$ in which case δ can be interpreted as the long-term level of volatility. In particular, by setting $\alpha = |b|$, $\beta = -\frac{a}{b}$, $\gamma = 0$, $\delta = d$, $\varepsilon = -\frac{a}{b}$, $\eta = \operatorname{sgn} b$ and $\lambda = c$ the volatility function corresponds to the Rebonato's *abcd* instantaneous volatility parameterisation given by equation (2.10). In particular, the Parameterisation 3.2 can capture both hump and long term-level of volatility.

Clearly, we can get extra flexibility by also varying the parameters γ, η, however it is often sensible to set $\varepsilon = \beta$ as its effect on the volatility function is relatively limited.

Parameterisation 3.3. Finally let us briefly discuss the instantaneous volatility function given by equation (5.15) corresponding to Parameterisation 3.3. Recall that the main reason for considering the three-factor models was the inability of the two-factor parameterisations to capture the hump and the long-term level of volatility simultaneously. However, note that Parameterisation 3.3 cannot capture the long-term level of volatility. Therefore it will in most case perform only marginally better over the Parameterisation 2.1 and 2.2 which does not justify the increase in the number of factors used.

5.2 Instantaneous Correlation

Let us now turn our attention to the instantaneous correlations. Recall that we are interested only in the time-homogeneous instantaneous correlations parameterisations, which can be represented by a function $\rho^{inst} : \mathbb{R}_+^2 \to [-1, 1]$ where $\rho^{inst}(x, y)$ is the instantaneous correlation between two forward rates with times to maturity x and y respectively.

Ideally one would take a similar approach as for instantaneous volatilities and determine the desirable properties of instantaneous correlations by relating them to prices of European swaptions. However, this turns out to be a difficult task as in general one cannot separate the effects of the instantaneous correlations from the effects of instantaneous volatilities on the European swaption prices (see Section 7.1 in Rebonato (2002)).

One therefore needs to take a different route and estimate the instantaneous correlations from historical data (see Section 7.2 in Rebonato (2002) and Section 14.3 in Andersen and Piterbarg (2010)). By doing so one usually observes that the resulting instantaneous correlation matrix satisfies the following stylised facts (see Section 7.2 in Rebonato (2002), Section 23.8 in Joshi (2011))

1. Instantaneous correlations are positive

$$\rho^{inst}(x, y) > 0; \tag{5.25}$$

2. Instantaneous correlations decrease as the absolute value of the difference between the two times to maturity increases

$$|x - y| < |x - z| \Rightarrow \rho^{inst}(x, y) > \rho^{inst}(x, z); \tag{5.26}$$

3. Instantaneous correlation between forward rates with the difference between their times to maturity increases as the time to maturity of the forward rate expiring earlier increases

$$x < x' \Rightarrow \rho^{inst}(x, x + y) < \rho^{inst}(x', x' + y); \tag{5.27}$$

The most basic example of an instantaneous correlation function satisfying the first two stylised facts is the *exponential instantaneous correlation function* given by parameter $\beta > 0$ and equation

$$\rho^{inst}(x, y) = \exp\big(-\beta\,|x - y|\big), \tag{5.28}$$

Note that the exponential instantaneous correlation violates the stylised fact 3. To correct for this violation one can introduce the *square-root exponential instantaneous correlation function* given by parameter $\beta' > 0$ and equation

$$\rho^{inst}(x, y) = \exp\big(-\beta'\,|\sqrt{x} - \sqrt{y}|\big). \tag{5.29}$$

Figure 5 shows plots of the exponential and square-root exponential instantaneous correlation functions. We used $\beta = 0.05$ to specify the exponential instantaneous correlation function and chose β' so that the two instantaneous correlation functions agree for the pair of forward rates with times to maturity 1 and 15 years. Observe that for both functions the correlations rapidly decrease as the difference between the times to maturity increases.

We will later observe that the instantaneous correlations in the two- and three-factor separable and time-homogeneous LMM cannot achieve such a rapid decrease in instantaneous correlations. This is not only the case for the separable LMMs but will be true for low-factor LMMs in general and is a necessary compromise one needs to make when using a low-factor LMM.

Another way of comparing the instantaneous volatility functions is by performing a principal component analysis on the $n \times n$ matrix of instantaneous correlations

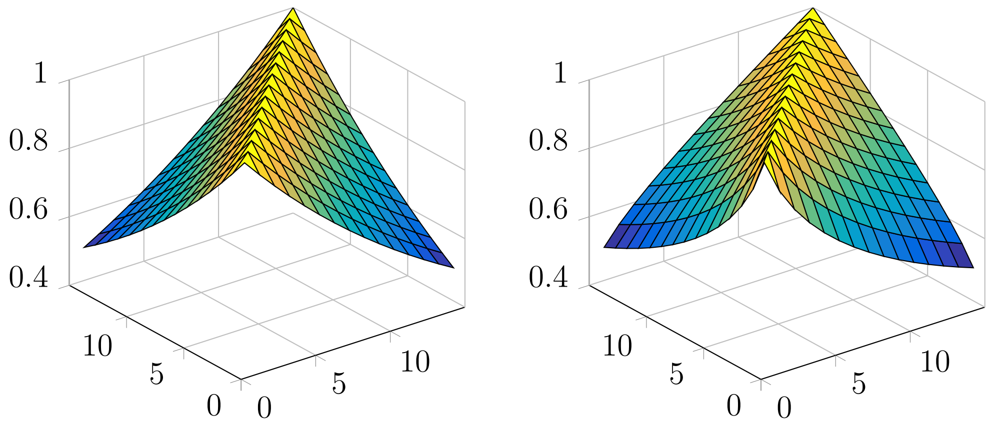

Figure 5: Plots of the exponential instantaneous correlation function (left) for $\beta = 0.05$ and the square-root exponential instantaneous correlation function (right) for $\beta' = 0.2436$.

between the rates with times to maturity $T_1, \ldots, T_n$. Empirical studies have shown that the first three components of such a matrix can be described as 'level', 'slope' and 'curvature' (see Lord and Pelsser (2007) Sections 1 and 2.2, and references within).

The Two-Factor Parameterisations

We now analyse the two-factor instantaneous correlation functions we obtained in Parameterisations 2.1 and 2.2. Note that in the two-factor case the instantaneous correlation matrix is of rank two or less and will therefore have at most two non-zero eigenvectors, which we would like to interpret as level (all elements of the same sign and approximately the same value) and slope (the elements are monotonically increasing or decreasing between the first and the last elements which are of opposite sign).

Parameterisation 2.1 We begin by considering the instantaneous correlation function given by equation (5.4). Without loss of generality we can assume that $\alpha_1, \alpha_2 > 0$, $\beta_1 \neq \beta_2$. Now recall that the parameter γ is the correlation between two components of the Brownian motions driving the separable LMM.

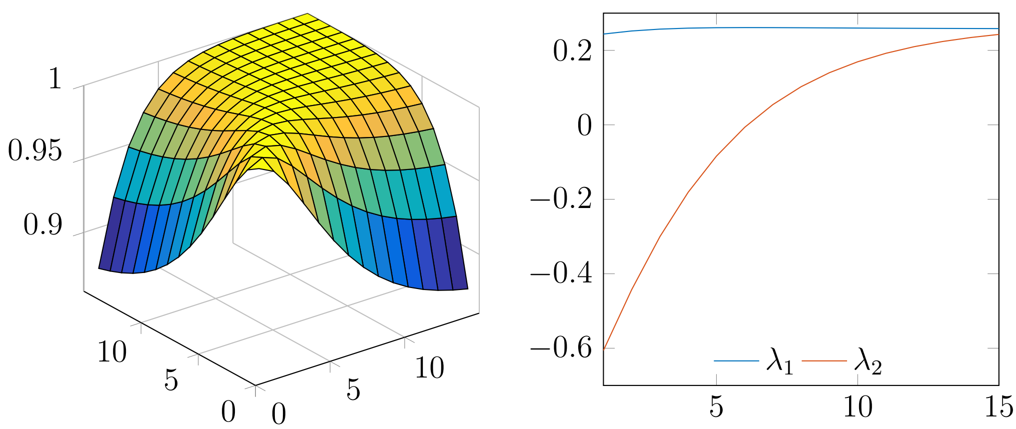

Figure 6: Plot of an instantaneous correlation function (left) and the first two principal components of the associated instantaneous correlation matrix for annual forward rates with times to maturity 1 to 15 years (right) corresponding to Parameterisation 2.1.

In particular when $\gamma \in \{-1, 1\}$ the components of the Brownian motion are perfectly (inversely) correlated. In this case the LMM is essentially a one-factor model and the forward rates are perfectly correlated. Note that when $\gamma \in \{-1, 1\}$ the resulting LMM is essentially driven by a single factor (see Remark 2.1), however it is separable in the dimension two and cannot be represented by a one-factor separable LMM.

On the other hand when $\gamma \in (-1, 1)$ the instantaneous correlation function is not identically equal to one and the resulting correlation matrix is of rank two. Moreover, the instantaneous correlations are strictly positive for every choice of parameters. However, it is in general difficult to analyse its dependence on the parameters due to complex interplay amongst them. Nevertheless, for a sensible choice of parameter values the correlation function results in mild-decorrelation between forward rates with short and long time to maturity and near perfect correlations between rates with longer times to maturity.

Figure 6 shows plots of a typical instantaneous correlation function (5.4) for a reasonable choice of parameter values and the first and second eigenvectors of the associated correlation matrix. Note that the forward rates with long maturities are nearly perfectly correlated, however there is some decorrelation between the rates of short to medium maturities and other rates. Moreover, the first two principal components of the correlation matrix can be interpreted as level and slope.

Parameterisation 2.2 We now turn our attention to the instantaneous correlation function given by equation (5.8). First observe that it only depends on the parameters β and γ.

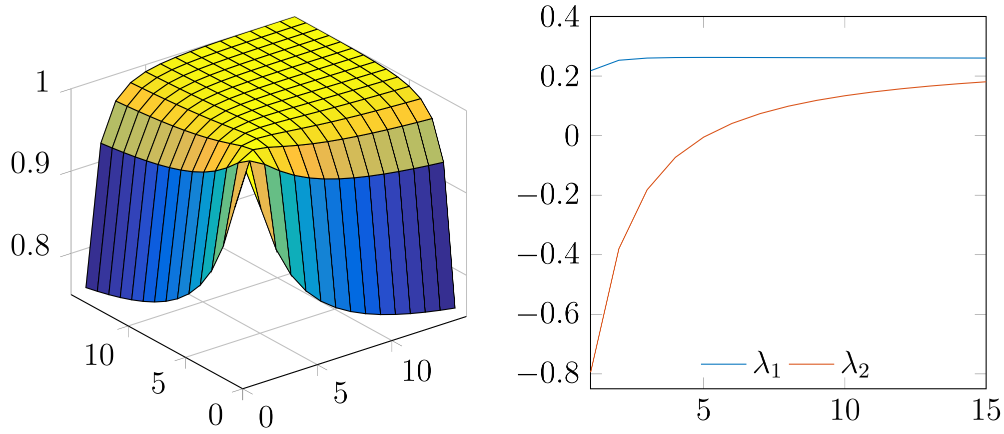

Figure 7: Plot of an instantaneous correlation function (left) and the first two principal components of the associated instantaneous correlation matrix for annual forward rates with times to maturity 1 to 15 years (right) corresponding to Parameterisation 2.2.

First note that when $\gamma = 0$ the instantaneous correlation function can be written as $\rho^{inst}(x_1, x_2) = \mathrm{sgn}((x_1 - \beta)(x_2 - \beta))$, in particular the model is effectively driven by a single factor. However when $\gamma > 0$ the instantaneous correlation function results in non-perfect correlations among forward rates.

On the other hand when $\beta > 0$ the instantaneous correlation function may attain negative values when one of the forward rates has time to maturity less then β and the other has time to maturity sufficiently greater than β. However, this turns out not to cause any problems from a practical perspective as $\beta > 0$ results in an unrealistic shape of the instantaneous volatility function. The more interesting scenario occurs when $\beta \leq 0$; and the instantaneous correlations are strictly positive. In this case increasing γ will decrease the correlations and decreasing β will increase the correlations amongst forward rates.

Figure 7 shows plots of a typical instantaneous correlation function for a reasonable choice of parameter values and the first and second principal component of a corresponding correlation matrix. Note the instantaneous correlations for rates of long-maturities are nearly perfect and there is some decorrelation between the forward rates of short and other times to maturity. Furthermore, the first two principal components can be interpreted as level and slope.

The Three-Factor Parameterisations

Having analysed the two-factor parameterisations let us now consider the three-factor Parameterisations 3.1 and 3.2. In the three-factor case we expect to observe curvature (the first and the last element are of the same sign but there is an element of the opposite sign which splits the elements into two monotonic sequences) in the principal component analysis of the instantaneous correlation matrix and higher levels of decorrelation.

Parameterisation 3.1 First consider the instantaneous correlation function given by equation (5.12). Recall that the matrix valued function ρ as defined in equation (5.22) is a correlation matrix describing the correlations amongst the components of driving Brownian motion. Therefore, the instantaneous correlation function will result in non-perfect instantaneous correlations only when the rank of matrix $\rho(t)$ is strictly greater than one.

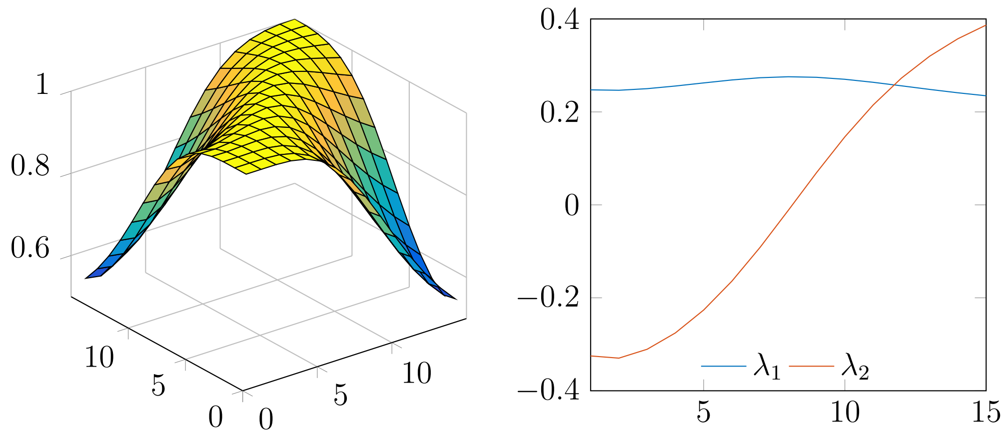

Figure 8: Plot of an instantaneous correlation function (left) and the first two principal components of the associated instantaneous correlation matrix for annual forward rates with times to maturity 1 to 15 years (right) corresponding to Parameterisation 3.1 when $\gamma_{1,2} = -1$.

Recall that from practical standpoint fixing $\gamma_{1,2} = -1$ is often desirable as it results in hump-shaped volatilities. It is then easy to see that the matrix $\rho(t)$ is a correlation matrix if and only if $\gamma_{1,3} = \gamma_{2,3} = 0$ and we have a three-factor separable parameterisation of a two-factor LMM. Nevertheless, as demonstrated by Figure 8, the instantaneous correlations obtained in such model are reasonable. In fact the decorrelation achieved is much greater than the ones observed in the two-factor separable models. Furthermore, the first two principal components of the correlation matrix can be interpreted as level and slope.

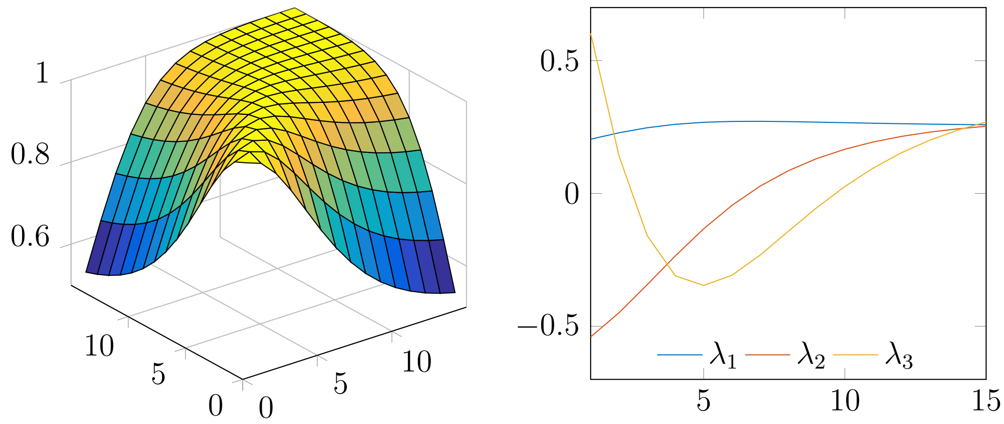

Figure 9: Plot of an instantaneous correlation function (left) and the first three principal components of the associated instantaneous correlation matrix for annual forward rates with times to maturity 1 to 15 years (right) corresponding to Parameterisation 3.1 when $\gamma_{1,2} = 0$.

On the other hand when $\rho(t)$ is a full rank matrix, the resulting model will be a proper three-factor LMM and the instantaneous correlation matrix will have three principal components corresponding to non-zero eigenvalues. Figure 9 shows plots of an instantaneous correlation function and the first three principal components of the associated instantaneous correlation matrix for a full rank $\rho(t)$. Note that the principal components can be interpreted as level, slope and curvature.

Observe the correlation functions in Figures 8 and 9 have significantly different shapes demonstrating the flexibility of the instantaneous correlation function (5.14).

Parameterisation 3.2 Finally let us consider the instantaneous correlation function given by equation (5.14) corresponding to perhaps the most interesting parameterisation of the three-factor separable and time-homogeneous LMM.

We begin by noting that in the special case when the parameters are chosen so that the instantaneous volatility function corresponds to the Reobnato's *abcd* parameterisation the resulting model is one-factor but it is represented by a three-factor separable parameterisation.

However, for a general parameterisation the instantaneous correlations will be non-perfect. Figure 10 shows plots of an instantaneous correlation function and the first three principal components of the associated correlation matrix for reasonable parameter values. Note that the instantaneous correlation function has shape similar to the one presented in Figure 8. Furthermore, observe that the principal components can be interpreted as the level, slope and curvature.

5.3 Remarks on Calibration and Implementation

Let us conclude this section by pointing out some practical remarks about the two- and three-factor separable parameterisations discussed above. Recall, that in all cases the instantaneous volatility and correlation function were determined by the same set of parameters. As a consequence, one has to simultaneously calibrate to the caplet and swaption prices.

In particular, to calibrate to caplet and swaption prices in the LMM one needs to be able to efficiently evaluate the *terminal covariance* elements between forward

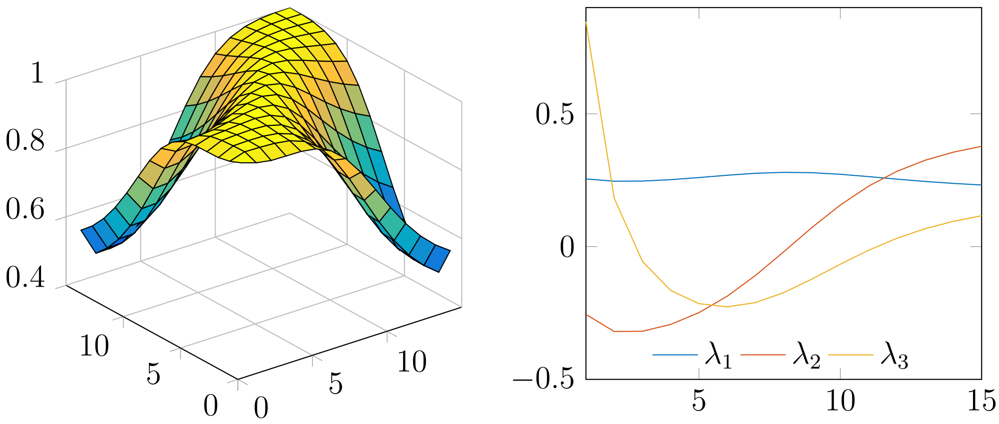

Figure 10: Plot of an instantaneous correlation function (left) and the first three principal components of the associated instantaneous correlation matrix for annual forward rates with times to maturity 1 to 15 years (right) corresponding to Parameterisation 3.2.

rates

$$C_{i,j}(T_k) = \int_0^{T_k} \rho^{inst}(T_i - t, T_j - t)\sigma^{inst}(T_i - t)\sigma^{inst}(T_j - t)dt, \quad k \leq i \wedge j. \quad (5.30)$$

It turns out that the terminal covariance elements of the parameterisations characterised in Theorems 4.1 and 4.2 have closed form representations, thus allowing for efficient calibration.

Furthermore, one could exploit the fact that the terminal covariance elements can be determined explicitly and perform a global calibration of a full-factor LMM and then calibrate the separable parameterisation to the terminal covariance elements that capture the dynamics of the forward rates relevant to the pricing of a particular instrument. However, the usefulness of such an approach is questionable (see Section 25.10 in Joshi (2011) for more detail).

6 Conclusion

In this paper we have addressed one of the main issues of the separable LMMs, their flexibility. We have generalised the separability condition and characterised the two- and three-factor separable LMM with time-homogeneous instantaneous volatilities. We then demonstrated that the obtained parameterisations are of practical interest by analysing their instantaneous volatilities and correlations. In particular, we have shown that a two-factor model can capture the long-term level of volatility when the instantaneous volatility function is decreasing or it can capture the hump shape in the instantaneous volatility function but not the long-term level of volatility. To capture the hump and the long term level of volatility we introduced the third factor, which can capture the popular Rebonato's *abcd* volatility function as a special case.

For all parameterisations we observed that the corresponding instantaneous correlation is also time-homogeneous, which was not directly imposed in the formulation of the problem. Furthermore, for reasonable parameterisations the instantaneous correlation functions were qualitatively in line with the stylised facts.

For ease of presentation we restricted our attention to the basic log-normal version of the LMM. However, one can generalise the separability condition further to the local-volatility LMM and reduce the dimension of the single time-step approximation, (see Gogala and Kennedy (2016) for more details, a similar approach was outlined in Section 12.8 in Joshi (2011)).

To conclude let us touch on the issue of the single time-step approximation admitting arbitrage and being useful only for time horizons up to 15 years. One way to avoid this issue is to use an appropriately specified *Markov-functional model* (MFM) instead of the single-time step approximation. The main idea of MFMs is to express forward rates at any given time as a function of some low-dimensional Markov process and the models are by construction arbitrage free and efficient to implement.

Recall, that in proving that a single-time step approximation of a separable LMM has Markovian dimension equal to the dimension of the Brownian motion driving the dynamics we have explicitly defined a Markov process x in equation (3.3). One can use this driving process to drive the dynamics of a MFM calibrated to the caplet prices from the separable LMM. Bennett and Kennedy (2005) have shown that in the case of a one-factor separable LMM the MFM specified as above has similar dynamics as the LMM. They believed that this observation also holds for separable LMM with higher number of factors.

Therefore, Theorems 4.1 and 4.2 are not only useful for characterising the two- and three-factor separable (log-normal) LMMs with time homogeneous instantaneous volatilities. But can be used in a more general local-volatility setting or to define two- and three-factor MFMs with dynamics similar to the LMM.

A Matrix separability

We have mentioned in Section 3 that the 'matrix separability' as described in Denson and Joshi (2009) is equivalent to our definition of separability. We now formalise this and furthermore, show that the number of free parameters in the approach by Denson and Joshi (2009) can be significantly reduced while not losing any generality, in particular we show that in two factor case the 'matrix separability' is equivalent to the two factor extension by Piterbarg (2004).

Definition A.1. *A volatility structure $\{\sigma^i : [0, T_i] \to \mathbb{R}^d\}_{i=1}^n$ is matrix separable if there exist a matrix valued function $C : [0, T_n] \to \mathbb{R}^{d\times d}$ and vectors $v^1 \dots, v^n \in \mathbb{R}^d$ such that*

$$\sigma^i(t) = C(t)v^i, \quad t \le T_i, i \in 1 \dots, n. \tag{A.1}$$

Furthermore, if $C(t)$ is upper triangular for every $t \in [0, T_n]$ we say that the volatility structure is upper-triangular separable.

Proposition A.2. *Suppose that $\mathcal{M}$ is a d-factor LMM then the following statements are equivalent,*

1. *$\mathcal{M}$ can be parametrised by a matrix separable volatility structure*
2. *$\mathcal{M}$ can be parametrised by an upper-triangular separable volatility structure driven by a d-dimensional Brownian motion with independent components*
3. *$\mathcal{M}$ can be parametrised by a separable volatility structure as in Definition 3.1.*

Proof. Recall that two parameterisations of LMM $(\{\sigma^i\}_{i=1}^n, \rho)$ and $(\{\tilde{\sigma}^i\}_{i=1}^n, \tilde{\rho})$ result in the same dynamics if the instantaneous volatility and correlations implied by them are the same or equivalently if

$$\langle \sigma^i(t), \rho(t)\sigma^j(t)\rangle = \langle \tilde{\sigma}^i(t), \tilde{\rho}(t)\tilde{\sigma}^j(t)\rangle, \quad t \le T_i \wedge T_j, i, j = 1, \dots, n. \tag{A.2}$$

1. ⇒ *2.* Let $(\{\sigma^i\}_{i=1}^n, \rho)$ be a matrix separable parameterisation of a d-factor LMM. Then there exist matrix valued function C and vectors $v^1, \dots, v^n$ such that (A.1) holds. Then

$$\langle \sigma^i(t), \rho(t)\sigma^j(t)\rangle = (v^i)^T C(t)^T \rho(t) C(t) v^j \tag{A.3}$$

and note that $C(t)^T \rho(t) C(t)$ is a positive semi-definite matrix. Then there exists an upper triangular matrix $\tilde{C}(t)$ (Cholesky decomposition) such that $C(t)^T \rho(t) C(t) = \tilde{C}(t)^T \tilde{C}(t)$ and therefore $(\{\sigma^i\}_{i=1}^n, \tilde{\rho})$, where $\tilde{\rho}(t)$ is the $d \times d$ identity matrix, is an upper-triangular separable parameterisation equivalent to $(\{\sigma^i\}_{i=1}^n, \rho)$.

2. ⇒ *3.* Let $(\{\sigma^i\}_{i=1}^n, \rho)$, where $\rho(t)$ is the $d \times d$ identity matrix, be an upper-triangular separable parameterisation of a d-factor LMM. Then

$$\langle \sigma^i(t), \rho(t)\sigma^j(t)\rangle = (v^i)^T C(t)^T C(t) v^j \tag{A.4}$$

and note that $B(t) := C(t)^T C(t)$ is a positive semidefinite matrix. Denote by $b_{i,i}(t)$ the ith element of the diagonal of matrix $B(t)$. Since $B(t)$ is positive semidefinite $b_{i,i}(t) \geq 0$ and we can define a diagonal matrix $A(t)$ with $a_{i,i}(t) := \sqrt{b_{i,i}(t)}$ as the ith element of its diagonal. Then there exists a correlation matrix $\tilde{\rho}$ such that $A(t)\tilde{\rho}(t)A(t) = B(t)$. We can then define a vector valued function $\tilde{\sigma}$ by

$$\tilde{\sigma}(t) := [a_{1,1}(t), \dots, a_{d,d}(t)]^T. \tag{A.5}$$

Then $(\{v^i * \tilde{\sigma}\}_{i=1}^n, \tilde{\rho})$ is a separable parameterisation equivalent to $(\{\sigma^i\}_{i=1}^n, \rho)$ since

$$\langle v^i * \tilde{\sigma}(t), \tilde{\rho}(t)(v^j * \tilde{\sigma}(t))\rangle = (v^i)^T A(t)\tilde{\rho}(t)A(t)v^j = \langle \sigma^i(t), \rho(t)\sigma^j(t)\rangle \tag{A.6}$$

3. ⇒ *1.* Obvious. □

B Proofs

Proof of Theorem 4.1. Let functions σ^{inst}, σ, v and $\rho_{1,2}$ be defined as in the statement of the theorem and assume that they satisfy Equation (4.10). Next we define functions $g_i, h_i, i = 1, 2, 3$ and f, as in Equations (4.11)–(4.15) and observe that they are a solution to the functional equation

$$f(T-t) = \sum_{i=1}^{3} g_i(t) h_i(T). \tag{B.1}$$

Since we assumed that the functions σ, v and $\rho_{1,2}$ are continuous this implies that functions $g_i, h_i, i = 1, 2, 3$ and f are continuous. Then any solution f to Equation (B.1) is of the form

$$f(y) = \sum_i P_i(y) \exp(-\lambda_i y), \tag{B.2}$$

where $\lambda_i \in \mathbb{C}$, P_i is a polynomial (with possibly complex coefficients) and $\sum_i (1 + \deg P_i) = 3$ (see discussion in Section 4). However, we are only interested in the solutions for which f is a non-negative real-valued function on $\mathbb{R}_+$. Therefore f has to be of one of the following forms:

1. $x_i, \lambda_i \in \mathbb{R}, i = 1, 2, 3$

$$f(y) = \sum_{i=1}^{3} x_i \exp(-\lambda_i y); \tag{B.3}$$

2. $x_1 \neq 0,\ x_2, x_3, \lambda_1, \lambda_2 \in \mathbb{R}$

$$f(y) = (x_1 y + x_2)\exp(-\lambda_1 y) + x_3 \exp(-\lambda_2 y); \tag{B.4}$$

3. $x_1 > 0,\ x_2, x_3, \lambda \in \mathbb{R}$

$$f(y) = x_1(y^2 + x_2 y + x_3)\exp(-\lambda_1 y); \tag{B.5}$$

4. $x_1, \lambda_1 \in \mathbb{C} \setminus \mathbb{R},\ x_2, \lambda_2 \in \mathbb{R}$

$$f(y) = x_1 \exp(-\lambda_1 y) + \overline{x}_1 \exp(-\overline{\lambda}_1 y) + x_2 \exp(-\lambda_2 y), \tag{B.6}$$

where $\overline{x}$ denotes the complex conjugate of x.

Before analysing the possibilities let us make an observation that Equations (4.14) and (4.15) imply

$$h_3(T)^2 = h_1(T)h_2(T), \tag{B.7}$$

which will be used throughout the proof.

Case 1. We first analyse the case when f is of the form as in Equation (B.3). We can assume without loss of generality that $\lambda_1 < \lambda_2 < \lambda_3$. Clearly f will be of desired form if the functions $g_i, h_i, i = 1, 2, 3$ are of the form,

$$g_i(t) = \alpha_i^2 \exp(2\beta_i t), \quad i = 1, 2 \tag{B.8}$$

$$g_3(t) = 2\gamma\alpha_1\alpha_2 \exp((\beta_1 + \beta_2)t), \tag{B.9}$$

$$h_i(T) = \exp(-2\beta_i T), \quad i = 1, 2, \tag{B.10}$$

$$h_3(T) = \exp(-(\beta_1 + \beta_2)T), \tag{B.11}$$

where $\alpha_1, \alpha_2 \geq 0,\ \beta_1, \beta_2 \in \mathbb{R}$ and $\gamma \in [-1, 1]$. Then f is given by

$$f(x) = \alpha_1^2 \exp(-2\beta_1 x) + \alpha_2^2 \exp(-2\beta_2 x) + 2\gamma\alpha_1\alpha_2 \exp(-(\beta_1 + \beta_2)x) \tag{B.12}$$

and functions v, σ and $\rho_{1,2}$ are given by

$$v(T) = \begin{bmatrix} \exp(-\beta_1 T) \\ \exp(-\beta_2 T) \end{bmatrix}, \tag{B.13}$$

$$\sigma(t) = \begin{bmatrix} \sqrt{\alpha_1}\exp(\beta_1 t) \\ \sqrt{\alpha_1}\exp(\beta_2 t) \end{bmatrix}, \tag{B.14}$$

$$\rho_{1,2}(t) = \gamma. \tag{B.15}$$

Note that this is indeed a valid parameterisation since $\rho_{1,2}(t) = \gamma \in [-1, 1]$.

Next we show that one cannot get a more general parameterisation. We will refer to the parameterisation given by Equations (B.13)–(B.14) as the 'original parameterisation'. Since f, $g_i, h_i, i = 1, 2, 3$, solve the Equation (4.16) and f is of the from as in Equation (B.1) $g_i, h_i, i = 1, 2, 3$, have to be of the form

$$g_i(t) = \alpha_i \exp(\lambda_1 t) + \beta_i \exp(\lambda_2 t) + \gamma_i \exp(\lambda_3 t), \tag{B.16}$$

$$h_i(T) = a_i \exp(-\lambda_1 T) + b_i \exp(-\lambda_2 T) + c_i \exp(-\lambda_3 T). \tag{B.17}$$

To show that a different parameterisation offers no generality over the one proposed above we can either show that the functions h_1, h_2, h_3 are linearly dependent or that $(\sum_{i=1}^n b_i\beta_i)^2 \leq 4(\sum_{i=1}^n a_i\alpha_i)(\sum_{i=1}^n c_i\gamma_i)$ and $(\sum_{i=1}^n c_i\gamma_i) > 0$.[4]

[4] Note that since f is non-negative any relevant solution will automatically satisfy $x_1 = \sum_{i=1}^n a_i\alpha_i \geq 0$

When functions h_1, h_2, h_3 are linearly dependent, then there exist constants $\xi_1, \xi_2, \xi_3 \in \mathbb{R}$ such that $\max\{\xi_1^2, \xi_2^2, \xi_3^2\} > 0$ and

$$\sum_{i=1}^{3} \xi_i h_i(T) = 0, \qquad T \geq 0. \tag{B.18}$$

Then at least one of constants ξ_1 ξ_2 is non-zero. Without loss of generality $\xi_1 = 1$. Then

$$f(T-t) = \sum_{i=1}^{3} g_i(t)h_i(T) = \sum_{i=2}^{3}(g_i(t) - \xi_i g_1(t))h_i(T) \tag{B.19}$$

and f can only be a sum of two exponential functions, thus less general than the original parameterisation.

We would now like to use the restriction in the Equation (B.7). Note that we have two consider two possibilities: $\lambda_2 = \frac{1}{2}(\lambda_1 + \lambda_2)$ and $\lambda_2 \neq \frac{1}{2}(\lambda_1 + \lambda_2)$.

First assume that $\lambda_2 \neq \frac{1}{2}(\lambda_1 + \lambda_3)$. Then Equation (B.7) implies

$$a_1 a_2 = a_3^2, \tag{B.20}$$
$$b_1 b_2 = b_3^2, \tag{B.21}$$
$$c_1 c_2 = c_3^2, \tag{B.22}$$
$$a_1 b_2 + a_2 b_1 = 2a_3 b_3, \tag{B.23}$$
$$a_1 c_2 + a_2 c_1 = 2a_3 c_3, \tag{B.24}$$
$$b_1 c_2 + b_2 c_1 = 2b_3 c_3. \tag{B.25}$$

Suppose that $a_1 = 0$, then $a_3 = 0$ and we are only interested in this parameterisation if $a_2 \neq 0$. Which in turn implies $b_1 = c_1 = 0$ and therefore $b_3 = c_3 = 0$. Thus $h_1 = h_3 \equiv 0$ and $f(T-t) = g_2(t)h_2(T)$ which leads to f being an exponential, thus offering no generality over our original parameterisation. Note that the problem is symmetrical thus we can assume that $a_i, b_i, c_i \neq 0, i = 1, 2, 3$.

Then it is easy to observe that

$$(a_1 b_2 - a_2 b_1)^2 = 0, \tag{B.26}$$
$$(a_1 c_2 - a_2 c_1)^2 = 0, \tag{B.27}$$
$$(b_1 c_2 - b_2 c_1)^2 = 0 \tag{B.28}$$

and therefore

$$a_1 b_2 = a_2 b_1, \qquad a_1 c_2 = a_2 c_1, \qquad b_1 c_2 = b_2 c_1. \tag{B.29}$$

Since $a_i, b_i, c_i \neq 0, i = 1, 2, 3$, then

$$\frac{a_2}{a_1} = \frac{b_2}{b_1} = \frac{c_2}{c_1} = x > 0 \tag{B.30}$$

and therefore $h_2 = xh_1$, hence functions h_1, h_2 and h_3 are linearly dependent, thus offering no generality over the original parameterisation.

therefore we need to assume $\lambda_2 = \frac{1}{2}(\lambda_1 + \lambda_3)$. In this case Equation (B.7) implies

$$a_1 a_2 = a_3^2, \tag{B.31}$$
$$c_1 c_2 = c_3^2, \tag{B.32}$$
$$a_1 b_2 + a_2 b_1 = 2a_3 b_3, \tag{B.33}$$
$$b_1 c_2 + b_2 c_1 = 2b_3 c_3, \tag{B.34}$$

$$b_1b_2 + a_1c_2 + a_2c_1 = b_3^2 + 2a_3c_3. \tag{B.35}$$

Without loss of generality we can assume that $a_1 = 1$ and split the analysis into two cases $a_2 \neq 0$ and $a_2 = 0$.

First let us consider the case when $a_2 = 0$. Then Equation (B.31) implies $a_3 = 0$ and consequently Equation (B.33) implies $b_2 = 0$. Then it is easy to deduce

$$c_2 = b_3^2, \qquad b_1c_2 = 2b_3c_3, \qquad c_1c_2 = c_3^2. \tag{B.36}$$

Note that if $b_3 = 0$ then $c_3 = 0$ and therefore $h_3 \equiv 0$, this would clearly offer no generality over original parameterisation. Therefore we can assume $b_3 \neq 0$ and we can express b_1, c_1, c_2 in terms of b_3 and c_3

$$b_1 = 2\frac{c_3}{b_3}, \qquad c_1 = \frac{c_3^2}{b_3^2}, \qquad c_2 = b_3^2. \tag{B.37}$$

We can now turn back to Equation (B.1), since we know that f is of the from as in Equation (B.3), we know that the sum of coefficients in front of term of the form $\exp(-\lambda_i T + \lambda_j t)$ is 0 when $i \neq j$. Therefore $\beta_1 = \gamma_1 = 0$ and

$$\frac{c_3}{b_3}\alpha_1 + b_3\alpha_3 = 0 \tag{B.38}$$

$$\frac{c_3^2}{b_3^2}\alpha_1 + b^2\alpha_2 + c_3\alpha_3 = 0 \tag{B.39}$$

then $\alpha_2 = 0$.

On the other hand we know that $4g_1(t)g_2(t) \geq g_3(t)^2$, in particular can compare the exponents in front of the largest exponentials

$$0 = 4\gamma_1\gamma_2 \geq \gamma_3^2 \tag{B.40}$$

thus $\gamma_3 = 0$. Then

$$4\alpha_1\gamma_2 = 4(\alpha_1\gamma_2 + \alpha_2\gamma_1 + \beta_1\beta_2) \geq \beta_3^2 + 2\alpha_3\gamma_3 = \beta_3^2. \tag{B.41}$$

Now recall that in the original parameterisation $4x_1x_3 \geq x_2^2$ and $x3 \geq 0$. Note that since g_2 is non-negative function the coefficient in front of largest exponential needs to be non-negative, i.e. $\gamma_2 \geq 0$, then $x_3 = \sum_{i=1}^{3} c_i\gamma_i = b_3^2\gamma_2 \geq 0$. Furthermore,

$$4x_1x_3 = \left(\sum_{i=1}^{3} a_i\alpha_i\right)\left(\sum_{i=1}^{3} c_i\gamma_i\right) = 4\alpha_1\gamma_2b_2^2 \geq b_3^2\beta_3^2 = \left(\sum_{i=1}^{3} b_i\beta_i\right)^2 = x_2^2, \tag{B.42}$$

showing that such parameterisation offers no generality over the original one.

Finally we need to consider the case when $\alpha_2 \neq 0$. In this case we can without loss of generality assume $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Next we show that when $b_1c_1 = 0$ and $b_2c_2 = 0$ this offers no generality over the original parameterisation.

Suppose that $b_1 = 0$ then $b_2 = 2b_3$ and clearly b_2 cannot be zero as this would offer no generality over the original parameterisation. Therefore we can assume that $b_2 \neq 0$. Now observe that if $c_2 = 0$ then $c_3 = 0$ and $b_2c_1 = 0$. Note that we assumed that $b_2 \neq 0$ therefore $c_1 = c_2 = c_3 = 0$ and the parameterisation offers no generality over the original parameterisation. In particular when $b_1 = 0$ this can assume that $b_2c_2 \neq 0$. By symmetry we can also assume that when $b_2 = 0$ then $b_1c_1 \neq 0$.

Now suppose that $b_1 \neq 0$ and $c_1 = 0$. Then $c_3 = 0$ and we can assume that $c_2 \neq 0$. On the other hand $b_1c_2 = 2b_3c_3 = 0$ contradicting our assumptions that b_1 and c_2 are non-zero. Similar argument holds when $b_2 \neq 0$ and $c_2 = 0$.

We can now assume without loss of generality that $b_1 c_1 \neq 0$. Next observe that $c_1 c_2 \geq 0$ and

$$
\begin{aligned}
b_1 c_2 + b_2 c_1 - 2\,\mathrm{sgn}(c_3)\sqrt{c_1 c_2} &= 0 && \text{(B.43)}\\
\mathrm{sgn}(c_1) b_1 \left|c_2\right| + \mathrm{sgn}(c_1) b_2 \left|c_1\right| - 2\,\mathrm{sgn}(c_3)\sqrt{c_1 c_2} &= 0 && \text{(B.44)}\\
b_1 \left|c_2\right| + b_2 \left|c_1\right| - 2\,\mathrm{sgn}(c_1 c_3)\sqrt{c_1 c_2} &= 0 && \text{(B.45)}\\
(b_1 \sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3) b_2 \sqrt{|c_1|})(\sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3)\sqrt{|c_1|}) &= 0 && \text{(B.46)}
\end{aligned}
$$

Therefore either $\sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3)\sqrt{|c_1|} = 0$ or $b_1 \sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3) b_2 \sqrt{|c_1|} = 0$.

Note that $\sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3)\sqrt{|c_1|} = 0$ if and only if $c_1 = c_2$ and $\mathrm{sgn}(c_1 c_3) = 1$, therefore $c_1 = c_2 = c_3$ and functions h_1, h_2, h_3 are linearly dependent and such parameterisation can offer no generality over the original parameterisation.

We can then assume that $b_1 \sqrt{|c_2|} - \mathrm{sgn}(c_1 c_3) b_2 \sqrt{|c_1|} = 0$. In particular this implies

$$
\begin{aligned}
b_1 \sqrt{|c_2|} &= \mathrm{sgn}(c_1 c_3) b_2 \sqrt{|c_1|} && \text{(B.47)}\\
\frac{b_2}{b_1} &= \mathrm{sgn}(c_1 c_3) \frac{\sqrt{|c_2|}}{\sqrt{|c_1|}} =: x \in \mathbb{R}. && \text{(B.48)}
\end{aligned}
$$

Then

$$
b_2 = b_1 x, \qquad b_3 = \frac{1}{2} b_1 (1 + x), \qquad c_2 = c_1 x^2, \qquad c_3 = c_1 x,
$$

where $b_1, c_1 \neq 0$ and $x \in \mathbb{R}$. Next we use the last remanning condition.

$$
\begin{aligned}
b_1 b_2 + c_1 + c_2 &= b_3^2 + 2 c_3 && \text{(B.49)}\\
4 c_1 (1 + x^2) + 4 b_1^2 x &= b_1^x (1 + x)^2 + 8 c_1 x && \text{(B.50)}\\
4 c_1 (1 - x)^2 &= b_1^2 (1 - x)^2 && \text{(B.51)}
\end{aligned}
$$

Note that when $x = 1$ then $h_1 = h_2 = h_3$ and therefore not offering any generality over the original parameterisation. We can then assume $x \neq 1$ and consequently

$$
c_1 = \frac{1}{4} b_1^2. \tag{B.52}
$$

Next we show that $x_3 \geq 0$ and $4 x_1 x_3 \geq x_2^2$. First note that coefficient in front of the term $\exp(-\lambda_i T + \lambda_j t)$ is 0 when $i \neq j$. In particular

$$
\begin{aligned}
0 &= \sum_{i=1}^{3} a_i \gamma_i = \gamma_1 + \gamma_2 + \gamma_3 && \text{(B.53)}\\
0 &= \sum_{i=1}^{3} b_i \gamma_i = b_1 \gamma_1 + b_1 x \gamma_2 + \frac{1}{2} b_1 \gamma_3 && \text{(B.54)}
\end{aligned}
$$

and therefore $\gamma_2 = \gamma_1$ and $\gamma_3 = -2\gamma_1$, furthermore since g_1 is non-negative $\gamma_1 \geq 0$ and

$$
x_3 = \sum_{i=1}^{3} c_i \gamma_i = \frac{1}{4} c_1 \gamma_1 (1 - x)^2 \geq 0. \tag{B.55}
$$

Next we prove that $4x_1x_3 \geq x_2^2$. Recall that $4g_1(t)g_2(t) \geq g_3(t)^2, T \geq 0$, then

$$\begin{aligned} 0 &\leq 4g_1(T)g_2(T) - g_3(T)^2 \\ &= (4\gamma_1\gamma_2 - \gamma_3^2)\exp(2\lambda_3 t) \\ &\quad + (4\beta_1\gamma_2 + 4\beta_2\gamma_1 - 2\beta_3\gamma_3)\exp((\lambda_2+\lambda_3)t) \\ &\quad + (4\beta_1\beta_2 + 4\alpha_1\gamma_2 + 4\alpha_2\gamma_1 - \beta_3^2 - 2\alpha_3\gamma_3)\exp(2\lambda_2 t) + \dots \end{aligned} \tag{B.56}$$

$$\begin{aligned} &= 4\gamma_1(\beta_1+\beta_2+\beta_3)\exp((\lambda_2+\lambda_3)t) \\ &\quad + (4\beta_1\beta_2 - \beta_3^2 + 4\gamma_1(\alpha_1 + 4\alpha_2 + 2\alpha_3))\exp(2\lambda_2 t) + \dots \end{aligned} \tag{B.57}$$

Next observe that $0 = \sum_{i=1}^3 a_1\beta_i = \sum_{i=1}^3 \beta_i$ and therefore

$$4\beta_1\beta_2 - \beta_3^2 + 4\gamma_1(\alpha_1 + 4\alpha_2 + 2\alpha_3) \geq 0 \tag{B.58}$$

$$4\gamma_1(\alpha_1 + 4\alpha_2 + 2\alpha_3) \geq (-\beta_1 - \beta_2)^2 - 4\beta_1\beta_2 \tag{B.59}$$

$$\gamma_1(\alpha_1 + 4\alpha_2 + 2\alpha_3) \geq \frac{1}{4}(\beta_1 - \beta_2)^2 \tag{B.60}$$

Then

$$x_2^2 = \left(\sum_{i=1}^3 b_i\beta_i\right)^2 \tag{B.61}$$

$$= b_1^2\Big(\beta_1 + \beta_2 x + \frac{1}{2}(-\beta_1 - \beta_2)(1+x)\Big)^2 \tag{B.62}$$

$$= \frac{1}{4}b_1^2(1-x)(\beta_1 - \beta_2)^2 \tag{B.63}$$

$$\leq b_1^2\gamma_1(\alpha_1 + \alpha_2 + \alpha_3)(1-x)^2. \tag{B.64}$$

On the other hand

$$4x_1x_3 = 4(\alpha_1 + \alpha_2 + \alpha_3)\frac{1}{4}c_1\gamma_1(1-x)^2 \tag{B.65}$$

$$= b_1^2\gamma_1(\alpha_1 + \alpha_2 + \alpha_3)(1-x)^2 \tag{B.66}$$

and therefore $4x_1x_3 \geq x_2^2$ and this parameterisation offers no generality over the original one.

Case 2. Next we analyse the case when f is of the form as in Equation (B.4). Then $h_i, 1, 2, 3$, is of the form

$$h_i(T) = (a_iT + b_i)\exp(-\lambda_1 T) + c_i\exp(-\lambda_2 T), \tag{B.67}$$

where $\lambda_1 \neq \lambda_2$. Then Equation (B.7) implies

$$(a_1T + b_1)(a_2T + b_2) = (a_3T + b_3)^2 \tag{B.68}$$

Note that $a_1 = 0$ implies $a_2 = a_3 = 0$ reducing the problem to the Case 1. Similarly $a_2 = 0$ implies $a_1 = a_3 = 0$. Then we can without loss of generality assume $a_1, a_2a_3 \neq 0$ and therefore

$$a_1a_2(T + \frac{b_1}{a_1})(T + \frac{b_2}{a_2}) = a_3^2(T + \frac{b_3}{a_3})^2. \tag{B.69}$$

(B.70)

Therefore

$$\frac{b_1}{a_1} = \frac{b_2}{a_2} = \frac{b_3}{a_3} = x \tag{B.71}$$

and

$$h_i(T) = a_i(T + x)\exp(-\lambda_1 T) + c_i \exp(-\lambda_2 T). \tag{B.72}$$

Now we can use (B.7) again and obtain

$$a_1c_2(T + x) + a2c_1(T + x) = 2a_3c_3(T + x) \tag{B.73}$$

$$a_1a_2 = a_3^2 \tag{B.74}$$

$$c_1c_2 = c_3^2 \tag{B.75}$$

And it is straightforward to deduce

$$(a_1c_2 - a_2c_1)^{=}0. \tag{B.76}$$

It then follows that

$$\frac{c_1}{a_1} = \frac{c_2}{a_2} = x' \tag{B.77}$$

and

$$a_1a_2 = a_3c_3x'. \tag{B.78}$$

Therefore $\frac{c_3}{a_3} = x'$, $h_2 = \frac{a_2}{a_1}h_1$ and $h_3 = \frac{a_3}{a_1}h_1$. Then f sovles

$$f(T - t) = g(t)h(T), \tag{B.79}$$

where $g = g_1 + \frac{a_2}{a_1}g_2 + \frac{a_3}{a_1}g_3$ and $h = h_1$. In particular, f is an exponential thus contradicting the assumption that f is of the form as in Equation (B.4) where $x_1 \neq 0$.

Case 3. Next we analyse the case when f is of the form as in Equation (B.5). Then clearly $g_i, h_i, i = 1, 2, 3$, are of the form

$$g_i(t) = p_i(t)\exp(\lambda_1 t), \tag{B.80}$$

$$h_i(T) = q_i(T)\exp(-\lambda_1 t), \tag{B.81}$$

where p_i and q_i are polynomials of degree two or less. Without loss of generality we can assume that the leading coefficient in polynomials $q_i, i = 1, 2, 3$, is equal to 1. Now denote by P the quadratic polynomial $P(y) = f(y)\exp(\lambda_1 y)$ and note that $P(T - t) = \sum_{i=1}^{3} p_i(t)q_i(T)$.

Furthermore, observe that Equation (B.7) implies that polynomials q_1 and q_2 have to be of the same parity and at least one of the polynomials q_1 and q_2 has to be quadratic. We can therefore without loss of generality assume that q_1 is quadratic polynomial. We then have two possibilities: $\deg q_2 = 0$ and $\deg q_2 = 0$.

We first explore the latter. When $\deg q_1 = 1$ and $\deg q_2 = 0$ the then Equation (B.7) implies that polynomials q_1, q_2, q_3 are of the form

$$q_1(T) = (T + c)^2, \quad q_2(T) = 1, \quad q_3(T) = T + c. \tag{B.82}$$

Then

$$P''(T - t) = \sum_{i=1}^{3} p_i(t)q_i''(T) = 2p_1(t) \tag{B.83}$$

and therefore $p_i(t) = a$. Furthermore since g_1 is non-negative function $a > 0$ (note that $a = 0$ leads to Case 1). Next observe that

$$P'(T - t) = \sum_{i=1}^{3} p_i(t)q_i'(T) = 2a(T + c) + p_3(t) \tag{B.84}$$

and p_2 has to be of the form $p_3(t) = -2a(t+b)$. Finally,

$$P(T-t) = a(T+c)^2 + p_2(t) - 2a(t+b)(T+c) \tag{B.85}$$
$$= a(T-t+c-b)^2 - a(t+b)^2 + p_2(t) \tag{B.86}$$

and $p_2(t) = a(t+b)^2 + d$. Then f, v and σ are of the form

$$f(T-t) = (a(T-t+c-b)+d)\exp(-\lambda_1(T-t)) \tag{B.87}$$
$$v(T) = \begin{bmatrix} (T+c)\exp(-\frac{1}{2}\lambda_1 T) \\ \exp(-\frac{1}{2}\lambda_1 T) \end{bmatrix}, \tag{B.88}$$
$$\sigma(t) = \begin{bmatrix} \sqrt{a}\exp(\frac{1}{2}\lambda_1 t) \\ \sqrt{a(t+b)^2+d}\exp(\frac{1}{2}\lambda_1 t) \end{bmatrix}. \tag{B.89}$$

and we can determine $\rho_{1,2}$ from Equation (4.13)

$$\rho_{1,2} = -\frac{(t+b)}{\sqrt{(t+b)^2+d}}. \tag{B.90}$$

Note that since $\rho_{1,2}(t) \in [-1,1]$ this implies $d \geq 0$, moreover since f depends only on the parameters difference b, c only through their difference we can set $c = 0$. And by introducing $\alpha = \sqrt{a}, \beta = b, \gamma = d$ and $\lambda = \frac{1}{2}\lambda_1$ we obtain the same parameterisation as in the statement of the theorem.

Now let us analyse the case when $\deg q_1 = \deg q_2 = 2$. Note that Equation (B.7) implies that when q_1 has two distinct roots then q_2 and q_3 have the same roots as q_1 and therefore functions h_1, h_2 and h_3 are linearly dependent which in turn implies that f cannot be a quadratic polynomial multiplied by an exponential. Then q_1, q_2 and q_3 have to be of the form

$$q_1 = (T+c)^2, \qquad q_2 = (T+d)^2, \qquad q_3 = (T+c)(T-d), \tag{B.91}$$

where $c \neq d$. We can now follow the same argument as in the previous case and observe that

$$P''(T-t) = 2(p_1(t)+p_2(t)+p_3(t)) = 2a > 0 \tag{B.92}$$

and

$$P'(T-t) = 2(T+c)p_1(t) + 2(T+d)p_2(t) + (2T+c+d)p_3(t) \tag{B.93}$$
$$= 2aT + 2cp_2(t)2dp_2(t) + (c+d)p_3(t) \tag{B.94}$$

therefore

$$2cp_2(t)2dp_2(t) + (c+d)p_3(t) = -2a(t+b). \tag{B.95}$$

Then

$$P(T-t) = (T+c)^2 p_1(t) + (T+d)^2 p_2(t) + (T+c)(T+d)p_3(t) \tag{B.96}$$
$$= aT^2 + 2aT(t+b) + c^2 p_1(t) + d^2 p_2(t) + cdp_3(t) \tag{B.97}$$

and therefore

$$c^2 p_1(t) + d^2 p_2(t) + cdp_3(t) = a((t+b)^2 + e). \tag{B.98}$$

Note that p_1, p_2, p_3 then solve the following system of linear equations

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 2c & 2d & c+d \\ c^2 & d^2 & cd \end{bmatrix}}_{:=A} \begin{bmatrix} p_1(t) \\ p_2(t) \\ p_3(t) \end{bmatrix} = \begin{bmatrix} a \\ -2a(t+b) \\ a((t+b)^2+e) \end{bmatrix} \tag{B.99}$$

The determinat of matrix A is $(c-d)^3$ and since we assumed $c \neq d$ matrix A is invertible and system is well defined and we can determine functions p_1, p_2, p_3

$$\begin{bmatrix} p_1(t) \\ p_2(t) \\ p_3(t) \end{bmatrix} = \frac{1}{(c-d)^2} \begin{bmatrix} a((t+b+d)^2+e) \\ a((t+c+d)^2+e) \\ -2a((t+b+c)(t+b+d)+e) \end{bmatrix}. \tag{B.100}$$

Finally, we need to find the conditions under which $4p_1(t)p_2(t) \geq p_3(t)^2$ so that $\rho_{1,2}$ will take values in the interval $[-1,1]$. In particular

$$0 \leq 4p_1(t)p_2(t) - p_3(t)^2 \tag{B.101}$$

$$= \frac{4a^2e}{(c-d)^2} \tag{B.102}$$

and therefore $e > 0$. Then f is of the form

$$f(y) = a((y+b)^2+e)\exp(-\lambda y), \tag{B.103}$$

where $a > 0$, $b \in \mathbb{R}$ and $e \geq 0$. Therefore this parameterisation offers no generality over the previous case.

Case 4 Finally we analyse the case when f is of the form as in Equation (B.6). We will denote the imaginary unit by ι to avoid confusion with index i. Without loss of generality we can take $\lambda_1 = \lambda - \iota\theta, \lambda, \theta \in \mathbb{R}$ and $x_1 = \frac{1}{2}(u - \iota v)$ and rewrite f as

$$f(y) = (u\cos(\theta y) + v\sin(\theta y))\exp(-\lambda y) + x_3\exp(-\lambda_2 y), \tag{B.104}$$

and clearly $g_i, h_i, i = 1,2,3$ have to be of the form

$$g_i(t) = (\alpha_i\cos(\theta t) + \beta_i\sin(\theta t))\exp(\lambda t) + \gamma_i\exp(\lambda_2 t), \tag{B.105}$$

$$h_i(T) = (a_i\cos(\theta T) + b_i\sin(\theta T))\exp(-\lambda T) + c_i\exp(-\lambda_2 T), \tag{B.106}$$

for some $\alpha_i, \beta_i, \gamma_i, a_i, b_i, c_i \in \mathbb{R}$.

If $\lambda \neq \lambda_2$ Equation (B.7) implies

$$a_1a_2 = a_3^2, \tag{B.107}$$

$$b_1b_2 = b_3^2, \tag{B.108}$$

$$c_1c_2 = c_3^2, \tag{B.109}$$

$$a_1b_2 + a_2b_1 = 2a_3b_3, \tag{B.110}$$

$$a_1c_2 + a_2c_1 = 2a_3c_3, \tag{B.111}$$

$$b_1c_2 + b_2c_1 = 2b_3c_3. \tag{B.112}$$

Recall that we got the same set of Equations in Case 1 ($\lambda_2 \neq \frac{1}{2}(\lambda_1+\lambda_3)$) in particular it is easy to observe that vectors $(a_i, b_i, c_i), i = 1,2,3$, are co-linear, thus implying that f is a solution of Pexider Equation, in particular $u = v = 0$. Thus it offers no generality over Case 1.

We can therefore assume $\lambda_2 = \lambda$, note that since $h_1, h_2 \geq 0$ this implies $c_1, c_2 > 0$[5] and we can without loss of generality assume $c_1 = c_2 = 1$. Then Equation (B.6) implies

$$a_1a_2 + 1 = a_3^2 + c_3^2, \tag{B.113}$$

[5] If $c_1 = 0$ it follows $h_1 = h_3 \equiv 0$ and similarly if $c_2 = 0$ then $h_2 = h_3 \equiv 0$ and in both cases it is trivial to observe that f is an exponential function.

$$b_1 b_2 + 1 = b_3^2 + c_3^2, \tag{B.114}$$
$$a_1 b_2 + a_2 b_1 = 2a_3 b_3, \tag{B.115}$$
$$a_1 + a_2 = 2a_3 c_3, \tag{B.116}$$
$$b_1 + b_2 = 2b_3 c_3. \tag{B.117}$$

Let us first consider the case when $a_1 + a_2 = 0$. Then either $a_3 = 0$ or $c_3 = 0$. Note, that $a_3 = 0$ would imply that functions h_1, h_2 and h_3 are linearly dependent which cannot be the case. Therefore $c_3 = 0$ and Equation (B.117) implies $b_1 + b_2 = 0$. Let us define $a := a_1 = -a_2$ and $b := b_1 = -b_2$, then it follows from Equations (B.113)–(B.115)

$$a^2 + b^2 = 1, \qquad a_3^2 = b^2, \qquad b_3^2 = a^2 \tag{B.118}$$

moreover $a_3 b_3 = -ab$. We can then assume without loss of generality that $b = \sqrt{1 - a^2}$ and

$$h_1(T) = (a\cos(\theta T) + b\sin(\theta T) + 1)\exp(-\lambda T), \tag{B.119}$$
$$h_2(T) = (-a\cos(\theta T) - b\sin(\theta T) + 1)\exp(-\lambda T), \tag{B.120}$$
$$h_3(T) = (b\cos(\theta T) - a\sin(\theta T))\exp(-\lambda T). \tag{B.121}$$

Next we would like to find the constraints on $\alpha_i, \beta_i, \gamma_i, i = 1, 2, 3$. First, recall that $cos(T - t)$ and $\sin(T - t)$ can be expanded as

$$\cos(T - t) = \cos T \cos t + \sin T \sin t, \tag{B.122}$$
$$\sin(T - t) = -\cos T \sin t + \sin T \cos t. \tag{B.123}$$

Therefore Equation (B.1) implies

$$a\gamma_1 - a\gamma_2 + b\gamma_3 = 0 \tag{B.124}$$
$$b\gamma_1 - b\gamma_2 - a\gamma_3 = 0 \tag{B.125}$$
$$\alpha_1 + \alpha_2 = 0 \tag{B.126}$$
$$\beta_1 + \beta_2 = 0 \tag{B.127}$$
$$a(\alpha_1 - \alpha_2) + b\alpha_3 = b(\beta_1 - \beta_2) - a\beta_3 \tag{B.128}$$
$$-a(\beta_1 - \beta_2) - b\beta_3 = b(\alpha_1 - \alpha_2) - a\alpha_3 \tag{B.129}$$

In particular, $\alpha := \alpha_1 = -\alpha_2$, $\beta := \beta_1 = -\beta_2$, $\gamma := \gamma_1 = \gamma_2$ and $\gamma_3 = 0$. Then we can find $\alpha_3 = 2\beta$ and $\beta_3 = -2\alpha$, hence

$$g_1(t) = (\alpha\cos(\theta t) + \beta\sin(\theta t) + \gamma)\exp(\lambda t), \tag{B.130}$$
$$g_2(t) = (-\alpha\cos(\theta t) - \beta\sin(\theta t) + \gamma)\exp(\lambda t), \tag{B.131}$$
$$g_3(t) = (2\beta\cos(\theta t) - 2\alpha\sin(\theta t))\exp(\lambda t). \tag{B.132}$$

Recall that g_2 and g_2 have to be non-negative, which implies that $\gamma > 0$ and $\alpha^2 + \beta^2 \leq \gamma^2$ and it is easy to verify that inequality $4g_1(t)g_2(t) \geq g_3(t)^2$.

Then, f is of the form

$$f(y) = 2((a\alpha + b\beta)\cos(\theta y) + (b\alpha - a\beta)\sin(\theta y) + \gamma)\exp(-\lambda y). \tag{B.133}$$

Moreover, f is a non-negative function if and only if $\gamma > 0$ and $\alpha^2 + \beta^2 \leq \gamma^2$ which is our assumption. Next observe that fixing $a = 0$ implies $b = 1$ and

$$f(y) = 2(\beta\cos(\theta y) + \alpha\sin(\theta y) + \gamma)\exp(-\lambda y). \tag{B.134}$$

Which clearly losses no generality. Finally, we need to determine functions v, σ and $\rho_{1,2}$. Note, that $v_i = \sqrt{h_i}, i = 1, 2$, will not work since $h_3 = v_1 v_2$ assumes values between $[-1, 1]$. We then have account the sign of the function h_3 while maintaining continuity of functions v_1 and v_2.

Note that

$$1 + \sin(\theta T) = 0 \Leftrightarrow \cos\frac{\theta T}{2} + \sin\frac{\theta T}{2} = 0, \tag{B.135}$$

$$1 - \sin(\theta T) = 0 \Leftrightarrow \cos\frac{\theta T}{2} - \sin\frac{\theta T}{2} = 0 \tag{B.136}$$

and

$$\cos(\theta x) = (\cos\frac{\theta x}{2} + \sin\frac{\theta x}{2})(\cos\frac{\theta x}{2} - \sin\frac{\theta x}{2}). \tag{B.137}$$

The v defined as

$$v(T) = \begin{bmatrix} \mathrm{sgn}(\cos\frac{\theta T}{2} + \sin\frac{\theta T}{2})\sqrt{1+\sin(\theta T)}\exp(-\frac{1}{2}\lambda T) \\ \mathrm{sgn}(\cos\frac{\theta T}{2} - \sin\frac{\theta T}{2})\sqrt{1-\sin(\theta T)}\exp(-\frac{1}{2}\lambda T) \end{bmatrix} \tag{B.138}$$

is a continuous function and $v_1(T)v_2(T) = h_3(T)$.

Now let us turn our attention to σ and $\rho_{1,2}$. When $\alpha^2 + \beta^2 < \gamma^2$ g_1 and g_2 are strictly positive functions and $4g_1(t)g_2(t) > g_3^2(t)$ for all t. Then σ and $\rho_{1,2}$ can be parametrised as

$$\sigma(t) = \begin{bmatrix} \sqrt{\alpha\cos(\theta t) + \beta\sin(\theta t) + \gamma}\exp(\frac{1}{2}\lambda t) \\ \sqrt{-\alpha\cos(\theta t) - \beta\sin(\theta t) + \gamma}\exp(\frac{1}{2}\lambda t) \end{bmatrix}, \tag{B.139}$$

$$\rho_{1,2} = \frac{\beta\cos(\theta t) - \alpha\sin(\theta t)}{\sqrt{\gamma^2 - (\alpha\cos(\theta t) + \beta\sin(\theta t))^2}} \tag{B.140}$$

Note that $\rho_{1,2}$ is well defined and continuous since $\alpha^2 + \beta^2 < \gamma^2$.

On the other hand when $\alpha^2 + \beta^2 = \gamma^2$ it follows that $4g_1(t)g_2(t) = g_3(t)^2$ for all t which implies $\rho_{1,2}(t)^2 = 1$, in particular since $\rho_{1,2}$ has to be a continuous function it has to be constant. We can then parameterise σ by first observing that a constant ϕ defined by

$$\phi = \begin{cases} \arccos\frac{\alpha}{\gamma}; & \beta \geq 0 \\ -\arccos\frac{\alpha}{\gamma}; & \beta < 0 \end{cases} \tag{B.141}$$

satisfies $\cos\phi = \frac{\alpha}{\gamma}$ and $\sin\phi = \frac{\beta}{\gamma}$. Then

$$\gamma + \alpha\cos(\theta t) + \beta\sin(\theta x) = 0 \Leftrightarrow \cos\frac{\theta t - \phi}{2} = 0, \tag{B.142}$$

$$\gamma - \alpha\cos(\theta t) - \beta\sin(\theta x) = 0 \Leftrightarrow \sin\frac{\theta t - \phi}{2} = 0 \tag{B.143}$$

and

$$\beta\cos(\theta t) - \alpha\sin(\theta t) = -2\cos\frac{\theta t - \phi}{2}\sin\frac{\theta t - \phi}{2}. \tag{B.144}$$

Then we can parametrise σ and $\rho_{1,2}$ as

$$\sigma(t) = \begin{bmatrix} \mathrm{sgn}(\cos\frac{\theta t - \phi}{2})\sqrt{\alpha\cos(\theta t) + \beta\sin(\theta t) + \gamma}\exp(\frac{1}{2}\lambda t) \\ -\mathrm{sgn}(\sin\frac{\theta t - \phi}{2})\sqrt{-\alpha\cos(\theta t) - \beta\sin(\theta t) + \gamma}\exp(\frac{1}{2}\lambda t) \end{bmatrix}, \tag{B.145}$$

$$\rho_{1,2} = 1 \tag{B.146}$$

in particular note that σ is a continuous function.

And we can get the parameterisation from the statement theorem by rescaling parameter $\lambda \mapsto 2\lambda$. □

Bibliography

J. Aczél. *Lectures on Functional Equations and Their Applications.* Academic Press, 1966.

L. Andersen and J. Andreasen. Volatility skews and extensions of the Libor market model. *Applied Mathematical Finance*, 7(1):1–32, 2000.

L. B. Andersen and V. V. Piterbarg. *Interest Rate Modeling.* Atlantic Financial Press, 2010.

M. N. Bennett and J. E. Kennedy. A Comparison of Markov-Functional and Market Models: The One-Dimensional Case. *The Journal of Derivatives*, 13(2):22–43, 2005.

F. Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3 (1-2):167–179, 1976.

A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7(2):127–155, 1997.

A. Carverhill. When is the short rate markovian? *Mathematical Finance*, 4(4): 305–312, 1994.

N. Denson and M. S. Joshi. Vega control. Avaliable at: http://dx.doi.org/10.2139/ssrn.1398523, 2009.

C. P. Fries and M. Rott. Cross currency and hybrid Markov functional models. Avaliable at: `http://ssrn.com/abstract=532122`, 2004.

J. Gogala and J. E. Kennedy. One-dimensional markov-functional models driven by non-gaussian markov processes. Submitted to Journal of Computational Finance, 2016.

D. Heath, R. Jarrow, and A. Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica: Journal of the Econometric Society*, 60(1):77–105, 1992.

P. J. Hunt, J. E. Kennedy, and A. Pelsser. Markov-functional interest rate models. *Finance and Stochastics*, 4(4):391–408, 2000.

C. Hunter, P. Jackel, and M. Joshi. Getting the drift. *Risk*, 14(7):81–84, 2001.

F. Jamshidian. LIBOR and swap market models and measures. *Finance and Stochastics*, 1(4):293–330, 1997.

M. S. Joshi. *More Mathematical Finance.* Pilot Whale, Melbourne [Australia], 2011.

M. S. Joshi and A. Stacey. New and robust drift approximations for the LIBOR market model. *Quantitative Finance*, 8(4):427–434, 2008.

R. Lord and A. Pelsser. Level–slope–curvature–fact or artefact? *Applied Mathematical Finance*, 14(2):105–130, 2007.

K. R. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with log-normal interest rates. *Journal of Finance*, 52(1): 409–430, 1997.

M. Musiela and M. Rutkowski. Continuous-time term structure models: Forward measure approach. *Finance and Stochastics*, 1(4):261–291, 1997.

R. Pietersz, A. Pelsser, and M. van Regenmortel. Fast drift-approximated pricing in the BGM model. *Journal of Computational Finance*, 8(1):93–124, 2004.

V. Piterbarg. Pricing and hedging callable libor exotics in forward libor models. *Journal of Computational Finance*, 8(2):65–119, 2004.

R. Rebonato. Calibrating the BGM model. *Risk*, 12(3):88–94, 1999.

R. Rebonato. *Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Model and Beyond*. Princeton University Press, 2002.