# FORWARD RATE DEPENDENT MARKOVIAN TRANSFORMATIONS OF THE HEATH-JARROW-MORTON TERM STRUCTURE MODEL

CARL CHIARELLA AND OH KANG KWON

*School of Finance and Economics*
*University ofTechnology Sydney*
*PO Box 123*
*Broadway NSW 2007*
*Australia*

`carl.chiarella@uts.edu.au`
`ohkang.kwon@uts.edu.au`

Abstract. In this paper, a class of forward rate dependent Markovian transformations of the Heath-Jarrow-Morton [HJM92] term structure model are obtained by considering volatility processes that are solutions of linear ordinary differential equations. These transformations generalise the Markovian systems obtained by Carverhill [Car94], Ritchken and Sankarasubramanian [RS95], Bhar and Chiarella [BC97], and Inui and Kijima [IK98], and also generalise the bond price formulae obtained therein.

## Introduction

In the risk neutral Heath-Jarrow-Morton [HJM92] term structure model, evolution of the forward rate process is completely determined by the forward rate volatilities. The HJM framework is very general and contains many of the earlier interest rate models as special cases, including [Vas77], [CIR85], [HW90], and [BK91], among others. One drawback of the generality, from a practical perspective, is that the model is non-Markovian in general and consequently does not readily lend itself to efficient solution techniques.

Suitable restrictions on volatility processes led Carverhill [Car94], Ritchken and Sankarasubramanian [RS95], Bhar and Chiarella [BC97], and Inui and Kijima [IK98] to transform the HJM model to *finite* dimensional Markovian systems. In [RS95] and [BC97] only the one-factor HJM models are considered, while, under a more transparent framework, [IK98] generalise the [RS95] models to the multifactor case. In [BG99] and [BS99], a theoretical framework is introduced for obtaining necessary and sufficient conditions under which HJM models are Markovian, and for constructing minimal realisation in such cases.

The [BC97] model has the feature that spot rate volatility may be an arbitrary function of the spot rate, and although the Markovian systems of [RS95] and [IK98] have the same feature, the bond price formulae obtained therein applies to a more general class of volatility processes, such as those which depend on a finite number of *fixed* tenor forward rates.[1] In each case, the forward rate volatility processes are expressible as a product of the spot rate volatility and a path-independent function. It should be noted that although [RS95] and [BC97] both consider the one-factor HJM model, they overlap only for a small set of volatility processes.

---

*Date*: First version January 6, 1998. Current revision March 30, 1999. Printed April 23, 1999.

[1] This fact does not appear to have been noted by the authors however.

In this paper, a *common* generalisation of the above models is obtained, in which the multifactor HJM model is transformed to a finite dimensional Markovian system, and in particular, a multifactor generalisation of the [BC97] model is obtained. Further, the volatility processes in the generalised models are allowed to be arbitrary functions of a finite number of fixed tenor forward rates. Consequently, they include finite dimensional forward rate dependent Markovian transformations of the multifactor HJM model.

The key observation in [IK98] was that when volatility processes[2] $\sigma_i(t,T,\omega)$, $1 \leq i \leq n$, satisfy the condition

$$\frac{\partial_T \sigma_i(t,T,\omega)}{\partial T} = \kappa_i(T)\sigma_i(t,T,\omega), \tag{1}$$

and $\kappa_i(T)$ are path independent functions of $T$, then the $n$-factor HJM model can be transformed to a $2n$-dimensional Markovian system, and the bond price can be obtained in terms of the $2n$ state variables. The condition (1) arises naturally from the desire to replace path dependent terms in the differential of the spot rate by an expression involving the spot rate itself, and is in fact sufficient to reduce the HJM model to a finite dimensional Markovian system, as described in [RS95] and [IK98].

Let $\mathcal{L}_i = \partial/\partial T - \kappa_i(T)$. Then (1) can be rewritten $\mathcal{L}_i\sigma_i(t,T,\omega) = 0$. That is, the [IK98] condition requires that, for each $t$, the volatility processes $\sigma_i(t,T,\omega)$ satisfy a first order, linear, homogeneous, ordinary differential equation in $T$. The essentially arbitrary initial condition for the differential equation then allows the spot volatility $\sigma_i(t,t,\omega)$ to be unrestricted. However, in order for the corresponding model to transform to a Markovian system with respect to state variables introduced in [RS95] and [IK98], the initial condition must be of the form

$$\sigma_i(t,t,\omega) = \sigma_i(t,t,r(t,\omega)). \tag{2}$$

That is, the spot volatility must be a function of the time variable $t$, and the spot rate $r(t,\omega)$. As mentioned earlier, although (2) is required to transform to a Markovian system,[3] the bond price formula obtained in [IK98] applies to a larger class of volatility processes.

In this paper, the approach of [IK98] is generalised by requiring that each $\sigma_i(t,T,\omega)$ is a function of $t$, $T$, and $m$ forward rates $f(t,t+\varsigma_1,\omega),\dots,f(t,t+\varsigma_m,\omega)$, so that

$$\sigma_i(t,T,\omega) = \sigma_i(t,T,f(t,t+\varsigma_1,\omega),\dots,f(t,t+\varsigma_m,\omega)), \tag{3}$$

and satisfies a differential equation of the form $\mathcal{L}_i\sigma_i(t,T,\omega) = 0$, where

$$\mathcal{L}_i = \frac{\partial^{m_i}}{\partial T^{m_i}} - \sum_{j=0}^{m_i-1} \kappa_{i,j}(T)\,\frac{\partial^j}{\partial T^j} \tag{4}$$

is an $m_i$-th order linear differential operator and the coefficients $\kappa_{i,j}(T)$ are path independent functions of $T$. The corresponding $n$-factor HJM model can then be transformed to a *finite* dimensional Markovian system of dimension at most $m\sum_{i=1}^{n} m_i^2(m_i+3)/2$. Further, for each $i$, the $m_i$ arbitrary boundary conditions for the differential equation, $\mathcal{L}_i\sigma_i(t,T,\omega) = 0$, allow $\sigma_i(t,t+T,\omega)$ to be arbitrary functions of the forward rates $f(t,t+\varsigma_1,\omega),\dots,f(t,t+\varsigma_m,\omega)$.

As in [RS95] and [IK98], although the bond price formulae given in §4 remains valid for a more general class of volatility processes, the restriction (3) is required to obtain a Markovian system.

---

[2] The $\omega$ in $\sigma_i(t,T,\omega)$ represents the dependence of the volatility process on the path followed by the underlying Wiener process. See §1.

[3] With respect to the state variables introduced in [IK98].

The outline of the paper is as follows. It begins with a brief review of the HJM term structure model and the parametrisation, $T = t + \varsigma$, due to Brace and Musiela [BM94] in §1. The main results of this paper are then presented in §2, in which the transformation of the multifactor HJM model to finite dimensional Markovian systems is outlined. Natural extensions of the [IK98] model are considered in §4, and explicit expressions for the bond price as a function of the state variables are obtained for these models. Finally, the paper concludes in §5.

## 1. RISK NEUTRAL HEATH-JARROW-MORTON MODEL AND THE BRACE-MUSIELA PARAMETRISATION

This section reviews in brief the HJM model and the parametrisation, $T = t + \varsigma$, introduced by Brace and Musiela [BM94]. For details, refer to [HJM92], [BM94], [MR97], or [Bjö96].

### 1.1. **Risk Neutral Heath-Jarrow-Morton Model.**

Fix a trading interval $[0, \tau]$, $\tau > 0$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where $\Omega$ is the set of states of the economy, $\mathcal{F}$ is the $\sigma$-algebra of measurable events, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

It is assumed that there are $n$ independent standard $\mathbb{P}$-Brownian motions $W_t^i$, $1 \le i \le n$, that generate a complete right continuous filtration $\{\mathcal{F}_t\}_{0 \le t \le \tau}$ on $(\Omega, \mathcal{F})$.

For each *maturity* $T \in [0, \tau]$, the time $t$ instantaneous forward rate $f(t, T, \omega)$ in the *risk-neutral* $n$-factor HJM model is a stochastic process determined by suitably well-behaved[4] volatility processes $\sigma_i(t, T, \omega)$, $1 \le i \le n$, and evolves according to the stochastic integral equation

$$f(t,T,\omega) = f(0,T,\omega) + \sum_{i=1}^{n} \int_0^t \sigma_i^*(s,T,\omega)\,ds + \sum_{i=1}^{n} \int_0^t \sigma_i(s,T,\omega)\,d\widetilde{W}_s^i, \tag{1.1}$$

where $0 \le t \le T$, $\sigma_i^*(s,T,\omega) = \sigma_i(s,T,\omega) \int_s^T \sigma_i(s,u,\omega)\,du$, and $\widetilde{W}_t^i$ are independent standard $\widetilde{\mathbb{P}}$-Brownian motions, where $\widetilde{\mathbb{P}}$ is a $\mathbb{P}$-equivalent martingale measure in the sense of [HK79] and [HP81].

Evolution of the spot rate process $r(t, \omega)$, where $r(t, \omega) = f(t, t, \omega)$, is determined by setting $T = t$ in (1.1), which yields

$$r(t,\omega) = f(0,t,\omega) + \sum_{i=1}^{n} \int_0^t \sigma_i^*(s,t,\omega)\,ds + \sum_{i=1}^{n} \int_0^t \sigma_i(s,t,\omega)\,d\widetilde{W}_s^i. \tag{1.2}$$

The price of a $T$-maturity pure discount bond, $P(t, T, \omega)$, is given by

$$P(t,T,\omega) = e^{-\int_t^T f(t,u,\omega)\,du} = \widetilde{\mathbb{E}}\left[ e^{-\int_t^T r(u)\,du} \middle| \mathcal{F}_t \right](\omega), \tag{1.3}$$

where $\widetilde{\mathbb{E}}$ is the expectation with respect to $\widetilde{\mathbb{P}}$.

---

[4] The main technical conditions are that for each $1 \le i \le n$,

(a) $\sigma_i(t,u,\omega)\colon [0,T]^2 \times \Omega \to \mathbb{R}$ is $[\mathcal{B}([0,T]^2) \otimes \mathcal{F}]/\mathcal{B}(R)$-measurable, and
(b) $\sigma_i(t,u,\omega)$ is $\{\mathcal{F}_t\}$-adapted for all $u$, and $\int_0^T \sigma_i(t,T,\omega)\,dt < \infty$ a.s. $\mathbb{P}$,

where $\mathcal{B}(X)$ denotes the Borel $\sigma$-algebra on $X$.

The corresponding stochastic differential equations are given by

$$df(t,T,\omega) = \sum_{i=1}^{n} \sigma_i^*(t,T,\omega)\,dt + \sum_{i=1}^{n} \sigma_i(t,T,\omega)\,d\widetilde{W}_t^i, \tag{1.4}$$

$$dr(t,\omega) = \left.\frac{\partial f(t,T,\omega)}{\partial T}\right|_{T=t} dt + \sum_{i=1}^{n} \sigma_i(t,t,\omega)\,d\widetilde{W}_t^i, \tag{1.5}$$

$$dP(t,T,\omega) = P(t,T,\omega)\,r(t,\omega)dt + P(t,T,\omega)\sum_{i=1}^{n}\left[\int_t^T \sigma_i(t,u,\omega)\,du\right] d\widetilde{W}_t^i, \tag{1.6}$$

where $0 \le t \le T \le \tau$, and

$$\begin{aligned}\left.\frac{\partial f(t,T,\omega)}{\partial T}\right|_{T=t} = \frac{\partial f(0,t,\omega)}{\partial t} &+ \sum_{i=1}^{n}\int_0^t \frac{\partial \sigma_i^*(s,t,\omega)}{\partial t}\,ds \\ &+ \sum_{i=1}^{n}\int_0^t \frac{\partial \sigma_i(s,t,\omega)}{\partial t}\,d\widetilde{W}_s^i.\end{aligned} \tag{1.7}$$

It can be seen from (1.4) that the forward rate process $f(t,T,\omega)$ is non-Markovian in general, since the volatility processes $\sigma_i(t,T,\omega)$ depend on the path $\omega$, and hence on the past. Even if $\sigma_i(t,T,\omega)$ did not depend on the past, (1.5) and (1.7) show that the spot rate process remains non-Markovian in general, due to the path dependent terms in (1.7) that involve integration over the past. Consequently, the general HJM model does not readily lend itself to practical implementations. If the HJM model can be transformed to a Markovian system, then the resulting system can be tackled more efficiently to obtain the bond price $P(t,T,\omega)$, either via the Monte Carlo simulation techniques, or by solving directly, or numerically, the resulting partial differential equation. The latter method, in a special case, is considered in [CK98b].

In the remainder of this paper, the risk-neutral $n$-factor HJM term structure model is used.

1.2. **Brace-Musiela Parametrisation.** The volatility processes we wish to consider have the form

$$\sigma_i(t,T,\omega) = \sigma_i(t,T,f(t,t+\varsigma_1,\omega),\ldots,f(t,t+\varsigma_m,\omega)).$$

That is, the dependence of $\sigma_i(t,T,\omega)$ on the path $\omega$ is absorbed into the dependence on $m$ forward rates $f(t,t+\varsigma_j,\omega)$, where $0 \le \varsigma_1 < \cdots < \varsigma_m$ are fixed *tenors*. In order to obtain a Markovian system, we are forced to introduce the forward rate processes $f(t,t+\varsigma_j,\omega)$ as state variables, and determine the stochastic differential equations governing their evolution over time. This in turn forces us to consider, for each $\varsigma \in [0,\infty)$, the process $f(t,t+\varsigma,\omega)$. The parametrisation, $T = t+\varsigma$, of the maturity variable was introduced by Brace and Musiela in [BM94]. They make the comment that the HJM parametrisation is suited to bonds while theirs is suited to swaps. In this vein, our method may be considered as one in which bonds are priced using a finite number of swap rates.

The parametrisation $T = t+\varsigma$ does not introduce anything new to the general HJM framework outlined in §1.1. However, it does lead to certain desirable properties. In particular, the parametrisation provides a symmetric treatment[5] of the forward rate process $f(t,t+\varsigma,\omega)$ and the spot rate process $r(t,\omega)$, and, under the parametrisation, the forward rate process $f(t,t+\varsigma,\omega)$ is valid for *all* $t$ rather than only for $t \le T$, which is the case in the HJM parametrisation. It is the

---

[5] As seen from (1.4) and (1.5), the differential of the spot rate process $r(t,\omega)$ and the differential of the forward rate process $f(t,T,\omega)$ are not treated 'symmetrically', in the sense that $dr(t,\omega)$ cannot be obtained from $df(t,T,\omega)$ by simply setting $T = t$, even though $r(t,\omega) = f(t,t,\omega)$.

latter property that is of greater importance for our purposes, since if $\sigma_i(t,T,\omega)$ were functions of $f(t,\varsigma_j,\omega)$ rather than being functions of $f(t,t+\varsigma_j,\omega)$, then the processes $f(t,T,\omega)$, $r(t,\omega)$, and $P(t,T,\omega)$ would be valid only for $t \leq \min_j(\varsigma_j)$.

The next step is to determine the stochastic differential equations governing the evolution of the forward rate process $f(t,t+\varsigma,\omega)$. For notational convenience, $\omega$ is omitted from $f(t,t+\varsigma,\omega)$, $\sigma_i(t,T,\omega)$, etc.

Fix $\varsigma \geq 0$, and consider the forward rate process $f(t,t+\varsigma)$. Then setting $T = t+\varsigma$ in (1.1) and (1.2), $f(t,t+\varsigma)$ and $r(t)$ are governed by equations

$$f(t,t+\varsigma) = f(0,t+\varsigma) + \sum_{i=1}^{n}\int_0^t \sigma_i^*(s,t+\varsigma)\,ds + \sum_{i=1}^{n}\int_0^t \sigma_i(s,t+\varsigma)\,d\widetilde{W}_s^i, \tag{1.8}$$

$$r(t) = f(0,t) + \sum_{i=1}^{n}\int_0^t \sigma_i^*(s,t)\,ds + \sum_{i=1}^{n}\int_0^t \sigma_i(s,t)\,d\widetilde{W}_s^i. \tag{1.9}$$

The stochastic differential equations for $f(t,t+\varsigma)$ and $r(t)$ are then given by

$$df(t,t+\varsigma) = \left[\frac{\partial f(t,t+\varsigma)}{\partial\varsigma} + \sum_{i=1}^{n}\sigma_i^*(t,t+\varsigma)\right]dt + \sum_{i=1}^{t}\sigma_i(t,t+\varsigma)\,d\widetilde{W}_t^i, \tag{1.10}$$

$$dr(t) = \left.\frac{\partial f(t,t+\varsigma)}{\partial\varsigma}\right|_{\varsigma=0} dt + \sum_{i=1}^{n}\sigma_i(t,t)\,d\widetilde{W}_t^i, \tag{1.11}$$

where $\sigma_i^*(t,t+\varsigma) = \sigma_i(t,t+\varsigma)\int_t^{t+\varsigma}\sigma(t,u)\,du$, and

$$\frac{\partial f(t,t+\varsigma)}{\partial\varsigma} = \left[\frac{\partial f(0,t+\varsigma)}{\partial\varsigma} + \sum_{i=1}^{n}\int_0^t \sigma_i^2(s,t+\varsigma)\,ds + f_i^\sharp(t,\varsigma)\right]dt, \tag{1.12}$$

where

$$f_i^\sharp(t,\varsigma) = \int_0^t \frac{\partial\sigma_i(s,t+\varsigma)}{\partial\varsigma}\int_s^{t+\varsigma}\sigma_i(s,u)\,du\,ds + \int_0^t \frac{\partial\sigma_i(s,t+\varsigma)}{\partial\varsigma}\,d\widetilde{W}_s^i. \tag{1.13}$$

As seen from (1.8)-(1.11), $r(t)$ is obtained from $f(t,t+\varsigma)$, and $dr(t)$ from $df(t,t+\varsigma)$, by setting $\varsigma = 0$. In particular, it is not necessary to compute both $df(t,t+\varsigma)$ and $dr(t)$. This was *not* the case in the standard HJM parametrisation, where $dr(t)$ required a *separate*, and more involved, computation than $df(t,T)$.

## 2. TRANSFORMATION TO A MARKOVIAN SYSTEM

This section outlines a method for transforming the $n$-factor HJM model described in §1 to a finite dimensional Markovian system. The variable $\omega$ continues to be omitted from $f(t,t+\varsigma,\omega)$, $\sigma_i(t,T,\omega)$, etc.

Let $m,n \in \mathbb{N}$, and assume given $n$ non-negative integers $m_1, m_2, \ldots, m_n$ and $m$ real numbers $0 \leq \varsigma_1 < \varsigma_2 < \cdots < \varsigma_m$. In addition to the standard HJM assumptions, it is furthermore assumed that:

[A1] For each $1 \leq i \leq n$, $\sigma_i(t,T)$ is a function of $t$, $T$, and the $m$ forward rates $f(t,t+\varsigma_1),\ldots,f(t,t+\varsigma_m)$, so that

$$\sigma_i(t,T,\omega) = \sigma_i(t,T,f(t,t+\varsigma_1,\omega),\ldots,f(t,t+\varsigma_m,\omega)), \tag{2.1}$$

[A2] For each $1 \leq i \leq n$, $\sigma_i(t,T)$ is $m_i$ times differentiable with respect to $T$ and satisfies the $m_i$-th order homogeneous linear differential equation

$$\mathcal{L}_i\sigma_i(t,T) = 0, \text{ where } \mathcal{L}_i = \frac{\partial^{m_i}}{\partial T^{m_i}} - \sum_{j=0}^{m_i-1}\kappa_{i,j}(T)\,\frac{\partial^j}{\partial T^j}, \tag{2.2}$$

and $\kappa_{i,j}(T)$ are *path independent* functions of $T$.

Under the assumptions [A1]–[A2], the volatility processes $\sigma_i(t,T)$ are arbitrary functions of $f(t,t+\varsigma_1),\ldots,f(t,t+\varsigma_m)$, and result in forward rate dependent HJM models. Note that the [RS95] and [IK98] models are obtained by taking $m_i=1$ for all $i$, and the [BC97] model is obtained by taking[6] $\mathcal{L}=(\partial/\partial t-\lambda)^m$. Clearly, these models do not allow the degree of flexibility permitted by [A1]–[A2] on $\sigma_i(t,T)$. The [RS95] and [IK98] models are restricted because $m_i=1$ for all $i$ in their model[7], while the [BC97] model is restricted because they begin by assuming that $\sigma(t,T)=G[r(t)]\,p_m(T-t)e^{-\lambda(T-t)}$, where $p_m(x)$ is a polynomial of degree $m$.

**Lemma 2.1.** *For $1\le i\le n$, $1\le j\le m$, and $p$, $q$, $r\in\mathbb{N}$, define the state variables*

$$\phi^{p,q}_{i,j}(t)=\int_0^t \frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\cdot\frac{\partial^{p+q}\sigma_i(s,t+\varsigma_j)}{\partial t^{p+q}}\,ds, \tag{2.3}$$

$$\begin{aligned}\psi^r_{i,j}(t)=&\int_0^t \frac{\partial^r\sigma_i(s,t+\varsigma_j)}{\partial t^r}\int_s^{t+\varsigma_j}\sigma_i(s,u)\,du\,ds\\&+\int_0^t\frac{\partial^r\sigma_i(s,t+\varsigma_j)}{\partial t^r}\,d\widetilde{W}^i_s,\end{aligned} \tag{2.4}$$

*and let*

$$\delta\phi^{p,q}_{i,j}(t)=\left[\frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\cdot\frac{\partial^{p+q}\sigma_i(s,t+\varsigma_j)}{\partial t^{p+q}}\right]_{s=t}.$$

*Then the following stochastic differential equations are satisfied:*

$$\begin{aligned}df(t,t+\varsigma_j)=&\left[\frac{\partial f(0,t+\varsigma_j)}{\partial t}+\sum_{i=1}^n\left(\phi^{0,0}_{i,j}(t)+\psi^1_{i,j}(t)\right)\right]dt\\&+\sum_{i=1}^n\sigma_i(t,t+\varsigma_j)\,d\widetilde{W}^i_t,\end{aligned} \tag{2.5}$$

$$d\phi^{p,q}_{i,j}(t)=\begin{cases}\left[\delta\phi^{p,0}_{i,j}(t)+2\phi^{p,1}_{i,j}(t)\right]dt, & \text{if } q=0\\ \left[\delta\phi^{p,q}_{i,j}(t)+\phi^{p,q+1}_{i,j}(t)+\phi^{p+1,q-1}_{i,j}(t)\right]dt, & \text{if } q>0,\end{cases} \tag{2.6}$$

$$d\psi^r_{i,j}(t)=\left[\phi^{0,r}_{i,j}(t)+\psi^{r+1}_{i,j}(t)\right]dt+\left[\frac{\partial^r\sigma_i(s,t+\varsigma_j)}{\partial t^r}\right]_{s=t}d\widetilde{W}^i_t. \tag{2.7}$$

*Proof.* The proofs are straight forward and the details are only provided for the first identity in (2.6).

$$\begin{aligned}d\phi^{p,0}_{i,j}(t)&=d\int_0^t\left(\frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\right)^2ds\\&=\left[\left.\frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\right|^2_{s=t}+\int_0^t\frac{\partial}{\partial t}\left(\frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\right)^2ds\right]dt\\&=\left[\delta\phi^{p,0}_{i,j}(t)+\int_0^t 2\,\frac{\partial^p\sigma_i(s,t+\varsigma_j)}{\partial t^p}\cdot\frac{\partial^{p+1}\sigma_i(s,t+\varsigma_j)}{\partial t^{p+1}}\,ds\right]dt\\&=\left[\delta\phi^{p,0}_{i,j}(t)+2\,\phi^{p,1}_{i,j}(t)\right]dt\,.\end{aligned}$$

The remaining identities are proved similarly. ☐

If $\sigma_i(t,T)$ are infinitely differentiable with respect to $T$, and satisfy [A1], then Lemma 2.1 implies that the resulting model is in general an *infinite* dimensional

---

[6] Recall that the model studied in [BC97] is a one factor HJM model.

[7] If $m_i=1$, then $(\partial/\partial T-\kappa_i(T))\sigma_i(t,T)=0$ has the solution $\sigma_i(t,T)=\sigma_i(t,t)e^{\int_t^T\kappa_i(u)\,du}$, and as a result $\sigma_i(t,T)$ can only depend on the spot volatility $\sigma_i(t,t)$.

Markovian system with respect to the state variables $f(t, t+\varsigma_j)$, $\phi^{p,q}_{i,j}(t)$, and $\psi^r_{i,j}(t)$. The purpose of (2.2) in assumption [A2] is to restrict the system to be *finite* dimensional.

Before proceeding with the proof that the HJM model under [A1]–[A2] can be transformed to a finite dimensional Markovian system, the central idea is illustrated with a diagram. A directed edge, $v(t) \to w(t)$, in Figure 2.1 indicates that $w(t)$ occurs in the expression for $dv(t)$. The edges are obtained from Lemma 2.1.

$$\begin{tikzcd}
 & & & & & \phi^{0,0}_{i,j}(t) \arrow[d] & \begin{matrix} f(t,t+\varsigma_j) \\ \psi^{0}_{i,j}(t) \end{matrix} \arrow[l] \arrow[d] \\
 & & & & \phi^{1,0}_{i,j}(t) \arrow[d] & \phi^{0,1}_{i,j}(t) \arrow[l] \arrow[d] & \psi^{1}_{i,j}(t) \arrow[l] \arrow[d] \\
 & & & \phi^{2,0}_{i,j}(t) \arrow[d] & \phi^{1,1}_{i,j}(t) \arrow[l] \arrow[d] & \phi^{0,2}_{i,j}(t) \arrow[l] \arrow[d] & \psi^{2}_{i,j}(t) \arrow[l] \arrow[d] \\
 & & & \cdots \arrow[d] & \cdots \arrow[d] & \cdots \arrow[d] & \cdots \arrow[d] \\
 & \phi^{m_i-1,0}_{i,j}(t) \arrow[d] & \cdots \arrow[l] & \phi^{2,m_i-3}_{i,j}(t) \arrow[l] \arrow[d] & \phi^{1,m_i-2}_{i,j}(t) \arrow[l] \arrow[d] & \phi^{0,m_i-1}_{i,j}(t) \arrow[l] \arrow[d] & \psi^{m_i-1}_{i,j}(t) \arrow[l] \arrow[d] \\
\phi^{m_i,0}_{i,j}(t) & \phi^{m_i-1,1}_{i,j}(t) \arrow[l] & \cdots \arrow[l] & \phi^{2,m_i-2}_{i,j}(t) \arrow[l] & \phi^{1,m_i-1}_{i,j}(t) \arrow[l] & \phi^{0,m_i}_{i,j}(t) \arrow[l] & \psi^{m_i}_{i,j}(t) \arrow[l]
\end{tikzcd}$$

FIGURE 2.1. Interdependence of State Variables.

If [A2] is assumed, then the variables in the lowest level can be expressed as linear combinations of the variables in higher levels, and the diagram can be terminated at depth $m_i - 1$. This observation essentially establishes that the HJM model can be transformed to a finite dimensional Markovian system, and is formalised in the following.

**Proposition 2.2.** *Let $\sigma_i(t, T)$ satisfy the assumptions [A1] and [A2].*

*(i) The variables $\psi^{m_i}_{i,j}(t)$ and $\phi^{\gamma, m_i-\gamma}_{i,j}(t)$, for $0 \le \gamma \le m_i$, can be written as linear combinations of $\psi^{\lambda}_{i,j}(t)$ and $\phi^{\mu,\nu}_{i,j}(t)$, with $\lambda < m_i$ and $\mu + \nu < m_i$, in the form*

$$\begin{aligned}
\psi^{m_i}_{i,j}(t) &= \sum_{\lambda=0}^{m_i-1} \kappa_{i,\lambda}(t+\varsigma_j)\, \psi^{\lambda}_{i,j}(t), \\
\phi^{m_i,0}_{i,j}(t) &= \sum_{0 \le \mu < \nu \le m_i-1} \kappa^{\mu,\nu}_{i}(t+\varsigma_j)\, \phi^{\mu,\nu-\mu}_{i,j}(t) \\
&\qquad + \sum_{0 \le \nu \le \mu \le m_i-1} \kappa^{\mu,\nu}_{i}(t+\varsigma_j)\, \phi^{\nu,\mu-\nu}_{i,j}(t), \\
\phi^{\gamma,m_i-\gamma}_{i,j}(t) &= \sum_{\mu=0}^{\gamma} \kappa_{i,\mu}(t+\varsigma_j)\, \phi^{\mu,\gamma-\mu}_{i,j}(t) + \sum_{\mu=\gamma+1}^{m_i-1} \kappa_{i,\mu}(t+\varsigma_j)\, \phi^{\gamma,\mu-\gamma}_{i,j}(t),
\end{aligned}$$

*where $\kappa^{\mu,\nu}_{i}(t+\varsigma_j) = \kappa_{i,\mu}(t+\varsigma_j)\, \kappa_{i,\nu}(t+\varsigma_j)$.*

*(ii) For each $1 \le j \le m$, the variable $\psi^0_{n,j}(t)$ can be written as a linear combination of $f(t, t+\varsigma_j)$, $f(0, t+\varsigma_j)$, and $\psi^0_{\lambda,j}(t)$, with $1 \le \lambda \le m_i$, in the form*

$$\psi^0_{n,j}(t) = f(t, t+\varsigma_j) - f(0, t+\varsigma_j) - \sum_{\lambda=1}^{n-1} \psi^0_{\lambda,j}(t). \tag{2.8}$$

*Proof.* (i) The result for $\psi_{i,j}^{m_i}(t)$ is first established.

$$\begin{aligned}
\psi_{i,j}^{m_i}(t) &= \int_0^t \frac{\partial^{m_i}\sigma_i(s,t+\varsigma_j)}{\partial t^{m_i}} \int_s^{t+\varsigma_j} \sigma_i(s,u)\,du\,ds \\
&\quad + \int_0^t \frac{\partial^{m_i}\sigma_i(s,t+\varsigma_j)}{\partial t^{m_i}}\, d\widetilde{W}_s^i \\
&= \int_0^t \left[\sum_{\lambda=0}^{m_i-1} \kappa_{i,\lambda}(t+\varsigma_j)\, \frac{\partial^{\lambda}\sigma_i(s,t+\varsigma_j)}{\partial t^{\lambda}}\right] \int_s^{t+\varsigma_j} \sigma_i(s,u)\,du\,ds \\
&\quad + \int_0^t \sum_{\lambda=0}^{m_i-1} \kappa_{i,\lambda}(t+\varsigma_j)\, \frac{\partial^{\lambda}\sigma_i(s,t+\varsigma_j)}{\partial t^{\lambda}}\, d\widetilde{W}_s^i, \quad \text{by [A2]} \\
&= \sum_{\lambda=0}^{m_i-1} \kappa_{i,\lambda}(t+\varsigma_j) \left[\int_0^t \frac{\partial^{\lambda}\sigma_i(s,t+\varsigma_j)}{\partial t^{\lambda}} \int_s^{t+\varsigma_j} \sigma_i(s,u)\,du\,ds \right. \\
&\quad \left. + \int_0^t \frac{\partial^{\lambda}\sigma_i(s,t+\varsigma_j)}{\partial t^{\lambda}}\, d\widetilde{W}_s^i\right] \\
&= \sum_{\lambda=0}^{m_i-1} \kappa_{i,\lambda}(t+\varsigma_j)\, \psi_{i,j}^{\lambda}(t).
\end{aligned}$$

Similar arguments establish the results for $\phi_{i,j}^{\gamma,m_i-\gamma}(t)$.

(ii) For $\psi_{n,j}^0(t)$, recall from (1.8) that $f(t,t+\varsigma_j) = f(0,t+\varsigma_j) + \sum_{\lambda=1}^n \psi_{\lambda,j}^0(t)$, whence

$$\psi_{n,j}^0(t) = f(t,t+\varsigma_j) - f(0,t+\varsigma_j) - \sum_{\lambda=1}^{n-1} \psi_{\lambda,j}^0(t).$$

This establishes the required identities. ☐

The main result is now stated.

**Theorem 2.3.** *Let $\sigma_i(t,T)$ satisfy the assumptions [A1] and [A2]. Then the HJM model transforms to a Markovian system with state variables $f(t,t+\varsigma_j)$, $\phi_{i,j}^{\lambda_i,\mu_i}(t)$, and $\psi_{i,j}^{\nu_i}(t)$, where the indices $i$, $j$, $\lambda_i$, $\mu_i$, and $\nu_i$ satisfy the restrictions*

- *(i) $1 \le i \le n$, $1 \le j \le m$, $0 \le \lambda_i$, $0 \le \mu_i$, and $0 \le \nu_i < m_i$,*
- *(ii) for each $i$, $\lambda_i + \mu_i < m_i$, and $(i,\nu_i) \ne (n,0)$.*

*In particular, the system has dimension at most $m\sum_{\gamma=1}^n \frac{1}{2}m_\gamma^2(m_\gamma+3)$.*

*Proof.* The finite dimensionality and the form of the state variables follow immediately from Proposition 2.2. For the dimension

$$\#\{f(t,t+\varsigma_j)\} = m, \quad \#\{\psi_{i,j}^{\nu}(t)\} \le m\Big[\Big(\textstyle\sum_{\mu=1}^n m_\mu\Big) - 1\Big],$$
$$\#\{\phi_{i,j}^{\lambda,\mu}(t)\} \le m\Big[\textstyle\sum_{\gamma=1}^n \frac{1}{2}m_\gamma(m_\gamma+1)\Big].$$

Hence the dimension is at most $m\sum_{\gamma=1}^n \frac{1}{2}m_\gamma^2(m_\gamma+3)$. ☐

The dependence of $\sigma_i(t,T)$ on the term structure can be increased by requiring that they satisfy a higher order differential equation, since this allows $\sigma_i(t,T)$ to be dependent on forward rates for a larger number of tenors. For smooth volatility functions, the general case in which volatilities depend on the entire term structure can be regarded as a limit of the model introduced in this paper in some sense.

The following result is useful in generating finite dimensional Markovian systems from other finite dimensional Markovian systems.

**Proposition 2.4.** *For each $1 \le i \le n$, fix a positive integer $n_i$, and for $1 \le j \le n_i$ define differential operators*

$$\mathcal{L}_i^{(j)} = \frac{\partial^{m_i^{(j)}}}{\partial T^{m_i^{(j)}}} - \sum_{k=0}^{m_i^{(j)}} \kappa_{i,k}^{(j)}(T)\,\frac{\partial^k}{\partial T^k} \quad \text{and} \quad \mathcal{L}_i = \prod_{j=1}^{n_i} \mathcal{L}_i^{(j)}.$$

*If $\sigma_i^{(j)}(t,T)$ satisfies $\mathcal{L}_i^{(j)}\sigma_i^{(j)}(t,T) = 0$, then $\sigma_i(t,T) = \sum_{j=1}^{n_i} x_i^{(j)}(t)\,\sigma_i^{(j)}(t,T)$ satisfies $\mathcal{L}_i\sigma_i(t,T) = 0$, where $x_i^{(j)}(t)$ are arbitrary functions of $t$.*

Loosely speaking, Proposition 2.4 implies that finite '$t$-linear' combinations of $\sigma_i(t,T)$ satisfying [A1]–[A2] again satisfy [A1]–[A2].

## 3. Examples

This short section lists some volatility functions to which Theorem 2.3 applies.

(i) *Polynomial Functions.* If $\mathcal{L}_i = \dfrac{\partial^{k+1}}{\partial T^{k+1}}$, then $\sigma_i(t,T) = \sum_{j=0}^{k} c_{i,j}(t)\,T^k$.

(ii) *Exponential Functions.* If $\mathcal{L}_i = \dfrac{\partial}{\partial T} - \lambda_i(T)$, then

$$\sigma_i(t,T) = c_i(t)\,e^{\int_0^T \lambda_i(x)\,dx}.$$

In particular, exponential, trigonometric, and hyperbolic functions are special cases.

(iii) *Bessel Functions.* If $\mathcal{L}_i = \dfrac{\partial^2}{\partial T^2} + \dfrac{1}{T}\dfrac{\partial}{\partial T} + \left[1 - \dfrac{m^2}{T^2}\right]$, then

$$\sigma_i(t,T) = c_i(t)\,J_m(T),$$

where $J_m(T)$ is the Bessel function of order $m$.

(iv) *Hypergeometric Functions.* If

$$\mathcal{L}_i = \frac{\partial^2}{\partial T^2} + \left[\frac{c - (a+b+1)\,T}{T\,(T-1)}\right]\frac{\partial}{\partial T} - \frac{ab}{T\,(T-1)},$$

then $\sigma_i(t,T) = c_i(t)\,F(a,b,c;T)$, where $F(a,b,c;T)$ is the hypergeometric function.

Additional $\sigma_i(t,T)$ satisfying the hypotheses of Theorem 2.3 may be obtained by using Proposition 2.4.

## 4. Generalisation of the Inui-Kijima Model

This section considers a subclass of Markovian systems introduced in §2 that are natural extensions of the [IK98] model. These are the higher order analogues of [IK98] in which the $\sigma_i(t,T)$ satisfy the differential equation $\mathcal{L}_i\sigma_i(t,T) = 0$, where $\mathcal{L}_i = \prod_{j=1}^{m_i}[\partial/\partial T - \lambda_{i,j}(T)]$. For certain special cases, explicit expressions for the bond price are obtained in terms of the state variables, and in particular, the [IK98] formula is obtained in Theorem 4.5.

As in the previous section, let $m_i$ be positive integers for $1 \le i \le n$, and let $\varsigma_j$ be given for $1 \le j \le m$. Throughout this section, assume [A1]–[A2] with

$$\mathcal{L}_i = \prod_{j=1}^{m_i}\left[\frac{\partial}{\partial T} - \lambda_{i,j}(T)\right], \tag{4.1}$$

and note that the [IK98] model is a special case with $m_i = 1$.

**Lemma 4.1.** *Let $\mathcal{L}_i$ be defined as in (4.1), and let $c_{i,j} : \mathbb{R} \times \Omega \to \mathbb{R}$ for $1 \le j \le m_i$. If $\sigma_i(t,T)$ are defined by*

$$\sigma_i(t,T) = \sum_{j=1}^{m_i} c_{i,j}(t)\, e^{\int_0^T \lambda_{i,j}(x)\,dx}, \tag{4.2}$$

*then $\mathcal{L}_i \sigma_i(t,T) = 0$.*

*Proof.* Let $\mathcal{L}_i^j = \prod_{k \ne j} [\partial/\partial T - \lambda_{i,k}(T)]$, where $1 \le k \le m_i$. Then

$$\mathcal{L}_i \sigma_i(t,T) = \sum_{j=1}^{m_i} c_{i,j}(t)\, \mathcal{L}_i^j\, [\partial_T - \lambda_{i,j}(T)]\, e^{\int_0^T \lambda_{i,j}(x)\,dx} = 0,$$

since $[\partial/\partial T - \lambda_{i,j}(T)]\, e^{\int_0^T \lambda_{i,j}(x)\,dx} = 0$. □

**Proposition 4.2.** *Let $\sigma_{i,j}(t,T) = \sum_{j=1}^{m_i} c_{i,j}(t)\, e^{\int_0^T \lambda_{i,j}(x)\,dx}$ as in Lemma 4.1. Then the corresponding HJM model transforms to a finite dimensional Markovian system with state variables as listed in Theorem 2.3.*

*Proof.* This follows from Theorem 2.3 and Lemma 4.1. □

Hence the HJM models considered in this section are Markovian. Considered now are two special cases of the present model for which a bond price formula is available either in terms of the state variables in Theorem 2.3, or a slightly larger set.

### 4.1. Inui-Kijima-Ritchken-Sankarasubramanian Bond Price Formula.

In this subsection, assume $m = 1$, $\varsigma_1 = 0$, and $m_i = 1$, for all $1 \le i \le n$, so that

$$\mathcal{L}_i = \frac{\partial}{\partial T} - \lambda_i(T) \quad \text{and} \quad \sigma_i(t,T) = c_i(t, r(t))\, e^{\int_0^T \lambda_i(x)\,dx}. \tag{4.3}$$

This is the general [IK98] model, and when $n = 1$ this is the [RS95] model. For this class of models, a formula for $P(t,T)$ can be obtained in terms of the state variables, and the following identity plays a crucial role.

**Lemma 4.3.** *The $\sigma_i(t,T)$ given by (4.3) satisfy the identities*

$$\sigma_i(s, t+u) = \sigma_i(s,t)\, e^{\int_t^{t+u} \lambda_i(x)\,dx}, \tag{4.4}$$

$$\sigma_i(s,t) = \sigma_i(s,s)\, e^{\int_s^t \lambda_i(x)\,dx}. \tag{4.5}$$

*Proof.* Since the second identity is a consequence of the first, it is sufficient to prove only the latter. Note, firstly, that

$$\begin{aligned}\sigma_i(s,t) &= c_i(s, r(s))\, e^{\int_0^t \lambda_i(x)\,dx} = \left[c_i(s, r(s))\, e^{\int_0^s \lambda_i(x)\,dx}\right] e^{\int_s^t \lambda_i(x)\,dx} \\ &= \tilde{c}_i(s, r(s))\, e^{\int_s^t \lambda_i(x)\,dx},\end{aligned}$$

where $\tilde{c}_i(s, r(s)) = c_i(s, r(s))\, e^{\int_0^s \lambda_i(x)\,dx}$. Hence

$$\begin{aligned}\sigma_i(s, t+u) &= \tilde{c}_i(s, r(s))\, e^{\int_s^{t+u} \lambda_i(x)\,dx} = \left[\tilde{c}_i(s, r(s))\, e^{\int_s^t \lambda_i(x)\,dx}\right] e^{\int_t^{t+u} \lambda_i(x)\,dx} \\ &= \sigma_i(s,t)\, e^{\int_t^{t+u} \lambda_i(x)\,dx}.\end{aligned}$$

The second identity follows from setting $t = s$ and $u = 0$ in the above identity. □

The following lemma is contained in [IK98, p37].

**Lemma 4.4.** *Let $\beta_i(t,T) = \int_t^T e^{-\int_t^u \lambda_i(x)\,dx} du$. Then the following identities hold:*

$$\int_t^T e^{-\int_t^u \lambda_i(x)\,dx} \int_t^u e^{-\int_t^v \lambda_i(x)\,dx} dv\,du = \frac{1}{2}\beta_i^2(t,T), \tag{4.6}$$

$$\int_t^T \sigma_i^*(s,u)\,du = \beta_i(t,T)\,\sigma_i^*(s,t) + \frac{1}{2}\beta_i^2(t,T)\,\sigma_i^2(s,t), \tag{4.7}$$

$$\int_t^T \sigma_i(s,u)\,du = \beta_i(t,T)\,\sigma_i(s,t), \tag{4.8}$$

*where $\sigma_i^*(s,u) = \sigma_i(s,u)\int_s^u \sigma_i(s,v)\,dv$.*

*Proof.* Denoting by $\chi_i(t,T)$ the term on the left hand side of (4.6)

$$\begin{aligned}
\chi_i(t,T) &= \int_t^T e^{-\int_t^u \lambda_i(x)\,dx} \int_t^u e^{-\int_t^v \lambda_i(x)\,dx} dv\,du \\
&= \int_t^T \frac{d}{du}\left[\int_t^u e^{-\int_t^v \lambda_i(x)\,dx} dv\right] \int_t^u e^{-\int_t^v \lambda_i(x)\,dx} dv\,du \\
&= \int_t^T \beta_i(t,u)\frac{d}{du}\beta_i(t,u)\,du = \int_t^T d\left[\frac{1}{2}\beta_i^2(t,u)\right] \\
&= \frac{1}{2}\left[\beta_i(t,T) - \beta_i(t,t)\right] \\
&= \frac{1}{2}\beta_i^2(t,T), \quad \text{since } \beta_i(t,t) = 0.
\end{aligned}$$

Next (4.7) is proved. From (4.4)

$$\begin{aligned}
\int_t^T \sigma_i^*(s,u)\,du &= \int_t^T \sigma_i(s,u) \int_s^u \sigma_i(s,v)\,dv\,du \\
&= \sigma_i(s,t)\int_t^T e^{-\int_t^u \lambda_i(x)\,dx}\left[\int_s^t \sigma_i(s,v)\,dv + \int_t^u \sigma_i(s,v)\,dv\right] du \\
&= \sigma_i(s,t)\int_t^T e^{-\int_t^u \lambda_i(x)\,dx} du \int_s^t \sigma_i(s,v)\,dv + \sigma_i^2(s,t)\,\chi_i(t,T),
\end{aligned}$$

which is (4.7). Similar arguments establish (4.8). ☐

Finally, the [IK98] formula for the bond price can now be obtained in terms of the state variables. Recall from (2.3) and (2.4) the definition of the state variables $\phi_{i,j}^{p,q}(t)$ and $\psi_{i,j}^r(t)$. In the present case, $j = 1$ is the only relevant second subscript, and so the simpler notation $\phi_i^{p,q}(t) = \phi_{i,1}^{p,q}(t)$ and $\psi_i^r(t) = \psi_{i,1}^r(t)$ is adopted. Note that $\varsigma_{i,1} = 0$.

In view of (1.1) and (1.3), the price of a pure discount bond is given by

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left[-\sum_{i=1}^n \int_t^T \psi_i^0(u)\,du\right]. \tag{4.9}$$

The following bond price formula is contained in [RS95, p60] for the one factor case, and [IK98, p431] for the multifactor case.

**Theorem 4.5.** *If $\sigma_i(t,T) = c_i(t)\,e^{\int_0^T \lambda_i(x)\,dx}$, then the bond price is given by the formula*

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left[-\Phi(t,T) - \Psi(t,T) - \beta_n(t,T)\,(r(t) - f(0,t))\right], \tag{4.10}$$

*where* $\beta_i(t,T) = \int_t^T e^{-\int_t^u \lambda_i(x)\,dx} du$, *for* $1 \le i \le n$,

$$\Phi(t,T) = \frac{1}{2}\sum_{i=1}^{n} \phi_i^{0,0}(t)\,\beta_i^2(t,T), \quad \textit{and}$$

$$\Psi(t,T) = \sum_{i=1}^{n-1} \psi_i^0(t)\,[\beta_i(t,T) - \beta_n(t,T)]\,.$$

*Proof.* Equation (4.9) required computation of $\int_t^T \psi_i^0(u)\,du$, which can be expressed as

$$\begin{aligned}
\int_t^T \psi_i^0(u)\,du = \int_t^T \psi_i^0(u)\,du &= \int_t^T \left[\int_0^t \sigma_i^*(s,u)\,ds + \int_0^t \sigma_i(s,u)\,d\widetilde{W}_s^i\right] du \\
&= \int_0^t \int_t^T \sigma_i^*(s,u)\,du\,ds + \int_0^t \int_t^T \sigma_i(s,u)\,du\,d\widetilde{W}_s^i \\
&= \int_0^t \left[\beta_i(t,T)\,\sigma_i^*(s,t) + \frac{1}{2}\beta_i^2(t,T)\,\sigma_i^2(s,t)\right] ds \\
&\qquad + \int_0^t \beta_i(t,T)\,\sigma_i(s,t)\,\widetilde{W}_s^i \quad \text{by (4.7) and (4.8)} \\
&= \beta_i(t,T)\,\psi_i^0(t) + \frac{1}{2}\beta_i^2(t,T)\,\phi_i^{0,0}(t),
\end{aligned}$$

where Fubini Theorem was used in the third equality. It follows that

$$\begin{aligned}
\sum_{i=1}^{n} \int_t^T \psi_i^0(u)\,du &= \sum_{i=1}^{n} \beta_i(t,T)\,\psi_i^0(t) + \frac{1}{2}\beta_i^2(t,T)\,\phi_i^{0,0}(t) \\
&= \Phi(t,T) + \sum_{i=1}^{n} \left[(\beta_i(t,T) - \beta_n(t,T)) + \beta_n(t,T)\right]\,\psi_i^0(t) \\
&= \Phi(t,T) + \Psi(t,T) + \beta_n(t,T)\sum_{i=1}^{n} \psi_i^0(t) \\
&= \Phi(t,T) + \Psi(t,T) + \beta_n(t,T)\,\left[r(t) - f(0,t)\right] \quad \text{by (2.8).}
\end{aligned}$$

This completes the proof. ☐

The following corollary shows that the bond price (4.10) extends to the forward rate dependent volatility case.

**Corollary 4.6.** *Let* $\sigma_i(t,T)$ *be given by (2.1), and satisfy (2.2) with* $m_1 = 1$ *for all* $i$, *so that*

$$\sigma_i(t,T) = c_i(t, f(t, t+\varsigma_1), \ldots, f(t, t+\varsigma_m))\,e^{-\int_t^T \lambda_i(u)\,du}. \tag{4.11}$$

*Then the bond price is again given by (4.10).*

*Proof.* The proof of Theorem 4.5 depends only on Lemma 4.3, which is satisfied by $\sigma_i(t,T)$ in (4.11). ☐

4.2. **Constant Coefficient Case.** In this subsection, assume $\lambda_{i,j}(T) = \lambda_{i,j}$ are constants, and $\lambda_{i,j} = \lambda_{i,k}$ if and only if $j = k$. Then

$$\mathcal{L}_i = \prod_{j=1}^{m_i}\left[\frac{\partial}{\partial T} - \lambda_{i,j}\right], \text{ and } \sigma_i(t,T) = \sum_{j=1}^{m_i} c_{i,j}(t)\,e^{\lambda_{i,j}T}. \tag{4.12}$$

When $m_i = 1$, (4.12) is a special case of the [IK98] model, but this assumption is not made here. For notational convenience, it is assumed that $\varsigma_j = 0$ for all $1 \le j \le m$, but the results extend trivially to the setting in which $\varsigma_j$ are arbitrary.

In order to obtain a bond price formula analogous to Theorem 4.5 for this model, additional state variables need to be introduced.

**Lemma 4.7.** *For each* $1 \le i \le n$, *define additional state variables* $\xi_i^{k,l}(t)$, *by*

$$\xi_i^{k,l}(t) = \int_0^t \frac{\partial^k \sigma_i(s,t)}{\partial t^k} \int_s^t \frac{\partial^l \sigma_i(s,u)}{\partial u^l}\, du\, ds. \tag{4.13}$$

*Then the variables* $r(t)$, $\{\psi_{i,1}^m(t) \mid 1 \le i \le n,\ 0 \le m \le m_i,\ (i,m) \ne (n,0)\}$, $\{\phi_{i,1}^{k,l}(t) \mid 1 \le i \le n,\ k \ge 0,\ l \ge 0,\ k+l < m_i\}$, *and* $\{\xi_i^{k,l}(t) \mid 0 \le k,\, l \le m_i\}$, *where* $\psi_{i,1}^m(t)$ *and* $\phi_{i,1}^{k,l}(t)$ *are as defined in (2.3) and (2.4), form a finite dimensional Markovian system.*

*Proof.* In view of Theorem 2.3, it suffices to show that the differentials $d\xi_i^{k,l}(t)$ can be expressed in terms of the state variables. But this follows from

$$\begin{aligned} d\xi_i^{k,l}(t) &= \left[ \int_0^t \left( \frac{\partial^k \sigma_i(s,t)}{\partial t^k} \cdot \frac{\partial^l \sigma_i(s,t)}{\partial t^l} + \frac{\partial^{k+1} \sigma_i(s,t)}{\partial t^{k+l}} \int_s^t \frac{\partial^l \sigma_i(s,u)}{\partial u^l}\, du \right) ds \right] dt \\ &= \left[ \phi_{i,1}^{k\wedge l,\, k\vee l - k\wedge l}(t) + \xi_i^{k+1,l}(t) \right] dt, \end{aligned}$$

where $k \wedge l = \min(k,l)$ and $k \vee l = \max(k,l)$. If $k \ge m_i$ or $l \ge m_i$, then as in Theorem 2.3 $\xi_i^{k,l}(t)$ can be written as linear combinations of $\xi_i^{\mu,\nu}(t)$ with $\mu$, $\nu < m_i$, since $\lambda_{i,j}$ are constants. ☐

The additional state variables introduced in Lemma 4.7 will enable us to obtain a bond price formula, analogous to (4.10).

Let $\Lambda_{i,j}(t,T) = c_{i,j}(t)\, e^{\lambda_{i,j} T}$. Then for each integer $k \ge 0$

$$\frac{\partial^k \sigma_i(t,T)}{\partial t^k} = \sum_{j=1}^{m_i} \lambda_{i,j}^k\, \Lambda_{i,j}(t,T), \tag{4.14}$$

and the state variables in (2.3) and (2.4) can be rewritten as follows, where the second and third subscripts have been omitted as in the previous subsection.

$$\phi_i^{k,l}(t) = \sum_{1 \le p,\, q \le m_i} \lambda_{i,p}^k \lambda_{i,q}^{k+l} \int_0^t \Lambda_{i,p}(s,t)\, \Lambda_{i,q}(s,t)\, ds, \tag{4.15}$$

$$\begin{aligned} \psi_i^m(t) = &\sum_{1 \le p,\, q \le m_i} \lambda_{i,p}^m \int_0^t \Lambda_{i,p}(s,t) \int_s^t \Lambda_{i,q}(s,u)\, du\, ds \\ &+ \sum_{1 \le p \le m_i} \lambda_{i,p}^m \int_0^t \Lambda_{i,p}(s,t)\, d\widetilde{W}_s^p. \end{aligned} \tag{4.16}$$

Note that by arguments similar to those used in Lemma 4.3, it is possible to obtain

$$\Lambda_{i,j}(s,t+u) = \Lambda_{i,j}(s,t)\, e^{\lambda_{i,j}(u-t)} \text{ and } \Lambda_{i,j}(s,t) = \Lambda_{i,j}(s,s)\, e^{\lambda_{i,j}(t-s)}. \tag{4.17}$$

For each $1 \le i \le n$, define the $m_i \times m_i$ van der Monde matrix $\mathcal{M}_i$ by

$$\mathcal{M}_i = \left[ M_i^{j,k} \right] = \left[ \begin{array}{ccccc} 1 & 1 & 1 & \cdots & 1 \\ \lambda_{i,1} & \lambda_{i,2} & \lambda_{i,3} & \cdots & \lambda_{i,m_i} \\ \lambda_{i,1}^2 & \lambda_{i,2}^2 & \lambda_{i,3}^2 & \cdots & \lambda_{i,m_i}^2 \\ \multicolumn{5}{c}{\dotfill} \\ \lambda_{i,1}^{m_i-1} & \lambda_{i,2}^{m_i-1} & \lambda_{i,3}^{m_i-1} & \cdots & \lambda_{i,m_i}^{m_i-1} \end{array} \right]. \tag{4.18}$$

Since $\lambda_{i,j} = \lambda_{i,k}$ if and only if $j = k$, by assumption, $\mathcal{M}_i$ are invertible. Denoting the inverse of $\mathcal{M}_i$ by $\mathcal{N}_i = \left[N_i^{j,k}\right]$, and writing

$$\dot{\partial}_T\sigma_i(t,T) = \left[\frac{\partial^0\sigma_i(t,T)}{\partial T^0}, \frac{\partial^1\sigma_i(t,T)}{\partial T^1}, \dots, \frac{\partial_T^{m_i-1}\sigma_i(t,T)}{\partial T^{m_i-1}}\right]^\tau, \quad \text{and} \tag{4.19}$$

$$\Lambda_{i,\cdot}(t,T) = \left[\Lambda_{i,1}(t,T), \Lambda_{i,2}(t,T), \dots, \Lambda_{i,m_i}(t,T)\right]^\tau, \tag{4.20}$$

where the superscript $^\tau$ represents matrix transposition, the following identities are immediately obtained from definitions and (4.14):

$$\mathcal{M}_i \times \mathcal{N}_i = I_{m_i\times m_i} = \mathcal{N}_i \times \mathcal{M}_i, \tag{4.21}$$

$$\dot{\partial}_t\sigma_i(t,T) = \mathcal{M}_i \times \Lambda_{i,\cdot}(t,T), \tag{4.22}$$

$$\Lambda_{i,\cdot}(t,T) = \mathcal{N}_i \times \dot{\partial}_T\sigma_i(t,T). \tag{4.23}$$

Using (4.23) $\Lambda_{i,j}(t,T)$ can be expressed as the linear combination

$$\Lambda_{i,j}(t,T) = \sum_{k=1}^{m_i} N_i^{j,k}\,\frac{\partial^{k-1}\sigma_i(t,T)}{\partial T^{k-1}}, \tag{4.24}$$

where $N_i^{j,k}$ are *constants*. The following lemma plays a role similar to Lemma 4.4 in the $m_i = 1$ case.

**Lemma 4.8.** *Let $\Lambda_{i,j}(t,T)$ be as defined above, and let*

$$\beta_i^j(t,T) = \int_t^T e^{\lambda_{i,j}(u-t)}du = \frac{1}{\lambda_{i,j}}\left[e^{\lambda_{i,j}(T-t)} - 1\right], \quad \textit{and} \tag{4.25}$$

$$\begin{aligned}\gamma_i^{j,k}(t,T) &= \int_t^T e^{\lambda_{i,j}(u-t)}\int_t^u e^{\lambda_{i,k}(v-t)}dv\,du \qquad (4.26)\\ &= \frac{1}{\lambda_{i,k}}\left[\frac{1}{\lambda_{i,j}+\lambda_{i,k}}\left(e^{(\lambda_{i,j}+\lambda_{i,k})(T-t)} - 1\right) - \frac{1}{\lambda_{i,j}}\left(e^{\lambda_{i,j}(T-t)} - 1\right)\right].\end{aligned}$$

*Then the following identities hold:*

$$\int_t^T \Lambda_i^{j,k}(s,u)\,du = \gamma_i^{j,k}(t,T)\,\Lambda_{i,j}(s,t)\,\Lambda_{i,k}(s,t) + \Gamma_i^{j,k}(s,t), \tag{4.27}$$

$$\int_t^T \Lambda_{i,j}(s,u)\,du = \beta_{i,j}(t,T)\Lambda_{i,j}(s,t), \tag{4.28}$$

*where $\Lambda_i^{j,k}(s,u) = \Lambda_{i,j}(s,u)\int_s^u \Lambda_{i,k}(s,v)\,dv$, and $\Gamma_i^{j,k}(s,t) = \beta_{i,j}(t,T)\,\Lambda_i^{j,k}(s,t)$.*

*Proof.* The arguments used in Lemma 4.4 apply here in view of (4.17). □

The bond price formula for the constant coefficient case can now be stated.

**Theorem 4.9.** *Let $\sigma_i(t,T) = \sum_{j=1}^{m_i} c_{i,j}(t)\,e^{\lambda_{i,j}T}$, where $\lambda_{i,j}$ are distinct constants, and let $\mathcal{N}_i = \left[N_i^{j,k}\right]$ be the inverse of $\mathcal{M}_i = \left[M_i^{j,k}\right] = \left[\lambda_{i,j}^{k-1}\right]$. Then the bond price is given by the formula*

$$P(t,T) = \frac{P(0,T)}{P(0,t)}\exp\left[-\Phi(t,T) - \Psi(t,T) - \Theta(t,T)\right], \tag{4.29}$$

*where*

$$\begin{aligned}
\Phi(t,T) &= \sum_{i=1}^{n} \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \beta_{i,j}(t,T)\, \xi_i^{l-1,m-1}(t), \\
\Psi(t,T) &= \sum_{i=1}^{n} \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \gamma_i^{j,k}(t,T)\, \phi_i^{l\wedge m-1,\, l\vee m - l\wedge m}(t), \\
\Theta(t,T) &= \sum_{i=1}^{n} \sum_{j,k} N_i^{j,k}\, \beta_{i,j}(t,T) \left[\psi_i^{k-1}(t) - \xi_i^{k-1,0}(t)\right],
\end{aligned}$$

*and* $1 \le j,\, k,\, l,\, m \le m_i$*. Here,* $l \wedge m = \min(l,m)$ *and* $l \vee m = \max(l,m)$*.*

*Proof.* As in Theorem 4.5, (4.9) dictates that $\sum_{i=1}^{n} \int_t^T \phi_i^0(u)\, du$ must be computed. Now for each $1 \le i \le n$

$$\int_t^T \phi_i^0(u)\, du = \int_t^T \left[\int_0^t \sigma_i(s,u) \int_s^t \sigma_i(s,v)\, dv\, ds + \int_0^t \sigma_i(s,u)\, d\widetilde{W}_s^i\right] du.$$

Applying Fubini Theorem, the above integral can be written

$$\int_0^t \left[\int_t^T \sigma_i(s,u) \int_s^t \sigma_i(s,v)\, dv\, du\right] ds + \int_0^t \left[\int_t^T \sigma_i(s,u)\, du\right] d\widetilde{W}_s^i.$$

Let $\mathcal{I}_i^1(t,T)$ and $\mathcal{I}_i^2(t,T)$ denote the two integrals above. Then

$$\begin{aligned}
\mathcal{I}_i^1(t,T) &= \sum_{j,k} \int_0^t \left[\int_t^T \Lambda_{i,j}(s,u) \int_s^u \Lambda_{i,k}(s,v)\, dv\, du\right] ds \qquad \text{by (4.14)} \\
&= \sum_{j,k} \int_0^t \left[\gamma_i^{j,k}(t,T)\Lambda_{i,j}(s,t)\Lambda_{i,k}(s,t) + \Gamma_i^{j,k}(s,t)\right] ds \qquad \text{by (4.27)} \\
&= \sum_{j,k} \gamma_i^{j,k}(t,T) \int_0^t \Lambda_{i,j}(s,t)\Lambda_{i,k}(s,t)\, ds \\
&\quad + \sum_{j,k} \beta_{i,j}(t,T) \int_0^t \Lambda_{i,j}(s,t) \int_s^t \Lambda_{i,k}(s,v)\, dv\, ds \\
&= \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \gamma_i^{j,k}(t,T) \int_0^t \frac{\partial^{l-1}\sigma_i(s,t)}{\partial t^{l-1}} \cdot \frac{\partial^{m-1}\sigma_i(s,t)}{\partial t^{m-1}}\, ds \\
&\quad + \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \beta_{i,j}(t,T) \int_0^t \frac{\partial^{l-1}\sigma_i(s,t)}{\partial t^{l-1}} \int_s^t \frac{\partial^{m-1}\sigma_i(s,v)}{\partial v^{m-1}}\, dv\, ds \\
&= \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \beta_{i,j}(t,T)\, \xi_i^{l-1,m-1}(t) \\
&\quad + \sum_{j,k,l,m} N_i^{j,l}\, N_i^{k,m}\, \gamma_i^{j,k}(t,T)\, \phi_i^{l\wedge m-1,\, l\vee m-l\wedge m}(t),
\end{aligned}$$

where the second last equality follows from (4.24). Summing over $i$, $\sum_{i=1}^{n} \mathcal{I}_i^1(t,T) = \Phi(t,T) + \Psi(t,T)$. Next $\mathcal{I}_i^2(t,T)$ is considered in a similar fashion.

$$\begin{aligned}
\mathcal{I}_i^2(t,T) &= \sum_j \int_0^t \int_t^T \Lambda_{i,j}(s,u)\, du\, d\widetilde{W}_s^i \quad \text{by (4.14)} \\
&= \sum_j \beta_{i,j}(t,T) \int_0^t \Lambda_{i,j}(s,t)\, d\widetilde{W}_s^i \quad \text{by (4.28)} \\
&= \sum_{j,k} N_i^{j,k}\, \beta_{i,j}(t,T) \int_0^t \frac{\partial^{k-1}\sigma_i(s,t)}{\partial t^{k-1}}\, d\widetilde{W}_s^i \quad \text{by (4.24)} \\
&= \sum_{j,k} N_i^{j,k}\, \beta_{i,j}(t,T) \left[\psi_i^{k-1}(t) - \xi_i^{k-1,0}(t)\right].
\end{aligned}$$

Summing over $i$, $\sum_{i=1}^{n} \mathcal{I}_i^2(t,T) = \Theta(t,T)$, and (4.29) follows. □

*Remark* 4.10. Note that all the arguments in this subsection remain valid even when $\lambda_{k,l} = x_{k,l} + i\, y_{k,l}$ are complex. So taking $m_k = 2m_k'$, and $\{\lambda_{k,l}\}$ consisting of conjugate pairs, Theorem 4.9 applies to volatility functions

$$\sigma_i(t,T) = \sum_{j=1}^{m_i'} c_{i,j}(t)\, e^{x_{i,j}T} \cos(y_{i,j}T) + d_{i,j}(t)\, e^{x_{i,j}T} \sin(y_{i,j}T),$$

which corresponds to taking the operator

$$\mathcal{L}_i = \prod_{j=1}^{m_i'} \left[\frac{\partial^2}{\partial T^2} - 2x_{i,j}\frac{\partial}{\partial T} + (x_{i,j}^2 + y_{i,j}^2)\right].$$

In particular, taking $x_{i,j} = 0$, the results of this subsection apply to $\sigma_i(t,T)$ that are Fourier series like in the variable $T$.

## 5. Conclusion

This paper has established very general conditions on the forward rate volatility processes under which the Heath-Jarrow-Morton model transforms to a finite dimensional Markovian system. The characterisations described will allow the construction of term structure models with volatility processes that depend on a set of fixed tenor forward rates. Given that only a certain number of fixed tenors are actively traded in most markets, such characterisations should suffice for most practical implementations.

It is evident from Theorem 2.3 that the size of finite dimensional Markovian representations can grow very rapidly, and consequently the $m_i$ would need to be fairly low. With regard to actual implementations, it has already been explained how the implementations of [BC97], [Car94], [RS95], and [IK98] can be represented in this framework. Once a finite dimensional Markovian representation is established, derivative prices can be obtained either by solving the PDE, obtained through the application of the Feynman-Kac Theorem, or by Monte Carlo simulation. As a general rule, it is convenient to apply the PDE approach for low dimensional systems, and to use the Monte Carlo methods for higher dimensional characterisations.

The fact that a formula for the bond price can be obtained is of great utility, since one need only solve the PDE or perform Monte Carlo simulations over the life of the derivative security of interest, which generally is of much shorter maturity than the underlying bond or swap.

Finally some numerical implementations of the models presented in this paper are mentioned. Firstly, Chiarella and El-Hassan [CEH99] have considered the [RS95] type volatility function and solved the PDE for the American bond option problem using the method of lines. In [BCEHZ99], Bhar, Chiarella, El-Hassan and Zheng consider a volatility process dependent on spot rate and one fixed tenor forward rate. They use the alternating direction implicit method to solve the three spatial variable PDE for the European bond option prices, and compare the results with those obtained using Monte Carlo simulation. This is probably the highest dimension that can be handled conveniently by the PDE methods. In [CEHZ99], Chiarella, El-Hassan and Zheng apply the Monte Carlo methods on forward rate volatilities that depend on the spot rate and a small number of fixed tenor forward rates to evaluate American bond options. Finally, in [CK98a] Chiarella and Kwon develop further the framework of this paper to obtain Markovian transformations of Heath-Jarrow-Morton models with stochastic volatility.

## References

[BC97] R. Bhar and C. Chiarella, *Transformation of Heath-Jarrow-Morton Models to Markovian Systems*, European Journal of Finance **3** (1997), 1–26.

[BCEHZ99] R. Bhar, C. Chiarella, N. El-Hassan, and X. Zheng, *Numerical Methods for Heath-Jarrow-Morton Models with Forward Rate Dependent Volatility Functions*, Working paper, School of Finance and Economics, University of Techonology Sydney, 1999.

[BG99] T. Björk and A. Gombani, *Minimal Realisations of Interest Models*, Working paper, Stockholm School of Economics, 1999.

[Bjö96] T. Björk, *Interest Rate Theory*, Financial Mathematics: Bressanone 1996, Springer Verlag Lecture Notes in Mathematics **1656** (1996), 53–122, ed. W. Runggaldier.

[BK91] F. Black and P. Karasinksi, *Bonds and Option Pricing when Short Rates are Lognormal*, Financial Analysts Journal (1991), 52–59.

[BM94] A. Brace and M. Musiela, *A Multifactor Gauss Markov Implementation of Heath, Jarrow, and Morton*, Mathematical Finance **4** (1994), no. 3, 259–283.

[BS99] T. Björk and L. Svensson, *On the Existence of Finite Dimensional Realisations for Nonlinear Forward Rate Models*, Working paper, Stockholm School of Economics, 1999.

[Car94] A. Carverhill, *When is Spot Rate Markovian?*, Mathematical Finance **4** (1994), no. 4, 305–312.

[CEH99] C. Chiarella and N. El-Hassan, *Pricing American Bond Options in the Heath-Jarrow-Morton Framework*, Working paper, School of Finance and Economics, University of Techonology Sydney, 1999.

[CEHZ99] C. Chiarella, N. El-Hassan, and X. Zheng, *American Options in the Heath-Jarrow-Morton Framework using Monte Carlo Simulation*, Working paper, School of Finance and Economics, University of Techonology Sydney, 1999.

[CIR85] J. Cox, J. Ingersoll, and S. Ross, *An Intertemporal General Equilibrium Model of Asset Prices*, Econometrica **53** (1985), no. 2, 386–384.

[CK98a] C. Chiarella and O. Kwon, *A Class of Heath-Jarrow-Morton Term Structure Models with Stochastic Volatility*, Working paper, School of Finance and Economics, University of Techonology Sydney, 1998.

[CK98b] C. Chiarella and O. Kwon, *Square Root Affine Transformations of the Heath-Jarrow-Morton Term Structure Model and Partial Differential Equations*, Working paper, School of Finance and Economics, University of Techonology Sydney, 1998.

[HJM92] D. Heath, R. Jarrow, and A. Morton, *Bond Princing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation*, Econometrica **60** (1992), no. 1, 77–105.

[HK79] J. Harrison and D. Kreps, *Martingales and Arbitrage in the Multiperiod Security Markets*, Journal of Economic Theory **20** (1979), 381–408.

[HP81] J. Harrison and S. Pliska, *Martingales and Stochastic Integrals in the Theory of Continuous Trading*, Stochastic Processes and Their Applications **11** (1981), 215–260.

[HW90] J. Hull and A. White, *Pricing Interest Rate Derivative Securities*, Review of Financial Studies **3** (1990), 573–592.

[IK98] K. Inui and M. Kijima, *A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models*, Journal of Financial and Quantitative Analysis **33** (1998), no. 3, 423–440.

[MR97] M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, first ed., Springer-Verlag, New York, 1997, Applications of Mathematics **30**.

[RS95] P. Ritchken and L. Sankarasubramanian, *Volatility Structures of Forward Rates and the Dynamics of the Term Structure*, Mathematical Finance **5** (1995), no. 1, 55–72.

[Vas77] O. Vasicek, *An Equilibrium Characterisation of the Term Structure*, Journal of Financial Economics **5** (1977), 177–188.