

A Note on the Dai–Singleton Canonical Representation of Affine Term Structure Models

Patrick Cheridito *
ORFE
Princeton University
dito@princeton.edu

Damir Filipović †
Vienna Institute of Finance,
University of Vienna, and
Vienna University of Economics and
Business Administration
damir.filipovic@vif.ac.at

Robert L. Kimmel ‡
Fisher College of Business
The Ohio State University
kimmel.42@osu.edu

September 2008

Abstract

Dai and Singleton (2000) study a class of term structure models for interest rates that specify the short rate as an affine combination of the components of an N -dimensional affine diffusion process. Observable quantities in such models are invariant under regular affine transformations of the underlying diffusion process. In their canonical form, the models in Dai and Singleton (2000) are based on diffusion processes with diagonal diffusion matrices. This motivates the following question: Can the diffusion matrix of an affine diffusion process always be diagonalized by means of a regular affine transformation?

We show that if the state space of the diffusion is of the form $D = \mathbb{R}_+^m \times \mathbb{R}^{N-m}$ for integers $0 \leq m \leq N$ satisfying $m \leq 1$ or $m \geq N - 1$, there exists a regular affine transformation of D onto itself that diagonalizes the diffusion matrix. So in this case, the Dai–Singleton canonical representation is exhaustive. On the other hand, we provide examples of affine diffusion processes with state space $\mathbb{R}_+^2 \times \mathbb{R}^2$ whose diffusion matrices cannot be diagonalized through regular affine transformation. This shows that for $2 \leq m \leq N - 2$, the assumption of diagonal diffusion matrices may impose unnecessary restrictions and result in an avoidable loss of generality.

Key words: affine diffusion processes, affine transformations, diagonal diffusion matrices.

1 Introduction

Continuous-time affine models have played a prominent role in both the term structure literature and the stochastic volatility literature. This prominence is no doubt largely due to their

An earlier version of this paper was circulated under the title “A Note on the Canonical Representation of Affine Diffusion Processes”. We thank Kenneth Singleton for helpful comments. Damir Filipović thanks Darrell Duffie and George Papanicolaou for their kind hospitality during his stay at Stanford University in September 2006.

*Supported by NSF Grant DMS-0505932, a Rheinstein Award and a Peek Fellowship.

†Supported by WWTF (Vienna Science and Technology Fund).

‡Supported by the Fisher College of Business and the Charles A. Dice Center for Research in Financial Economics.

analytic tractability. The early literature focused on specific models; see, for example, Vasicek (1977) and Cox et al. (1985) for single-factor term structure models, Chen (1996), Balduzzi et al. (1996) for multiple-factor term structure models, and Hull and White (1987) and Heston (1993) for asset price models with stochastic volatility. A more recent strand of the literature, however, focuses on broad classes of models rather than specific cases. For the case of affine term structure models (ATSMs), see, for example, Duffie and Kan (1996) and Duffie et al. (2003) for systematic treatments, Dai and Singleton (2000) for an empirical investigation and classification scheme, Duffie (2002) and Cheridito et al. (2007) for extended market price of risk specifications, Collin-Dufresne et al. (2008) and Joslin (2006) for alternate classification schemes, and Duffie et al. (2000) for additional applications of affine processes.

Dai and Singleton (2000) group ATSMs into families $\mathbb{A}_m(N)$, where $N \geq 1$ denotes the dimension of the underlying state process, and $0 \leq m \leq N$ the number of linearly independent components determining the conditional variances and covariances. The conditional variances and covariances appear as the components of the diffusion matrix in the stochastic differential equation that governs the dynamics of the state process. Any regular affine transformation of the state process of an ATSM leads to another ATSM that produces the same short rates and bond prices. That is, the observable implications of the model are invariant to regular affine transformation of the state process. ATSMs that may thus appear to be distinct from the stochastic differential equation of the state process and the interest rate specification, can in fact generate identical term structure implications. Kwon (2007) provides a group theoretic interpretation of this fact.

Dai and Singleton (2000) specify a canonical form for ATSMs and impose parameter restrictions to ensure that the underlying affine state variable process exists and that ATSMs with the same observable implications have a unique canonical representation. In the canonical form, the ATSMs in Dai and Singleton (2000) are based on affine diffusion processes with state space $D := \mathbb{R}_+^m \times \mathbb{R}^{N-m}$ and diagonal diffusion matrix.

In this paper we investigate the question, whether any N -dimensional affine diffusion process can be brought into this form through a regular affine transformation. Obviously, this is only possible for affine diffusion processes defined on state spaces of the form $\Lambda D + \lambda$ for a regular $N \times N$ -matrix Λ and $\lambda \in \mathbb{R}^N$. Such state spaces are natural for affine processes. But there also exist affine processes with different state spaces. Some examples are discussed in Section 12 of Duffie et al. (2003), and Gouriéroux and Sufana (2006) provide a classification of two-factor affine diffusion models with general state spaces. But until recently, such models have not been common in applications; see, however, Da Fonseca et al. (2008) or Buraschi et al. (2006). To answer the question for affine diffusion processes on state spaces of the form $\Lambda D + \lambda$, it is sufficient to consider the case D because $\Lambda D + \lambda$ can be mapped onto D with the regular affine transformation $x \mapsto \Lambda^{-1}(x - \lambda)$. We prove that for all $0 \leq m \leq N$ such that $m \leq 1$ or $m \geq N - 1$, every affine diffusion process with state space D can be turned into an affine diffusion process with diagonal diffusion matrix through a regular affine transformation which leaves D invariant, and we provide two counter-examples in $\mathbb{A}_2(4)$ which show that these conditions cannot be weakened in general. Many of the early studies restrict attention to the case $N \leq 3$. Then at least one of our conditions is always satisfied, and every affine diffusion process with state space D can be diagonalized. However, models with $N \geq 4$ are becoming more and more common in the literature; see, for example, Thompson (2008), Collin-Dufresne et al. (2008), and Egorov et al. (2008). Four factor models also arise naturally when the stochastic volatility model of Hull and White (1987) and Heston (1993) are extended to

include two stocks. Requiring that the state variables have a diagonal diffusion matrix imposes unnecessary restrictions on these models that may impede their ability to match important features of the data. By comparing the ATSMs of Dai and Singleton (2000) to the maximal representations in their own paper, Collin-Dufresne et al. (2008) also conclude that for $N \leq 3$, the Dai–Singleton specification covers all ATSMs with state space D , while for $N = 4$, this is not the case. Our result implies the specific findings of Collin-Dufresne et al. (2008) for $N \leq 4$.

The remainder of the paper is organized as follows: In Section 2 we formally introduce affine diffusion processes and state our main result. Although we focus on the case of ATSMs, the results apply to other cases, such as affine stochastic volatility models. Section 3 contains the proof of the main result. In Section 4 we provide the two counter-examples showing that affine processes cannot in general be diagonalized. Finally, Section 5 concludes.

2 The problem and the main result

ATSMs specify the short rate $r(t)$ as an affine function of an N -dimensional affine diffusion process $X(t) = (X_1(t), \dots, X_N(t))$:

$$(2.1) \quad r(t) = d_0 + d^T X(t).$$

Here, $d_0 \in \mathbb{R}$, $d \in \mathbb{R}^N$, and T denotes transposition. Prices of zero-coupon bonds are given by

$$P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(u) du} \right],$$

where $\mathbb{E}_t^{\mathbb{Q}}$ denotes conditional expectation under a risk neutral probability measure \mathbb{Q} . By “affine diffusion process” it is meant that $X(t)$ is a solution of a stochastic differential equation of the form

$$(2.2) \quad dX(t) = \mu(X(t)) dt + \sigma(X(t)) dW(t),$$

where $W(t)$ is an N -dimensional \mathbb{Q} -Brownian motion and the drift μ and the diffusion matrix $\alpha := \sigma\sigma^T$ are affine functions of $X(t)$:

$$\mu(X(t)) = b + \beta X(t),$$

for $b \in \mathbb{R}^N$ and $\beta \in \mathbb{R}^{N \times N}$, and

$$\alpha(X(t)) = \alpha^0 + X_1(t)\alpha^1 + \dots + X_N(t)\alpha^N$$

for symmetric $N \times N$ -matrices $\alpha^0, \dots, \alpha^N$.

It is known from Duffie and Kan (1996) that zero coupon bond prices in ATSMs are of the form

$$P(t, T) = \exp(-A(T-t) - \langle B(T-t), X(t) \rangle),$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in \mathbb{R}^N and A and B are deterministic functions satisfying certain ordinary differential equations. For a detailed study of affine processes in a more general context, we refer to Duffie et al. (2003).

It can easily be checked that for every regular $N \times N$ -matrix Λ and $\lambda \in \mathbb{R}^N$, the affine transform

$$(2.3) \quad Y(t) = \Lambda X(t) + \lambda,$$

satisfies

$$dY(t) = [\Lambda b - \Lambda\beta\Lambda^{-1}\lambda + \Lambda\beta\Lambda^{-1}Y(t)] dt + \Lambda\sigma(\Lambda^{-1}[Y(t) - \lambda]) dW(t).$$

Hence, $Y(t)$ is again an affine diffusion process with affine diffusion matrix

$$(2.4) \quad \Lambda\alpha(\Lambda^{-1}[Y(t) - \lambda])\Lambda^T = \Lambda\alpha(X(t))\Lambda^T.$$

The short rate process can be expressed in terms of $Y(t)$ as

$$(2.5) \quad r(t) = d_0 - d^T\Lambda^{-1}\lambda + d^T\Lambda^{-1}Y(t).$$

This shows that $Y(t)$ and (2.5) specify an ATSM producing the same short rates and bond prices as $X(t)$ and (2.1). That is, a regular affine transformation of the state process changes the particular form of the stochastic differential equation (2.2). But it leaves observable quantities, such as short rates and bond prices invariant. In their canonical form, the models in Dai and Singleton (2000) are based on diffusion processes with state space $D := \mathbb{R}_+^m \times \mathbb{R}^{N-m}$ and diagonal diffusion matrix. This motivates the question whether all affine diffusion processes can be brought into such a form via a regular affine transformation. Of course, this is only possible for affine diffusion processes on state spaces of the form $\Lambda D + \lambda$, and to cover this case, it is enough to consider D . Our main result is the following:

Theorem 2.1 *Let X be an affine diffusion process with state space D and diffusion matrix $\alpha : D \rightarrow \mathbb{R}^{N \times N}$. If $m \leq 1$ or $m \geq N - 1$, then there exists a regular $N \times N$ -matrix Λ such that*

$$(2.6) \quad \Lambda D = D$$

and

$$(2.7) \quad \Lambda\alpha(x)\Lambda^T \quad \text{is diagonal for all } x \in D.$$

Thus, in view of (2.4), the diffusion matrix of the D -valued affine diffusion process $Y = \Lambda X$ is diagonal.

Note that we only require Λ to be regular and not orthogonal; see also Remark 3.1 below. Since at least one of the conditions in Theorem 2.1 is always satisfied for $N \leq 3$, we obtain the following corollary:

Corollary 2.2 *For $N \leq 3$, every affine diffusion process with state space D can be turned into an affine diffusion process with diagonal diffusion matrix through a regular $N \times N$ matrix Λ such that $\Lambda D = D$.*

3 Proof of Theorem 2.1

From Theorem 2.7 of Duffie et al. (2003) we know that a function $\alpha : D \rightarrow \mathbb{R}^{N \times N}$ is equal to the diffusion matrix $\sigma\sigma^T$ of an affine diffusion process with state space D if and only if it is of the form

$$(3.1) \quad \alpha(x) = \alpha^0 + x_1\alpha^1 + \cdots + x_m\alpha^m, \quad x \in D,$$

where $\alpha^0, \dots, \alpha^m$ are positive semi-definite symmetric $N \times N$ -matrices satisfying

$$(3.2) \quad \alpha_{kk}^j = 0, \quad \text{for } 0 \leq j \leq m \text{ and } 1 \leq k \leq m \text{ with } k \neq j.$$

By positive semi-definiteness and symmetry, condition (3.2) immediately implies that

$$(3.3) \quad \alpha_{kl}^j = \alpha_{lk}^j = 0, \quad \text{for } 0 \leq j \leq m \text{ and } 1 \leq k \leq m \text{ with } k \neq j \text{ and } 1 \leq l \leq N.$$

We illustrate conditions (3.1) and (3.3) for $N = 3$. The first case is $m = 0$. Then,

$$\alpha(x) \equiv \alpha^0$$

for an arbitrary positive semi-definite symmetric $N \times N$ -matrix α^0 . For $m = 1$, we have

$$\alpha^0 = \begin{pmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{pmatrix}, \quad \alpha^1 = \begin{pmatrix} + & * & * \\ & + & * \\ & & + \end{pmatrix},$$

for $m = 2$,

$$\alpha^0 = \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{pmatrix}, \quad \alpha^1 = \begin{pmatrix} + & 0 & * \\ & 0 & 0 \\ & & + \end{pmatrix}, \quad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{pmatrix},$$

and for $m = 3$,

$$\alpha^0 = 0, \quad \alpha^1 = \begin{pmatrix} + & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix}, \quad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 \\ & + & 0 \\ & & 0 \end{pmatrix}, \quad \alpha^3 = \begin{pmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{pmatrix},$$

where we leave the lower triangle of symmetric matrices blank, + denotes a non-negative real number and * any real number such that positive semi-definiteness holds.

It is immediate from (3.1) that a regular $N \times N$ matrix Λ fulfills (2.7) if and only if

$$(3.4) \quad \Lambda \alpha^0 \Lambda^T, \dots, \Lambda \alpha^m \Lambda^T \quad \text{are diagonal.}$$

Denote by e^1, \dots, e^N the standard basis in \mathbb{R}^N . Then $\Lambda \alpha^j \Lambda^T$ is diagonal if and only if

$$(3.5) \quad \langle \Lambda^T e^k, \alpha^j \Lambda^T e^l \rangle = \langle e^k, \Lambda \alpha^j \Lambda^T e^l \rangle = 0 \quad \text{for all } 1 \leq k \neq l \leq N.$$

Hence, the existence of a regular Λ satisfying (2.7) is equivalent to the existence of linearly independent vectors f^1, \dots, f^N in \mathbb{R}^N such that

$$(3.6) \quad \langle f^k, \alpha^j f^l \rangle = 0 \quad \text{for all } 0 \leq j \leq m \text{ and all } 1 \leq k \neq l \leq N.$$

To show (2.6) and (3.4), we consider the four cases $m = N$, $m = 0$, $m = N - 1$ and $m = 1$ separately.

Case $m = N$: By (3.3), $\alpha^0, \dots, \alpha^N$ are already diagonal. So Λ can be taken to be the $N \times N$ -identity matrix.

Case $m = 0$: By (3.1), $\alpha \equiv \alpha^0$ for a positive semi-definite symmetric $N \times N$ matrix α^0 . So

there exists an orthogonal $N \times N$ matrix Λ such that $\Lambda\alpha^0\Lambda^T$ is diagonal (see for instance, Theorem 2 in Section 16.2 of Gelfand, 1961). Since $D = \mathbb{R}^N$, condition (2.6) is also satisfied.

Case $1 \leq m = N - 1$: It follows from (3.3) that the only component of α^0 that can be non-zero is α_{NN}^0 . For $1 \leq j \leq N - 1$, only the entries $\alpha_{jj}^j, \alpha_{jN}^j, \alpha_{Nj}^j, \alpha_{NN}^j$ of the matrix α^j can be non-zero. Hence, if we define $f^k = e^k$, for $1 \leq k \leq N - 1$, and f^N component-wise as

$$f_i^N = \begin{cases} -\frac{\alpha_{iN}^i}{\alpha_{ii}^i}, & \text{for } 1 \leq i \leq N - 1 \text{ if } \alpha_{ii}^i \neq 0 \\ 0, & \text{for } 1 \leq i \leq N - 1 \text{ if } \alpha_{ii}^i = 0 \\ 1, & \text{for } i = N \end{cases},$$

then f^1, \dots, f^N are linearly independent vectors in \mathbb{R}^N that satisfy (3.6). By (3.5), the regular $N \times N$ -matrix Λ having f^1, \dots, f^N as row vectors, satisfies (3.4), and it follows from the special form of the vectors f^1, \dots, f^N that $\Lambda(\mathbb{R}_+^{N-1} \times \mathbb{R}) = \mathbb{R}_+^{N-1} \times \mathbb{R}$.

Case $m = 1$: By condition (3.3), the matrix α^0 is of the form

$$\alpha^0 = \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix}$$

for a positive semi-definite symmetric $(N - 1) \times (N - 1)$ -matrix A . If $\alpha_{11}^1 \neq 0$, we define the symmetric $N \times N$ -matrix $\tilde{\alpha}^1$ by

$$\tilde{\alpha}_{jk}^1 := \alpha_{jk}^1 - \frac{\alpha_{j1}^1 \alpha_{k1}^1}{\alpha_{11}^1}, \quad 1 \leq j, k \leq N.$$

Notice that, for any $x \in \mathbb{R}^N$, we have

$$\langle x, \tilde{\alpha}^1 x \rangle = \langle x, \alpha^1 x \rangle - \frac{\langle x, \alpha^1 e^1 \rangle^2}{\langle e^1, \alpha^1 e^1 \rangle} \geq \langle x, \alpha^1 x \rangle - \langle x, \alpha^1 x \rangle = 0$$

by the Cauchy–Schwarz inequality $|\langle x, \alpha^1 e^1 \rangle| \leq \sqrt{\langle x, \alpha^1 x \rangle} \sqrt{\langle e^1, \alpha^1 e^1 \rangle}$. This shows that $\tilde{\alpha}^1$ is positive semi-definite. If $\alpha_{11}^1 = 0$, we set $\tilde{\alpha}^1 := \alpha^1$. In either case, $\tilde{\alpha}^1$ is of the form

$$\tilde{\alpha}^1 = \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix}$$

for a positive semi-definite symmetric $(N - 1) \times (N - 1)$ -matrix B . It follows from Theorem 8.7.1 in Golub and Van Loan (1996) that there exists a regular $(N - 1) \times (N - 1)$ -matrix Q such that QAQ^T and QBQ^T are both diagonal. Consider the regular $N \times N$ -matrix

$$(3.7) \quad \Lambda := \begin{pmatrix} 1 & 0 & \dots & 0 \\ \Lambda_{21} & & & \\ \cdot & & Q & \\ \cdot & & & \\ \Lambda_{N1} & & & \end{pmatrix},$$

where the entries $\Lambda_{21}, \dots, \Lambda_{N1}$ are chosen such that

$$(3.8) \quad \langle \Lambda^T e^k, \alpha^1 e^1 \rangle = 0 \quad \text{for all } k \geq 2$$

(note that if $\alpha_{11}^1 = 0$, then $\alpha^1 e^1 = 0$ and (3.8) holds true for arbitrary $\Lambda_{21}, \dots, \Lambda_{N1}$). Then $\Lambda \alpha^0 \Lambda^T$ and $\Lambda \tilde{\alpha}^1 \Lambda^T$ are diagonal. If $\alpha_{11}^1 = 0$, then α^1 is equal to $\tilde{\alpha}^1$. In the case $\alpha_{11}^1 \neq 0$, it follows from the definition of $\tilde{\alpha}^1$ and (3.8) that

$$\langle \Lambda^T e^k, \alpha^1 \Lambda^T e^l \rangle = \langle e^k, \Lambda \tilde{\alpha}^1 \Lambda^T e^l \rangle + \frac{\langle \Lambda^T e^k, \alpha^1 e^1 \rangle \langle \Lambda^T e^l, \alpha^1 e^1 \rangle}{\langle e^1, \alpha^1 e^1 \rangle} = 0$$

for all $k \neq l$. Hence, (3.4) is satisfied. Finally, it can readily be seen from (3.7) that Λ maps $\mathbb{R}_+ \times \mathbb{R}^{N-1}$ onto $\mathbb{R}_+ \times \mathbb{R}^{N-1}$.

Remark 3.1 It is well-known that there exists an *orthogonal* $N \times N$ -matrix Λ satisfying (3.4), and thus (2.7), if and only if $\alpha^0, \dots, \alpha^m$ commute. However, in Theorem 2.1 we do not require Λ to be orthogonal. This makes the transformation more general. For instance, the positive semi-definite symmetric matrices, satisfying condition (3.2) for $N = 3$ and $m = 1$,

$$\alpha^0 = \begin{pmatrix} 0 & 0 & 0 \\ & 1 & -1 \\ & & 1 \end{pmatrix} \quad \text{and} \quad \alpha^1 = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 4 \end{pmatrix}$$

do not commute, but can nevertheless be diagonalized with the regular, non-orthogonal matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 4 & -1 \end{pmatrix}.$$

4 Counter-examples for $2 \leq m \leq N - 2$

In this section, we give two examples of diffusions on $\mathbb{R}_+^2 \times \mathbb{R}^2$ whose diffusion matrix cannot be diagonalized through regular affine transformation. For the first example, we provide a complete proof of non-diagonalizability. But one might object that this example is degenerate because it consists of a diffusion on $\mathbb{R}_+^2 \times \mathbb{R}^2$ whose diffusion matrix has only rank 2. Our second example is a diffusion on $\mathbb{R}_+^2 \times \mathbb{R}^2$ with a diffusion matrix of full rank. Non-diagonalizability can be proved with analogous arguments. But a full-fledged proof is so long and tedious that we omit the details.

4.1 Degenerate counter-example

Let $0 < \gamma < 1$. Then, the matrices

$$\alpha^0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 0 \\ & & & \gamma \end{pmatrix}, \quad \alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}, \quad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}.$$

satisfy condition (3.2) for $N = 4$ and $m = 2$. Hence, we know from Theorem 2.7 of Duffie et al. (2003) that there exists an affine diffusion on $\mathbb{R}_+^2 \times \mathbb{R}^2$ with diffusion matrix

$$\alpha(x) = \alpha^0 + x_1 \alpha^1 + x_2 \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 1 + x_1 + x_2 & x_2 \\ & & & \gamma + x_1 + x_2 \end{pmatrix}.$$

For instance, $\alpha(x)$ can be written as $\sigma(x)\sigma^T(x)$ for

$$\sigma(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \sqrt{\gamma+x_1} & 0 & \sqrt{x_2} & \sqrt{1-\gamma} \\ 0 & \sqrt{\gamma+x_1} & \sqrt{x_2} & 0 \end{pmatrix}.$$

The corresponding diffusion process X is of the form

$$\begin{aligned} dX_1 &= (b_1 + \langle \beta_1, X \rangle) dt \\ dX_2 &= (b_2 + \langle \beta_2, X \rangle) dt \\ dX_3 &= (b_3 + \langle \beta_3, X \rangle) dt + \sqrt{\gamma + X_1} dW_1 + \sqrt{X_2} dW_3 + \sqrt{1-\gamma} dW_4 \\ dX_4 &= (b_4 + \langle \beta_4, X \rangle) dt + \sqrt{\gamma + X_1} dW_2 + \sqrt{X_2} dW_3. \end{aligned}$$

Since X lives on $\mathbb{R}_+^2 \times \mathbb{R}^2$, $\langle \beta_1, x \rangle$ and $\langle \beta_2, x \rangle$ must not depend on x_3 and x_4 . Hence, X_1 and X_2 are deterministic. Still, X is of the form (9) in Dai and Singleton (2000) because $\sigma(x)$ can be written as $\sigma(x) = \Sigma \sqrt{S(x)}$ for

$$\Sigma = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad S(x) = \begin{pmatrix} \gamma + x_1 & 0 & 0 & 0 \\ 0 & \gamma + x_1 & 0 & 0 \\ 0 & 0 & x_2 & 0 \\ 0 & 0 & 0 & 1 - \gamma \end{pmatrix}.$$

Now, assume that there exists a regular 4×4 -matrix Λ such that $\Lambda \alpha(x) \Lambda^T$ is diagonal for all $x \in \mathbb{R}_+^2 \times \mathbb{R}^2$. Then, by (3.6), there exist four linearly independent vectors f^1, \dots, f^4 in \mathbb{R}^4 such that

$$(4.1) \quad f_3^k f_3^l + \gamma f_4^k f_4^l = 0$$

$$(4.2) \quad f_3^k f_3^l + f_4^k f_4^l = 0$$

$$(4.3) \quad (f_3^k + f_4^k)(f_3^l + f_4^l) = 0.$$

for all $k \neq l$. From (4.1) and (4.2) we get

$$(4.4) \quad f_3^k f_3^l = f_4^k f_4^l = 0.$$

By linear independence there can be at most two vectors among f^1, \dots, f^4 with $f_3^k = f_4^k = 0$. We may assume that f^3 and f^4 are not of this form. We then deduce from (4.4) that either

$$f_3^3 = 0 \text{ and } f_4^4 = 0 \quad \text{or} \quad f_3^4 = 0 \text{ and } f_4^3 = 0.$$

The first case together with (4.3) implies $f_4^3 f_3^4 = 0$, the second one $f_3^3 f_4^4 = 0$. Both contradict the assumption of linear independence of f_1, \dots, f_4 . This shows that $\alpha(x)$ cannot be diagonalized by regular affine transformation.

4.2 Non-degenerate counter-example

For $0 < \gamma < 1$, the matrices

$$\alpha^0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & & \\ & & & \gamma \end{pmatrix}, \quad \alpha^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}, \quad \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & \\ & 1 & 1 & \\ & & & 1 \end{pmatrix}.$$

satisfy condition (3.2). By Duffie et al. (2003), there exists an affine diffusion on $\mathbb{R}_+^2 \times \mathbb{R}^2$ with diffusion matrix

$$\alpha(x) = \alpha^0 + x_1\alpha^1 + x_2\alpha^2 = \begin{pmatrix} x_1 & 0 & 0 & 0 \\ & x_2 & 0 & 0 \\ & & 1 + x_1 + x_2 & x_2 \\ & & & \gamma + x_1 + x_2 \end{pmatrix}.$$

$\alpha(x)$ has rank 4 in the interior of $\mathbb{R}_+^2 \times \mathbb{R}^2$, and there exists no 4×4 -matrix Λ such that $\Lambda\alpha(x)\Lambda^T$ is diagonal for all $x \in \mathbb{R}_+^2 \times \mathbb{R}^2$. This can be proved along the lines of Subsection 4.1. But the argument is much longer and not given here. Alternatively, non-diagonalizability can be checked with software like Mathematica that allows for symbolic calculations.

We point out that now, $\sigma(x)$ cannot be of the form $\Sigma\sqrt{S(x)}$ for a 4×4 -matrix Σ and a diagonal 4×4 -matrix $S(x)$ because if it were, then Σ would have to be regular and

$$\Lambda\alpha\Lambda^T = \Lambda\Sigma S(x)\Sigma^T\Lambda^T$$

would be diagonal for $\Lambda = \Sigma^{-1}$.

5 Conclusion

We have demonstrated that for all $0 \leq m \leq N$ with $m \leq 1$ or $m \geq N - 1$, any affine diffusion process on $\mathbb{R}_+^m \times \mathbb{R}^{N-m}$ can be transformed into one on the same state space with diagonal diffusion matrix by way of a regular affine transformation. This shows that in this case, the Dai–Singleton specification covers all affine diffusion term structure models with state spaces of the form $\Lambda(\mathbb{R}_+^m \times \mathbb{R}^{N-m}) + \lambda$. However, we also showed that there exist affine diffusion processes on $\mathbb{R}_+^2 \times \mathbb{R}^2$ with non-diagonalizable diffusion matrix. Hence, for ATSMs with four risk factors or more, the assumption of instantaneously uncorrelated state variables may result in an unnecessary loss of generality with possible consequences of poor fit to data or non-optimal model selection.

A study of the potential of affine diffusions with non-affine state spaces or the practical benefits from the use of affine diffusions with more than three risk factors and non-diagonalizable diffusion matrix are beyond the scope of this paper and left for future research.

References

- Balduzzi, P., Das, S., Foresi, S., Sundaram, R. (1996). A Simple Approach to Three Factor Affine Term Structure Models. *Journal of Fixed Income* 6, 43–53.
- Buraschi, A., Porchia, P., Trojani, F. (2006). Correlation Risk and Optimal Portfolio Choice. AFA 2008 New Orleans Meetings Paper. Available at SSRN: <http://ssrn.com/abstract=908664>.
- Chen, L. (1996). Stochastic Mean and Stochastic Volatility – A Three-Factor Model of the Term Structure of Interest Rates and its Applications to the Pricing of Interest Rate Derivatives. *Financial Markets, Institutions, and Instruments* 5, 1–88.
- Cheridito, P., Filipović, D., Kimmel, R. L. (2007). Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics* 83(1), 123–170.

- Collin-Dufresne, P., Goldstein, R.S., Jones, C.S. (2008). Identification of Maximal Affine Term Structure Models. *Journal of Finance* 63(2), 743–795.
- Cox, J.C., Ingersoll, J.E., Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* 53(2), 385–408.
- Da Fonseca, J., Grasselli, M., Tebaldi, C. (2008). Option pricing when correlations are stochastic: an analytical framework. *Forthcoming in Review of Derivatives Research*.
- Dai, Q., Singleton, K. J. (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance* 55(5), 1943–1978.
- Duffee, G. R. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance* 57(1), 405–443.
- Duffie, D., Filipović, D., Schachermayer, W. (2003). Affine Processes and Applications in Finance. *Annals of Applied Probability* 13(3), 984–1053.
- Duffie, D., Kan, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6(4), 379–406.
- Duffie, D., Pan, J., Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68(6), 1343–1376.
- Egorov, A., Li, H., Ng, D. (2008). A Tale of Two Yield Curves: Modeling the Joint Term Structure of Dollar and Euro Interest Rates. *Forthcoming in Journal of Econometrics*.
- Gelfand, I. M. (1961). *Lectures on Linear Algebra*. Translated by A. Shenitzer. Dover Publications, Inc. New York.
- Golub, G. H., Van Loan, C. F. (1996). *Matrix Computations*. Third Edition. John Hopkins University Press.
- Gourieroux, C., Sufana, R. (2006) A Classification of Two-Factor Affine Diffusion Term Structure Models. *Journal of Financial Econometrics* 4(1), 31–52.
- Heston, S. (1993). A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bonds and Currency Options. *Review of Financial Studies* 6(2), 327–343.
- Hull, J., White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance* 42(2), 281–300.
- Joslin, S. (2006). Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities? *Working Paper*.
- Kwon, K. (2007). On the Equivalence of a Class of Affine Term Structure Models. *Forthcoming in Annals of Finance*.
- Thompson, S. (2008). Identifying Term Structure Volatility from the LIBOR-Swap Curve. *Review of Financial Studies* 21(2), 819–854.
- Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5(2), 177–188.