

## Linear-Quadratic Jump-Diffusion Modeling

Peng Cheng <sup>a</sup> and Olivier Scaillet <sup>b 1</sup>

<sup>a</sup> Barclays Capital and Swiss Finance Institute, 5 The North Colonnade, Canary Wharf, London E14 4BB, the United Kingdom. PC@fame.ch

<sup>b</sup> HEC Genève and Swiss Finance Institute, Université de Genève, 102 Bd. Carl Vogt, CH - 1211, Genève 4, Switzerland. olivier.scaillet@hec.unige.ch

First version: November, 2002

This version: September, 2006

### Abstract

We aim at accommodating the existing affine jump-diffusion and quadratic models under the same roof, namely the linear-quadratic jump-diffusion (LQJD) class. We give a complete characterization of the dynamics of this class by stating explicitly the structural constraints, as well as the admissibility conditions. This allows us to carry out a specification analysis for the 3-factor LQJD models. We compute the standard transform of the state vector relevant to asset pricing up to a system of ordinary differential equations. We show that the LQJD class can be embedded into the affine class through use of an augmented state vector. This establishes a one-to-one equivalence relationship between both classes in terms of transform analysis.

KEYWORDS: Linear-quadratic models, affine models, jump-diffusions, generalized Fourier transform, option pricing.

JEL CLASSIFICATION: G12.

---

<sup>1</sup>We would like to thank the Editor, Dilip Madan, and the two referees for constructive criticism and comments, which have led to a substantial improvement over the previous version. We also thank P. Collin-Dufresne, J. Detemple, G. Dhaene, D. Duffie, J.D. Fermanian, S. Galluccio, C. Gouriéroux, A. Kaul, M. Musiela, and R. Stulz for many stimulating discussions. We have received fruitful comments from participants at the FAME doctoral workshop, AFFI meeting 2003, ESEM 2003, EFA 2003, and finance seminars at Imperial College, BNP Paribas, Zürich, Lugano, and CREST (Paris). The authors have received support by the Swiss National Science Foundation through the National Center of Competence Research (NCCR): Financial Valuation and Risk Management. The first author also receives a Doctoral Research Grant from International Centre FAME for completing the project. Part of this research was done when the second author was visiting THEMA and IRES.

Downloadable at

[http://www.hec.unige.ch/professeurs/SCAILLET\\_Olivier/pages\\_web/Home\\_Page.htm](http://www.hec.unige.ch/professeurs/SCAILLET_Olivier/pages_web/Home_Page.htm)

## 1 Introduction

Current research using jump-diffusion processes relies mostly on two classes of models: the affine jump-diffusion (AJD) class in the sense of, e.g., Duffie and Kan (1996) and Duffie, Pan, and Singleton (2000), and the quadratic Gaussian (QG) class in the sense of, e.g., Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2002). The popularity of both classes rests not only in their modeling flexibility, but also in their technical tractability in deriving the standard transform<sup>2</sup> of the state vector, defined in Duffie, Pan, and Singleton (2000), for option and bond pricing. Filipović (2002) proves that the maximal consistent order of a separable, polynomial model is two. Naturally, one would wonder whether the AJD and the QG models exhaust the set of models with tractable solutions for the transform, or a maximal flexible, polynomial model is yet to be found.

We find the existing AJD and QG classes, as they are currently specified, both belong to a more general, quadratic framework, which will be established in this paper. Within this framework, the state vector is constructed from a jump-diffusion vector, which is restricted to linear form only, and a pure diffusion vector, which enters both the linear and the quadratic forms. We detail the minimal sufficient conditions for an admissible state vector, as well as the structural constraints for obtaining the standard transform as an exponential quadratic function. We call this structure *linear-quadratic* (LQ) to reflect its construction, and such a process *linear-quadratic jump-diffusion* (LQJD). By Filipović (2002), the LQJD class is, in fact, the maximal flexible dynamic structure for a separable, quadratic model with tractable transforms of the underlying process.

The LQJD framework, as a generalization of the AJD and QG classes, joins their modeling flexibility together. An AJD model, while capable of capturing jumps, is inherently linear and may not capture nonlinearity in the underlying process adequately. See Ait-Sahalia (1996), Ahn and Gao (1999), and Dai and Singleton (2000) for evidence on nonlinearity in interest rates and swap yield curves, respectively. Moreover, Dai and Singleton (2000) and Backus, Foresi, Mozumda, and Wu (2001) point out that the fitting performance of an AJD model can be significantly improved only at the cost of introducing negative correlations among the state variables, and thus losing positivity of the underlying, e.g., interest rates. The QG class seems to be a neat solution to these problems. Examples can be found in Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003) on the term structure of interest rates, and Leippold and Wu (2006) on a multi-currency framework. However, a consistent QG model does not permit jumps (Chen, Bayraktar, and Poor (2005)). An LQJD model, in

---

<sup>2</sup>Duffie, Pan, and Singleton (2000) also define the extended transform of the state vector. The solution procedure, however, is the same as that of the standard transform, and will not be discussed in this paper.

contrast, would accommodate these important features, i.e., nonlinearity, jumps, positivity of the underlying process, under the same roof.

An additional, striking feature of the LQJD state vector, in comparison to the pure AJD and QG state vectors, is its nonlinear specifications for the drift, the diffusion, and the market price of risk. For example, let  $\mu^{\mathbb{P}}$  and  $\mu^{\mathbb{Q}}$  be the drift vectors under the two equivalent probability measures  $\mathbb{P}$  and  $\mathbb{Q}$ . If the underlying process is either AJD or QG under both measures, the quantity  $\mu^{\mathbb{P}} - \mu^{\mathbb{Q}}$  would only admit an affine specification. See Dai and Singleton (2003). Under the LQJD framework, however,  $\mu^{\mathbb{P}} - \mu^{\mathbb{Q}}$  can be LQ in the state vector. This is a marked improvement over the AJD and QG models. In particular, an investigation of nonlinearity in the market price of risk as suggested by Pan (2002), which is not possible in the AJD or QG framework, is now feasible.

Despite the additional modeling flexibility, the LQJD class is still as tractable as the QG and the AJD classes. Like Duffie, Pan, and Singleton (2000), we derive the transforms of the LQJD process up to a system of ordinary differential equations (ODE), which we discover to be a system of *non-symmetric Riccati differential equations* (RDEs). Results from the literature of RDEs (i.e., Radon's lemma) suggest that the initial value problem for a RDE system that we have to solve in the LQJD setting is (locally) equivalent to the one for a linear system. Hence, there is a standard routine for solving the differential system.

A further finding, resulting from the solution scheme of the LQJD standard transform, is the valuation equivalence between linear-quadratic and affine models. Intuitively, by a simple change-of-variables technique, a quadratic form  $aX^2 + bX + c$  in  $X$  can be replaced by  $aZ + bX + c$ , which is now affine in an *augmented, constrained* vector  $(X, Z)$ . We prove that the system of ODEs obtained for a LQJD process is identical to that for its augmented affine version, hence an LQJD model can in fact be rewritten as a constrained AJD model. This, together with the fact that the AJD class is nested in the LQJD class, establishes a *one-to-one* relationship between the two classes of models in terms of their transforms. In other words, the set of LQJD models that is absolutely distinct from (augmented) AJD models is *empty* when considering asset pricing by transform analysis.

However, while the augmented LQ model is affine in the sense of Duffie, Pan, and Singleton (2000), it may not fit into the affine specifications of Dai and Singleton (2000). We show, with an example, that the causes of the problem are the presence of the quadratic factors and their deterministic relationship with the augmented factor. Therefore, the equivalence between LQJD and AJD models stated here concerns the valuation side of their transforms only.

The relationship between tractable affine and quadratic models has drawn little research attention until recently. See, for example, Pan and Wu (2005) and Chen, Filipović, and

Poor (2004), where special quadratic term structure models with affine equivalence are given. See also Gouriéroux and Sufana (2003), who are aware of such equivalence in a study of Wishart autoregressive processes in discrete time. However, without a unified ground on which this problem can be examined, the affine and the quadratic classes are easily, and indeed commonly taken as two separate sets. The LQJD framework provides such a unified ground, and we consider our finding a valuable theoretical contribution to the debate of affine versus quadratic classes.

The rewriting of an LQJD to its affine equivalence also permits analyses done previously for the AJD class to be carried out for the LQJD class. One example is Collin-Dufresne, Goldstein and Jones (2003), in which a three-factor maximal affine term structure model with stochastic volatility is studied. Their technique of rotating latent state variables into observables is similarly applicable in the LQJD setting. We can also conduct a specification analysis for an LQJD model in a similar manner to that in Dai and Singleton (2000). In particular, in the paper we give the sufficient conditions for an LQJD specification to be admissible, and provide the maximal flexible specifications for three-factor LQJD models.

To the best of our knowledge, the first piece of work that coins the term LQJD is Piazzesi (2001)<sup>3</sup> in interest rate modeling. The paper initiates the issue of including jumps with quadratic arrival intensity for pricing bonds in the quadratic class, and fulfils the task by constructing the state vector from two parts: one being pure Gaussian-Markov without jumps, and the other being square-root process with jumps. The drift and the covariance matrices are still affine in the state vector. Clearly, the state vector in Piazzesi (2001) is an AJD vector and a QG vector knitted together, and the possibility of integrating nonlinearity in the drift of the linear vector is not dealt with. We believe that this specification of the state vector is not rich enough to render the model an extension fully capable of integrating the modeling strengths of both classes in one.

We are also aware that Liu (2006) proposes some state vector dynamics fairly akin to ours. While Liu's analysis is oriented toward solving optimal dynamic portfolio selection problems, ours aims at pricing issues via transform analysis. Consequently, we are able to prove that the AJD class is more general than conventionally thought, and is actually equivalent to the LQJD class. In contrast, Liu (2006) still treats the affine class as a non-trivial subset of the LQJD class. Furthermore, the specification of Liu (2006) (his equation (11)) is not fully compatible with our LQJD setting. In particular, the third inequality of his equation (11) does not allow for a proper solution for the transforms.

The rest of the paper is structured as follows. Section 2 describes the specification of the LQJD framework and the structural constraints. Section 3 discusses the admissibility

---

<sup>3</sup>A published version of this paper is Piazzesi (2004), where the term LQJD is dropped.

conditions and carries out a specification analysis of the LQJD model. Section 4 computes the standard transform and shows the link and the equivalence between the LQJD and AJD classes in terms of their transforms. A discussion of option pricing via transform analysis in the LQJD setting is also included. Section 5 concludes. Technical details are collected in the Appendices. Further details, explanations and references, as well as a numerical application of LQJD modeling to stochastic volatility, can also be found in an extended version of this paper, Cheng and Scaillet (2002).

## 2 The LQJD settings

In this section we provide a general definition of the LQJD state vector and its standard transform. We write down the partial-integro differential equation (PIDE) that must be satisfied by the standard transform, and discuss the structural constraints on the state vector such that the PIDE can be solved up to the solution of a system of ODEs. It turns out that a LQJD process admits a nonlinear drift and a nonlinear diffusion, which is a marked improvement over the pure AJD and QG classes.

The complete solution to the standard transform, however, will be presented after a discussion of admissibility conditions in Section 3.

### 2.1 The LQJD state vector and its standard transform

The  $n$ -dimension, *càdlàg* state vector  $X_t$  is drawn from some state space  $\mathbb{D}$ , and follows the stochastic differential equation (SDE):

$$dX_t = \mu(X_t, t) dt + \sigma(X_t, t) dW_t + dJ_t, \quad (2.1)$$

where:

- (i)  $W_t$  is a standard  $n^\circ$ -dimension Brownian motion vector;
- (ii)  $J_t$  is a pure jump process with independent increments  $dJ_t$ , whose size distributions and arrival intensities are given by  $\Pi(dy, t)$  and  $\lambda(X_t, t)$ , respectively;
- (iii)  $(\mathcal{E}, \mathcal{F}, \mathbb{P})$  is the usual probability space with  $(W, J)$ -augmented filtration  $(\mathcal{F}_t)_{t \geq 0}$ , meaning that:
  - $\mathcal{F}_0$  contains all the  $\mathbb{P}$ -null sets of  $\mathcal{F}$ ;
  - the filtration  $\mathcal{F}$  is right-continuous.

See, e.g., Protter (2004).

For identification, we require that  $n \geq n^\circ$ . Moreover, we assume that  $\mu$ ,  $\sigma$ ,  $\Pi$ , and  $\lambda$  satisfy the regularity conditions that guarantee a unique solution to (2.1) for every  $X_0$ .

In the LQJD setting, the drift matrix  $\mu(X_t, t)$ , the covariance matrix  $\Omega(X_t, t) = \sigma(X_t, t)\sigma(X_t, t)^\top$ , and the jump arrival intensity  $\lambda(X_t, t)$  are all LQ in the state vector by construction. That is, each entry of  $\mu$ ,  $\Omega$ , and  $\lambda$  is of the following form:

$$\varkappa(X_t, t) = \frac{1}{2}X^\top \Lambda_\varkappa(t) X_t + b_\varkappa(t)^\top X_t + c_\varkappa(t), \quad (2.2)$$

where the superscript  $\top$  denotes matrix transpose, and the coefficient matrix  $\Lambda_\varkappa$  is block diagonal with the following representation:

$$\Lambda_\varkappa = \begin{pmatrix} A_\varkappa & 0 \\ 0 & 0 \end{pmatrix}, \quad (2.3)$$

with  $A_\varkappa$  being symmetric. Note that the dimension of  $A_\varkappa$  depends on the number of state variables that enters the quadratic forms of  $\mu$ ,  $\Omega$ , and  $\lambda$ . We can now partition the state vector  $X_t$  and the coefficient vector  $b_\varkappa$  accordingly as:

$$X = \begin{pmatrix} \bar{X}_t \\ \underline{X}_t \end{pmatrix}, \quad b_\varkappa = \begin{pmatrix} k_\varkappa \\ l_\varkappa \end{pmatrix}, \quad (2.4)$$

and rewrite (2.2) as:

$$\varkappa(X, t) = \underbrace{l_\varkappa(t)^\top \underline{X}_t}_{\text{(linear part)}} + \underbrace{\frac{1}{2}\bar{X}_t^\top A_\varkappa(t)\bar{X}_t + k_\varkappa(t)^\top \bar{X}_t + c_\varkappa(t)}_{\text{(quadratic part)}}, \quad (2.5)$$

hence the name ‘linear-quadratic’.

The standard transform of a state vector  $X_t$  is defined, in general, as:

$$\phi(g; X_t, t, T) = E_t \left[ \exp \left( - \int_t^T R(X_s, s) ds \right) e^{g(X_T, T)} \right], \quad (2.6)$$

where  $E_t[\cdot] = E[\cdot | \mathcal{F}_t]$  is the expectation conditional on information up to  $t$ :  $\mathcal{F}_t$ ,  $t \leq T < \infty$ . If the process  $\Phi(X_t, t)$ , defined as:

$$\Phi(X_t, t) = \exp \left( - \int_0^t R(X_s, s) ds \right) e^{g(X_t, t)}, \quad (2.7)$$

is a martingale, then the standard transform can be solved as:

$$\begin{aligned}
\phi(g; X_t, t, T) &= e^{g(X_t, t)} \frac{E_t[\Phi(X_T, T)]}{\Phi(X_t, t)} \\
&= e^{g(X_t, t)}.
\end{aligned} \tag{2.8}$$

This implies that  $g(X_t, t)$  satisfies the following PIDE:

**Lemma 1** *If the technical integrability conditions hold and the function  $g(x, t)$  satisfies the following PIDE (the Cauchy problem):*

$$R = \frac{\partial g}{\partial t} + \mu^\top \frac{\partial g}{\partial x} + \frac{1}{2} \text{tr} \left[ \left( \frac{\partial^2 g}{\partial x^2} + \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \right) \Omega \right] + \lambda(\theta - 1), \tag{2.9}$$

where  $\theta = \int_{\mathbb{D}} e^{g(x+y, t) - g(x, t)} \Pi(x, dy, t)$ , then  $\Phi_t$  is a martingale.

**Proof.** See Appendix A. ■

Given the result in Filipović (2002) that the maximal consistent order of a separable, polynomial model is two, we take both  $g(X_t, t)$  and  $R(X_t, t)$  as LQ functions in  $X_t$ .

We use  $\mathbb{LQ}_m^q(n)$  to denote an  $n$ -factor LQJD model where the first  $q$  members of the state vector  $X_t$  appear at least once in a quadratic form (i.e.,  $\bar{X}_t$  has dimension  $q$ ), and  $m$  members of  $\underline{X}_t$  appear in either  $\lambda$  or  $\Omega$ . Without loss of generality, we assume these  $m$  factors are the first  $m$  members of  $\underline{X}_t$ . Hence, for an  $\mathbb{LQ}_m^q(n)$  model,

$$g(X, t) = l(t)^\top \underline{X}_t + \frac{1}{2} \bar{X}_t^\top A(t) \bar{X}_t + k(t)^\top \bar{X}_t + c(t),$$

where  $A$  is symmetric with rank  $q$ . Clearly, the LQJD model reduces to a QG model when  $q = n$ , and to an AJD model (such as the  $\mathbb{A}_m(n)$  model of Dai and Singleton (2000) with jumps) when  $q = 0$ . By reference to these two classes, the variables  $\bar{X}$  and  $\underline{X}$  are named quadratic and affine vectors, respectively.

## 2.2 The structural constraints

For both  $g(X_t, t)$  and  $R(X_t, t)$  to be LQ in  $X_t$ , an  $\mathbb{LQ}_m^q(n)$  model must satisfy the following structural constraints (See Appendix B.1 for more details):

[SC1] The first  $q$  entries of the jump component  $J$  are zeros.

That is, jumps are restricted to the affine vector  $\underline{X}$  only. This comes from  $\lambda$ ,  $g$  and  $R$  being all LQ function by definition, which, together with (2.9), implies that  $\theta$  is in fact

independent of  $x$ . The minimal constraint for this to hold true is that the first  $q$  entries of the jump size distribution is zero.

[SC2] The drift matrix of  $X_t$  is:

$$\mu(X_t, t) = \begin{pmatrix} \bar{\mu}(\bar{X}_t, t) \\ \underline{\mu}(\bar{X}_t, \underline{X}_t, t) \end{pmatrix}_{n \times 1}. \quad (2.10)$$

In particular:

- (a) The drift matrix of the quadratic vector  $\bar{X}$ ,  $\bar{\mu}(\bar{X}_t, t)$ , is only an affine function of  $\bar{X}$ ;
- (b) The drift matrix of the affine vector  $\underline{X}$ ,  $\underline{\mu}(\bar{X}_t, \underline{X}_t, t)$ , is LQ in  $\underline{X}$  (i.e., affine in  $\underline{X}$  and quadratic in  $\bar{X}$ ).

The restriction that  $\bar{\mu}(\bar{X}_t, t)$  is an affine function of  $\bar{X}$  only could be justified heuristically as follows. By definition,  $\underline{X}$  should remain in linear terms only. If  $\bar{\mu}$  were an affine function of  $\underline{X}$  as well, it would bring powers of  $\underline{X}$  to two as the quadratic vector  $\bar{X}$  pass through the quadratic terms. Constraint [SC2](a) looks quite restrictive, because it does not allow linking members of  $\underline{X}$  with  $\bar{X}$  through the drift of  $\bar{X}$ . For instance, if  $\underline{X}$  is the logarithm of stock price and  $\bar{X}$  is the state vector describing the dynamics of stock price volatility, it is indeed desirable to let the logarithm of the stock price  $\underline{X}$  play a ‘feedback’ role on the volatility state vector  $\bar{X}$ . Constraint [SC2](a) rules out the possibility of having this type of ‘feedback’ effect through the drift of  $\bar{X}$  in the LQJD setting. However, we can still model this effect through the correlation structure of  $\underline{X}$ .

In comparison to an AJD process, the affine vector  $\underline{X}$  in the LQJD framework has a nonlinear component that is generated by a quadratic form of the quadratic vector  $\bar{X}$ . As already mentioned, we believe that the incorporated nonlinearity is a marked improvement from the pure AJD and QG classes. In particular, the quantity  $\mu^{\mathbb{P}} - \mu^{\mathbb{Q}}$ , where  $\mathbb{P}$  and  $\mathbb{Q}$  are two equivalent probability measures, is LQ in the state vector in the LQJD setting. In contrast, as pointed out by Dai and Singleton (2003), both the AJD and QG classes admit only affine specifications for  $\mu^{\mathbb{P}} - \mu^{\mathbb{Q}}$ . Apparently, such restriction prevents an investigation of nonlinear specification, as suggested in Pan (2002).

[SC3] The covariance matrix of  $X_t$  is:

$$\Omega(X_t, t) = \begin{pmatrix} \bar{\Omega}(t) & \tilde{\Omega}(\bar{X}_t, t) \\ \tilde{\Omega}(\bar{X}_t, t)^\top & \underline{\Omega}(\bar{X}_t, \underline{X}_t, t) \end{pmatrix}_{n \times n}. \quad (2.11)$$

In particular:

- (a) the covariance matrix of the quadratic vector  $\bar{\mathbf{X}}, \bar{\Omega}(t)$ , is deterministic in  $t$ ;
- (b) the covariance matrix of the affine vector  $\mathbf{X}, \underline{\Omega}(\bar{\mathbf{X}}_t, \mathbf{X}_t, t)$ , is LQ in  $\mathbf{X}$  (i.e., affine in  $\underline{\mathbf{X}}$  and quadratic in  $\bar{\mathbf{X}}$ );
- (c) from (a) and (b), the covariance matrix between  $\underline{\mathbf{X}}$  and  $\bar{\mathbf{X}}, \tilde{\Omega}(\bar{\mathbf{X}}_t, t)$ , is affine in  $\bar{\mathbf{X}}$  only.

Constraint [SC3](b) differs from the usual practice in affine modeling, which restricts  $\sigma$  to be square-root affine in the state vector. See, for example, Dai and Singleton (2000). However, it is the covariance matrix  $\Omega$ , not  $\sigma$  itself, that matters for solving the PIDE (2.9). See Duffie, Pan and Singleton (2000). Moreover, restrictions on  $\sigma$  would in fact exclude an important group of models, which distinguish themselves from the square-root affine models by carrying *sign* information in  $\sigma$ .

Also note that option pricing models based on a quadratic specification for  $\sigma$  of the stock price as in Rady (1997) are not embedded in the LQJD class. This is because the logarithm of the stock price is usually a member of the affine vector  $\underline{\mathbf{X}}$ , and as such it is excluded from any quadratic forms in the LQJD setting by construction.

There are only a small number of LQJD models existing in the current literature that are distinct from the pure AJD or QG models. The more distant ones include Stein and Stein (1991), which is  $\mathbb{LQ}_0^1(2)$  with  $\tilde{\Omega} = 0$  and  $\lambda = 0$  (i.e., no jumps), and Schöbel and Zhu (1999), which is an extension of Stein and Stein (1991) with  $\tilde{\Omega} \neq 0$ . The more recent ones include Piazzesi (2001) and (2004), which are  $\mathbb{LQ}_1^1(4)$  with  $\underline{\mu}$  being only affine in the state vector. Santa-Clara and Yan (2005) propose an  $\mathbb{LQ}_0^2(3)$  model for modeling stock market index with stochastic volatility and jumps. It corresponds, in our notation, to  $\bar{\mathbf{X}}_t = \begin{pmatrix} X_{1t} & X_{2t} \end{pmatrix}^\top$ , which models the instantaneous volatility ( $V_t = X_{1t}^2$ ) and the jump intensity ( $\lambda_t = X_{2t}^2$ ), together with  $\mathbf{X}_t = (X_{3t})$ , which is the logarithm of the stock price. The drift and the covariance matrices for this  $\mathbb{LQ}_0^2(3)$  model are, respectively

$$\mu_t = \begin{pmatrix} k_{\mu_1} X_{1t} + c_{\mu_1} \\ k_{\mu_2} X_{2t} + c_{\mu_2} \\ \frac{1}{2} \bar{\mathbf{X}}_t^\top A_{\mu_3} \bar{\mathbf{X}}_t + c_{\mu_3} \end{pmatrix}, \quad \Omega_t = \begin{pmatrix} c_{11} & c_{12} & k_{13} X_{1t} \\ . & c_{22} & k_{23} X_{1t} \\ . & . & X_{1t}^2 \end{pmatrix}, \quad (2.12)$$

where  $k$  and  $c$ 's are constants, and  $A_{\mu_3}$  is a constant,  $2 \times 2$  diagonal matrix.

We will use the Santa-Clara and Yan (2005) model as a running example, and show that it is not the maximal flexible model in the  $\mathbb{LQ}_0^2(3)$  class. In the next section, we will discuss the LQJD admissibility conditions and the relevant specification analysis, which address the issue of maximal flexible models.

### 3 Admissibility conditions and specification analysis

In addition to the structural constraints which are necessary conditions for obtaining a tractable transform, we also need admissibility conditions that guarantee a positive jump intensity  $\lambda$  and a positive semi-definite covariance matrix  $\Omega$ , such that a solution to SDE (2.1) exists and is unique. The structural constraints and the admissibility conditions allows us to extend the classification scheme and specification analysis of Dai and Singleton (2000) to the LQJD setting. Specifically, an  $n$ -factor LQJD model with  $q$  quadratic factors can be classified into  $n - q + 1$  subfamilies (indexed by  $m$ ), and the maximal flexible  $\mathbb{LQ}_m^q(n)$  model can be identified accordingly. We illustrate such analysis with three-factor LQJD models.

In a pure affine diffusion setting, Gouriéroux and Sufana (2006) points out that the analysis of Dai and Singleton (2000) provides sufficient but not necessary conditions, and excludes some admissible, non-linear state space. Their analysis, however, involves the study of the dynamics of  $\Omega$  and its determinant on the boundary of the state space, which will become extremely involved when the number of factors increases. We will still follow the recipe provided by Dai and Singleton (2000), and leave the Gouriéroux and Sufana (2006) type of analysis for  $n$ -factor models in a jump-diffusion setting for future research. However, our results on the equivalence relationship between LQJD and AJD classes, and in particular the rewriting of an LQJD model as its AJD equivalence (see Section 4.2 below), actually provides examples mentioned in Gouriéroux and Sufana (2006), which do not fit into the standard Dai and Singleton (2000) classification.

#### 3.1 Admissibility conditions

For an LQJD process to be admissible, its jump intensity  $\lambda$  should be positive and its covariance matrix  $\Omega$  should be positive semi-definite. When  $q = n$ , an  $\mathbb{LQ}_m^q(n)$  model reduces to a QG model and no admissibility conditions are needed. When  $q = 0$ , the LQJD model becomes an AJD model, and the admissibility conditions, as well as the corresponding specification analysis, have been detailed in Dai and Singleton (2000). For an  $\mathbb{LQ}_m^q(n)$  model with  $0 < q < n$ , we have:

[AC1] Up to invariant transforms defined in Dai and Singleton (2000), the covariance matrix of the first  $m$  factors of the affine vector  $\underline{X}_t$  is diagonal, with the  $i^{th}$  diagonal term linear in  $X_{q+i,t}$ , the  $i^{th}$  factor of  $\underline{X}_t$

The invariant transforms of Dai and Singleton (2000) are applicable to the first  $m$  factors of  $\underline{X}_t$ , because they only appear in affine terms by definition, and their diffusions do not

contain any quadratic terms. [AC1] ensures that when  $X_{q+i,t}$ ,  $i \leq m$ , reaches zero, its diffusion also becomes zero and  $X_{q+i,t}$  is locally deterministic. In the presentation we have applied the invariant transforms to get rid of the constant term in the diffusion of  $X_{q+i,t}$ , so that we can discuss its drift separately. In effect, imposing [AC1] is equivalent to represent  $X_{q+i,t}$  in a canonical form as discussed in Dai and Singleton (2000).

At zero boundary,  $X_{q+i,t}$  should have a positive drift. Recall that the drift of  $X_{q+i,t}$  can have the following specification:

$$\mu_{q+i}(X, t) = l_{\mu_{q+i}}(t)^\top \underline{X}_t + \underbrace{\frac{1}{2} \bar{X}_t^\top A_{\mu_{q+i}}(t) \bar{X}_t + k_{\mu_{q+i}}(t)^\top \bar{X}_t + c_{\mu_{q+i}}(t)}_{\text{(quadratic part)}}. \quad (3.1)$$

To guarantee a positive  $\mu_{q+i}$  when  $X_{q+i,t}$  is at its zero boundary, we must have:

[AC2] The  $j^{\text{th}}$  entry of  $l_{\mu_{q+i}}$  is zero when  $j > m$ , and non-negative when  $j \leq m$  and  $j \neq i$ .

[AC3] The quadratic part of (3.1) is non-negative, i.e.,

- (a)  $A_{\mu_{q+i}}$  is positive semi-definite;
- (b)  $k_{\mu_{q+i}}$  belongs to the column space of  $A_{\mu_{q+i}}$ ;
- (c)  $c_{\mu_{q+i}} \geq \frac{1}{2} k_{\mu_{q+i}}^\top A_{\mu_{q+i}}^+ k_{\mu_{q+i}}$ , where the superscript  $+$  denotes the Moore-Penrose, or generalized inverse of a matrix (e.g., when a matrix  $A$  is non-singular,  $A^+ = A^{-1}$ ).

The admissibility conditions above closely resembles the ones of an AJD model. In particular, [AC1] corresponds to conditions **C2** and **C3** in Dai and Singleton (2000), and [AC2] and [AC3] to conditions **C1**, **C4** and **C5**. See Appendix B of Dai and Singleton (2000).

If  $X_{q+i,t}$  has a jump component, its jump size must be non-negative. Specifically:

[AC4] For all  $i \leq m$ , the jump size distribution of  $X_{q+i,t}$  has support on  $\mathbb{R}^+ \cup \{0\}$ .

Finally, we need the volatilities of the last  $n - q - m$  factors of  $\underline{X}_t$  (i.e.,  $X_{q+m+i,t}$  where  $i \leq n - q - m$ ), as well as members of the jump intensity vector  $\lambda$ , to be non-negative. Let  $\vartheta$  denote either  $\Omega_{q+m+i, q+m+i}$  where  $i \leq n - q - m$ , or a member of  $\lambda$ . Note that  $\vartheta$  is LQ in the state vector and can be represented as:

$$\vartheta(X, t) = l_\vartheta(t)^\top \underline{X}_t + \underbrace{\frac{1}{2} \bar{X}_t^\top A_\vartheta(t) \bar{X}_t + k_\vartheta(t)^\top \bar{X}_t + c_\vartheta(t)}_{\text{(quadratic part)}}. \quad (3.2)$$

We have:

[AC5] The  $j^{\text{th}}$  entry of  $l_{\vartheta}$  is zero when  $j > m$  (by definition), and non-negative when  $j \leq m$ .

[AC6] The quadratic part of (3.2) is non-negative.

There is no need to discuss the specifications of the off-diagonal terms of  $\Omega$ , because they can be constructed by applying structural constraint [SC3] and admissibility conditions [AC5] - [AC6].

### 3.2 Specification analysis of three-factor LQJD models

The structural constraints and admissibility conditions allow us to carry out a specification analysis of the LQJD models that identifies the maximal dynamic structure for a given number of factors. We concentrate on the three-factor LQJD models, i.e.,  $\mathbb{LQ}_m^q(3)$ . As mentioned above, the cases of  $q = 0$  and  $q = 3$  have already been studied in the AJD and QG literature, respectively. Hence, we will only consider  $q = 1$  and  $q = 2$ .

From the previous subsection we know that an  $\mathbb{LQ}_m^q(n)$  family can be further classified into  $n - q + 1$  subfamilies. For  $n = 3$  and  $q = 1, 2$ , this results in five subfamilies of LQJD models. For each subfamily, we consider the maximal flexible structure for the drift and the covariance matrices. We will not discuss the structure of the jump intensity matrix, for it very much resembles that of the drift matrix.

#### 3.2.1 The maximal $\mathbb{LQ}_m^1(3)$ models

For an  $\mathbb{LQ}_m^1(3)$  model,  $\bar{X}_t = (X_{1t})$  and  $\mathbf{X}_t = \begin{pmatrix} X_{2t} & X_{3t} \end{pmatrix}^\top$ . The maximal flexible structure for the drift of all  $\mathbb{LQ}_m^1(3)$  models can be represented as:

$$\mu_t = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ \mu_{3t} \end{pmatrix} = \begin{pmatrix} k_{\mu_1} X_{1t} + c_{\mu_1} \\ l_{\mu_2}^\top \mathbf{X}_t + \frac{1}{2} A_{\mu_2} X_{1t}^2 + k_{\mu_2} X_{1t} + c_{\mu_2} \\ l_{\mu_3}^\top \mathbf{X}_t + \frac{1}{2} A_{\mu_3} X_{1t}^2 + k_{\mu_3} X_{1t} + c_{\mu_3} \end{pmatrix}.$$

If  $m = 1$ , then the admissibility conditions [AC2] and [AC3] are binding on  $\mu_{2t}$ , i.e., the second entry of  $l_{\mu_2}$  is zero, and  $c_{\mu_2} \geq \frac{1}{2} k_{\mu_2}^2 A_{\mu_2}^{-1}$ . If  $m = 2$ , then the admissibility conditions [AC2] and [AC3] are binding on both  $\mu_{2t}$  and  $\mu_{3t}$ .

The maximal flexible structure for the diffusion (covariance) matrices of  $\mathbb{LQ}_m^1(3)$  models are, respectively:

-  $\mathbb{LQ}_0^1(3)$ :

There is no binding admissibility conditions on the diffusion of an  $\mathbb{LQ}_0^1(3)$  model. Hence, its diffusion may look like:

$$\sigma_t = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \beta_{21}X_{1t} + \alpha_{21} & \beta_{22}X_{1t} + \alpha_{22} & \beta_{23}X_{1t} + \alpha_{23} \\ \beta_{31}X_{1t} + \alpha_{31} & \beta_{32}X_{1t} + \alpha_{32} & \beta_{33}X_{1t} + \alpha_{33} \end{pmatrix}.$$

The resulting covariance matrix of the  $\mathbb{LQ}_0^1(3)$  would be:

$$\Omega_t = \begin{pmatrix} c_{11} & k_{12}X_{1t} + c_{12} & k_{13}X_{1t} + c_{13} \\ \cdot & \frac{1}{2}A_{22}X_{1t}^2 + k_{22}X_{1t} + c_{22} & \frac{1}{2}A_{23}X_{1t}^2 + k_{23}X_{1t} + c_{23} \\ \cdot & \cdot & \frac{1}{2}A_{33}X_{1t}^2 + k_{33}X_{1t} + c_{33} \end{pmatrix}.$$

-  $\mathbb{LQ}_1^1(3)$ :

The admissibility conditions [AC5] and [AC6] are binding on the diffusion matrix of an  $\mathbb{LQ}_1^1(3)$  model. One possible specification is:

$$\sigma_t = \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \sigma_{22}\sqrt{X_{2t}} & 0 \\ \beta_{31}X_{1t} + \alpha_{31} & \sigma_{32}\sqrt{X_{2t}} & \sqrt{\beta_{33}X_{2t} + \alpha_{33}} \end{pmatrix}$$

The covariance matrix is, in general:

$$\Omega_t = \begin{pmatrix} c_{11} & 0 & k_{13}X_{1t} + c_{13} \\ \cdot & l_{22}X_{2t} & l_{23}X_{2t} \\ \cdot & \cdot & l_{33}X_{2t} + \frac{1}{2}A_{33}X_{1t}^2 + k_{33}X_{1t} + c_{33} \end{pmatrix}.$$

-  $\mathbb{LQ}_2^1(3)$ :

By [AC1], the covariance matrix can only be:

$$\Omega_t = \begin{pmatrix} c_{11} & 0 & 0 \\ \cdot & l_{22}X_{2t} & 0 \\ \cdot & \cdot & l_{33}X_{3t} \end{pmatrix}.$$

This corresponds to the diffusion matrix:

$$\sigma_t = \pm \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \sigma_{22}\sqrt{X_{2t}} & 0 \\ 0 & 0 & \sigma_{33}\sqrt{X_{3t}} \end{pmatrix}.$$

Note that regardless of the sign of  $\sigma_t$ , the covariance is identical, as well as the resulting standard transform.

Similar to the AJD class, as more affine variables enter the specification of the diffusion matrix, the structure of  $\Omega_t$  becomes increasingly restrictive.

### 3.2.2 The maximal $\mathbb{LQ}_m^2(3)$ models

For an  $\mathbb{LQ}_m^2(3)$  model,  $\bar{X}_t = \begin{pmatrix} X_{1t} & X_{2t} \end{pmatrix}^\top$  and  $X_t = (X_{3t})$ . The maximal flexible structure for the drift of all  $\mathbb{LQ}_m^2(3)$  models can be represented as:

$$\mu_t = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \\ \mu_{3t} \end{pmatrix} = \begin{pmatrix} k_{\mu_1}^\top \bar{X}_t + c_{\mu_1} \\ k_{\mu_2}^\top \bar{X}_t + c_{\mu_2} \\ l_{\mu_3} X_t + \frac{1}{2} \bar{X}_t^\top A_{\mu_3} \bar{X}_t + k_{\mu_3}^\top \bar{X}_t + c_{\mu_3} \end{pmatrix}.$$

If  $m = 1$ , the admissibility condition [AC3] is binding on  $\mu_{3t}$ , i.e.,  $c_{\mu_3} \geq \frac{1}{2} k_{\mu_3}^\top A_{\mu_3}^+ k_{\mu_3}$ .

The maximal flexible structure for the diffusion (covariance) matrices of  $\mathbb{LQ}_m^2(3)$  models are, respectively:

-  $\mathbb{LQ}_0^2(3)$ :

There is no binding admissibility conditions on the diffusion of an  $\mathbb{LQ}_0^2(3)$  model. Hence, its diffusion may look like:

$$\sigma_t = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \beta_{31}^\top \bar{X}_t + \alpha_{31} & \beta_{32}^\top \bar{X}_t + \alpha_{32} & \beta_{33}^\top \bar{X}_t + \alpha_{33} \end{pmatrix}.$$

The resulting covariance matrix of the  $\mathbb{LQ}_0^2(3)$  would be:

$$\Omega_t = \begin{pmatrix} c_{11} & c_{12} & k_{13} X_{1t} + c_{13} \\ . & c_{22} & k_{23} X_{1t} + c_{23} \\ . & . & \frac{1}{2} \bar{X}_t^\top A_{33} \bar{X}_t + k_{33}^\top \bar{X}_t + c_{33} \end{pmatrix}.$$

In Section 2 we presented the Santa-Clara and Yan (2005) model (Equation (2.12)), which is an  $\mathbb{LQ}_0^2(3)$ . By comparing the drift and the covariance matrices with the maximal flexible specifications of an  $\mathbb{LQ}_0^2(3)$  model above, we see that the Santa-Clara and Yan (2005) model is not maximal flexible, since it sets the second entry of  $k_{\mu_1}$ ,

the first entry of  $k_{\mu_2}$ ,  $k_{\mu_3}$ ,  $c_{13}$ ,  $c_{23}$ ,  $c_{33}$ , and  $k_{33}$  to zero, and:

$$A_{\mu_3} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_{33} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

-  $\mathbb{LQ}_1^2(3)$ :

The structure of  $\mathbb{LQ}_1^2(3)$  is less flexible than that of  $\mathbb{LQ}_0^2(3)$  in that the admissibility condition [AC1] becomes binding. The maximal flexible diffusion of  $\mathbb{LQ}_1^2(3)$  is:

$$\sigma_t = \pm \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ & & \sigma_{33}\sqrt{X_{3t}} \end{pmatrix},$$

and the corresponding covariance matrix is:

$$\Omega_t = \begin{pmatrix} c_{11} & c_{12} & 0 \\ . & c_{22} & 0 \\ . & . & k_{33}X_{3t} \end{pmatrix}.$$

## 4 The solution to the standard transform and the equivalence between linear-quadratic and affine models

Provided that the structural constraints and the admissibility conditions are satisfied, we can derive the solution of the standard transform as an exponential LQ function. A very important result of this section is the equivalence relationship between the LQJD and the AJD classes in terms of their transforms. That is, an LQJD process can actually be reformulated as a constrained AJD process.

### 4.1 The solution to the standard transform

Recall that the solution to the standard transform is:

$$\phi(g; X_t, t, T) = e^{g(X_t, t)},$$

provided that the LQ function  $g(x, t)$  satisfies the PIDE (2.9) in Lemma 1:

$$R = \frac{\partial g}{\partial t} + \mu^\top \frac{\partial g}{\partial x} + \frac{1}{2} \text{tr} \left[ \left( \frac{\partial^2 g}{\partial x^2} + \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \right) \Omega \right] + \lambda(\theta - 1),$$

where  $\theta(x, t) = \int_{\mathbb{D}} e^{g(x+y, t) - g(x, t)} \Pi(x, dy, t)$ , and  $R$  and  $\lambda$  are both LQ functions, i.e., for  $\varkappa = R$  and  $\lambda$ ,  $\varkappa(x, t) = \frac{1}{2}x^\top \Lambda_\varkappa(t)x + b_\varkappa(t)^\top x + c_\varkappa(t)$ . We may solve the PIDE up to a system of ODEs using the method of undetermined coefficients. To achieve this, first note that the drift  $\mu(x, t)$  and the covariance matrix  $\Omega(x, t)$  of the state vector can be written as:

$$\begin{aligned}\mu &= \frac{1}{2} \left( \mathbf{I}_n \otimes x^\top \right) \mathcal{A}x + \mathcal{B}x + \mathcal{C}, \\ \Omega &= \frac{1}{2} \left( \mathbf{I}_n \otimes x^\top \right) \mathfrak{A}(\mathbf{I}_n \otimes x) + \mathfrak{B}(\mathbf{I}_n \otimes x) + \mathfrak{C},\end{aligned}$$

where  $\otimes$  is the Kronecker product operator,  $\mathbf{I}_n$  is an  $n$ -dimensional identity matrix, and the coefficient matrices  $\begin{pmatrix} \mathcal{A} & \mathcal{B} & \mathcal{C} \end{pmatrix}$  and  $\begin{pmatrix} \mathfrak{A} & \mathfrak{B} & \mathfrak{C} \end{pmatrix}$  satisfy the structural and admissibility constraints. For example,  $\mu$  and  $\Omega$  of the Santa-Clara and Yan (2005) model can be represented with:

$$\mathcal{A} = \begin{pmatrix} 0_{3 \times 3} \\ 0_{3 \times 3} \\ \Lambda_{\mu_3} \end{pmatrix}, \quad \text{with } \Lambda_{\mu_3} = \begin{pmatrix} A_{\mu_3} & 0_{2 \times 1} \\ 0_{1 \times 2} & 0 \end{pmatrix};$$

$$\mathfrak{A} = \begin{pmatrix} 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & \Lambda_{33} \end{pmatrix}, \quad \text{with } \Lambda_{33} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

$$\mathfrak{B} = \begin{pmatrix} 0_{1 \times 3} & 0_{1 \times 3} & B_{13} \\ 0_{1 \times 3} & 0_{1 \times 3} & B_{23} \\ B_{31} & B_{32} & 0_{1 \times 3} \end{pmatrix}, \quad \text{with } B_{ij} = \begin{pmatrix} k_{ij} & 0 & 0 \end{pmatrix};$$

and

$$\mathcal{B} = \begin{pmatrix} k_{\mu_1} & 0 & 0 \\ 0 & k_{\mu_2} & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad \mathcal{C} = \begin{pmatrix} c_{\mu_1} \\ c_{\mu_2} \\ c_{\mu_3} \end{pmatrix}; \quad \mathfrak{C} = \begin{pmatrix} c_{11} & c_{12} & 0 \\ c_{12} & c_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We have the following proposition:

**Proposition 1** *Suppose that the technical integrability conditions hold, and that the LQ function  $g(x, t) = \frac{1}{2}x^\top \Lambda(t)x + b(t)^\top x + c(t)$  with terminal condition  $g(x, T) = \hat{g}(x)$ ,*

admits a unique solution through the following system of ODEs:

$$-\frac{db}{dt} = b_\lambda (\theta - 1) - b_R + b_\mu + \frac{1}{2}b_\Omega, \quad (4.1)$$

$$-\frac{d\Lambda}{dt} = \Lambda_\lambda (\theta - 1) - \Lambda_R + \Lambda_\mu + \frac{1}{2}\Lambda_\Omega, \quad (4.2)$$

$$-\frac{dc}{dt} = c_\lambda (\theta - 1) - c_R + c_\mu + \frac{1}{2}c_\Omega + \frac{1}{2}tr [A\bar{\Omega}], \quad (4.3)$$

where

$$\begin{aligned} \Lambda_\mu &= \mathcal{A}^\top (b \otimes \mathbf{I}_n) + 2\mathcal{B}^\top \Lambda, \\ b_\mu &= \Lambda \mathcal{C} + \mathcal{B}^\top b, \\ c_\mu &= \mathcal{C}^\top b, \end{aligned} \quad (4.4)$$

and

$$\begin{aligned} \Lambda_\Omega &= 2\Lambda \mathfrak{C} \Lambda + 4\Lambda \mathfrak{B} (b \otimes \mathbf{I}_n) + (b^\top \otimes \mathbf{I}_n) \mathfrak{A} (b \otimes \mathbf{I}_n), \\ b_\Omega &= 2\Lambda \mathfrak{C} b + (b^\top \otimes \mathbf{I}_n) \mathfrak{B}^\top b, \\ c_\Omega &= b^\top \mathfrak{C} b. \end{aligned} \quad (4.5)$$

Then the solution to the standard transform  $\phi(g; X_t, t, T)$ , defined by (2.6), is:

$$\phi(g; X_t, t, T) = e^{g(X_t, t)}.$$

**Proof.** See Appendix B.2. ■

One may have observed that the ODEs (4.2) yield a system of *non-symmetric matrix Riccati differential equations* (RDE), which takes the form:

$$\frac{d}{d\tau} \varpi = M_{21}(\tau) + M_{22}(\tau) \varpi - \varpi M_{11}(\tau) - \varpi M_{12}(\tau) \varpi. \quad (4.6)$$

Take the specifications of the Santa-Clara and Yan (2005) model as an example. The model is  $\mathbb{LQ}_0^2(3)$ , so we only have to consider the ODE for the leading diagonal block,  $A$ , of  $\Lambda$ . Moreover,  $A_R = 0$ ,  $\theta$  is calculated from a log-normal distribution, and:

$$A_\lambda = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Using the  $\begin{pmatrix} \mathcal{A} & \mathcal{B} & \mathcal{C} \end{pmatrix}$  and  $\begin{pmatrix} \mathfrak{A} & \mathfrak{B} & \mathfrak{C} \end{pmatrix}$  matrices of the Santa-Clara and Yan (2005) model presented above, we have:

$$\frac{d}{d\tau}A = M_{21}(\tau) + M_{22}(\tau)A - AM_{11}(\tau) - AM_{12}(\tau)A,$$

where:

$$M_{21}(\tau) = A_\lambda(\theta - 1) + A_{\mu_3}b_3 + \begin{pmatrix} b_3^2 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$M_{22}(\tau) = 2 \begin{pmatrix} k_{\mu_1} & 0 \\ 0 & k_{\mu_2} \end{pmatrix}, \quad M_{11}(\tau) = -2 \begin{pmatrix} k_{13}b_3 & 0 \\ k_{23}b_3 & 0 \end{pmatrix}, \quad M_{12}(\tau) = - \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix},$$

with  $b_3$  being the  $3^{rd}$  entry of  $b$  in ODE (4.1). In fact, the ODE of  $b$  can also be written in the form of (4.6).

A standard solution procedure, called *Radon's lemma*, exists for such systems:

**Theorem 1** (*Radon's lemma*)

Let:

$$M(\tau) = \begin{pmatrix} M_{11}(\tau) & M_{12}(\tau) \\ M_{21}(\tau) & M_{22}(\tau) \end{pmatrix}.$$

If  $Y(\tau) = \begin{pmatrix} Q(\tau) \\ P(\tau) \end{pmatrix}$  is, on some interval  $U \subset \mathbb{R}$ , a solution of the linear system:

$$\frac{d}{d\tau}Y(\tau) = M(\tau)Y(\tau), \quad (4.7)$$

such that  $\det Q(\tau) \neq 0$  for  $\tau \in U$ , then:

$$\varpi(\tau) = P(\tau)Q(\tau)^{-1}$$

is a solution of (4.6); in particular,  $\varpi(\tau_0) = P(\tau_0)Q(\tau_0)^{-1}$ .

By Radon's lemma, the initial problem for a matrix RDE system that we have to solve in the LQJD setting is (locally) equivalent to an initial value problem for the linear system defined in (4.7). Since standard procedures exist for solving linear systems of ODEs such as (4.7), the added computational burden is limited. For instance, in solving for the standard transform of the Santa-Clara and Yan (2005) model, one may apply Radon's lemma to obtain  $b$  and  $\Lambda$ .  $c$  is then obtained by integrating the right hand side of (4.3).

We make no further efforts on the discussion about the existence and uniqueness of the solutions of RDE systems. First, there does not exist a general theory on these issues for matrix Riccati systems (see Freiling (2002)). This implies that such discussions must be case specific. Second, all models for financial applications are simple enough to admit unique solutions. For further details on RDE systems, see, for instance, Freiling (2002) and the references therein.

Finally, we note that some authors tend to call all quadratic matrix differential equations matrix RDEs. However, not all quadratic differential equations can be represented in a form similar to (4.6). Previous studies on AJD and QG models have mentioned that the resulting ODE systems are Riccati equations, but none of them has clarified their views on this point.

## 4.2 The equivalence between LQJD and AJD models

While it is straightforward to recognize AJD class as a subset of the LQJD class through the definitions, it might come as a surprise that the LQJD class can in fact be accommodated in the AJD class. Indeed, by introducing some *pseudo-factors* to replace the quadratic terms, one may re-formulate an LQJD model as a constrained AJD model. This reformulation leads to a *one-to-one* equivalence relationship between LQJD and AJD classes in terms of their transforms.

To see this, first note that all quadratic terms are affine in elements of  $\bar{X}\bar{X}^\top$ . Therefore, we introduce the vector  $Z$  of pseudo-factors, which is defined as:

$$Z = v \left[ \bar{X}\bar{X}^\top \right], \quad (4.8)$$

where  $v$  is the vector-half operator. This operator, also denoted by  $\text{vech}$ , stacks the lower elements of a square matrix into a vector. Hence,  $v \left[ \bar{X}\bar{X}^\top \right]$  only collects the distinct elements of the symmetric matrix  $\bar{X}\bar{X}^\top$ .

Note that the drift of the quadratic vector  $\bar{X}$  is affine in  $\bar{X}$ , and the covariance matrix is deterministic. Hence the drift of  $Z$  is affine in  $\left( Z^\top \quad \bar{X}^\top \right)^\top$ , and the covariance of  $Z$  is affine in  $\bar{X}$ . We can now rephrase the LQJD setting in terms of the augmented state vector:

$$X^a = \begin{pmatrix} Z \\ \bar{X} \\ \mathbf{X} \end{pmatrix}. \quad (4.9)$$

Using notations from Duffie, Pan and Singleton (2000), we have:

$$dX^a = \mu^a(X_t^a) dt + \sigma^a(X_t^a) dW_t + dJ_t^a, \quad (4.10)$$

where:

$$\begin{aligned}\mu^a &= \mathbf{K}_1 X^a + \mathbf{K}_0, \\ \Omega^a &= \mathbf{H}_1 (\mathbf{I}_{N+n} \otimes X^a) + \mathbf{H}_0.\end{aligned}$$

The above expressions show that all terms in the LQJD setting can be represented as affine forms of  $X^a$ , which means that LQJD and AJD classes are in fact nested within each other. We further have an equivalence relationship between the LQJD and AJD classes in terms of their standard transforms:

**Proposition 2** *The standard transform of an  $n$ -factor LQJD model with an  $m$ -dimension, quadratic vector  $\bar{X}$  is equal to the standard transform of a constrained,  $(N+n)$ -factor,  $N = q(q+1)/2$ , affine model of Duffie, Pan, and Singleton (2000), where the state vector is augmented by an additional  $N \times 1$  pseudo state vector  $Z = v(\bar{X}\bar{X}^\top)$  with  $Z_0 = v(\bar{X}_0\bar{X}_0^\top)$ .*

**Proof.** Applying the results from the LQJD setting and the results from the AJD setting, respectively, to  $X$  and  $X^a$  and their transforms. This leads to two sets of ODEs, which can be shown to be equivalent<sup>4</sup>. ■

The Santa-Clara and Yan (2005) model is again well suited as an example here. For simplicity, we suppress the jump component, so that the model becomes  $\mathbb{LQ}_0^1(2)$ :

$$\mu_t = \begin{pmatrix} k_{\mu_1} X_{1t} + c_{\mu_1} \\ \frac{1}{2} A_{\mu_3} X_{1t}^2 + c_{\mu_3} \end{pmatrix}, \quad \Omega_t = \begin{pmatrix} c_{11} & k_{13} X_{1t} \\ . & X_{1t}^2 \end{pmatrix}. \quad (4.11)$$

We introduce the pseudo state vector, which consists of the instantaneous variance,  $V_t$ , of the stock price logarithm  $s_t$ :

$$Z_t = V_t = X_{1t}^2.$$

Note that the augmented state vector is  $X_t = \begin{pmatrix} V_t & X_{1t} & s_t \end{pmatrix}^\top$  and  $\Omega_t$  is affine in both  $V_t$  and  $X_{1t}$ , corresponding to an  $\mathbb{A}_2(3)$  model. We write the augmented drift  $\mu_t^a$  and covariance  $\Omega_t^a$  alongside the maximal flexible drift  $\mu_t^{\mathbb{A}}$  and covariance  $\Omega_t^{\mathbb{A}}$  of an  $\mathbb{A}_2(3)$  model

---

<sup>4</sup>For a more detailed proof, please refer to the extended version of this paper, Cheng and Scaillet (2002).

(  $X_t = \begin{pmatrix} V_t^\mathbb{A} & X_{1t}^\mathbb{A} & X_{2t}^\mathbb{A} \end{pmatrix}^\top$ , where  $V_t^\mathbb{A}$  and  $X_{1t}^\mathbb{A}$  enter the diffusion term):

$$\begin{aligned} \mu_t^a &= \begin{pmatrix} 2k_{\mu_1} & 2c_{\mu_1} & 0 \\ 0 & k_{\mu_1} & 0 \\ \frac{1}{2}A_{\mu_3} & 0 & 0 \end{pmatrix} \begin{pmatrix} V_t \\ X_{1t} \\ s_t \end{pmatrix} + \begin{pmatrix} c_{11} \\ c_{\mu_1} \\ c_{\mu_3} \end{pmatrix}, \\ \mu_t^\mathbb{A} &= \begin{pmatrix} k_{11} & k_{12} & 0 \\ k_{21} & k_{22} & 0 \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} V_t \\ X_{1t} \\ X_{2t} \end{pmatrix} + \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}, \end{aligned}$$

and:

$$\Omega_t^a = \begin{pmatrix} 4c_{11}V_t & 2c_{11}X_{1t} & 2k_{13}V_t \\ \cdot & c_{11} & k_{13}X_{1t} \\ \cdot & \cdot & V_t \end{pmatrix}, \quad \Omega_t^\mathbb{A} = \begin{pmatrix} \beta_{11}V_t^\mathbb{A} & 0 & \beta_{13}V_t^\mathbb{A} \\ \cdot & \beta_{22}X_{1t}^\mathbb{A} & \beta_{23}X_{1t}^\mathbb{A} \\ \cdot & \cdot & \beta_{33}V_t^\mathbb{A} + \beta_{33}X_{1t}^\mathbb{A} + \alpha_{33} \end{pmatrix}.$$

The most obvious difference is that in the augmented  $\mathbb{LQ}_0^1(2)$  model  $V_t = X_{1t}^2$ , where  $X_{1t}$  follows a Gaussian process, but in the  $\mathbb{A}_2(3)$  model  $V_t^\mathbb{A}$  and  $X_{1t}^\mathbb{A}$  are distinct square-root processes. Moreover, the drift  $\mu_t^\mathbb{A}$  has to satisfy the constraints:

$$k_{12} \geq 0, \quad \theta_1 \geq 0, \quad k_{21} \geq 0, \quad \theta_2 \geq 0,$$

while  $\mu_t^a$  satisfies a different set of constraints, equivalent to relaxing the constraints on  $\begin{pmatrix} k_{12} & \theta_1 & k_{21} & \theta_2 \end{pmatrix}$ , but imposing  $k_{21} = k_{32} = k_{33} = 0$ , which is unnecessary, and:

$$k_{11} = 2k_{22}, \quad k_{12} = 2\theta_2,$$

due to the deterministic relationship  $V_t = X_{1t}^2$ .

The constraints imposed on  $\Omega_t^\mathbb{A}$  and  $\Omega_t^a$  are also different. For  $\Omega_t^\mathbb{A}$ :

$$\beta_{11} > 0, \quad \beta_{22} > 0, \quad \beta_{33} > 0, \quad \beta_{33} > 0, \quad \alpha_{33} > 0.$$

For  $\Omega_t^a$ ,  $\beta_{22} = \beta_{33} = \alpha_{33} = 0$ , which is unnecessary. However, it allows correlation between  $V_t$  and  $X_{1t}$  (the correlation is  $\pm 1$ , depending on the sign of  $X_{1t}$ ), while in  $\Omega_t^\mathbb{A}$ ,  $V_t^\mathbb{A}$  and  $X_{1t}^\mathbb{A}$  have zero correlation. It is also interesting to look at the  $(2,3)^{th}$  term of the covariance matrices. In the reduced Santa-Clara and Yan (2005) model, this term is closely linked to the correlation between the volatility factor  $X_{1t}$  and the logarithm of the stock price  $s_t$ . One may observe that in the LQJD model, the correlation actually depends on the sign of  $X_{1t}$ , while in the affine model  $\mathbb{A}_2(3)$  the correlation is either positive or negative, but never

shift signs.

In summary, while an LQ model can be re-written as an affine model by augmenting the state vector, it may not fit into an affine class defined by Dai and Singleton (2000). This is due to the presence of the Gaussian factors, e.g.,  $X_{1t} = \pm\sqrt{V_t}$ , which are particular to the quadratic models. The equivalence relationship stated here is, therefore, on the valuation side. Put it differently, while an LQJD model can be solved using techniques normally applied to AJD models, it extends the modeling flexibility of affine models to the non-linear territory. This is conveniently evidenced by the discussion of the correlation terms above.

Proposition 2 is a strong result, for the quadratic class has always been taken to be a *separate* group from the affine class in asset pricing methodology. We have just shown that this perception is not valid. Indeed, the proposition above concerns a full valuation equivalence, i.e., the numerical schemes necessary for computing the transforms will deliver exactly the same results. A straightforward consequence of Proposition 2 is that the analysis techniques of affine models also become applicable to LQJD models. An example is Collin-Dufresne, Goldstein and Jones (2003), in which a three-factor maximal affine term structure model with stochastic volatility is studied. Their technique of rotating latent state variables into observables is similarly applicable in the LQJD setting. Moreover, the reformulation can be done in an automatic way through matrix algebra manipulations and easily implemented in a symbolic calculus package (see the extended version of this paper, Cheng and Scaillet (2002), for a fully worked example).

Further note that the technique of change of variables, which we use to derive the valuation equivalence between AJD and LQJD models, does not apply for general structures of the state vector. This equivalence is really due to the presence of LQ structures at every step of the identification procedure. For example, the change of variables technique breaks down in the Hull and White (1987) model, where the instantaneous variance is the exponential of the volatility factor.

### 4.3 The LQJD dynamics under the risk-neutral measure

It is well known that option prices are not derived from the data generating process under the historical (objective) measure  $\mathbb{P}$ , but from some risk-adjusted process under an equivalent measure  $\mathbb{Q}$ . Therefore, for pricing purpose, one needs to know the specification of the state-price density,  $\xi_t$ , which is defined in the LQJD framework as:

$$\xi(X_t, t) = \exp\left(-\int_0^t R^{\mathbb{P}}(X_s, s) ds\right) e^{g\xi(X_t, t)}, \quad (4.12)$$

where

$$g_\xi(X_t, t) = \frac{1}{2} X_t^\top \Lambda_\xi(t) X_t + b_\xi^\top(t) X_t + c_\xi(t),$$

satisfies the PIDE (2.9). Without loss of generality, we assume  $\xi(X_0, 0) = 1$ , which gives the initial condition for the PIDE. By Lemma 1,  $\xi(X_t, t)$  is a positive  $\mathbb{P}$ -martingale. Furthermore, by restricting  $\xi(X_t, t)$  to be exponential LQ in  $X_t$ , we have ensured that the structure of the state vector remains LQJD under the new measure  $\mathbb{Q}$ .

The equivalent martingale measure  $\mathbb{Q}$  is defined via:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_t = \frac{\xi(X_T, T)}{\xi(X_t, t)}. \quad (4.13)$$

Let:

$$W_t^\mathbb{Q} = W_t^\mathbb{P} - \int_0^t \sigma(X_s, s)^\top [\Lambda_\xi(s) X_s + b_\xi(s)] ds. \quad (4.14)$$

The following Lemma, which is similar to Lemma 2 in Appendix C of Duffie, Pan and Singleton (2000), states that  $\xi(X_t, t) W^\mathbb{Q}$  is a  $\mathbb{P}$  local martingale. It then follows that  $W^\mathbb{Q}$  is a standard Brownian Motion under  $\mathbb{Q}$ .

**Lemma 2** *Provided that all technical integrability conditions are satisfied,  $\xi(X_t, t) W^\mathbb{Q}$  is a  $\mathbb{P}$ -martingale.*

**Proof.** See Appendix C. ■

Moreover, let  $N$  be the jump-counting process with intensity  $\lambda^\mathbb{P}(X_t, t)$  under  $\mathbb{P}$  and  $\lambda^\mathbb{Q}(X_t, t)$  under  $\mathbb{Q}$ . Define:

$$M^\mathbb{Q} = N_t^\mathbb{P} - \int_0^t \theta(l_\xi) \lambda^\mathbb{P}(X_s, s) ds. \quad (4.15)$$

Since jumps are restricted to affine variables  $\underline{X}$  only, results from Duffie, Pan and Singleton (2000) concerning jumps are directly applicable in the LQJD setting. Specifically, by Lemma 3 in Appendix C of Duffie, Pan and Singleton (2000), and provided that the technical integrability conditions are satisfied,  $\xi(X_t, t) M^\mathbb{Q}$  is a  $\mathbb{P}$ -martingale. It follows that  $M^\mathbb{Q}$  is a compensated jump-counting process under  $\mathbb{Q}$ .

The structure of the state vector under the measure  $\mathbb{Q}$  is now:

$$dX_t = \mu^\mathbb{Q}(X_t, t) dt + \sigma(X_t, t) dW_t^\mathbb{Q} + dJ_t^\mathbb{Q}, \quad (4.16)$$

with the drift being:

$$\mu^\mathbb{Q}(X_t, t) = \mu^\mathbb{P}(X_t, t) + \Omega(X_t, t) [\Lambda_\xi(t) X_t + b_\xi(t)], \quad (4.17)$$

and the jump intensity being:

$$\lambda^{\mathbb{Q}}(X_t, t) = \theta(l_\xi) \lambda^{\mathbb{P}}(X_t, t). \quad (4.18)$$

The diffusion part remains unchanged.

One may now easily infer from (4.16) the market price of risk relative to the  $\mathbb{Q}$  drift  $\mu^{\mathbb{Q}}(X_t, t)$ . In particular, the quantity  $\mu^{\mathbb{Q}} - \mu^{\mathbb{P}}$  is LQ in the state vector  $X_t$ . In contrast,  $\Lambda_\xi(t)$  is zero in an AJD model, and  $\Omega$  is independent of  $X_t$  in a QG model. Consequently,  $\mu^{\mathbb{Q}} - \mu^{\mathbb{P}}$  is only affine in both latter cases.

Since the state-price density is obtained explicitly, one may estimate jointly the objective and the risk-neutral measures in the LQJD settings and extract information content from the option markets. An analysis of this kind can be found, for instance, in Chernov and Ghysels (2000) and Pan (2002).

Given the dynamics of the state vector under the risk-neutral measure  $\mathbb{Q}$ , option pricing via transform analysis is then straightforward. See e.g. Duffie, Pan and Singleton (2000), Carr and Madan (1999), and Lewis (2000).

## 5 Conclusion

We have generalized the transform analysis methods existing for the AJD and QG classes to the LQJD case. We present in detail the characterization of the LQJD structure, derive the structural restrictions and the admissibility conditions, and carry out a specification analysis for the three-factor LQJD models. We solve the standard transform up to a system of ODEs, which is identified as a system of non-symmetric Riccati differential equations with standard solution routines. Finally, we prove that an LQJD model can be converted to an AJD model by introducing a vector of pseudo factors. The notion is quite intuitive, but has never been demonstrated in full generality before. This is a strong result, for researchers have always taken affine and quadratic models as two separate classes, whereas we show that the set of the quadratic models that is absolutely distinct from the affine ones is actually empty in terms of asset pricing by transform analysis.

The LQJD model also provides the theoretical basis for future research on nonlinear specifications of the market price of risk, as well as its estimation using joint observations of prices and options. Such study with affine models has already been carried out by, for instance, Chernov and Ghysels (2000) and Pan (2002). Since the LQJD framework is very flexible, selecting an appropriate model is of great concern. We leave this issue for future research.

## Appendix A. Proof of Lemma 1

By Ito's lemma for semi-martingales, we have

$$\Phi_t = \Phi_0 + \int_0^t \mathcal{D}\Phi_s ds + \int_0^t \eta_s dW_s + \mathcal{J}_t, \quad (\text{A.1})$$

where the infinitesimal operator  $\mathcal{D}$  is defined as:

$$\begin{aligned} \mathcal{D}f(x, t) &= \frac{\partial f}{\partial t}(x, t) + \mu(x)^\top \frac{\partial f}{\partial x}(x, t) + \frac{1}{2} \text{tr} \left[ \Omega(x) \frac{\partial^2 f}{\partial x^2}(x, t) \right] \\ &\quad + \lambda(x, t) \int_{\mathbb{D}} [f(x+y, t) - f(x, t)] \Pi(x, dy, t), \end{aligned} \quad (\text{A.2})$$

and

$$\begin{aligned} \eta_t &= \left( \frac{\partial \Phi_t}{\partial x} \right)^\top \sigma_t, \\ \mathcal{J}_t &= \sum_{0 < \tau(i) \leq t} (\Phi_{\tau(i)} - \Phi_{\tau(i)-}) - \int_0^t \gamma_s ds, \end{aligned}$$

with  $\tau(i)$  denoting the  $i^{\text{th}}$  jump time of  $X$ , and

$$\gamma_t = \lambda(x, t) \int_{\mathbb{D}} [\Phi(x+y, t) - \Phi(x, t)] \Pi(x, dy, t).$$

Suppose the following technical integrability conditions hold:

- (i)  $E_0[|\Phi_T|] < \infty$ ;
- (ii)  $E_0 \left[ \left( \int_0^T \eta_s \eta_s^\top ds \right)^{1/2} \right] < \infty$ ;
- (iii)  $E_0 \left[ \int_0^T |\gamma_s| ds \right] < \infty$ .

By integrability condition (ii),  $\int_0^t \eta_s dW_s$  is a martingale. Furthermore,  $\mathcal{J}_t$  is a martingale

as well. This is because:

$$\begin{aligned}
E_0 \left[ \sum_{0 < \tau^{(i)} \leq t} (\Phi_{\tau^{(i)}} - \Phi_{\tau^{(i)-}}) \right] &= E_0 \left[ \sum_{0 < \tau^{(i)} \leq t} E_{\tau^{(i)-}} [\Phi_{\tau^{(i)}} - \Phi_{\tau^{(i)-}}] \right] \\
&= E_0 \left[ \sum_{0 < \tau^{(i)} \leq t} \Phi_{\tau^{(i)-}} (\theta_{\tau^{(i)}} - 1) \right] \\
&= E_0 \left[ \sum_{0 < \tau^{(i)} \leq t} \int_{\tau^{(i)-}}^{\tau^{(i)}} \Phi_u (\theta_u - 1) dN_u \right] \\
&= E_0 \left[ \int_0^t \Phi_u (\theta_u - 1) dN_u \right] \\
&= E_0 \left[ \int_0^t \Phi_u (\theta_u - 1) \lambda_u du \right] \\
&= E_0 \left[ \int_0^t \gamma_u du \right],
\end{aligned}$$

where the last but one equality is due to the fact that the jump-counting process  $N_t$  has intensity  $\lambda_t$ , and the last equality is by definition of  $\gamma_t$ . By integrability condition (iii), we have:

$$\begin{aligned}
E_0 [\mathcal{J}_t] &= E_0 \left[ \sum_{0 < \tau^{(i)} \leq t} (\Phi_{\tau^{(i)}} - \Phi_{\tau^{(i)-}}) - \int_0^t \gamma_u du \right] \\
&= 0.
\end{aligned}$$

Hence  $\mathcal{J}_t$  is a martingale.

The PIDE is obtained by computing  $\mathcal{D}\Phi_t$ , setting it to zero, and dividing through the resulting equation by  $\Phi_t (\neq 0)$ .

## Appendix B. Identification restrictions and ODEs

Appendix B gives details about the derivation of the structural constraints underlying the LQJD modeling, as well as the computations leading to the ODEs of Proposition 1. They both rely on the PIDE of Lemma 1:

$$R = \frac{\partial g}{\partial t} + \mu^\top \frac{\partial g}{\partial x} + \frac{1}{2} tr \left[ \left( \frac{\partial^2 g}{\partial x^2} + \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \right) \Omega \right] + \lambda (\theta - 1),$$

where all terms should be LQ in  $x$ .

## B.1 Justification of the structural constraints

Structural constraint [SC1] on the jump components is justified as follows. Note that the ODEs are identified by imposing  $\mathcal{D}\Phi_t \equiv 0$ . The last component in  $\mathcal{D}\Phi_t$  is:

$$\begin{aligned} & \lambda(x, t) \int_{\mathbb{D}} [\Phi(x + y, t) - \Phi(x, t)] \Pi(x, dy, t) \\ = & \Phi(x, t) \lambda(x, t) \int_{\mathbb{D}} \exp\left(\frac{1}{2}x^\top \Lambda(t) y + \frac{1}{2}y^\top \Lambda(t) x + \frac{1}{2}y^\top \Lambda(t) y + b(t)^\top y\right) \Pi(dy, t). \end{aligned}$$

Since  $\Phi(x, t)$  can be cancelled throughout  $\mathcal{D}\Phi(x, t) = 0$ , and since we want the remaining terms all be LQ functions for identification, it is necessary that the integral term in the above equation be independent of  $x$ . Given the structure of  $\Lambda(t)$  imposed by Assumption 1, the minimal restriction on  $y$  is then [SC1], namely its first  $q$  entries are zeros.

For [SC2] and [SC3], note that the right hand side of the PIDE above contains the following two quantities:

- i)  $\mu^\top \frac{\partial g}{\partial x}$ ;
- ii)  $tr \left[ \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \Omega \right]$ , where  $g(x, t) = \frac{1}{2}x^\top \Lambda(t) x + b(t)^\top x + c(t)$ .

i) and ii) depend on  $\mu$  and  $\Omega$ , respectively, and must be LQ in  $x$  to permit the use of the method of undetermined coefficients.

We start by making no assumptions on  $\mu$  and  $\Omega$ . First, we stack  $\mu$  as:

$$\mu = \begin{pmatrix} \frac{1}{2}x^\top \Lambda_{\mu_1} x + b_{\mu_1}^\top x + c_{\mu_1} \\ \vdots \\ \frac{1}{2}x^\top \Lambda_{\mu_n} x + b_{\mu_n}^\top x + c_{\mu_n} \end{pmatrix}_{n \times 1}.$$

Let:

$$\mathcal{A} = \begin{pmatrix} \Lambda_{\mu_1} \\ \vdots \\ \Lambda_{\mu_n} \end{pmatrix}_{n^2 \times n}, \quad \mathcal{B} = \begin{pmatrix} b_{\mu_1}^\top \\ \vdots \\ b_{\mu_n}^\top \end{pmatrix}_{n \times n}, \quad \mathcal{C} = \begin{pmatrix} c_{\mu_1} \\ \vdots \\ c_{\mu_n} \end{pmatrix}_{n \times 1},$$

then  $\mu$  can be compactly written as:

$$\mu = \frac{1}{2} \left( \mathbf{I}_n \otimes x^\top \right) \mathcal{A} x + \mathcal{B} x + \mathcal{C}.$$

Note that  $\frac{\partial g}{\partial x} = \Lambda x + b$ . For  $\mu^\top \frac{\partial g}{\partial x}$  to be LQ in  $x$ , we must have:

$$\left[ \left( \mathbf{I}_n \otimes x^\top \right) \mathcal{A} x \right]^\top \Lambda x = 0,$$

which leads to  $\Lambda_{\mu_i} \equiv 0$ , for all  $i = 1, 2, \dots, m$ . This justifies [SC2].

Similarly, we can write  $\Omega$  as:

$$\Omega = \frac{1}{2} \left( \mathbf{I}_n \otimes x^\top \right) \mathfrak{A} (\mathbf{I}_n \otimes x) + \mathfrak{B} (\mathbf{I}_n \otimes x) + \mathfrak{C}.$$

For  $tr \left[ \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \Omega \right]$  to be LQ in  $x$ , first note that:

$$tr \left[ \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \Omega \right] = \left( \frac{\partial g}{\partial x} \right)^\top \Omega \frac{\partial g}{\partial x}.$$

Hence we need:

$$\begin{aligned} x^\top \Lambda^\top \left( \mathbf{I}_n \otimes x^\top \right) \mathfrak{A} (\mathbf{I}_n \otimes x) \Lambda x &= 0, \\ x^\top \Lambda^\top \mathfrak{B} (\mathbf{I}_n \otimes x) \Lambda x &= 0, \end{aligned}$$

which yields [SC3].

## B.2 Obtaining the ODEs

From above, we can compute:

$$\mu^\top \frac{\partial g}{\partial x} = \frac{1}{2} x^\top \Lambda_\mu x + b_\mu^\top x + c_\mu,$$

where  $\left( \Lambda_\mu \quad b_\mu \quad c_\mu \right)$  are given by (4.4), and

$$tr \left[ \frac{\partial g}{\partial x} \left( \frac{\partial g}{\partial x} \right)^\top \Omega \right] = \frac{1}{2} x^\top \Lambda_\Omega x + b_\Omega^\top x + c_\Omega,$$

where  $\left( \Lambda_\Omega \quad b_\Omega \quad c_\Omega \right)$  are given by (4.5).

Moreover,

$$\frac{\partial^2 g}{\partial x^2} = \Lambda = \begin{pmatrix} A_{m \times m} & 0 \\ 0 & 0 \end{pmatrix},$$

and by [SC3](a), the leading  $m \times m$  block,  $\bar{\Omega}$ , of  $\Omega$  is deterministic in  $t$ . Hence:

$$tr \left[ \frac{\partial^2 g}{\partial x^2} \Omega \right] = tr [A \bar{\Omega}],$$

and is deterministic in  $t$  as well.

Knowing that both  $R$  and  $\lambda$  are LQ functions of  $x$ , i.e., for  $\varkappa = R, \lambda$ ,

$$\varkappa(x, t) = \frac{1}{2}x^\top \Lambda_\varkappa(t) x + b_\varkappa(t)^\top x + c_\varkappa(t),$$

we can easily apply the method of undetermined coefficients and obtain the ODEs in Proposition 1.

## Appendix C. Proof of Lemma 2

By Ito's formula, for  $0 \leq s \leq t \leq T$ ,

$$\begin{aligned} \xi_t W_t^{\mathbb{Q}} &= \xi_s W_s^{\mathbb{Q}} + \int_s^t W_u^{\mathbb{Q}} d\xi_u + \int_s^t \xi_{u-} dW_u^{\mathbb{Q}} + \int_s^t d\langle \xi, W^{\mathbb{Q}} \rangle_u^c \\ &= \xi_s W_s^{\mathbb{Q}} + \int_s^t W_u^{\mathbb{Q}} d\xi_u + \int_s^t \xi_{u-} \left( dW_u^{\mathbb{P}} - \sigma(X_u, u)^\top [\Lambda_\xi(u) X_u + b_\xi(u)] du \right) \\ &\quad + \int_s^t \xi_u \sigma(X_u, u)^\top [\Lambda_\xi(u) X_u + b_\xi(u)] du \\ &= \xi_s W_s^{\mathbb{Q}} + \int_s^t W_u^{\mathbb{Q}} d\xi_u + \int_s^t \xi_{u-} dW_u^{\mathbb{P}}, \end{aligned}$$

where  $\xi_t$  stands for  $\xi(X_t, t)$ , and  $\langle \xi, W^{\mathbb{Q}} \rangle_u^c$  denotes the continuous part of  $\langle \xi, W^{\mathbb{Q}} \rangle$ . Since  $W^{\mathbb{P}}$  and  $\xi$  are both  $\mathbb{P}$ -martingales,  $\int_0^t W_u^{\mathbb{Q}} d\xi_u$  and  $\int_0^t \xi_{u-} dW_u^{\mathbb{P}}$ ,  $t \geq 0$ , are  $\mathbb{P}$ -martingales as well. Hence  $\xi_t W_t^{\mathbb{Q}}$  is a  $\mathbb{P}$ -martingale.

## References

- [1] Ahn, D-H., and B., Gao (1999): *A Parametric Nonlinear Model of Term Structure Dynamics*. **Review of Financial Studies**, 12, 721-762.
- [2] Ahn, D-H., R.F. Dittmar, and A.R. Gallant (2002): *Quadratic Term Structure Models: Theory and Evidence*. **Review of Financial Studies**, 15, 243-288.
- [3] Ait-Sahalia, Y. (1996): *Testing Continuous-Time Models of the Spot Interest Rate*. **Review of Financial Studies**, 9, 385-426.
- [4] Backus, D., S., Foresi, A., Mozumdar, and L., Wu (2001): *Predictable Changes in Yields and Forward Rates*. **Journal of Financial Economics**, 59, 281-311.
- [5] Carr, P., and D.B. Madan (1999): *Option Valuation Using the Fast Fourier Transform*. **Journal of Computational Finance**, 2, 61-73.
- [6] Chen, L., E. Bayraktar, and H.V. Poor (2005): *Consistency Problems for Jump-Diffusion Models*. **Applied Mathematical Finance**, 12(2), 101-119.
- [7] Chen, L., D. Filipovic, and H.V. Poor (2004): *Quadratic Term Structure Models for Risk-free and Defaultable Rates*. **Mathematical Finance**, 14(4), 515-536.
- [8] Cheng, P., and O. Scaillet (2002): *Linear-Quadratic Jump-Diffusion Modeling with Application to Stochastic Volatility*. FAME working paper series, No. 67.
- [9] Chernov, M. and E. Ghysels (2000): *A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation*. **Journal of Financial Economics**, 56, 407-458.
- [10] Collin-Dufresne, P., B. Goldstein, and C. Jones (2003): *Identification and Estimation of maximal affine term structure models: an application to stochastic volatility*. Working paper, Carnegie Mellon University.
- [11] Dai, Q., and K.J. Singleton (2000): *Specification Analysis of Affine Term Structure Models*. **Journal of Finance**, 55(5), 1943-1978.
- [12] Dai, Q., and K.J. Singleton (2003): *Term Structure Dynamics in Theory and Reality*. **Review of Financial Studies**, 16, 631-678.
- [13] Duffie, D., and R. Kan (1996): *A Yield-Factor Model of Interest Rates*. **Mathematical Finance**, 6, 379-406.
- [14] Duffie, D., J. Pan, and K. Singleton (2000): *Transform Analysis and Asset Pricing for Affine Jump-Diffusions*. **Econometrica**, 68, 1343-1376.

- [15] Freiling, G. (2002): *A Survey of Nonsymmetric Riccati Equations*. **Linear Algebra and Its Applications**, 243-270.
- [16] Filipović, D. (2002): *Separable Term Structures and the Maximal Degree Problem*. **Mathematical Finance**, 12(4), 341-349.
- [17] Gouriéroux, C. and R. Sufana (2003): *Wishart quadratic term structure models*. Working paper, CREST.
- [18] Gouriéroux, C. and R. Sufana (2006): *A Classification of Two-Factor Affine Diffusion Term Structure Models*. **Journal of Financial Econometrics**, 4(1), 31-52.
- [19] Hull, J. and A. White (1987): *The Pricing of Options on Assets with Stochastic Volatilities*. **Journal of Finance**, 42, 281-300.
- [20] Lewis, A.L. (2000): **Option Valuation under Stochastic Volatility with Mathematica Code**. Finance Press, USA.
- [21] Leippold, M. and L. Wu (2002): *Asset Pricing under the Quadratic Class*. **Journal of Financial and Quantitative Analysis**, 37, 271-295.
- [22] Leippold, M. and L. Wu (2003): *Design and Estimation of Quadratic Term Structure Models*. **European Finance Review**, 7, 47-73.
- [23] Leippold, M. and L. Wu (2006): *Design and Estimation of Multi-Currency Quadratic Models*. **Review of Finance**, forthcoming.
- [24] Liu, J. (2006): *Portfolio Selection In Stochastic Environments*. **The Review of Financial Studies**, forthcoming.
- [25] Piazzesi, M. (2001): *An Econometric Model of the Yield Curve with Macroeconomic Jump Effects*. NBER Working Paper 8246.
- [26] Piazzesi, M. (2004): *Bond yields and the Federal Reserve*. **Journal of Political Economy**, 113(2), 311-344.
- [27] Pan, J. (2002): *The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time-series Study*. **Journal of Financial Economics**, 63, 3-50.
- [28] Pan, E. and L. Wu (2005): *Taking Positive interest Rates Seriously*. **Advances in Quantitative Analysis of Finance and Accounting**, 4, Chapter 14.
- [29] Protter, P.E. (2004): **Stochastic Integration and Differential Equations**. Springer.

- [30] Rady, S. (1997): *Option Pricing in the Presence of Natural Boundaries and a Quadratic Diffusion Term*. **Finance and Stochastics**, 1, 331-344.
- [31] Santa-Clara, P. and S. Yan (2005): *Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options*. Working Paper, UCLA.
- [32] Schöbel, R. and J. Zhu (1999): *Stochastic Volatility With an Ornstein-Uhlenbeck Process: An Extension*. **European Finance Review**, 3, 23-46.
- [33] Stein, E.M. and J.C., Stein (1991): *Stock Price Distributions with Stochastic Volatility: An Analytic Approach*. **The Review of Financial Studies**, 4(4), 727-752.