Linear-Quadratic Jump-Diffusion Modelling with Application to Stochastic Volatility

Peng Cheng[a] and Olivier Scaillet[b] [1]

[a] HEC Lausanne and FAME, Université de Lausanne, CH - 1015 Lausanne Dorigny, Suisse. Peng.Cheng@hec.unil.ch.

[b] HEC Genève and FAME, Université de Genève, 102 Bd Carl Vogt, CH - 1211 Genève 4, Suisse. olivier.scaillet@hec.unige.ch.

First version : November 2002

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 underlying this class of models as well as identification constraints, and compute standard and extended transforms relevant to asset pricing. We also show that the LQJD class can be embedded into the affine class through use of an augmented state vector. We further establish that an equivalence relationship holds between both classes in terms of transform analysis. An option pricing application to multifactor stochastic volatility models reveals that adding nonlinearity into the model significantly reduces pricing errors, and further addition of a jump component in the stock price largely improves goodness-of-fit for in-the-money calls but less for out-of-the-money ones.

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

JEL CLASSIFICATION: G12.

[1]We would like to thank J. Detemple, G. Dhaene, J.D. Fermanian, S. Galluccio, M. Musiela, and R. Stulz for many stimulating discussions. We have received fruitful comments from participants at the FAME doctoral workshop. 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.

1 Introduction

In modelling dynamics of state variables in an arbitrage free economy, researchers are inevitably confronted with the tension between the comprehensiveness of the underlying structure for matching stylized properties of the underlying processes driven by economic intuition, and the complexity involved in subsequent computations and calibrations.

Not surprisingly, the first, and probably still the most exploited models come from the class of affine jump-diffusions. Originated from the seminal works of Vasicek (1977) and Cox, Ingersoll, and Ross (1985), the affine class has found its way into many branches of finance, especially in the modelling of the term structure of interest rates. In answering requests from empirical analysis for more versatile structures, it has grown rapidly from single factor models without jumps to higher dimensions with non-trivial correlation structures between the state variables and jumps with even stochastic arrivals. A generalization of the affine class without jumps can be found in Duffie and Kan (1996), where closed-form solutions are found for zero-coupon-bond prices. Dai and Singleton (2000) carry out a specification analysis and sort out individual models into 'maximal flexible' subgroups. Duffie, Filipovic, Schachermayer (2002) give a complete characterization of regular affine processes and discuss their applications in the term structure of interest rates, default risks, and option pricing theory. With paramount generality, Duffie, Pan, and Singleton (2000) fully characterize the structure of affine diffusion models with jumps (AJD models), and using techniques dating back to Heston (1993), derive closed-form expressions for an 'extended transform' of such processes up to the solutions of a system of ordinary differential equations.

Despite the level of sophistication of the affine class of models and its relative computational simplicity, there is strong pressure to go beyond. The research efforts are again triggered by empirical studies from various areas of finance but mostly from the term structure of interest rates. The utmost concerns are (i) to achieve better goodness-of-fits while maintaining the positiveness of the underlying process and (ii) to capture nonlinearity in the state price density. It has been shown, empirically, that neither can rest well under the roof of the affine class. Notable works include Dai and Singleton (2000) where it is found that adding Gaussian variables in the affine state vector allows for much more flexibility in modelling and consequently improves the fitting performance; and Backus, Foresi, Mozumda, and Wu (2001) where it is shown that adding negative square-root variables help better explaining anomalies in interest rates. However, the only affine model that guarantees the positiveness of the short rates is the one where all state variables are square-root processes. Incorporating additional Gaussian or negative square-root factors will not permit almost surely positive interest rates. Furthermore, the empirical study by Dai and Singleton (2000) shows that pricing errors are sensitive to the slopes of the swap yield curves, suggesting omitted nonlinearity in the affine models.

The first step towards nonlinear models is, naturally, to consider the quadratic class where the drift and covariance matrices are quadratic forms of the state vector. Such attempts start in the early 1990s and involve the double square-root model of Longstaff (1989), the univariate and multivariate quadratic models of Beaglehole and Tenney (1992) and Beaglehole and Tenney (1991), the squared-autoregressive-independent-variable normal term structure (SAINTS) model of Constantinides (1992), the quadratic model of El Karoui, Myneni, and Viswanathan (1992), and the generalized SAINTS model of Ahn (1995). However, the general theory of the quadratic class has not taken shape until recently. The pioneering studies that lay down the framework of the quadratic class are Ahn, Dittmar, and Gallant (2002), Leippold and Wu (2002), and Chen and Poor (2002). The first two pieces of works characterize the structure of the state vector under the quadratic class, clarify the identification constraints, and develop pricing formulas for discount bonds. The specification that the short rates are quadratic in the state vector then guarantees its positivity without sacrificing modelling flexibility, and nonlinearity is taken into account by construction. It has also been noted by Leippold and Wu (2002) that the computational burdens are not much heavier than in the affine class. Chen and Poor (2002) approach the problem from a different perspective. Similar in spirits to Duffie, Filipovic, Schachermayer (2002), they define a regular quadratic Gaussian process and characterize with mathematical rigor quadratic term structure models in a gerneral Markov setting. On the empirical side, Ahn, Dittmar, and Gallant (2002), and Leippold and Wu (2001) calibrate the quadratic models against the term structure of interest rates. Their results strongly favour the quadratic models over the affine ones.

However, the theory is still far from complete. First of all, the quadratic class is developed with modelling the yield curve in mind. When it comes to the pricing of equity options with stochastic volatility, for instance, the assumption that the stock price and the state variables underlying the volatility dynamics are independent has to be made. Since correlation between the stock price and the volatility processes is instrumental to generate smile skewness (see, e.g., Renault and Touzi (1996)), and indeed a skew is commonly found empirically (see, e.g., Bates (1991, 1997)), such independence assumption is obviously a very undesirable feature. Secondly, the empirical literature of stochastic volatility also reveals that having jumps in the stock price poses a strong explanation for the magnitude of the smile effects. For example, the study of Bates (1996) on the dynamics of exchange rates embedded in Deutsche Mark option prices shows that the stochastic volatility parameters are implausible given the time series properties of implied volatilities, while a process with jumps yields much more consistent estimates. Study of the S&P500 futures option market in Bates (2000) further confirms these findings. Moreover, Bakshi, Cao, and Chen (1997) show that incorporating stochastic volatility and jumps is important for pricing and internal consistency of the model. Albeit that these models are affine, they all point to the fact that jumps play an important role in modelling the underlying price process. However, such feature cannot be reconciled in the quadratic class. Last but not least, it is intuitive to think that, by a simple change-of-variables technique the quadratic variables can be replaced by affine ones and there must be an intimate link between the quadratic and affine classes. No such efforts have been conducted in Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2002).

It is therefore our aim to establish a theoretical framework which could fulfill the unaccomplished tasks of the quadratic class in terms of asset pricing. It turns out that the only way to proceed is to partition the state vector into two parts: the quadratic vector and the affine vector. The quadratic vector is allowed to enter the quadratic forms, and the affine vector is restricted to affine forms. Moreover, the jump components are attached only to the affine vector. This structure is called *linear-quadratic* to reflect its construction. The drift matrix, the covariance matrix and the jump intensity are specified to be linear-quadratic in the state vector, while extra conditions are required to ensure identifiability of linear-quadratic diffusion model with jumps (LQJD models).

Following the methodology of Heston (1993) and Duffie, Pan, and Singleton (2000), we are able to derive standard and extended transforms in linear-quadratic processes, which prove to be useful in deriving various pricing formulas for a substantial set of financial claims.

The identification procedure leads to a system of ordinary differential equations, which we discover to be *non-symmetric Riccati differential equations*. We compute explicitly the coefficient matrices of the non-symmetric Riccati differential equations, and cite results from the literature of differential equations on how the system can be solved. Given that there is a standard routine for such purpose, we demonstrate rigorously that the computational burden is not much augmented in the linear-quadratic class.

We also recover in detail the intimate link between the quadratic and the affine classes. Specifically, we show how a linear-quadratic model can be converted to its affine counterpart in an *automatic* manner through use of an augmented state vector. Since this rewriting can be done in an automatic way through use of matrix algebra, it also means that the procedure can be easily implemented in a symbolic calculus package. Furthermore, we prove that the system of ordinary equations obtained for the quadratic class is identical to that from its affine version using the results from the affine class. This, together with the fact that the affine class is nested in the quadratic class, establishes an equivalence relationship between the two classes of models in terms of their transforms. In other words, the set of quadratic models that is absolutely distinct from the affine class is *empty* when considering asset pricing by transform analysis.

However, when an equal number of state variables is considered in both classes, quadratic models have obvious advantages over affine ones by their intrinsic parsimony as well as their ability to accommodate nonlinearity. We demonstrate this through a numerical example on stochastic volatility. The 'horse-race' of the linear-quadratic model against the affine ones of Bates (2000), Duffie, Pan, and Singleton (2000) and Heston (1993), both with and without jumps, shows that adding a supplementary state variable into the volatility process and introducing nonlinearity into the structure significantly improves goodness-of-fit.

The first LQJD model, to the best of our knowledge, appears in Piazzesi (2001), where again the problem of modelling the yield curve is considered. The paper initiates the issue of including jumps with quadratic arrival intensity in the quadratic class, and fulfil the task by partitioning the state vector into two parts, one being pure Gaussian-Markov without jumps, and the other being square-root process with jumps. The drift and variance-covariance matrices are still constrained to be affine in the state vector and no further discussion is carried out in the structural constaints on these matrices, which render the paper falling out of being a complete characterization of the LQJD class.

In a study of separable term structure models, Filipović (2002) proves that the maximal consistent order of the polynomial term structure models is two. In this sense, our study of the LQJD model actually serves as the final touch to the whole picture.

The rest of the paper is structured as follows. Section 2 describes the specification of the LQJD framework and conditions for identification. Section 3 computes the standard and extended transforms, discusses numerical solution procedures, and shows the link and equivalence between the linear-quadratic and affine classes. Section 4 presents the option pricing theory in the linear-quadratic setting, followed by Section 5 on a numerical application of LQJD modelling to stochastic volatility. Section 6 concludes. Four appendices contain technicalities and an example of affine reformulation of an LQJD model.

2 Characterization of LQJD modelling

In this section we give a detailed description of LQJD modelling.

2.1 The general LQJD setting

Suppose an n-dimension vector X_t, characterized by the stochastic differential equation (SDE):

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

is drawn from some state space $\mathbb{D}$, with

(i) W_t a standard n°-dimension Brownian motion;

(ii) J_t a pure jump process with size distribution $\Pi(X, dy, t)$ and a non-negative arrival intensity $\lambda(X_t, t)$, both continuous in X and depending only on $X_{t^-} = \lim_{s\uparrow t} X_s$ so that the process is Markovian; and

(iii) $(\mathcal{E}, \mathcal{F}, \mathbb{P})$ the usual probability space with (W, J)-augmented filtration $(\mathcal{F}_t)_{t\geq 0}$.

For identification, we require that $n \geq n^\circ$, i.e. the dimension of the state vector is at least as large as that of the Brownian motion.

In the LQJD setting, it is assumed that the drift matrix $\mu(X_t)$, 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 linear-quadratic (LQ) in the state vector X_t, namely each entry of the coefficient matrices are of the following form:

$$\varkappa(X, t) = \frac{1}{2} X^\top \Lambda(t)\, X + b^\top(t)\, X + c(t), \tag{2.2}$$

where the superscript $^\top$ denotes the transpose of the underlying matrix, and the coefficient matrices $\Lambda_{n\times n}$, $b_{n\times 1}$ and c (possibly complex-valued) are all deterministic in t.

More specifically, Λ is block diagonal with only the leading square block being non-singular and all remaining entries being zeros:

Assumption 1 *For all* $\varkappa$*,*

$$\Lambda = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \tag{2.3}$$

where $A_{m\times m}$*,* $m \leq n$*, is non-singular and symmetric.*

By Assumption 1, only the first m components of the state vector will enter the quadratic term of (2.2). To see this, partition the state vector X and the coefficient vector of the affine term b as:

$$X = \begin{pmatrix} \bar{\mathrm{X}} \\ \underline{\mathrm{X}} \end{pmatrix}, \qquad b = \begin{pmatrix} k \\ l \end{pmatrix}, \tag{2.4}$$

where the upper part are $m \times 1$ and the lower part are $(n - m) \times 1$, respectively. We can now rewrite (2.2) as:

$$\varkappa(X, t) = \frac{1}{2} \bar{\mathrm{X}}^\top A(t) \bar{\mathrm{X}} + k^\top(t)\, \bar{\mathrm{X}} + l^\top(t)\, \underline{\mathrm{X}} + c(t). \tag{2.5}$$

We will switch between the representations (2.2) and (2.5) in the sequel, whichever is more convenient.

It is obvious that models of the AJD class are obtained by setting $A \equiv 0$ and $k \equiv 0$ in (2.5), while models of the quadratic class correspond to the absence of a jump component and $l \equiv 0$. By reference to these two classes, the variables $\bar{X}$, resp. $\underline{X}$, are named quadratic, resp. affine. With the additional quadratic components introduced through $\bar{X}$, LQJD models can capture nonlinearity in the drift, diffusion, and jump intensity of the state vector, a property that is often found empirically (see, e.g., Ait-Sahalia (1996), Ahn and Gao (1999)), but missing from AJD models.

2.2 The identification constraints

For an LQJD model to be identifiable (in the sense of permitting use of the method of undetermined coefficients discussed in Section 3), we need specific constraints on the drift and covariance matrices as well as the jump component. We state these constraints as assumptions and explain the necessity of their existence in an intuitive manner. Appendix A gives a formal treatment of this issue.

As already mentioned the state vectors of quadratic and AJD models correspond to $\bar{X}$ and $\underline{X}$ of X, respectively. Given that the maximum order of X in the LQJD setting is the same as that in the quadratic class and higher than that in the AJD class, we expect the identification constraints of the quadratic class to be binding (on $\bar{X}$) in the LQJD class as well, and those of the AJD class to be relaxed.

In the quadratic case, it is necessary (and sufficient) to have the drift matrix be affine in the state vector and the diffusion matrix be constant. See, for example, Leippold and Wu (2002) for details. Therefore:

Assumption 2 *The drift matrix $\bar{\mu}$ of the quadratic variables $\bar{X}$ is affine in the quadratic variables $\bar{X}$.*

Assumption 3 *The diffusion matrix $\bar{\sigma}$ of the quadratic variables $\bar{X}$ is deterministic in t.*

We further require the next restriction on the drift.

Assumption 4 *The drift matrix $\bar{\mu}$ of the quadratic variables $\bar{X}$ is independent of the affine variables $\underline{X}$.*

Assumption 4 could be justified heuristically by the fact that $\underline{X}$ should only remain in affine terms, hence must not enter the drift of $\bar{X}$ which will pass through the quadratic terms. This might look quite restrictive because it might be of interest in some models to link members of $\underline{X}$ with $\bar{X}$. 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 $\bar{X}$. Assumption 4 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 X. The numerical example in Section 5 demonstrates that this could be effectively done by including a factor that is (almost) perfectly correlated with the logarithm of the stock price.

In the AJD class, the model is identifiable as long as the drift and the covariance matrices of the state vector are affine. Since the order of X is raised to two in the LQJD framework, the drift and the covariance matrices of $\underline{X}$ gain further flexibility in specification, i.e.,

Assumption 5 *The drift matrix $\underline{\mu}$ of the affine variables $\underline{X}$ is LQ in X.*

Assumption 6 *Each entry in the diffusion matrix $\underline{\sigma}$ of the affine variables $\underline{X}$ can be either affine in the quadratic variables $\bar{X}$ or square-root affine in an LQ function of X, but not both.*

Assumption 6 differs from the usual practice in affine modelling, which tends to restrict the diffusion matrix to be square-root affine in the state vector. See, for example, Dai and Singleton (2000). The motivation for such restriction is to ensure that the underlying process is positive almost surely. In the LQJD environment, however, positiveness of the underlying process is naturally guaranteed through the quadratic state vector $\bar{\mathrm{X}}$. Moreover, as is also pointed out in Duffie, Pan and Singleton (2000), it is the covariance matrix, not the diffusion matrix itself, that matters in the procedure of identification. Such restriction in the LQJD setting would in fact exclude an important group of model candidates, which distinguish themselves from square-root affine models by carrying *sign* information in the diffusion matrix.

If a square-root affine process does enter the affine variables $\underline{\mathrm{X}}$, there are some extra restrictions on other members of X such that no square roots in the state vector show up in Ω and its LQ property is maintained:

Assumption 7 *If the $(i,j)^{th}$ entry of the diffusion matrix $\underline{\sigma}$ of the affine variables $\underline{X}$ is square-root affine in an LQ function of X, then each entry in the j^{th} column must also be square-root affine in the same function.*

Assumption 8 *If the j^{th} column of the diffusion matrix $\underline{\sigma}$ of the affine variables $\underline{X}$ is square-root affine, then the quadratic variables $\bar{X}$ are independent of the corresponding Brownian motion W_j.*

To complete the LQJD setup, we note that jumps will be restricted to the affine variables $\underline{\mathrm{X}}$. This assumption avoids quadratic jump components impeding identification.

Assumption 9 *The first m entries of the jump component J are zeros.*

We further need positivity of the jump intensity, as well as the variances of the state variables which are diagonal terms of Ω. We use the superscript $^+$ to denote the Moore-Penrose, or generalized inverse of a matrix. When a matrix Λ is non-singular, $\Lambda^+ = \Lambda^{-1}$. The necessary and sufficient conditions for the LQ jump intensity and variances to be positive semi-definite are:

Assumption 10 *The jump intensity and the variances of the state variables are LQ in the state vector, and their coefficient matrices $\begin{pmatrix} \Lambda. & b. & c. \end{pmatrix}$ satisfy:*
i) $\Lambda.$ is positive semidefinite;
ii) $b.$ belongs to the column space of $\Lambda.$; and
iii) $c. \geq \frac{1}{2} b.^{\top} \Lambda.^{+} b.$.

We may now restate the structure of the LQJD state vector X more clearly by stacking the coefficient matrices.

a) For the quadratic variables $\bar{\mathrm{X}}$:

$$d\bar{\mathrm{X}}_t = \bar{\mu}\left(\bar{\mathrm{X}}_t, t\right) dt + \bar{\sigma} dW_t, \tag{2.6}$$

where $\bar{\mu}\left(\bar{\mathrm{X}}_t, t\right)$ is $m \times 1$ and $\bar{\sigma}(t)$ is $m \times n$. By Assumption 3, the drift is equal to:

$$\bar{\mu}\left(\bar{\mathrm{X}}_t, t\right) = \begin{pmatrix} k_1^\top \bar{\mathrm{X}}_t + c_1 \\ \vdots \\ k_m^\top \bar{\mathrm{X}}_t + c_m \end{pmatrix}. \tag{2.7}$$

To streamline notations, we use calligraphic letters $\begin{pmatrix} \mathcal{A} & \mathcal{K} & \mathcal{L} & \mathcal{C} \end{pmatrix}$ for coefficient matrices associated with the drift:

$$\bar{\mathcal{K}} = \begin{pmatrix} \bar{k}_1^\top \\ \vdots \\ \bar{k}_m^\top \end{pmatrix}_{m \times m}, \quad \bar{\mathcal{C}} = \begin{pmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_m \end{pmatrix}_{m \times 1}.$$

An upper bar denotes that the coefficients are those of $\bar{\mu}\left(\bar{\mathrm{X}}_t, t\right)$, which can now be written compactly as:

$$\bar{\mu}\left(\bar{\mathrm{X}}_t, t\right) = \bar{\mathcal{K}}\bar{\mathrm{X}}_t + \bar{\mathcal{C}}. \tag{2.8}$$

Moreover, the covariance matrix $\bar{\Omega}(t) = \bar{\sigma}(t)\bar{\sigma}(t)^\top$ of $\bar{\mathrm{X}}_t$ is deterministic in t by Assumption 3.

b) For the affine variables $\underline{\mathrm{X}}$:

$$d\underline{\mathrm{X}}_t = \underline{\mu}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right) dt + \underline{\sigma}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right) dW_t + d\underline{J}, \tag{2.9}$$

where $\underline{J}$ is the lower $(n-m) \times 1$ block of the jump vector J, $\underline{\mu}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right)$ is $(n-m) \times 1$ and $\underline{\sigma}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right)$ is $(n-m) \times n$.

Similar analysis leads to:

$$\underline{\mu}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right) = \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}_t^\top\right) \underline{\mathcal{A}}\bar{\mathrm{X}}_t + \underline{\mathcal{K}}\bar{\mathrm{X}}_t + \underline{\mathcal{L}}\underline{\mathrm{X}}_t + \underline{\mathcal{C}}, \tag{2.10}$$

and:

$$\underline{\Omega}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t, t\right) = \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}^\top\right) \underline{\mathfrak{A}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \underline{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \underline{\mathfrak{L}}\left(\mathbf{I}_{n-m} \otimes \underline{\mathrm{X}}\right) + \underline{\mathfrak{C}}, \tag{2.11}$$

where Fractur style letters $\begin{pmatrix} \mathfrak{A} & \mathfrak{K} & \mathfrak{L} & \mathfrak{C} \end{pmatrix}$ are used for coefficients of the covariance matrix and an under bar denotes their position, $\mathbf{I}_n$ denotes the identity matrix of order n, $\otimes$ is the Kronecker product operator,

$$\underline{\mathcal{M}} = \begin{pmatrix} \underline{M}_1 \\ \vdots \\ \underline{M}_{n-m} \end{pmatrix},$$

with $\mathcal{M} = \mathcal{A}, \mathcal{K}, \mathcal{L}, \mathcal{C}$ $(M = A, k, l, c)$ and dimensions $m(n-m) \times m$, $(n-m) \times m$, $(n-m) \times$

$(n-m)$, $(n-m)\times 1$, respectively, and

$$\underline{\mathfrak{M}} = \begin{pmatrix} \underline{M}_{1,1} & \cdots & \underline{M}_{1,n-m} \\ \vdots & \ddots & \vdots \\ \underline{M}_{n-m,1} & \cdots & \underline{M}_{n-m,n-m} \end{pmatrix},$$

with $\underline{\mathfrak{M}} = \underline{\mathfrak{A}}, \underline{\mathfrak{K}}, \underline{\mathfrak{L}}, \underline{\mathfrak{C}}$ $(M = A, k, l, c)$ and dimensions $m(n-m) \times m(n-m)$, $(n-m) \times m(n-m)$, $(n-m) \times (n-m)^2$, $(n-m) \times (n-m)$, respectively.

Finally, by Assumptions 3, 7 and 8, the covariance matrix between the quadratic and affine variables will be affine in $\bar{\mathrm{X}}$ only. Hence, it can be represented as:

$$\tilde{\Omega}\left(\bar{\mathrm{X}}_t, t\right) = \tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \tilde{\mathfrak{C}}, \tag{2.12}$$

where:

$$\tilde{\mathfrak{M}} = \begin{pmatrix} \tilde{M}_{1,1}^{\top} & \cdots & \tilde{M}_{1,n-m}^{\top} \\ \vdots & \ddots & \vdots \\ \tilde{M}_{m,1}^{\top} & \cdots & \tilde{M}_{m,n-m}^{\top} \end{pmatrix},$$

with $\tilde{\mathfrak{M}} = \tilde{\mathfrak{K}}, \tilde{\mathfrak{C}}$ $(M = k, c)$, and dimensions $m \times m(n-m)$ and $m \times (n-m)$, respectively.

c) Eventually the complete drift and covariance matrices of the state vector X are given by:

$$\mu\left(X_t, t\right) = \begin{pmatrix} \bar{\mu}\left(\bar{\mathrm{X}}_t\right) \\ \underline{\mu}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t\right) \end{pmatrix}_{n\times 1}, \tag{2.13}$$

and

$$\Omega\left(X_t, t\right) = \begin{pmatrix} \bar{\Omega} & \tilde{\Omega}\left(\bar{\mathrm{X}}_t\right) \\ \tilde{\Omega}\left(\bar{\mathrm{X}}_t\right)^{\top} & \underline{\Omega}\left(\bar{\mathrm{X}}_t, \underline{\mathrm{X}}_t\right) \end{pmatrix}_{n\times n}. \tag{2.14}$$

3 The transforms

Given that the underlying state vector is specified so that it satisfies all identification constraints, we are able to compute the expected value of some discounted payoffs up to the solution of a system of ordinary differential equations (ODE).

The procedure of identifying the system of ODEs is similar to that used in the AJD and quadratic classes. We first use Ito's lemma to decompose a random payoff into a finite variation part and a martingale. If the finite variation part is null, the conditional expectation of the random payoff is a martingale. This yields a partial differential equation (PDE) characterization of the payoff. The system of ODEs can then be obtained by conjecturing a solution, passing it through the PDE, and applying the method of undetermined coefficients.

The above discussion also indicates that payoffs that can be priced analytically are not arbitrary. The class of payoffs (possibly after some kind of transforms) that can be handled by such procedure must be one of the following:

$$V^{s}\left(X_{T},T\right)=e^{g_{1}\left(X_{T},T\right)}, \tag{3.1}$$

or

$$V^{e}\left(X_{T},T\right)=g_{0}\left(X_{T},T\right)e^{g_{1}\left(X_{T},T\right)}, \tag{3.2}$$

where the discount rate is $R\left(X_{t},t\right)$, and $g_{i}\left(X_{t},t\right)$, $i=0,1$, and $R\left(X_{t},t\right)$ are all LQ functions. Note that adequate retrictions as in Assumption 10 may be imposed to ensure positivity of interest rate (and default rate) processes modelled as LQ functions.

The payoff V^{s}, resp. V^{e}, is termed the *standard*, resp. *extended*, *transform* by Duffie, Pan and Singleton (2000) in addressing AJD models. We adopt the same terminology here, but note that the set of payoffs that we work on has been enlarged. Typical examples of this extended class of payoffs are those that are quadratic in $\bar{\mathrm{X}}$.

Before going for the computations, we introduce the following notations for differentiation: for any $C^{1,1}$ function $f:\mathbb{D}\times\mathbb{R}^{+}\rightarrow\mathbb{C}$, define:

$$\begin{aligned}
\dot{f}\left(x,t\right) &= \frac{\partial f}{\partial t}\left(x,t\right),\\
f_{x_{i}}\left(x,t\right) &= \frac{\partial f}{\partial x_{i}}\left(x,t\right),\\
\bigtriangledown_{x}f\left(x,t\right) &= \left[f_{x_{1}}\left(x,t\right),...,f_{x_{n}}\left(x,t\right)\right]^{\top},\\
\bigtriangledown_{xx}f\left(x,t\right) &= \left[\bigtriangledown_{x}f_{x_{1}}\left(x,t\right),...,\bigtriangledown_{x}f_{x_{n}}\left(x,t\right)\right].
\end{aligned}$$

The infinitesimal operator $\mathcal{D}$ is then defined as:

$$\begin{aligned}
\mathcal{D}f\left(x,t\right) &= \dot{f}\left(x,t\right)+\mu\left(x\right)^{\top}\bigtriangledown_{x}f\left(x,t\right)+\frac{1}{2}tr\left[\Omega\left(x\right)\bigtriangledown_{xx}f\left(x,t\right)\right]\\
&\quad+\lambda\left(x,t\right)\int_{\mathbb{D}}\left[f\left(x+y,t\right)-f\left(x,t\right)\right]\Pi\left(x,dy,t\right).
\end{aligned} \tag{3.3}$$

To alleviate the notational burden, we will often omit all function arguments in the following sections as well as in the appendices.

3.1 The standard transform

We aim at computing the standard transform defined as follows:

$$\phi^{s}\left(g_{1};X_{t},t,T\right)=E_{t}\left[\exp\left(-\int_{t}^{T}R\left(X_{s},s\right)ds\right)e^{g_{1}\left(X_{T},T\right)}\right]. \tag{3.4}$$

One may have noticed that the standard transform gives the time t price of a future payoff $e^{g_{1}\left(X_{T},T\right)}$ where $R\left(X_{t},t\right)$ is the appropriate discount rate. For instance, when $g_{1}\left(X_{t},t\right)\equiv0$, (3.4) yields the price of a zero-coupon bond. For truncated payoffs such as European and Asian options, (3.4) is not directly applicable. However, many authors have shown that, after some transformations, the prices assume the form of (3.4) and they can then be solved in the transformed space using the results of this section. See also Carr and

Madan (1999) and Lewis (2001) for details on option pricing using transform techniques.

To solve for the standard transform, first consider the process:

$$\Phi_t = \exp\left(-\int_0^t R(X_s, s)ds\right) e^{g_1(X_t,t)}. \tag{3.5}$$

By Ito's lemma, we have

$$\Phi_t = \Phi_0 + \int_0^t \mathcal{D}\Phi_s ds + \int_0^t \eta_s dW_s + J_t. \tag{3.6}$$

The processes η_t and J_t are, respectively,

$$\eta_t = (\bigtriangledown_x \Phi_t)^\top \sigma_t, \tag{3.7}$$

and

$$J_t = \sum_{0<\tau(i)\leq t} \left(\Phi_{\tau(i)} - \Phi_{\tau(i)-}\right) - \int_0^t \gamma_s ds, \tag{3.8}$$

where $\tau(i)$ denotes the i^{th} jump time of X and:

$$\gamma_t = \lambda \int_{\mathbb{D}} \left[\Phi\left(X_t + y\right) - \Phi\left(X_t\right)\right] \Pi\left(X_t, dy\right). \tag{3.9}$$

Suppose the technical integrability conditions given below are satisfied:

$$\begin{array}{ll} (i) & E\left[|\Phi_T|\right] < \infty, \\ (ii) & E\left[\left(\int_0^T \eta_s \eta_s' ds\right)^{1/2}\right] < \infty, \\ (iii) & E\left[\int_0^T |\gamma_s| ds\right] < \infty. \end{array} \tag{3.10}$$

By Lemma 1 in Appendix A of Duffie, Pan and Singleton (2000), both J_t and $\int_0^t \eta_s dW_s$ are martingales. Furthermore, if $\mathcal{D}\Phi_t \equiv 0$, Φ is driftless and hence a martingale as well, namely:

$$\Phi_t = E_t^{\mathcal{X}}[\Phi_T]. \tag{3.11}$$

Multiplying both sides of this last equality by $\exp\left(\int_0^t R(X_s, s)ds\right)$ yields our standard transform (3.4), i.e.,

$$\phi^s\left(g_1; X_t, t, T\right) = e^{g_1(X_t,t)}. \tag{3.12}$$

What remains now is to find a suitable function $g_1\left(X_t, t\right)$ such that the condition $\mathcal{D}\Phi_t \equiv 0$ holds. Let:

$$\theta_1\left(l_1\right) = \int_{\mathbb{D}} e^{l_1^\top \underline{y}} \Pi\left(x, dy\right), \tag{3.13}$$

where $\underline{y}$ is the lower $(n-m) \times 1$ block of y. We have:

Lemma 1 *If the technical integrability conditions* (3.4) *hold and the LQ function* $g_1(X_t, t)$ *satisfies the partial integro-differential equation (PIDE) (the Cauchy problem):*

$$R = \dot{g}_1 + \mu^\top \triangledown_x g_1 + \frac{1}{2} tr\left[\left(\triangledown_{xx} g_1 + \triangledown_x g_1 (\triangledown_x g_1)^\top\right)\Omega\right] + \lambda\left[\theta_1(l_1) - 1\right], \tag{3.14}$$

then Φ_t *is a martingale.*

Proof. Since the first m components of y are zeros by construction, we conclude that for any LQ function $\varkappa$:

$$\varkappa(x+y) = \varkappa(x) + l_\varkappa^\top \underline{\mathrm{y}}.$$

Hence:

$$\Phi_t(x+y) - \Phi_t(x) = \Phi_t\left(e^{l_1^\top \underline{\mathrm{y}}} - 1\right).$$

The rest of the proof is then straightforward from computing the derivative terms in $\mathcal{D}\Phi_t$. ■

Since the functional form g_1 is known, we can identify g_1 by the method of undetermined coefficients. This procedure results in a set of ODEs which have to be satisfied by the coefficient matrices $\begin{pmatrix} A_1 & k_1 & l_1 & c_1 \end{pmatrix}$ of g_1. The problem of (3.4) is henceforth reduced to solving the system of ODEs in the following proposition.

Proposition 1 *Suppose the technical integrability conditions (3.10) hold and the following system of ODEs, with initial conditions given by* $g_1(X_\tau, \tau)|_{\tau=0} = g_1(X_t, t)|_{t=T}$, $\tau = T - t$, *admits a unique solution:*

$$\frac{d}{d\tau} l_1 = \underline{\mathcal{L}}^\top l_1 + \frac{1}{2}\left(l_1^\top \otimes \mathbf{I}_{n-m}\right)\underline{\mathfrak{L}}^\top l_1 + \left[\theta_1(l_1) - 1\right]\, l_\lambda - l_R, \tag{3.15}$$

$$\frac{d}{d\tau}\varpi_1 = M_{21} + M_{22}\varpi_1 - \varpi_1 M_{11} - \varpi_1 M_{12}\varpi_1, \tag{3.16}$$

$$\frac{d}{d\tau} c_1 = b_1^\top \mathcal{C} + \frac{1}{2}\left(tr\left[A_1 \bar{\Omega}^\top\right] + b_1^\top \mathfrak{C} b_1\right) + \left[\theta_1(l_1) - 1\right] c_\lambda - c_R, \tag{3.17}$$

where $(A, k, l, c)_{\lambda,R}$ *are the coefficient matrices of LQ functions* λ *and* R*, respectively,* $\varpi_1 = \begin{pmatrix} A_1 & k_1 \end{pmatrix}$*, and matrices* $\begin{pmatrix} M_{11} & M_{12} & M_{21} & M_{22} \end{pmatrix}$ *are functions of* τ *and* l_1 *with explicit expressions detailed in Appendix A. Then the standard transform* ϕ^s *defined by (3.4) is given by (3.12).*

Proof. Straightforward from previous discussion and the computations made in Appendix A during the identification process. ■

One may have noticed that ODEs (3.15) and (3.17) bear close resemblance to the ones of the AJD class (see Equations (2.5) and (2.6) of Duffie, Pan and Singleton (2000)). This heavily suggests that there is some intimate link between LQJD and AJD models. Furthermore, (3.15) is a system of its own and thus can be solved independently of others. This further suggests that the appropriate procedure of solving the whole system is to follow the sequence of the sub-systems (3.15), (3.16) and (3.17). Finally, the ODE (3.16) is of the standard form of a system of *non-symmetric matrix Riccati differential equations* (RDE). Since standard procedures exist for solving non-symmetric RDEs, we have a complete routine for disentangling the whole system, and the level of complexity obviously depends on the sub-system (3.15). We will elaborate on these points after a brief discussion of the extended transform.

3.2 The extended transform

By the same rationale as that in the case of the standard transform, we can identify the system of ODEs for the extended transform:

$$\phi^e\left(g_0, g_{1;} X_t, t, T\right) = E_t\left[\exp\left(-\int_t^T R\left(X_s, s\right) ds\right) g_0\left(X_T, T\right) e^{g_1\left(X_T, T\right)}\right]. \tag{3.18}$$

Consider the random process:

$$\tilde{\Phi}_t = \exp\left(-\int_0^t R_s ds\right) g_0 e^{g_1}. \tag{3.19}$$

Assume that following technical integrability conditions hold:

$$\begin{array}{ll} (i) & E\left[|\tilde{\Phi}_T|\right] < \infty\ , \\ (ii) & E\left[\left(\int_0^T \tilde{\eta}_s \tilde{\eta}_s' ds\right)^{1/2}\right] < \infty\ , \\ (iii) & E\left[\int_0^T |\tilde{\gamma}_s| ds\right] < \infty\ , \end{array} \tag{3.20}$$

where:

$$\tilde{\eta}_t = \left(\nabla_x \tilde{\Phi}_t\right)^\top \sigma_t, \tag{3.21}$$

and

$$\tilde{\gamma}_t = \lambda \int_{\mathbb{D}} \left[\tilde{\Phi}\left(X_t + y\right) - \tilde{\Phi}\left(X_t\right)\right] \Pi\left(X_t, dy\right). \tag{3.22}$$

If functions $g_i\left(X_t, t\right)$, $i = 0, 1$, uniquely exist such that $\mathcal{D}\tilde{\Phi}_t = 0$, then $\tilde{\Phi}_t$ is a martingale. We then have a solution to the extended transform:

$$\phi^e\left(g_0, g_{1;} X_t, t, T\right) = g_0\left(X_t, t\right) e^{g_1\left(X_t, t\right)}. \tag{3.23}$$

Applying results obtained for g_1, we can identify the PIDE for g_0. Let:

$$\theta_0\left(l_0, l_1\right) = \int_{\mathbb{D}} \left(l_0^\top \underset{\sim}{y}\right) e^{l_1^\top \underset{\sim}{y}} \Pi\left(x, dy\right). \tag{3.24}$$

We have:

Lemma 2 *If the technical integrability conditions* (3.20) *hold and the LQ function* $g_0\left(x, t\right)$

satisfies the PIDE:

$$0 = \dot{g}_0 + \mu^\top \bigtriangledown_x g_0 + \frac{1}{2} tr\left[\left(\bigtriangledown_{xx} g_0 + 2 \bigtriangledown_x g_0 \left(\bigtriangledown_x g_1\right)^\top\right)\Omega\right] + \lambda\theta_0\left(l_0, l_1\right), \tag{3.25}$$

then $\tilde{\Phi}_t$ *is a martingale.*

Proof. Recall that, for any LQ function $\varkappa$:

$$\varkappa(x+y) = \varkappa(x) + l_\varkappa^\top \underline{y}.$$

So:

$$\begin{aligned}\tilde{\Phi}(x+y) - \tilde{\Phi}(x) &= g_0(x+y)\Phi(x+y) - g_0(x)\Phi(x) \\ &= g_0(x)\left[\Phi(x+y) - \Phi(x)\right] + l_\varkappa^\top \underline{y}\Phi(x+y)\end{aligned},$$

which gives

$$\begin{aligned}&\int_{\mathbb{D}} \left[\tilde{\Phi}(x+y) - \tilde{\Phi}(x)\right]\Pi(x,dy) \\ = \ &\tilde{\Phi}\left[\int_{\mathbb{D}} e^{l_1^\top \underline{y}}\Pi(x,dy) - 1\right] + \Phi\left[\int_{\mathbb{D}} \left(l_0^\top \underline{y}\right) e^{l_1^\top \underline{y}}\Pi(x,dy)\right].\end{aligned}$$

The rest of the proof is straightforward from computing terms of derivatives in $\mathcal{D}\tilde{\Phi}_t$. ■

Now by the same technique of coefficient identification as the one used for Proposition 1, we get a set of ODEs for the coefficient matrices $\begin{pmatrix} A_0 & k_0 & l_0 & c_0 \end{pmatrix}$ of g_0:

Proposition 2 *Suppose the technical integrability conditions (3.20) hold and the following system of ODEs, with initial conditions given by* $\left[g_0\left(X_\tau, \tau\right)\right]_{\tau=0} = \left[g_0\left(X_t, t\right)\right]_{t=T}$, $\tau = T - t$, *admits a unique solution:*

$$\frac{d}{d\tau} l_0 = \underline{\mathcal{L}}^\top l_0 + \left(l_0^\top \otimes \mathbf{I}_{n-m}\right)\underline{\mathfrak{L}}^\top l_1 + \theta_0\left(l_0, l_1\right) l_\lambda, \tag{3.26}$$

$$\frac{d}{d\tau}\varpi_0 = M_{21}^e + M_{22}^e\varpi_0 - \varpi_0 M_{11}^e - \varpi_0 M_{12}^e \varpi_0, \tag{3.27}$$

$$\frac{d}{d\tau} c_0 = b_0^\top \mathcal{C} + \frac{1}{2} tr\left[A_0 \bar{\Omega}^\top\right] + b_0^\top \mathfrak{C} b_1 + \theta_0\left(l_0, l_1\right) c_\lambda, \tag{3.28}$$

where $\varpi_0 = \begin{pmatrix} A_0 & k_0 \end{pmatrix}$, *and matrices* $\begin{pmatrix} M_{11}^e & M_{12}^e & M_{21}^e & M_{22}^e \end{pmatrix}$ *are functions of* τ, l_0 *and* l_1 *with explicit forms detailed in Appendix A. Then the extended transform* ϕ^e *defined by (3.18) is given by (3.23) where* g_1 *is determined as in Proposition 1.*

Proof. Similar to that of Proposition 1. ■

Note that ODES (3.26) and (3.28) are also akin to ODEs encountered in AJD modelling (see Duffie, Pan and Singleton (2000)).

3.3 Solving non-symmetric matrix Riccati equations

As we have previously remarked, the ODEs satisfied by ϖ_i, $i = 0, 1$, are non-symmetric matrix RDEs, which are, in general, of the form:

$$\frac{d}{d\tau}\varpi = M_{21}(\tau) + M_{22}(\tau)\varpi - \varpi M_{11}(\tau) - \varpi M_{12}(\tau)\varpi, \tag{3.29}$$

where ϖ is a matrix and the coefficients,

$$M(\tau) = \begin{pmatrix} M_{11}(\tau) & M_{12}(\tau) \\ M_{21}(\tau) & M_{22}(\tau) \end{pmatrix}, \tag{3.30}$$

can be real or complex. A powerful result on RDEs, the *Radon's lemma* (see Freiling (2002)), says that each non-symmetric matrix RDE system is equivalent to a linear system in the following sense:

Theorem 1 *(Radon's lemma)*

1. *Let* $\varpi(\tau)$ *be on some interval* $U \subset \mathbb{R}$ *a solution of the RDE* (3.29) *with* $\varpi(\tau_0) = \varpi_0$. *If* Q *is for* $\tau \in U$ *the unique solution of the initial value problem*

$$\begin{aligned} \frac{d}{d\tau}Q(\tau) &= (M_{11}(\tau) + M_{12}(\tau_0)\varpi(\tau))\,Q(\tau), \\ Q(\tau_0) &= \mathbf{I}, \end{aligned}$$

and $P(\tau) := \varpi(\tau)\,Q(\tau)$, *then*

$$Y(\tau) = \begin{pmatrix} Q(\tau) \\ P(\tau) \end{pmatrix}$$

defines for $\tau \in U$ *the solution of* (3.31) *with* $Y(\tau_0) = \begin{pmatrix} \mathbf{I} \\ \varpi_0 \end{pmatrix}$.

2. *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:*

$$\frac{d}{d\tau}Y(\tau) = M(\tau)\,Y(\tau), \tag{3.31}$$

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

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

is a solution of (3.29); *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 (3.31). Since standard procedures exist for solving linear systems of ODEs such as (3.31), the added computation burden relative to that of an AJD model is limited, whereas restrictions are unleashed on the structural flexibility of the state vector, which had to be sacrificed for computational efficiency.

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. One example is that l_i is forced to be a constant, i.e. $\frac{d}{d\tau}l_i = 0$, so that the coefficients of the Riccati systems ϖ_i are constant and that unique solutions are guaranteed[2]. Although such a restriction is not necessary, it does ease the computations, and the resulting structures of the state vector are still sophisticated enough to capture stylized facts found for the underlying processes. 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 (3.16) and (3.27). Previous studies on AJD and quadratic models have mentioned that the resulting ODE systems are Riccati equations, but none of them has clarified their own views on this point. We are the first to represent the system in the form of non-symmetric matrix RDEs with all coefficient matrices explicitly identified. See Appendix A for details.

3.4 The equivalence between LQJD and AJD classes

As already mentioned the AJD class is a subset of the LQJD class. It will soon be shown that the LQJD class can be accommodated in the AJD class by replacing the quadratic terms by some new variables or *pseudo-factors*. This way of proceeding should not come as a surprise if one recalls how a quadratic fit can be easily achieved in the usual linear regression model. Quadratic and cross-products terms only need to be treated as additional factors (explanatory variables in the regression case). This reformulation will lead to an 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 $X_t X_t^\top$. Therefore, we introduce the vector Z of pseudo-factors, which is defined as:

$$Z = v\left[\bar{\mathrm{X}}\bar{\mathrm{X}}^\top\right], \tag{3.32}$$

where v is the vector-half operator. This operator, also sometimes denoted by vech, stacks the lower elements of an $m \times m$ matrix into an $m(m+1)/2 \times 1$ vector. Hence $v\left[\bar{\mathrm{X}}\bar{\mathrm{X}}^\top\right]$ only collects the distinct elements of the squared symmetric matrix $\bar{\mathrm{X}}\bar{\mathrm{X}}^\top$. See Appendix B for further details. The v or vech operator is well known by empirical finance researchers, for example in multivariate GARCH modelling.

We can now rephrase the LQJD setting in terms of the augmented state vector:

$$X^a = \begin{pmatrix} Z \\ \bar{\mathrm{X}} \\ \underline{\mathrm{X}} \end{pmatrix}. \tag{3.33}$$

[2] If we conceive a square-root process as quadratic of an arithmetic process, then most of the existing affine models (e.g., Heston (1993)) are of this kind, with l_i corresponding to the multiplicator of the stock price logarithm.

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

$$dX^a = \mu^a (X_t^a)\, dt + \sigma^a (X_t^a)\, dW_t + dJ_t^a, \tag{3.34}$$

where:

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

All matrices $\mathbf{H}_0$, $\mathbf{H}_1$, $\mathbf{K}_0$, $\mathbf{K}_1$ are explicitly given in terms of the drift and covariance matrices of the initial state vectors $\bar{\text{X}}$ and $\underline{\text{X}}$ of the LQJD model in Appendix B. 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:

Proposition 3 *The standard and extended transforms in an n-factor LQJD model are equal to the transforms in an* $(N+n)$*-factor* $\left(N = \frac{1}{2}m\,(m+1)\right)$ *AJD model, where the state vector is augmented by an additional* $N \times 1$ *pseudo state vector* $Z = v\left(\bar{X}\bar{X}^\top\right)$.

Proof. See Appendix C. ■

Note that the equivalence relationship of AJD and LQJD classes does not hold for general payoffs. Neither does the technique of change of variables 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 precedure. This, however, does not hold in general. For example, the equivalence relationship breaks down when the pseudo factors correspond to exponentials of the initial state variables.

Proposition 3 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 at all, and demonstrated how an LQJD model can be transformed into an AJD one in an *automatic* manner. The above proposition concerns both a theoretical and a numerical equivalence. Indeed we show in the proof that the underlying ODEs are rigorously the same in both settings. This implies that numerical schemes necessary for computing their solutions will deliver exactly the same results. To illustrate this equivalence, we present in Appendix D an example of affine reformulation of a two-factor LQ stochastic volatility model, which is to be discussed in Section 5. Besides, a straightforward consequence of Proposition 3 is that it allows applying the specification analysis (with slight modifications due to Assumption 6) developed by Dai and Singleton (2000) once quadratic term structure models are reformulated as affine term structure models.

Nevertheless, residing in the LQJD setting has advantages of its own. An obvious one is its parsimony over the AJD models: for the *same* number of factors, an LQJD model has the extra capacity to accommodate effects that cannot be handled in an AJD model without introducing the pseudo-factors Z. Put it differently, N additional factors have to be introduced for drawing up equivalence between an n-factor LQJD model and its AJD counterpart. Apparently, N grows fast with m, e.g., for $m = 1$, $N = 1$; $m = 2$, $N = 3$; $m = 3$, $N = 6$; etc.. Furthermore, since the ODEs satisfied by l_i are the same in both frameworks, and the procedures for solving the Riccati systems are standard, it costs little but rewards much by shifting directly to LQJD modelling.

4 Option pricing in the LQJD setting

In this section we show how future payoffs of given types can be priced in the LQJD framework. Results are akin to other results obtained for asset pricing using transform analysis (see e.g. Duffie, Pan and Singleton (2000) and Carr and Madan (1999)).

4.1 The state-price density

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 one of the following three terms: the state-price density, the numéraire, or the market price of risk.

We use the state-price density to pin down the issue. Suppose the data generating process under the measure $\mathbb{P}$ is that of (2.1). We define the state-price density ξ_t as:

$$\xi_t = \exp\left(-\int_0^t R^{\mathbb{P}}(X_s, s)\, ds\right) e^{g_\xi(X_t,t)}, \tag{4.1}$$

where

$$\begin{aligned} g_\xi(X_t,t) &= \frac{1}{2}X_t^\top \Lambda_\xi X_t + b_\xi^\top(t) X_t + c_\xi(t) \\ &= \frac{1}{2}\bar{\mathrm{X}}^\top A_\xi \bar{\mathrm{X}} + k_\xi^\top(t)\bar{\mathrm{X}} + l_\xi^\top(t)\underline{\mathrm{X}} + c_\xi(t), \end{aligned}$$

satisfies the PIDE (3.14). Without loss of generality, we assume $\xi_0 = 1$, which gives the initial condition for the PIDE. By Lemma 1, ξ is a positive $\mathbb{P}$-martingale. Furthermore, by restricting ξ_t to be exponential LQ in X_t, we have ensured that the structure of the state vector remains LQ under the new measure $\mathbb{Q}$.

The equivalent risk-adjusted measure $\mathbb{Q}$ is defined via:

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_t = \frac{\xi_T}{\xi_t}. \tag{4.2}$$

Let:

$$W_t^{\mathbb{Q}} = W_t^{\mathbb{P}} - \int_0^t \sigma(X_s,s)^\top \left[\Lambda_\xi(s) X_s + b_\xi(s)\right] ds. \tag{4.3}$$

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

Lemma 3 *Provided that all technical integrability conditions are satisfied, $\xi W^{\mathbb{Q}}$ is a $\mathbb{P}$-martingale.*

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

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

where $\left\langle \xi, W^{\mathbb{Q}} \right\rangle^c$ denotes the continuous part of $\left\langle \xi, W^{\mathbb{Q}} \right\rangle$. Since $W^{\mathbb{P}}$ and ξ 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. ■

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

$$M^{\mathbb{Q}} = N_t^{\mathbb{P}} - \int_0^t \theta_1\left(l_\xi\right) \lambda^{\mathbb{P}} ds. \tag{4.4}$$

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, $\xi M^{\mathbb{Q}}$ 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:

$$dX_t = \mu^{\mathbb{Q}}\left(X_t, t\right) dt + \sigma\left(X_t, t\right) dW_t^{\mathbb{Q}} + dJ_t^{\mathbb{Q}}, \tag{4.5}$$

with the drift being:

$$\mu^{\mathbb{Q}}\left(X_t, t\right) = \mu\left(X_t, t\right) + \Omega\left(X_t, t\right)\left[\Lambda_\xi\left(t\right) X_t + b_\xi\left(t\right)\right], \tag{4.6}$$

and the jump intensity:

$$\lambda^{\mathbb{Q}}\left(X_t, t\right) = \theta_1\left(l_\xi\right) \lambda^{\mathbb{P}}\left(X_t, t\right). \tag{4.7}$$

The diffusion part remains unchanged.

One may now easily infer from (4.5) the market price of risk relative to the $\mathbb{Q}$ drift $\mu^{\mathbb{Q}}\left(X_t, t\right)$ and the numéraire under $\mathbb{Q}$. It can also be shown that both have incorporated nonlinearity as well as jumps in their structures.

Since the state-price density is obtained explicitly, one may estimate jointly the objective and the risk-adjusted measures in the LQJD settings and extract information content from the options market. An analysis of this kind can be found in Chernov and Ghysels (2000), where the Heston model is applied.

4.2 The expected present value of a truncated future value

We now lend some space to the computation of the present value of a truncated future value, which is especially relevant to option pricing problems and is defined, in general, as follows:

$$G_{g_1,g_2}(k; X_t, t, T) = E_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T R(X_s, s)\, ds\right) e^{g_1(X_T,T)} \mathbb{I}_{\{g_2(X_T,T)\leq k\}}\right], \tag{4.8}$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function, and the superscript $\mathbb{Q}$ denote the risk-adjusted measure. For ease of exposition, we will use g_{1t} for $g_1(X_t, t)$ and suppress the arguments (X_t, t, T) in the following.

The equation (4.8) has to be transformed before results from previous sections on standard and extended transforms can be applied. Two transform procedures exist for such purpose: the Fourier-Stieltjes transform of $G_{g_1,g_2}(k)$, defined as:

$$\mathcal{G}_{g_1,g_2}(v) = \int_{\mathbb{R}} e^{ivk} dG_{g_1,g_2}(k), \tag{4.9}$$

and the generalized Fourier transform, defined as:

$$\mathbb{G}_{g_1,g_2}(v) = \int_{\mathbb{R}} e^{ivk} G_{g_1,g_2}(k)\, dk. \tag{4.10}$$

The Fourier-Stieltjes transform (4.9) has been used in Duffie, Pan and Singleton (2000), where the transform variable v is restricted to take real values, and it is shown that:

$$\mathcal{G}_{g_1,g_2}(v) = \phi^s(g_{1T} + ivg_{2T}). \tag{4.11}$$

We can easily extend v to the complex domain: $v = v_r + iv_i$, where v_r, $v_i \in \mathbb{R}$ and $i^2 = -1$. By choosing properly the imaginary part of v such that $e^{ivk}G_{g_1,g_2}(k)$ vanishes as $k \to \pm\infty$, we can establish an equivalence relationship between $\mathcal{G}_{g_1,g_2}(v)$ and $\mathbb{G}_{g_1,g_2}(v)$:

$$\mathbb{G}_{g_1,g_2}(v) = -\frac{1}{iv}\mathcal{G}_{g_1,g_2}(v). \tag{4.12}$$

Then, by the Fourier inversion formula and the facts that $G_{g_1,g_2}(k)$ is real and thus $\mathbb{G}_{g_1,g_2}(v)$ is odd in its imaginary part and even in its real part, we have

$$G_{g_1,g_2}(k) = \frac{e^{v_i k}}{\pi}\int_0^\infty \mathrm{Re}\left[e^{-iv_r k}\mathbb{G}_{g_1,g_2}(v)\right] dv_r. \tag{4.13}$$

See also Proposition 4 of Leippold and Wu (2002) for similar results.

One may consider the generalized Fourier transfrom of $G_{g_1,g_2}(k)$ as the Fourier transform of $e^{-v_i k}G_{g_1,g_2}(k)$, which corresponds to the modified call price in Carr and Madan (1999). The factor $e^{-v_i k}$ is applied such that the modified price satisfies the integrability conditions. One may check that $v_i = -\alpha$, where α is the dampening coefficient in Carr and Madan (1999).

In practice, one may choose between the Fourier-Stieltjes transform and the generalized Fourier transform. The advantage of the generalized Fourier transform is that it is much more efficient in terms of numerical computations. In some cases, as we will show in the following section, the option prices are obtained by one integration, whereas using the Fourier-Stieltjes transform involves two. Moreover, one may easily apply the fast Fourier transform (FFT) algorithm once the generalized Fourier transform is obtained. See, for instance, Carr and Madan (1999) for details on how this could be done.

However, to apply the generalized Fourier transform, one always has to consider a proper choice for v_i for numerical efficiency. Carr and Madan (1999) give a heuristic discussion about the choice of v_i in the context of pricing simple European options. They suggest to find bounds on v_i such that higher moments of the stock price $S_t = e^{s_t}$ are finite and that singularities are avoided. However, the most efficient value for v_i within these bounds remains a choice of experience.

4.3 The generalized Fourier transform of some payoffs

The following table presents examples of the generalized Fourier transforms of some option prices:

——Table I. Fourier transform of some payoffs ——

One may have noticed that all the transforms are linked to the standard transform of the stock price logarithm $\phi^s\left((iv+1)\,s_T\right)$, which can be computed by Proposition 1. It is also shown that simple European calls and puts have exactly the same transforms. The difference, however, lies in their restrictions on v_i: to obtain the price of a call, the transform has to be integrated along a contour parallel to the v_r-axis and in the negative v_i part of the v-plane, and for a put the contour must be above $v_i = 1$.

A similar table can be found on page 37 of Lewis (2000). There the transform is carried out for the terminal stock price logarithm $s_T = \ln S_T$, whereas here the transform is for the strike price logarithm $k = \ln K$. The resulting transforms are identical, except for the expectation $E[\cdot]$ that has to be taken through the standard transofrm of the stock price logarithm $\phi^s\left((iv+1)\,s_T\right)$ in this context. This is not surprising, for in these claims s_T and k are homogeneous of degree one. The difference, again, lies in the restrictions on v_i: for simple European calls and puts the restrictions are reversed here versus those in Lewis (2000); for covered calls, the restrictions are the same.

We have left out from Table I a special claim that does not exist in the market but is of practical importance numerically: the 'out-of-the-money' (OTM) claims, defined as:

$$
\begin{aligned}
(4.14) \qquad C(k) &= E_t\left[e^{-\int_t^T R_s ds}\left(\left(e^k - e^{g_{2T}}\right)\mathbb{I}_{g_{2T}<k}\right)\right]\mathbb{I}_{g_{2t}>k} \\
&\quad + E_t\left[e^{-\int_t^T R_s ds}\left(\left(e^{g_{2T}} - e^k\right)\mathbb{I}_{-g_{2T}<-k}\right)\right]\mathbb{I}_{-g_{2t}>-k}.
\end{aligned}
$$

As Carr and Madan (1999) have pointed out, options of short maturities approach their non-analytic intrinsic values quickly, and the resulting integrands in the Fourier inversion are highly oscillatory. It is therefore computationally more efficient to consider claims with only time values, such as those of (4.14). The prices of 'in-the-money' (ITM) claims can then be derived via parity relationships such as the Put-Call Parity.

To compute the generalized Fourier transform of (4.14), we can first write it as:

$$\begin{aligned} C(k) &= \left[e^{k} G_{0,g_2}(k) - G_{g_2,g_2}(k)\right] \mathbb{I}_{g_{2t}>k} \\ &\quad + \left[G_{g_2,-g_2}(-k) - e^{k} G_{0,-g_2}(-k)\right] \mathbb{I}_{-g_{2t}>-k}. \end{aligned}$$

The following lemma states the transform of $G_{g_1,g_2}(k)\,\mathbb{I}_{\{g_{3t}\geq k\}}$, where $g_i, i = 1,2,3$ are all LQ functions:

Lemma 4 *Provided all technical integration conditions are satisfied, the generalized Fourier transform of* $G_{g_1,g_2}(k)\,\mathbb{I}_{\{g_{3t}\geq k\}}$ *is:*

$$\begin{aligned} \mathbb{G}_{g_1,g_2,g_3}(v) &= \int_{\mathbb{R}} e^{ivk} G_{g_1,g_2}(k)\, \mathbb{I}_{\{g_{3t}\geq k\}} dk \\ &= \frac{1}{iv}\left[e^{ivg_{3t}} G_{g_1,g_2}(g_{3t}) - G_{g_1+ivg_2,g_2}(g_{3t})\right]. \end{aligned} \tag{4.15}$$

Proof. The result is easily derived by applying Fubini Theorem. ■

Note that $G_{g_1,g_2}(k)$ is actually a limiting case of $G_{g_1,g_2}(k)\,\mathbb{I}_{\{g_{3t}\geq k\}}$, taken at $g_{3t}\to\infty$. Moreover, $G_{g_1,g_2}(y)$ converges to $\phi^s(g_{1T})$ as $y\to\infty$. It can be checked that, by choosing an appropriate value for the imaginary part of v, $e^{ivg_{3t}}G_{g_1,g_2}(g_{3t})$ vanishes in the limits and (4.15) reduces to (4.12).

The generalized Fourier transform can now be computed for an OTM claim

Proposition 4 *The generalized Fourier transform of an OTM claim is:*

$$\begin{aligned} \mathbb{C}(v) &= \int_{\mathbb{R}} e^{ivk} C(k)\, dk \\ &= \frac{1}{1+iv} e^{(1+iv)g_{2t}} \phi^s(0) - \frac{1}{iv} e^{ivg_{2t}} \phi^s(g_{2T}) - \frac{1}{v^2 - iv} \phi^s((1+iv)\, g_{2T}). \end{aligned} \tag{4.16}$$

Proof. The result of Lemma 4 is directly applicable:

$$\begin{aligned} \mathbb{C}(v) &= \frac{1}{1+iv}\left[e^{(1+iv)g_{2t}} G_{0,g_2}(g_{2t}) - G_{(1+iv)g_2,g_2}(g_{2t})\right] \\ &\quad - \frac{1}{iv}\left[e^{ivg_{2t}} G_{g_2,g_2}(g_{2t}) - G_{(1+iv)g_2,g_2}(g_{2t})\right] \\ &\quad + \frac{1}{iv}\left[G_{(1+iv)g_2,-g_2}(-g_{2t}) - e^{ivg_{2t}} G_{g_2,-g_2}(-g_{2t})\right] \\ &\quad - \frac{1}{1+iv}\left[G_{(1+iv)g_2,-g_2}(-g_{2t}) - e^{(1+iv)g_{2t}} G_{0,-g_2}(-g_{2t})\right]. \end{aligned}$$

Further note that:

$$G_{g_1,g_2}(k) + G_{g_1,-g_2}(-k) = \phi^s(g_{1T}).$$

A simplification yields the result. ■

5 A two-factor stochastic volatility model: a numerical example

In this section we consider a simple 2-factor stochastic volatility model as a numerical application of LQJD modelling.

5.1 The stochastic dynamics of the state vector

Under the risk-adjusted measure, the instantaneous variance, V_t, is specified to be the square of the sum of two random variables such that it is guaranteed to be positive:

$$V_t = 2\left(X_{1t} + X_{2t}\right)^2, \tag{5.1}$$

where the two factors are both mean-reverting processes:

$$d\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} \kappa_1\left(\theta - X_{1t}\right) \\ -\kappa_2 X_{2t} \end{pmatrix} dt + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} d\begin{pmatrix} W_{1t} \\ W_{2t} \end{pmatrix}, \tag{5.2}$$

with W_{1t} being independent of W_{2t}. We force the long-term mean of X_2 to be zero, such that it acts as a correction to the driving factor X_1. One may then view the instantaneous volatility $\sqrt{V_t}$ as behaving mostly like X_1, which captures major information contents on the market, with small extra variations from X_2, which reflect small frequent shifts from time to time.

To ensure that the correlation between the instantaneous variance and the stock return processes is a constant, we let $s_t = \ln S_t$, where S_t is the stock price, and:

$$ds_t = \left(r - \delta - \frac{1}{2}V_t - \lambda\bar{m}\right)dt + \left(X_1 + X_2\right)\left[\rho_1 dW_{1t} + \rho_2 dW_{2t} + \sqrt{2 - \rho_1^2 - \rho_2^2}\, dW_{3t}\right] + dJ_t, \tag{5.3}$$

where:

$$\begin{aligned} r &= \text{constant risk-free interest rate,} \\ \delta &= \text{continuously compounded dividend rate,} \\ \lambda &= \text{jump arrival intensity,} \\ \bar{m} &= \text{jump risk premium.} \end{aligned}$$

The correlation, $\rho_{s,V}$, between the price logarithm and the instantaneous variance is then:

$$\rho_{s,V} = \frac{\rho_1\sigma_1 + \rho_2\sigma_2}{\sqrt{2\left(\sigma_1^2 + \sigma_2^2\right)}}. \tag{5.4}$$

An interesting property of the model is that, in contrast to existing two-factor models, the diffusion term in (5.3) is allowed to be either positive or negative. We do not think it brings about any problem, for on the one hand there is indeed an equilibrium economy to support such dynamics[3], and on the other hand the correlation structure between the price logarithm and the factors has more variations, which we consider as an advantage.

[3] We thank J. Detemple for pointing this out to us.

To see this point, let:

$$\rho_{s,X_i}(t) = \frac{(X_{1t} + X_{2t})\,\rho_i}{\sqrt{V_t}} \tag{5.5}$$

be the correlation of s_t and X_{it}. Obviously, the sign of $\rho_{s,X_i}(t)$ now depends on the sign of $(X_{1t} + X_{2t})$. In practice, we usually expect ρ_i to be negative, which will be verified once we calibrate the model to a real data set. Assume this is true for the moment, and consider the following cases:

1. $(X_{1t} + X_{2t}) > 0$ at time t, and the correlation $\rho_{s,X_i}(t)$ is negative. Therefore, an increase (decrease) in magnitude of s_t will lead to a decrease (increase) in $(X_{1t} + X_{2t})$ and thus V_t, which is indeed consistent with common thinking that a price fall (soar) will drive up (down) the volatility.

2. $(X_{1t} + X_{2t}) \leq 0$ at time t, and the correlation $\rho_{s,X_i}(t)$ is positive. Since X_1 is quickly reverting to a positive number θ, it is expected that in this case, the absolute value of $(X_{1t} + X_{2t})$ is likely to be small, and the market is in a *quiet* state. Now:

   (a) If s_t decreases in the next instant, $(X_{1t} + X_{2t})$ will decrease as well. And, since $(X_{1t} + X_{2t})$ is negative, V_t will actually increase, which is again consistent with common sense.

   (b) If s_t increases in the next instant, $(X_{1t} + X_{2t})$ will increase as well. Two scenarios might occur here: (i) s_t soars by a small amount such that $(X_{1t} + X_{2t})$ actually decreases in absolute terms, then V_t decreases and the market continues in the quiet state; but (ii) s_t might increase by a large scale such that $(X_{1t} + X_{2t})$ increases in absolute terms. Then V_t increases and the market turns towards a more *disturbed* state.

3. Of course, it is also possible that $(X_{1t} + X_{2t})$ is significantly negative. Mimicking the analysis in the last point, we see that in this case increases in s_t will probably bring the market back to normal, i.e., $(X_{1t} + X_{2t})$ and V_t will be pulled back to zero; and only further decreases in s_t or increases of extreme scales in s_t would push $(X_{1t} + X_{2t})$ and V_t away from zero, which probably corresponds to a market turmoil.

Since both X_i's are modelled as Ornstein-Uhlenbeck (OU) processes, and since they are independent, it is actually possible to compute explicitly the probability of $X_{1T} + X_{2T} < 0$. Knowing that, for an OU process X with parameters $\begin{pmatrix} k & \theta & \sigma \end{pmatrix}$,

$$X_T \mid \mathcal{F}_t \sim \mathcal{N}\left((X_t - \theta)\,e^{-k(T-t)} + \theta, \frac{\sigma^2}{2k}\left(1 - e^{-2k(T-t)}\right)\right), \tag{5.6}$$

we have:

$$X_{1T} + X_{2T} \mid \mathcal{F}_t \sim \mathcal{N}\left(\eta, \xi^2\right), \tag{5.7}$$

where:

$$\eta = X_{1t}e^{-k_1(T-t)} + X_{2t}e^{-k_2(T-t)} + \theta_1\left(1 - e^{-k_1(T-t)}\right) \tag{5.8}$$

$$\xi^2 = \frac{\sigma_1^2}{2k_1}\left(1 - e^{-2k_1(T-t)}\right) + \frac{\sigma_2^2}{2k_2}\left(1 - e^{-2k_2(T-t)}\right) \tag{5.9}$$

The probability that $X_{1T} + X_{2T}$ is negative is then:

$$P\left(X_{1T} + X_{2T} < 0\right) = \mathbf{N}\left(\frac{-\eta}{\xi}\right), \tag{5.10}$$

where $\mathbf{N}(\cdot)$ is the cumulative distribution function of the standard normal distribution.

5.2 Other related models

Our LQJD model extends the Stein and Stein model in three directions by introducing two volatility factors, correlation and jumps. In Stein and Stein (1991), the volatility of the stock price is modelled as the square of an OU process X with parameters $\begin{pmatrix} k & \theta & \sigma \end{pmatrix}$, which is assumed to be independent of s. A correlated version of the Stein and Stein model with closed form solutions for call option prices can be found in Schöbel and Zhu (1999).

A simple application of Ito's lemma on $V_t = X_t^2$ reveals that

$$dV_t = 2k\left[\left(\frac{\sigma^2}{2k} + \theta X_t\right) - V_t\right]dt + 2X_t\sigma_t dW_{1t}. \tag{5.11}$$

When $\theta \equiv 0$, the correlated Stein and Stein model looks very similar to that of Heston (1993) where V_t is a square-root process. However, it is obvious that if the volatility of the stock price is modelled as the square of an OU process, the diffusion of the volatility factor plays an important role in the drift of V_t, while in the Heston model, the diffusion of the volatility factor, now V_t itself, has nothing to do with the drift of V_t. This property also holds asymptotically, namely on the stationary distribution of the variance process. In the correlated Stein and Stein case, the stationary distribution is a chi-square distribution with mean dependent on the diffusion term of the volatility factor, while in the Heston case it is a Gamma distribution with mean independent of the diffusion term. We believe that the former is more intuitively appealing, for it is indeed expected that, when volatility of volatility is high, the level of the volatility itself is large as well. Clearly the scaled integrated variance $\int_t^T V_s ds/(T-t)$ will also inherit this dependence feature for either finite or infinite maturities.

Preliminary calibration exercise indicates that, for the data set we have used, the long term level θ of the volatility factor is essentially zero. The model is then almost identical to the Heston one and is short of interest for subsequent comparisons. For a detailed discussion of the boundary behavior of the instantaneous volatility at $V_t = 0$, see Schöbel and Zhu (1999).

Our aim is then to compare the performance of multi-factor volatility models. The first two factor volatility model we take for comparison is the affine model proposed by Duffie, Pan, and Singleton (2000). The instantaneous variance V_t is now made mean-reverting to a stochastic long-term trend $\bar{V}_t$ and is correlated with s_t at a constant rate ρ:

(5.12)

$$d\begin{pmatrix} V_t \\ \bar{V}_t \\ s_t \end{pmatrix} = \begin{pmatrix} \kappa_1\left(\bar{V}_t - V_t\right) \\ \kappa_2\left(\theta - \bar{V}_t\right) \\ r - \delta - \frac{1}{2}V_t \end{pmatrix} dt + \begin{pmatrix} \sigma\sqrt{1-\rho^2}\sqrt{V_t} & 0 & \sigma\rho\sqrt{V_t} \\ 0 & \sigma_0\sqrt{\bar{V}_t} & 0 \\ 0 & 0 & \sqrt{V_t} \end{pmatrix} d\begin{pmatrix} W_{1t} \\ W_{2t} \\ W_{3t} \end{pmatrix}.$$

The model resembles ours in that the instantaneous variance process has constant correlation with the price logarithm and that a stochastic factor has been identified to capture the long-term level of the variance. The difference between V_t and $\bar{V}_t$ can thus be considered as the correction factor, which we try to model with X_2. However, besides being affine, the model also differs from ours in the correlation structure between s_t and the volatility factors. In this case, the correlation is fully captured by V_t alone, and the long-term level of the volatility is allowed to drift around freely. Our calibration results in the following reveal that such specification dampens the benefits of introducing a second factor for the volatility.

A more closely related model to ours is the affine model of Bates (2000), where the two factors are modelled as square-root processes:

$$dV_{it} = \kappa_i\left(\theta_i - X_{it}\right)dt + \sigma_i\sqrt{V_{it}}dW_{it}, \qquad i = 1, 2, \tag{5.13}$$

each having a constant correlation ρ_i with the stock price logarithm, and the instantaneous variance is simply the sum of factors: $V_t = V_{1t} + V_{2t}$.

Several differences exist between this affine model and our quadratic model. First, since the volatility factors in the Bates model have the same structure, they are not clearly identifiable. This gives rise to confusion in both estimating and interpreting the results. For instance, the magnitudes of long term means θ_i and the mean-reverting speeds κ_i alternate in order completely when different estimation methods are applied. See Bates (2000) (e.g., Table 2 on page 203 and Table 6 on page 215). To amend for this problem, we restrict θ_2 to be zero as in our two-factor LQJD model. Such restriction assigns the same role to V_2 as that of X_2 in our model, and facilitates meaningful comparisons of the fitted parameters.

Second, the Bates' model is essentially linear, while ours is nonlinear. Obviously, the factors in Bates (2000) correspond to the squares of our factors, i.e. $V_i \sim X_i^2$ for $i = 1, 2$. What is missing is the cross products of the factors, X_1X_2, which can be considered as an additional factor by Proposition 3. Considering that nonlinearity has been documented in many empirical studies, the existence of such factor in our model seems to a great advantage as will also be revealed in the numerical section of this paper.

The third difference lies in the correlation structure. The correlation between the price logarithm and the instantaneous variance in Bates' model is:

$$\rho_{s,V}^B(t) = \frac{\rho_1\sigma_1V_{1t} + \rho_2\sigma_2V_{2t}}{\sqrt{V_t\left(\sigma_1^2V_{1t} + \sigma_2^2V_{2t}\right)}}, \tag{5.14}$$

which is not a constant. Although the time varying feature of $\rho_{s,V}^B(t)$ is desirable, the facts that it stays negative over time and that the correlations between the price logarithm and the factors are constant make the model less versatile in handling different states of market as compared to ours.

To our best knowledge the three multi-factor volatility models we have discussed so far represent the *status quo* of research in modelling stochastic volatility. We note that Leippold and Wu (2002) also mention how stochastic volatility can be modelled in the quadratic class under the stringent assumption that the volatility and the stock price are independent of each other. It has been shown in several works that a negative correlation, in particular, leads to a downward volatility skew, which is related to the premium in option prices for large down-side risk that has existed since the 1987 crash. See, for instance, Renault and Touzi (1996) for further discussions on the relationship between the correlation and the shape of implied volatility, and Bates (1991) for relevant empirical evidence. The independence assumption thus constitutes an intrinsic drawback for modelling stochastic volatility.

In the following calibration exercise, we will use the one-factor affine stochastic volatility model developed by Heston (1993) as benchmark and confront all models with it in various aspects. Such exercise would reveal, in hierarchical order, the necessity of having the second factor in the model, the desirability of a more flexible correlation structure between the state variables, and the advantages of adding nonlinearity in the structure of the state vector.

We note that the models under discussion can be further extended to include jumps (with even stochastic intensity of arrivals, either linear or nonlinear) in the state vector. To keep matters at bay, we present estimation results for models where jumps are only affecting s_t. The jump component has a constant arrival time intensity λ with jump size distributed as a lognormal variable with mean μ_J and standard deviation σ_J. More sophisticated jump specifications are left to future research.

5.3 Data and estimation methods

Our data set of option prices corresponds to prices inferred from implied volatilities of S&P500 index options across strikes and maturities on November 2, 1993 (87 entries in total). This set extracted from Ait-Sahalia and Lo (1998) is exactly the same as the one studied in Duffie, Pan and Singleton (2000). This enables us to use their estimates as guidance in calibration and as benchmark for comparison. We also opt for their choice of setting the interest rate $r = 3.19\%$, the dividend rate $\delta = 0$, and the jump risk premium

$$\bar{m} = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1,$$

ensuring risk-neutrality. Figure 1 displays the market implied volatility smiles for all maturities.

—— Figure 1. Market implied smiles ——

The generalized Fourier transform of puts and calls are readily available from Table I. Hence the prices are easily obtained by Fourier inversion, and we only need to choose a suitable value for the imaginary part of the transform variable v_i. Using the technique discussed at the end of Section 4.2, we decide that $v_i = 1.5$ for puts (the value of v_i for calls can be determined in the same way). A plot of the modulus of the generalized Fourier transform against the imaginary and the real parts of the transform variable gives a visual justification of our choice. Figure 2 corresponds to this plot for the LQ model. It can be seen that the modulus is almost nil for $v_i = 1.5$. Plots of other models are similar and will not be presented:

—— Figure 2. The modulus of the transform integrand ——

Apparently, singularities arise in the neighborhood of zero. In other areas the transform is well-behaved. Therefore, any value of v_i that is far enough from zero would suffice for the Fourier inversion to work, provided that it satisfies the restrictions in Table I. A further observation is that the plot is symmetric about $v_i = 0$. This property is termed *reflection symmetry* in Lewis (2000).

Once the transform variable is determined, we calibrate the theoretical prices to the market data by minimizing the mean square errors (MSE)[4]. Since the parameter sets are of large scale, optimization program may terminate at local optima. We adopt a simple but effective approach to resolve this issue. First, we set appropriate ranges for the individual parameters such that absurd values would not show up. Then we optimize MSEs over randomly chosen initial values inside subsets of the set of parameters, and we select the set of parameters achieving the best MSE after convergence. Thanks to that procedure we are convinced that the resulting parameters are what we are after.

We also use the fitted parameters to compute mean absolute errors (MAD) of prices, as well as MSE and MAD of implied volatilities. Any one of these four criteria can serve as yardstick for measuring the performance of various models. Note however that market practitioners usually work and think in terms of fitted smiles, and will probably favour use of criteria based on implied volatilities.

5.4 Estimation results

The following table presents fitted parameter values for the LQ, Bates, Duffie, Pan and Singleton (DPS), and Heston models, with and without jumps.

—— Table II. Fitted parameters ——

As could have been expected, the most sophisticated models exhibit better fits: the Heston and DPS model stick to the lower spectrum of the performance scale, and the Bates and LQ models compete at the higher end. Moreover, models with jumps significantly improve the fits in all cases. In terms of MSE of fitted prices, our LQ model emerges as best among all considered models. The Bates model gives a better fit in the domain of implied volatilities when calibrated without jumps. However this is no longer the case when jumps are added.

To verify the results, we further divide the fitted option prices into three subsamples according to moneyness. An option is at-the-money (ATM) if its strike price is within the $\pm 2\%$ of the spot, and a call (resp. put) is out-of-the-money (OTM) if its strike is above 102% (resp. below 98%) of the spot. In almost all cases, the LQ model beats others by a fairly large margin. However, the Bates model performs better in terms of MSE in the cases of ATM and OTM options without jumps, winning by 0.007 and 0.0003, respectively, over the LQ model. However, it loses by 0.015 in the case of ITM options.

[4]Programs have been developed in Matlab and rely on the routines ode15s.m for solving the ODEs, quad.m for numerical integration, and lsqnonlin.m for minimization of MSEs. They are available upon request from the authors.

—— Table III. The MSE and MAD ——

It is interesting to observe that the DPS and Heston models are almost identical in all aspects. Our first observation is that the volatility of the *stochastic* long-term mean of the volatility is very close to zero, which in fact renders it to be deterministic for this specific set of data. Re-examination of the structure of the model then leads us to believe that the problem arises out of the fact that the long term mean $\bar{V}$ is made independent of the stock price process: the stochastic volatility factor V_t, which is correlated with s_t, has largely gauged the effects of randomness and hence adding an independent variable would not give rise to further improvements. This, however, also unveils the other side of the story. If there is indeed a need to add an extra factor, such factor should be correlated with the stock price so as to capture what the first volatility factor V_{1t} has missed out. In other words, it should really play a supplementary role in the model. Adding too many factors may well lead to over-fitting and must be cautioned against in the name of the principle of parsimony.

Thus, we are more convinced (than ever) that the second volatility factor, X_{2t}, resp. V_{2t}, in the LQ, resp. Bates, model, should act as a *corrector*: its mean should be small, and its variations be far less drastic than the first volatility factor. Our restriction that $\theta_2 = 0$ is thus appropriate, if not absolutely necessary.

The calibration exercise indeed yields what has been expected. For example, the volatility of the second volatility factor is about $\frac{1}{5}$, resp. $\frac{1}{10}$, of that of the first one in the LQ, resp. Bates, model without jumps, and about $\frac{1}{6}$, resp. $\frac{1}{200}$, with jumps.

Another interesting parameter estimate is the correlation of the correction factor with s_t. In either cases of Bates and LQ models, the correlation is very close to -1. This points to a possibility of modelling the *feedback effect*. In previous sections we have shown that such effect cannot be modelled through the drift matrix in the LQJD setting. However, the feedback factor can be well incorporated through the diffusion matrix. The data set we have used has provided an empirical justification for such presumption.

Furthermore, we note that the mean-reverting speed of the second volatility factor is small. In the LQ model without jumps and the Bates model with jumps, it is virtually zero. This makes the correction factor a pure innovation that is (perfectly) inversely correlated with the stock price logarithm. A by-product of having an innovation of this kind is the justification of *order shifting* in parameter estimates in the Bates model. Recall that V_i in the Bates model corresponds to the square of X_i in the LQ model. By Ito's lemma, it can be shown that the long term mean θ_2 of the second volatility factor is of order $\sigma_2^2 / 2k_2$, where σ_2 and k_2 are the parameters of the LQ model. A small k_2 would then lead to large θ_2 estimates, provided that the LQ model is the correct one.

The following figures give a visual presentation of the fitted results.

—— Figures 3. Implied volatility surface ——

Figure 3 presents the calibrated implied volatility surfaces of the LQ model, with and without jumps. One can easily observe that the 'steepness' of the surface increases with jump components. Moreover, the slopes of the smiles level out as time to maturity increases.

—— Figures 4. 17-day-to-maturity implied volatility smiles ——

Figures 4 plots the fitted 17-day-to-maturity smiles of all models againt the market implied smile. Again, adding jumps significantly improves goodness-of-fit. The LQ and Bates models come quite close to each other. Both of them display under-pricing in the far OTM (ITM) ends, and slight over-pricing ATM. The LQ model has more curvature than the Bates model, by virtue of non-linearity.

—— Figures 5. 318-day-to-maturity implied volatility smiles ——

Figure 5 is similar in spirits to Figure 4, except that the 318-day-to-maturity smiles are plotted. In this case, the jump component no longer plays an important role in fitting performance, and all models calibrate the market fairly well. One can again observe that the LQ model demonstrates more curvature than the rest.

6 Conclusion

We have generalized transform analysis methods existing for the AJD and quadratic classes to the LQJD case. We present in detail the characterization of the LQJD structure, and derive in a rigorous manner restrictions for identification. The standard and extended transforms, as well as pricing formulas for standard financial claims, are also obtained. Furthermore, we show that the system of ODEs, which is identified as an intermediate step to solving the transforms, is a system of non-symmetric Riccati differential equations, and find there exists a standard routine to resolve the pricing problem in the LQJD setting. 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 rigorously before. This result proves to be very strong, 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.

Unlike previous research on quadratic models, all of them oriented towards the modelling of the term structure of interest rates, we consider the issue of multifactor stochastic volatility. Our calibration exercise reveals that incorporating nonlinearity into the instantaneous volatility process significantly improves goodness-of-fit over affine multifactor stochastic volatility models with the same number of factors.

Since our model is very flexible, selecting an appropriate one for the modelling of various stochastic processes in finance will be of great concern. We leave this issue, as well as further econometric analysis of the performance of LQJD modelling, for future research.

Appendix A. Identification restrictions and ODEs

This first appendix gives details about the derivation of the identification restrictions underlying the LQJD modelling as well as the computation leading to the ODEs of Proposition 1

(ODEs of Proposition 2 can be computed along similar lines). They both rely on the PIDE of Lemma 1:

$$R = \dot{g}_1 + \mu^\top \bigtriangledown_x g_1 + \frac{1}{2} tr \left[\left(\bigtriangledown_{xx} g_1 + \bigtriangledown_x g_1 \left(\bigtriangledown_x g_1 \right)^\top \right) \Omega \right] + \lambda \left[\theta_1 \left(l_1 \right) - 1 \right],$$

where all terms should be LQ in x.

A.1 Justification of identification restrictions

The right hand side of the above PIDE contains the following two quantities:

(i) $\mu^\top \bigtriangledown_x g_1$, and

(ii) $tr \left[\bigtriangledown_x g_1 \bigtriangledown_x g_1^\top \Omega \right]$, where

$$\begin{aligned} g_1 &= \frac{1}{2} x^\top \Lambda_1 x + b_1^\top x + c_1 \\ &= \frac{1}{2} \bar{\mathrm{x}}^\top A_1 \bar{\mathrm{x}} + k_1^\top \bar{\mathrm{x}} + l_1^\top \underline{\mathrm{x}} + c_1. \end{aligned}$$

The first, resp. second, one depends on μ, resp. Ω, and will need to be LQ in x to achieve identification. We start by making no assumptions on μ and Ω.

First, we stack μ as:

$$\mu = \begin{pmatrix} \frac{1}{2} \bar{\mathrm{x}}^\top A_{\mu_1} \bar{\mathrm{x}} + k_{\mu_1}^\top \bar{\mathrm{x}} + l_{\mu_1}^\top \underline{\mathrm{x}} + c_{\mu_1} \\ \vdots \\ \frac{1}{2} \bar{\mathrm{x}}^\top A_{\mu_n} \bar{\mathrm{x}} + k_{\mu_n}^\top \bar{\mathrm{x}} + l_{\mu_n}^\top \underline{\mathrm{x}} + c_{\mu_n} \end{pmatrix}_{n \times 1}.$$

Let:

$$\mathcal{A} = \begin{pmatrix} A_{\mu_1} \\ \vdots \\ A_{\mu_n} \end{pmatrix}_{mn \times m}, \quad \mathcal{K} = \begin{pmatrix} k_{\mu_1}^\top \\ \vdots \\ k_{\mu_n}^\top \end{pmatrix}_{n \times m}, \quad \mathcal{L} = \begin{pmatrix} l_{\mu_1}^\top \\ \vdots \\ l_{\mu_n}^\top \end{pmatrix}_{n \times (n-m)}, \quad \mathcal{C} = \begin{pmatrix} c_{\mu_1} \\ \vdots \\ c_{\mu_n} \end{pmatrix}_{n \times 1},$$

then μ can be compactly written as:

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

Now:

$$\begin{aligned} 2\mu^\top \bigtriangledown_x g_1 &= \left[\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}^\top \right) \mathcal{A} \bar{\mathrm{x}} + \mathcal{K} \bar{\mathrm{x}} + \mathcal{L} \underline{\mathrm{x}} + \mathcal{C} \right]^\top \left(\Lambda_1 x + b_1 \right) \\ &= \left[\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}^\top \right) \mathcal{A} \bar{\mathrm{x}} + \mathcal{K} \bar{\mathrm{x}} + \mathcal{L} \underline{\mathrm{x}} + \mathcal{C} \right]^\top \left[\begin{pmatrix} A_1 \bar{\mathrm{x}} \\ 0 \end{pmatrix} + \begin{pmatrix} k_1 \\ l_1 \end{pmatrix} \right]. \end{aligned}$$

For it to be LQ in x, we must have:

$$\left[\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}^\top \right) \mathcal{A} \bar{\mathrm{x}} \right]^\top \begin{pmatrix} A_1 \bar{\mathrm{x}} \\ 0 \end{pmatrix} = 0, \tag{A.1}$$

which leads to $A_{\mu_i} \equiv 0$, for all $i = 1, 2, \cdots, m$. This justifies Assumption 2.

Similarly, we write Ω as:

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

For $tr\left[\bigtriangledown_x g_1 \bigtriangledown_x g_1^\top \Omega\right]$ to be LQ in x, we must have:

$$(\Lambda_1 x)^\top \left[\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}^\top\right)\mathfrak{A}\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}\right) + \mathfrak{K}\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}\right) + \mathfrak{L}\left(\mathbf{I}_n \otimes \underline{\mathrm{x}}\right)\right](\Lambda_1 x) = 0,$$

or individually,

$$x^\top \Lambda_1^\top \left(\mathbf{I}_n \otimes \bar{\mathrm{x}}^\top\right)\mathfrak{A}\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}\right)\Lambda_1 x = 0, \tag{A.2}$$

$$x^\top \Lambda_1^\top \mathfrak{K}\left(\mathbf{I}_n \otimes \bar{\mathrm{x}}\right)\Lambda_1 x = 0, \tag{A.3}$$

$$x^\top \Lambda_1^\top \mathfrak{L}\left(\mathbf{I}_n \otimes \underline{\mathrm{x}}\right)\Lambda_1 x = 0, \tag{A.4}$$

which justify Assumptions 3, 4, and 5, respectively.

Assumptions 6, 7 and 8 are there to rule out possible appearance of $x^{\frac{1}{2}}$ in the process of identifying ODEs.

Assumption 9 on the jump components is justified as follows. Note that the ODEs are identified by imposing $\mathcal{D}\Phi_t = 0$, where the infinitesimal operator $\mathcal{D}$ is defined in (3.3) and Φ_t is exponential-LQ in the state vector as in (3.5). The last component in $\mathcal{D}\Phi_t$ is:

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

Since $\Phi_t(x)$ can be cancelled throughout $\mathcal{D}\Phi_t = 0$, 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 Λ imposed by Assumption 1, the minimal restriction on y is then Assumption 9, namely its first m entries are zeros.

A.2 Obtaining the ODEs: details of computations

Now that the identification restrictions have been derived, we may use them and derive the ODEs of Proposition 1. We proceed term by term and put $g = g_i$, $i = 1, 2$, for notational convenience.

A.2.1 Computation of $\mu^\top \bigtriangledown_x g$

First note that:

$$\bigtriangledown_x g = \Lambda x + b = \begin{pmatrix} A\bar{\mathrm{x}} + k \\ l \end{pmatrix}.$$

Then:

$$\begin{aligned}\mu^{\top}\bigtriangledown_x g &= \begin{pmatrix}\bar{\mu}^{\top} & \underline{\mu}^{\top}\end{pmatrix}\begin{pmatrix}A\bar{\mathrm{x}}+k\\ l\end{pmatrix}\\ &= \bar{\mu}^{\top}\left(A\bar{\mathrm{x}}+k\right)+\underline{\mu}^{\top}l.\end{aligned}$$

Knowing that:

$$\bar{\mu}\left(\bar{\mathrm{x}},t\right)=\bar{\mathcal{K}}\bar{\mathrm{x}}+\bar{\mathcal{C}},$$

and that

$$\underline{\mu}\left(\bar{\mathrm{x}},\underline{\mathrm{x}},t\right)=\frac{1}{2}\left(\mathbf{I}_{n-m}\otimes\bar{\mathrm{x}}^{\top}\right)\underline{\mathcal{A}}\bar{\mathrm{x}}+\underline{\mathcal{K}}\bar{\mathrm{x}}+\underline{\mathcal{L}}\underline{\mathrm{x}}+\underline{\mathcal{C}},$$

we have:

$$\begin{aligned}\bar{\mu}^{\top}\left(A\bar{\mathrm{x}}+k\right) &= \left(\bar{\mathcal{K}}\bar{\mathrm{x}}+\bar{\mathcal{C}}\right)^{\top}\left(A\bar{\mathrm{x}}+k\right)\\ &= \frac{1}{2}\bar{\mathrm{x}}^{\top}\left(\bar{\mathcal{K}}^{\top}A+A\bar{\mathcal{K}}\right)\bar{\mathrm{x}}+\left(\bar{\mathcal{K}}^{\top}k+A\bar{\mathcal{C}}\right)^{\top}\bar{\mathrm{x}}+\bar{\mathcal{C}}^{\top}k,\end{aligned}\tag{A.5}$$

where the quadratic coefficient in the second equality has been symmetrized, and

$$\begin{aligned}\underline{\mu}^{\top}l &= \left[\frac{1}{2}\left(\mathbf{I}_{n-m}\otimes\bar{\mathrm{x}}^{\top}\right)\underline{\mathcal{A}}\bar{\mathrm{x}}+\underline{\mathcal{K}}\bar{\mathrm{x}}+\underline{\mathcal{L}}\underline{\mathrm{x}}+\underline{\mathcal{C}}\right]^{\top}l_1\\ &= \frac{1}{2}\bar{\mathrm{x}}^{\top}\left[\frac{1}{2}\underline{\mathcal{A}}^{\top}\left(l\otimes\mathbf{I}_m\right)+\frac{1}{2}\left(l^{\top}\otimes\mathbf{I}_m\right)\underline{\mathcal{A}}\right]\bar{\mathrm{x}}+\left(\underline{\mathcal{K}}^{\top}l\right)^{\top}\bar{\mathrm{x}}+\left(\underline{\mathcal{L}}^{\top}l\right)^{\top}\underline{\mathrm{x}}+\underline{\mathcal{C}}^{\top}l.\end{aligned}\tag{A.6}$$

The second equality results from the fact that, for matrices $A_{m\times n}$ and $B_{n\times q}$,

$$ABd=\left(d^{\top}\otimes A\right)vec\left[B\right]=\left(A\otimes b^{\top}\right)vec\left[B^{\top}\right],\tag{A.7}$$

where d is $q\times 1$. See equation (8) on page 31 of Magnus and Neudecker (1988). Taking $A=\mathbf{I}_{n-m}$, $B=l_1$ and $d=\bar{x}$ yields the result. The following computations use frequently (A.7).

For identification purpose we would like to represent the sum of (A.5) and (A.6) in the standard form of an LQ function with coefficients $\begin{pmatrix}A^* & k^* & l^* & c^*\end{pmatrix}$. Let:

$$\varpi=\begin{pmatrix}A & k\end{pmatrix},$$

we have:

$$\begin{pmatrix}A^* & k^*\end{pmatrix}=M_{21}^*+M_{22}^*\varpi-\varpi M_{11}^*-\varpi M_{12}^*\varpi,\tag{A.8}$$

where:

$$\begin{aligned}M_{21}^* &= \begin{pmatrix}\frac{1}{2}\underline{\mathcal{A}}^{\top}\left(l\otimes\mathbf{I}_m\right)+\frac{1}{2}\left(l^{\top}\otimes\mathbf{I}_m\right)\underline{\mathcal{A}} & \underline{\mathcal{K}}^{\top}l\end{pmatrix},\\ M_{22}^* &= \underline{\mathcal{K}}^{\top},\\ M_{11}^* &= -\begin{pmatrix}M_{22}^{*\top} & \bar{\mathcal{C}}\\ 0 & 0\end{pmatrix},\\ M_{12}^* &= 0.\end{aligned}$$

Identification of the remaining coefficients is straightforward:

$$\begin{pmatrix} l^* & c^* \end{pmatrix} = \begin{pmatrix} \underline{\mathcal{L}}^\top l & b^\top \mathcal{C} \end{pmatrix} \tag{A.9}$$

where

$$b = \begin{pmatrix} k \\ l \end{pmatrix}, \qquad \mathcal{C}_\mu = \begin{pmatrix} \bar{\mathcal{C}} \\ \underline{\mathcal{C}} \end{pmatrix}.$$

A.2.2 Computation of $tr\left[\nabla_x g \, \nabla_x (g_1)^\top \Omega\right]$

Note that:

$$\begin{aligned}
\nabla_x g \, \nabla_x (g_1)^\top &= \begin{pmatrix} A\bar{\mathrm{x}} + k \\ l \end{pmatrix} \begin{pmatrix} \bar{\mathrm{x}}^\top A_1 + k_1^\top & l_1^\top \end{pmatrix} \\
&= \begin{pmatrix} A\bar{\mathrm{x}}\bar{\mathrm{x}}^\top A_1 + A\bar{\mathrm{x}}k_1^\top + k\bar{\mathrm{x}}^\top A_1 + kk_1^\top & A\bar{\mathrm{x}}l_1^\top + kl_1^\top \\ l\bar{\mathrm{x}}^\top A_1 + lk_1^\top & ll_1^\top \end{pmatrix},
\end{aligned}$$

and

$$\Omega = \begin{pmatrix} \bar{\Omega} & \tilde{\Omega} \\ \tilde{\Omega}^\top & \underline{\Omega} \end{pmatrix},$$

where $\bar{\Omega} = \bar{\sigma}\bar{\sigma}^\top$ is constant, and $\tilde{\Omega}$ and $\underline{\Omega}$ are given by (2.12) and (2.11), respectively. Hence:

$$tr\left[\nabla_x g_0 \, \nabla_x (g_1)^\top \Omega\right] = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4,$$

where, by property of the tr operator as well as (**??**),

$$\begin{aligned}
\Theta_1 &= tr\left[\left(A\bar{\mathrm{x}}\bar{\mathrm{x}}^\top A_1 + A\bar{\mathrm{x}}k_1^\top + k\bar{\mathrm{x}}^\top A_1 + kk_1^\top\right)\bar{\Omega}\right] \\
&= \frac{1}{2}\bar{\mathrm{x}}^\top \left(A\bar{\Omega}A_1 + A_1\bar{\Omega}A\right)\bar{\mathrm{x}} + \left(A\bar{\Omega}k_1 + A_1\bar{\Omega}k\right)^\top \bar{\mathrm{x}} + k_1^\top \bar{\Omega}k, \\
\Theta_2 &= tr\left[\left(A\bar{\mathrm{x}}l_1^\top + kl_1^\top\right)\tilde{\Omega}^\top\right] \\
&= \frac{1}{2}\bar{\mathrm{x}}^\top \left[\left(l_1^\top \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top A + A\tilde{\mathfrak{K}}\left(l_1 \otimes \mathbf{I}_m\right)\right]\bar{\mathrm{x}} + \left[A\tilde{\mathfrak{C}}l_1 + \left(l_1 \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top k\right]^\top \bar{\mathrm{x}} + k^\top \tilde{\mathfrak{C}}l_1, \\
\Theta_3 &= tr\left[\left(l\bar{\mathrm{x}}^\top A_1 + lk_1^\top\right)\tilde{\Omega}\right] \\
&= \frac{1}{2}\bar{\mathrm{x}}^\top \left[\left(l^\top \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top A_1 + A_1\tilde{\mathfrak{K}}^\top\left(l \otimes \mathbf{I}_m\right)\right]\bar{\mathrm{x}} + \left[A_1\tilde{\mathfrak{C}}l + \left(l \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top k_1\right]^\top \bar{\mathrm{x}} + k_1^\top \tilde{\mathfrak{C}}l, \\
\Theta_4 &= tr\left[\left(ll_1^\top\right)\underline{\Omega}\right] \\
&= \frac{1}{2}\bar{\mathrm{x}}^\top \left(l_1^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{A}}\left(l \otimes \mathbf{I}_m\right)\bar{\mathrm{x}} + \left[\left(l^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{K}}^\top l_1\right]^\top \bar{\mathrm{x}} + \left[\left(l^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{L}}^\top l_1\right]^\top \underline{\mathrm{x}} + l_1^\top \underline{\mathfrak{C}}l.
\end{aligned}$$

Again, we wish to represent the above in the standard form of an LQ function with coefficients $\begin{pmatrix} A^{**} & k^{**} & l^{**} & c^{**} \end{pmatrix}$. Let:

$$\varpi_i = \begin{pmatrix} A_i & k_i \end{pmatrix},$$

then:

$$\begin{aligned}\begin{pmatrix} A^{**} & k^{**} \end{pmatrix} &= M_{21}^{**}(l,l_1) \\ &\quad + M_{22}^{**}(l)\varpi_1 - \varpi_1 M_{11}^{**}(l) - \varpi_0 M_{12}^{**}\varpi_1 \\ &\quad + M_{22}^{**}(l_1)\varpi_0 - \varpi_0 M_{11}^{**}(l_1) - \varpi_1 M_{12}^{**}\varpi_0, \end{aligned} \tag{A.10}$$

where

$$\begin{aligned} M_{21}^{**}(l,l_1) &= \begin{pmatrix} \left(l_1^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{A}}\left(l \otimes \mathbf{I}_m\right) & \left(l^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{K}}^\top l_1 \end{pmatrix}, \\ M_{22}^{**}(l_i) &= \left(l_i^\top \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top, \\ M_{11}^{**}(l_i) &= -\begin{pmatrix} M_{22}^{**}(l_i)^\top & \tilde{\mathfrak{C}} l_i \\ 0 & 0 \end{pmatrix}, \\ M_{12}^{**} &= -\begin{pmatrix} \bar{\Omega} \\ 0 \end{pmatrix}. \end{aligned}$$

The remaining terms are:

$$\begin{pmatrix} l^{**} & c^{**} \end{pmatrix} = \begin{pmatrix} \left(l^\top \otimes \mathbf{I}_{n-m}\right)\underline{\mathfrak{L}}^\top l_1 & b^\top \mathfrak{C} b_1 \end{pmatrix}, \tag{A.11}$$

where

$$b_i = \begin{pmatrix} k_i \\ l_i \end{pmatrix}, \qquad \mathfrak{C} = \begin{pmatrix} \bar{\Omega} & \tilde{\mathfrak{C}} \\ \tilde{\mathfrak{C}}^\top & \underline{\mathfrak{C}} \end{pmatrix}.$$

A.2.3 Identifying ODEs in Proposition 1

One can now easily identify the ODEs using (A.8), (A.9), (A.10) and (A.11). For Proposition 1, setting $g = g_1$ in the above equations yields:

$$\begin{aligned} M_{21} &= \begin{pmatrix} \frac{1}{2}\left[\underline{\mathcal{A}}^\top (l_1 \otimes \mathbf{I}_m) + \left(\underline{\mathcal{A}}^\top (l_1 \otimes \mathbf{I}_m)\right)^\top + \left(l_1^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{A}}(l_1 \otimes \mathbf{I}_m)\right] + \left[\theta_1(l_1) - 1\right] A_\lambda - A_R \\ l_1^\top \left(\underline{\mathcal{K}} + \frac{1}{2}\underline{\mathfrak{K}}(l_1 \otimes \mathbf{I}_m)\right) + \left[\theta_1(l_1) - 1\right] k_\lambda^\top - k_R^\top \end{pmatrix}^\top, \\ M_{22} &= \left[\bar{\mathcal{K}} + \tilde{\mathfrak{K}}(l_1 \otimes \mathbf{I}_m)\right]^\top, \\ M_{11} &= -\begin{pmatrix} (M_{22}^s)^\top & \bar{\mathcal{C}} + \tilde{\mathfrak{C}} l_1 \\ 0 & 0 \end{pmatrix}, \\ M_{12} &= -\begin{pmatrix} \bar{\Omega} \\ 0 \end{pmatrix}. \end{aligned}$$

Appendix B. The structure of the augmented state vector

This second appendix explains how we can rewrite any LQJD model as an AJD model with an augmented state vector. As already mentioned, this rewriting can be done in an *automatic* way through use of matrix algebra, which also means that the procedure can be easily implemented in a symbolic calculus package. Before deriving the dynamics of the augmented state vector, we first introduce a special matrix that helps to extract distinct entries from a symmetric matrix.

B.1 The duplication matrix

For a symmetric $m \times m$ matrix A, let $v[A]$ denote the $N \times 1$ $\left(N = \frac{1}{2}m(m+1)\right)$ vector obtained from $vec[A]$ by eliminating the supradiagonal entries of A. For example, for

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

we have:

$$v[A] = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{22} \end{pmatrix}.$$

It turns out that there uniquely exists an $m^2 \times N$ matrix D_m such that:

$$D_m v[A] = vec[A]. \tag{B.1}$$

The matrix D_m is called the duplication matrix in linear algebra, and is easy to build from the algorithm described in Magnus and Neudecker (1988). Its Moore-Penrose (MP) inverse is:

$$D_m^+ = \left(D_m^\top D_m\right)^{-1} D_m^\top. \tag{B.2}$$

Then

$$v[A] = D_m^+ vec[A]. \tag{B.3}$$

The following properties of the duplication matrix prove to be useful in the following computations:

$$\begin{aligned} K_{mm} D_m &= D_m, && \text{(B.4)} \\ D_m^+ D_m &= \mathbf{I}_{N^2}, && \text{(B.5)} \\ D_m D_m^+ &= \frac{1}{2}\left(\mathbf{I}_{m^2} + K_{mm}\right), && \text{(B.6)} \\ D_m D_m^+ (b \otimes A) &= \frac{1}{2}(b \otimes A + A \otimes b), && \text{(B.7)} \end{aligned}$$

where b is a vector, and K_{mn} is the communication matrix that permutes the vectorization of an $m \times n$ matrix A to the vectorization of its transpose. That is,

$$K_{mn} vec[A] = vec\left[A^\top\right].$$

See Theorems 5 and 12 on pages 33 and 49, respectively, of Magnus and Neudecker (1988).

All quadratic forms $\bar{X}^\top A\bar{X}$ can now be written in terms of $Z = v\left[\bar{X}\bar{X}^\top\right]$:

$$\begin{aligned} \bar{X}^\top A\bar{X} &= vec\,[A]^\top\, vec\left[\bar{X}\bar{X}^\top\right] \\ &= vec\,[A]^\top\, D_m Z. \end{aligned} \tag{B.8}$$

B.2 The dynamics of the pseudo state vector Z

Applying Ito's lemma to $Z = v\left(\bar{X}\bar{X}^\top\right)$ yields:

$$dZ_t = v\left[\bar{X}_t\left(d\bar{X}_t\right)^\top + \left(d\bar{X}_t\right)\bar{X}_t^\top + \bar{\Omega}dt\right]. \tag{B.9}$$

The dynamics of $\bar{X}$ is given by (2.6). We have:

$$dZ_t = \mu_Z\left(Z_t, \bar{X}_t\right)dt + \sigma_Z\left(\bar{X}_t\right)dW_t, \tag{B.10}$$

where:

$$\begin{aligned} \mu_Z &= D_m^+\left[\left(\bar{\mathcal{K}}\otimes \mathbf{I}_m + \mathbf{I}_m\otimes\bar{\mathcal{K}}\right) vec\left[\bar{X}_t\bar{X}_t^\top\right] + \left(\bar{\mathcal{C}}\otimes\mathbf{I}_m + \mathbf{I}_m\otimes\bar{\mathcal{C}}\right)\bar{X}_t + vec\left[\bar{\Omega}\right]\right], \\ \sigma_Z &= D_m^+\left[\bar{\sigma}\otimes\bar{X}_t + \bar{X}_t\otimes\bar{\sigma}\right]. \end{aligned}$$

We want to represent μ_Z and $\Omega_Z = \sigma_Z\sigma_Z^\top$ in (affine) terms of Z_t and $\bar{X}_t$:

$$\mu_Z = \mathcal{Q}_Z Z_t + \mathcal{K}_Z\bar{X}_t + \mathcal{C}_Z, \tag{B.11}$$

$$\Omega_Z = \mathfrak{Q}_Z\left(\mathbf{I}_N\otimes Z_t\right) + \mathfrak{K}_Z\left(\mathbf{I}_N\otimes\bar{X}_t\right) + \mathfrak{C}_Z, \tag{B.12}$$

From the properties (B.4)-(B.7) of the duplication matrix as well as (B.3) we get:

$$\begin{aligned} D_m^+\left(\bar{\mathcal{K}}\otimes\mathbf{I}_m + \mathbf{I}_m\otimes\bar{\mathcal{K}}\right) vec\left[\bar{X}_t\bar{X}_t^\top\right] &= 2D_m^+\left(\bar{\mathcal{K}}\otimes\mathbf{I}_m\right)D_m Z_t, \\ D_m^+\left(\bar{\mathcal{C}}\otimes\mathbf{I}_m + \mathbf{I}_m\otimes\bar{\mathcal{C}}\right)\bar{X}_t &= 2D_m^+\left(\bar{\mathcal{C}}\otimes\mathbf{I}_m\right)\bar{X}_t, \end{aligned}$$

Hence:

$$\mathcal{Q}_Z = 2D_m^+\left(\bar{\mathcal{K}}\otimes\mathbf{I}_m\right)D_m, \tag{B.13}$$

$$\mathcal{K}_Z = 2D_m^+\left(\bar{\mathcal{C}}\otimes\mathbf{I}_m\right), \tag{B.14}$$

$$\mathcal{C}_Z = v\left[\bar{\Omega}\right]. \tag{B.15}$$

Similarly,

$$D_m^+\left[\bar{\sigma}\otimes\bar{X}_t + \bar{X}_t\otimes\bar{\sigma}\right] = 2D_m^+\left(\bar{\sigma}\otimes\bar{X}_t\right).$$

Hence

$$\Omega_Z\left(\bar{X}_t\right) = 4D_m^+\left(\bar{\Omega}\otimes\bar{X}_t\bar{X}_t^\top\right)D_m^{+\top}. \tag{B.16}$$

Obviously, $\mathfrak{K}_Z = 0$ and $\mathfrak{C}_Z = 0$. It is, however, difficult to visualize $\mathfrak{Q}_Z$. Nevertheless, for our identification purpose it suffices to know that, for *any* $m \times 1$ vector γ,

$$\mathfrak{Q}_Z\left(\mathbf{I}_N \otimes v\left[\gamma\gamma^\top\right]\right) = \Omega_Z(\gamma). \tag{B.17}$$

For identifying $\mathfrak{Q}_Z$, one may then choose γ to be, say, a vector of ones.

B.3 The augmented drift matrix μ^a

Apparently:

$$\mu^a = \begin{pmatrix} \mu_Z \\ \bar{\mu} \\ \underline{\mu} \end{pmatrix}, \tag{B.18}$$

where μ_Z, $\bar{\mu}$ and $\underline{\mu}$ are given by (B.11), (2.8) and (2.10), respectively. We need to rewrite the quadratic form in $\underline{\mu}$ in terms of Z. Recall that:

$$\underline{\mathcal{A}} = \begin{pmatrix} \underline{A}_1 \\ \vdots \\ \underline{A}_{n-m} \end{pmatrix}.$$

Then the quadratic form in $\underline{\mu}$ can be reformulated:

$$\begin{aligned} \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}_t^\top\right)\underline{\mathcal{A}}\bar{\mathrm{X}}_t &= \frac{1}{2}\begin{pmatrix} \bar{\mathrm{X}}_t^\top \underline{A}_1 \bar{\mathrm{X}}_t \\ \vdots \\ \bar{\mathrm{X}}_t^\top \underline{A}_{n-m} \bar{\mathrm{X}}_t \end{pmatrix} \\ &= \underline{\mathcal{Q}}Z, \end{aligned}$$

where

$$\underline{\mathcal{Q}} = \frac{1}{2}\begin{pmatrix} vec\left[\underline{A}_1\right]^\top \\ \vdots \\ vec\left[\underline{A}_{n-m}\right]^\top \end{pmatrix} D_m. \tag{B.19}$$

It is actually easy to see that, for an arbitrary $m \times 1$ vector γ,

$$\underline{\mathcal{Q}}v\left[\gamma\gamma^\top\right] = \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \gamma^\top\right)\underline{\mathcal{A}}\gamma. \tag{B.20}$$

Summarizing, we have:

$$\mu^a = \mathbf{K}_1 X^a + \mathbf{K}_0, \tag{B.21}$$

where:

$$\text{(B.22)} \qquad \mathbf{K}_1 = \begin{pmatrix} \mathcal{Q}_Z & \mathcal{K}_Z & 0 \\ 0 & \bar{\mathcal{K}} & 0 \\ \underline{\mathcal{Q}} & \underline{\mathcal{K}} & \underline{\mathcal{L}} \end{pmatrix},$$

$$\text{(B.23)} \qquad \mathbf{K}_0 = \begin{pmatrix} \mathcal{C}_Z \\ \bar{\mathcal{C}} \\ \underline{\mathcal{C}} \end{pmatrix}.$$

B.4 The augmented covariance matrix Ω^a

First partition Ω^a as follows:

$$\text{(B.24)} \qquad \Omega^a = \begin{pmatrix} \Omega_Z & \bar{\Omega}_Z & \underline{\Omega}_Z \\ \bar{\Omega}_Z^\top & \bar{\Omega} & \tilde{\Omega} \\ \underline{\Omega}_Z^\top & \tilde{\Omega}^\top & \underline{\Omega} \end{pmatrix}.$$

The matrices Ω_Z, $\bar{\Omega}$, $\underline{\Omega}$ and $\tilde{\Omega}$ are known, but with $\underline{\Omega}$ to be re-written in (affine) terms of the augmented state vector X^a:

$$\begin{aligned} \text{(B.25)} \quad \underline{\Omega} &= \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}^\top\right) \underline{\mathfrak{A}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \underline{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \underline{\mathfrak{L}}\left(\mathbf{I}_{n-m} \otimes \underline{\mathrm{X}}\right) + \underline{\mathfrak{C}} \\ &= \underline{\mathfrak{Q}}\left(\mathbf{I}_{n-m} \otimes Z_t\right) + \underline{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right) + \underline{\mathfrak{L}}\left(\mathbf{I}_{n-m} \otimes \underline{\mathrm{X}}\right) + \underline{\mathfrak{C}}, \end{aligned}$$

where $\underline{\mathfrak{Q}}$ is determined as follows. Recall that the $(i,j)^{th}$ term of $\frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}^\top\right) \underline{\mathfrak{A}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}\right)$ is $\frac{1}{2}\bar{\mathrm{X}}^\top \underline{\mathfrak{A}}_{ij} \bar{\mathrm{X}}$, which equals $\frac{1}{2} vec\left[\underline{\mathfrak{A}}_{ij}\right]^\top D_m Z$. $\underline{\mathfrak{Q}}$ is then:

$$\text{(B.26)} \qquad \underline{\mathfrak{Q}} = \frac{1}{2}\begin{pmatrix} vec\left[\underline{\mathfrak{A}}_{11}\right]^\top D_m & \cdots & vec\left[\underline{\mathfrak{A}}_{1,n-m}\right]^\top D_m \\ \vdots & \ddots & \vdots \\ vec\left[\underline{\mathfrak{A}}_{n-m,1}\right]^\top D_m & \cdots & vec\left[\underline{\mathfrak{A}}_{n-m,n-m}\right]^\top D_m \end{pmatrix}.$$

However, for our reformulation purpose, it suffices to observe that, for *any* $m \times 1$ vector γ,

$$\text{(B.27)} \qquad \underline{\mathfrak{Q}}\left(\mathbf{I}_{n-m} \otimes v\left[\gamma\gamma^\top\right]\right) = \frac{1}{2}\left(\mathbf{I}_{n-m} \otimes \gamma^\top\right) \underline{\mathfrak{A}}\left(\mathbf{I}_{n-m} \otimes \gamma\right).$$

What remains are the covariance matrices $\bar{\Omega}_Z$ and $\underline{\Omega}_Z$, which can be determined to be:

$$\begin{aligned} \text{(B.28)} \qquad \bar{\Omega}_Z &= 2D_m^+\left(\bar{\sigma} \otimes \bar{\mathrm{X}}_t\right)\bar{\sigma}^\top \\ &= 2D_m^+\left(\bar{\Omega} \otimes \bar{\mathrm{X}}_t\right), \end{aligned}$$

and

$$\begin{aligned} \text{(B.29)} \qquad \underline{\Omega}_Z &= 2D_m^+\left(\bar{\sigma} \otimes \bar{\mathrm{X}}_t\right)\underline{\sigma}^\top \\ &= 2D_m^+\left(\tilde{\Omega} \otimes \bar{\mathrm{X}}_t\right). \end{aligned}$$

Using (2.12) for $\tilde{\Omega}$, we have:

$$\text{(B.30)} \qquad \begin{aligned} \underline{\Omega}_Z &= 2D_m^+ \left[\tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}_t\right) + \tilde{\mathfrak{C}}\right] \otimes \bar{\mathrm{X}}_t \\ &= \underline{\mathfrak{Q}}_Z \left(\mathbf{I}_{n-m} \otimes Z_t\right) + \underline{\mathfrak{K}}_Z \left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}_t\right), \end{aligned}$$

where

$$\text{(B.31)} \qquad \underline{\mathfrak{K}}_Z = 2D_m^+ \left(\tilde{\mathfrak{C}} \otimes \mathbf{I}_m\right),$$

and $\underline{\mathfrak{Q}}_Z$ is determined as follows. Note that the $(i,j)^{th}$ term of $\tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \bar{\mathrm{X}}_t\right)$ is $\tilde{\mathfrak{K}}_{ij}\bar{\mathrm{X}}_t$, which is a scalar. Multiplying it by $\bar{\mathrm{X}}$ yields:

$$\left(\tilde{\mathfrak{K}}_{ij}\bar{\mathrm{X}}_t\right)\bar{\mathrm{X}}_t = \left(\bar{\mathrm{X}}_t\bar{\mathrm{X}}_t^\top\right)\tilde{\mathfrak{K}}_{ij}^\top,$$

which is a vector. Therefore it does not change anything if we apply the *vec* operator to it. By the property of the *vec* operator (see also Theorem 2 on page 30 of Magnus and Neudecker (1988)), we have:

$$\left(\bar{\mathrm{X}}_t\bar{\mathrm{X}}_t^\top\right)\tilde{\mathfrak{K}}_{ij}^\top = \left(\tilde{\mathfrak{K}}_{ij} \otimes \mathbf{I}_{n-m}\right) D_m Z.$$

Letting $\underline{\mathfrak{Q}}_Z$ be the following yields the result:

$$\text{(B.32)} \qquad \underline{\mathfrak{Q}}_Z = 2D_m^+ \begin{pmatrix} \left(\tilde{\mathfrak{K}}_{11} \otimes \mathbf{I}_{n-m}\right) D_m & \cdots & \left(\tilde{\mathfrak{K}}_{1,n-m} \otimes \mathbf{I}_{n-m}\right) D_m \\ \vdots & \ddots & \vdots \\ \left(\tilde{\mathfrak{K}}_{m^2,1} \otimes \mathbf{I}_{n-m}\right) D_m & \cdots & \left(\tilde{\mathfrak{K}}_{m^2,n-m} \otimes \mathbf{I}_{n-m}\right) D_m \end{pmatrix}.$$

However, like $(B.27)$, it actually suffices to know that, for *any* $m \times 1$ vector γ,

$$\text{(B.33)} \qquad \underline{\mathfrak{Q}}_Z \left(\mathbf{I}_{n-m} \otimes v\left[\gamma\gamma^\top\right]\right) = 2D_m^+ \left[\tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m} \otimes \gamma\right)\right] \otimes \gamma.$$

We can finally represent the augmented covariance matrix in terms of X^a as:

$$\text{(B.34)} \qquad \Omega^a = \mathbf{H}_1\left(\mathbf{I}_{N+n} \otimes X^a\right) + \mathbf{H}_0.$$

It is easy to determine $\mathbf{H}_0$:

$$\text{(B.35)} \qquad \mathbf{H}_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \bar{\sigma}\bar{\sigma}^\top & \tilde{\mathfrak{C}} \\ 0 & \tilde{\mathfrak{C}}^\top & \underline{\mathfrak{C}} \end{pmatrix}.$$

Writing out $\mathbf{H}_1$ explicitly is difficult, but computing $\mathbf{H}_1\left(\mathbf{I}_{N+n} \otimes \gamma\right)$ is straightforward for any $(N+n) \times 1$ vector γ. The trick to compute this last expression is to partition γ into three blocks:

$$\text{(B.36)} \qquad \gamma = \begin{pmatrix} \gamma_Z \\ \gamma_{\bar{\mathrm{X}}} \\ \gamma_{\underline{\mathrm{X}}} \end{pmatrix},$$

with dimensions $N \times 1$, $m \times 1$, and $(n-m) \times 1$, respectively. $\mathbf{H}_1(\mathbf{I}_{N+n} \otimes \gamma)$ is then obtained by replacing everywhere Z by γ_Z, $\bar{\mathrm{X}}$ by $\gamma_{\bar{\mathrm{X}}}$, and $\underline{\mathrm{X}}$ by $\gamma_{\underline{\mathrm{X}}}$ in $\mathbf{H}_1(\mathbf{I}_{N+n} \otimes X^a)$. Now, if we wish to compute $\mathbf{H}_1$, we only need to take γ equal to a vector of ones.

Appendix C. Proof of Proposition 3

In this section we prove Proposition 3, which states the equivalence relationship between LQJD and AJD classes in terms of their transforms. We aim to check that coefficient matrices in the two classes of models satisfy the same system of ODEs. Given that they also have the same initial conditions, the identification procedures are identical and therefore the two classes are equivalent. We only proceed with the ODEs of the standard transform. The results can also be shown to hold for the extended transform.

Proof. Note that the coefficient matrices to be identified in the LQJD model are $\begin{pmatrix} A & k & l & c \end{pmatrix}$. Their counterparts in the LQJD-transformed AJD model are $\begin{pmatrix} \beta & \alpha \end{pmatrix}$, where $\alpha = c$ and

$$\beta = \begin{pmatrix} \beta_A \\ \beta_k \\ \beta_l \end{pmatrix}, \tag{C.1}$$

with

$$\beta_A = \frac{1}{2} D_m^\top vec\left[A\right], \tag{C.2}$$

$$\beta_k = k, \tag{C.3}$$

$$\beta_l = l. \tag{C.4}$$

We will show that the system of ODEs obtained for $\begin{pmatrix} \beta & \alpha \end{pmatrix}$ using the AJD procedure is identical to the one for $\begin{pmatrix} A & k & l & c \end{pmatrix}$.

First note that, since the specifications of the jumps and the discount rate are the same in the LQJD model and its AJD transformation, they can be neglected without invalidating the arguments of the proof.

The ODEs satisfied by α and β without the jumps and the discount rate are obtained by directly applying (2.5) and (2.6) of Duffie, Pan and Singleton (2000):

$$\frac{d}{d\tau}\alpha = \mathbf{K}_0^\top \beta + \frac{1}{2}\beta^\top \mathbf{H}_0 \beta, \tag{C.5}$$

$$\frac{d}{d\tau}\beta = \mathbf{K}_1^\top \beta + \frac{1}{2}\left(\beta^\top \otimes \mathbf{I}_{N+n}\right) \mathbf{H}_1^\top \beta. \tag{C.6}$$

1. Using (B.23) for $\mathbf{K}_0$ and (B.35) for $\mathbf{H}_0$, we have:

$$\frac{d}{d\tau}\alpha = \mathcal{C}_Z^\top \beta_A + \begin{pmatrix} \bar{\mathcal{C}}^\top & \underline{\mathcal{C}}^\top \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_l \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \beta_k^\top & \beta_l^\top \end{pmatrix} \begin{pmatrix} \bar{\sigma}\bar{\sigma}^\top & \tilde{\mathfrak{C}} \\ \tilde{\mathfrak{C}}^\top & \underline{\mathfrak{C}} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_l \end{pmatrix}.$$

We only need to show that $\mathcal{C}_Z^\top \beta_u = \frac{1}{2} tr\left[A\bar{\Omega}\right]$, for the remaining terms of $d\alpha/d\tau$ are in exact conformity with corresponding terms in $dc/d\tau$. Using (B.15) for $\mathcal{C}_Z$ and (C.2) for β_A as well as (B.3) yields:

$$\begin{aligned}\mathcal{C}_Z^\top \beta_A &= \frac{1}{2} v\left[\bar{\Omega}\right]^\top D_m^\top vec\left[A\right] \\ &= \frac{1}{2} vec\left[\bar{\Omega}\right]^\top \left(D_m D_m^+\right)^\top vec\left[A\right].\end{aligned}$$

Since $D_m D_m^+ = \frac{1}{2}\left(\mathbf{I}_{m^2} + K_m\right)$ by (B.6) and $\bar{\Omega}$ is symmetric,

$$D_m D_m^+ vec\left[\bar{\Omega}\right] = vec\left[\bar{\Omega}\right].$$

Given that $\bar{\Omega}$ and A are symmetric matrices of the same order, by property of the trace operator:

$$tr\left[A\bar{\Omega}\right] = vec\left[\bar{\Omega}\right]^\top vec\left[A\right].$$

Therefore:

$$\mathcal{C}_Z^\top \beta_u = \frac{1}{2} tr\left[A\bar{\Omega}\right].$$

2. To ease identification of the individual ODEs of the system $d\beta/d\tau$, we introduce the canonocal vector $(N+n)\times 1$ vector ε_i whose i^{th} component is one and zero elsewhere. Obviously,

$$\frac{d}{d\tau}\beta_i = \varepsilon_i^\top \frac{d}{d\tau}\beta.$$

From (C.6), we get:

$$\frac{d}{d\tau}\beta_i = \varepsilon_i^\top \mathbf{K}_1^\top \beta + \frac{1}{2}\beta^\top \left[\mathbf{H}_1\left(\mathbf{I}_{N+n}\otimes \varepsilon_i\right)\right]^\top \beta.$$

The first term of the right hand side can be computed explicitly:

$$\mathbf{K}_1^\top \beta = \begin{pmatrix} \mathcal{Q}_Z^\top \beta_A + \underline{Q}^\top \beta_l \\ \mathcal{K}_Z^\top \beta_A + \bar{\mathcal{K}}^\top \beta_k + \underline{\mathcal{K}}^\top \beta_l \\ \underline{\mathcal{L}}^\top \beta_l \end{pmatrix}.$$

The second term $\mathbf{H}_1\left(\mathbf{I}_{N+n}\otimes \varepsilon_i\right)$ can also be computed using the trick described in the ending paragraph of Appendix B. We only need to partition ε_i conformably with β:

$$\varepsilon_i = \begin{pmatrix} \varepsilon_A \\ \varepsilon_k \\ \varepsilon_l \end{pmatrix}_i$$

- For identifying $d\beta_l/d\tau$, note that $i \in [N+m+1, N+n]$ so that $\varepsilon_A = 0$ and $\varepsilon_k = 0$. Thus:

$$\mathbf{H}_1\left(\mathbf{I}_{N+n}\otimes \varepsilon_i\right) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \underline{\mathfrak{L}}\left(\mathbf{I}_{n-m}\otimes \varepsilon_l\right) \end{pmatrix}.$$

This immediately leads to:

$$\frac{d}{d\tau}\beta_l = \underline{\mathfrak{L}}^\top \beta_l + \frac{1}{2}\left(\beta_l^\top \otimes \mathbf{I}_{n-m}\right)\underline{\mathfrak{L}}\beta_l,$$

which is also the ODE system satisfied by l.

- For $d\beta_k/d\tau$, $i \in [N+1, N+m]$ so that $\varepsilon_A = 0$ and $\varepsilon_l = 0$. Thus:

$$\mathbf{H}_1\left(\mathbf{I}_{N+n}\otimes\varepsilon_i\right) = \begin{pmatrix} 0 & 2D_m^+\left(\bar{\Omega}\otimes\varepsilon_k\right) & \underline{\mathfrak{K}}_Z\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right) \\ \left[2D_m^+\left(\bar{\Omega}\otimes\varepsilon_k\right)\right]^\top & 0 & \tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right) \\ \left[\underline{\mathfrak{K}}_Z\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right)\right]^\top & \left[\tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right)\right]^\top & \underline{\mathfrak{K}}\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right) \end{pmatrix},$$

and

$$\frac{1}{2}\beta^\top\left[\mathbf{H}_1\left(\mathbf{I}_{N+n}\otimes\varepsilon_i\right)\right]^\top\beta = (i) + (ii) + (iii) + (iv),$$

where:

  - By (A.7),

$$\begin{aligned} (i) &= \frac{1}{2}\beta_l^\top\left[\underline{\mathfrak{K}}\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right)\right]^\top\beta_l \\ &= \frac{1}{2}\varepsilon_k^\top\left(\beta_l^\top\otimes\mathbf{I}_m\right)\underline{\mathfrak{K}}^\top\beta_l. \end{aligned}$$

  -

$$\begin{aligned} (ii) &= \beta_A^\top\left[2D_m^+\left(\bar{\Omega}\otimes\varepsilon_k\right)\right]\beta_k \\ &= vec\left[A\right]^\top D_m D_m^+\left(\bar{\Omega}\otimes\varepsilon_k\right)\beta_k \\ &= \frac{1}{2}\left[\left(\bar{\Omega}\otimes\varepsilon_k^\top\right)vec\left[A\right] + \left(\varepsilon_k^\top\otimes\bar{\Omega}\right)vec\left[A\right]\right]^\top\beta_k \\ &= \varepsilon_k^\top A\bar{\Omega}\beta_k. \end{aligned}$$

    The second equality is by (C.2), the third by (B.6) and the property of the communication matrix, and the fourth by (A.7).

  -

$$\begin{aligned} (iii) &= \beta_A^\top\left[\underline{\mathfrak{K}}_Z\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right)\right]\beta_l \\ &= \varepsilon_k^\top A\tilde{\mathfrak{C}}\beta_l. \end{aligned}$$

    The result is obtained by using (C.2) for β_A, (B.31) for $\underline{\mathfrak{K}}_Z$, and the same procedure as in (ii).

  - Similarly,

$$\begin{aligned} (iv) &= \beta_l^\top\left[\tilde{\mathfrak{K}}\left(\mathbf{I}_{n-m}\otimes\varepsilon_k\right)\right]^\top\beta_k \\ &= \varepsilon_k^\top\left(\beta_l^\top\otimes\mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top\beta_k. \end{aligned}$$

Finally, note that, by (B.14), (C.2),

$$\mathcal{K}_Z^\top \beta_A = A\bar{\mathcal{C}}.$$

Gathering results:

$$\begin{aligned}\frac{d}{d\tau}\beta_k &= A\bar{\mathcal{C}} + \bar{\mathcal{K}}^\top \beta_k + \underline{\mathcal{K}}^\top \beta_l + \\ &\frac{1}{2}\left(\beta_l^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{K}}^\top \beta_l + A\bar{\Omega}\beta_k + A\tilde{\mathfrak{C}}\beta_l + \left(\beta_l^\top \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top \beta_k,\end{aligned}$$

which is identical to $\frac{d}{d\tau}k$

- For $d\beta_A/d\tau$, $i \in [1, N]$ so that $\varepsilon_k = 0$ and $\varepsilon_l = 0$. Thus:

$$\mathbf{H}_1\left(\mathbf{I}_{N+n} \otimes \varepsilon_i\right) = \begin{pmatrix} \mathfrak{Q}_Z\left(\mathbf{I}_N \otimes \varepsilon_A\right) & 0 & \underline{\mathfrak{Q}}_Z\left(\mathbf{I}_{n-m} \otimes \varepsilon_A\right) \\ 0 & 0 & 0 \\ \left[\underline{\mathfrak{Q}}_Z\left(\mathbf{I}_{n-m} \otimes \varepsilon_A\right)\right]^\top & 0 & \underline{\mathfrak{Q}}\left(\mathbf{I}_{n-m} \otimes \varepsilon_A\right) \end{pmatrix}.$$

The system of ODEs satisfied by β_A is:

$$\frac{d}{d\tau}\beta_A = \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4 + \Theta_5,$$

where:

$$\begin{aligned}\Theta_1 &= \underline{\mathcal{Q}}^\top \beta_l, \\ \Theta_2 &= \mathcal{Q}_Z^\top \beta_A, \\ \Theta_3 &= \left(\mathbf{I}_N \otimes \beta_l^\top\right)\underline{\mathfrak{Q}}_Z^\top \beta_A, \\ \Theta_4 &= \frac{1}{2}\left(\mathbf{I}_N \otimes \beta_l^\top\right)\underline{\mathfrak{Q}}^\top \beta_l, \\ \Theta_5 &= \frac{1}{2}\left(\mathbf{I}_N \otimes \beta_A^\top\right)\mathfrak{Q}_Z^\top \beta_A.\end{aligned}$$

The system of ODEs satisfied by A can be obtained from results of Appendix A:

$$\frac{d}{d\tau}A = \hat{\Theta}_1 + \hat{\Theta}_2 + \hat{\Theta}_3 + \hat{\Theta}_4 + \hat{\Theta}_5,$$

where:

$$\begin{aligned}\hat{\Theta}_1 &= \frac{1}{2}\left[\left(l^\top \otimes \mathbf{I}_m\right)\underline{\mathcal{A}} + \underline{\mathcal{A}}^\top\left(l \otimes \mathbf{I}_m\right)\right], \\ \hat{\Theta}_2 &= \bar{\mathcal{K}}^\top A + A\bar{\mathcal{K}}, \\ \hat{\Theta}_3 &= \left(l^\top \otimes \mathbf{I}_m\right)\tilde{\mathfrak{K}}^\top A + A\tilde{\mathfrak{K}}\left(l \otimes \mathbf{I}_m\right), \\ \hat{\Theta}_4 &= \frac{1}{2}\left(l^\top \otimes \mathbf{I}_m\right)\underline{\mathfrak{A}}^\top\left(l \otimes \mathbf{I}_m\right), \\ \hat{\Theta}_5 &= A\bar{\Omega}A.\end{aligned}$$

Since β_A and A are linked by (C.2), we expect the same relationship exists between Θ_i and its counterpart $\hat{\Theta}_i$ for $i = 1, ..., 5$, i.e.

$$\Theta_i = \frac{1}{2} D_m^\top vec\left[\hat{\Theta}_i\right].$$

This is indeed true, for

$$v\left[\gamma\gamma^\top\right]^\top \Theta_i = \frac{1}{2} v\left[\gamma\gamma^\top\right]^\top D_m^\top vec\left[\hat{\Theta}_i\right],$$

for *any* $m \times 1$ vector γ. For instance, for $i = 1$,

$$\begin{aligned} v\left[\gamma\gamma^\top\right]^\top \Theta_1 &= \left(\underline{\mathcal{Q}} v\left[\gamma\gamma^\top\right]\right)^\top \beta_l \\ &= \frac{1}{2}\left[\left(\mathbf{I}_{n-m} \otimes \gamma^\top\right) \underline{\mathcal{A}} \gamma\right]^\top \beta_l \\ &= \frac{1}{2} \gamma^\top \underline{\mathcal{A}}^\top \left(\beta_l \otimes \mathbf{I}_m\right) \gamma \\ &= \frac{1}{2} v\left[\gamma\gamma^\top\right]^\top D_m^\top vec\left[\underline{\mathcal{A}}^\top \left(\beta_l \otimes \mathbf{I}_m\right)\right], \end{aligned}$$

by (B.20) for the second equality, (A.7) for the third, and (B.8) for the fourth. On the other hand,

$$\frac{1}{2} v\left[\gamma\gamma^\top\right]^\top D_m^\top vec\left[\hat{\Theta}_1\right] = \frac{1}{2} v\left[\gamma\gamma^\top\right]^\top D_m^\top vec\left[\underline{\mathcal{A}}^\top \left(l \otimes \mathbf{I}_m\right)\right].$$

Since γ is arbitrary, we must have $\Theta_1 = \frac{1}{2} D_m^\top vec\left[\hat{\Theta}_1\right]$. The remaining equalities can be derived in exactly the same manner.

Thereby we have established that the ODEs obtained for $\begin{pmatrix} A & k & l & c \end{pmatrix}$ by the LQJD procedure are in total conformity with those for $\begin{pmatrix} \beta & \alpha \end{pmatrix}$ by the AJD procedure. Furthermore, they have the same set of initial conditions. Therefore, asset pricing in the LQJD seting may be performed inside the AJD setting. Together with the fact that AJD models form a subset of the LQJD class, we conclude that the two classes are equivalent in terms of transform analysis. ■

Appendix D. Affine reformulation of the two-factor LQ stochastic volatility model

For better understanding of the equivalence relationship between the LQJD and AJD classes, we demonstrate how the LQJD dynamics can be formulated as affine dynamics with a concrete example, namely the two-factor LQ stochastic volatility model in Section 5.

Recall that the dynamics of the state vector in the two-factor LQ stochastic volatility model are given by (5.2) and (5.3), with constant interest rate, zero dividend rate, and constant jump intensity. The coefficients of the standard transform:

$$\phi^{s}\left(g_{1}\right)=e^{-r\tau}E_{t}\left[\exp\left(\frac{1}{2}\bar{\mathrm{X}}_{T}^{\top}A\bar{\mathrm{X}}_{T}+k^{\top}\bar{\mathrm{X}}_{T}+l^{\top}\underline{\mathrm{X}}_{T}+c\right)\right],$$

are $\left(\begin{array}{ccc}\varpi & l & c\end{array}\right)$ where:

$$\varpi=\left(\begin{array}{cc}A & k\end{array}\right)=\left(\begin{array}{ccc}a_{1} & a_{2} & k_{1}\\ a_{2} & a_{3} & k_{2}\end{array}\right).$$

By Proposition 1, $\left(\begin{array}{ccc}\varpi & l & c\end{array}\right)$ satisfy the system of ODEs:

$$\begin{aligned}\frac{d}{d\tau}l &= 0,\\ \frac{d}{d\tau}\varpi &= M_{21}+M_{22}\varpi-\varpi M_{11}-\varpi M_{12}\varpi,\\ \frac{d}{d\tau}c &= -\left(r-\lambda\mu\right)l_{1}+\theta\kappa_{1}k_{1}+\frac{1}{2}\left[\sigma_{1}^{2}\left(a_{1}+k_{1}^{2}\right)+\sigma_{2}^{2}\left(a_{3}+k_{2}^{2}\right)\right],\end{aligned}$$

where:

$$M_{11}=-\left(\begin{array}{ccc}-\kappa_{1} & 0 & \theta\kappa_{1}+\rho_{1}l_{1}\\ \kappa_{2} & -\kappa_{2} & +\rho_{2}l_{2}\\ 0 & 0 & 0\end{array}\right),\qquad M_{12}=-\left(\begin{array}{cc}\sigma_{1}^{2} & 0\\ 0 & \sigma_{2}^{2}\\ 0 & 0\end{array}\right),$$

$$M_{21}=2\left(\begin{array}{ccc}-l_{1}+l_{1}^{2} & -l_{1}+l_{1}^{2} & 0\\ -l_{1}+l_{1}^{2} & -l_{1}+l_{1}^{2} & 0\end{array}\right),\qquad M_{22}=-\left(\begin{array}{cc}-\kappa_{1} & \kappa_{2}\\ 0 & -\kappa_{2}\\ 0 & 0\end{array}\right).$$

To get the affine formulation of the two-factor LQ stochastic volatility model,, we need to add the following pseudo factors to the initial state vector:

$$\begin{aligned}Z_{1} &= X_{1}^{2},\\ Z_{2} &= X_{2}^{2},\\ Z_{12} &= X_{1}X_{2}.\end{aligned}$$

By Ito's lemma:

$$\begin{aligned}dZ_{1t} &= \left(\sigma_{1}^{2}+2\kappa_{1}\theta X_{1t}-2\kappa_{1}Z_{1t}\right)dt+2\sigma_{1}X_{1t}dW_{1t},\\ dZ_{2t} &= \left(\sigma_{2}^{2}-2\kappa_{2}Z_{2t}\right)dt+2\sigma_{2}X_{2t}dW_{2t},\\ dZ_{12t} &= \left[\kappa_{1}\theta X_{2t}-\left(\kappa_{1}+\kappa_{2}\right)Z_{12t}\right]dt+\sigma_{1}X_{2t}dW_{1t}+\sigma_{2}X_{1t}dW_{2t}.\end{aligned}$$

Apparently, the augmented state vector belongs to the AJD class. One may now apply results from the AJD class and check that the resulting system of ODEs is identical to the one obtained from the LQJD class.

References

[1] Ahn, D-H. (1995): *A Generalized Squared Autoregressive Intertemporal Term Structure Model*. Working paper, University of North Carolina, Chapel Hill.

[2] Ahn, D-H., and B., Gao (1999): *A Parametric Nonlinear Model of Term Structure Dynamics.* **Review of Financial Studies,** 12, 721-762.

[3] 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.

[4] Ait-Sahalia, Y. (1996): *Testing Continuous-Time Models of the Spot Interest Rate.* **Review of Financial Studies**, 9, 385-426.

[5] Ait-Sahalia, Y., and A. Lo (1998): *Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices.* **Journal of Finance**, 53(2), 499-548.

[6] Backus, D., S., Foresi, A., Mozumdar, and L., Wu (2001): *Predictable Changes in Yields and Forward Rates.* **Journal of Financial Economics**, 59, 281-311.

[7] Backus, D., S., Foresi, and C. Telmer (2001): *Affine Term Structure Models and the Forward Premium Anomaly.* **Journal of Finance**, 56(1), 279-304.

[8] Bakshi G., C. Cao, and Z. Chen (1997): *Empirical Performance of Alternative Option Pricing Models.* **Journal of Finance**, 52(5), 2003-2049.

[9] Bates, D. (1991): *The Crash of '87: Was It expected? The Evidence from Options Markets.* **Journal of Finance,** 46(3), 1009-1044.

[10] Bates, D. (1996): *Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options.* **Review of Financial Studies**, 9, 69-107.

[11] Bates, D. (2000): *Post-'87 Crash Fears in the S&P 500 Futures Option Market.* **Journal of Econometrics**, 94, 181-238.

[12] Beaglehole, D.R., and M. Tenney (1991): *General Solution of Some Interest Rate-Contingent Claim Pricing Equations.* **Journal of Fixed Income**, 1, 69-83.

[13] Beaglehole, D.R., and M. Tenney (1992): *A Nonlinear Equilibrium Model of Term Structures of Interest Rates: Corrections and Additions.* **Journal of Financial Economics**, 32, 345-354.

[14] Carr, P., and D.B. Madan (1999): *Option Valuation Using the Fast Fourier Transform.* **Journal of Computational Finance**, 2, 61-73.

[15] Chen, L., and H.V. Poor (2002): *A General Characterization of Quadratic Term Strucutre Models.* Working paper, Princeton University.

[16] 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.

[17] Constantinides, G.M. (1992): *A Theory of the Nominal Term Structure of Interest Rates.* **Review of Financial Studies**, 5, 531-552.

[18] Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985): *A Theory of the Term Structure of Interest Rates.* **Econometrica,** 53, 385-407.

[19] Dai, Q., and K.J. Singleton (2000): *Specification Analysis of Affine Term Structure Models.* **Journal of Finance**, 55(5), 1943-1978.

[20] Duffie, D., D. Filipovic, and W. Schachermayer (2002): *Affine Processes and Applications in Finance.* **The Annals of Applied Probability**, forthcoming.

[21] Duffie, D., and R. Kan (1996): *A Yield-Factor Model of Interest Rates.* **Mathematical Finance**, 6, 379-406.

[22] Duffie, D., J. Pan, and K. Singleton (2000): *Transform Analysis and Asset Pricing for Affine Jump-Diffusions.* **Econometrica,** 68, 1343-1376.

[23] El Karoui, N., R. Myneni, and R. Viswanatha (1992): *Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: I. Theory.* Working paper, Universite de Paris VI and Stanford University.

[24] Freiling, G. (2002): *A Survey of Nonsymmetric Riccati Equations.* **Linear Algebra and Its Applications**, 243-270.

[25] Filipović, D. (2001): *A General Characterization of One Factor Affine Term Structure Models.* **Finance and Stochastics**, 5, 389-412.

[26] Filipović, D. (2002): *Separable Term Structures and the Maximal Degree Problem.* **Mathematical Finance**, 12(4), 341-349.

[27] Hélyette Geman, N. El Karoui, and J-C. Rochet (1995): *Changes of Numéraire, Changes of Probability Measure and Option Pricing.* **Journal of Applied Probability**, 32, 443-458.

[28] Heston, S. (1993): *A Closed-form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options.* **Review of Financial Studies,** 6, 327-343.

[29] Jamshidian, F. (1996): *Bond, Futures and Option Valuation in the Quadratic Interest Rate Model.* **Applied Mathematical Finance**, 3, 93-115.

[30] Lewis, A.L. (2000): **Option Valuation under Stochastic Volatility with Mathematica Code**. Finance Press, USA.

[31] Lewis, A.L. (2001): *A Simple Option Formula for General Jump-Diffusion and Other Exponential Lévy Processes.* Envision Financial Systems and OptionCity.net. Working paper.

[32] Leippold, M. and L. Wu (2001): *Design and Estimation of Quadratic Term Structure Models.* Working paper, University of Zurich and Fordham University.

[33] Leippold, M. and L. Wu (2002): *Asset Pricing under the Quadratic Class.* **Journal of Financial and Quantitative Analysis**, forthcoming.

[34] Longstaff, F.A. (1989): *A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates.* **Journal of Financial Economics**, 23, 195-224.

[35] Magnus, J.R., and H. Neudecker (1988): **Matrix Differential Calculus with Applications in Statistics and Econometrics**. John Wiley and Sons, GB.

[36] Piazzesi, M. (2001): *An Econometric Model of the Yield Curve with Macroeconomic Jump Effects.* NBER Working Paper 8246.

[37] Renault, M. and N. Touzi (1996): *Option Hedging and Implied Volatilities in a Stochastic Volatility Model.* **Mathematical Finance**, 6(3), 279-302.

[38] Schöbel, R. and J. Zhu (1999): *Stochastic Volatility With an Ornstein-Uhlenbeck Process: An Extension.* **European Finance Review**, 3, 23-46.

[39] 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.

[40] Vasicek, O. (1977): *An Equilibrium Characterization of the Term Structure.* **Journal of Financial Economics,** 5, 177-188.

Table I

Generalized Fourier Transforms of Prices of Some Financial Claims

Claim	Terminal payoff	Price Transform	Restrictions on v_i
Digital (1)	$e^{s_T}\mathbb{I}_{\{s_T<k\}}$	$-\frac{1}{iv}\phi^s((iv+1)s_T)$	$v_i > 0$
Digital (2)	$e^k\mathbb{I}_{\{s_T>k\}}$	$\frac{1}{iv+1}\phi^s((iv+1)s_T)$	$v_i < 1$
Call	$(e^{s_T}-e^k)\mathbb{I}_{\{s_T>k\}}$	$\frac{1}{iv-v^2}\phi^s((iv+1)s_T)$	$v_i < 0$
Put	$(e^k-e^{s_T})\mathbb{I}_{\{s_T<k\}}$	$\frac{1}{iv-v^2}\phi^s((iv+1)s_T)$	$v_i > 1$
Money market	1	$2\pi\delta(v)$	none

The table presents the generalized Fourier transforms of prices of some typical financial claims, as well as restrictions on the imaginary part of the transform variable v_i. $\phi^s(\bullet)$ stands for the standard transform discussed in Proposition 1, and $\delta(\bullet)$ is a delta Dirac function.

Table II.

Fitted Parameter Values of Various Stochastic Volatility Models

			Heston	DPS	Bates	LQ
(Without Jumps)						
	$\rho_{s,V}$		–0.6965	–0.6944	–0.6289	–0.6271
	ρ_{s,X_1}		–0.6965	–0.6944	–0.7025	–0.6973
	ρ_{s,X_2}		–	–	–0.8946	–1.0000
	σ_1		0.4257	0.4179	0.6134	0.2054
	σ_2		–	0.0000	0.0683	0.0429
	κ_1		4.3526	4.1904	5.9838	6.9470
	κ_2		–	6.3123	0.0085	0.0000
	θ		0.0129	0.0129	0.0108	0.0823
	$\sqrt{V}$		0.0843	0.0831	0.0819	0.0772
Option Price	*MSE*		0.0124	0.0123	0.0086	0.0079
	MAD		0.0957	0.0956	0.0764	0.0767
Implied Vol.	*MSE*	$(\times 10^{-3})$	0.0750	0.0801	0.0420	0.0545
	MAD		0.0043	0.0044	0.0032	0.0036
(With Jumps)						
	$\rho_{s,V}$		–0.7831	–0.7836	–0.5774	–0.5786
	ρ_{s,X_1}		–0.7831	–0.7836	–0.7143	–0.6741
	ρ_{s,X_2}		–	–	–1.0000	–0.9361
	σ_1		0.1939	0.2023	0.4974	0.1857
	σ_2		–	0.0000	0.0027	0.0308
	κ_1		2.7945	3.0315	3.2436	13.8801
	κ_2		–	4.6868	0.0000	0.1000
	θ		0.0099	0.0100	0.0091	0.1031
	λ		0.0748	0.0682	0.1015	0.0356
	μ_J		–0.1384	–0.1467	–0.1236	0.2191
	σ_J		0.1503	0.1526	0.0359	0.1459
	$\sqrt{V}$		0.0787	0.0787	0.0787	0.0735
Option Price	*MSE*		0.0070	0.0070	0.0063	0.0036
	MAD		0.0645	0.0647	0.0663	0.0484
Implied Vol.	*MSE*	$(\times 10^{-3})$	0.0097	0.0092	0.0195	0.0063
	MAD		0.0019	0.0022	0.0022	0.0015

Table III

The MSE and MAD of Fitted ITM, ATM, and OTM Option Prices

		Heston	DPS	Bates	LQ
(Without Jumps)					
MSE	ITM	0.0127	0.0135	0.0097	0.0082
	ATM	0.0137	0.0118	0.0073	0.0080
	OTM	0.0086	0.0078	0.0063	0.0066
MAD	ITM	0.1007	0.1041	0.0856	0.0772
	ATM	0.0949	0.0878	0.0609	0.0792
	OTM	0.0753	0.0724	0.0640	0.0695
(With Jumps)					
MSE	ITM	0.0082	0.0080	0.0058	0.0033
	ATM	0.0065	0.0069	0.0074	0.0045
	OTM	0.0026	0.0027	0.0066	0.0036
MAD	ITM	0.0702	0.0704	0.0638	0.0466
	ATM	0.0612	0.0629	0.0701	0.0520
	OTM	0.0452	0.0413	0.0708	0.0498

Figure 1. Market implied smiles

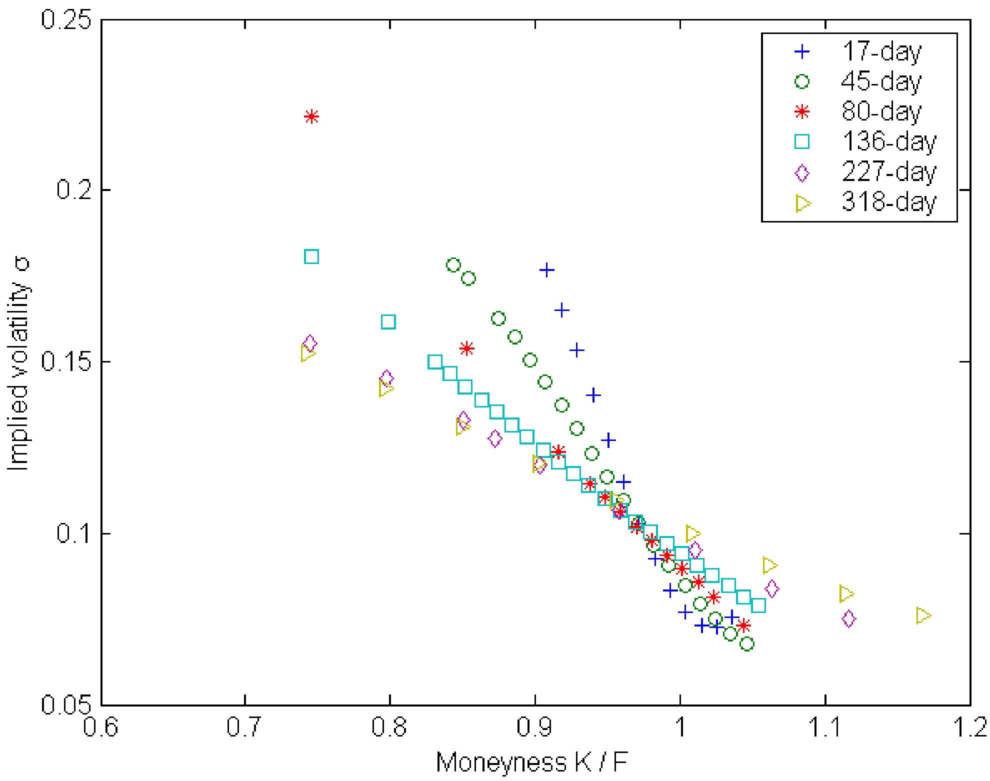

Figure 2. The modulus of the generalized Fourier transform of calls

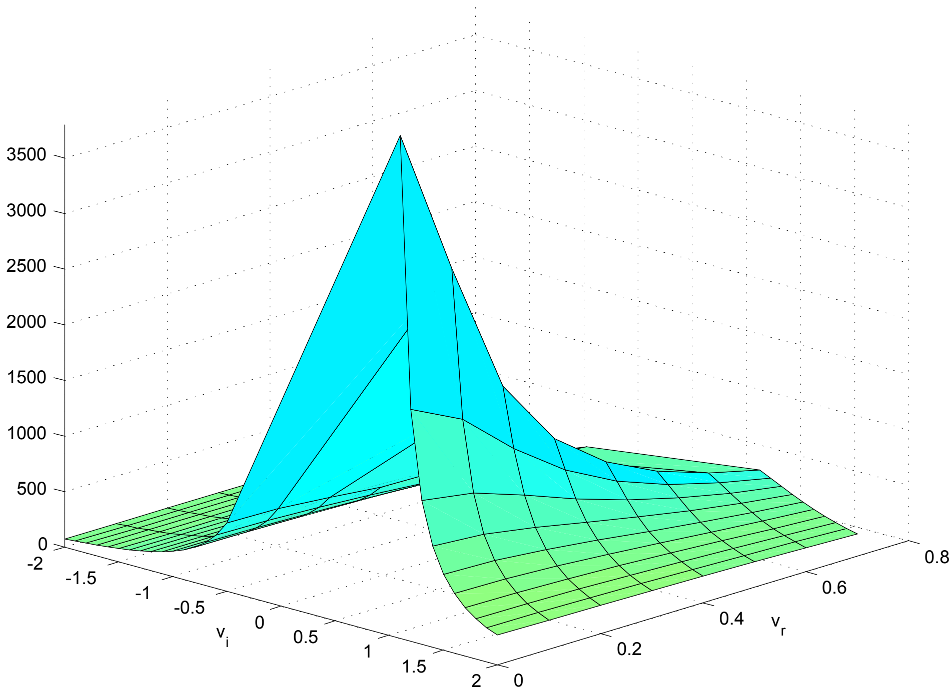

Figure 3a. Implied volatility surface of the LQ model without jumps

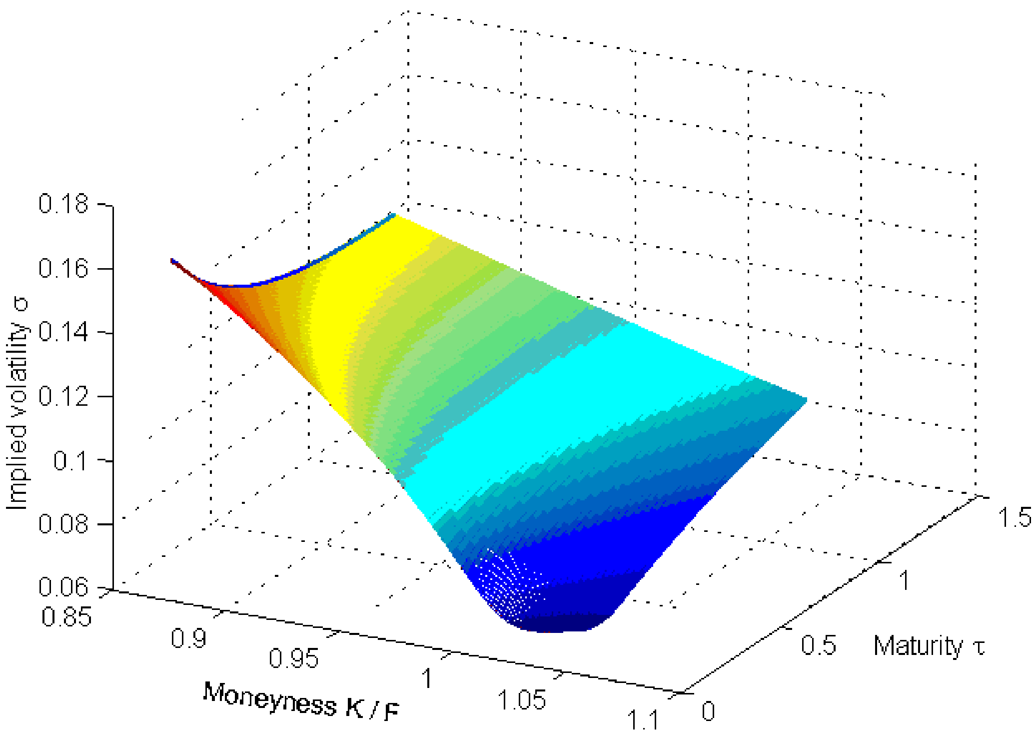

Figure 3b. Implied volatility surface of the LQ model with jumps

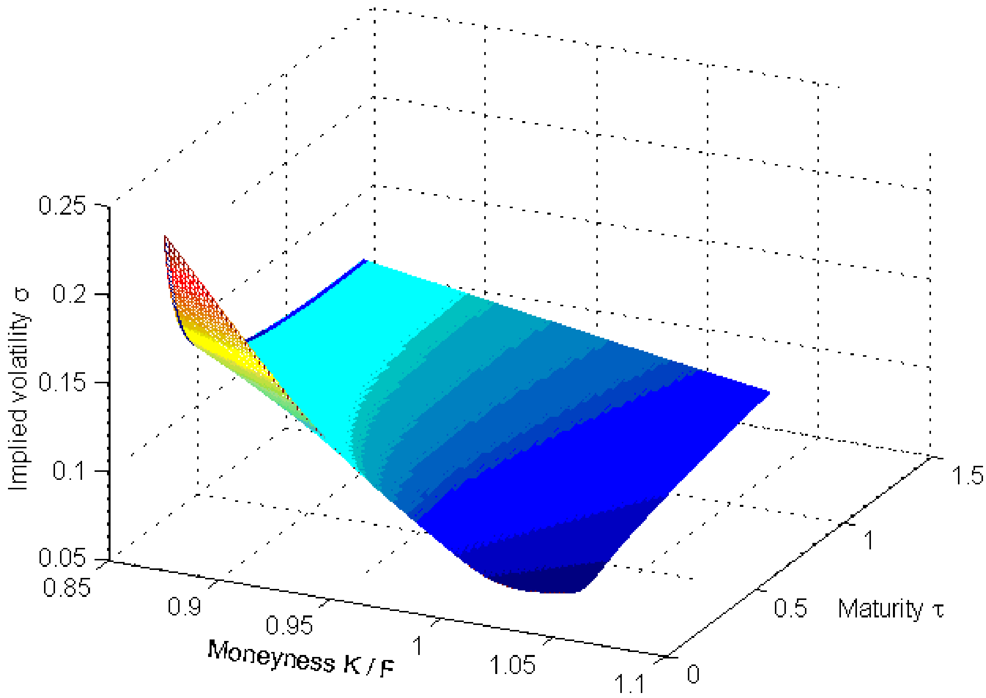

Figure 4a. 17-day implied volatility smiles (without jumps)

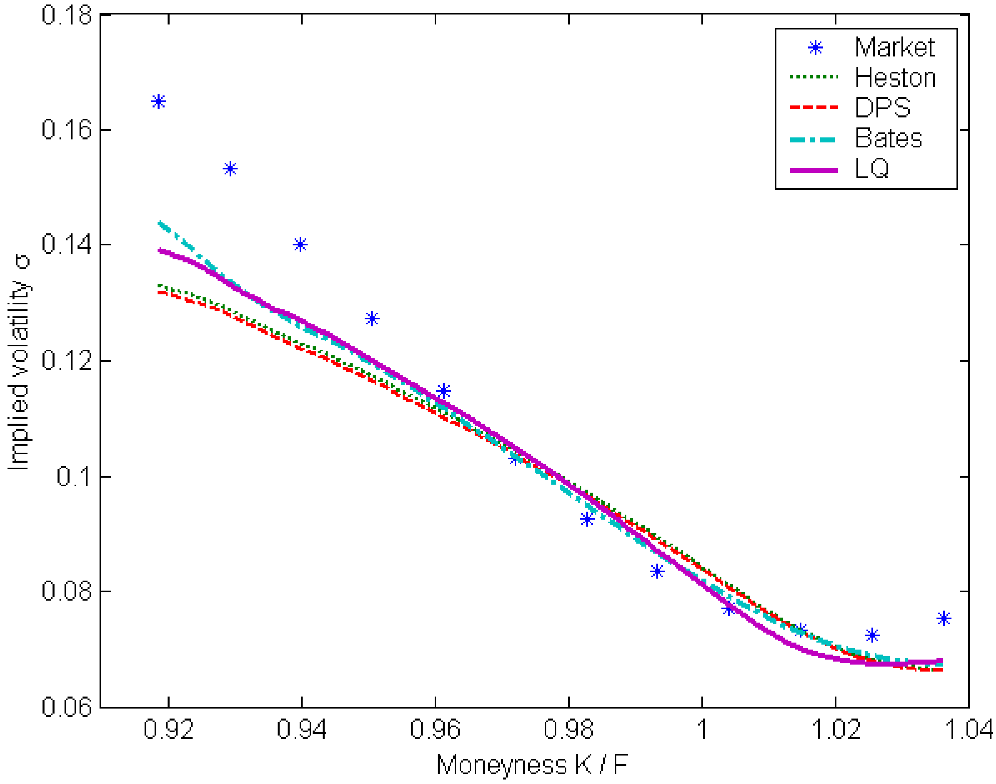

Figure 4b. 17-day implied volatility smiles (with jumps)

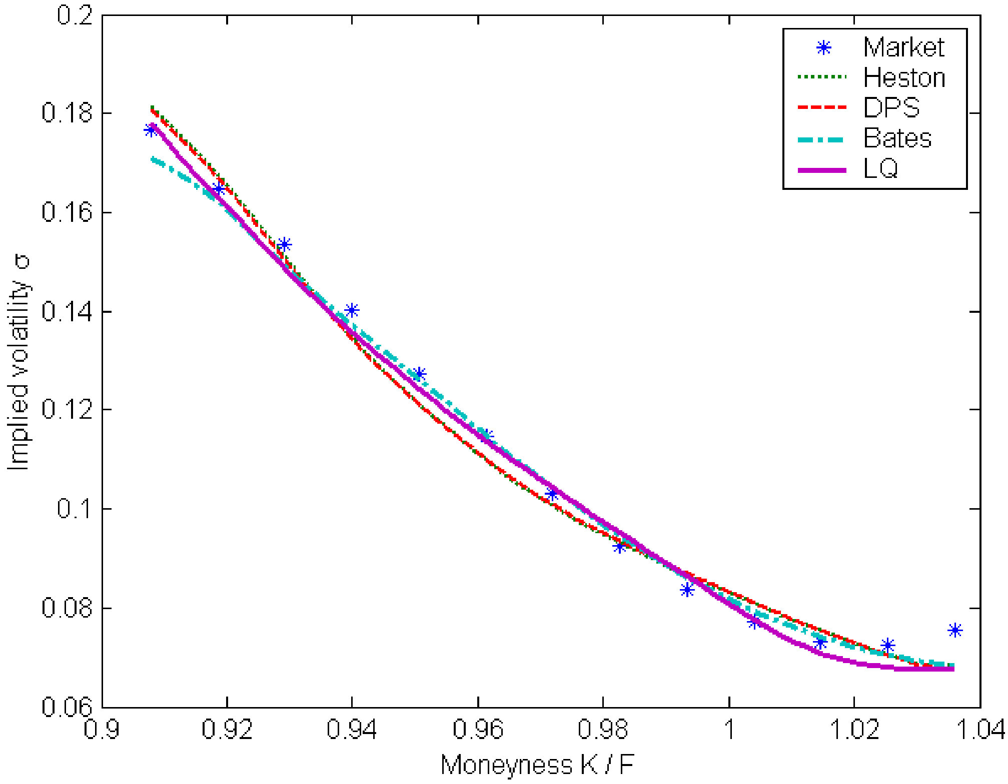

Figure 5a. 318-day implied volatility smiles (without jumps)

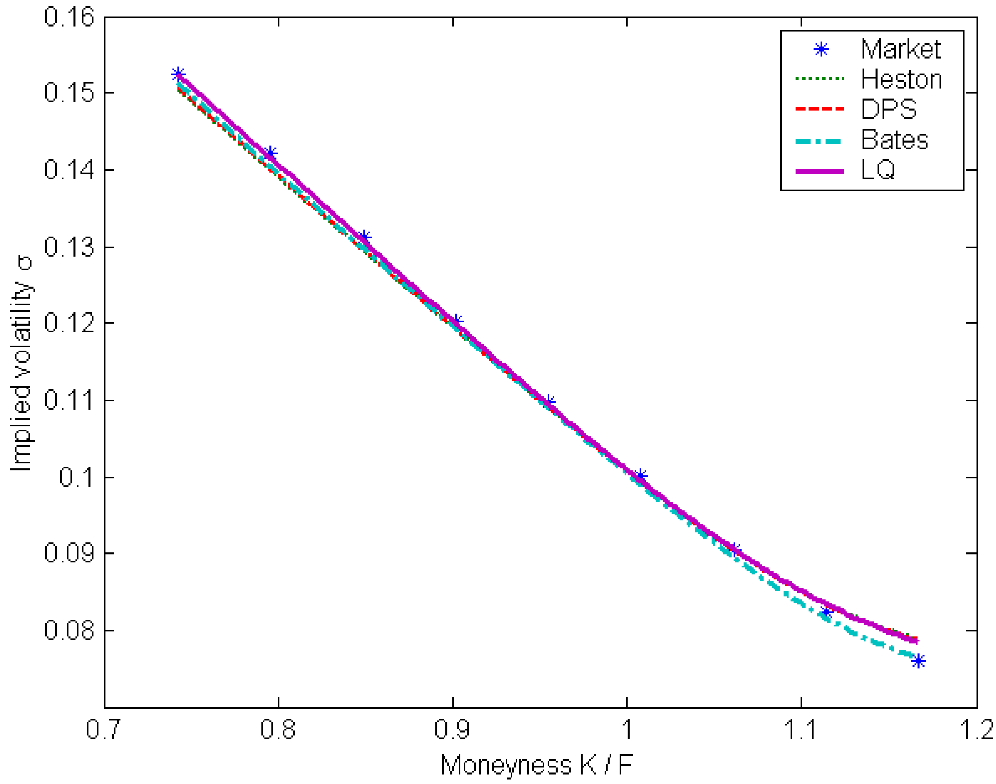

Figure 5b. 318-day implied volatility smiles (with jumps)