Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

A multifactor volatility Heston model

JosÉ Da Fonseca , Martino Grasselli & Claudio Tebaldi

To cite this article: JosÉ Da Fonseca , Martino Grasselli & Claudio Tebaldi (2008) A multifactor volatility Heston model, Quantitative Finance, 8:6, 591-604, DOI: 10.1080/14697680701668418

To link to this article: https://doi.org/10.1080/14697680701668418

Published online: 17 Sep 2008.

Quantitative Finance, Vol. 8, No. 6, September 2008, 591–604

Routledge
Taylor & Francis Group

A multifactor volatility Heston model

JOSÉ DA FONSECA†‡, MARTINO GRASSELLI*†§
and CLAUDIO TEBALDI¶

†ESILV, Ecole Supérieure d'Ingénieurs Léonard de Vinci, Département Mathématiques
et Ingénierie Financière, 92916 Paris La Défense, France
‡Zeliade Systems, 56 Rue Jean-Jacques Rousseau, 75001 Paris, France
§Dipartimento di Matematica Pura ed Applicata, Universitá degli Studi di Padova,
Via Trieste 63, 35121 Padova, Italy
¶Department of Finance, Univ. L. Bocconi, Viale Isonzo 25, 20136 Milano, Italy

(*Received 28 March 2006; in final form 6 September 2007*)

We model the volatility of a single risky asset using a multifactor (matrix) Wishart affine process, recently introduced in finance by Gourieroux and Sufana. As in standard Duffie and Kan affine models the pricing problem can be solved through the Fast Fourier Transform of Carr and Madan. A numerical illustration shows that this specification provides a separate fit of the long-term and short-term implied volatility surface and, differently from previous diffusive stochastic volatility models, it is possible to identify a specific factor accounting for the stochastic leverage effect, a well-known stylized fact of the FX option markets analysed by Carr and Wu.

Keywords: Stochastic volatility; Financial derivatives; Volatility modelling; Options pricing; Options volatility

1. Introduction

Accurate volatility modelling is a crucial step for the implementation of realistic and efficient risk-minimizing strategies for financial and insurance companies. For example, pension plans usually attach guarantees to their products that are linked to equity returns. Hedging of such guarantees involves, beyond plain vanilla options, also exotic contracts, such as, for example, cliquet options. These instruments, also called ratchet options, periodically 'lock in' profits in a manner somewhat analogous to a mechanical ratchet. Exotic contracts such as cliquet options require an accurate modeling of the true realized variance process. In fact, a cliquet option can be seen as a series of consecutive forward start options, the prices of which depend only on realized volatility (see, e.g., Hipp (1996)). As well explained by Bergomi (2004), there is a structural limitation that prevents one-factor stochastic volatility models pricing consistently these types of options jointly with plain vanilla options. A possible reconciliation requires that the volatility process is driven by at least two factors, even in a single asset framework, as supported by empirical tests such as the principal component analysis investigated by Cont and Fonseca (2002).

Among one factor stochastic volatility models, the most popular and easy to implement is certainly the Heston (1993) model, in which the volatility satisfies a (positive) single factor square root process, where the pricing and hedging problem can be efficiently solved by performing a fast Fourier transform (FFT hereafter, see, e.g., Carr and Madan (1999)).

Within the Heston model an accurate modeling of the smile–skew effect for the implied volatility surface is usually obtained assuming a (negative) correlation between the noise driving the stock return and a suitable calibration of the parameters driving the volatility. It is indeed a common observation that a single factor diffusive model is not sufficiently flexible to take into account the risk component introduced by the variability of the skew, also known as *stochastic skew* (see, e.g., Carr and Wu (2004)). In the case of FX options this risk factor is directly priced in the quotes of 'risk reversal' strategies.

*Corresponding author. Email: grassell@math.unipd.it

Quantitative Finance
ISSN 1469–7688 print/ISSN 1469–7696 online © 2008 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/14697680701668418

The aim of this paper is to extend the Heston model to a multifactor specification for the volatility process in a single asset framework. While standard multifactor modeling of stochastic volatility is based on the class of affine term structure models introduced by Duffie and Kan (1996) and classified by Dai and Singleton (2000), in our model the factor process driving volatility is based on the matrix Wishart process, developed mathematically by Bru (1991). Our model takes inspiration from the multi-asset market model analysed by Gourieroux and Sufana (2004): in their model the Wishart process describes the dynamics of the covariance matrix and is assumed to be independent of the asset noise. On the contrary, we show that a symmetric matrix specification is potentially very useful for improving the affine factor modeling of the implied volatility curve. In fact, the introduction of matrix notation provides a simple and powerful parametrization of the dependence between the asset noise and each volatility factor. In particular, using a square 2×2 matrix of factors, we show that the expression of the return-volatility covariance is linear in the off-diagonal factor, which can be directly identified as the 'stochastic skew risk factor': in fact, such a factor can specifically be used to generate *a stochastic leverage effect*, which, in the case of an FX option, can be directly calibrated on Risk Reversal quotes.

Summing up, the present single asset model achieves the following goals:

(i) the term structure of the realized volatilities is described by a (matrix) multifactor model;
(ii) a stochastic leverage effect appears and can be used to describe stochastic skew effects as required in FX option markets (Carr and Wu 2004);
(iii) analytic tractability, i.e. the pricing problem, can be handled through the FFT methodology as in Carr and Madan (1999).

We provide a numerical illustration that motivates the introduction of the Wishart (multifactor) volatility process: we show that our model, differently from the traditional Heston (single factor) model, can fit separately the long-term volatility level and the short-term volatility skew. Moreover, the correlation between assets' returns and their volatility turns out to be stochastic, so that, in our model, we can deal with a stochastic skew effect as in Carr and Wu (2004).

The paper is organized as follows. In section 2 we introduce the stochastic (Wishart) volatility market model together with the correlation structure. In section 3 we solve the general pricing problem by determining the explicit expression of the Laplace–Fourier transforms of the relevant processes. In addition, we explicitly compute the price of the forward-start options, i.e. the building blocks of cliquet options. Section 4 provides a numerical illustration that shows the advantages of the Wishart specification with respect to the single factor Heston specification as well as the $\mathbb{A}_2(3)$ (in the terminology of Dai and Singleton (2000)) multi-Heston model. In section 5 we provide the conclusions and future developments. We gather in appendix A some technical proofs, and in appendix B we develop the computations in the two-dimensional case for the reader's convenience. Finally, appendix C discusses the general affine correlation structure in the two-dimensional case.

2. The Wishart volatility process

In an arbitrage-free frictionless financial market we consider a risky asset, the price of which is

$$\frac{\mathrm{d}S_t}{S_t} = r\,\mathrm{d}t + \mathrm{Tr}[\sqrt{\Sigma_t}\,\mathrm{d}Z_t], \tag{1}$$

where r denotes the (not necessarily constant) risk-free interest rate, Tr is the trace operator, $Z_t \in M_n$ (the set of square matrices) is a matrix Brownian motion (i.e. composed of n^2 independent Brownian motions) under the risk-neutral measure and Σ_t belongs to the set of symmetric $n \times n$ positive-definite matrices (as well as its square root $\sqrt{\Sigma_t}$).

From equation (1) it follows that the quadratic variation of the risky asset is the trace of the matrix Σ_t: that is, in this specification the volatility is multi-dimensional since it depends on the elements of the matrix process Σ_t, which is assumed to satisfy the following dynamics:

$$\begin{aligned}\mathrm{d}\Sigma_t &= (\Omega\Omega^{\mathrm{T}} + M\Sigma_t + \Sigma_t M^{\mathrm{T}})\mathrm{d}t \\ &\quad + \sqrt{\Sigma_t}\mathrm{d}W_t Q + Q^{\mathrm{T}}(\mathrm{d}W_t)^{\mathrm{T}}\sqrt{\Sigma_t},\end{aligned} \tag{2}$$

with Ω, M, $Q \in M_n$, Ω invertible, and $W_t \in M_n$ is a matrix Brownian motion. Equation (2) characterizes the Wishart process introduced by Bru (1991), and represents the matrix analogue of the square root mean-reverting process. In order to grant the strict positivity and the typical mean reverting feature of the volatility, the matrix M is assumed to be negative semi-definite, while Ω satisfies

$$\Omega\Omega^{\mathrm{T}} = \beta Q^{\mathrm{T}} Q,$$

with the real parameter $\beta > n-1$ (see Bru (1991), p. 747). Wishart processes have recently been applied in finance by Gourieroux and Sufana (2004): they considered a multi-asset stochastic volatility model:

$$\mathrm{d}S_t = \mathrm{diag}[S_t](r\mathbf{1}\mathrm{d}t + \sqrt{\Sigma_t}\mathrm{d}Z_t),$$

where S_t, $Z_t \in \mathbb{R}^n$, $\mathbf{1} = (1, \ldots, 1)^T$ *and the* (*Wishart*) *volatility matrix is assumed to be independent of* Z_t. In our (single-asset) specification we relax the independency assumption: in particular, in order to take into account the skew effect of the (implied) volatility smile, we assume correlation between the noises driving the asset and the noises driving the volatility process.

2.1. The correlation structure

We correlate the two matrix Brownian motions W_t, Z_t in such a way that all the (scalar) Brownian motions belonging to the column i of Z_t and the corresponding Brownian motions of the column j of W_t have the same correlation, say R_{ij}. This leads to a constant matrix $R \in M_n$ (identified up to a rotation) that completely describes the correlation structure, in such a way that Z_t can be written as $Z_t := W_t R^{\mathrm{T}} + B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}$ ($\mathbb{I}$ represents the identity matrix and $^{\mathrm{T}}$ denotes transposition) where B_t is a (matrix) Brownian motion independent of W_t.

Proposition 2.1: *The process* $Z_t := W_t R^{\mathrm{T}} + B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}$ *is a matrix Brownian motion.*

Proof: It is well known that Z_t is a matrix Brownian motion iff for any $\alpha,\beta \in \mathbb{R}^n$,

$$\mathrm{Cov}_t(\mathrm{d}Z_t\alpha, \mathrm{d}Z_t\beta) = \mathbb{E}_t[(\mathrm{d}Z_t\alpha)(\mathrm{d}Z_t\beta)^{\mathrm{T}}] = \alpha^{\mathrm{T}}\beta\mathbb{I}\mathrm{d}t.$$

Here

$$\begin{aligned}\mathrm{Cov}_t(\mathrm{d}Z_t\alpha, \mathrm{d}Z_t\beta) &= \mathbb{E}_t[(\mathrm{d}W_tR^{\mathrm{T}}\alpha \\ &\quad + \mathrm{d}B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}\alpha)(\mathrm{d}W_tR^{\mathrm{T}}\beta \\ &\quad + \mathrm{d}B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}\beta)^{\mathrm{T}}] \\ &= \mathrm{Cov}_t(\mathrm{d}W_tR^{\mathrm{T}}\alpha, \mathrm{d}WR^{\mathrm{T}}\beta) \\ &\quad + \mathrm{Cov}_t(\mathrm{d}B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}\alpha, \\ &\quad \times \mathrm{d}B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}\beta) \\ &= \alpha^{\mathrm{T}}RR^{\mathrm{T}}\beta\mathbb{I}\mathrm{d}t + \alpha^{\mathrm{T}}(\mathbb{I} - RR^{\mathrm{T}})\beta\mathbb{I}\mathrm{d}t \\ &= \alpha^{\mathrm{T}}\beta\mathbb{I}\mathrm{d}t. \qquad \square\end{aligned}$$

In principle, one should allow for a $n^2 \times n^2$ matrix corresponding to the (possibly different) correlations between W_t and Z_t. However, in order to give analytical tractability to the model (in particular, in order to preserve the affinity) some constraints should be imposed on the correlation structure. It turns out that such (nonlinear) constraints are quite binding: in order to give an idea we classify in the appendix all the possibilities for the case $n = 2$. Our choice can be seen as a parsimonious way (using only n^2 parameters) to introduce a simple correlation structure in the model.

3. The pricing problem

In this section we deal with the pricing problem of plain vanilla contingent claims, in particular the European call with payoff

$$(S_T - K)^+.$$

We shall see that, within the Wishart specification, analytical tractability is preserved exactly as in the (one-dimensional) Heston model. In fact, it is well known that in order to solve the pricing problem of plain vanilla options, it is sufficient to compute the conditional characteristic function (under the risk-neutral measure) of the underlying (see, e.g., Duffie *et al.* (2000)) or, equivalently, of the return process $Y_t = \ln S_t$, which satisfies the following SDE:

$$\begin{aligned}\mathrm{d}Y_t &= \left(r - \frac{1}{2}\mathrm{Tr}[\Sigma_t]\right)\mathrm{d}t \\ &\quad + \mathrm{Tr}[\sqrt{\Sigma_t}(\mathrm{d}W_tR^{\mathrm{T}} + \mathrm{d}B_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}})]. \qquad (3)\end{aligned}$$

We will first compute the infinitesimal generator of the relevant processes and we will show that the computation of the characteristic function involves the solution of a Matrix Riccati ODE. We will linearize such equations and we will then provide the closed-form solution to the pricing problem via the FFT methodology.

3.1. The Laplace transform of the asset returns

Following Duffie *et al.* (2000), in order to solve the pricing problem for plain vanilla options we just need the Laplace transform of the process (3). Since the Laplace transform of Wishart processes is exponentially affine (see, e.g., Bru (1991)), we surmise that the conditional moment generating function of the asset returns is the exponential of an affine combination of Y and the elements of the Wishart matrix. In other words, we look for three deterministic functions $A(t) \in M_n$, $b(t) \in \mathbb{R}$, $c(t) \in \mathbb{R}$ that parametrize the Laplace transform:

$$\begin{aligned}\Psi_{\gamma,t}(\tau) &= \mathbb{E}_t \exp\{\gamma Y_{t+\tau}\} \\ &= \exp\{\mathrm{Tr}[A(\tau)\Sigma_t] + b(\tau)Y_t + c(\tau)\}, \qquad (4)\end{aligned}$$

where $\mathbb{E}_t$ denotes the conditional expected value with respect to the risk-neutral measure and $\gamma \in \mathbb{R}$. By applying the Feynman–Kac argument, we have

$$\begin{aligned}\frac{\partial\Psi_{\gamma,t}}{\partial\tau} &= \mathcal{L}_{Y,\Sigma}\Psi_{\gamma,t}, \\ \Psi_{\gamma,t}(0) &= \exp\{\gamma Y_t\}. \qquad (5)\end{aligned}$$

The matrix setting for the Wishart dynamics implies a non-standard definition of the infinitesimal generator for the couple (Y_t, Σ_t). The infinitesimal generator for the Wishart process, Σ_t, has been computed by Bru (1991, p. 746, formula (5.12)):

$$\mathcal{L}_\Sigma = \mathrm{Tr}[(\Omega\Omega^{\mathrm{T}} + M\Sigma + \Sigma M^{\mathrm{T}})D + 2\Sigma DQ^{\mathrm{T}}QD], \qquad (6)$$

where D is a matrix differential operator with elements

$$D_{i,j} = \left(\frac{\partial}{\partial\Sigma^{ij}}\right).$$

For the reader's convenience, we develop the computations in the two-dimensional case in appendix B. With the previous result, we can now find the infinitesimal generator of the couple (Y_t,Σ_t).

Proposition 3.1: *The infinitesimal generator of* (Y_t, Σ_t) *is given by*

$$\mathcal{L}_{Y,\Sigma} = \left(r - \frac{1}{2}\,\mathrm{Tr}[\Sigma]\right)\frac{\partial}{\partial y} + \frac{1}{2}\,\mathrm{Tr}[\Sigma]\frac{\partial^2}{\partial y^2} + \mathrm{Tr}[(\Omega\Omega^T + M\Sigma + \Sigma M^T)D + 2\Sigma D Q^{\mathrm{T}} Q D] + 2\,\mathrm{Tr}[\Sigma R Q D]\frac{\partial}{\partial y}. \quad (7)$$

Proof: See appendix A. □

Thus the explicit expression of (5) is

$$\frac{\partial \Psi_{\gamma,t}}{\partial \tau} = \left(r - \frac{1}{2}\mathrm{Tr}[\Sigma]\right)\frac{\partial \Psi_{\gamma,t}}{\partial y} + \frac{1}{2}\mathrm{Tr}[\Sigma]\frac{\partial^2 \Psi_{\gamma,t}}{\partial y^2} + \mathrm{Tr}[(\Omega\Omega^{\mathrm{T}} + M\Sigma + \Sigma M^{\mathrm{T}})D\Psi_{\gamma,t} + 2(\Sigma D Q^{\mathrm{T}} Q D)\Psi_{\gamma,t}] + 2\mathrm{Tr}[\Sigma R Q D]\frac{\partial \Psi_{\gamma,t}}{\partial y},$$

and by replacing the candidate (4) we obtain

$$0 = -\mathrm{Tr}\left[\frac{\mathrm{d}}{\mathrm{d}\tau}A(\tau)\Sigma\right] - \frac{\mathrm{d}}{\mathrm{d}\tau}b(\tau)Y - \frac{\mathrm{d}}{\mathrm{d}\tau}c(\tau) + \mathrm{Tr}[(\Omega\Omega^{\mathrm{T}} + M\Sigma + \Sigma M^{\mathrm{T}})A(\tau) + 2\Sigma A(\tau)Q^{\mathrm{T}}QA(\tau) + 2\Sigma RQA(\tau)b(\tau)] + \left(r - \frac{1}{2}\mathrm{Tr}[\Sigma]\right)b(\tau) + \frac{1}{2}\mathrm{Tr}[\Sigma]b^2(\tau), \quad (8)$$

with boundary conditions

$$A(0) = 0 \in M_n, \quad b(0) = \gamma \in \mathbb{R}, \quad c(0) = 0.$$

By identifying the coefficients of Y we deduce

$$\frac{\mathrm{d}}{\mathrm{d}\tau}b(\tau) = 0,$$

hence

$$b(\tau) = \gamma, \quad \text{for all } \tau.$$

By identifying the coefficients of Σ we obtain the Matrix Riccati ODE satisfied by $A(\tau)$:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}A(\tau) = A(\tau)M + (M^{\mathrm{T}} + 2\gamma RQ)A(\tau) + 2A(\tau)Q^{\mathrm{T}}QA(\tau) + \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n$$
$$A(0) = 0. \quad (9)$$

Finally, as usual, the function $c(\tau)$ can be obtained by direct integration:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}c(\tau) = \mathrm{Tr}[\Omega\Omega^{\mathrm{T}}A(\tau)] + \gamma r, \quad c(0) = 0. \quad (10)$$

3.2. Matrix Riccati linearization

Matrix Riccati equations such as (9) have several nice properties (see, e.g., Freiling (2002)): the most remarkable property is that their flow can be linearized by doubling the dimension of the problem, due to the fact that Riccati ODE belong to a quotient manifold (see Grasselli and Tebaldi (2008) for further details). For sake of completeness, we now recall the linearization procedure, and provide the closed-form solution to (9) and (10).

Put

$$A(\tau) = F(\tau)^{-1}G(\tau) \quad (11)$$

for $F(\tau) \in GL(n)$, $G(\tau) \in M_n$. Then

$$\frac{\mathrm{d}}{\mathrm{d}\tau}[F(\tau)A(\tau)] - \frac{\mathrm{d}}{\mathrm{d}\tau}[F(\tau)]A(\tau) = F(\tau)\frac{\mathrm{d}}{\mathrm{d}\tau}A(\tau),$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\tau}G(\tau) - \frac{\mathrm{d}}{\mathrm{d}\tau}[F(\tau)]A(\tau) = \frac{\gamma(\gamma-1)}{2}F(\tau) + G(\tau)M + (F(\tau)(M^{\mathrm{T}} + 2\gamma RQ) + 2G(\tau)Q^{\mathrm{T}}Q)A(\tau).$$

The last ODE leads to a system of $(2n)$ *linear* equations:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}G(\tau) = \frac{\gamma(\gamma-1)}{2}F(\tau) + G(\tau)M,$$
$$\frac{\mathrm{d}}{\mathrm{d}\tau}F(\tau) = -F(\tau)(M^{\mathrm{T}} + 2\gamma RQ) - 2G(\tau)Q^{\mathrm{T}}Q, \quad (12)$$

which can be written as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} G(\tau) & F(\tau) \end{pmatrix} = \begin{pmatrix} G(\tau) & F(\tau) \end{pmatrix}\begin{pmatrix} M & -2Q^{\mathrm{T}}Q \\ \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n & -(M^{\mathrm{T}} + 2\gamma RQ) \end{pmatrix}.$$

Its solution is simply obtained through exponentiation:

$$\begin{pmatrix} G(\tau) & F(\tau) \end{pmatrix} = \begin{pmatrix} G(0) & F(0) \end{pmatrix}\exp\tau\begin{pmatrix} M & -2Q^{\mathrm{T}}Q \\ \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n & -(M^{\mathrm{T}} + 2\gamma RQ) \end{pmatrix}$$
$$= \begin{pmatrix} A(0) & \mathbb{I}_n \end{pmatrix}\exp\tau\begin{pmatrix} M & -2Q^{\mathrm{T}}Q \\ \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n & -(M^{\mathrm{T}} + 2\gamma RQ) \end{pmatrix}$$
$$= \begin{pmatrix} A(0)A_{11}(\tau) + A_{21}(\tau) & A(0)A_{12}(\tau) + A_{22}(\tau) \end{pmatrix},$$

where

$$\begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} = \exp\tau\begin{pmatrix} M & -2Q^{\mathrm{T}}Q \\ \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n & -(M^{\mathrm{T}} + 2\gamma RQ) \end{pmatrix}. \quad (13)$$

In conclusion, we obtain

$$A(\tau) = (A(0)A_{12}(\tau) + A_{22}(\tau))^{-1}(A(0)A_{11}(\tau) + A_{21}(\tau)),$$

and since $A(0)=0$,

$$A(\tau) = A_{22}(\tau)^{-1}A_{21}(\tau), \tag{14}$$

which represents the closed-form solution of the matrix Riccati (9). Let us now turn our attention to equation (10). We can improve its computation by the following trick: from (12) we obtain

$$G(\tau) = -\frac{1}{2}\left(\frac{\mathrm{d}}{\mathrm{d}\tau}F(\tau) + F(\tau)(M^{\mathrm{T}} + 2\gamma RQ)\right)(Q^{\mathrm{T}}Q)^{-1},$$

and plugging into (11) and using the properties of the trace we deduce

$$\frac{\mathrm{d}}{\mathrm{d}\tau}c(\tau) = -\frac{\beta}{2}\mathrm{Tr}\left[F(\tau)^{-1}\frac{\mathrm{d}}{\mathrm{d}\tau}F(\tau) + (M^{\mathrm{T}} + 2\gamma RQ)\right] + \gamma r.$$

Now we can integrate the last equation and obtain

$$c(\tau) = -\frac{\beta}{2}\mathrm{Tr}\big[\log F(\tau) + (M^{\mathrm{T}} + 2\gamma RQ)\tau\big] + \gamma r\tau.$$

This result is very interesting because it avoids the numerical integration involved in the computation of $c(\tau)$.

Remark 1: The computation of the Laplace transform for both asset returns and variance factors,

$$\begin{aligned}\Psi_{\gamma,\Gamma,t}(\tau) &= \mathbb{E}_t \exp\{\gamma Y_{t+\tau} + \mathrm{Tr}[\Gamma\Sigma_{t+\tau}]\}\\ &= \exp\{\mathrm{Tr}[\widetilde{A}(\tau)\Sigma_t] + \widetilde{b}(\tau)Y_t + \widetilde{c}(\tau)\},\end{aligned} \tag{15}$$

can be easily handled by replacing the corresponding boundary conditions and repeating the above procedure.

3.3. *The characteristic function and the FFT method*

Let us now return to the pricing problem of a call option, and let us briefly recall the fast Fourier transform (FFT) method as in Carr and Madan (1999). For a fixed $\alpha > 0$, let us consider the scaled call price at time 0 as

$$\begin{aligned}c_T(k) &:= \exp\{\alpha k\}\mathbb{E}[\exp\{-rT\}(S_T - K)^+]\\ &= \exp\{\alpha k\}\mathbb{E}[\exp\{-rT\}(\exp\{Y_T\} - \exp\{k\})^+],\end{aligned}$$

where $k = \log K$. The modified call price $c_T(\alpha)$ is introduced in order to obtain a square integrable function (Carr and Madan 1999), and its Fourier transform is given by

$$\begin{aligned}\psi_T(v) &:= \int_{-\infty}^{+\infty} \exp\{ivk\}c_T(k)\mathrm{d}k\\ &= \exp\{-rT\}\frac{\Phi_{(v-(\alpha+1)i),0}(T)}{(\alpha+iv)(\alpha+1+iv)},\end{aligned}$$

which involves the characteristic function Φ. Recall that, from the Laplace transform, the characteristic function is easily derived by replacing γ with $i\gamma$, where $i = \sqrt{-1}$. The inverse fast Fourier transform is an efficient method for computing the following integral:

$$Call(0) = \frac{\exp\{-\alpha k\}}{2\pi}\int_{-\infty}^{+\infty} \exp\{-ivk\}\psi_T(v)\mathrm{d}v,$$

which represents the inverse transform of $\psi_T(v)$, that is the price of the (non-modified) call option. In conclusion, the call option price is known once the parameter α is chosen (typically $\alpha = 1.1$; Carr and Madan 1999) and the characteristic function Φ is found explicitly, which is the case of the (Heston as well as of the) Wishart volatility model.

3.4. *Explicit pricing for the forward-start option*

In this section we apply the methodology developed in the previous section in order to determine the price of a forward-start contract. This contract represents the building block for both cliquet options and variance swaps. All these contracts share the common feature of being pure variance contracts. The first step consists in considering a forward-start call option whose payoff at maturity T is defined as

$$FS\ \mathrm{Call}(T) = \left(\frac{S_T}{S_t} - K\right)^+,$$

where S_t is the stock price at a fixed date t, $0 \le t \le T$. In the following, we follow the (single volatility factor) presentation of Hong (2004). By risk-neutral valuation, the initial price of this option is given by

$$FS\ \mathrm{Call}(0) = \mathbb{E}\left[\exp\{-rT\}\left(\frac{S_T}{S_t} - K\right)^+\right].$$

In particular, in the Black and Scholes framework where volatility is constant, one obtains

$$FS\ \mathrm{Call}(0) = \exp\{-rt\}B\&S(K, 1, T-t, \sigma_{\mathrm{BS}}),$$

where $B\&S(K, 1, T-t, \Sigma_{BS})$ denotes the Black–Scholes price formula of the corresponding call option computed with spot price (at time t) $S_t = 1$: note that, in this way, the forward-start contract price is independent of the level of the underlying asset and depends only on the volatility. Let us consider the forward log-return

$$Y_{t,T} = \ln\frac{S_T}{S_t} = Y_T - Y_t,$$

so that the price of the forward-start call option is given by

$$FS\ \mathrm{Call}(0) = \mathbb{E}[\exp\{-rT\}(\exp\{Y_{t,T}\} - \exp\{k\})^+],$$

with, as before, $k = \ln K$. Let us denote by $\Phi_{\gamma,0}(t,T)$ the characteristic function of the log-return $Y_{t,T}$, i.e. the so-called *forward characteristic function*, defined as

$$\Phi_{\gamma,0}(t,T) := \mathbb{E}[\exp\{i\gamma Y_{t,T}\}]. \tag{16}$$

The modified option price is given by

$$c_{t,T}(k) = \exp\{\alpha k\} FS\ \mathrm{Call}(0),$$

and its Fourier transform

$$\begin{aligned}\psi_{t,T}(v) &= \int_{-\infty}^{+\infty} \exp\{ivk\} c_{t,T}(k)\mathrm{d}k \\ &= \exp\{-rT\} \frac{\Phi_{(v-(\alpha+1)i),0}(t,T)}{(\alpha+iv)(\alpha+1+iv)}.\end{aligned} \tag{17}$$

Therefore, again, we realize that in order to price a forward-start call option, it is sufficient to compute the forward characteristic function $\Phi_{\gamma,0}(t,T)$. This computation will involve the characteristic function of the Wishart process, which is given in the following.

Proposition 3.2: *Given a real symmetric matrix D, the conditional characteristic function of the Wishart process Σ_t is given by*

$$\begin{aligned}\Phi^{\Sigma}_{D,t}(\tau) &= \mathbb{E}_t \exp\{i\, Tr[D\Sigma_{t+\tau}]\} \\ &= \exp\{Tr[B(\tau)\Sigma_t] + C(\tau)\},\end{aligned} \tag{18}$$

where the deterministic complex-valued functions $B(\tau) \in M_n(\mathbb{C}^n)$, $C(\tau) \in \mathbb{C}$ are given by

$$\begin{aligned}B(\tau) &= (iDB_{12}(\tau) + B_{22}(\tau)t)^{-1}(iDB_{11}(\tau) + B_{21}(\tau)), \\ C(\tau) &= \mathrm{Tr}[\Omega\Omega^T \int_0^\tau B(s)\mathrm{d}s],\end{aligned} \tag{19}$$

with

$$\begin{pmatrix} B_{11}(\tau) & B_{12}(\tau) \\ B_{21}(\tau) & B_{22}(\tau) \end{pmatrix} = \exp \tau \begin{pmatrix} M & -2Q^TQ \\ 0 & -M^T \end{pmatrix}.$$

Proof: See appendix A. □

We now have all the ingredients to compute the forward characteristic function of the log-returns $\Phi_{\gamma,0}(t,T)$:

$$\begin{aligned}\Phi_{\gamma,0}(t,T) &= \mathbb{E}[\exp\{i\gamma Y_{t,T}\}] \\ &= \mathbb{E}[\mathbb{E}_t[\exp\{i\gamma(Y_T - Y_t)\}]] \\ &= \mathbb{E}[\exp\{-i\gamma Y_t\}\mathbb{E}_t[\exp\{i\gamma Y_T\}]] \\ &= \mathbb{E}[\exp\{-i\gamma Y_t\}\exp\{\mathrm{Tr}[A(T-t)\Sigma_t] \\ &\quad + i\gamma Y_t + c(T-t)\}] \\ &= \exp\{c(T-t)\}\mathbb{E}[\exp\{\mathrm{Tr}[A(T-t)\Sigma_t]\}] \\ &= \exp\{\mathrm{Tr}[B(t)\Sigma_0] + C(t) + c(T-t)\},\end{aligned}$$

where the last equality comes from (18), and $B(t)$ is given by (19) with $\tau = t$ and boundary condition

$$B(0) = A(T-t).$$

With the function $\Phi_{\gamma,0}(t,T)$, it suffices to plug into (17) and apply the FFT in order to determine the forward-start call price.

4. Numerical illustration

In this section we provide some examples proving that the Wishart specification for the volatility has greater flexibility than the (single-factor as well as multi-factor) Heston specification. We quote option prices using Black&Scholes volatility, which is common practice in the market. Let us denote by V_t the instantaneous volatility in the (single factor) Heston model, the dynamics of which are given by

$$\mathrm{d}V_t = \kappa(\theta - V_t)\mathrm{d}t + \epsilon\sqrt{V_t}\mathrm{d}W_t^2,$$

where θ represents the long-term volatility, κ is the mean reversion parameter, ϵ is the volatility of volatility parameter (also called vol-of-vol), ρ is the correlation between the volatility and the stock, V_0 is the initial spot volatility and W_t^2 is the (scalar) Brownian motion of the volatility process, which in the Heston model is assumed to be correlated with the Brownian motion W_t^1 driving the asset returns.

We proceed as follows:

1. We consider the simplest modification of the previous choice which allows us to reproduce a volatility surface that cannot be generated by the (single factor) Heston model;
2. We compare our model with the multi-dimensional version of the Heston model when the volatility is driven by a two-dimensional affine process, the state space domain of which is $\mathbb{R}_+^2$ as classified by Dai and Singleton (2000): in particular, we show that, within our Wishart specification, we have an additional degree of freedom in order to capture the stochasticity of the skew effect by preserving analytic tractability.

4.1. *Wishart embedding Heston volatility*

The Heston model can easily be nested in the Wishart model for a specific choice of the parameters. When all matrices involved in the Wishart dynamics are proportional to the identity matrix, it is straightforward to see that $\mathrm{Tr}(\Sigma_t)$ follows a square root process and both models produce the same smile at different maturities.

The original motivation for introducing multifactor models comes from the observation that the dynamics of the implied volatility surface, as well as the realized volatility process, are driven by at least two stochastic factors. The simplest example of an implied volatility pattern that cannot be reproduced by a single factor model is obtained by considering a diagonal model while specifying two different mean reversion parameters in the (diagonal) matrix M. In particular, if we choose $M_{11}=-3$ and $M_{22}=-0.333$, then we can associate with the element Σ_{11} the meaning of a short-term factor, while Σ_{22} has an impact on the long-term volatility. Let us take the following values:

$$M=\begin{pmatrix}-3 & 0\\ 0 & -0.333\end{pmatrix},\qquad R=\begin{pmatrix}-0.7 & 0\\ 0 & -0.7\end{pmatrix}$$
$$Q=\begin{pmatrix}0.25 & 0\\ 0 & 0.25\end{pmatrix},\qquad \Sigma_0=\begin{pmatrix}0.01 & 0\\ 0 & 0.01\end{pmatrix},\qquad (20)$$

and $\beta=3$. In this case we see that, in the Wishart model, the long-term volatility increases. This additional degree of freedom is interesting from a practical point of view because, in the market, there are some long-term products such as forward-start options and cliquet options, whose maturity can be much longer than one year. It is then important to obtain prices for such contracts in closed form, in order to investigate the properties of the long-term smile. Observe that, typically, long-term volatility is higher than short-term volatility. Now we want to generate the same volatility smile with the Heston model, so in order to fit the implied volatility at 2 years we have to set $\theta=0.38^2$, while the other parameters are $\kappa=6$, $\sigma_0=0.15$, $\epsilon=0.5$ $\rho=-0.7$. However, an increase in the long-term volatility also induces an increase in the 3 month volatility, so that the short-term fit for the implied volatility is unsatisfactory, as illustrated in figure 1.

On the other hand, we can fit perfectly the short-term volatility produced by the Wishart model by setting $\theta=0.295^2$. However, in this case the long-term volatility decreases and this time we arrive at an unsatisfactory fit of the long-term implied volatility level, as shown in figure 2.

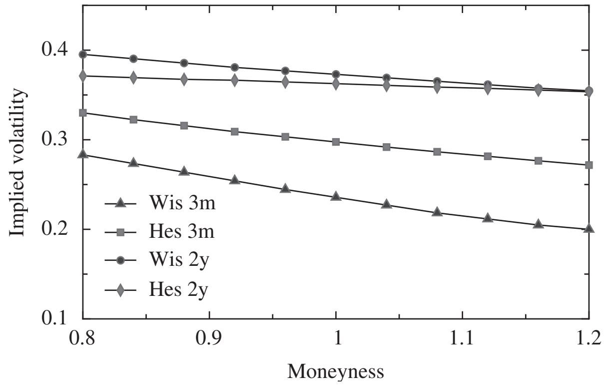

Figure 1. Implied volatility for the Wishart model (Wis) and Heston (Hes) model. Option maturities are 3 months (3m) and 2 years (2y). Moneyness is defined as K/S_0 where S_0 is the initial spot value.

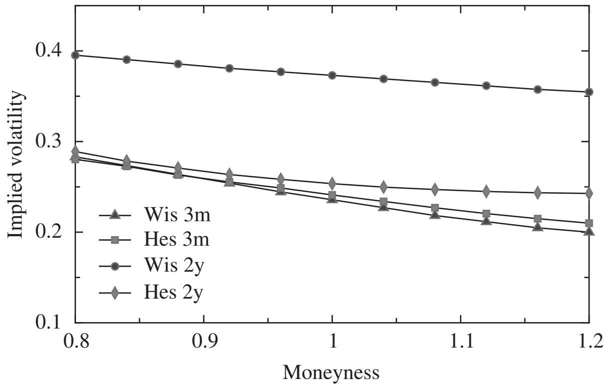

Figure 2. Implied volatility for the Wishart model (Wis) and Heston (Hes) model.

4.2. Wishart versus $\mathbb{A}_2(3)$-Heston volatility

Note that the above observation is not sufficient to justify the introduction of the previous Wishart (matrix) affine model given by (20), the covariance matrix of which can also be reproduced† using the following (vector) affine model, which belongs to the canonical class $\mathbb{A}_2(3)$ of Dai and Singleton (2000):

$$\begin{aligned}
\mathrm{d}X_t^1 &= \kappa_1(\theta_1-X_t^1)\mathrm{d}t+\epsilon_1\sqrt{X_t^1}\mathrm{d}W_t^1,\\
\mathrm{d}X_t^2 &= \kappa_2(\theta_2-X_t^2)\mathrm{d}t+\epsilon_2\sqrt{X_t^2}\mathrm{d}W_t^2,\\
\mathrm{d}Y_t &= \left(r-\frac{1}{2}(X_t^1+X_t^2)\right)\mathrm{d}t+\rho_1\sqrt{X_t^1}\mathrm{d}W_t^1+\rho_2\sqrt{X_t^2}\mathrm{d}W_t^2\\
&\quad+\sqrt{(1-\rho_1^2)X_t^1+(1-\rho_2^2)X_t^2}\mathrm{d}B_t,
\end{aligned}$$

where W_t^1, W_t^2 and B_t are independent Brownian motions. In fact, both models lead to the same covariance matrix where the state space domain of the positive factors Σ_t^{11}, Σ_t^{22} (respectively X_t^1, X_t^2 for the $\mathbb{A}_2(3)$ model) is $\mathbb{R}_2^+$.

We remark that, in this $\mathbb{A}_2(3)$ model, once the short-term and long-term implied volatility levels are fitted, there are no more free parameters in order to describe the stochasticity of the leverage effect (which leads to a stochastic skew in the spirit of Carr and Wu (2004)): in fact, it turns out that the correlation between the asset's returns and their volatilities is stochastic, but it depends (only) on the volatility factors:

$$\begin{aligned}
&\mathrm{Corr_t}(\mathrm{Noise}(\mathrm{d}Y),\mathrm{Noise}(\mathrm{Vol}(\mathrm{d}Y)))\\
&\quad=\frac{\langle Y, X^1+X^2\rangle_t}{\sqrt{\langle Y\rangle_t\langle X^1+X^2\rangle_t}}\\
&\quad=\frac{\mathbb{E}_t[(\sqrt{X_t^1}\mathrm{d}Z_t^1+\sqrt{X_t^2}\mathrm{d}Z_t^2)(\epsilon_1\sqrt{X_t^1}\mathrm{d}W_t^1+\epsilon_2\sqrt{X_t^2}\mathrm{d}W_t^2)]}{\sqrt{X_t^1+X_t^2}\sqrt{\epsilon_1^2X_t^1+\epsilon_2^2X_t^2}}\\
&\quad=\frac{\rho_1\epsilon_1X_t^1+\rho_2\epsilon_2X_t^2}{\sqrt{X_t^1+X_t^2}\sqrt{\epsilon_1^2X_t^1+\epsilon_2^2X_t^2}}. \qquad (21)
\end{aligned}$$

†We thank an anonymous referee for this observation.

4.2.1. Stochastic leverage effect in the Wishart model. In order to compute the analogue of (21) in the general Wishart model, let us now consider the correlation between the stock noise and the noise driving its *scalar* volatility, represented by $\mathrm{Tr}(\Sigma_t)$: this is computed in the following.

Proposition 4.1: *The stochastic correlation between the stock noise and the volatility noise in the Wishart model is given by*

$$\rho_t = \frac{\mathrm{Tr}[RQ\Sigma_t]}{\sqrt{\mathrm{Tr}[\Sigma_t]}\sqrt{\mathrm{Tr}[Q^{\mathrm{T}}Q\Sigma_t]}}. \tag{22}$$

Proof: See Appendix A. □

The previous proposition highlights the analytical tractability of the Wishart specification: in fact, within the Wishart model it is possible to handle the (stochastic) correlation (and, in turn, the stochastic skew effect) by means of the product RQ.

- When the product RQ is a multiple of the identity matrix, we recover the usual constant correlation parameter as in the (single factor) Heston model as well as in the multi-Heston model with $\rho_1 = \rho_2$ and $\epsilon_1 = \epsilon_2$.
- When the product RQ is diagonal, then the Wishart model is qualitatively equivalent to a $\mathbb{A}_2(3)$ multi-Heston model, in the sense that the stochastic correlation depends only on the volatility factors Σ_t^{11}, Σ_t^{22} (while the off-diagonal factor Σ_t^{12} does not appear): in fact, in this case (22) reads

$$\rho_t^{\mathbb{A}_2(3)} = \frac{R_{11}Q_{11}\Sigma_t^{11} + R_{22}Q_{22}\Sigma_t^{22}}{\sqrt{\Sigma_t^{11}+\Sigma_t^{22}}\sqrt{Q_{11}^2\Sigma_t^{11}+Q_{22}^2\Sigma_t^{22}}},$$

which is exactly the analogue of (21).
- When the product RQ is not diagonal (i.e. when R or Q is not diagonal), from (22) it turns out that ρ_t depends also on the off-diagonal volatility term Σ_t^{12}:

$$\rho_t^{\mathrm{Wis}} = \rho_t^{\mathbb{A}_2(3)} + \frac{Q_{22}R_{12}}{\sqrt{\Sigma_t^{11}+\Sigma_t^{22}}\sqrt{Q_{11}^2\Sigma_t^{11}+Q_{22}^2\Sigma_t^{22}}}\Sigma_t^{12};$$

that is, in the Wishart specification, the off-diagonal elements of the vol-of-vol matrix Q and the correlation matrix R are additional degrees of freedom w.r.t. the $\mathbb{A}_2(3)$ multi-Heston model in order to control the stochasticity of (the correlation and, in turn, of) the leverage effect once the short-term and long-term implied volatility levels are fitted. This represents a suitable feature of a stochastic volatility model that can be calibrated on Risk Reversal quotes in the spirit of Carr and Wu (2004).

This model cannot be nested in a $\mathbb{A}_2(3)$ since the admissible domains of $\mathbb{A}_2(3)$ and the Wishart model are crucially different: while the former has the linear structure of $\mathbb{R}_+^m \times \mathbb{R}^{(n-m)}$, the Wishart domain is the symmetric cone of positive definite matrices (see also Grasselli and Tebaldi (2008)), which is nonlinear in the factors (the domain of Σ_t^{12} is given by the set $\Sigma_t^{11}\Sigma_t^{22} - (\Sigma_t^{12})^2 > 0$). This nonlinearity allows the Wishart specification to reproduce new effects w.r.t. the classic (vector) affine models.

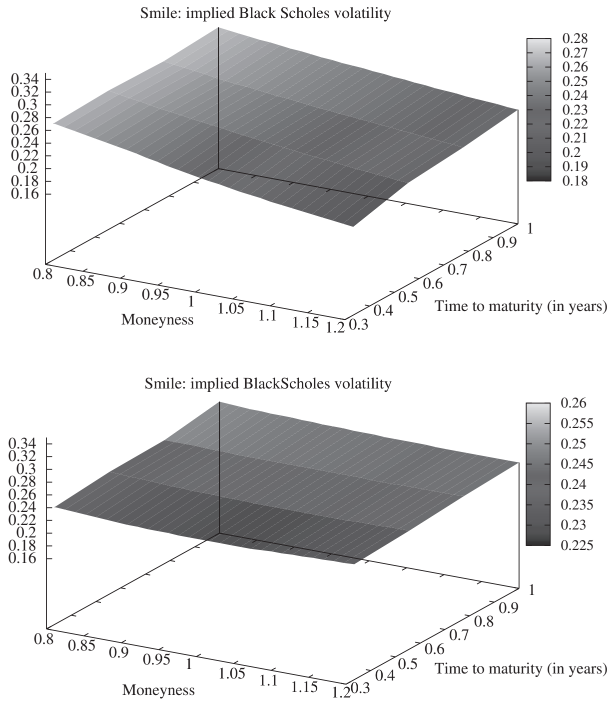

Figure 3. Wishart implied volatility for R_1 (left) and R_2 (right).

4.2.2. The impact of R on the stochastic leverage effect. In the following examples we compare the Wishart specification with diagonal matrix parameters (equivalent to the $\mathbb{A}_2(3)$ multi-Heston model) with a non-diagonal one, in order to highlight the additional flexibility introduced by off-diagonal terms.

It is well known that, in the Heston model, the skew is related to the (negative) correlation between the volatility and the stock price. Taking the matrices

$$M = \begin{pmatrix} -5 & 0 \\ 0 & -3 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.35 & 0 \\ 0 & 0.25 \end{pmatrix},$$
$$\Sigma_0 = \begin{pmatrix} 0.02 & 0 \\ 0 & 0.02 \end{pmatrix}, \quad R_1 = \begin{pmatrix} -0.7 & 0 \\ 0 & -0.5 \end{pmatrix},$$
$$R_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{23}$$

and $\beta = 3$ in the Wishart model, we obtain for R_1 (R_2) the left (right)-hand side of figure 3, which confirms that R is strictly related to the leverage effect in both the Wishart and $\mathbb{A}_2(3)$ multi-Heston models. In particular, the short- and long-term implied volatility levels can be

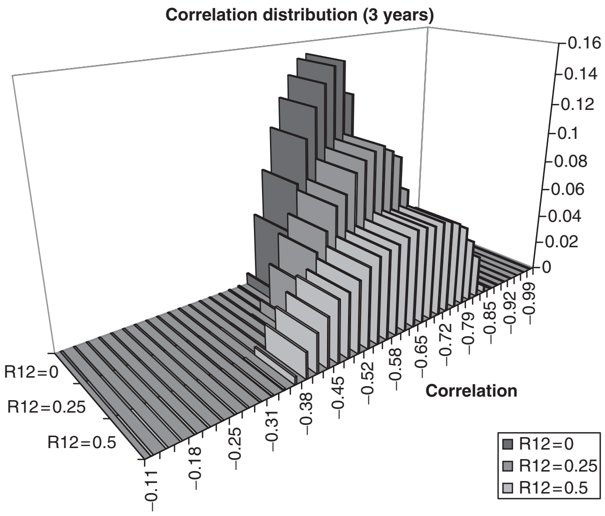

Figure 4. Distribution functions of the correlation process in the Wishart model with non-diagonal matrix R.

fitted using the diagonal terms in matrix R_1 as well as the parameters ρ_1, ρ_2 in the $\mathbb{A}_2(3)$ multi-Heston model.

Now let us consider the Wishart model with the non-diagonal correlation matrix R_3 given by

$$R_3 = \begin{pmatrix} -0.7 & R_{12} \\ 0 & -0.5 \end{pmatrix}.$$

From (22) we obtain

$$\rho_t^{\mathrm{Wis}} = \frac{-0.7(0.35)\Sigma_t^{11} + 0.25(R_{12}\Sigma_t^{12} - 0.5\Sigma_t^{22})}{\sqrt{\Sigma_t^{11} + \Sigma_t^{22}}\sqrt{(0.35)^2\Sigma_t^{11} + (0.25)^2\Sigma_t^{22}}}.$$

The presence of the off-diagonal parameter R_{12} introduces the new factor Σ_t^{12} into the correlation, which is described by an additional source of uncertainty. In figure 4 we consider the distribution of the correlation process for different values of R_{12}: note that the distribution becomes more sparse as R_{12} increases, a new effect that cannot be reproduced by the $\mathbb{A}_2(3)$ multi-Heston model.

5. Conclusion

We have shown that the multifactor volatility extension of the Heston model considered in this paper is sufficiently flexible to take into account correlations with the underlying asset returns. It preserves analytical tractability, i.e. a closed form for the conditional characteristic function, and a linear factor structure that can potentially be very useful in the calibration procedure. Finally, our numerical results show that the flexibility induced by the additional factors allows a better fit of the smile-skew effect at both long and short maturities. In particular, in contrast to the Heston model, the Wishart specification permits a separate fit of both long-term and short-term skew (or volatility level), so that we can allow for more complex term structures for the implied volatility surface. Future work will be devoted to the calibration of this model to option prices and further studies are needed in order to illustrate the improvements in calibration with respect to the (scalar and $\mathbb{A}_2(3)$ multi-factor) Heston model. From a financial econometric perspective, on the other hand, this model seems to be a natural candidate for the analysis and description of volatility and stochastic correlation effects on risk premia valued by the market.

Acknowledgements

We thank Fabio Mercurio and the participants at the 2006 Ascona Meeting and the Séminaire Bachelier of Paris for helpful comments. We also thank two anonymous referees for careful reading of the paper. An earlier draft of this paper appeared under the title 'Wishart multi-dimensional stochastic volatility'.

References

Bergomi, L., Smile dynamics. *Risk*, 2004, **September**, 117–123.
Bergomi, L., Smile dynamics II. *Risk*, 2005, **October**, 67–73.
Bru, M.F., Wishart Processes. *J. Theor. Probab.*, 1991, **4**, 725–743.
Cont, R. and da Fonseca, J., Dynamics of implied volatility surfaces. *Quant. Finan.*, 2002, **2**, 45–60.
Carr, P. and Madan, D.B., Option valuation using the fast Fourier transform. *J. Comput. Finan.*, 1999, **2**, 61–73.
Carr, P. and Wu, L., Stochastic skew in currency options, 2004 (Preprint).
Dai, Q. and Singleton, K., Specification analysis of affine term structure models. *J. Finan.*, 2000, **55**, 1943–1978.
Duffie, D. and Kan, R., A yield-factor model of interest rates. *Math. Finan.*, 1996, **6**, 379–406.
Duffie, D., Pan, J. and Singleton, K., Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 2000, **68**, 1343–1376.
Freiling, G., A survey of nonsymmetric Riccati equations. *Linear Algebra Applic.*, 2002, **351–352**, 243–270.
Gourieroux, C. and Sufana, R., *Wishart quadratic term structure models* CREF, pp. 03–10, 2003 (HEC Montreal).
Gourieroux, C. and Sufana, R. Derivative pricing with multivariate stochastic volatility: application to credit risk, 2004, *Working paper*, CREST.
Grasselli, M. and Tebaldi, C., Soluable affine term structure models. *Math. Finan.*, 2008, **18**, 135–153.
Heston, S.L., A closed-form solution for option with stochastic volatility with applications to bond and currency options. *Rev. Finan. Stud.*, 1993, **6**, 327–343.
Hipp, C., Options for guaranteed index-linked life insurance, *in AFIR 1996 Proceedings*, Vol II, 1996, pp. 1463–1483.
Hong, G., Forward smile and derivative pricing, 2004 (UBS Preprint).

Appendix A: Proofs

Proof: (proof of proposition 3.1) The only non-trivial term in (7) comes from the covariation

$$\mathrm{d}\langle \Sigma^{ij}, Y\rangle_t, \quad \text{for } i,j = 1, \ldots, n.$$

It will be useful to introduce the square root matrix $\sigma_t := \sqrt{\Sigma_t}$, so that

$$\Sigma_t^{ij} = \sum_{l=1}^{n} \sigma_t^{il}\sigma_t^{lj} = \sum_{l=1}^{n} \sigma_t^{il}\sigma_t^{jl},$$

where the last equality follows from the symmetry of σ_t. Now we identify the covariation terms with the coefficients of $(\partial^2/\partial x_{ij}, \partial y)$, thus obtaining

$$\begin{aligned}
\mathrm{d}\langle \Sigma^{ij}, Y\rangle_t &= \mathbb{E}_t\left[\left(\sum_{l,k=1}^{n} \sigma_t^{il}\mathrm{d}W_{lk}Q_{kj} + \sum_{l,k=1}^{n} \sigma_t^{jl}\mathrm{d}W_{lk}Q_{ki}\right)\right.\\
&\qquad \left.\times\left(\sum_{l,k,h=1}^{n} \sigma_t^{lk}\mathrm{d}W_{kh}R_{lh}\right)\right]\\
&= \sum_{l,k,h=1}^{n} (\sigma_t^{il}Q_{kj} + \sigma_t^{jl}Q_{ki})\sigma_t^{hl}R_{hk}\mathrm{d}t\\
&= \sum_{k,h=1}^{n}\left(\left(\sum_{l=1}^{n}\sigma_t^{il}\sigma_t^{hl}\right)Q_{kj} + \left(\sum_{l=1}^{n}\sigma_t^{jl}\sigma_t^{hl}\right)Q_{ki}\right)R_{hk}\mathrm{d}t\\
&= \sum_{k,h=1}^{n} (\Sigma_t^{ih}Q_{kj} + \Sigma_t^{jh}Q_{ki})R_{hk}\mathrm{d}t,
\end{aligned}$$

which corresponds to the coefficient of the term $(\partial^2/\partial x_{ij}, \partial yt)$, since

$$2\mathrm{Tr}[\Sigma RQD]\frac{\partial}{\partial y} = 2\sum_{i,j,k,h=1}^{n} D^{ij}\Sigma^{jh}R_{hk}Q_{ki}\frac{\partial}{\partial y},$$

and since, by definition, D is symmetric. □

Proof (proof of proposition 3.2)**:** We repeat the reasoning as in (4) where this time there is no dependence on Y_t, so that the (complex-valued non-symmetric) Matrix Riccati ODE satisfied by $B(\tau)$ becomes

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}B(\tau) &= B(\tau)M + M^{\mathrm{T}}B(\tau) + 2B(\tau)Q^{\mathrm{T}}QB(\tau),\\
B(0) &= iD,
\end{aligned}$$

while

$$C(\tau) = \mathrm{Tr}[\Omega\Omega^{\mathrm{T}}\int_0^{\tau} B(s)\,\mathrm{d}s].$$

Applying the linearization procedure, we arrive at the explicit solution $B(\tau) = F(\tau)^{-1}G(\tau)$, with

$$\begin{aligned}
(\,G(\tau)\quad F(\tau)\,) &= (\,G(0)\quad F(0)\,)\exp\tau\begin{pmatrix} M & -2Q^{\mathrm{T}}Q\\ 0 & -M^{\mathrm{T}}\end{pmatrix}\\
&= (\,B(0)\quad \mathbb{I}_n\,)\exp\tau\begin{pmatrix} M & -2Q^{\mathrm{T}}Q\\ 0 & -M^{\mathrm{T}}\end{pmatrix}\\
&= (\,iDB_{11}(\tau) + B_{21}(\tau)\quad iDB_{12}(\tau) + B_{22}(\tau)\,),
\end{aligned}$$

which completes the proof. □

Proof: (proof of proposition 4.1) The first step consists of finding the stock noise:

$$\begin{aligned}
\frac{\mathrm{d}S_t}{S_t} &= r\mathrm{d}t + \mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}Z_t]\\
&= r\mathrm{d}t + \sqrt{\mathrm{Tr}[\Sigma_t]}\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}Z_t]}{\sqrt{\mathrm{Tr}[\Sigma_t]}}\\
&= r\mathrm{d}t + \sqrt{\mathrm{Tr}[\Sigma_t]}\mathrm{d}z_t,
\end{aligned}$$

where z_t is a standard Brownian motion. We now compute the (scalar) standard Brownian motion w_t driving the process $\mathrm{Tr}[\Sigma_t]$:

$$\begin{aligned}
\mathrm{dTr}[\Sigma_t] &= (\mathrm{Tr}[\Omega\Omega^{\mathrm{T}}] + 2\mathrm{Tr}[M\Sigma_t])\,\mathrm{d}t + 2\mathrm{Tr}[\sqrt{\Sigma_t}\,\mathrm{d}W_tQ]\\
&= \ldots\mathrm{d}t + 2\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}W_tQ]}{\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\\
&= \ldots\mathrm{d}t + 2\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}\,\mathrm{d}w_t,
\end{aligned}$$

where we have used the fact that

$$\begin{aligned}
\mathrm{d}\langle\mathrm{Tr}[\Sigma_.]\rangle_t &= \sum_{ij}\mathrm{Cov}_t(e_i^{\mathrm{T}}\mathrm{d}\Sigma_te_i, e_j^{\mathrm{T}}\,\mathrm{d}\Sigma_te_j)\\
&= 4\sum_{ij}\mathrm{Cov}_t(e_i^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{d}W_tQe_i, e_j^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{d}W_tQe_j)\\
&= 4\sum_{ij}\mathbb{E}_t[e_i^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{d}W_tQe_ie_j^{\mathrm{T}}Q^{\mathrm{T}}\mathrm{d}W_t^{\mathrm{T}}\sqrt{\Sigma_t}e_j]\\
&= 4\sum_{ij}e_i^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{Tr}[Qe_ie_j^{\mathrm{T}}Q^{\mathrm{T}}]\sqrt{\Sigma_t}e_j\mathrm{d}t\\
&= 4\sum_{ij}\mathrm{Tr}[Q^{\mathrm{T}}Qe_ie_j^{\mathrm{T}}]e_i^{\mathrm{T}}\Sigma_te_j\mathrm{d}t\\
&= 4\sum_{ij}e_j^{\mathrm{T}}Q^{\mathrm{T}}Qe_ie_i^{\mathrm{T}}\Sigma_te_j\mathrm{d}t\\
&= 4\sum_{j}e_j^{\mathrm{T}}Q^{\mathrm{T}}Q\Sigma_te_j\mathrm{d}t\\
&= 4\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]\mathrm{d}t,
\end{aligned}$$

and we have used

$$\mathbb{E}_t[\mathrm{d}W_tQe_ie_j^{\mathrm{T}}Q^{\mathrm{T}}\mathrm{d}W_t^{\mathrm{T}}] = \mathrm{Tr}[Qe_ie_j^{\mathrm{T}}Q^{\mathrm{T}}]\mathrm{d}t,$$

since from proposition 2.1

$$\mathbb{E}_t[\mathrm{d}W_t\alpha\beta^{\mathrm{T}}\mathrm{d}W_t^{\mathrm{T}}] = \alpha^{\mathrm{T}}\beta\mathbb{I}\mathrm{d}t = \mathrm{Tr}[\alpha\beta^{\mathrm{T}}]\mathbb{I}\mathrm{d}t.$$

In conclusion, the correlation between the stock noise and the volatility noise in the Wishart model is *stochastic* and corresponds to the correlation between the

Brownian motions z_t and w_t, the covariation of which is given by

$$\begin{aligned}
\mathrm{Cov}_t(\mathrm{d}z_t,\mathrm{d}w_t) &= \mathrm{Cov}_t\left(\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}Z_t]}{\sqrt{\mathrm{Tr}[\Sigma_t]}},\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}W_tQ]}{\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\right)\\
&= \mathbb{E}_t\left[\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\mathrm{d}W_tR^{\mathrm{T}}]}{\sqrt{\mathrm{Tr}[\Sigma_t]}}\frac{\mathrm{Tr}[\sqrt{\Sigma_t}\,\mathrm{d}W_tQ]}{\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\right]\\
&= \frac{\sum_{ij}\mathrm{Cov}_t(e_i^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{d}WR^{\mathrm{T}}e_i,e_j^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{d}W_tQe_j)}{\sqrt{\mathrm{Tr}[\Sigma_t]}\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\\
&= \frac{1}{\sqrt{\mathrm{Tr}[\Sigma_t]}\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\\
&\quad\times\sum_{ij}e_i^{\mathrm{T}}\sqrt{\Sigma_t}\mathrm{Tr}[R^{\mathrm{T}}e_ie_j^{\mathrm{T}}Q^{\mathrm{T}}]\sqrt{\Sigma_t}e_j\mathrm{d}t\\
&= \frac{1}{\sqrt{\mathrm{Tr}[\Sigma_t]}\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}R^{\mathrm{T}}]\mathrm{d}t\\
&= \frac{\mathrm{Tr}[\Sigma_tRQ]}{\sqrt{\mathrm{Tr}[\Sigma_t]}\sqrt{\mathrm{Tr}[\Sigma_tQ^{\mathrm{T}}Q]}}\,\mathrm{d}t
\end{aligned}$$

□

Appendix B: The two-dimensional case

In this appendix we develop the computations in (6) in the case $n=2$. This means that the Wishart process Σ_t satisfies the following SDE:

$$\begin{aligned}
\mathrm{d}\Sigma_t &= \mathrm{d}\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}\\
&= \left(\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}\begin{pmatrix}\Omega_{11} & \Omega_{21}\\ \Omega_{12} & \Omega_{22}\end{pmatrix}\right.\\
&\quad+\begin{pmatrix}M_{11} & M_{12}\\ M_{21} & M_{22}\end{pmatrix}\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}\\
&\quad\left.+\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}\begin{pmatrix}M_{11} & M_{21}\\ M_{12} & M_{22}\end{pmatrix}\right)\mathrm{d}t\\
&\quad+\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}^{1/2}\begin{pmatrix}\mathrm{d}W_t^{11} & \mathrm{d}W_t^{12}\\ \mathrm{d}W_t^{21} & \mathrm{d}W_t^{22}\end{pmatrix}\begin{pmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}\\
&\quad+\begin{pmatrix}Q_{11} & Q_{21}\\ Q_{12} & Q_{22}\end{pmatrix}\begin{pmatrix}\mathrm{d}W_t^{11} & \mathrm{d}W_t^{21}\\ \mathrm{d}W_t^{12} & \mathrm{d}W_t^{22}\end{pmatrix}\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}^{1/2}.
\end{aligned}$$

Let

$$\begin{pmatrix}\sigma_t^{11} & \sigma_t^{12}\\ \sigma_t^{12} & \sigma_t^{22}\end{pmatrix}:=\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}^{1/2},$$

so that

$$\sigma_t^2=\Sigma_t=\begin{pmatrix}(\sigma_t^{11})^2+(\sigma_t^{12})^2 & \sigma_t^{11}\sigma_t^{12}+\sigma_t^{12}\sigma_t^{22}\\ \sigma_t^{11}\sigma_t^{12}+\sigma_t^{12}\sigma_t^{22} & (\sigma_t^{12})^2+(\sigma_t^{22})^2\end{pmatrix}. \tag{B1}$$

We obtain

$$\begin{aligned}
\mathrm{d}\Sigma_t^{11} &= (.)\mathrm{d}t+2\sigma_t^{11}(Q_{11}\mathrm{d}W_t^{11}+Q_{21}\mathrm{d}W_t^{12})\\
&\quad+2\sigma_t^{12}(Q_{11}\mathrm{d}W_t^{21}+Q_{21}\mathrm{d}W_t^{22}),\\
\mathrm{d}\Sigma_t^{12} &= (.)\mathrm{d}t+\sigma_t^{11}(Q_{12}\mathrm{d}W_t^{11}+Q_{22}\mathrm{d}W_t^{12})\\
&\quad+\sigma_t^{12}(Q_{12}\mathrm{d}W_t^{21}+Q_{22}\mathrm{d}W_t^{22})\\
&\quad+\sigma_t^{12}(Q_{11}\mathrm{d}W_t^{11}+Q_{21}\mathrm{d}W_t^{12})\\
&\quad+\sigma_t^{22}(Q_{11}\mathrm{d}W_t^{21}+Q_{21}\mathrm{d}W_t^{22}),\\
\mathrm{d}\Sigma_t^{22} &= (.)\mathrm{d}t+2\sigma_t^{12}(Q_{12}\mathrm{d}W_t^{11}+Q_{22}\mathrm{d}W_t^{12})\\
&\quad+2\sigma_t^{22}(Q_{12}\mathrm{d}W_t^{21}+Q_{22}\mathrm{d}W_t^{22}),
\end{aligned}$$

and using (B1):

$$\begin{aligned}
\mathrm{d}\langle\Sigma^{11},\Sigma^{11}\rangle_t &= 4\Sigma_t^{11}(Q_{11}^2+Q_{21}^2)\mathrm{d}t,\\
\mathrm{d}\langle\Sigma^{12},\Sigma^{12}\rangle_t &= (\Sigma_t^{11}(Q_{12}^2+Q_{22}^2)\\
&\quad+2\Sigma_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})\\
&\quad+\Sigma_t^{22}(Q_{11}^2+Q_{21}^2))\mathrm{d}t,\\
\mathrm{d}\langle\Sigma^{22},\Sigma^{22}\rangle_t &= 4\Sigma_t^{22}(Q_{12}^2+Q_{22}^2)\mathrm{d}t,\\
\mathrm{d}\langle\Sigma^{11},\Sigma^{12}\rangle_t &= (2\Sigma_t^{11}(Q_{11}Q_{12}+Q_{21}Q_{22})\\
&\quad+2\Sigma_t^{12}(Q_{11}^2+Q_{21}^2))\mathrm{d}t,\\
\mathrm{d}\langle\Sigma^{11},\Sigma^{22}\rangle_t &= 4\Sigma_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})\mathrm{d}t,\\
\mathrm{d}\langle\Sigma^{12},\Sigma^{22}\rangle_t &= 2(\Sigma_t^{12}(Q_{12}^2+Q_{22}^2)\\
&\quad+\Sigma_t^{22}(Q_{11}Q_{12}+Q_{21}Q_{22}))\mathrm{d}t.
\end{aligned}$$

On the other hand, from (6) we can identify the coefficient of $(\partial^2/\partial\Sigma_{ij},\ \partial\Sigma_{lk})$ in the trace of the matrix $2\Sigma_tDQ^TQD$, that is

$$\begin{aligned}
&2\begin{pmatrix}\Sigma_t^{11} & \Sigma_t^{12}\\ \Sigma_t^{12} & \Sigma_t^{22}\end{pmatrix}\begin{pmatrix}\dfrac{\partial}{\partial\Sigma^{11}} & \dfrac{\partial}{\partial\Sigma^{12}}\\ \dfrac{\partial}{\partial\Sigma^{12}} & \dfrac{\partial}{\partial\Sigma^{22}}\end{pmatrix}\begin{pmatrix}Q_{11} & Q_{21}\\ Q_{12} & Q_{22}\end{pmatrix}\\
&\begin{pmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}\begin{pmatrix}\dfrac{\partial}{\partial\Sigma^{11}} & \dfrac{\partial}{\partial\Sigma^{12}}\\ \dfrac{\partial}{\partial\Sigma^{12}} & \dfrac{\partial}{\partial\Sigma^{22}}\end{pmatrix}.
\end{aligned}$$

After some computations, we obtain

$$\begin{aligned}
&\mathrm{Tr}[2\Sigma_tDQ^{\mathrm{T}}QD]\\
&\quad=2\mathrm{Tr}[\Sigma_tDQ^{\mathrm{T}}QD]=2\Sigma_t^{11}(Q_{11}^2+Q_{21}^2)\frac{\partial^2}{(\partial\Sigma^{11})^2}\\
&\qquad+2(\Sigma_t^{11}(Q_{12}^2+Q_{22}^2)+2\Sigma_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})\\
&\qquad+\Sigma_t^{22}(Q_{11}^2+Q_{21}^2))\frac{\partial^2}{(\partial\Sigma^{12})^2}+2\Sigma_t^{22}(Q_{12}^2+Q_{22}^2)\frac{\partial^2}{(\partial\Sigma^{22})^2}\\
&\qquad+4(\Sigma_t^{11}(Q_{11}Q_{12}+Q_{21}Q_{22})+\Sigma_t^{12}(Q_{11}^2+Q_{21}^2))\frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}}\\
&\qquad+4\Sigma_t^{12}(Q_{11}Q_{12}+Q_{21}Q_{22})\frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{22}}+4(\Sigma_t^{12}(Q_{12}^2+Q_{22}^2)\\
&\qquad+\Sigma_t^{22}(Q_{11}Q_{12}+Q_{21}Q_{22}))\times\frac{\partial^2}{\partial\Sigma^{12}\partial\Sigma^{22}},
\end{aligned}$$

thus proving the equality in (6), since

$$\begin{aligned}\mathcal{L}_\Sigma &= \mathrm{Tr}[(\Omega\Omega^{\mathrm{T}} + M\Sigma + \Sigma M^{\mathrm{T}})D] \\ &\quad + \frac{1}{2}\left\{\langle\Sigma^{11},\Sigma^{11}\rangle_t \frac{\partial^2}{(\partial\Sigma^{11})^2}\right. \\ &\quad + 4\langle\Sigma^{12},\Sigma^{12}\rangle_t \frac{\partial^2}{(\partial\Sigma^{12})^2} + \langle\Sigma^{22},\Sigma^{22}\rangle_t \frac{\partial^2}{(\partial\Sigma^{22})^2}. \\ &\quad + 4\langle\Sigma^{11},\Sigma^{12}\rangle_t \frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}} \\ &\quad + 2\langle\Sigma^{11},\Sigma^{22}\rangle_t \frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{22}} \\ &\quad \left. + 4\langle\Sigma^{12},\Sigma^{22}\rangle_t \frac{\partial^2}{\partial\Sigma^{12}\partial\Sigma^{22}}\right\},\end{aligned}$$

where we recall that

$$\begin{aligned}2\langle\Sigma^{12},\Sigma^{12}\rangle_t \frac{\partial^2}{(\partial\Sigma^{12})^2} &= \langle\Sigma^{12},\Sigma^{12}\rangle_t \frac{\partial^2}{(\partial\Sigma^{12})^2} \\ &\quad + \langle\Sigma^{21},\Sigma^{21}\rangle_t \frac{\partial^2}{(\partial\Sigma^{21})^2}, \\ 4\langle\Sigma^{11},\Sigma^{12}\rangle_t \frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}} &= 2\langle\Sigma^{11},\Sigma^{12}\rangle_t \frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}} \\ &\quad + 2\langle\Sigma^{11},\Sigma^{21}\rangle_t \frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{21}}.\end{aligned}$$

Appendix C: The affinity constraints on the correlation structure

In this appendix we study the general correlation structure in the case $n=2$. We introduce four matrices $R11$, $R12$, $R21$, $R22 \in M_2$ representing the correlations among the matrix Brownian motions (in total, $16 = n^2 \times n^2$ correlations: Rab_{ij} denotes the correlation between Z_t^{ab} and W_t^{ij}). In this way we can write

$$Z_t^{11} = \mathrm{Tr}[W_t R11^{\mathrm{T}}] + \sqrt{1 - \mathrm{Tr}[R11 R11^{\mathrm{T}}]}B_t^{11}, \quad \text{(C1)}$$

$$Z_t^{12} = \mathrm{Tr}[W_t R12^{\mathrm{T}}] + \sqrt{1 - \mathrm{Tr}[R12 R12^{\mathrm{T}}]}B_t^{12}, \quad \text{(C2)}$$

$$Z_t^{21} = \mathrm{Tr}[W_t R21^{\mathrm{T}}] + \sqrt{1 - \mathrm{Tr}[R21 R21^{\mathrm{T}}]}B_t^{21}, \quad \text{(C3)}$$

$$Z_t^{22} = \mathrm{Tr}[W_t R22^{\mathrm{T}}] + \sqrt{1 - \mathrm{Tr}[R22 R22^{\mathrm{T}}]}B_t^{22}. \quad \text{(C4)}$$

First we note that there are some constraints on the parameters in order to grant that Z_t is indeed a matrix Brownian motion.

Proposition C.1: *Z_t is a matrix Brownian motion iff*

$$\begin{aligned}&\mathrm{Tr}[Rij Rlm^T] = 0, \quad \textit{for } (i,j) \neq (l,m), \\ &i,j,l,m \in \{1,2\}.\end{aligned} \quad \text{(C5)}$$

Proof: Let us consider the first element of the matrix $\mathrm{Cov}_t(\mathrm{d}Z_t\alpha, \mathrm{d}Z_t\beta)$:

$$\begin{aligned}&\mathrm{Cov}_t(\mathrm{d}Z_t\alpha,\mathrm{d}Z_t\beta)_{11} \\ &\quad = \mathbb{E}_t(\mathrm{Tr}[\mathrm{d}W_t R11^{\mathrm{T}}]\alpha_1 + \sqrt{1-\mathrm{Tr}[R11R11^{\mathrm{T}}]}\,\mathrm{d}B_t^{11}\alpha_1 \\ &\qquad + \mathrm{Tr}[\mathrm{d}W_t R12^{\mathrm{T}}]\alpha_2 + \sqrt{1-\mathrm{Tr}[R12R12^{\mathrm{T}}]}\mathrm{d}B_t^{12}\alpha_2) \\ &\qquad \times (\mathrm{Tr}[\mathrm{d}W_t R11^{\mathrm{T}}]\beta_1 + \sqrt{1-\mathrm{Tr}[R11R11^{\mathrm{T}}]}\mathrm{d}B_t^{11}\beta_1 \\ &\qquad + \mathrm{Tr}[\mathrm{d}W_t R12^{\mathrm{T}}]\beta_2 + \sqrt{1-\mathrm{Tr}[R12R12^{\mathrm{T}}]}\mathrm{d}B_t^{12}\beta_2) \\ &\quad = \alpha_1\beta_1\mathrm{d}t + \alpha_2\beta_2\mathrm{d}t \\ &\qquad + (\alpha_1\beta_2 + \alpha_2\beta_1)(R11_{11}R12_{11} \\ &\qquad + R11_{12}R12_{12} + R11_{21}R12_{21} + R11_{22}R12_{22})\,\mathrm{d}t.\end{aligned}$$

Since we have to prove that $\mathrm{Cov}_t(\mathrm{d}Z_t\alpha, \mathrm{d}Z_t\beta) = \alpha^T\beta\mathbb{I},\mathrm{d}t$ for all vectors α, β, it must be that

$$\begin{aligned}&R11_{11}R12_{11} + R11_{12}R12_{12} \\ &\quad + R11_{21}R12_{21} + R11_{22}R12_{22} = 0,\end{aligned}$$

that is $\mathrm{Tr}[R11R12^{\mathrm{T}}] = 0$. Similar computations for the other components complete the proof. □

Now we look for the additional constraints on the matrices Rij in order to grant the affinity of the model, that is such that $\mathcal{L}_{Y,\Sigma}$ is affine on the elements of Σ_t. Let us consider the first element:

$$\begin{aligned}&\mathrm{d}\langle\Sigma^{11}, Y\rangle_t \\ &\quad = \mathbb{E}_t[(\sigma_t^{11}\mathrm{d}Z_t^{11} + \sigma_t^{12}\mathrm{d}Z_t^{12} + \sigma_t^{12}\mathrm{d}Z_t^{21} + \sigma_t^{22}\mathrm{d}Z_t^{22})\mathrm{d}\Sigma_t^{11}] \\ &\quad = 2((\sigma_t^{11})^2 Q_{11}R11_{11} + (\sigma_t^{11})^2 Q_{21}R11_{12} + \sigma_t^{11}\sigma_t^{12}Q_{11}R11_{21} \\ &\qquad + \sigma_t^{11}\sigma_t^{12}Q_{21}R11_{22} + \sigma_t^{11}\sigma_t^{12}Q_{11}R12_{11} + \sigma_t^{11}\sigma_t^{12}Q_{21}R12_{12} \\ &\qquad + (\sigma_t^{12})^2 Q_{11}R12_{21} + (\sigma_t^{12})^2 Q_{21}R12_{22} + \sigma_t^{11}\sigma_t^{12}Q_{11}R21_{11} \\ &\qquad + \sigma_t^{11}\sigma_t^{12}Q_{21}R21_{12} + (\sigma_t^{12})^2 Q_{11}R21_{21} + (\sigma_t^{12})^2 Q_{21}R21_{22} \\ &\qquad + \sigma_t^{11}\sigma_t^{22}Q_{11}R22_{11} + \sigma_t^{11}\sigma_t^{22}Q_{21}R22_{12} + \sigma_t^{12}\sigma_t^{22}Q_{11}R22_{21} \\ &\qquad + \sigma_t^{12}\sigma_t^{22}Q_{21}R22_{22})\mathrm{d}t.\end{aligned}$$

It follows that

$$\begin{aligned}R22_{11} &= 0, \\ R22_{12} &= 0, \\ R11_{11} &= R12_{21} + R21_{21}, \\ R11_{12} &= R12_{22} + R21_{22}, \\ R22_{21} &= R11_{21} + R12_{11} + R21_{11}, \\ R22_{22} &= R11_{22} + R12_{12} + R21_{12}.\end{aligned}$$

From the expression $\mathrm{d}\langle\Sigma^{22}, Y\rangle_t$ we obtain

$$\begin{aligned}R11_{21} &= 0, \\ R11_{22} &= 0,\end{aligned}$$

and it turns out that the other conditions are redundant, as well as those coming from $\mathrm{d}\langle\Sigma^{12}, Y\rangle_t$. In conclusion, the affinity constraints lead to the following specification for the four correlation matrices:

$$R12 = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$
$$R21 = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$
$$R11 = \begin{pmatrix} c+g & d+h \\ 0 & 0 \end{pmatrix},$$
$$R22 = \begin{pmatrix} 0 & 0 \\ a+e & b+f \end{pmatrix}.$$

Now we impose (C5) and obtain

$$R11 \perp R21 \longrightarrow e(c+g) + f(d+h) = 0, \tag{C6}$$

$$R11 \perp R12 \longrightarrow a(c+g) + b(d+h) = 0, \tag{C7}$$

$$R22 \perp R21 \longrightarrow g(a+e) + h(b+f) = 0, \tag{C8}$$

$$R22 \perp R12 \longrightarrow c(a+e) + d(b+f) = 0, \tag{C9}$$

$$R12 \perp R21 \longrightarrow ae + bf + cg + dh = 0. \tag{C10}$$

After some manipulations we arrive at

$$\frac{ae}{(a+e)^2} + \frac{cg}{(b+f)^2} = 0. \tag{C11}$$

Here we see that there are eight parameters but subject to five (nonlinear) constraints, allowing only a few compatible choices for the parameters. We are now ready to write down the infinitesimal generator associated with the general (affine) two-dimensional case.

Proposition C.2: *The infinitesimal generator of* (Y_t, Σ_t) *is given by*

$$\begin{aligned}\mathcal{L}_{Y,\Sigma} &= \left(r - \frac{1}{2}\operatorname{Tr}[\Sigma]\right)\frac{\partial}{\partial y} + \frac{1}{2}\operatorname{Tr}[\Sigma]\frac{\partial^2}{\partial y^2} \\ &\quad + \operatorname{Tr}[(\Omega\Omega^T + M\Sigma + \Sigma M^T)D \\ &\quad + 2\Sigma D Q^T Q D] + 2\operatorname{Tr}[\Sigma(R11 + R22)QD]\frac{\partial}{\partial y}.\end{aligned} \tag{C12}$$

Proof: We focus on the covariation terms $\mathrm{d}\langle \Sigma^{ij}, Y\rangle_t$, for $i,j = 1, \ldots, 2$:

$$\begin{aligned}\mathrm{d}\langle \Sigma^{11}, Y\rangle_t &= 2Q_{11}((c+g)\Sigma^{11} + (a+e)\Sigma^{12})\mathrm{d}t \\ &\quad + 2Q_{21}((\mathrm{d}+h)\Sigma^{11} + (b+f)\Sigma^{12})\mathrm{d}t, \\ \mathrm{d}\langle \Sigma^{22}, Y\rangle_t &= 2Q_{12}((a+e)\Sigma^{22} + (c+g)\Sigma^{12})\mathrm{d}t \\ &\quad + 2Q_{22}((\mathrm{d}+h)\Sigma^{12} + (b+f)\Sigma^{22})\,\mathrm{d}t, \\ \mathrm{d}\langle \Sigma^{12}, Y\rangle_t &= Q_{12}((c+g)\Sigma^{11} + (a+e)\Sigma^{12})\,\mathrm{d}t \\ &\quad + Q_{22}((\mathrm{d}+h)\Sigma^{11} + (b+f)\Sigma^{12})\,\mathrm{d}t \\ &\quad + Q_{11}((c+g)\Sigma^{12} + (a+e)\Sigma^{22})\mathrm{d}t \\ &\quad + Q_{21}((\mathrm{d}+h)\Sigma^{12} + (b+f)\Sigma^{22})\,\mathrm{d}t,\end{aligned}$$

and we obtain the proof, since $\mathrm{d}\langle \Sigma^{ij}, Y\rangle_t$ corresponds to the coefficient of the term $(\partial^2/\partial\, x_{ij}\,\partial y)$, and

$$\begin{aligned}&\operatorname{Tr}[\Sigma(R11 + R22)QD]\frac{\partial}{\partial y} \\ &= \operatorname{Tr}\left[\begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix}\begin{pmatrix} c+g & d+h \\ a+e & b+f \end{pmatrix}\begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \right. \\ &\qquad \left. \times \begin{pmatrix} \frac{\partial}{\partial\Sigma^{11}} & \frac{\partial}{\partial\Sigma^{12}} \\ \frac{\partial}{\partial\Sigma^{12}} & \frac{\partial}{\partial\Sigma^{22}} \end{pmatrix}\right]\frac{\partial}{\partial y},\end{aligned}$$

and by definition D is symmetric. □

By applying the Feynman–Kac argument to the Laplace transform

$$\begin{aligned}\Psi_{\gamma,t}(\tau) &= \mathbb{E}_t \exp\{\gamma Y_{t+\tau}\} & \text{(C13)} \\ &= \exp\{\operatorname{Tr}[A(\tau)\Sigma_t] + b(\tau)Y_t + c(\tau)\}, & \text{(C14)}\end{aligned}$$

we obtain $b(\tau) \equiv \gamma$ and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}A(\tau) &= A(\tau)M + (M^{\mathrm{T}} + 2\gamma(R11 + R22)Q)A(\tau) \\ &\quad + 2A(\tau)Q^{\mathrm{T}}QA(\tau) \\ &\quad + \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n, \\ A(0) &= 0.\end{aligned} \tag{C15}$$

We have proved the following.

Proposition C.3: *The Riccati equations satisfied by the matrix coefficient* $A(\tau)$ *associated with the Laplace transform* (*C14*) *are given by* (*C15*), *where*

$$R11 = \begin{pmatrix} c+g & d+h \\ 0 & 0 \end{pmatrix},$$
$$R22 = \begin{pmatrix} 0 & 0 \\ a+e & b+f \end{pmatrix},$$

and the parameters a, b, c, d, e, f, g, h *satisfy the (nonlinear) constraints* (*C6*), (*C7*), (*C8*), (*C9*), (*C10*) *and* (*C11*).

Remark C.1: Our model corresponds to choosing $c = d = e = f = 0$ (or, equivalently, $a = b = g = h = 0$): we obtain

$$R12 = \begin{pmatrix} \rho_{21} & \rho_{22} \\ 0 & 0 \end{pmatrix},$$
$$R21 = \begin{pmatrix} 0 & 0 \\ 2\rho_{11} & \rho_{12} \end{pmatrix},$$
$$R11 = \begin{pmatrix} \rho_{11} & \rho_{12} \\ 0 & 0 \end{pmatrix},$$
$$R22 = \begin{pmatrix} 0 & 0 \\ \rho_{21} & \rho_{22} \end{pmatrix},$$

and

$$Z_t^{11} = W_t^{11}\rho_{11} + W_t^{12}\rho_{12} + \sqrt{1-\rho_{11}^2-\rho_{12}^2}B_t^{11}, \quad \text{(C16)}$$

$$Z_t^{12} = W_t^{11}\rho_{21} + W_t^{12}\rho_{22} + \sqrt{1-\rho_{21}^2-\rho_{22}^2}B_t^{12}, \quad \text{(C17)}$$

$$Z_t^{21} = W_t^{21}\rho_{11} + W_t^{22}\rho_{12} + \sqrt{1-\rho_{11}^2-\rho_{12}^2}B_t^{21}, \quad \text{(C18)}$$

$$Z_t^{22} = W_t^{21}\rho_{21} + W_t^{22}\rho_{22} + \sqrt{1-\rho_{21}^2-\rho_{22}^2}B_t^{22}. \quad \text{(C19)}$$

We can then introduce the matrix

$$R = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix}$$

in such a way that $Z_t := W_t R^{\mathrm{T}} + \widetilde{B}_t\sqrt{\mathbb{I} - RR^{\mathrm{T}}}$, where $\widetilde{B}_t$ is a matrix Brownian motion which can be deduced from B_t.