

## A Multifactor Volatility Heston Model

Grasselli, Martino; Da Fonseca, Jose; Tebaldi, Claudio

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

[www.peerproject.eu](http://www.peerproject.eu)

### Empfohlene Zitierung / Suggested Citation:

Grasselli, M., Da Fonseca, J., & Tebaldi, C. (2008). A Multifactor Volatility Heston Model. *Quantitative Finance*, 8(6), 591-604. <https://doi.org/10.1080/14697680701668418>

### Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**gesis**  
Leibniz-Institut  
für Sozialwissenschaften

### Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der  
  
Leibniz-Gemeinschaft

## A Multifactor Volatility Heston Model

Journal:	<i>Quantitative Finance</i>
Manuscript ID:	RQUF-2006-0066.R2
Manuscript Category:	Research Paper
Date Submitted by the Author:	28-Jul-2007
Complete List of Authors:	grasselli, martino; Università di Padova, Matematica Pura ed Applicata Da Fonseca, Jose tebaldi, claudio; Università di Verona, Scienze Economiche
Keywords:	Stochastic Volatility, Options Pricing, Options Volatility, Financial Derivatives
JEL Code:	G12 - Asset Pricing < G1 - General Financial Markets < G - Financial Economics, G13 - Contingent Pricing Futures Pricing < G1 - General Financial Markets < G - Financial Economics
<p>Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.</p> <p>QFrevisionfinal.tex figures.zip</p>	

# A Multifactor Volatility Heston Model

José Da Fonseca, ESILV \*and Zeliade Systems †  
Martino Grasselli, Università degli Studi di Padova ‡and ESILV  
Claudio Tebaldi, Università degli Studi di Verona §

July 28, 2007

## Abstract

We consider a model for a single risky asset whose volatility follows a multifactor (matrix) Wishart affine process, recently introduced in finance by Gourieroux and Sufana (2004). As in standard Duffie and Kan (1996) affine models the pricing problem can be solved through the Fast Fourier Transform of Carr and Madan (1999). 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 a stochastic leverage effect, a well known stylized fact of FX option markets analyzed in Carr and Wu (2004).

**Keywords:** Wishart processes, Stochastic volatility, Matrix Riccati ODE, FFT. **JEL:** G12, G13

## 1 Introduction

An accurate volatility modelling is a crucial step in order to implement 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, like for example cliquet options. These

\*Ecole Supérieure d'Ingénieurs Léonard de Vinci, Département Mathématiques et Ingénierie Financière, 92916 Paris La Défense, France. Email: jose.da\_fonseca@devinci.fr

†Zeliade Systems, 56, Rue Jean-Jacques Rousseau 75001 Paris

‡Dipartimento di Matematica Pura ed Applicata, Via Belzoni 7, Padova, Italy. E-mail: grassell@math.unipd.it

§Department of Economics, Via Giardino Giusti 2, 37129 Verona, Italy. E-mail: claudio.tebaldi@univr.it. Acknowledgements: we thank Fabio Mercurio as well as the participants of the 2006 Ascona meeting and the Séminaire Bachelier of Paris for helpful comments. We also thank two anonymous referees for the careful reading of the paper. An earlier draft of this paper appeared under the title "Wishart Multi-Dimensional Stochastic Volatility".

1  
2  
3  
4  
5  
6  
7  
8 instruments, also called ratchet options, periodically "lock in" profits in a man-  
9 ner somewhat analogous to a mechanical ratchet. Exotic contracts like cliquet  
10 options, require an accurate modeling of the true realized variance process. In  
11 fact a cliquet option can be seen as a series of consecutive forward start op-  
12 tions whose prices depend only on realized volatility (see e.g. Hipp 1996). As  
13 well explained in Bergomi (2004), there is a structural limitation which pre-  
14 vents one-factor stochastic volatility models to price consistently these types  
15 of options jointly with plain vanilla options. A possible reconciliation requires  
16 that the volatility process is driven by at least 2 factors, even in a single as-  
17 set framework, as supported by empirical tests like the principal component  
18 analysis investigated in Cont and Fonseca (2002).

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

24 Within the Heston model an accurate modeling of the smile-skew effect for  
25 the implied volatility surface is usually obtained assuming a (negative) corre-  
26 lation between the noise driving the stock return and a suitable calibration of  
27 the parameters driving the volatility. It is indeed a common observation that a  
28 single factor diffusive model is not flexible enough to take into account the risk  
29 component introduced by the variability of the skew, also known as *stochastic*  
30 *skew* (see e.g. Carr and Wu 2004). In the case of FX options this risk factor is  
31 directly priced in the quotes of "risk reversals" strategies.

32 The aim of this paper is to extend the Heston model to a multifactor spec-  
33 ification for the volatility process in a single asset framework. While standard  
34 multifactor modeling of stochastic volatility is based on the class of affine term  
35 structure models introduced in Duffie and Kan (1996) and classified in Dai and  
36 Singleton (2000), in our model the factor process driving volatility is based  
37 on the matrix Wishart process, mathematically developed in Bru (1991). Our  
38 model takes inspiration from the multi-asset market model analyzed in Gourier-  
39oux and Sufana (2004): in their model the Wishart process describes the dy-  
40 namics of the covariance matrix and is assumed to be independent of the assets  
41 noises. On the contrary, we show that a symmetric matrix specification is poten-  
42 tially very useful to improve the affine factor modeling of the implied volatility  
43 curve. In fact, the introduction of the matrix notation provides a simple and  
44 powerful parametrization of the dependence between the asset noise and each  
45 volatility factor. In particular, using a square  $2 \times 2$  matrix of factors, we show  
46 that the expression of the return-volatility covariance is linear in the off diagonal  
47 factor, which can be directly identified as the "stochastic skew risk factor": in  
48 fact, such factor can be specifically used to generate a *stochastic leverage ef-*  
49 *fect*, which in the case of FX option can be directly calibrated on Risk Reversal  
50 quotes.  
51

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

- 53  
54 i) the term structure of the realized volatilities is described by a (matrix) mul-  
55

tifactor model;

- ii) a stochastic leverage effect appears and can be used to describe stochastic skew effects as required in FX option markets (see 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 which shows the advantages carried by the Wishart specification with respect to the single factor Heston one as well as the  $A_2(3)$  (in the terminology of Dai and Singleton 2000) multi-Heston model. In Section 5 we provide some conclusions and future developments. We gather in Appendix A some technical proofs, while in Appendix B we develop the computations in the 2-dimensional case for the reader's convenience. Finally, Appendix C discusses the general affine correlation structure in the 2-dimensional case.

## 2 The Wishart volatility process

In an arbitrage-free frictionless financial market we consider a risky asset whose price follows:

$$\frac{dS_t}{S_t} = rdt + Tr \left[ \sqrt{\Sigma_t} dZ_t \right], \quad (1)$$

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 by  $n^2$  independent Brownian motions) under the risk-neutral measure and  $\Sigma_t$  belongs to the set of symmetric  $n \times n$  positive-definite matrices (as well as its square root  $\sqrt{\Sigma_t}$ ).

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

$$d\Sigma_t = (\Omega\Omega^T + M\Sigma_t + \Sigma_t M^T) dt + \sqrt{\Sigma_t} dW_t Q + Q^T (dW_t)^T \sqrt{\Sigma_t}, \quad (2)$$

with  $\Omega, M, Q \in M_n$ ,  $\Omega$  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  $\Omega$  satisfies

$$\Omega\Omega^T = \beta Q^T Q$$

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

$$dS_t = \text{diag}[S_t] \left( r\mathbf{1}dt + \sqrt{\Sigma_t}dZ_t \right),$$

where  $S_t, Z_t \in \mathbb{R}^n$ ,  $\mathbf{1} = (1, \dots, 1)^T$  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) which completely describes the correlation structure, in such a way that  $Z_t$  can be written as  $Z_t := W_t R^T + B_t \sqrt{\mathbb{I} - RR^T}$ , ( $\mathbb{I}$  represents the identity matrix and  $T$  denotes transposition) where  $B_t$  is a (matrix) Brownian motion independent of  $W_t$ .

**Proposition 1** *The process  $Z_t := W_t R^T + B_t \sqrt{\mathbb{I} - RR^T}$  is a matrix Brownian motion.*

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

$$\begin{aligned} \text{Cov}_t(dZ_t\alpha, dZ_t\beta) &= \mathbb{E}_t \left[ (dZ_t\alpha)(dZ_t\beta)^T \right] \\ &= \alpha^T \beta \mathbb{I} dt. \end{aligned}$$

Here

$$\begin{aligned} \text{Cov}_t(dZ_t\alpha, dZ_t\beta) &= \mathbb{E}_t \left[ \left( dW_t R^T \alpha + dB_t \sqrt{\mathbb{I} - RR^T} \alpha \right) \left( dW_t R^T \beta + dB_t \sqrt{\mathbb{I} - RR^T} \beta \right)^T \right] \\ &= \text{Cov}_t(dW_t R^T \alpha, dW_t R^T \beta) + \text{Cov}_t \left( dB_t \sqrt{\mathbb{I} - RR^T} \alpha, dB_t \sqrt{\mathbb{I} - RR^T} \beta \right) \\ &= \alpha^T RR^T \beta \mathbb{I} dt + \alpha^T (\mathbb{I} - RR^T) \beta \mathbb{I} dt \\ &= \alpha^T \beta \mathbb{I} dt. \end{aligned}$$

■ In line of 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 grant analytical tractability of the model (in particular in order to preserve the affinity) some constraints should be imposed on the correlation structure. It turns out that such (non linear) constraints are quite binding: in order to give an idea we classify in the Appendix all the possibilities in 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 (1-dimensional) Heston model. In fact, it is well known that in order to solve the pricing problem of plain vanilla options, it is enough to compute the conditional characteristic function (under the risk-neutral measure) of the underlying (see e.g. Duffie, Pan and Singleton 2000) or equivalently of the return process  $Y_t = \ln S_t$ , which satisfies the following SDE:

$$dY_t = \left( r - \frac{1}{2} \text{Tr} [\Sigma_t] \right) dt + \text{Tr} \left[ \sqrt{\Sigma_t} \left( dW_t R^T + dB_t \sqrt{\mathbb{I} - RR^T} \right) \right]. \quad (3)$$

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, Pan and Singleton (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 guess 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 terms, 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:

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

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

$$\frac{\partial \Psi_{\gamma,t}}{\partial \tau} = \mathcal{L}_{Y,\Sigma} \Psi_{\gamma,t} \quad (5)$$

$$\Psi_{\gamma,t}(0) = \exp \{ \gamma Y_t \},$$

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

$$\mathcal{L}_{\Sigma} = Tr [(\Omega \Omega^T + M \Sigma + \Sigma M^T) D + 2 \Sigma D Q^T Q D], \quad (6)$$

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 2-dimensional case in Appendix B. Endowed with the previous result, we can now find the infinitesimal generator of the couple  $(Y_t, \Sigma_t)$ :

**Proposition 2** *The infinitesimal generator of  $(Y_t, \Sigma_t)$  is given by*

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

**Proof:** See Appendix A.

Thus the explicit expression of (5) is:

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

and by replacing the candidate (4) we obtain

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

with boundary conditions

$$\begin{aligned} A(0) &= 0 \in M_n, \\ b(0) &= \gamma \in \mathbb{R}, \\ c(0) &= 0. \end{aligned}$$



By identifying the coefficients of  $Y$  we deduce

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

hence

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

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

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

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

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

### 3.2 Matrix Riccati linearization

Matrix Riccati Equations like (9) have several nice properties (see e.g. Freiling 2002): the most remarkable one is that their flow can be linearized by doubling the dimension of the problem, this due to the fact that Riccati ODE belong to a quotient manifold (see Grasselli and Tebaldi 2004 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

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

and

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

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

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

which can be written as follows:

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

Its solution is simply obtained through exponentiation:

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

where

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

In conclusion, we get

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

and since  $A(0) = 0$ ,

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

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

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

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

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

Now we can integrate the last equation and obtain

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

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

**Remark 3** *The computation of the Laplace Transform for both asset returns and variance factors*

$$\begin{aligned} \Psi_{\gamma,t}(\tau) &= \mathbb{E}_t \exp \{ \gamma Y_{t+\tau} + \text{Tr} [\Gamma \Sigma_t] \} \\ &= \exp \left\{ \text{Tr} \left[ \tilde{A}(\tau) \Sigma_t \right] + \tilde{b}(\tau) Y_t + \tilde{c}(\tau) \right\}, \end{aligned} \quad (15)$$

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 come back 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

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

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

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

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

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

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  $\alpha$  is chosen (typically  $\alpha = 1.1$ , see Carr and Madan 1999) and the characteristic function  $\Phi$  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 find out 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 to be pure variance contracts. The first step consists in considering a Forward-Start call option, whose payoff at the maturity  $T$  is defined as follows:

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

where  $S_t$  is the stock price at a fixed date  $t$ ,  $0 \leq t \leq T$ . In the following, we follow the (single volatility factor) presentation of Wong (2004). By risk-neutral

valuation, the initial price of this option is given by

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

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

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

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  $t$ )  $S_t = 1$ : notice 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-starts call option is given by

$$FSCall(0) = \mathbb{E} \left[ \exp \{-rT\} (\exp \{Y_{t,T}\} - \exp \{k\})^+ \right],$$

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 by

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

The modified option price is given by

$$c_{t,T}(k) = \exp \{\alpha k\} FSCall(0)$$

and its Fourier transform

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

therefore here again we realize that in order to price a forward-starts 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 4** *Given a real symmetric matrix  $D$ , the conditional characteristic function of the Wishart process  $\Sigma_t$  is given by:*

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

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

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

with

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

**Proof:** See Appendix A.

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

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

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

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

Endowed with the function  $\Phi_{\gamma,0}(t, T)$ , it suffices to plug into (17) and apply the FFT in order to find 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 one. We quote option prices using Black&Scholes volatility, which is a common practice in the market. Let us denote by  $V_t$  the instantaneous volatility in the (single factor) Heston model, whose dynamics is given by

$$dV_t = \kappa(\theta - V_t)dt + \epsilon\sqrt{V_t}dW_t^2,$$

where  $\theta$  represents the long-term volatility,  $\kappa$  is the mean reversion parameter,  $\epsilon$  is the volatility of volatility parameter (also called vol-of-vol),  $\rho$  is the correlation between the volatility and the stock,  $V_0$  is the initial spot volatility and  $W_t^2$  is (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 to reproduce a volatility surface which 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 2 dimensional affine process whose state space domain is  $\mathbb{R}_+^2$  as classified in 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 be easily nested into 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  $Tr(\Sigma_t)$  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 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 to the element  $\Sigma_{11}$  the meaning of a short-term factor, while  $\Sigma_{22}$  has an impact on the long-term volatility. Let us take the following values:

$$M = \begin{pmatrix} -3 & 0 \\ 0 & -0.333 \end{pmatrix}, R = \begin{pmatrix} -0.7 & 0 \\ 0 & -0.7 \end{pmatrix} \quad (20)$$

$$Q = \begin{pmatrix} 0.25 & 0 \\ 0 & 0.25 \end{pmatrix}, \Sigma_0 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix},$$

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 on the market there are some long-term products such as forward start option and cliquet options whose maturity can be much higher 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 one. 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 of the long-term volatility induces also an increase of the 3 months volatility, so that the short-term fit for the implied volatility is unsatisfactory, as illustrated in Figure (1).

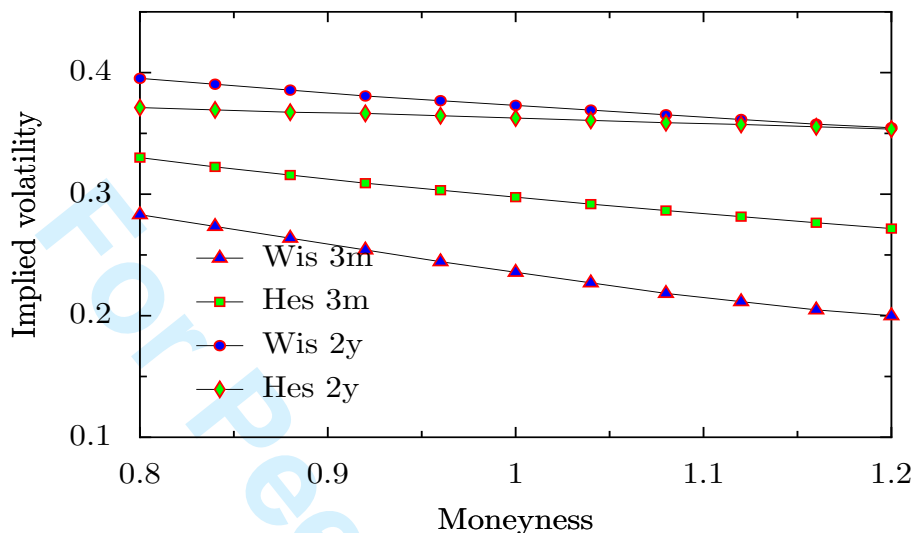


Figure 1: Implied volatility for the Wishart model (Wis) and Heston (Hes) model. Option maturities are 3 months (3m) and 2 years (2y). Moneyness is defined by  $\frac{K}{S_0}$  where  $S_0$  is the initial spot value.

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 to an unsatisfactory fit of the long-term implied volatility level as shown in Figure (2).

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

Notice that the above observation is not sufficient to justify the introduction of the previous Wishart (matrix) affine model given by (20), whose covariance matrix can be also reproduced<sup>1</sup> using the following (vector) affine model, which belongs to the canonical class  $\mathbb{A}_2(3)$  of Dai and Singleton (2000):

$$dX_t^1 = \kappa_1(\theta_1 - X_t^1)dt + \epsilon_1\sqrt{X_t^1}dW_t^1,$$

$$dX_t^2 = \kappa_2(\theta_2 - X_t^2)dt + \epsilon_2\sqrt{X_t^2}dW_t^2,$$

$$dY_t = \left(r - \frac{1}{2}(X_t^1 + X_t^2)\right)dt + \rho_1\sqrt{X_t^1}dW_t^1 + \rho_2\sqrt{X_t^2}dW_t^2 + \sqrt{(1 - \rho_1^2)X_t^1 + (1 - \rho_2^2)X_t^2}dB_t,$$

<sup>1</sup>We thank an anonymous referee for the observation

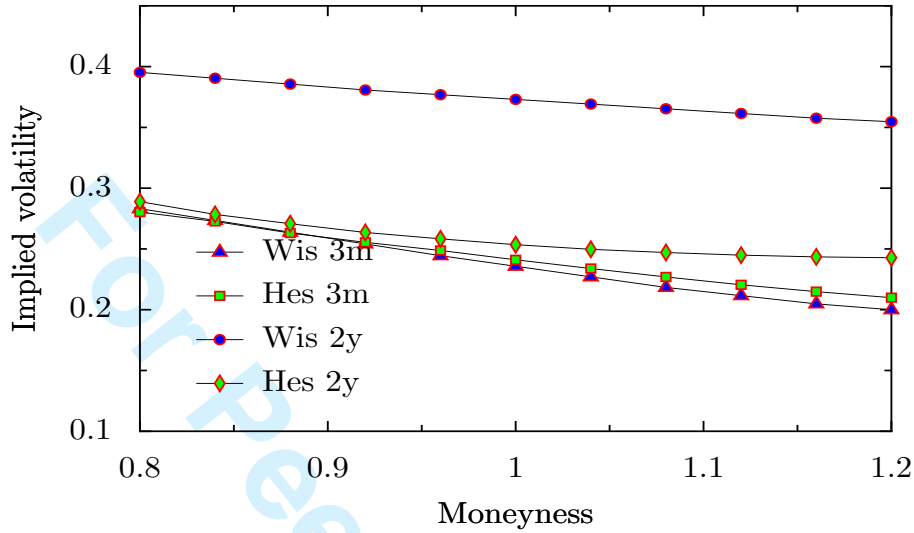


Figure 2: Implied volatility for the Wishart model (Wis) and Heston (Hes) model.

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

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:

$$\begin{aligned}
 \text{Corr}_t(\text{Noise}(dY), \text{Noise}(\text{Vol}(dY))) &= \frac{\langle Y, X^1 + X^2 \rangle_t}{\sqrt{\langle Y \rangle_t \langle X^1 + X^2 \rangle_t}} \\
 &= \frac{\mathbb{E}_t \left[ \left( \sqrt{X_t^1} dZ_t^1 + \sqrt{X_t^2} dZ_t^2 \right) \left( \epsilon_1 \sqrt{X_t^1} dW_t^1 + \epsilon_2 \sqrt{X_t^2} dW_t^2 \right) \right]}{\sqrt{X_t^1 + X_t^2} \sqrt{\epsilon_1^2 X_t^1 + \epsilon_2^2 X_t^2}} \\
 &= \frac{\rho_1 \epsilon_1 X_t^1 + \rho_2 \epsilon_2 X_t^2}{\sqrt{X_t^1 + X_t^2} \sqrt{\epsilon_1^2 X_t^1 + \epsilon_2^2 X_t^2}}. \quad (21)
 \end{aligned}$$

#### 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  $\text{Tr}(\Sigma_t)$ : this is computed in the following

**Proposition 5** *The stochastic correlation between the stock noise and the volatil-*



ity noise in the Wishart model is given by

$$\rho_t = \frac{\text{Tr}[RQ\Sigma_t]}{\sqrt{\text{Tr}[\Sigma_t]}\sqrt{\text{Tr}[Q^T Q\Sigma_t]}}. \quad (22)$$

**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  $\Sigma_t^{11}, \Sigma_t^{22}$  (while the off diagonal factor  $\Sigma_t^{12}$  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  $\rho_t$  depends also on the off diagonal volatility term  $\Sigma_t^{12}$ :

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

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 which can be calibrated on Risk Reversal quotes in the spirit of Carr and Wu (2004).

This model cannot be nested into 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 2004), which is non linear in the factors (the domain of  $\Sigma_t^{12}$  is given by the set  $\Sigma_t^{11}\Sigma_t^{22} - (\Sigma_t^{12})^2 > 0$ ). This non linearity allows the Wishart specification to reproduce new effects w.r.t. the classic (vector) affine models.

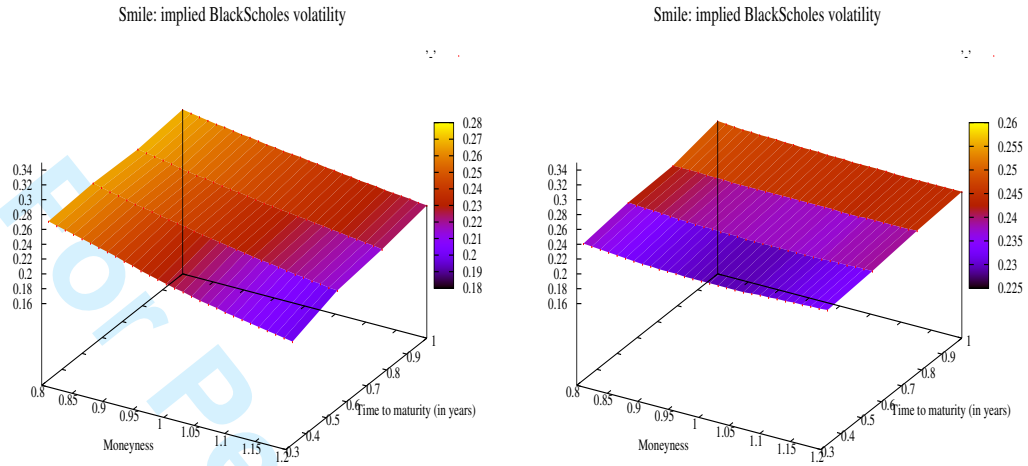


Figure 3: Wishart implied volatility for  $R_1$  (left) and  $R_2$  (right).

#### 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 a (negative) correlation between the volatility and the stock price. Taking the matrices

$$M = \begin{pmatrix} -5 & 0 \\ 0 & -3 \end{pmatrix}, Q = \begin{pmatrix} 0.35 & 0 \\ 0 & 0.25 \end{pmatrix}, \Sigma_0 = \begin{pmatrix} 0.02 & 0 \\ 0 & 0.02 \end{pmatrix} \quad (23)$$

$$R_1 = \begin{pmatrix} -0.7 & 0 \\ 0 & -0.5 \end{pmatrix}, R_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and  $\beta = 3$  in the Wishart model, we get for  $R_1$  (resp.  $R_2$ ) the left (resp. 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 fitted by using the diagonal terms in the matrix  $R_1$  as well as the parameters  $\rho_1, \rho_2$  in the  $\mathbb{A}_2(3)$  multi-Heston model.

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

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

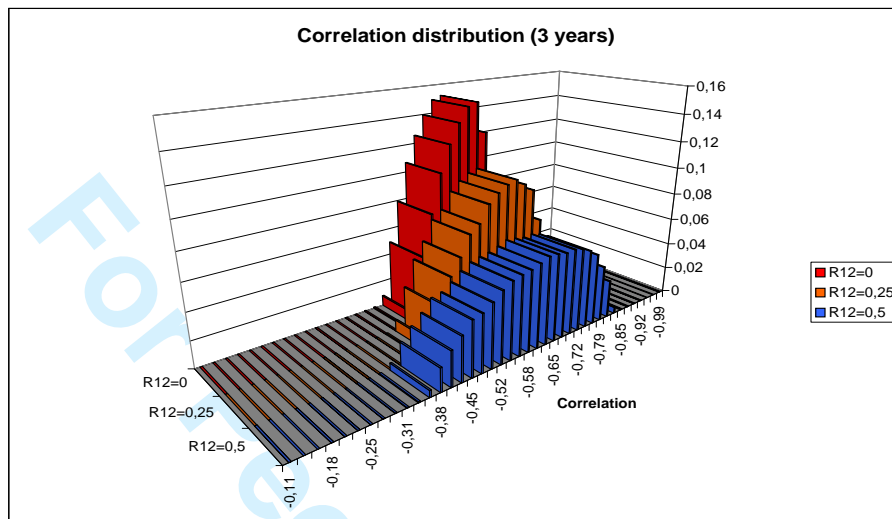


Figure 4: Distribution functions of the correlation process in the Wishart model with non diagonal matrix  $R$ .

From (22) we obtain:

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

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

## 5 Conclusion

We showed that the multifactor volatility extension of the Heston model considered in this paper is flexible enough to take into account correlations with the underlying asset returns. In the meanwhile it preserves analytical tractability, i.e. a closed form for the conditional characteristic function, and a linear factor structure which can be potentially very useful in the calibration procedure. Finally, our numerical results show that the flexibility induced by the additional factors allow a better fit of the smile-skew effect at both long and short maturities. In particular, contrarily to the Heston model, the Wishart specification does permit 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 on 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 to analyze and describe volatility and stochastic correlations' effects on the risk premia valued by the market.

## References

- [1] Bergomi, L. (2004) "Smile dynamics", Risk September, 117-123.
- [2] Bergomi, L. (2005) "Smile dynamics II", Risk October.
- [3] Bru, M. F. (1991) "Wishart Processes". *Journal of Theoretical Probability*, **4**, 725-743.
- [4] Cont, R. and J. da Fonseca (2002) "Dynamics of implied volatility surfaces". *Quantitative Finance*, **2**, No 1, 45-60.
- [5] Carr, P. and D. B. Madan (1999) "Option valuation using the fast Fourier transform". *Journal of Computational Finance*, **2** No 4.
- [6] Carr, P. and L. Wu.(2004) "Stochastic Skew in Currency Options", preprint.
- [7] Dai, Q. and K. Singleton (2000) "Specification Analysis of Affine Term Structure Models". *Journal of Finance*, **55**, 1943-1978.
- [8] Duffie, D. and R. Kan (1996) "A Yield-Factor Model of Interest Rates". *Mathematical Finance*, **6** (4), 379-406.
- [9] Duffie, D., J. Pan and K. Singleton (2000) "Transform analysis and asset pricing for affine jump-diffusions". *Econometrica*, **68**, 1343-1376.
- [10] Freiling, G. (2002): "A Survey of Nonsymmetric Riccati Equations". *Linear Algebra and Its Applications*, 243-270.
- [11] Gouriéroux, C. and R. Sufana (2003) "Wishart Quadratic Term Structure Models". CREF 03-10, HEC Montreal.
- [12] Gouriéroux, C. and R. Sufana (2004) "Derivative Pricing with Multivariate Stochastic Volatility: Application to Credit Risk". Working paper CREST.
- [13] Grasselli M. and C. Tebaldi (2004) "Solvable Affine Term Structure Models". *Mathematical Finance*, to appear.
- [14] Heston, S. L. (1993) "A Closed-Form Solution for Option with Stochastic Volatility with Applications to Bond and Currency Options ". *Review of Financial Studies*, **6**, (2), 327-343.

- [15] C. Hipp (1996) "Options for Guaranteed Index-linked Life Insurance". AFIR 1996 Proceedings, Vol. II, S. 1463-1483.
- [16] Wong, G. (2004) "Forward Smile and Derivative Pricing ". Working paper UBS.

## 6 Appendix A: Proofs

**Proof of Proposition 2:** The only non trivial term in (7) comes from the covariation

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

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  $\sigma_t$ . Now we identify the covariation terms with the coefficients of  $\left(\frac{\partial^2}{\partial x_{ij} \partial y}\right)$ , thus obtaining

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

which corresponds to the coefficient of the term  $\left(\frac{\partial^2}{\partial x_{ij} \partial y}\right)$ , since

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

and since by definition  $D$  is symmetric. ■

**Proof of Proposition 4:** 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

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

while

$$C(\tau) = Tr \left[ \Omega \Omega^T \int_0^\tau B(s) ds \right].$$

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

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

which gives the statement. ■

**Proof of Proposition 5:** The first step consists in finding the stock noise:

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + Tr \left[ \sqrt{\Sigma_t} dZ_t \right] \\ &= rdt + \sqrt{Tr[\Sigma_t]} \frac{Tr[\sqrt{\Sigma_t} dZ_t]}{\sqrt{Tr[\Sigma_t]}} \\ &= rdt + \sqrt{Tr[\Sigma_t]} dz_t, \end{aligned}$$

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

$$\begin{aligned} dTr[\Sigma_t] &= (Tr[\Omega \Omega^T] + 2Tr[M\Sigma_t]) dt + 2Tr[\sqrt{\Sigma_t} dW_t Q] \\ &= \dots dt + 2\sqrt{Tr[\Sigma_t Q^T Q]} \frac{Tr[\sqrt{\Sigma_t} dW_t Q]}{\sqrt{Tr[\Sigma_t Q^T Q]}} \\ &= \dots dt + 2\sqrt{Tr[\Sigma_t Q^T Q]} dw_t, \end{aligned}$$

where we used the fact that

$$\begin{aligned}
d \langle Tr [\Sigma.] \rangle_t &= \sum_{ij} Cov_t (e_i^T d\Sigma_t e_i, e_j^T d\Sigma_t e_j) \\
&= 4 \sum_{ij} Cov_t (e_i^T \sqrt{\Sigma_t} dW_t Q e_i, e_j^T \sqrt{\Sigma_t} dW_t Q e_j) \\
&= 4 \sum_{ij} \mathbb{E}_t [e_i^T \sqrt{\Sigma_t} dW_t Q e_i e_j^T Q^T dW_t^T \sqrt{\Sigma_t} e_j] \\
&= 4 \sum_{ij} e_i^T \sqrt{\Sigma_t} Tr [Q e_i e_j^T Q^T] \sqrt{\Sigma_t} e_j dt \\
&= 4 \sum_{ij} Tr [Q^T Q e_i e_j^T] e_i^T \Sigma_t e_j dt \\
&= 4 \sum_{ij} e_j^T Q^T Q e_i e_i^T \Sigma_t e_j dt \\
&= 4 \sum_j e_j^T Q^T Q \Sigma_t e_j dt \\
&= 4 Tr [\Sigma_t Q^T Q] dt,
\end{aligned}$$

where we used that

$$\mathbb{E}_t [dW_t Q e_i e_j^T Q^T dW_t^T] = Tr [Q e_i e_j^T Q^T] dt$$

since from Proposition 1:

$$\begin{aligned}
\mathbb{E}_t [dW_t \alpha \beta^T dW_t^T] &= \alpha^T \beta \mathbb{I} dt \\
&= Tr [\alpha \beta^T] \mathbb{I} dt.
\end{aligned}$$

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  $w_t$ , whose covariation is given by:

$$\begin{aligned}
Cov_t (dz_t, dw_t) &= Cov_t \left( \frac{Tr [\sqrt{\Sigma_t} dZ_t]}{\sqrt{Tr [\Sigma_t]}}, \frac{Tr [\sqrt{\Sigma_t} dW_t Q]}{\sqrt{Tr [\Sigma_t Q^T Q]}} \right) \\
&= \mathbb{E}_t \left[ \frac{Tr [\sqrt{\Sigma_t} dW_t R^T]}{\sqrt{Tr [\Sigma_t]}} \frac{Tr [\sqrt{\Sigma_t} dW_t Q]}{\sqrt{Tr [\Sigma_t Q^T Q]}} \right] \\
&= \frac{\sum_{ij} Cov_t (e_i^T \sqrt{\Sigma_t} dW_t R^T e_i, e_j^T \sqrt{\Sigma_t} dW_t Q e_j)}{\sqrt{Tr [\Sigma_t]} \sqrt{Tr [\Sigma_t Q^T Q]}} \\
&= \frac{1}{\sqrt{Tr [\Sigma_t]} \sqrt{Tr [\Sigma_t Q^T Q]}} \sum_{ij} e_i^T \sqrt{\Sigma_t} Tr [R^T e_i e_j^T Q^T] \sqrt{\Sigma_t} e_j dt \\
&= \frac{1}{\sqrt{Tr [\Sigma_t]} \sqrt{Tr [\Sigma_t Q^T Q]}} Tr [\Sigma_t Q^T R^T] dt \\
&= \frac{Tr [\Sigma_t R Q]}{\sqrt{Tr [\Sigma_t]} \sqrt{Tr [\Sigma_t Q^T Q]}} dt
\end{aligned}$$

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

■

## 7 Appendix B: The 2-dimensional case

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

$$\begin{aligned} d\Sigma_t &= d \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix} \\ &= \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} \Omega_{11} & \Omega_{21} \\ \Omega_{12} & \Omega_{22} \end{pmatrix} \\ &\quad + \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix} + \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix} \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix} dt \\ &\quad + \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix}^{1/2} \begin{pmatrix} dW_t^{11} & dW_t^{12} \\ dW_t^{21} & dW_t^{22} \end{pmatrix} \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \\ &\quad + \begin{pmatrix} Q_{11} & Q_{21} \\ Q_{12} & Q_{22} \end{pmatrix} \begin{pmatrix} dW_t^{11} & dW_t^{21} \\ dW_t^{12} & dW_t^{22} \end{pmatrix} \begin{pmatrix} \Sigma_t^{11} & \Sigma_t^{12} \\ \Sigma_t^{12} & \Sigma_t^{22} \end{pmatrix}^{1/2}. \end{aligned}$$

Let be

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

so that

$$\sigma_t^2 = \Sigma_t = \begin{pmatrix} (\sigma_t^{11})^2 + (\sigma_t^{12})^2 & \sigma_t^{11}\sigma_t^{12} + \sigma_t^{12}\sigma_t^{22} \\ \sigma_t^{11}\sigma_t^{12} + \sigma_t^{12}\sigma_t^{22} & (\sigma_t^{12})^2 + (\sigma_t^{22})^2 \end{pmatrix}. \quad (24)$$

We obtain

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



and using (24):

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

On the other hand, from (6) we can identify the coefficient of  $\left(\frac{\partial^2}{\partial \Sigma_{ij} \partial \Sigma_{ik}}\right)$  in the trace of the matrix  $2\Sigma_t DQ^T QD$ , that is

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

After some computations, we obtain:

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

thus proving the equality in (6), since

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

where we recall that

$$2 \langle \Sigma^{12}, \Sigma^{12} \rangle_t \frac{\partial^2}{(\partial \Sigma^{12})^2} = \langle \Sigma^{12}, \Sigma^{12} \rangle_t \frac{\partial^2}{(\partial \Sigma^{12})^2} + \langle \Sigma^{21}, \Sigma^{21} \rangle_t \frac{\partial^2}{(\partial \Sigma^{21})^2};$$

$$4 \langle \Sigma^{11}, \Sigma^{12} \rangle_t \frac{\partial^2}{\partial \Sigma^{11} \partial \Sigma^{12}} = 2 \langle \Sigma^{11}, \Sigma^{12} \rangle_t \frac{\partial^2}{\partial \Sigma^{11} \partial \Sigma^{12}} + 2 \langle \Sigma^{11}, \Sigma^{21} \rangle_t \frac{\partial^2}{\partial \Sigma^{11} \partial \Sigma^{21}}.$$

## 8 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 4 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

$$Z_t^{11} = Tr [W_t R11^T] + \sqrt{1 - Tr [R11 R11^T]} B_t^{11} \quad (25)$$

$$Z_t^{12} = Tr [W_t R12^T] + \sqrt{1 - Tr [R12 R12^T]} B_t^{12} \quad (26)$$

$$Z_t^{21} = Tr [W_t R21^T] + \sqrt{1 - Tr [R21 R21^T]} B_t^{21} \quad (27)$$

$$Z_t^{22} = Tr [W_t R22^T] + \sqrt{1 - Tr [R22 R22^T]} B_t^{22} \quad (28)$$

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

**Proposition 6**  $Z_t$  is a matrix Brownian motion iff

$$Tr [RijRlm^T] = 0 \text{ for } (i, j) \neq (l, m), \quad i, j, l, m \in \{1, 2\}. \quad (29)$$

**Proof:** Let us consider the first element of the matrix  $Cov_t(dZ_t \alpha, dZ_t \beta)$  :

$$\begin{aligned} Cov_t(dZ_t \alpha, dZ_t \beta)_{11} &= \mathbb{E}_t \left( Tr [dW_t R11^T] \alpha_1 + \sqrt{1 - Tr [R11 R11^T]} dB_t^{11} \alpha_1 \right. \\ &\quad \left. + Tr [dW_t R12^T] \alpha_2 + \sqrt{1 - Tr [R12 R12^T]} dB_t^{12} \alpha_2 \right) \\ &\quad \cdot \left( Tr [dW_t R11^T] \beta_1 + \sqrt{1 - Tr [R11 R11^T]} dB_t^{11} \beta_1 \right. \\ &\quad \left. + Tr [dW_t R12^T] \beta_2 + \sqrt{1 - Tr [R12 R12^T]} dB_t^{12} \beta_2 \right) \\ &= \alpha_1 \beta_1 dt + \alpha_2 \beta_2 dt \\ &\quad + (\alpha_1 \beta_2 + \alpha_2 \beta_1) (R11_{11} R12_{11} + R11_{12} R12_{12} + R11_{21} R12_{21} + R11_{22} R12_{22}) dt. \end{aligned}$$

Since we have to prove that  $Cov_t(dZ_t \alpha, dZ_t \beta) = \alpha^T \beta \mathbb{1} dt$  for all vectors  $\alpha, \beta$ , it must be that

$$R11_{11} R12_{11} + R11_{12} R12_{12} + R11_{21} R12_{21} + R11_{22} R12_{22} = 0,$$

that is  $Tr [R_{11}R_{12}^T] = 0$ . Similar computations for the other components lead to the conclusion. ■

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

$$\begin{aligned} d \langle \Sigma^{11}, Y \rangle_t &= \mathbb{E}_t \left[ (\sigma_t^{11} dZ_t^{11} + \sigma_t^{12} dZ_t^{12} + \sigma_t^{12} dZ_t^{21} + \sigma_t^{22} dZ_t^{22}) d\Sigma_t^{11} \right] \\ &= 2 \left( (\sigma_t^{11})^2 Q_{11} R_{11_{11}} + (\sigma_t^{11})^2 Q_{21} R_{11_{12}} + \sigma_t^{11} \sigma_t^{12} Q_{11} R_{11_{21}} \right. \\ &\quad + \sigma_t^{11} \sigma_t^{12} Q_{21} R_{11_{22}} + \sigma_t^{11} \sigma_t^{12} Q_{11} R_{12_{11}} + \sigma_t^{11} \sigma_t^{12} Q_{21} R_{12_{12}} \\ &\quad + (\sigma_t^{12})^2 Q_{11} R_{12_{21}} + (\sigma_t^{12})^2 Q_{21} R_{12_{22}} + \sigma_t^{11} \sigma_t^{12} Q_{11} R_{21_{11}} \\ &\quad + \sigma_t^{11} \sigma_t^{12} Q_{21} R_{21_{12}} + (\sigma_t^{12})^2 Q_{11} R_{21_{21}} + (\sigma_t^{12})^2 Q_{21} R_{21_{22}} \\ &\quad \left. + \sigma_t^{11} \sigma_t^{22} Q_{11} R_{22_{11}} + \sigma_t^{11} \sigma_t^{22} Q_{21} R_{22_{12}} + \sigma_t^{12} \sigma_t^{22} Q_{11} R_{22_{21}} + \sigma_t^{12} \sigma_t^{22} Q_{21} R_{22_{22}} \right) dt \end{aligned}$$

It follows that

$$\begin{aligned} R_{22_{11}} &= 0 \\ R_{22_{12}} &= 0 \\ R_{11_{11}} &= R_{12_{21}} + R_{21_{21}} \\ R_{11_{12}} &= R_{12_{22}} + R_{21_{22}} \\ R_{22_{21}} &= R_{11_{21}} + R_{12_{11}} + R_{21_{11}} \\ R_{22_{22}} &= R_{11_{22}} + R_{12_{12}} + R_{21_{12}} \end{aligned}$$

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

$$\begin{aligned} R_{11_{21}} &= 0 \\ R_{11_{22}} &= 0 \end{aligned}$$

and it turns out that the other conditions are redundant, as well as those coming from  $d \langle \Sigma^{12}, Y \rangle_t$ . In conclusion, the affinity constraints lead to the following specification for the 4 correlation matrices:

$$\begin{aligned} R_{12} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ R_{21} &= \begin{pmatrix} e & f \\ g & h \end{pmatrix} \\ R_{11} &= \begin{pmatrix} c+g & d+h \\ 0 & 0 \end{pmatrix} \\ R_{22} &= \begin{pmatrix} 0 & 0 \\ a+e & b+f \end{pmatrix}. \end{aligned}$$

Now we impose (29) and obtain:

$$R11 \perp R21 \longrightarrow e(c + g) + f(d + h) = 0 \quad (30)$$

$$R11 \perp R12 \longrightarrow a(c + g) + b(d + h) = 0 \quad (31)$$

$$R22 \perp R21 \longrightarrow g(a + e) + h(b + f) = 0 \quad (32)$$

$$R22 \perp R12 \longrightarrow c(a + e) + d(b + f) = 0 \quad (33)$$

$$R12 \perp R21 \longrightarrow ae + bf + cg + dh = 0. \quad (34)$$

After some manipulations we arrive to

$$\frac{ae}{(a + e)^2} + \frac{cg}{(b + f)^2} = 0. \quad (35)$$

Here we see that there are 8 parameters but subject to 5 (nonlinear) constraints, allowing only a few compatible choices for the parameters. Now we are ready to write down the infinitesimal generator associated to the general (affine) 2-dimensional case:

**Proposition 7** *The infinitesimal generator of  $(Y_t, \Sigma_t)$  is given by*

$$\begin{aligned} \mathcal{L}_{Y, \Sigma} = & \left( r - \frac{1}{2} Tr [\Sigma] \right) \frac{\partial}{\partial y} + \frac{1}{2} Tr [\Sigma] \frac{\partial^2}{\partial y^2} \\ & + Tr [(\Omega \Omega^T + M \Sigma + \Sigma M^T) D + 2 \Sigma D Q^T Q D] + 2 Tr [\Sigma (R11 + R22) Q D] \frac{\partial}{\partial y}. \end{aligned} \quad (36)$$

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

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

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

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

and by definition  $D$  is symmetric. ■

By applying the Feynman-Kac argument to the Laplace transform

$$\Psi_{\gamma,t}(\tau) = \mathbb{E}_t \exp \{ \gamma Y_{t+\tau} \} \quad (37)$$

$$= \exp \{ Tr [ A(\tau) \Sigma_t ] + b(\tau) Y_t + c(\tau) \}, \quad (38)$$

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

$$\frac{d}{d\tau} A(\tau) = A(\tau) M + (M^T + 2\gamma (R11 + R22) Q) A(\tau) + 2A(\tau) Q^T Q A(\tau) + \frac{\gamma(\gamma - 1)}{2} \mathbb{I}_n \quad (39)$$

$$A(0) = 0.$$

We have proved the following

**Proposition 8** *The Riccati equations satisfied by the matrix coefficient  $A(\tau)$  associated to the Laplace transform (37) are given by (39), where*

$$R11 = \begin{pmatrix} c + g & d + h \\ 0 & 0 \end{pmatrix}$$

$$R22 = \begin{pmatrix} 0 & 0 \\ a + e & b + f \end{pmatrix},$$

where the parameters  $a, b, c, d, e, f, g, h$  satisfy the (non-linear) constraints (30), (31), (32), (33), (34), (35).

**Remark 9** *Our model corresponds to choosing  $c = d = e = f = 0$  (or equivalently  $a = b = g = h = 0$ ): we obtain*

$$R12 = \begin{pmatrix} \rho_{21} & \rho_{22} \\ 0 & 0 \end{pmatrix}$$

$$R21 = \begin{pmatrix} 0 & 0 \\ \rho_{11} & \rho_{12} \end{pmatrix}$$

$$R11 = \begin{pmatrix} \rho_{11} & \rho_{12} \\ 0 & 0 \end{pmatrix}$$

$$R22 = \begin{pmatrix} 0 & 0 \\ \rho_{21} & \rho_{22} \end{pmatrix},$$

and

$$Z_t^{11} = W_t^{11} \rho_{11} + W_t^{12} \rho_{12} + \sqrt{1 - \rho_{11}^2 - \rho_{12}^2} B_t^{11} \quad (40)$$

$$Z_t^{12} = W_t^{11} \rho_{21} + W_t^{12} \rho_{22} + \sqrt{1 - \rho_{21}^2 - \rho_{22}^2} B_t^{12} \quad (41)$$

$$Z_t^{21} = W_t^{21} \rho_{11} + W_t^{22} \rho_{12} + \sqrt{1 - \rho_{11}^2 - \rho_{12}^2} B_t^{21} \quad (42)$$

$$Z_t^{22} = W_t^{21} \rho_{21} + W_t^{22} \rho_{22} + \sqrt{1 - \rho_{21}^2 - \rho_{22}^2} B_t^{22} \quad (43)$$

we can then introduce a matrix

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

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