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 §

March 16, 2006

Abstract

We consider a model for a single risky asset whose volatility follows a multifactor (matrix) Wishart process. Our model appears to be the most natural and genuine multifactor extension of the Heston model while preserving its analytical tractability: in fact, we completely solve the pricing problem through the Fast Fourier Transform as in Carr and Madan (1999). A simple numerical test shows that the Wishart (multifactor) process is more flexible than the Heston (single factor) one in providing a consistent pricing for options on both the underlying and its realized volatility.

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 usally 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 instruments, also called ratchet options, periodically "lock in" profits in a manner somewhat analogues to a mechanical ratchet. Exotic contracts like 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 whose prices depend only on realized volatility (see e.g. Hipp 1996). As

*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

well explained in Bergomi (2004), there is a structural limitation which prevents one-factor stochastic volatility models to price consistently these types of options jointly with plain vanilla options. A possible reconciliation requires that the volatility process is driven by at least 2 factors, even in a single asset framework, as supported by empirical tests like the principal component analysis investigated in Cont and Fonseca (2002).

Among one factor stochastic volatility models, the most popular and easy to implement is certainly the Heston (1992) one, in which the volatility satisfies a (positive) single factor square root process, where the pricing and hedging problem can be efficiently solved 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 model is not flexible enough to take into account the variability of the skew, also known as *correlation risk*.

The aim of this paper is to extend the Heston model to a multifactor specification for the volatility process in a single asset framework. The multifactor volatility process is a multi-dimensional version of the square root model which is called matrix Wishart process, mathematically developed in Bru (1991). Our model takes inspiration from the multi-asset market model analyzed in 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 assets noises. On the contrary, we introduce a correlation structure between the single-asset noise and volatility factors, in order to reproduce the skew effect on the implied volatility curve. Within our specification:

i) the term structure of the realized volatilities is described by a (correlated) multifactor model.

ii) the covariation between the asset's noise and each volatility factor can be parametrized and controlled separately.

iii) the analytic tractability of the Heston model is fully preserved: in fact we can provide an explicit solution for the Laplace transform, which completely solves the pricing problem through the FFT.

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 model, can fit independently the long-term volatility level and the short-term volatility skew. 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 standard Heston one. In Section 5 we provide some conclusions and future developments.

2 The Wishart volatility process

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

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

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 motions (i.e. composed by 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 (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:

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

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

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

with the real parameter $\beta \geq n-1$ (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:

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

where $S_t, Z_t \in \mathbb{R}^n$, $\mathbf{1} = (1, ..., 1)^T$ 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 (implicit) 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).

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 $\alpha, \beta \in \mathbb{R}^n$,

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

Here

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

■

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.

2.2 The stochastic correlation between stock and volatility

In order to compute the analogue of the correlation as in the original Heston model, we must consider the correlation between the stock noise and the noise driving its *scalar* volatility, represented by $Tr(\Sigma_t)$: this is computed in the following

Proposition 2 *The stochastic correlation between the stock noise and the volatility noise in the Wishart model is given by*

$$\rho_t = \frac{Tr\left[RQ\Sigma_t\right]}{\sqrt{Tr\left[\Sigma_t\right]}\sqrt{Tr\left[Q^TQ\Sigma_t\right]}}. \tag{3}$$

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 compute explicitly the dynamics of the stochastic correlation. This gives us a direct way to handle the correlation (and in turn the correlation risk) by mean of the matrix R.

Obviously, when both matrices R and Q are multiple of the identity matrix, we recover the usual constant correlation parameter. However, notice that even in the 'scalar' case $R = c\mathbb{I}$ where c is a real number, we have that ρ_t will be stochastic in the Wishart model: this opens the possibility to obtain a stochastic structure for the implied skew, which is a suitable feature of a stochastic volatility model in the spirit of Carr and Wu (2004).

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:

$$dY_t = \left(r - \frac{1}{2}Tr\left[\Sigma_t\right]\right)dt + Tr\left[\sqrt{\Sigma_t}\left(dW_t R^T + dB_t\sqrt{\mathbb{I} - RR^T}\right)\right]. \tag{4}$$

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 (4). 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 combinations 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:

$$\begin{aligned}\Psi_{\gamma,t}(\tau) &= \mathbb{E}_t \exp\left\{\gamma Y_{t+\tau}\right\} \\ &= \exp\left\{Tr\left[A(\tau)\Sigma_t\right] + b(\tau)Y_t + c(\tau)\right\},\end{aligned} \tag{5}$$

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

$$\begin{aligned}\frac{\partial \Psi_{\gamma,t}}{\partial \tau} &= \mathcal{L}_{Y,\Sigma} \Psi_{\gamma,t} \\ \Psi_{\gamma,t}(0) &= \exp\left\{\gamma Y_t\right\},\end{aligned} \tag{6}$$

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):

$$\mathcal{L}_{\Sigma} = Tr\left[\left(\Omega\Omega^T + M\Sigma + \Sigma M^T\right) D + 2\Sigma D Q^T Q D\right], \tag{7}$$

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, it is possible to find the infinitesimal generator of the couple (Y_t, Σ_t):

Proposition 3 *The infinitesimal generator of* (Y_t, Σ_t) *is given by*

$$\begin{aligned}\mathcal{L}_{Y,\Sigma} &= \left(r - \frac{1}{2} Tr\left[\Sigma\right]\right)\frac{\partial}{\partial y} + \frac{1}{2} Tr\left[\Sigma\right]\frac{\partial^2}{\partial y^2} \\ &\quad + Tr\left[\left(\Omega\Omega^T + M\Sigma + \Sigma M^T\right) D + 2\Sigma D Q^T Q D\right] + 2Tr\left[\Sigma R Q D\right]\frac{\partial}{\partial y}.\end{aligned} \tag{8}$$

Proof: See Appendix A.

Thus the exlicit expression of (6) is:

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

and by replacing the candidate (5) we obtain

$$\begin{aligned}0 &= -Tr\left[\frac{\partial}{\partial \tau} A(\tau)\Sigma\right] - \frac{\partial}{\partial \tau} b(\tau) Y - \frac{\partial}{\partial \tau} c(\tau) \\ &\quad + Tr\left[\left(\Omega\Omega^T + M\Sigma + \Sigma M^T\right) A(\tau) + 2\Sigma A(\tau) Q^T Q A(\tau) + 2\Sigma R Q A(\tau) b(\tau)\right] \\ &\quad + \left(r - \frac{1}{2} Tr\left[\Sigma\right]\right) b(\tau) + \frac{1}{2} Tr\left[\Sigma\right] b^2(\tau),\end{aligned} \tag{9}$$

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

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

hence

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

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

$$\begin{aligned}\frac{\partial}{\partial\tau}A(\tau)&=A(\tau)M+\left(M^T+2\gamma RQ\right)A(\tau)+2A(\tau)Q^TQA(\tau)+\frac{\gamma(\gamma-1)}{2}\mathbb{I}_n \\ A(0)&=0.\end{aligned} \tag{10}$$

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

$$\begin{aligned}\frac{\partial}{\partial\tau}c(\tau)&=Tr\left[\Omega\Omega^TA(\tau)\right]+\gamma r, \\ c(0)&=0.\end{aligned} \tag{11}$$

3.2 Matrix Riccati linearization

Matrix Riccati Equations like (10) 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 (10) and (11). Put

$$A(\tau)=F(\tau)^{-1}G(\tau) \tag{12}$$

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

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

and in the new variables the Riccati ODE leads to the system of $(2n)$ *linear* equations:

$$\begin{aligned}\frac{d}{d\tau}G(\tau)&=\frac{\gamma(\gamma-1)}{2}F(\tau)+G(\tau)M \\ \frac{d}{d\tau}F(t,\tau)&=-F(\tau)\left(M^T+2\gamma RQ\right)-2G(\tau)Q^TQ,\end{aligned} \tag{13}$$

which is solved by:

$$\begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau)\end{pmatrix}=\exp\tau\begin{pmatrix} M & -2Q^TQ \\ \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n & -\left(M^T+2\gamma RQ\right)\end{pmatrix} \tag{14}$$

In conclusion, we get

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

and since $A(0)=0$,

$$A\left(\tau\right)=A_{22}\left(\tau\right)^{-1}A_{21}\left(\tau\right), \tag{15}$$

which represents the closed-form solution of the Matrix Riccati (10). Let us now turn our attention to equation (11). We can improve its computation by the following trick: from (13) we obtain

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

and plugging into (12) and using the proprieties of the trace we deduce

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

Now we can integrate the last equation and obtain

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

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

Remark 4 *The computation of the Laplace Transform for both asset returns and variance factors*

$$\begin{aligned}\Psi_{\gamma,t}(\tau)&=\mathbb{E}_t\exp\left\{\gamma Y_{t+\tau}+Tr\left[\Gamma\Sigma_t\right]\right\}\\&=\exp\left\{Tr\left[\widetilde{A}(\tau)\Sigma_t\right]+\widetilde{b}(\tau)Y_t+\widetilde{c}(\tau)\right\},\end{aligned} \tag{16}$$

can be easily handled 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

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

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

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

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:

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

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 α is chosen (typically $\alpha = 1.1$, see 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 above 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 follow:

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

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

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

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

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

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

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

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

$$\Phi_{\gamma,0}(t,T) := \mathbb{E}\left[e^{i\gamma Y_{t,T}}\right]. \tag{17}$$

The modified option price is given by

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

and its Fourier transform

$$\begin{aligned}\psi_{t,T}(v) &= \int_{-\infty}^{+\infty} e^{ivk} c_{t,T}(k)dk \\ &= e^{-rT} \frac{\Phi_{(v-(\alpha+1)i),0}(t,T)}{(\alpha+iv)(\alpha+1+iv)},\end{aligned} \tag{18}$$

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

$$\begin{aligned}\Phi^{\Sigma}_{D,t}(\tau) &= \mathbb{E}_t \exp\{iTr[D\Sigma_{t+\tau}]\} \\ &= \exp\{Tr[B(\tau)\Sigma_t] + C(\tau)\},\end{aligned} \tag{19}$$

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

$$\begin{aligned}B(\tau) &= (iDB_{12}(\tau) + B_{22}(\tau))^{-1}(iDB_{11}(\tau) + B_{21}(\tau)) \\ C(\tau) &= Tr\left[\Omega\Omega^T \int_0^\tau B(s)ds\right],\end{aligned} \tag{20}$$

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.

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

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

where the last equality comes from (19), where $B(t)$ is given by (20) 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 (18) and apply the FFT in order to find the forward-starts 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 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 Heston model, whose dynamic is given by

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

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 (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: first we show that the Heston model can be nested into the Wishart model by a suitable choice of the parameters, then we consider the simplest modification of this choice which allows to reproduce a volatility surface which cannot be generated by the Heston model.

4.1 Case One: Heston-like Wishart model

In this first example we show how the Heston model can be nested into the Wishart model for a specific choice of the parameters. In fact, if all matrices involved in the Wishart dynamics are proportional to the identity matrix, it is

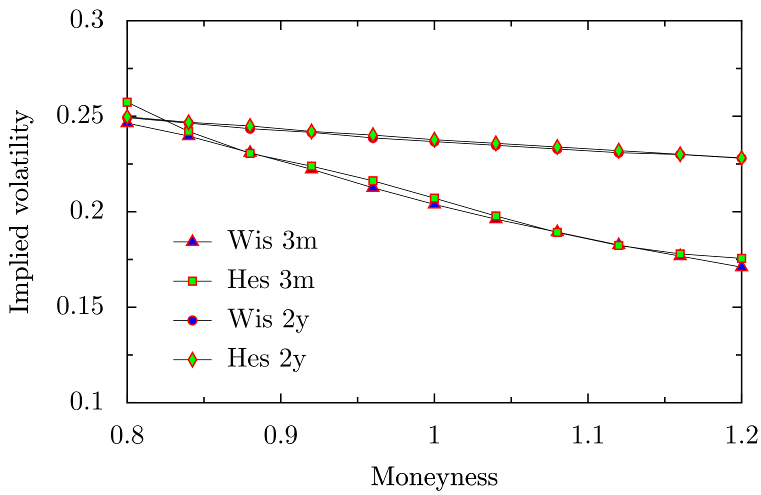

Figure 1: Implied volatility for the Wishart model (Wis) and Heston (Hes) model. Option maturities are 3 months (3m) and 2 years (2y). Moneyness is defined by $\frac{K}{S_0}$ where S_0 is the initial spot value.

straightforward to see that $Tr(\Sigma_t)$ follows a square root process. In particular, if we choose

$$M = \begin{pmatrix} -3 & 0 \\ 0 & -3 \end{pmatrix} \; R = \begin{pmatrix} -0.7 & 0 \\ 0 & -0.7 \end{pmatrix}$$

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

and $\beta = 3$, then the Wishart model is equivalent to a Heston model with parameters: $\kappa = 6$, $\theta = 0.25^2$, $\sigma_0 = 0.15$, $\epsilon = 0.5$ $\rho = -0.7$, which are consistent with those observed in the market. Figure (1) confirms that both models produce the same smile at different maturities.

4.2 Case Two: Wishart versus Heston volatility

In this example we show the flexibility of the Wishart model in describing some implied volatility patterns that cannot be reproduced by the Heston model. In fact we have the possibility to specify separately the two mean reversion parameters of the (diagonal) matrix M. In particular, if we leave the same value for $M_{11} = -3$ and we choose $M_{22} = -0.333$, then we can associate to the element Σ_{11} the meaning of a short-term factor, while Σ_{22} has an impact to the long-term volatility. With respect to our first example, let us take the following values:

$$M = \begin{pmatrix} -3 & 0 \\ 0 & -0.333 \end{pmatrix} \; R = \begin{pmatrix} -0.7 & 0 \\ 0 & -0.7 \end{pmatrix}$$

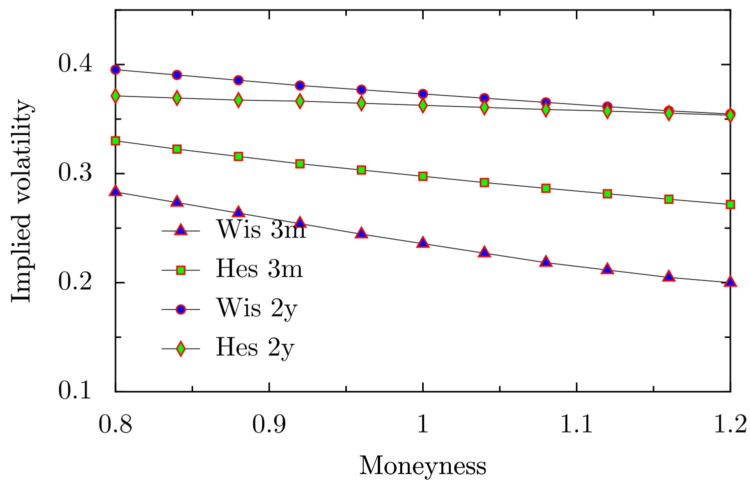

Figure 2: Implied volatility for the Wishart model (Wis) and Heston (Hes) model. Option maturities are 3 months (3m) and 2 years (2y). Moneyness is defined by $\frac{K}{S_0}$ where S_0 is the initial spot value.

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

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 remain unchanged: $\kappa = 6$, , $\sigma_0 = 0.15$, $\epsilon = 0.5$ $\rho = -0.7$. However, increasing the long-term volatility makes also increase the 3 months volatility, so that the short-term fit for the implied volatility is unsatisfactory, as illustrated in Figure (2).

On the other hand, we can fit perfectly the short-term volatility produced by 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 (3).

4.3 The impact of R on the smile

In the following figure we show the impact of the matrix R on the skew. It is well known that in the Heston model the skew is related to a (negative) correlation

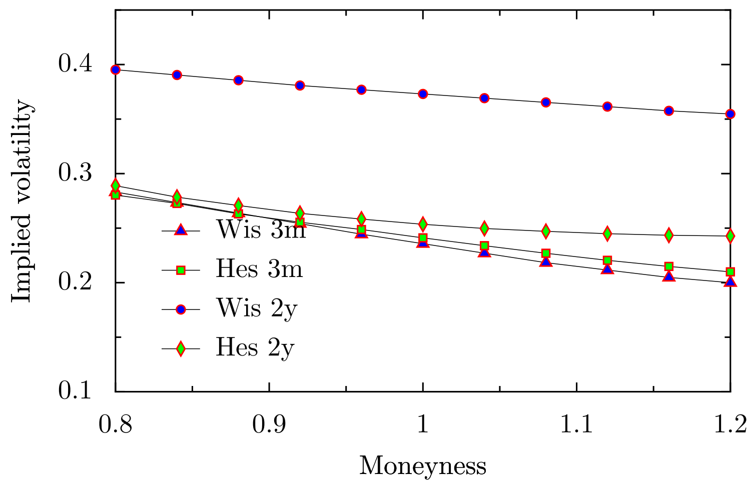

Figure 3: Implied volatility for the Wishart model (Wis) and Heston (Hes) model. Option maturities are 3 months (3m) and 2 years (2y). Moneyness is defined by $\frac{K}{S_0}$ where S_0 is the initial spot value.

between the volatility and the stock price. Taking the matrices

$$M = \begin{pmatrix} -5 & 0 \\ 0 & -3 \end{pmatrix} \; Q = \begin{pmatrix} 0.35 & 0 \\ 0 & 0.25 \end{pmatrix} \; \Sigma_0 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$$

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

and $\beta = 3$ in the Wishart model, we get for R_1 (resp. R_2) the left (resp. right) side hand of Figure (4).

The previous figures confirm that R is strictly related to the correlation structure and they show the flexibility of the Wishart model in describing the skew by mean of several parameters.

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

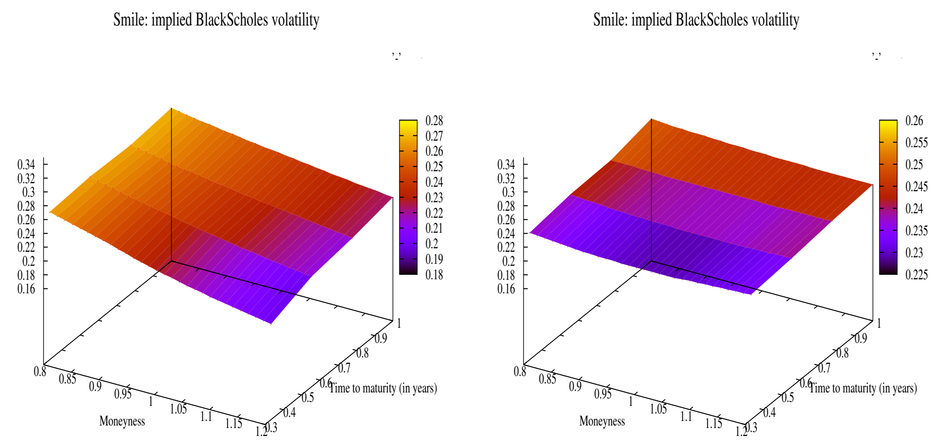

Figure 4: Implied volatility for the Wishart model corresponding to R_1 (left) and R_2 (right).

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) 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] Duffie, D. and R. Kan (1996) "A Yield-Factor Model of Interest Rates". *Mathematical Finance*, **6** (4), 379-406.

[8] Duffie, D., J. Pan and K. Singleton (2000) "Transform analysis and asset pricing for affine jump-diffusions". *Econometrica*, **68**, 1343-1376.

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

[10] Gourieroux, C. and R. Sufana (2003) "Wishart Quadratic Term Structure Models". CREF 03-10, HEC Montreal.

[11] Gourieroux, C. and R. Sufana (2004) "Derivative Pricing with Multivariate Stochastic Volatility: Application to Credit Risk". Working paper CREST.

[12] Grasselli M. and C. Tebaldi (2004) "Solvable Affine Term Structure Models". *Mathematical Finance*, to appear.

[13] C. Hipp (1996) "Options for Guaranteed Index-linked Life Insurance". AFIR 1996 Proceedings, Vol. II, S. 1463-1483.

[14] Wong, G. (2004) "Forward Smile and Derivative Pricing ". Working paper UBS.

6 Appendix A: Proofs

Proof of Proposition 2: The first step consists in finding the stock noise:

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

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

$$\begin{aligned}dTr\left[\Sigma_t\right] &= \left(Tr\left[\Omega\Omega^T\right] + 2Tr\left[M\Sigma_t\right]\right)dt + 2Tr\left[\sqrt{\Sigma_t}dW_tQ\right]\\ &= ...dt + 2\sqrt{Tr\left[\Sigma_tQ^TQ\right]}\frac{Tr\left[\sqrt{\Sigma_t}dW_tQ\right]}{\sqrt{Tr\left[\Sigma_tQ^TQ\right]}}\\ &= ...dt + 2\sqrt{Tr\left[\Sigma_tQ^TQ\right]}dw_t,\end{aligned}$$

where we used the fact that

$$
\begin{aligned}
d\left\langle Tr\left[\Sigma_{.}\right]\right\rangle_t &= \sum_{ij} Cov_t\left(e_i^T d\Sigma_t e_i, e_j^T d\Sigma_t e_j\right)\\
&= 4\sum_{ij} Cov_t\left(e_i^T\sqrt{\Sigma_t} dW_t Q e_i, e_j^T\sqrt{\Sigma_t} dW_t Q e_j\right)\\
&= 4\sum_{ij} e_i^T\sqrt{\Sigma_t} dW_t Q e_i e_j^T Q^T dW_t^T\sqrt{\Sigma_t} e_j\\
&= 4\sum_{ij} e_i^T\sqrt{\Sigma_t} Tr\left[Q e_i e_j^T Q^T\right]\sqrt{\Sigma_t} e_j dt\\
&= 4\sum_{ij} Tr\left[Q^T Q e_i e_j^T\right] e_i^T \Sigma_t e_j dt\\
&= 4\sum_{ij} e_j^T \Sigma_t e_i e_i^T Q^T Q e_j dt\\
&= 4\sum_{j} e_j^T \Sigma_t Q^T Q e_j dt\\
&= 4Tr\left[\Sigma_t Q^T Q\right] dt.
\end{aligned}
$$

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:

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

■

Proof of Proposition 3: The only non trivial term in (8) comes from the covariation

$$d < \Sigma^{ij}, Y >_t, \text{ for } i, j = 1, ..., n.$$

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 $\left(\frac{\partial^2}{\partial x_{ij}\partial y}\right)$, thus obtaining

$$\begin{aligned}
d < \Sigma^{ij}, Y >_t &= 2\left(\sum_{l,k=1}^{n} \sigma_t^{il} dW_{lk} Q_{kj}\right)\left(\sum_{l,k,h=1}^{n} \sigma_t^{lk} dW_{kh} R_{lh}\right) \\
&= 2\sum_{l,k,h=1}^{n} \sigma_t^{il} Q_{kj} \sigma_t^{hl} R_{hk} dt \\
&= 2\sum_{k,h=1}^{n}\left(\sum_{l=1}^{n} \sigma_t^{il}\sigma_t^{hl}\right) Q_{kj} R_{hk} dt \\
&= 2\sum_{k,h=1}^{n} \Sigma_t^{ih} Q_{kj} R_{hk} dt,
\end{aligned}$$

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

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

and since by definition D is symmetric. ∎

Proof of Proposition 5: We repeat the reasoning as in (5) 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

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

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

$$\begin{aligned}
\begin{pmatrix} G(\tau) & F(\tau) \end{pmatrix} &= \begin{pmatrix} G(0) & F(0) \end{pmatrix} \exp\tau \begin{pmatrix} M & -2Q^T Q \\ 0 & -M^T \end{pmatrix} \\
&= \begin{pmatrix} B(0) & \mathbb{I}_n \end{pmatrix} \exp\tau \begin{pmatrix} M & -2Q^T Q \\ 0 & -M^T \end{pmatrix} \\
&= \begin{pmatrix} iDB_1^1(\tau) + B_1^2(\tau) & iDB_2^1(\tau) + B_2^2(\tau) \end{pmatrix},
\end{aligned}$$

which gives the statement. ∎

7 Appendix B: The 2-dimensional case

In this Appendix we develop the computations in (7) in the case $n = 2$. This means that the Wishart process Σ_t satisfies the following SDE:

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

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

$$\sigma_t^2 = \Sigma_t = \begin{pmatrix} \left(\sigma_t^{11}\right)^2 + \left(\sigma_t^{12}\right)^2 & \sigma_t^{11}\sigma_t^{12} + \sigma_t^{12}\sigma_t^{22} \\ \sigma_t^{11}\sigma_t^{12} + \sigma_t^{12}\sigma_t^{22} & \left(\sigma_t^{12}\right)^2 + \left(\sigma_t^{22}\right)^2 \end{pmatrix}. \tag{21}$$

We obtain

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

and using (21):

$$\begin{aligned}
&d < \Sigma^{11}, \Sigma^{11} >_t = 4\Sigma_t^{11}(Q_{11}^2 + Q_{21}^2)dt, \\
&d < \Sigma^{12}, \Sigma^{12} >_t = \left(\Sigma_t^{11}\left(Q_{12}^2 + Q_{22}^2\right) + 2\Sigma_t^{12}\left(Q_{11}Q_{12} + Q_{21}Q_{22}\right) + \Sigma_t^{22}(Q_{11}^2 + Q_{21}^2)\right) dt, \\
&d < \Sigma^{22}, \Sigma^{22} >_t = 4\Sigma_t^{22}(Q_{12}^2 + Q_{22}^2)dt, \\
&d < \Sigma^{11}, \Sigma^{12} >_t = \left(2\Sigma_t^{11}\left(Q_{11}Q_{12} + Q_{21}Q_{22}\right) + 2\Sigma_t^{12}\left(Q_{11}^2 + Q_{21}^2\right)\right) dt, \\
&d < \Sigma^{11}, \Sigma^{22} >_t = 4\Sigma_t^{12}\left(Q_{11}Q_{12} + Q_{21}Q_{22}\right) dt, \\
&d < \Sigma^{12}, \Sigma^{22} >_t = 2\left(\Sigma_t^{12}\left(Q_{12}^2 + Q_{22}^2\right) + \Sigma_t^{22}\left(Q_{11}Q_{12} + Q_{21}Q_{22}\right)\right) dt.
\end{aligned}$$

On the other hand, from (7) we can identify the coefficient of $\left(\frac{\partial^2}{\partial\Sigma_{ij}\partial\Sigma_{lk}}\right)$ 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:

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

thus proving the equality in (7), since

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

where we recall that

$$2<\Sigma^{12},\Sigma^{12}>_t\frac{\partial^2}{\left(\partial\Sigma^{12}\right)^2}=<\Sigma^{12},\Sigma^{12}>_t\frac{\partial^2}{\left(\partial\Sigma^{12}\right)^2}+<\Sigma^{21},\Sigma^{21}>_t\frac{\partial^2}{\left(\partial\Sigma^{21}\right)^2};$$

$$4<\Sigma^{11},\Sigma^{12}>_t\frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}}=2<\Sigma^{11},\Sigma^{12}>_t\frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{12}}+2<\Sigma^{11},\Sigma^{21}>_t\frac{\partial^2}{\partial\Sigma^{11}\partial\Sigma^{21}}.$$

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

$$Z_t^{11} = Tr\left[W_t R11^T\right] + \sqrt{1 - Tr\left[R11R11^T\right]} B_t^{11} \tag{22}$$

$$Z_t^{12} = Tr\left[W_t R12^T\right] + \sqrt{1 - Tr\left[R12R12^T\right]} B_t^{12} \tag{23}$$

$$Z_t^{21} = Tr\left[W_t R21^T\right] + \sqrt{1 - Tr\left[R21R21^T\right]} B_t^{21} \tag{24}$$

$$Z_t^{22} = Tr\left[W_t R22^T\right] + \sqrt{1 - Tr\left[R22R22^T\right]} B_t^{22} \tag{25}$$

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*

$$Tr\left[RijRlm^T\right] = 0 \text{ for } (i,j) \neq (l,m), \quad i,j,l,m \in \{1,2\}. \tag{26}$$

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

$$\begin{aligned}
Cov_t\left(dZ_t\alpha, dZ_t\beta\right)_{11} = & \Big(Tr\left[dW_t R11^T\right]\alpha_1 + \sqrt{1 - Tr\left[R11R11^T\right]} dB_t^{11}\alpha_1 \\
& + Tr\left[dW_t R12^T\right]\alpha_2 + \sqrt{1 - Tr\left[R12R12^T\right]} dB_t^{12}\alpha_2\Big) \\
& . \Big(Tr\left[dW_t R11^T\right]\beta_1 + \sqrt{1 - Tr\left[R11R11^T\right]} dB_t^{11}\beta_1 \\
& + Tr\left[dW_t R12^T\right]\beta_2 + \sqrt{1 - Tr\left[R12R12^T\right]} dB_t^{12}\beta_2\Big) \\
& = \alpha_1\beta_1 + \alpha_2\beta_2 \\
& + \left(\alpha_1\beta_2 + \alpha_2\beta_1\right)\left(R11_{11}R12_{11} + R11_{12}R12_{12} + R11_{21}R12_{21} + R11_{22}R12_{22}\right).
\end{aligned}$$

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

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

that is $Tr\left[R11R12^T\right] = 0$. Similar computations for the other components lead to the conclusion. ■

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:

$$
\begin{aligned}
d < \Sigma^{11}, Y >_t &= \left(\sigma_t^{11} dZ_t^{11} + \sigma_t^{12} dZ_t^{12} + \sigma_t^{12} dZ_t^{21} + \sigma_t^{22} dZ_t^{22}\right) d\Sigma_t^{11} \\
&= 2\Big(\left(\sigma_t^{11}\right)^2 Q_{11} R11_{11} + \left(\sigma_t^{11}\right)^2 Q_{21} R11_{12} + \sigma_t^{11}\sigma_t^{12} Q_{11} R11_{21} \\
&+ \sigma_t^{11}\sigma_t^{12} Q_{21} R11_{22} + \sigma_t^{11}\sigma_t^{12} Q_{11} R12_{11} + \sigma_t^{11}\sigma_t^{12} Q_{21} R12_{12} \\
&+ \left(\sigma_t^{12}\right)^2 Q_{11} R12_{21} + \left(\sigma_t^{12}\right)^2 Q_{21} R12_{22} + \sigma_t^{11}\sigma_t^{12} Q_{11} R21_{11} \\
&+ \sigma_t^{11}\sigma_t^{12} Q_{21} R21_{12} + \sigma_t^{11}\sigma_t^{12} Q_{11} R21_{21} + \left(\sigma_t^{12}\right)^2 Q_{21} R21_{22} \\
&+ \sigma_t^{11}\sigma_t^{22} Q_{11} R22_{11} + \sigma_t^{11}\sigma_t^{22} Q_{21} R22_{12} + \sigma_t^{12}\sigma_t^{22} Q_{11} R22_{21} + \sigma_t^{12}\sigma_t^{22} Q_{21} R22_{22} \}
\end{aligned}
$$

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 of $d < \Sigma^{22}, Y >_t$ 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 $d < \Sigma^{12}, Y >_t$. In conclusion, the affinity constraint lead to the following specification for the 4 correlation matrices:

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

Now we impose (26) and obtain:

$$
\begin{aligned}
R11 \perp R21 &\longrightarrow e(c+g) + f(d+h) = 0 && (27) \\
R11 \perp R12 &\longrightarrow a(c+g) + b(d+h) = 0 && (28) \\
R22 \perp R21 &\longrightarrow g(a+e) + h(b+f) = 0 && (29) \\
R22 \perp R12 &\longrightarrow c(a+e) + d(b+f) = 0 && (30) \\
R12 \perp R21 &\longrightarrow ae + bf + cg + dh = 0. && (31)
\end{aligned}
$$

After some manipulations we arrive to

$$\frac{ae}{(a+e)^2} + \frac{cg}{(b+f)^2} = 0. \tag{32}$$

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, Σ_t) *is given by*

$$\begin{aligned}\mathcal{L}_{Y,\Sigma} &= \left(r - \frac{1}{2}Tr\left[\Sigma\right]\right)\frac{\partial}{\partial y} + \frac{1}{2}Tr\left[\Sigma\right]\frac{\partial^2}{\partial y^2} \\ &\quad + Tr\left[\left(\Omega\Omega^T + M\Sigma + \Sigma M^T\right)D + 2\Sigma D Q^T Q D\right] + 2Tr\left[\Sigma\left(R11 + R22\right)QD\right]\frac{\partial}{\partial y}.\end{aligned} \tag{33}$$

Proof: We focus on the covariation terms $d < \Sigma^{ij}, Y >_t$, for $i, j = 1, ..., 2:$

$$\begin{aligned}
d &< \Sigma^{11}, Y >_t = 2Q_{11}\left(\left(c+g\right)\Sigma^{11} + \left(a+e\right)\Sigma^{12}\right) \\
&\quad + 2Q_{21}\left(\left(d+h\right)\Sigma^{11} + \left(b+f\right)\Sigma^{12}\right) \\
d &< \Sigma^{22}, Y >_t = 2Q_{12}\left(\left(a+e\right)\Sigma^{22} + \left(c+g\right)\Sigma^{12}\right) \\
&\quad + 2Q_{22}\left(\left(d+h\right)\Sigma^{12} + \left(b+f\right)\Sigma^{22}\right) \\
d &< \Sigma^{12}, Y >_t = Q_{12}\left(\left(c+g\right)\Sigma^{11} + \left(a+e\right)\Sigma^{12}\right) \\
&\quad + Q_{22}\left(\left(d+h\right)\Sigma^{11} + \left(b+f\right)\Sigma^{12}\right) \\
&\quad + Q_{11}\left(\left(c+g\right)\Sigma^{12} + \left(a+e\right)\Sigma^{22}\right) \\
&\quad + Q_{21}\left(\left(d+h\right)\Sigma^{12} + \left(b+f\right)\Sigma^{22}\right)
\end{aligned}$$

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

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

and by definition D is symmetric. ■

By applying the Feynman-Kac argument to the Laplace transform

$$\begin{aligned}\Psi_{\gamma,t}(\tau) &= \mathbb{E}_t \exp\left\{\gamma Y_{t+\tau}\right\} && (34)\\ &= \exp\left\{Tr\left[A(\tau)\Sigma_t\right] + b(\tau)Y_t + c(\tau)\right\}, && (35)\end{aligned}$$

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

$$\frac{\partial}{\partial\tau}A(\tau) = A(\tau)M + \left(M^T + 2\gamma\left(R11+R22\right)Q\right)A(\tau) + 2A(\tau)Q^TQA(\tau) + \frac{\gamma(\gamma-1)}{2}\mathbb{I}_n \tag{36}$$

$$A(0) = 0.$$

We have proved the following

Proposition 8 *The Riccati equations satisfied by the matrix coefficient $A(\tau)$ associated to the Laplace transform (34) are given by (36), 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 (27), (28), (29), (30), (31), (32).

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

$$Z_t^{11} = W_t^{11}\rho_{11} + W_t^{12}\rho_{12} + \sqrt{1-\rho_{11}^2-\rho_{12}^2}B_t^{11} \quad (37)$$
$$Z_t^{12} = W_t^{11}\rho_{21} + W_t^{12}\rho_{22} + \sqrt{1-\rho_{21}^2-\rho_{22}^2}B_t^{12} \quad (38)$$
$$Z_t^{21} = W_t^{21}\rho_{11} + W_t^{22}\rho_{12} + \sqrt{1-\rho_{11}^2-\rho_{12}^2}B_t^{21} \quad (39)$$
$$Z_t^{22} = W_t^{21}\rho_{21} + W_t^{22}\rho_{22} + \sqrt{1-\rho_{21}^2-\rho_{22}^2}B_t^{22} \quad (40)$$

we can then introduce a matrix

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

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