INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2005-43

Whishart Autoregressive Model for Stochastic Risk

C. GOURIEROUX[1]

---

[1] CREST, CEPREMAP (Paris) and University of Toronto.
Prepared for the special issue of Econometric Review on stochastic volatility. This version has benefited of helpful comments by J. Jasiak, H. Massam, A. Monfort, C. Robert, R. Sufana.

WISHART AUTOREGRESSIVE MODEL FOR STOCHASTIC RISK

C. GOURIEROUX [1]

(May, 2005)

[1]CREST, CEPREMAP and University of Toronto.
Prepared for the special issue of Econometric Review on stochastic volatility. This version has benefited of helpful comments by J. Jasiak, H., Massam, A., Monfort, C., Robert, R., Sufana.

Wishart Autoregressive Model for Stochastic Risk

Abstract

Risks are usually represented and measured by volatility-covolatility matrices. Wishart processes are models for a dynamic analysis of multivariate risk, that describe the evolution of stochastic volatility-covolatility matrices, constrained to be symmetric positive definite. The autoregressive Wishart process (WAR) is the multivariate extension of the Cox, Ingersoll, Ross (CIR) process introduced for scalar stochastic volatility. As the CIR process it allows for closed form solutions for a number of financial problems, such as term structure of T-bonds and corporate bonds, derivative pricing in multivariate stochastic volatility model and structural model for credit risk. Moreover the Wishart dynamics are very flexible and are serious competitors for less structural multivariate ARCH models.

Keywords : Wishart Process, Stochastic Volatility, Factor, Quadratic Term Structure, Credit Risk.

JEL Number :G12, G13.

Le modèle Wishart autorégressif pour risque stochastique

Résumé

Les risques sont habituellement représentés et mesurés par des matrices de volatilité. Les processus Wishart sont des modèles permettant une analyse dynamique des risques multivariés; ils décrivent l'évolution de matrices de volatilité stochastique, matrices symétriques positives. Le processus de Wishart autorégressif (WAR) est l'extension multivariée du processus de Cox, Ingersoll, Ross (CIR) introduit pour une volatilité scalaire. Comme le processus CIR, il conduit à des solutions explicites pour un grand nombre de problèmes financiers, tels les structures par terme des obligations d'état et d'entreprises, la valorisation des produits dérivés dans les modèles multivariés à volatilité stochastique et les modèles structurels de risque de crédit. De plus la dynamique Wishart est très flexible et apparaît comme un concurrent sérieux des modèles ARCH multivariés moins structurels.

Mots Clés : Processus Wishart, volatilité stochastique, facteur, structure par terme quadratique, risque de crédit.

1 INTRODUCTION

Risks are usually represented and measured by volatility-covolatility matrices. Wishart processes are models for a dynamic analysis of multivariate risk, that describe the evolution of stochastic volatility-covolatility matrices, constrained to be symmetric positive definite. The Wishart process is the direct multivariate extension of the Cox-Ingersoll-Ross (CIR) process introduced for a scalar volatility. Its importance is easily understood if we remind the advantages and limitations of the CIR process. Loosely speaking the CIR process allows for closed form solutions for a number of financial problems. These problems concern the term structure of riskfree interest rate [Cox-Ingersoll-Ross model, Cox-Ingersoll-Ross (1985)], the analysis of default risk [the so-called stochastic default intensity model considered by Lando (1998)], the derivative pricing in stochastic volatility model [the so-called Heston's model considered in Heston (1993) and Ball, Roma (1994)], the analysis of intertrade durations [Gourieroux, Jasiak (2005)] and the associated models of time deformation. The counterpart of a search for closed form solutions is generally a lack of flexibility of the model and possible misspecification. This limitation exists for the CIR based financial models and is well-documented, especially for applications to term structure. By increasing the dimensionality of the risk, that is by replacing a scalar volatility by a volatility-covolatility matrix, Wishart models increase significantly the flexibility of the associated models while keeping the advantage of closed form solutions. Moreover the structural interpretations of the Wishart based models make them direct competitors of flexible descriptive approaches such as multinomial trees [e.g. Ho, Lee (2004)], or dynamic conditional correlation GARCH models [Engle, Sheppard (2001)].

In Section 2 we first consider the CIR process and its main financial applications to term structure and derivative pricing. The aim of Section 3 is to introduce the Wishart process as the multivariate extension of the CIR process and explain why it allows to extend the financial models based on the CIR process to a multiasset framework, while keeping closed form solutions. In Section 4 we discuss the flexibility of the Wishart process for representing various dynamics of volatilities-covolatilities. We also explain how to reduce the number of parameters appearing in the unconstrained Wishart process by looking for factor representations and we discuss statistical inference. Section 5 concludes. The proofs are gathered in appendices.

2 The Cox-Ingersoll-Ross Process

The Cox-Ingersoll-Ross model specifies a dynamics for a positive scalar process. This process, denoted by (y_t) in this section, can represent a return volatility, an interest rate, a stochastic discount factor, the difference between ask and bid prices, or a latent risk factor. We first describe the dynamics and the distributional properties of the process. In a second subsection we discuss the links with the Ornstein-Uhlenbeck process. Finally we present the applications of the Cox-Ingersoll-Ross process to the term structure of interest rates and to derivative pricing in a stochastic volatility framework.

2.1 Definition and Distributional Properties.

2.1.1 Definition

Definition 1 : The Cox-Ingersoll-Ross process satisfies the diffusion equation :

$$dy_t = -k(y_t - \theta)dt + (\eta^2 y_t)^{1/2} dW_t,$$

where (W_t) is a univariate Brownian motion and the parameters satisfy :
$\eta > 0, k\theta > 0$.

This process is also called square root process due to the expression $(\eta^2 y_t)^{1/2}$ of the process volatility. The condition $\eta > 0$ can be always imposed since only the square η^2 matters. The inequality restriction $k\theta > 0$ ensures the positivity of the process (if the initial value $y_o > 0$). Indeed at a date t where the process vanishes $y_t = 0$, the process becomes locally deterministic since $\eta^2 y_t = 0$, whereas the drift is equal to $k\theta dt > 0$. The positivity of the drift implies a rejection towards positivity and explains why the process satisfies the expected positivity restriction.

Finally note that the drift $E_t dy_t = -k(y_t - \theta)dt$, and the volatility $V_t(dy_t) = \eta^2 y_t dt$ are both affine functions of the current process value y_t. Thus the Cox-Ingersoll-Ross process is an example of so-called affine process [Duffie, Kan (1996)]. This explains the possibility to get closed form expressions for the process distribution by means of the conditional Laplace transform.

2.1.2 Conditional Laplace Transform

In Definition 1 the dynamics is described by the recursive equation, which explains how to generate y_{t+dt} given y_t. More generally it is useful to analyze the transitions at any horizon h, that is to consider the conditional distribution of y_{t+h} given y_t. This can be done in various ways. First we will consider the (conditional) Laplace transform (or moment generating function) defined by :

$$\psi_{t,h}(u) = E\left(\exp[-uy_{t+h}]|y_t\right), u \geq 0,$$

which gives the conditional moment of any decreasing exponential transform of y_{t+h}. Since (y_t) is a positive process, the conditional Laplace transform characterizes the transition between dates t and $t+h$ [Feller (1971)]. Alternatively we can look for the conditional density of y_{t+h} given y_t, or for a practical way to simulate y_{t+h} for given y_t. In this subsection we first focus on the Laplace transform and consider the other approaches in the following subsections.

Proposition 1 : The (conditional) Laplace transform of the CIR process is :

$$\begin{aligned} \psi_{t,h}(u) &= E_t[\exp(-uy_{t+h})] \\ &= \exp -[a(h,u)y_t + b(h,u)], \end{aligned}$$

where functions a, b satisfy the differential equations:

$$\begin{aligned} \frac{\partial a(h,u)}{\partial h} &= -ka(h,u) - \frac{\eta^2}{2}a(h,u)^2, h \geq 0, \\ \frac{\partial b(h,u)}{\partial h} &= k\theta a(h,u), h \geq 0, \end{aligned}$$

with initial conditions : $a(0,u) = u, b(0,u) = 0$.
The solutions of the above differential system are :

$$
\begin{aligned}
a(h,u) &= \frac{u\exp(-kh)}{1+\dfrac{\eta^2 u}{2k}[1-\exp(-kh)]}, \\
b(h,u) &= \frac{2k\theta}{\eta^2}\log\left[1+u\frac{\eta^2}{2k}(1-\exp(-kh))\right].
\end{aligned}
$$

Proof : See Appendix 1.1 and Appendix 2.1.

The conditional Laplace transform is an exponential affine function of the current value y_t, which is another characterization of affine diffusion models [see Duffie, Filipovic, Schachermayer (2003)]. Note also that $\lim_{h\to\infty} a(h,u) = 0$, if and only if parameter k is strictly positive. The condition : $k > 0$, is a stationarity condition of the CIR process. Under this stationarity condition the transition at horizon h tends to the stationary distribution (or marginal distribution) of the process in the long run ($h \to \infty$). This marginal distribution admits the Laplace transform :

$$
\begin{aligned}
E\exp(-uy_t) &= \lim_{h\to\infty} \exp[-b(h,u)] \\
&= \exp\left\{-\frac{2k\theta}{\eta^2}\log\left[1+u\frac{\eta^2}{2k}\right]\right\},
\end{aligned}
$$

or :

$$
E\exp(-uy_t) = \left(1+u\frac{\eta^2}{2k}\right)^{-\frac{2k\theta}{\eta^2}}. \tag{2.1}
$$

2.1.3 The Transition Distribution

The transition distribution of y_{t+h} given y_t has been characterized in the previous subsection in terms of its conditional Laplace transform. The proposition below provides the interpretation of the transition in terms of standard families of distributions and gives the expression of the conditional density.

Proposition 2 :
i) The conditional distribution of y_{t+h} given y_t is a noncentered gamma distribution up to a scale factor. This distribution denoted by $\gamma[\nu, \frac{\rho(h)}{\lambda(h)}y_t, \lambda(h)]$ is such that : $y_{t+h}/\lambda(h)$ follows a centered gamma distribution $\gamma(\nu+Z_{t+h})$, with

a stochastic degree of freedom such that Z_{t+h} follows the Poisson distribution $\mathcal{P}\left[\frac{\rho(h)}{\lambda(h)}y_t\right]$.

ii) The parameters are :

$$\nu = \frac{2k\theta}{\eta^2}, \rho(h) = \exp(-kh), \lambda(h) = \frac{\eta^2}{2k}[1 - \exp(-kh)].$$

iii) The conditional pdf is given by :

$$\begin{aligned} f(y_{t+h}|y_t) \quad &= \exp\left[-\left(y_{t+h} - \rho(h)y_t\right)/\lambda(h)\right] \frac{y_{t+h}^{\nu-1}}{\lambda(h)^\nu} \\ &\sum_{z=0}^{\infty}\left[\frac{1}{z!\Gamma(\nu+z)}\left[\frac{y_{t+h}\rho(h)y_t}{\lambda(h)^2}\right]^z\right\}. \end{aligned}$$

iv) When $k > 0$ and h tends to infinity, we get the marginal distribution, which is a centered gamma distribution up to a scale factor :

$$\gamma(\nu, 0, \eta^2/2k).$$

Proof : See Appendix 3.

The first part of Proposition 2 provides a simple approach for simulating discrete time paths $y_1, y_2, \dots, y_t, y_{t+1} \dots$, say, of a Cox-Ingersoll-Ross process. Once y_t^s has been drawn , the intermediate stochastic degree of freedom Z_{t+1}^s is drawn in the appropriate Poisson distribution, and then y_{t+1}^s is drawn in the appropriate centered gamma distribution $\lambda(\nu + Z_{t+1}^s)$, up to the scale factor $\lambda(1)$.

The conditional p.d.f. involves a series expansion, whose terms tend quickly to zero when z increases. Thus this pdf can be approximated by its truncated version with a finite number of terms only.

2.1.4 Integrated Process

In financial applications, we are often interested in the integrated process, that is in quantities such as :

$$\int_t^{t+h} y_\tau d\tau.$$

The conditional distribution of the integrated process is also easily derived by means of its conditional Laplace transform.

Proposition 3 :

i) The conditional Laplace transform of the integrated CIR process (for $u = 1$) :

$$\psi^*_{t,h}(1) = E_t \left[\exp(-\int_t^{t+h} y_\tau d\tau) \right]$$

can be written as :

$$\psi^*_{t,h}(1) = \exp\left(-a^*(h) y_t - b^*(h)\right),$$

where functions a^*, b^* satisfy the partial differential system :

$$\begin{cases} \dfrac{\partial a^*(h)}{\partial h} &= 1 - k a^*(h) - \dfrac{\eta^2}{2}(a^*(h))^2, \\ \\ \dfrac{\partial b^*(h)}{\partial h} &= k\theta a^*(h), \end{cases}$$

with initial conditions : $a^*(0) = 0, b^*(0) = 0$.

The solutions of this system are :

$$\begin{aligned} a^*(h) &= \frac{2}{\gamma + k} - \frac{4\gamma}{\gamma + k} \frac{1}{(\gamma + k)\exp(\gamma h) + \gamma - k}, \\ b^*(h) &= -\frac{k\theta(\gamma + k)}{\eta^2} h + \frac{2k\theta}{\eta^2} \log[(\gamma + k)\exp(\gamma h) + \gamma - k] \\ &- \frac{2k\theta}{\eta^2} \log(2\gamma), \\ \text{where : } \gamma &= \sqrt{k^2 + 2\eta^2}. \end{aligned}$$

Proof : See Appendix 1.2 and Appendix 2.1.

The general expression of the conditional Laplace transform of the integrated CIR process is easily deduced from Proposition 3. Indeed this Laplace transform is :

$$\begin{aligned}
\psi^*_{t,h}(u) &= E_t \exp(-u \int_t^{t+h} y_\tau d\tau) \\
&= E_t[\exp - \int_t^{t+h} (uy_\tau) d\tau].
\end{aligned}$$

This is the Laplace transform of the process $y^*_t = uy_t$, which is still a Cox-Ingersoll-Ross process, since :

$$\begin{aligned}
d(y^*_t) &= -uk(y_t - \theta)dt + (u^2\eta^2 y_t)^{1/2} dW_t \\
&= -k(y^*_t - u\theta)dt + (\eta^{*2} y^*_t)^{1/2} dW_t, \\
\text{with : } \quad k^* &= k, \theta^* = u\theta, \eta^{*2} = u\eta^2.
\end{aligned}$$

Therefore : $\psi^*_{t,h}(u) = E_t\left[-a^*(h,u)y_t - b^*(h,u)\right]$

$$\begin{aligned}
\text{where : } \quad a^*(h,u) &= \frac{2u}{\gamma(u)+k} - \frac{4u\gamma(u)}{\gamma(u)+k} \frac{1}{(\gamma(u)+k)\exp[\gamma(u)h] + \gamma(u) - k}, \\
b^*(h,u) &= -\frac{k\theta}{\eta^2}[\gamma(u)+k]h + \frac{2k\theta}{\eta^2} \log[(\gamma(u)+k)\exp[\gamma(u)h] + \gamma(u) - k] \\
&\quad - \frac{2k\theta}{\eta^2} \log(2\gamma(u)), \\
\text{where : } \quad \gamma(u) &= \sqrt{k^2 + 2u\eta^2}.
\end{aligned}$$

2.1.5 Link with the Ornstein-Uhlenbeck Process.

Let us consider an Ornstein-Uhlenbeck process defined by :

$$dx_t = ax_t dt + w dW_t, \text{ say.}$$

By applying Ito's formula, the dynamics followed by the square of this process $y_t = x_t^2$ is :

$$\begin{aligned} d(x_t^2) &= 2x_t(ax_t dt + wdW_t) + w^2 dt, \\ \text{or : } \quad dy_t &= (2ay_t + w^2)dt + 2wy_t^{1/2}dW_t. \end{aligned}$$

More generally let us consider J independent Ornstein-Uhlenbeck processes with identical parameters :

$$dx_{jt} = ax_{jt}dt + wdW_{jt}, j = 1, \ldots, J,$$

where the Brownian motions $(W_{jt}), j = 1, \ldots, J$ are independent.

The process $y_t = x_{1t}^2 + \ldots + x_{Jt}^2$ is such that :

$$dy_t = \sum_{j=1}^{J} d(x_{jt}^2) = \sum_{j=1}^{J}(2ax_{jt}^2 + w^2)dt + \sum_{j=1}^{J} 2wx_{jt}dW_{jt},$$

or equivalently by aggregating the Brownian motions :

$$dy_t = (2ay_t + Jw^2)dt + 2wy_t^{1/2}dW_t.$$

Proposition 4 :

The sum of squares of J independent Ornstein-Uhlenbeck processes with identical parameters a, w^2 is a Cox-Ingersoll-Ross process with :

$$k = -2a, \eta = 2w, k\theta = Jw^2.$$

In this way we generate all CIR processes such that : $4k\theta/\eta^2$ is integer.

2.2 Financial Applications

Two equivalent alternative approaches can be followed for asset pricing based on no arbitrage arguments. On the one hand we can consider the risk-neutral world and define the asset price as the risk neutral expectation of the sum of future discounted payoffs. On the second hand we can implement the computations in the historical (real) world and introduce the dynamic risk correction by means of a stochastic discount factor. In the first subsection both approaches are applied to the analysis of the term structure of interest rates. In the second subsection we consider derivative pricing in the framework of stochastic volatility models.

2.2.1 Term Structure of Interest Rates.

Let us denote by $B(t, t+h) = \exp[-hr(t, t+h)]$ the price at t of a zero-coupon bond with time to maturity h and by $r(t, t+h)$ the interest rate (geometric yield) at time t for term h. The infinitesimal riskfree rate at time t is : $r_t = \lim_{h\to 0} r(t, t+h)$.

i) Risk-Neutral Approach

In the risk-neutral approach the price of the zero-coupon bond is given by :

$$B(t, t+h) = \overset{Q}{E_t}[\exp(-\int_t^{t+h} r_\tau d\tau)],$$

where $\overset{Q}{E_t}$ denotes the conditional expectation computed with the risk-neutral probability. Let us assume that, under the risk-neutral probability Q, the process of infinitesimal rate satisfies the Cox-Ingersoll-Ross process :

$$dr_t = -\tilde{k}[r_t - \tilde{\theta}]dt + (\tilde{\eta}^2 r_t)^{1/2} d\tilde{W}_t,$$

where $\sim$ is introduced for risk-neutral parameters. The formula for the zero-coupon bond price is directly deduced from Proposition 3. We get :

$$B(t, t+h) = \exp[-\tilde{a}^*(h) r_t - \tilde{b}^*(h)],$$

$$\begin{aligned}
\text{where : } \tilde{a}^*(h) &= \frac{2}{\tilde{\gamma}+\tilde{k}} - \frac{4\tilde{\gamma}}{\tilde{\gamma}+\tilde{k}} \frac{1}{(\tilde{\gamma}+\tilde{k})\exp(\tilde{\gamma}h)+\tilde{\gamma}-\tilde{k}}, \\
\tilde{b}^*(h) &= -\frac{\tilde{k}\tilde{\theta}(\tilde{\gamma}+\tilde{k})h}{\tilde{\eta}^2} + \frac{2\tilde{k}\tilde{\theta}}{\tilde{\eta}^2}\log[(\tilde{\gamma}+\tilde{k})\exp(\tilde{\gamma}h)+\tilde{\gamma}-\tilde{k}] \\
&- \frac{2\tilde{k}\tilde{\theta}}{\tilde{\eta}^2}\log(2\tilde{\gamma}),
\end{aligned}$$

where : $\tilde{\gamma} = \sqrt{\tilde{k}^2 + 2\tilde{\eta}^2}$.

This is the well-known Cox-Ingersoll-Ross model [Cox-Ingersoll-Ross (1985)] and an example of affine term structure model (ATSM), where all rates

$r(t, t+h)$ are affine functions of r_t, with sensitivity coefficients depending on time to maturity.

We know that :

$$r_t = \lim_{h \to 0} r(t, t+h) = \frac{\partial \tilde{a}^*(0)}{\partial h} r_t + \frac{\partial \tilde{b}^*(0)}{\partial h}.$$

It is easily checked that the limiting conditions $\frac{\partial \tilde{a}^*(0)}{\partial h} = 1, \frac{\partial \tilde{b}^*(0)}{\partial h} = 0$ are satisfied by the solution above for any choice of risk-neutral parameters.

ii) Historical Approach

In this approach the price of the zero-coupon bond is :

$$B(t, t+h) = \overset{P}{E_t} [\exp -(\int_t^{t+h} m_\tau d\tau)],$$

where $\overset{P}{E_t}$ denotes the conditional expectation computed with the historical probability and m_t denotes the stochastic discount factor (sdf). Let us assume that, under the historical probability P, the sdf (m_t) satisfies the CIR process :

$$dm_t = -k(m_t - \theta)dt + (\eta^2 m_t)^{1/2} dW_t.$$

The formula for the zero-coupon prices becomes :

$$B(t, t+h) = \exp[-a^*(h) m_t - b^*(h)],$$

where $a^*(h)$ and $b^*(h)$ are given in Proposition 3. The different rates $r(t, t+h)$ are now affine functions of m_t. In particular the short term rate r_t is also an affine function of m_t :
$r_t = \alpha_o + \alpha_1 m_t$, say. Thus the zero-coupon prices can also be written as :

$$B(t, t+h) = \exp\left[-a^*(h) \frac{r_t - \alpha_o}{\alpha_1} - b^*(h)\right],$$

and are exponential affine functions of r_t. We deduce that both historical and risk-neutral approaches with CIR processes provide identical results, whenever the historical and risk-neutral parameters are chosen appropriately.

2.2.2 Derivative Pricing in Stochastic Volatility Model

Let us consider a standard Black-Scholes model with stochastic drift and volatility independent of (W_t) for the dynamics of the asset price :

$$d \log S_t = \mu_t dt + \sigma_t dW_t, \qquad \text{(say)}. \tag{2.2}$$

Conditional on drift and volatility trajectories, the distribution of $\log S_{t+h}$ given $\log S_t$ is Gaussian with mean $\log S_t + \int_t^{t+h} \mu_\tau d\tau$, and variance : $\int_t^{t+h} \sigma_\tau^2 d\tau$. In particular the conditional Laplace transform of $\log S_{t+h}$ is :

$$\begin{aligned} & E[\exp(u \log S_{t+h}) | \underline{S_t}, (\mu_t), (\sigma_t^2)] \\ = \;& \exp[u \log S_t + u \int_t^{t+h} \mu_\tau d\tau + \frac{u^2}{2} \int_t^{t+h} \sigma_\tau^2 d\tau]. \end{aligned}$$

Let us now assume that the drift includes a risk premium which is an affine function of the volatility :

$$\mu_t = \alpha_o + \alpha_1 \sigma_t^2. \tag{2.3}$$

Then the conditional Laplace transform becomes an exponential affine function of the integrated volatility.

$$\begin{aligned} & E[\exp(u \log S_{t+h}) | \underline{S_t}, (\mu_t), (\sigma_t^2)] \\ = \;& \exp[u \log S_t + u\alpha_o h + (u\alpha_1 + \frac{u^2}{2}) \int_t^{t+h} \sigma_\tau^2 d\tau]. \end{aligned}$$

The Laplace transform above is computed taking into account the future paths of drift and volatility. Usually it is assumed that the (investor's) available information includes only the current and lagged values of S and σ^2. Thus the Laplace transform conditional on this restricted information set is :

$$\begin{aligned} \Phi_{t,h}(u) &= E[\exp(u \log S_{t+h}) | \underline{S_t}, \underline{\mu_t}, \underline{\sigma_t^2}] \\ &= E[\exp[u \log S_t + u\alpha_o h + (u\alpha_1 + \frac{u^2}{2}) \int_t^{t+h} \sigma_\tau^2 d\tau] | \underline{S_t}, \underline{\sigma_t^2}] \\ &= \exp(u \log S_t + u\alpha_o h) \psi_{t,h}^*(-u\alpha_1 - \frac{u^2}{2}), \end{aligned}$$

if (σ_t^2) follows a CIR process.

If the riskfree rate is set equal to zero and the dynamics above are under the risk neutral probability, $\Phi_{t,h}(u)$ is simply the price at t of a derivative paying $\exp(u \log S_{t+h})$ at $t+h$. The prices are defined for any value of u, including possibly complex values, and admits closed form expression. Then the price of a more standard derivative such as for instance a European call written on S_{t+h} is easily deduced by inverting appropriately the Fourier transform (that is the Laplace transform with pure imaginary argument u), which admits a closed-form expression. These closed form expressions of derivative prices in CIR stochastic volatility models have been first derived in Heston (1993), Ball, Roma (1994) [see also Pearson, Sun (1994) for the introduction of a risk premium in the drift of the price equation and Duffie, Pan, Singleton (2000) for a general presentation of transform analysis].

Financial applications to term structure and to derivative pricing in stochastic volatility model show that the expressions of the prices are directly related to the Laplace transform of an integrated latent factor, which corresponds to either the short term interest rate, the sdf, or the latent stochastic volatility. The existence of a closed form expression of the conditional Laplace transform of the integrated CIR process explains the closed form expressions of the prices derived in the CIR framework.

3 The Wishart Process

The Wishart process is the multivariate extension of the Cox-Ingersoll-Ross process. Like the CIR it can be defined in various ways. In Section 3.1, we first use the interpretation of the Wishart process as sum of matrix squares of Ornstein- Uhlenbeck processes to derive the expression of the conditional Laplace transform of this process and of its conditional pdf. Then we write a matrix diffusion system satisfied by the Wishart process. Finally we derive the closed form expression of the conditional Laplace transform of the integrated Wishart process. The financial applications are presented in Section 3.2. We first explain why some linear combinations of the elements of a Wishart process can be scalar positive processes. Then we discuss Wishart quadratic term structure (QTSM) models, derivative pricing in multivariate stochastic volatility model and structural models for credit risk.

3.1 Construction of the Wishart Process

The simplest way to derive the distribution of a Wishart process is to start from multivariate Ornstein-Uhlenbeck processes. By considering the sum of matrix squares of independent Ornstein-Uhlenbeck processes with identical dynamics, we derive the conditional Laplace transform of the (multivariate) Wishart process with integer degree of freedom. Then this process is extended to fractional degree of freedom.

Let us introduce K independent n-dimensional Ornstein-Uhlenbeck processes :

$$dx_{k,t} = Ax_{k,t}dt + QdW_{k,t}, k = 1, \dots, K,$$

where A, Q are (n, n) matrices with Q invertible, and let us consider the matrix process :

$$Y_t = \sum_{k=1}^{K} x_{k,t}x'_{k,t}.$$

By construction this process is such that Y_t is a (n,n) (stochastic) symmetric positive semi-definite matrix. Moreover the matrix Y_t is positive definite whenever $K \geq n$.

Then it is useful to define the conditional Laplace transform of the elements of Y_{t+h} given the elements of Y_t under a form appropriate for matrix processes. For this purpose let us remark that for any symmetric matrix Γ we get :

$$\begin{aligned} Tr(\Gamma Y) &= \sum_{i=1}^{n}(\Gamma Y)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}Y_{ji} \\ &= \sum_{i=1}^{n}\gamma_{ii}Y_{ii} + 2\sum_{i<j}\gamma_{ij}Y_{ij}. \end{aligned}$$

Thus any linear combination of the elements of a symmetric matrix Y can be written as $Tr(\Gamma Y)$, and the conditional Laplace transform can be defined as :

$$\psi_{t,h}(\Gamma) = E[\exp Tr(\Gamma Y_{t+h})|Y_t],$$

where Γ is a symmetric matrix. The closed form expression of this function is given below.

Proposition 4 : The process (Y_t) is a Markov process of order 1, with conditional Laplace transform :

$$\begin{aligned} \psi_{t,h}(\Gamma) &= \frac{\exp Tr[M(h)'\Gamma(Id - 2\Sigma(h)\Gamma)^{-1}M(h)Y_t]}{(det[Id - 2\Sigma(h)\Gamma])^{K/2}}, \\ \text{where : } M(h) &= \exp(Ah), \\ \Sigma(h) &= \int_0^h \exp(As)QQ'[\exp(As)]'ds. \end{aligned}$$

Proof : See Appendix 3.3.

The matrices $M(h), \Sigma(h)$ are easily interpreted in terms of Ornstein-Uhlenbeck processes. Indeed the time discretized Ornstein-Uhlenbeck process is simply a Gaussian Vector autoregressive process of order 1 [VAR (1)]. At horizon h its conditional mean is : $E(x_{k,t+h}|x_{k,t}) = M(h)x_{k,t}$ and its conditional variance is : $V(x_{k,t+h}|x_{k,t}) = \Sigma(h)$.

3.2 Distribution of the Wishart Process

The conditional Laplace transform of Proposition 4 can be recognized as the Laplace transform of a noncentered Wishart distribution. It is known that this distribution exists also for noninteger degree of freedom K [see e.g. Muirhead (1982)].

Definition 1 : An autoregressive Wishart process (WAR) is a Markov process of (n, n) stochastic symmetric positive definite matrices with conditional Laplace transform :

$$\psi_{t,h}(\Gamma) = \frac{\exp Tr[M(h)'\Gamma(Id - 2\Sigma(h)\Gamma)^{-1}M(h)Y_t]}{[det(Id - 2\Sigma(h)\Gamma)]^{K/2}},$$

where : $M(h) = \exp(Ah), \Sigma(h) = \int_0^h \exp(As)QQ'[\exp(As)]'ds$ and $K > 0$. Its conditional distribution is a noncentered Wishart distribution.

In particular the transition pdf at horizon h admits a closed form expression which involves a series expansion [see Anderson, Girshick (1944), Muirhead (1982), p442].

$$
\begin{aligned}
f(Y_{t+h}|Y_t) &= \frac{1}{2^{Kn/2}}\frac{1}{\Gamma_n(K/2)}(det\ \Sigma(h))^{-K/2} \\
& (det\, Y_{t+h})^{(K-n-1)/2}\ \exp\{-\frac{1}{2}Tr[\Sigma(h)^{-1}(Y_{t+h}+M(h)Y_tM(h)')]\} \\
& {}_oF_1(K/2, \frac{1}{4}M(h)Y_tM(h)'Y_{t+h}), \\
\text{where}:\ \ \Gamma_n(K/2) &= \int_{A>>0} \exp[Tr(-A)](detA)^{(K-n-1)/2}
\end{aligned}
$$

is the multidimensional gamma function and ${}_oF_1$ is the hypergeometric function of matrix arguments. The hypergeometric function admits a series expansion which involves the so-called zonal polynomials. These polynomials have no closed form expressions, but can be computed recursively [James (1968), Muirhead (1982), Chapter 7.2].

The noncentered Wishart transition for the volatility process differs from Wishart specifications usually introduced in a Bayesian analysis of multivariate volatility. In this literature the stochastic volatility is assumed to follow an inverse Wishart distribution with a scale parameter function of lagged volatility [see Philipov, Glickman (2004) for a comparison of such an approach with the DCC model]. This specification is appropriate for Bayesian updating, but does not provide an affine process and simple prediction formulas for integrated volatility process.

3.3 Diffusion Representation of the Wishart Process

The link between the Wishart and Ornstein-Uhlenbeck processes can also be used to guess the diffusion representation of a Wishart process. Let us consider the case $K=1$, where : $Y_t = x_{1,t}x'_{1,t}$, with : $dx_{1,t} = Ax_{1,t}dt + QdW_{1,t}$. We get :

$$
\begin{aligned}
dY_t &= Y_{t+dt} - Y_t \\
&= x_{1,t+dt}x'_{1,t+dt} - x_{1,t}x'_{1,t} \\
&\simeq (x_{1,t} + Ax_{1,t}dt + QdW_{1,t})(x_{1,t} + Ax_{1,t}dt + QdW_{1,t})' - x_{1,t}x'_{1,t} \\
&\simeq x_{1,t}x'_{1,t}A'dt + Ax_{1,t}x'_{1,t}dt + QE(dW_{1,t}dW'_{1,t})Q' \\
&+ x_{1,t}dW'_{1,t}Q' + QdW_{1,t}x'_{1,t} \\
&= [x_{1,t}x'_{1,t}A' + Ax_{1,t}x'_{1,t} + QQ']dt \\
&+ x_{1,t}dW'_{1,t}Q' + QdW_{1,t}x'_{1,t},
\end{aligned}
$$

by keeping only the relevant terms of the expansion (this is the well-known Ito's formula). However this expression is difficult to use, especially since the process (Y_t) does not clearly appear in the stochastic part of the right hand side. A better representation involves a matrix Brownian motion (W_t), that is a (n, n) stochastic matrix whose elements are independent standard scalar Brownian motions.

Proposition 5 : The Wishart process satisfies the (matrix) diffusion system :

$$
\begin{aligned}
dY_t &= (KQQ' + Y_tA' + AY_t)dt \\
&+ Y_t^{1/2}dW_tQ + Q'dW_tY_t^{1/2},
\end{aligned}
$$

where (W_t) is a (n, n) standard Brownian motion.

Proof : It can be checked that the drift and volatility of the process satisfying the diffusion system above are identical to the drift and volatility deduced from the conditional Laplace transform of Definition 1 (see Gourieroux, Sufana (2004)).

QED

This diffusion representation is easy to use, for instance to compute the first and second order moments of the stochastic process. The drift is :

$$E_t dY_t = (KQQ' + Y_t A' + AY_t)dt,$$

and is an affine function of Y_t.

Of course it is more difficult to represent the volatility matrix of dY_t (or of the vector obtained by stacking the different elements of dY_t), which has generally a large dimension $\left(\frac{n(n+1)}{2}, \frac{n(n+1)}{2}\right)$. Fortunately it is equivalent to know dY_t or to know the square of the norm [2] associated with dY_t : $\alpha \to \alpha' dY_t \alpha$. We get :

$$\begin{aligned}
& Cov_t(\alpha' dY_t \alpha, \beta' dY_t \beta) \\
= \quad & Cov_t(\alpha' Y_t^{1/2} dW_t Q\alpha + \alpha' Q' dW_t' Y_t^{1/2} \alpha, \beta' Y_t^{1/2} dW_t Q\beta + \beta' Q' dW_t' Y_t^{1/2} \beta) \\
= \quad & E_t[(\alpha' Y_t^{1/2} dW_t Q\alpha + \alpha' Q' dW_t' Y_t^{1/2} \alpha)(\beta' Q' dW_t' Y_t^{1/2} \beta + \beta' Y_t^{1/2} dW_t Q\beta).
\end{aligned}$$

By noting that for any n-dimensional vectors u, v :

$$\begin{aligned}
E_t(dW_t uv' dW_t') &= E_t(dW_t' uv' dW_t) = v'u \; Id \; dt, \\
E_t[dW_t uv' dW_t] &= E_t(dW_t' uv' dW_t'] = vu' dt,
\end{aligned}$$

we deduce that :

$$\begin{aligned}
& Cov_t(\alpha' dY_t \alpha, \beta' dY_t \beta) \\
= \quad & [4\alpha' Y_t \beta \; \alpha' Q' Q\beta] dt,
\end{aligned}$$

for any n-dimensional vectors α, β. Thus all volatilities-covolatilities between elements of dY_t are also affine functions of Y_t.

Since the drift and volatility are both affine functions of Y_t [or equivalently the conditional Laplace transform is an exponential affine function of Y_t], the Wishart process is an affine process. However it is not a member of the

[2] For any pair of vectors α, β we get : $\alpha' dY_t \beta = \frac{1}{2}[(\alpha+\beta)' dY_t (\alpha+\beta) - \alpha' dY_t \alpha - \beta' dY_t \beta]$. Therefore it is equivalent to know the inner product, or the norm associated with dY_t.

standard class of affine processes introduced by Duffie, Kan (1996) and later studied in details by Dai, Singleton (2000).

Remark 1 : The Wishart dynamics can be extended by considering the processes which satisfy the differential system :

$$\begin{aligned} dY_t &= (\Omega\Omega' + Y_t A' + AY_t)dt \\ &+ Y_t^{1/2} dW_t Q + Q' dW_t' Y_t^{1/2}, \end{aligned}$$

in which Ω is a (n, n) invertible matrix and the constraint $\Omega\Omega' = KQQ'$ is not imposed. Indeed the right hand side of the equation is a symmetric matrix. Moreover we know that :

$$\begin{aligned} E_t(\alpha' Y_t \alpha) &= (\alpha'\Omega\Omega'\alpha + \alpha' Y_t A'\alpha + \alpha' A Y_t \alpha)dt, \\ V_t(\alpha' Y_t \alpha) &= (4\alpha' Y_t \alpha\ \alpha' Q' Q \alpha)dt. \end{aligned}$$

It the matrix (Y_t) reaches the boundary of the set of symmetric positive semi-definite matrices, there exists a vector α such that $\alpha' Y_t \alpha = 0 \Longleftrightarrow Y_t \alpha = 0$ (since Y_t is positive semi-definite). At this date the process $\alpha' Y_t \alpha$ becomes locally deterministic, since $V_t(\alpha' Y_t \alpha) = 0$, with a positive drift $E_t(\alpha' Y_t \alpha) = \alpha'\Omega\Omega'\alpha dt > 0$. Thus the process is rejected towards positivity. This ensures that the solution of the extended Wishart equation corresponds to a process of symmetric positive matrices.

3.4 Integrated Wishart Process

In the Cox-Ingersoll-Ross framework, the most important result for financial applications is the closed-form expression of the conditional Laplace transform of the integrated process. A similar result can be derived in the Wishart framework.

Proposition 6 : The conditional Laplace transform of the integrated Wishart process can be written as :

$$\begin{aligned}
\psi^*_{t,h}(\Gamma) &= E_t(\exp Tr[\Gamma \int_t^{t+h} Y_\tau d\tau]) \\
&= \exp[Tr[A^*(h,\Gamma)Y_t] + b^*(h,\Gamma)], \\
\text{where : } \frac{\partial A^*}{\partial h}(h,\Gamma) &= \Gamma + A(h,\Gamma)A + A'A(h,\Gamma) + 2A(h,\Gamma)Q'QA(h,\Gamma), \\
\frac{\partial b^*}{\partial h}(h,\Gamma) &= KTr[A(h,\Gamma)QQ'],
\end{aligned}$$

with initial conditions : $A^*(0,\Gamma) = 0, b^*(0,\Gamma) = 0$. The solution of this (matrix) Riccati differential systems is :

$$A^*(h,\Gamma) = A^*(\Gamma) + \exp[(A + 2Q'QA^*(\Gamma))h]'$$

$$\{-(A^*(\Gamma))^{-1} + 2\int_0^h \exp[A + 2Q'QA^*(h,\Gamma)u]Q'Q\exp[A + 2Q'QA^*(h,\Gamma)u]'du\}$$

$$\exp[[A + 2Q'QA^*(\Gamma)]h],$$

where $A^*(\Gamma)$ satisfies :

$$A'A^*(\Gamma) + A^*(\Gamma)A + 2A^*(\Gamma)Q'QA^*(\Gamma) + \Gamma = 0.$$

Proof : See Appendix 1.3 and Appendix 2.2.

In general the Riccati differential systems do not admit closed form solution. In the Wishart framework these equations can be solved partly [see Grasselli, Tebaldi (2004) for a discussion of solvable ATSM].

3.5 Financial Applications

We have seen in Section 2.2 that closed form solutions for term structure and derivative pricing were easily obtained in models with a single factor following a Cox-Ingersoll-Ross process. The aim of this section is to extend these results to a multifactor Wishart framework. These extensions provide :

i) a general presentation of the so-called Wishart quadratic term structure models (WQTSM);

ii) the generalization of Heston's stochastic volatility model to the multi-asset framework;

iii) a structural model of credit risk which can be applied to stochastic assets and liabilities, and to any number of corporations.

3.5.1 Positive Affine Transformations of the Wishart Process.

We have already noted that any linear combination of the elements of a Wishart process (Y_t) can be written as $Tr(CY_t)$, where C is a symmetric matrix. The property below is the keypoint for financial applications.

Lemma 1 : If (Y_t) is a Wishart process and C is a symmetric positive definite matrix, the scalar process $Tr(CY_t)$ is positive.

Proof : Indeed the matrix C admits a spectral decomposition $C = \sum_{i=1}^{n} \lambda_i u_i u_i'$, where $\lambda_i, i = 1, \ldots, n$, are the positive eigenvalues and $u_i, i = 1, \ldots, n$, are the associated eigenvectors. Then we get :

$$\begin{aligned}
Tr(CY_t) &= Tr[\sum_{i=1}^{n} \lambda_i u_i u_i' Y_t] \\
&= \sum_{i=1}^{n} \lambda_i Tr(u_i u_i' Y_t) \\
&= \sum_{i=1}^{n} \lambda_i u_i' Y_t u_i \text{ (since we can commute within the trace)} \\
&> 0,
\end{aligned}$$

since (Y_t) is a symmetric positive definite matrix. QED.

Therefore the Wishart process can be used to define affine factor models for positive variables. Typically if $z_{jt}, j = 1, \ldots, J$ are nonnegative processes, a joint affine factor representation of these processes is.

$$z_{jt} = c_j + Tr(C_j Y_t), j = 1, \ldots, J,$$

where c_j are nonnegative scalars and C_j are symmetric positive semi-definite matrices. This representation is used below for different types of nonnegative processes such as interest rates, default intensities, volatilities, volatility-in-mean effects and risk premia.

3.5.2 Joint Analysis of Term Structures for T-bonds and Corporate Bonds.

The term structure model of Section 2.2.1 is easily extended to consider jointly T-bond and corporate bonds. In the basic model with zero recovery rate in case of default, the prices of T-bonds and corporate bonds can be written as :

$$\begin{aligned} B(t, t+h) &= \overset{Q}{E}_t\, [\exp(-\int_t^{t+h} r_\tau d\tau)], \\ B_i(t, t+h) &= \overset{Q}{E}_t\, [\exp[-\int_t^{t+h} r_\tau d\tau - \int_t^{t+h} \lambda_{i,\tau} d\tau]), i = 1, \ldots, I, \end{aligned}$$

where i is the firm index, r_t the infinitesimal riskfree rate, $\lambda_{i,t}$ the infinitesimal default intensity for firm i, and the conditional expectations are computed under the risk-neutral distribution.

Let us now assume that :

$$r_t = c + Tr(CY_t), \lambda_{i,t} = d_i + Tr(D_i Y_t), i = 1, \ldots, I,$$

where matrices $C, D_i, i = 1, \ldots, n$ are symmetric positive semi-definite, the scalars $c, d_i, i = 1, \ldots, n$ are nonnegative and (Y_t) is a Wishart process under the risk-neutral distribution. Then we get :

$$B(t, t+h) = \exp(-hc)\, \overset{Q}{E}_t\, (\exp -Tr[C \int_t^{t+h} Y_\tau d\tau]),$$

$$B_i(t, t+h) = \exp(-h(c + d_i))\, \overset{Q}{E}_t\, [\exp -Tr((C + D_i) \int_t^{t+h} Y_\tau d\tau)], i = 1, \ldots, I.$$

As in the one factor case considered in Section 2.2.1, the prices of T-bonds and corporate bonds depend on the conditional Laplace transform of the integrated volatility of the Wishart process, and admit closed form expression.

The model above is rather flexible to represent various term structure patterns, since :

i) it allows for a large number of factors, equal to $n(n+1)/2$.

ii) Some factors can be specific of a given firm, common to several firms, or can appear in both infinitesimal riskfree rate and default intensities, according to the zero elements of the matrices of sensitivity coefficients $C, D_i, i = 1, \ldots, I$ (more precisely of their ranks and null spaces).

iii) It allows for a large variety of patterns for the term structures, which increases with the number of factors, that is with n.

The model includes as special cases the quadratic term structure models introduced in the literature for riskfree bonds [Constantinides (1992), Leippold, Wu (2002), Ahn, Dittmar, Gallant (2002), Gourieroux, Sufana (2003)], and the so-called reduced form model for credit risk, in which r_t and $\lambda_{i,t}$ are linear combinations of independent CIR processes [Lando (1998), Duffie, Singleton (1999)].

3.5.3 Derivative Pricing in Multivariate Stochastic Volatility Model.

The closed form expression of derivative prices obtained in the CIR stochastic volatility model can be extended to the multiasset framework in a similar way. Let us assume a zero riskfree rate and consider two risky assets with prices $S_{i,t}, i = 1, 2$ such that :

$$\begin{bmatrix} d\log \quad S_{1,t} \\ \\ d\log \quad S_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 + Tr(C_1 Y_t) \\ \\ c_2 + Tr(C_2 Y_t) \end{bmatrix} dt + Y_t^{1/2} dW_t,$$

where the stochastic volatility matrix Y_t follows a (2,2) Wishart process independent of the bivariate Brownian motion (W_t) (under the risk-neutral probability). The stochastic differential system above can be integrated conditional on $S_{1,t}, S_{2,t}$ and on the volatility-covolatility path.

The conditional distribution of $\log S_{1,t+h}, \log S_{2,t+h}$ given $S_{1,t}, S_{2,t}, (Y_t)$ is a Gaussian distribution with mean :

$$\begin{pmatrix} \log S_{1,t} \\ \\ \log S_{2,t} \end{pmatrix} + \begin{pmatrix} c_1 \\ \\ c_2 \end{pmatrix} h + \begin{bmatrix} Tr(C_1 \int_t^{t+h} Y_\tau d\tau) \\ \\ Tr(C_2 \int_t^{t+h} Y_\tau d\tau) \end{bmatrix},$$

and variance-covariance matrix :

$$\int_t^{t+h} Y_\tau d\tau.$$

Therefore the conditional Laplace transform of $\log S_{1,t+h}, \log S_{2,t+h}$, given $S_{1,t}, S_{2,t}$ and the volatility path is :

$$\begin{aligned} & E[\exp(u_1 \log S_{1,t+h} + u_2 \log S_{2,t+h}) | S_{1,t}, S_{2,t}, (Y_t)] \\ = & \exp\{(u_1 \log S_{1,t} + u_2 \log S_{2,t} + h(u_1 c_1 + u_2 c_2) \\ + & Tr\left[\left[u_1 C_1 + u_2 C_2 + \frac{1}{2}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(u_1, u_2)\right] \int_t^{t+h} Y_\tau d\tau\right]\right\}. \end{aligned}$$

The conditional Laplace transform of $\log S_{1,t+h}, \log S_{2,t+h}$ given $S_{1,t}, S_{2,t}$ and the current factor value only is directly deduced from the conditional Laplace transform of the integrated Wishart process [see Proposition 6]. We get :

$$\begin{aligned} & E[\exp(u_1 \log S_{1,t+h} + u_2 \log S_{2,t+h}) | S_{1,t}, S_{2,t}, Y_t] \\ = & \exp[u_1 \log S_{1,t} + u_2 \log S_{2,t} + h(u_1 c_1 + u_2 c_2)] \\ & \psi^*_{t,h}\left[u_1 C_1 + u_2 C_2 + \frac{1}{2}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}(u_1, u_2)\right]. \end{aligned}$$

3.5.4 Structural Credit Risk Model

The extension of Heston's model to the multiasset framework can be used to extend the structural model of credit risk initially proposed by Merton (1974). In the structural model, default occurs when the asset value of the firm is below its total liability. The basic Merton's model assumes :
i) a predetermined time to default;
ii) an asset value following a geometric Brownian motion;

iii) a predetermined evolution of the debt which can be assumed constant [Leland, Toft (1992)], or calibrated to some market or firm specific observables as in the KMV approach.

The use of Wishart process allows for a stochastic debt and stochastic volatility models, while keeping closed form solutions for derivative prices. Moreover the analysis can be performed in a similar way for any number of firms [see Gourieroux, Sufana (2004)]. The model for one firm is :

$$\begin{bmatrix} d\log A_t \\ \\ d\log L_t \end{bmatrix} = \begin{bmatrix} \mu_A + Tr(C_A Y_t) \\ \\ \mu_L + Tr(C_L Y_t) \end{bmatrix} dt + Y_t^{1/2} dW_t,$$

and the derivative pricing formula derived for the multiasset framework can be applied directly.

In this specification the terms $Tr(C_A Y_t)$ and $Tr(C_L Y_t)$ included in the drift don't admit interpretation as risk premia since the firm's asset value and liability are not directly traded on a market. However such effects can be expected for the following reason. Let us for a while consider Merton's model with a constant debt level, but stochastic volatility. Default can occur at least for two very different reasons :
i) the asset value has a decreasing trend, and reaches the liability level.
ii) the asset value has a constant mean, but there is a large volatility increase, which implies a crossing of the liability level .
In the first situation the decreasing trend is likely due to an economic fundamental of the firm. But in the second situation, the fundamental is stable, whereas asset value will pass below debt level during a transitory period only. In the second situation the medium term rating of the firm allows it to increase its debt in order to increase its investments and jointly the asset value. This explains why correlated evolutions of asset value and liability are observed when risk increases. This is exactly what is assumed in the joint specification above of asset value and liability. Finally note that this model is a solution to one of the unresolved issues in modelling credit risk listed by Turnbull (2004).

4 From General to Specific

We have seen in Section 3 the importance and flexibility of the Wishart process for structural financial applications such as derivative pricing or term structure analysis. It is also necessary to check if the model is appropriate for modelling the dynamics of latent or observable volatility-covolatility matrices. For this purpose it has clearly to be compared to the competing multivariate ARCH models proposed in the literature[3], especially to the Dynamic Conditional Correlation (DCC) model introduced by Engle, Sheppard (2001), Engle (2002)a. Note that this comparison concerns the dynamics of volatility-covolatility matrices, not the dynamics of asset returns. For the asset return dynamics, it has been seen in Section 3.5.3 that in the multivariate stochastic volatility model the volatility is not assumed to be a deterministic function of past return innovations (as in ARCH approach), but has its own dynamics.

First we have to consider the unconstrained Wishart process, where parameters K, M, Σ are not constrained a priori (except the symmetry and positivity of Σ and the positivity of K), and to discuss the dynamics of volatilities-covolatilities that can be reproduced. In this respect the Wishart process appears very flexible. However this flexibility could be due to the large number of parameters, which are involved. In a second step we explain how to solve for the curse of dimensionality, while keeping the flexibility required by the data. This will be done by means of factor representations. Finally we discuss statistical inference for observed or latent, constrained or unconstrained Wishart processes.

4.1 Dynamic Properties of the Wishart Process.

We list below some dynamic properties which can be reproduced with a Wishart process and refer to Gourieroux, Jasiak, Sufana (2004) for more details. In each case, we give examples of other dynamic models, which cannot reproduce these features. Before discussing these features, it has to be mentioned that :

i) A Wishart process can be defined in discrete time.

The discrete time Wishart process is a Markov process with conditional Laplace transform :

[3] To facilitate the comparison, the main stochastic volatility models proposed in the literature are reviewed in Appendix 4.

$$\psi_{t,1}(\Gamma) = \frac{\exp Tr[\Gamma(Id - 2\Sigma\Gamma)^{-1} M Y_t M']}{[\det(Id - 2\Sigma\Gamma)]^{K/2}}.$$

The dynamics of this process is characterized by parameters K, M, Σ, where $K > 0$, and Σ is symmetric positive semi-definite. The "latent" autoregressive matrix is not constrained.

ii) Any time discretized Wishart diffusion process is a discrete time Wishart.

The time discretized Wishart diffusion processes are obtained when the "latent" autoregressive matrix can be written as : $M = \exp A$. This is a rather restrictive constraint, which for instance implies that M is invertible and does not allow for recursive systems or factor representations [see Section 4.2]. To summarize there exist many more discrete time Wishart processes than continuous time Wishart processes.

iii) In discrete time, Wishart processes can be defined with any autoregressive order.

For instance a Wishart process of autoregressive order p (denoted WAR(p)) admits a conditional Laplace transform :

$$\psi_{t,1}(\Gamma) = \frac{\exp Tr[\Gamma(Id - 2\Sigma\Gamma)^{-1} \sum_{j=0}^{p-1} M_j Y_{t-j} M_j']}{[det\ (Id - 2\Sigma\Gamma)]^{K/2}}.$$

Thus the assumption of Markov process of order one can be weakened and will not be discussed later on.

Some dynamic properties of the Wishart process are the following ones :

i) A Wishart process is not necessarily time reversible, that is its dynamic properties can differ in the usual and reverse times. This possibility is not surprising since the continuous time Wishart process [resp. discrete time] can be constructed from multivariate Ornstein-Uhlenbeck processes [resp. Gaussian VAR(1)], which are not reversible in general. The multivariate ARCH models are also generally not reversible. However the specifications which involve linear combinations of independent one-dimensional CIR processes (as the Duffie-Kan model [Duffie, Kan (1996)]) imply the time reversibility.

ii) The Wishart process allows for negative dynamic dependence between volatilities. As above this feature cannot be reproduced by the basic Duffie-Kan model, but can be by several multivariate ARCH models.

iii) The Wishart process is compatible with dynamic conditional correlations and these dynamic correlations can evoluate "independently" of the basic volatilities. The need for this feature justified the introduction of DCC model in the multivariate ARCH literature and explained why the other multivariate ARCH models are less interesting than the DCC. For instance the CCC model of Bollerslev (1987), (1990) is simple to estimate, but assumes unrealistic constant correlations [Tse (2000), Engle, Sheppard (2001)].

iv) In financial analysis it is important to consider the spectral decomposition of a volatility-covolatility matrix. Loosely speaking the eigenvector corresponding to the largest eigenvalue provides the portfolio allocation with the largest risk. Symmetrically the eigenvector corresponding to the smallest eigenvalue provides the portfolio allocation with the smallest risk. In particular if the smallest eigenvalue is close to zero, we get an almost riskfree portfolio. This portfolio is the key tool for the so-called arbitragists.

The Wishart process allows for both stochastic eigenvectors and stochastic eigenvalues, and the evolutions of eigenvectors and eigenvalues can be weakly dependent. Typically it is easy to reproduce situations in which the largest eigenvalue is very high whereas the smallest one is close to zero (This requires more than one factor, that is $n \geq 2$) [see Gourieroux, Jasiak, Sufana (2004)]. In such a case two very different types of risk have to be hedged, that are the increasing risk due to the largest eigenvalue and the risk created by the arbitragist strategies trying to profit of the quasi-arbitrage.

Note that the Duffie-Kan model assumes eigenvectors constant in time, and that the standard ARCH factor models or stochastic volatility factor models [Jacquier, Marcus (2001)] suppose a restrictive singular value decomposition of the volatility matrix.

v) The Wishart process allows for simple nonlinear prediction formulas at any horizon. This is the direct consequence of the closed form of the conditional Laplace transform of the Wishart. Usually the ARCH type models are presented as semi-parametric models with simple linear prediction formulas. It is easy to compute multistep ahead forecast $E(Y_{t+h}|Y_t)$, and sometimes the second order conditional moments. However linear prediction formulas are not so useful for financial applications. First the existence of derivative assets with nonlinear payoffs require nonlinear prediction formulas. Second for risk management the regulator ask for the computation of conditional quantiles. The distributional assumption is generally introduced in ARCH models by means of the conditionally standardized error term, but nonlinear prediction requires simulations of the future volatility path.

vi) In a Wishart process representing the asset return volatilities and covolatilities, the volatility of asset 1 can depend on the lagged volatility of asset 2. This is clearly not the case in the DCC models, which considers separately the volatilities per asset, before considering the dynamics of correlations, which can involve several assets. Loosely speaking the DCC approach uses different information sets for the volatility analysis and for the correlation analysis.

vii) The Wishart process ensures symmetric positive definite matrices. If the symmetry constraint is satisfied by the major part of multivariate ARCH models, the positivity is realized by a limited number of them.

4.2 Factor Representation of the Wishart Process

The factor representations of linear dynamic models are well-defined and frequently used in both applied and academic literature [see e.g. Chapter 9 in Gourieroux, Jasiak (2001) and the references therein]. More precisely let us consider a n-dimensional Gaussian vector autoregressive process :

$$X_{t+1} = MX_t + \Sigma^{1/2}\varepsilon_{t+1},$$

where (ε_t) is a standard Gaussian white noise.

Let us assume that the rank of the autoregressive matrix M is strictly less than n, for instance equal to $L = 2 < n$. Then matrix M can be decomposed as :

$$M = \beta_1\alpha_1' + \beta_2\alpha_2',$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are n-dimensional vectors such that β_1 and β_2 [resp. α_1 and α_2] are linearly independent.

Then the linear autoregressive model can be written as :

$$\begin{aligned} X_{t+1} &= \beta_1\alpha_1' X_t + \beta_2\alpha_2' X_t + \Sigma^{1/2}\varepsilon_{t+1} \\ &= \beta_1 F_{1,t} + \beta_2 F_{2,t} + \Sigma^{1/2}\varepsilon_{t+1}, \\ \text{where : } F_{j,t} &= \alpha_j' X_t, j = 1, 2. \end{aligned}$$

All past information is represented by a small number of summary statistics, that are the two factor values : $F_{j,t} = \alpha_j' X_t, j = 1, 2$. Moreover for any vector γ orthogonal to β_1 and β_2, we get :

$$\begin{aligned} \gamma' X_{t+1} &= \gamma'(\beta_1 F_{1,t} + \beta_2 F_{2,t}) + \gamma' \Sigma^{1/2} \varepsilon_{t+1} \\ &= \gamma' \Sigma^{1/2} \varepsilon_{t+1}. \end{aligned}$$

Any linear combination $\gamma' X_{t+1}$ is a white noise.Thus in general the dynamic linear model can be rewriten into a simplified form. More precisely, let us assume that α_1 and α_2 are not orthogonal to β_1 and β_2. Then we can consider the transformed process :

$$\tilde{X}_{t+1} = \begin{pmatrix} \alpha_1' \\ \alpha_2' \\ \gamma_1' \\ \vdots \\ \gamma_{n-2}' \end{pmatrix} X_{t+1},$$

where $\gamma_1, \ldots \gamma_{n-2}$ are linearly independent vectors orthogonal to β_1 and β_2. The transformed process satisfies the linear dynamic model :

$$\tilde{X}_{t+1} = \begin{pmatrix} \tilde{M}_{11} & 0 \\ & \\ 0 & 0 \end{pmatrix} \tilde{X}_t + \tilde{\Sigma}^{1/2} \varepsilon_{t+1},$$

say, where $\tilde{M}_{11}$ is a (2,2) matrix. Thus, by considering the spectral decomposition (rank, null space and range) of matrix M, we can separate the directions independent of the past and the directions summarizing all past information. Such an analysis is the basis for all factor representations introduced in the literature, including the so-called factor ARCH representation [see Appendix 4]. It can also be noted that a model with "reduced rank" of the autoregressive matrix involves a much smaller number of parameters than the unconstrained linear dynamic model.

At this step it is important to mention that such a vector representation cannot be defined in continuous time, where the multivariate Ornstein-Uhlenbeck process is the analogue of the Gaussian VAR(1). Indeed for a time discretized Ornstein-Uhlenbeck process the autoregressive matrix is : $M = M(1) = \exp A$ and cannot admit zero eigenvalues. Equivalently the notion of white noise process does not exist in continuous time. Thus in this section we focus on discrete time Wishart processes for which the autoregressive matrix is not necessarily constrained.

Let us now consider a discrete time multivariate Wishart process and assume that the rank of M is equal to L. Then matrix M can be written as :

$$M = \beta\alpha',$$

where β and α are (n, L) matrices with rank L. Let us also introduce a matrix γ with dimension $(n, n-L)$ such that $\gamma'\beta = 0$. In general (α, γ) is an invertible matrix and it is equivalent to consider the initial volatility-covolatility matrix Y_t, or the transformed matrix :

$$\tilde{Y}_t = (\alpha, \gamma)' Y_t (\alpha, \gamma) = \begin{bmatrix} \alpha' Y_t \alpha & \alpha' Y_t \gamma \\ \\ \gamma' Y_t \alpha & \gamma' Y_t \gamma \end{bmatrix}.$$

When Y_t is the return volatility matrix of n financial assets, $\tilde{Y}_t$ is nothing else than the return volatility matrix of the n asset portfolios with allocations defined by the columns of matrix (α, γ).

By considering the expression of the conditional Laplace transform :

$$\psi_{t,1}(\Gamma) = \frac{\exp Tr[(Id - 2\Sigma\Gamma)^{-1}\beta\alpha' Y_t \alpha\beta']}{[det\ (Id - 2\Sigma\Gamma)]^{K/2}},$$

the following results are easily deduced [Gourieroux, Jasiak, Sufana (2004)].

i) The past information is summarized by the subvolatility matrix $\alpha' Y_t \alpha$. In nonlinear dynamic models, such sufficient summaries are called factors. Thus we get $L(L+1)/2$ factors when the rank of M is L.

ii) The process $(\alpha' Y_t \alpha)$ is also a Wishart process with reduced dimension L.

iii) The process $(\gamma' Y_t \gamma)$ is a white noise process.

The minimal expected number of factors depends on the application.

When we analyse stock returns, the basic equilibrium model [CAPM] says that the volatility of the market portfolio drives the volatilities and covolatilities of the different stocks. This is a one factor model $(L = 1)$. However in practice other volatility factors can appear, in particular to account for the different economic sectors represented in stock markets.

For applications to term structure, at least 3 factors have generally to be introduced to get flexible term structure patterns. They drive intuitively the dynamics of the level, slope and curvature of the term structure. A three factor model corresponds to $L = 2$.

These rather small expected numbers of factors are important since the expression of the transition density of the Wishart process is greatly simplified for $L = 1$ or $L = 2$ [see e.g. Anderson, Girshick (1944), Glejser (1976)].

4.3 Statistical Inference

4.3.1 Observable Processes

Before discussing estimation methods and testing procedures, it is necessary to define the observable processes. In this respect, three types of problems have to be distinguished.

i) **Discrete time observations of a Wishart process are available**

A typical example is the analysis of daily intraday realized volatility-covolatility matrices. For instance Gourieroux, Jasiak, Sufana (2004) consider a series of historical volatility matrices at horizon 5mn. More precisely, they consider the asset returns at 5mn inverval, $r_{t,n}$, say, where t denotes the day and n the index of the 5mn period within the day. For each day t, they compute the historical volatility at 5mn :

$$Y_t = \frac{1}{N}\sum_{n=1}^{N}(r_{t,n} - \frac{1}{N}\sum_{n=1}^{N} r_{t,n})(r_{t,n} - \frac{1}{N}\sum_{n=1}^{N} r_{t,n})',$$

where $N = 72$ is the number of 5mn spells within a trading day. A discrete time Wishart process is assumed and fitted for this time series.

ii) Discrete time observations of a Wishart process are available up to some parameters.

This situation arises in stochastic risk models when derivative prices are observed. Let us consider the example of T-bonds. Note that all coupon bonds can be seen as derivatives written on the short term interest rate. We have seen in Section 3.5.2 that the Wishart Quadratic Term Structure model is a special case of affine model. Therefore let us assume for instance $L = 2$ and, at any discrete date t, three observable yields $r(t, t + h_j), j = 1, 2, 3$, say. Since the model is affine, there exists an affine relationship between the yields and the factors $Y_{1,1,t}, Y_{1,2,t}, Y_{2,2,t}$:

$$\begin{bmatrix} r(t, t+h_1) \\ r(t, t+h_2) \\ r(t, t+h_3) \end{bmatrix} = \alpha(\theta) + \beta(\theta) \begin{pmatrix} Y_{1,1,t} \\ Y_{1,2,t} \\ Y_{2,2,t} \end{pmatrix}, \text{say,}$$

where θ denotes the parameter. Thus :

$$\begin{pmatrix} Y_{1,1,t} \\ Y_{1,2,t} \\ Y_{2,2,t} \end{pmatrix} = \beta(\theta)^{-1} \left[\begin{bmatrix} r(t, t+h_1) \\ r(t, t+h_2) \\ r(t, t+h_3) \end{bmatrix} - \alpha(\theta) \right],$$

are observed up to some parametrized transformation.

In this situation of partial observability it is important to check for the identifiable parameters.

iii) The Wishart process is partially observed through a measurement equation.

The situation arises in multivariate stochastic volatility model [see Section 3.5.3], when the returns of the underlying assets $\Delta \log S_{1,t}, \Delta \log S_{2,t}$, say, are observed, but the underlying stochastic volatilities or the derivative prices are not observed.

4.3.2 Estimation Methods

All situations described in the section above correspond to different types of dynamic factor models for which standard estimation methods can be applied, at least from a theoretical point of view. Such methods are the maximum likelihood method, simulation based methods as indirect inference and the generalized method of moment. From a practical point of view, these methods are more or less easy to implement and more or less efficient in finite samples. It is important to mention that, since the introduction of the CIR process in pricing models, the methods of moment appear as the most appropriate.

i) Maximum likelihood approach.

In the situation where the Wishart process is observed possibly up to some parameterized transformations the expression of the likelihood function can be derived. However it involves the series expansion in terms of the zonal polynomials and some parameter can be involved in the support of the distribution in case of a parameterized transformation [Gourieroux, Monfort (2005)]. Practically it can be implemented when the rank of M is equal to 1, or 2, that is for models with at most 2 dynamic factors. But the standard asymptotic theory, that are the rates of convergence and the type of asymptotic distributions, are modified when the support is parameterized.

ii) Simulation Based Methods

Due to their interpretations as sum of squares of Gaussian VAR(1) processes, the discrete time WAR processes with integer degree of freedom are easy to simulate. However the simulation methods are much more difficult to implement for fractional degree of freedom whenever the rank of L is larger than two [see for a survey on simulation methods for noncentered Wishart distributions].

iii) Generalized Method of Moment

Let us first consider discrete time observations $y_1, \ldots, y_T$ of a Cox-Ingersoll-Ross process. Different conditional moments admit closed form expressions.

a) The moment of real exponential transformations of y_t :

$$E_t \left[\exp(-\gamma y_{t+1})\right], \gamma \in I\!R^+.$$

b) The moment of imaginary exponential transformation of y_{t+1} :

$$E_t \exp(i\gamma y_{t+1}), \gamma \in I\!R,$$

that are the sine and cosine transformations of γy_{t+1}.

c) The moment of the Laguerre polynomials [see Gourieroux, Jasiak (2005)] :

$$E_t P_n(y_{t+1}), n \in \mathbb{N}^*.$$

Each set above defines an infinite number of moments conditions. Moreover each set of functions, that are $\{\exp(-\gamma y), \gamma \in I\!R^+\}, (\exp(i\gamma y), \gamma \in$

$I\!R$), $(P_n(y), n \in \mathbb{N}^*)$, can be used to approximate as closely as possible the score function. Thus, from a theoretical point of view, the optimal use of any of this set allows reaching the asymptotic efficiency [see Carrasco, Florens (2000)] (except when the support is parameterized).

However in practice, the moment estimation is based on only a finite number of conditional moments of the set, 5-7, say. The set of moment conditions providing the smallest loss of efficiency with this limited number of well-chosen moments will be prefered. In the applications to CIR process the approach by the empirical characteristic function (set b) has been privilegiated [see Singleton (2001), Bates (2002) for applications to CIR process and Feuerverger, Mc Dunnough (1981), Feuerverger (1990) for the first introduction of this approach]. It is not clear that the choice of set b) is the most appropriate. Firstly the Laguerre polynomials are the keystone in the nonlinear canonical decomposition of the transition pdf of the CIR process. By this decomposition it is possible to know exactly and then to control the loss of efficiency, when we use only the first N polynomials [see Kessler, Sorensen (1999) for an application of this methodology]. Second the sine and cosine transformations have no appealing financial interpretation. This is not the case of the real exponential transformations, which were systematically involved in all financial computations to derivative pricing and term structure, and are likely closer to the parameters of interest. Typically it is easier to approximate the transition pdf in the tail (that are the extreme risks) by a small number of decreasing exponential functions than by a small number of sine and cosine functions.

For an observable Wishart process, it is difficult to use the approach based on zonal polynomials appearing in the expression of the transition pdf, since their closed form is not well-known and requires time consuming recursive computations (except for small rank of autoregressive matrix M). The arguments for the choices between sets a) and b) are the same in the multivariate Wishart framework than for the CIR process.

Finally we have already mentioned that the conditional Laplace transforms of the observable processes still have closed form expressions, when (Y_t) is not directly observable. For instance, when the Wishart process is known up to a parameterized transformation, as in the example of WQTSM, the conditional expectation :

$$E_t \exp\left\{(\gamma_1,\gamma_2,\gamma_3)\beta(\theta)^{-1}\left[\left[\begin{array}{c} r(t,t+h_1) \\ r(t,t+h_2) \\ r(t,t+h_3) \end{array}\right] - \alpha(\theta\right]\right\},$$

has a closed form expression in terms of observables rates.

When the Wishart process is known through a measurement equation as in the multivariate stochastic volatility model, we have already derived the conditional Laplace transform of the yield process [see Section 3.5.3]. Along the same lines it is possible to get the closed form expression of a joint marginal Laplace transform of the type $E\exp(\sum_{h=0}^{H-1} u_{1,h}\log S_{1,t+h} + u_{2,h}\log S_{2,t+h})$. This form can be the basis of a moment method based on the observations of asset prices only.

Finally it has been proposed to invert recursively the associated joint Fourier transform in order to apply maximum likelihood [Bates (2002)]. This approach, which extends the standard Kalman filter, can be followed if the number of assets in rather small and the support does not depend on parameters.

5 Concluding Remarks

The Wishart process is the multivariate extension of the Cox, Ingersoll, Ross process. It has been shown that the Wishart process is appropriate for modelling multivariate risk in various financial problems and provides closed form derivative prices. Moreover it is sufficiently flexible to compete with multivariate ARCH models.

The field of applications of Wishart process seems quite large. For instance it can be used :

i) to define joint dynamics of interest rates, exchange rates, market returns, and to study international volatility transmission [Gourieroux, Monfort, Sufana (2005)];

ii) to construct a new class of stochastic intensity models for duration. This class can be used to extend the Archimedean family of copulas, or to propose joint modelling of trading times and prices for high frequency data.

iii) to develop a multivariate mean-variance causality analysis.

Appendix 1
Derivation of the Riccati Equations

1.1 Transition of the CIR process.

Let us consider the (conditional) Laplace transform of the CIR process and assume an exponential affine form in the current value :

$$\psi_{t,h}(u) = E_t[\exp(-uY_{t+h})] = \exp[-a(h,u)y_t - b(h,u)], \text{say.}$$

By iterated expectations theorem we get :

$$\begin{aligned}
\psi_{t,h}(u) &= E_t E_{t+dt} \exp -(uY_{t+h}) \\
&= E_t[\psi_{t+dt,h-dt}(u)] \\
&= E_t \exp[-a(h-dt,u)y_{t+dt} - b(h-dt,u)] \\
&\sim E_t \exp[-a(h-dt,u)[y_t - k(y_t-\theta)dt + (\eta^2 y_t)^{1/2} dW_t] \\
&\quad -b(h-dt,u)] \\
&= \exp[-a(h-dt,u)y_t + a(h-dt,u)k(y_t-\theta)dt - b(h-dt,u)] \\
&\quad E_t \exp[-a(h-dt,u)(\eta^2 y_t)^{1/2} dW_t] \\
&\sim \exp\{-a(h-dt,u)y_t + a(h,u)k(y_t-\theta)dt - b(h-dt,u) \\
&\quad +\frac{1}{2}a^2(h,u)\eta^2 y_t dt].
\end{aligned}$$

By identifying with the assumed expression of $\psi_{t,h}(u)$, we get :

$$\begin{cases}
a(h,u) \sim a(h-dt,u) - ka(h,u)dt - \frac{1}{2}\eta^2 a^2(h,u)dt, \\
b(h,u) \sim b(h-dt,u) + k\theta a(h,u)dt.
\end{cases}$$

By taking dt close to zero, we get the two functions as solutions of :

$$\begin{cases}
\dfrac{\partial a(h,u)}{\partial h} = -ka(h,u) - \frac{1}{2}\eta^2 a^2(h,u), \\
\dfrac{\partial b(h,u)}{dh} = k\theta a(h,u).
\end{cases} \tag{a.1}$$

Moreover, since $E_t \exp(-uY_t) = \exp(-uy_t)$, the functions satisfy the initial restrictions :

$$a(0,u) = u, b(0,u) = 0.$$

1.2 The Integrated CIR Process

Let us assume that :

$$\psi^*_{t,h}(u) = E_t \left[\exp\left(-u \int_t^{t+h} Y_\tau d\tau \right) \right] = \exp\left[-a^*(h,u)y_t - b^*(h,u) \right].$$

By iterated expectation, we get :

$$\begin{aligned}
\psi^*_{t,h}(u) &= E_t E_{t+dt} \left[\exp\left(-u \int_t^{t+h} Y_\tau d\tau \right) \right] \\
&= E_t \left\{ \exp\left(-u \int_t^{t+dt} Y_\tau d\tau \right) E_{t+dt} \left[\exp\left(-u \int_{t+dt}^{t+h} Y_\tau d\tau \right) \right] \right\} \\
&\simeq E_t \left[\exp(-uy_t dt) \psi^*_{t+dt,h-dt}(u) \right] \\
&= E_t \exp \left\{ -uy_t dt - a^*(h-dt,u)[y_t - k(y_t-\theta)dt + (\eta^2 y_t)^{1/2} dW_t \right. \\
&\quad \left. -b^*(h-dt,u) \right\} \\
&\simeq \exp \left[-uy_t dt - a^*(h-dt,u)y_t + a^*(h,u)k(y_t-\theta) \right. \\
&\quad \left. + \frac{a^*(h,u)^2}{2} \eta^2 y_t dt - b^*(h-dt,u) \right].
\end{aligned}$$

By identification both expressions of the Laplace transform, we get for dt tending to zero.

$$\left\{ \begin{array}{lcl} \dfrac{\partial a^*(h,u)}{\partial h} &=& u - ka^*(h,u) - \dfrac{\eta^2}{2}[a^*(h,u)]^2, \\ \dfrac{\partial b^*(h,u)}{\partial h} &=& k\theta a^*(h,u). \end{array} \right. \tag{a.2}$$

The initial conditions are : $a^*(0,u)=0, b^*(0,u)=0$, since :
$\psi^*_{t,0}(u) = E_t \exp[-u \int_t^t Y_\tau d\tau] = 1.$

1.3 The Integrated Wishart Process

We have :

$$\begin{aligned}
\psi^*_{t,h+dt}(\Gamma) &= E_t\{\exp[Tr(\int_t^{t+dt} \Gamma Y_\tau d\tau)]\psi^*_{t+dt,h}(\Gamma)\} \\
&\simeq \exp\{Tr(\Gamma Y_t)dt + b^*(h,\Gamma)\}E_t \exp Tr(A^*(h,\Gamma)Y_{t+dt}) \\
&\simeq \exp\{Tr(\Gamma Y_t)dt + b^*(h,\Gamma) + E_t Tr(A^*(h,\Gamma)Y_{t+dt}) \\
&+ \frac{1}{2}V_t Tr(A^*(h,\Gamma)Y_{t+dt})\} \\
&= \exp\{Tr(\Gamma Y_t)dt + b^*(h,\Gamma) + Tr[A^*(h,\Gamma)Y_t + (KQQ' + Y_t A' + AY_t)dt)\} \\
&+ 2Tr[A^*(h,\Gamma)Y_t A^*(h,\Gamma)Q'Q]dt].
\end{aligned}$$

The result follows by identifying both expressions of the Laplace transform and let t tend to zero.

Appendix 2
Closed Form Solutions of Riccati Equations

2.1 One Dimensional Riccati Equation.

Let us consider a univariate Riccati equation :

$$\frac{da(h)}{dh} = b[a(h) - c_o][a(h) - c_1].$$

This equation can also be written as :

$$da(h)\left[\frac{1}{a(h) - c_o} - \frac{1}{a(h) - c_1}\right] = b(c_o - c_1)dh.$$

It can be integrated as :

$$\frac{a(h) - c_o}{a(h) - c_1} = \frac{a(0) - c_o}{a(0) - c_1}\exp[b(c_o - c_1)h],$$

or equivalently, as :

$$a(h) = c_1 + \frac{[a(0) - c_1](c_o - c_1)}{a(0) - c_1 - [a(0) - c_o]\exp[b(c_o - c_1)h]} \tag{a.3}$$

i) Application to the transition of the CIR process

From Appendix 1.1 equation (a.1) the function $a(h, u)$ satisfies a Riccati equation with : $b = \frac{-\eta^2}{2}, c_1 = 0, c_o = -\frac{2k}{\eta^2}$ and initial condition $a(0, u) = u$. By applying formula (a.3), we deduce that :

$$\begin{aligned} a(h, u) &= \frac{-\frac{2k}{\eta^2}u}{u - [u + \frac{2k}{\eta^2}]\exp(kh)} \\ &= \frac{u\exp(-kh)}{1 + \frac{\eta^2 u}{2k}[1 - \exp(-kh)]}. \end{aligned}$$

Finally function $b(h, u)$ can be derived by integrating the second part of system (a.1). We get :

$$\begin{aligned}\frac{\partial b(h,u)}{\partial h} &= k\theta a(h,u) = \frac{ku\theta\exp(-kh)}{1+\frac{\eta^2}{2k}u[1-\exp(-kh)]} \\ &= \frac{2k\theta}{\eta^2}\frac{\frac{\eta^2}{2k}uk\exp(-kh)}{1+\frac{\eta^2}{2k}u[1-\exp(-kh)]},\end{aligned}$$

and deduce :

$$b(h,u) = \frac{2k\theta}{\eta^2}\log\left\{1+\frac{\eta^2}{2k}u[1-\exp(-kh)]\right\}.$$

ii) **Application to the integrated CIR process**

The right hand side of the multidimensional Riccati equation satisfied by $a^*(h,u)$ can be written as :

$$1 - ka^*(h) - \frac{\eta^2}{2}(a^*(h))^2 = b(a^*(h)-c_o)(a^*(h)-c_1),$$

$$\begin{aligned}\text{where} \quad b &= -\eta^2/2, c_o = -k/\eta^2 - \gamma/\eta^2, c_1 = -k/\eta^2 + \gamma/\eta^2, \\ \gamma &= \sqrt{k^2+2\eta^2}.\end{aligned}$$

By applying the general formula (.) providing the solution of a Riccati equation, we get :

$$\begin{aligned}a^*(h) &= \frac{\gamma-k}{\eta^2} + \frac{2(\gamma-k)\gamma/\eta^2}{k-\gamma-(k+\gamma)\exp(\gamma h)} \\ &= \frac{2}{\gamma+k} - \frac{4\gamma}{\gamma+k}\frac{1}{\gamma-k+(\gamma+k)\exp(\gamma h)},\end{aligned}$$

since : $\gamma^2 - k^2 = 2\eta^2$.

2.2 Multidimensional Riccati Equation

Generally multidimensional Riccati equations have no closed form solutions. However the multidimensional Riccati equations involved in Wishart processes have a special form. This explains why their solution can be easily derived [see Grasselli, Tebaldi (2004) for a discussion of Riccati equations with closed form solution].

Let us consider a (matricial) Riccati differential system :

$$\frac{dA(h)}{dh} = B'A(h) + A(h)B + 2A(h) \wedge A(h) + C, \tag{a.4}$$

where $A(h), \wedge, C$ are symmetric (n, n) matrices and B is a square (n, n) matrix. The solution of the multidimensional equation (a.4) is [see Gourieroux, Sufana (2004)] :

$$A(h) = A^* + \exp[(B + 2 \wedge A^*)h]'$$

$$\{(A(0) - A^*)^{-1} + 2\int_0^h \exp[(B + 2 \wedge A^*)u] \wedge \exp[(B + 2 \wedge A^*)u]'du\}.$$

$$\exp[(B + 2 \wedge A^*)h],$$

where A^* satisfies :

$$B'A^* + A^*B + 2A^*BA^* + C = 0.$$

This result can be directly applied to the multidimensional partial Riccati equation of Proposition 6 to get the closed form solution.

Appendix 3
The Noncentered Gamma and Wishart Distributions

3.1 Laplace transform of the noncentered gamma distribution

Let us assume that Y follows $\gamma(\nu, \beta, \lambda)$, say. Then Y/λ follows $\gamma(\nu + Z)$, where Z follows $\mathcal{P}(\beta)$.We get :

$$\begin{aligned}
& E[\exp(-uY)] \\
= \; & EE\left\{\exp[-u\lambda(Y/\lambda)]|Z\right\} \\
= \; & E\left[\frac{1}{(1+u\lambda)^{\nu+Z}}\right] \\
= \; & \sum_{z=0}^{\infty}\left[\exp(-\beta)\frac{\beta^z}{z!}\frac{1}{(1+u\lambda)^{\nu+z}}\right] \\
= \; & \frac{1}{(1+u\lambda)^{\nu}}\exp(-\beta)\exp\left(\frac{\beta}{1+u\lambda}\right) \\
= \; & \exp[-\beta u\lambda/(1+u\lambda)]\exp[-\nu\log(1+u\lambda)]. \qquad (a.??)
\end{aligned}$$

Application to the CIR process.

By comparing this expression with the (conditional) Laplace transform of the CIR process at horizon h, we deduce Proposition 2, with :

$$\lambda(h) = \frac{\eta^2}{2k}[1-\exp(-kh)], \beta(h) = \frac{\exp(-kh)}{\lambda(h)}Y_t, \nu = \frac{2k\theta}{\eta^2}.$$

3.2 Density function of the noncentered gamma distribution

The density function of the $\gamma(\nu, \beta, \lambda)$ distribution is :

$$f(y) = \sum_{z=0}^{\infty}\left[\exp(-y/\lambda)\frac{y^{\nu+z-1}}{\lambda^{\nu+z}\Gamma(\nu+z)}\exp(-\beta)\frac{\beta^z}{z!}\right].$$

We deduce the formula given in Proposition 2 iii).

3.3 Conditional Laplace transform of the square of a Gaussian VAR process

The proof is based on the following lemma :

Lemma : For any symmetric positive semi-definite matrix Ω and any vector $\mu \in I\!R^n$, we get :

$$\int_{R^n} \exp(-x'\Omega x + \mu' x) dx = \frac{\pi^{n/2}}{(det\Omega)^{1/2}} \exp(\frac{1}{4}\mu'\Omega^{-1}\mu).$$

Let us now consider an Ornstein-Uhlenbeck process :

$$dx_t = Ax_t dt + QdW_t.$$

The conditional distribution of x_{t+1} given x_t is Gaussian, with mean $M(h)x_t$ and variance-covariance matrix $\Sigma(h)$. For $Y_{t+h} = x_{t+h}x'_{t+h}$, we get :

$$\begin{aligned}
& E_t \exp Tr(\Gamma Y_{t+h}) \\
= \; & E_t \exp(x'_{t+h}\Gamma x_{t+h}) \\
= \; & \int_{R^n} \exp\{x'[\Gamma - \frac{\Sigma(h)^{-1}}{2}]x + x'\Sigma(h)^{-1}M(h)x_t\}dx \\
& \frac{1}{(2\pi)^{n/2}} \frac{1}{(det\Sigma(h))^{1/2}} \exp[-\frac{1}{2}x'_t M(h)'\Sigma(h)^{-1}M(h)x_t].
\end{aligned}$$

By applying the lemma we get :

$$\begin{aligned}
& E_t[\exp Tr(\Gamma Y_{t+h})] \\
= \; & \frac{1}{det(Id - 2\Sigma(h)\Gamma)^{1/2}} \\
& \exp\{-\frac{1}{2}x'_t M(h)'\Sigma(h)^{-1}M(h)x_t + \frac{1}{2}x'_t M(h)'\Sigma(h)^{-1}(\Sigma(h)^{-1} - 2\Gamma)^{-1}\Sigma(h)^{-1}M(h)x_t \\
= \; & \frac{\exp Tr[\Gamma(Id - 2\Sigma(h)\Gamma)^{-1}M(h)Y_t M(h)']}{det(Id - 2\Sigma(h)\Gamma)^{1/2}}.
\end{aligned}$$

This is the result of Proposition 4 for $K = 1$. The general case is immediately deduced.

Appendix 4
Competing Multivariate Dynamic Volatility Models

In this appendix we review the structure of the main multivariate models proposed for stochastic volatility-covolatility matrices. For expository purpose we assume $n = 2$ and a small number of lags.

1. Duffie-Kan model [Duffie, Kan (1996), Dai, Singleton (2000)]

This model is usually written in continuous time. A typical example is :

$$Y_t = Q \begin{pmatrix} a_{1,1} f_{1,t} + a_{1,2} f_{2,t} + c_1 & 0 \\ 0 & a_{2,1} f_{1,t} + a_{2,2} f_{2,t} + c_2 \end{pmatrix} Q',$$

where $a_{ij} \geq 0, c_j \geq 0, (f_{1,t})$ and $(f_{2,t})$ are independent CIR processes.

2. The Constant Conditional Correlation Model (CCC) [Bollerslev (1987)]

A typical example is :

$$\begin{aligned} y_{1,1,t} &= c_1 + a_1 y_{1,1,t-1} + b_1 r_{1,t}^2, \\ y_{2,2,t} &= c_2 + a_2 y_{2,2,t-1} + b_2 r_{2,t}^2, \\ y_{1,2,1} &= \rho (y_{1,1,t})^{1/2} (y_{2,2,t})^{1/2}, \end{aligned}$$

where $r_t = (r_{1,t}, r_{2,t})' = Y_t^{1/2} \varepsilon_t$ and $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ is a Gaussian white noise.

3. The Factor ARCH Model [Diebold, Nerlove (1989), Engle, Ng, Rothschild (1990), King, Sentana, Whadwani (1994)].

This model is generally written on asset returns. In the one factor case, the model is defined by :

$$\begin{aligned} r_{1,t} &= \beta_1 f_t + \varepsilon_{1,t}, \\ r_{2,t} &= \beta_2 f_t + \varepsilon_{2,t}, \end{aligned}$$

where $(\varepsilon_{1,t}, \varepsilon_{2,t})$ is a Gaussian white noise, independent of the factor process, and the factor process is ARCH. The structure of the return volatility-covolatility matrix is :

$$\begin{aligned} V_{t-1}\begin{pmatrix} r_{1,t} \\ r_{2,t} \end{pmatrix} &= V\varepsilon + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} (\beta_1, \beta_2) V_{t-1} f_t \\ &= \Omega + \eta_{t-1} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} (\beta_1, \beta_2), \text{ say.} \end{aligned}$$

4. Baba, Engle , Kraft, Kroner Model (BEKK) [Baba, Engle, Kraft, Kroner (1987), Engle, Kroner (1995)]

A typical example is :

$$Y_t = C + \sum_{j=1}^{J} \lambda_j A_j' r_{t-1} r_{t-1}' A_j,$$

where (r_t) denotes the vector of asset returns, A_j are $(2,2)$ matrices, λ_j positive scalars, and $C >> 0$. This specification ensures symmetric positive semi-definite matrices Y_t.

5. The Dynamic Conditional Correlation Model (DCC) [Engle, Sheppard (2001), Engle (2002)].

This model distinguishes the dynamics of asset volatilities and asset conditional correlations. Typically a one-dimensional GARCH specification is introduced for each volatility :

$$\begin{aligned} y_{1,1,t} &= c_1 + a_1 y_{1,1,t-1} + b_1 r_{1,t}^2, \\ y_{2,2,t} &= c_2 + a_2 y_{2,2,t-1} + b_2 r_{1,t}^2, \\ \frac{y_{1,2,t}}{\sqrt{y_{1,1,t}}\sqrt{y_{2,2,t}}} &= \frac{q_{1,2,t}}{\sqrt{q_{1,1,t}}\sqrt{q_{2,2,t}}}, \\ \text{where : } \quad q_{i,j,t} &= (1-\alpha-\beta) q_{i,j} + \alpha \frac{r_{i,t-1}}{\sqrt{y_{i,i,t}}} \frac{r_{j,t-1}}{\sqrt{y_{j,j,t}}} + \beta q_{i,j,t-1}. \end{aligned}$$

The CCC model corresponds to the special case $\alpha = \beta = 0$.

REFERENCES

Ahn, C., Dittmar, R., and A., Gallant (2002) : "Quadratic Term Structure Models : Theory and Evidence", Review of Financial Studies, 15, 243-288.

Anderson, T., and M., Girshick (1944) : "Some Extensions of the Wishart Distribution", Annals of Mathematical Statistics, 15, 345-357.

Ang, A., and J., Chen (2002) : "Asymmetric Correlations in Equity Portfolios", Journal of Financial Economics, 63, 443-494.

Baba, Y., Engle, R., Kraft, D. and K., Kroner (1987) : "Multivariate Simultaneous Generalized ARCH", DP UCSD.

Bates, D. (2002) : "Maximum Likelihood Estimation of Latent Affine Processes", DP Iowa Univ.

Ball, C., and A., Roma (1994) : "Stochastic Volatility Option Pricing", Journal of Financial and Quantitative Analysis, 29, 589-607.

Bollerslev, T. (1987) : "A Multivariate GARCH Model with Constant Conditional Correlations for a Set of Exchange Rates", DP. Northwestern University.

Bollerslev, T. (1990) : "Modelling the Coherence in Short Run Nominal Exchange Rates : A Multivariate Generalized ARCH Model", Review of Economics and Statistics, 72, 498-505.

Bollerslev, T., Chou, R., and K., Kroner (1992) : "ARCH Modelling in Finance : A Review of the Theory and Empirical Evidence", Journal of Econometrics, 52, 5-59.

Carrasco, M., and J.P., Florens (2000) : "Generalization of a GMM in a Continuum of Moment Conditions", Econometric Theory, 11, 797-834.

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

Cox, J., Ingersoll, J., and S., Ross (1985) : "A Theory of the Term Structure of Interest Rates", Econometrica, 53, 385-407.

Dai, Q., and K., Singleton (2000) : "Specification Analysis of Affine Term Structure Models", Journal of Finance, 55, 1943-1978.

Diebold, F., and M., Nerlove (1989) : "The Dynamics of Exchange Rate Volatility : A Multivariate Latent Factor ARCH Model", Journal of Applied Econometrics, 4, 1-22

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

Duffie, D., Filipovic, D., and W., Schachermayer (2003) : "Affine Processes and Application in Finance", Annals of Applied Probability, Vol 13, 984-1053.

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

Duffie, D., and K., Singleton (1999) : "Simulated Correlated Defaults", DP Stanford University.

Engle, R. (2002)a : "Dynamic Conditional Correlation : A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models", Journal of Business and Economic Statistics, 20, 122-150.

Engle, R. (2002) b : "New Frontiers for ARCH Models", Journal of Applied Econometrics, 17, 425-446.

Engle, R., and K., Kroner (1995) : "Multivariate Simultaneous GARCH", Econometric Theory, 11, 122-150.

Engle, R., Ng, V., and M., Rothschild (1990) : "Asset Pricing with a Factor ARCH Covariance Structure : Empirical Estimates for Treasury Bills", Journal of Econometrics, 45, 213-237.

Engle, R., and K., Sheppard (2001) : "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", NBER Working Paper 8554.

Feller, W. (1971) : "An Introduction to Probability Theory and Its Applications", Vol 2, Wiley.

Feuerverger, A. (1990) : "An Efficient Result for the Empirical Characteristic Function on Stationary Time Series Models", The Canadian Journal of Statistics, 18, 155-161.

Feuerverger, A., and P., Mc Dunnough (1981) : "On the Efficiency of the Empirical Characteristic Function Procedures", Journal of the Royal Statistical Society, B, 43, 20-77.

Glejser, L. (1976) : "A Canonical Representation of the Noncentral Wishart Distribution Useful for Simulation", Journal of the American Statistical Association, 71, 690-697.

Gourieroux, C., and J., Jasiak (2001) : "Financial Econometrics", Princeton University Press.

Gourieroux, C., and J., Jasiak (2005) : "Autoregressive Gamma Processes", forthcoming International Journal of Forecasting.

Gourieroux, C., Jasiak, J., and R., Sufana (2004) : "The Wishart Autoregressive Process of Multivariate Stochastic Volatility", DP University of Toronto.

Gourieroux, C., Monfort, A. and R., Sufana (2005) : "International Money and Stock Market Contingent Claims", DP, University of Toronto.

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

Gourieroux, C., and R., Sufana (2004) : "Derivative Pricing with Multivariate Stochastic Volatility", CREF, 04-09, HEC, Montreal.

Grasselli, M., and C., Tebaldi (2004) : "Solvable Affine Term Structure Models", DP, Verone University.

Han, Y. (2002) : "The Economic Value of Volatility Modelling : Asset Allocation with a High Dimensional Dynamic Latent Factor Multivariate Stochastic Volatility Model", DP, Olin School of Business, St Louis.

Harvey, A., Ruiz, E., and N., Shephard (1994) : "Multivariate Stochastic Variance Models", Review of Economic Studies, 61, 247-264.

Heston, S. (1993) : "A Closed-Form Solution for Options with Stochastic Volatility with Application to Bond and Currency Options", Review of Financial Studies, 6, 327-343.

Ho, T., and S., Lee (2004) : "A Multifactor Binomial Interest Rate Model with State Time Dependent Volatilities", DP, Hanyang Univ.

Jacquier, E., and A., Marcus (2001) : "Asset Allocation Models and Market Volatility", Financial Analysts Journal, 57,16-30.

James, A. (1968) : "Calculation of Zonal Polynomial Coefficients by Use of the Laplace Beltrami Operator", Ann. Math. Stat., 39, 1711-1718.

Kessler, M., and M., Sorensen (1999) : "Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process", Bernoulli, 5, 299-314.

King, M., Sentana, E., and S., Wadhwani (1994) : "Volatility Links between National Stock Markets", Econometrica, 62, 901-933.

King, M., and S., Wadhwani (1990) : "Transmission of Volatility Between Stock Markets", Review of Financial Studies, 3, 5-33.

Kraft, E., and R., Engle (1983) : "Multiperiod Forecast Error Variances of Inflation Estimated from ARCH Models", in Zellner, eds : Applied Time Series Analysis of Economic Data, Bureau of Census, Washington.

Lando, D. (1998) : "On Cox Processes and Credit Risky Securities", Review of Derivatives Research, 2, 99-120.

Ledoit, O., Santa-Clara, P., and M., Wolf (2003) : "Flexible Multivariate GARCH Modelling with an Application to International Stock Markets", Review of Economics and Statistics, 81, forthcoming.

Leland, M., and K., Toft (1996) : "Optimal Capital Structure, Endogenous Bankruptcy and the Term Structure of Credit Spreads", Journal of Finance, 51, 987-1019.

Leippold, M., and L., Wu (2002) : "Asset Pricing Under the Quadratic Class", Journal of Financial and Quantitative Analysis, 37, 274-295.

Lin, J. (1977) : "Generalized Method of Moments Estimation of Affine Diffusion Processes", Standford Univ DP.

Merton, R. (1974) : "On the Pricing of Corporate Debt : The Risk Structure of Interest Rates", Journal of Finance, 29, 449-470.

Muirhead, R. (1982) : "Aspect of Multivariate Statistical Theory", Wiley.

Pearson, N., and T., Sun (1994) : "Exploiting the Conditional Density in Estimating the Term Structure : An Application to the CIR Model", Journal of Finance, 46, 1279-1304.

Philipov, A., and M., Glickman (2004) : "Multivariate Stochastic Volatility via Wishart Processes", DP Boston Univ.

Singleton, K. (2001) : "Estimation of Affine Diffusion Models Based on the Empirical Characteristic Function", Journal of Econometrics, 102, 111-141.

Turnbull, S. (2004) : "Unresolved Issues in Modelling Credit Risky Assets", DP. Univ. of Houston.

Tse, Y., (2000) : "A Test for Constant Correlations in a Multivariate GARCH Model", Journal of Econometrics, 98, 107-127.

Wishart, J. (1928 a,b) : "The Generalized Product Moment Distribution in Samples from a Multinomial Population", Biometrika, 20, 32 and 424.