Efficient Simulation of the Wishart model

Pierre Gauthier [*], Dylan Possamaï[†]

September 22, 2009

Abstract. In financial mathematics, Wishart processes have emerged as an efficient tool to model stochastic covariance structures. Their numerical simulation may be quite challenging since they involve matrix processes. In this article, we propose an extensive study of financial applications of Wishart processes. First, we derive closed-form formulas for option prices in the single-asset case. Then, we show the relationship between Wishart processes and Wishart law. Finally, we review existing discretization schemes (Euler and Ornstein-Uhlenbeck) and propose a new scheme, adapted from Heston's QEM discretization scheme. Extensive numerical results support our comparison of these three schemes.

Keywords: Stochastic Volatility, Equity options, Multifactor model, Wishart model, Discretization scheme, Random Matrix, Heston model

JEL Classification: C15, G13
MSC 2000: 60-99, 65C05

[*]pierre.gauthier@pricingpartners.com - Pricing Partners, Paris Cybervillage, 204 Rue de Crimee, 75019 Paris, France
[†]dylan.possamai@pricingpartners.com - Pricing Partners, Paris Cybervillage, 204 Rue de Crimee, 75019 Paris, France

1 Introduction

Since Black and Scholes introduced their option valuation model, an extensive literature focuses on pointing out its limitations. More precisely, two major features of the equity index options market cannot be captured through Black-Scholes model. First of all, observed market prices for both in-the-money and out-the-money options are higher than Black-Scholes prices with at-the-money volatilities. This effect is known as the volatility smile: the volatility depends both on the option expiry and the option strike. Second, there exists a term structure of implied volatilities. As a matter of fact, a constant volatility parameter does not enable to model correctly this behavior.

In order to model the smile efficiently, stochastic volatility models are a popular approach. They enable to have distinct processes for the stock return and its variance. Thus, they may generate volatility smiles. Moreover, if the variance process embeds a mean reversion term, these models can capture the term-structure in the variance dynamics. Popular stochastic volatility models include Heston [28], SABR and square-root models to name but a few.

Efficient calibration of stochastic volatility models requires an analytical formula for option prices. For numerous models, including Heston, this is achieved through the Fourier-transform technique described in Carr and Madan [11]. Further refinements have been discussed in Lewis [36] and Lord and Kahl [38].

In order to match precisely the market implied volatility surface, it turns out that Heston model does not have enough parameters. Therefore, we may wish to add degrees of freedom (i.e additional parameters) while retaining analytical tractability. Many academicians and practitioners have tackled this challenge by considering time-dependent extensions of the original Heston model. Another direction is to model the variance by a variable of higher dimension. Such a path has been followed by Gouriéroux et al and Da Fonseca et al., who replace the Cox-Ingersoll-Ross variance process by a Wishart process. Their model enables to have a tighter control of the covariance dynamics as it is represented by a matrix of size n (2 is often enough in practice). In this paper, we follow their approach, taking a closer look at numerical issues.

Regarding exotic products, the Fourier transform does not exist in most cases. Consequently, we have to resort to numerical simulation, Monte Carlo being a popular choice. Monte Carlo simulation may be challenging as convergence speed can be quite low for naive Euler scheme. Therefore, there is a need for more accurate discretization of both stock return and variance processes. In [44], van Haastrecht and Pelsser propose an extensive discussion of existing discretization schemes for Heston model. However, regarding numerical simulation of Wishart processes, literature is less abundant. Bru and Yor et al. give an interesting and thorough theoretical approach of Wishart processes. These references contain most of Wishart processes properties which are recalled hereafter.

In this article, we consider the framework introduced by Da Fonseca et al. Our objective is two-fold. First, we study the numerical behavior of the characteristic function. Notably, we show evidence that the complex log issue arises and we explain how to solve it. Second, we get interested into convergence speed of Monte-Carlo simulation using different discretization schemes.

The paper is structured as follows. First, in section 2, we introduce notations that will be used throughout the article. In section 3, we recall the model dynamics. In section 4, we recall important (for our purpose) properties of Wishart processes. Then, in section 5, we look at the characteristic function of Wishart processes. Section 6 is devoted to Monte-Carlo discretization schemes. Our numerical results are presented in section 8 and conclusion concludes.

2 Notation

Unless otherwise stated, from now on we will work with a probability space denoted $\left(\Omega, (\mathcal{F}_t)_{t\in\mathbb{R}^+}, \mathbb{Q}\right)$, where $\mathbb{Q}$ is supposed to be a so-called risk-neutral probability. It means that unless otherwise stated, all expectations are taken under the probability $\mathbb{Q}$. Besides, an expectation noted with the index t as in $\mathbb{E}_t$, will be equivalent to taking the expectation conditionally on the filtration $\mathcal{F}$ up to time t.

- $M_{n,m}(\mathbb{R}), M_{n,m}(\mathbb{C})$: the sets of $n \times m$ real and complex matrix.
- $M_n(\mathbb{R}), M_n(\mathbb{C})$: the sets of real and complex square matrix.
- $S_n(\mathbb{R})$: the sets of real symmetric square matrix.
- $S_n^+(\mathbb{R})$: the sets of real symmetric positive square matrix.
- $S_n^-(\mathbb{R})$: the sets of real symmetric negative square matrix.
- $\widetilde{S}_n^+(\mathbb{R})$: the sets of real symmetric positive definite square matrix.
- $\widetilde{S}_n^-(\mathbb{R})$: the sets of real symmetric negative definite square matrix.
- A^T: the transpose of the matrix A.
- $\mathrm{Tr}(A)$: the trace of the matrix A.
- $\det(A)$: the determinant of the matrix A.
- A^{ij} is the element on the i^{th} row and j^{th} column of the matrix A.
- If $A \in S_n^+$, $\sqrt{A}$ is the unique symmetric positive matrix which square equals A.
- O_n: the orthogonal group.
- $GL_n(\mathbb{R}), GL_n(\mathbb{C})$: the sets of real and complex invertible matrix.
- Let I be a subset of $\{1, ..., n\}$ and $A \in M_n(\mathbb{R})$. We will note A_{II} the principal minor of the matrix A associated to I, that is to say $A_{II} = \left(A^{ij}\right)_{\min(I)\leq i,j\leq\max(I)}$.
- For a given matrix M, $\lambda(M)$, $\lambda_{min}(M)$ and $\lambda_{max}(M)$ will denote respectively any, the smallest and the largest eigenvalue of M.
- $\forall\, x \in \mathbb{R}$, $[x]$ will be the integer part of x.
- For a countable set $\mathcal{N}$, $\#\mathcal{N}$ will denote its cardinal.
- If the random variable X defined on our probability space has the probability distribution $\mathcal{D}$, we will note $X \sim \mathcal{D}$.

3 Model presentation

Following our previous work on the Double Heston model in [23], we intend to study here its natural multifactor generalization, the Wishart Models. They were first introduced by Bru in [9] and extensively studied in the form that we are interested in by Da Fonseca and al. in [18], [19], [20] and [21]. We stay in the framework of a single-asset model and assume the following dynamics for the price of the asset S_t:

$$\frac{dS_t}{S_t} = r_t dt + \mathrm{Tr}\left(\sqrt{V_t} dZ_t\right)$$

$$dV_t = \left(\beta Q^T Q + M V_t + V_t M^T\right) dt + \sqrt{V_t} dW_t Q + Q^T dW_t^T \sqrt{V_t}, \tag{1}$$

where:

- r_t is the short rate.
- $Q \in GL_n(\mathbb{R})$.
- M is a negative matrix.
- $\beta > n - 1$.
- W_t and Z_t are matrix Brownian motions.
- $V_0 \in \widetilde{S}_n^+$.

We also assume the following correlation structure, which has been explicitly chosen so that the model could belong to the general class of affine processes (see [14] for a comprehensive study of affine processes) and thus could be analytically tractable:

$$dZ_t = dW_t R + dB_t \sqrt{I_n - R^T R}, \tag{2}$$

where R is a matrix which spectral radius is below 1 and B_t a matrix Brownian motion independent of W_t.

Therefore, the Wishart model here appears as a multi-factor extension of the Heston model and a generalization of the Double Heston, where the volatility process is a symmetric positive definite matrix which follows the dynamic of a Wishart process.

4 Properties of Wishart processes

4.1 Existence and unicity

The existence and unicity of a solution to the SDE (1) is given by the following theorem (see [16] and [9])

Theorem 1. *Let V_t be a process satisfying (1)*

- *If $\beta \in]n-1, n+1[$ then (1) has a unique weak solution in S_n^+*
- *If $\beta \geq n+1$ then (1) has a unique strong solution in $\widetilde{S}_n^+$*
- *The eigenvalue process $\left\{\lambda_t^i, 1 \leq i \leq n\right\}_{t\geq 0}$ almost surely never collide. That is to say that almost surely*

$$\forall t \geq 0,\ \lambda_t^1 > ... > \lambda_t^n \geq 0,$$

the last inequation being strict as soon as $\beta \geq n+1$.

Thus we are dealing with a matrix process with values in S_n^+ or $\widetilde{S}_n^+$, depending on the value of β.

4.2 Link between Ornstein-Uhlenbeck and Wishart processes

In the particular cases where β is an integer, the following proposition (which can be found in [27]) links Ornstein-Uhlenbeck and Wishart processes. It will prove useful especially when considering discretization schemes.

Proposition 1. *Let $\beta \in \mathbb{N}^*$ and let $\{X_{k,t}, t \geq 0\}_{1\leq k\leq \beta}$ be independent vectorial Ornstein-Uhlenbeck processes in $\mathbb{R}^n$ with dynamic:*

$$dX_{k,t} = MX_{k,t}dt + Q^T dW_{k,t},$$

where $\{W_{k,t}, t \geq 0, 1 \leq k \leq \beta\}$ are independent vectorial Brownian motions.

Then,

$$Y_t = \sum_{k=1}^{\beta} X_{k,t}X_{k,t}^T,$$

has the dynamic (1).

Proof. Thanks to Itō calculus, we easily get:

$$dY_t = \left(\beta Q^T Q + MY_t + Y_t M^T\right) dt + \sum_{k=1}^{\beta} X_{k,t}dW_{k,t}^T Q + Q^T dW_{k,t}X_{k,t}^T.$$

Then, using Lévy criterion for Brownian motions, we can define a matrix-valued Brownian motion W so that:

$$\sqrt{Y_t}dW_t = \sum_{k=1}^{\beta} X_{k,t}dW_{k,t}^T,$$

which ends the proof. □

4.3 Dynamic of the determinant of a Wishart Process

Since the dynamic of the determinant of the Wishart process will prove useful when we consider some discretization scheme, it is interesting to obtain it right now. Of course, for our analysis to be pertinent, we need to restrain to the case $\beta \geq n+1$. In this case the Wishart process remains in the cone of symmetric definite positive matrix and its determinant is always strictly positive. We will use the following theorem which can be found in [9].

Theorem 2. *Let $\{\xi_t\}_{t\geq 0}$ be a matrix valued process in S_n which entries are continuous semi-martingales. For $t \geq 0$, we know from the spectral theorem that there exists $U_t \in O_n$ such that*

$$U_t^T \xi_t U_t = Diag\left(\lambda_t^1, \lambda_t^2, ..., \lambda_t^n\right),$$

where the $\left(\lambda_t^j\right)_{1\leq j\leq n}$ are the eigenvalues of the process ξ_t.

Let us now define the two processes A_t and Γ_t such that

$$A_t = \int_0^t U_s^T d\xi_s U_s$$

$$\Gamma_t^{ij} dt = d\left\langle A^{ij}, A^{ji}\right\rangle_t.$$

Then we have

$$\forall\ 1 \leq i \leq n,\ d\lambda_t^i = dM_t^i + dJ_t^i$$

$$dJ_t^i = \sum_{j=1}^n \frac{\Gamma_t^{ij}}{\lambda_t^i - \lambda_t^j} \mathbb{1}_{\lambda_t^i \neq \lambda_t^j} dt + dY_t^i,$$

where dM_t^i and dY_t^i are respectively the martingale part and the finite variation part of dA^{ii}.

Proof. See Appendix A □

We can now state the following proposition of [9] and [16]

Proposition 2. *Let $\{V_t\}_{t\geq 0}$ be the solution of (1) in the case $\beta \geq n+1$. Then we have*

$$\frac{d\left(\det V_t\right)}{\det V_t} = \left[(\beta - n + 1)\operatorname{Tr}\left[V_t^{-1}Q^TQ\right] + 2\operatorname{Tr}(M)\right]dt + 2\operatorname{Tr}\left[\sqrt{V_t}^{-1} dW_t Q\right] \quad (3)$$

$$d\log\left(\det V_t\right) = \left[(\beta - n - 1)\operatorname{Tr}\left[V_t^{-1}Q^TQ\right] + 2\operatorname{Tr}(M)\right]dt + 2\operatorname{Tr}\left[\sqrt{V_t}^{-1} dW_t Q\right] \quad (4)$$

Proof. Let $U_t \in O_n$ so that

$$U_t^T V_t U_t = Diag\left(\lambda_t^1, \lambda_t^2, ..., \lambda_t^n\right).$$

Let us also note $\widetilde{Q}_t = U_t^T Q U_t$, $\widetilde{M}_t = U_t^T M U_t$ and $Z_t = \int_0^t U_s^T dW_s U_s$. Thanks to Lévy criterion for Brownian motions and the fact that U_t is an orthogonal matrix, it is clear that Z_t is still a matrix Brownian motion.

Using theorem 2 we have after some calculations

$$d\lambda_t^i = \left[\beta\left(\widetilde{Q}_t^T \widetilde{Q}_t\right)^{ii} + 2\lambda_t^i \widetilde{M}_t^{ii} + \sum_{j\neq i} \frac{\lambda_t^i \left(\widetilde{Q}_t^T \widetilde{Q}_t\right)^{jj} + \lambda_t^j \left(\widetilde{Q}_t^T \widetilde{Q}_t\right)^{ii}}{\lambda_t^i - \lambda_t^j}\right] dt$$

$$+ 2\sqrt{\lambda_t^i} \sum_{k=1}^n \widetilde{Q}_t^{ki} dZ_t^{ik}$$

Besides, according to Itō's lemma, the fact that $d\left\langle \lambda_t^i, \lambda_t^i \right\rangle = 0$ for $i \neq j$ and that the trace of a matrix is similarity-invariant, we have

$$\begin{aligned}
\frac{d\left(\det V_t\right)}{\det V_t} &= \frac{1}{\det V_t} d\left(\prod_{i=1}^n \lambda_t^i\right) \\
&= \sum_{i=1}^n \frac{d\lambda_t^i}{\lambda_t^i} \\
&= \left[\sum_{j=1}^n \left(\widetilde{Q}_t^T \widetilde{Q}_t\right)^{jj} \sum_{i\neq j} \frac{1}{\lambda_t^i - \lambda_t^j} - \sum_{j=1}^n \frac{\left(\widetilde{Q}_t^T \widetilde{Q}_t\right)^{jj}}{\lambda_t^j} \sum_{i\neq j} \frac{\lambda_t^i}{\lambda_t^i - \lambda_t^j}\right] dt \\
&+ \left[\beta \sum_{i=1}^n \frac{(\widetilde{Q}_t^T \widetilde{Q}_t)^{ii}}{\lambda_t^i} + 2\,\mathrm{Tr}(\widetilde{M}_t)\right] dt + 2\frac{1}{\sqrt{\lambda_t^i}} \sum_{k=1}^n \widetilde{Q}_t^{ki} dZ_t^{ik} \\
&= \left[(\beta - n + 1)\,\mathrm{Tr}\left[V_t^{-1} Q^T Q\right] + 2\,\mathrm{Tr}(\widetilde{M}_t)\right] dt + 2\,\mathrm{Tr}\left[\sqrt{V_t}^{-1} dW_t Q\right].
\end{aligned}$$

The result for $\log\left(\det V_t\right)$ is a simple consequence of Itō's lemma. □

4.4 Infinetesimal generator of the Wishart process

When we will deal later on with the pricing problem of European options in this framework, we will generalize the characteristic function technique already used for Heston and Double Heston models (see [23]). In order to compute the characteristic function (or the Laplace Transform) of the Wishart process, we will then need its infinitesimal generator. It has been computed, for example, in [10].

Proposition 3. *The infinitesimal generator associated with the Wishart process (1) is given by, for* $V \in S_n^+$*:*

$$\mathcal{L}^V = \operatorname{Tr}\left[\left(\beta Q^T Q + MV + VM^T\right) D + 2VDQ^TQD\right], \tag{5}$$

where

$$D = \left(\frac{\partial}{\partial V^{ij}}\right)_{1 \le i,j \le n} \tag{6}$$

Proof. See Appendix A □

4.5 Characteristic function of the Wishart Process

Proposition 4. *Let* $\Theta \in S_n(\mathbb{R})$ *and* $t, h \geq 0$ *so that* $I_n - 2i\Sigma(h)\Theta \in \widetilde{S}_n^+(\mathbb{C})$*, then:*

$$\mathbb{E}\left[e^{i\operatorname{Tr}(\Theta V_{t+h})} | V_t\right] = \frac{e^{i\operatorname{Tr}\left[\Theta\left(I_n - 2i\Sigma(h)\Theta\right)^{-1}\Delta(h)V_t\Delta(h)^T\right]}}{\left(\det\left[I_n - 2i\Sigma(h)\Theta\right]\right)^{\frac{\beta}{2}}}, \tag{7}$$

where

$$\Delta(h) = e^{hM} \tag{8}$$

$$\Sigma(h) = \int_0^h e^{sM}Q^TQe^{sM^T}ds \tag{9}$$

Proof. Let $\psi(h, V_t) = \mathbb{E}\left[e^{i\operatorname{Tr}(\Theta V_{t+h})}|V_t\right]$. As usual when dealing with models involving affine processes, we assume that the characteristic function is exponentially affine, that is to say:

$$\psi(h, V_t) = e^{\operatorname{Tr}(B(h)V_t) + \mu(h)}, \tag{10}$$

where $B(h)$ is a symmetric complex matrix with $B(0) = i\Theta$ and $\mu(0) = 0$.

Then thanks to Feynman-Kac theorem, we know that ψ is the solution of the following PDE:

$$\frac{\partial \psi}{\partial h}(h, V_t) = \mathcal{L}^V \psi(h, V_t). \tag{11}$$

Replacing (5) in (11) leads then to:

$$\operatorname{Tr}\left[B'(h)V_t\right] + \mu'(h) = \operatorname{Tr}\left[\left(B(h)M + M^T B(h) + 2B(h)Q^T Q B(h)\right) V_t\right]$$
$$+ \beta \operatorname{Tr}\left[Q^T Q B(h)\right].$$

We identify the terms containing V_t and get the following differential system:

$$B'(h) = B(h)M + M^T B(h) + 2B(h)Q^T Q B(h) \tag{12}$$

$$\mu'(h) = \beta \operatorname{Tr}\left[Q^T Q B(h)\right] \tag{13}$$

The first equation above is a matrix Ricatti equation (see [22]) which can be easily solved with a linearization procedure, as presented by Da Fonseca and al. in [21].

Put

$$B(h) = F(h)^{-1} G(h), \tag{14}$$

for $F(h) \in GL_n(\mathbb{C})$, $G(h) \in M_n(\mathbb{C})$, $F(0) = I_n$ and $G(0) = i\Theta$.

With these new variables the equation (12) becomes

$$-F'(h)B(h) + G'(h) = G(h)M + F(h)M^T B(h) + 2G(h)Q^T Q B(h), \tag{15}$$

which in turn leads to:

$$G'(h) = G(h)M$$
$$-F'(h) = F(h)M^T + 2G(h)Q^T Q, \tag{16}$$

which is easily solved by

$$\begin{pmatrix} G(h) & F(h) \end{pmatrix} = \begin{pmatrix} i\Theta & I_n \end{pmatrix} \exp\left[h \begin{pmatrix} M & -2Q^T Q \\ 0 & -M^T \end{pmatrix}\right].$$

Besides, using techniques for calculating integrals involving matrix exponentials (notably developed in [37]) we can easily get :

$$\exp\left[h \begin{pmatrix} M & -2Q^T Q \\ 0 & -M^T \end{pmatrix}\right] = \begin{pmatrix} e^{hM} & -2e^{hM}\int_0^h e^{-sM} Q^T Q e^{-sM^T} \\ 0 & e^{-hM^T} \end{pmatrix}. \tag{17}$$

After some calculations, we finally get :

$$F(h) = (I_n - 2i\Theta\Sigma(h))e^{-hM^T}$$

$$G(h) = i\Theta e^{hM},$$

and therefore:

$$B(h) = ie^{hM^T}(I_n - 2i\Theta\Sigma(h))^{-1}\Theta e^{hM}$$

$$= ie^{hM^T}\Theta(I_n - 2i\Sigma(h)\Theta)^{-1}e^{hM}.$$

Then, instead of solving the equation for $\mu(h)$ we use the following remark. From (16) we obtain:

$$G(h) = -\frac{1}{2}\left(F'(h) + F(h)M^T\right)\left(Q^TQ\right)^{-1},$$

and plugging into (14) and using the properties of the trace we deduce

$$\mu'(h) = -\frac{\beta}{2}\operatorname{Tr}\left(F(h)^{-1}F'(h) + M^T\right).$$

We can easily integrate to get:

$$\mu(h) = -\frac{\beta}{2}\left(h\operatorname{Tr}(M) + \operatorname{Tr}\left[\log F(h)\right]\right).$$

The logarithm of the complex matrix $F(h)$ is well-defined since by definition $F(h) \in GL_n(\mathbb{C})$, and using properties of the matrix logarithm we get:

$$\operatorname{Tr}(\log F(h)) = -h\operatorname{Tr}(M^T) + \operatorname{Tr}\left[\log(I_n - 2i\Theta\Sigma(h))\right]$$

$$= -h\operatorname{Tr}(M) + \log\left[\det(I_n - 2i\Theta\Sigma(h))\right].$$

Thus

$$\mu(h) = -\frac{\beta}{2}\log\left[\det(I_n - 2i\Theta\Sigma(h))\right],$$

which ends the proof □

Remark It is rather evident that for the sake of efficiency, it is less time-consuming to compute $B(h)$ using the matrix exponential form than the integral one, since the latter is only semi-closed. The matrix exponential form will then be preferred when implementing the model.

In addition to having obtained the characteristic function of the Wishart process, we will see in the next section that the particular form obtained allows us to identify the conditional law of the process, namely the non-central Wishart distribution.

4.6 Wishart processes and Wishart Distributions

4.6.1 Definition of the non-central Wishart Distribution

Whereas central Wishart distributions have been extensively studied in the literature, the related non-central distributions have received less attention. We will here recall some of the interesting (for our purpose) facts about those distributions. For a more comprehensive study, see the recent papers by Letac et al. ([35] and [43]) or the pioneer works of Anderson ([7]) and Anderson ([8]).

Definition Consider $k \geq n$ independent random vectors $X_1, ..., X_k$ of $\mathbb{R}^n$ distributed according to multivariate Gaussian distributions $\mathcal{N}(0, \Sigma)$. Consider also k non-random vectors of $\mathbb{R}^n$ denoted by $\mu_1, ..., \mu_k$. The distribution of

$$W = \sum_{i=1}^{k} (X_i + \mu_i)(X_i + \mu_i)^T, \tag{18}$$

is a non-central Wishart distribution with k degrees of freedom, denoted as $\mathcal{W}_k(\mu, \Sigma)$ where $\mu = \sum_{i=1}^{k} \mu_i \mu_i^T$.

4.6.2 Laplace Transform and Moments

By standard calculations, the Laplace Transform of the Wishart distribution can be easily computed.

Proposition 5. *Let $W \sim \mathcal{W}_k(\mu, \Sigma)$ and let $\Theta \in S_n$ so that $I_n + 2\Sigma S \in \widetilde{S}_n^+$,then*

$$\mathbb{E}\left[e^{-\operatorname{Tr}(\Theta W)}\right] = \frac{e^{-\operatorname{Tr}\left[\Theta(I_n+2\Sigma\Theta)^{-1}\mu\right]}}{\left[\det(I_n + 2\Sigma\Theta)\right]^{k/2}}. \tag{19}$$

That result allows to extend the definition of Wishart distribution to real values of k, as long as $k > n - 1$ by defining it as the unique probability measure on the cone of symmetric positive matrix which Laplace Transform is given by (19).

Then the moments of the non-central Wishart distribution can be classically obtained by differentiating the Laplace transform. After some tedious calculations and the use of results of [35] and [43], we can find that:

Proposition 6. *Let $W \sim \mathcal{W}_k(\mu, \Sigma)$. Then we have*

$\forall (i, j, p, l) \in [\![1, n]\!]^4$

$$\mathbb{E}(W^{ij}W^{pl}) = \Sigma^{ij}\Sigma^{pl}k^2 + \left(\Sigma^{il}\Sigma^{pj} + \Sigma^{ip}\Sigma^{jl} + \mu^{ij}\Sigma^{pl} + \mu^{pl}\Sigma^{ij}\right)k$$
$$+ \mu^{ij}\mu^{pl} + \mu^{il}\Sigma^{pj} + \mu^{ip}\Sigma^{jl} + \mu^{jl}\Sigma^{ip} + \mu^{pj}\Sigma^{il} \tag{20}$$

$$\mathbb{E}(W) = \mu + k\Sigma. \tag{21}$$

As a corollary we can then get the Alam and Mitra formula (see [3])

Proposition 7. *Let $W \sim \mathcal{W}_k(\mu, \Sigma)$ and note $m = \mu + k\Sigma$ its expectation matrix. Then we have*

$$\mathbb{E}\left[(W-m)^2\right] = \frac{1}{k}\left(m^2 + m\,\mathrm{Tr}(m) - \mu^2 - \mu\,\mathrm{Tr}(\mu)\right). \tag{22}$$

Proof. Let $A = \mathbb{E}\left[(W-m)^2\right]$. Using proposition 6 we have :

$$\begin{aligned} A^{ij} &= \sum_{l=1}^{n} \mathbb{E}\left(W^{il}W^{lj}\right) - \sum_{l=1}^{n}\left(\mu^{il} + k\Sigma^{il}\right)\left(\mu^{lj} + k\Sigma^{lj}\right) \\ &= k\left(\Sigma^{ij}\,\mathrm{Tr}(\Sigma) + \left(\Sigma^2\right)^{ij}\right) + (\mu\Sigma)^{ij} + (\Sigma\mu)^{ij} + \mu^{ij}\,\mathrm{Tr}(\Sigma) + \Sigma^{ij}\,\mathrm{Tr}(\mu). \end{aligned}$$

Thus we have

$$\begin{aligned} \mathbb{E}\left[(W-m)^2\right] &= k\left(\Sigma\,\mathrm{Tr}(\Sigma) + \Sigma^2\right) + \mu\Sigma + \Sigma\mu + \mu\,\mathrm{Tr}(\Sigma) + \Sigma\,\mathrm{Tr}(\mu) \\ &= \frac{1}{k}\left((\mu + k\Sigma)^2 - \mu^2 + (\mu + k\Sigma)\,\mathrm{Tr}\left(\mu + k\Sigma\right) - \mu\,\mathrm{Tr}(\mu)\right) \\ &= \frac{1}{k}\left(m^2 + m\,\mathrm{Tr}(m) - \mu^2 - \mu\,\mathrm{Tr}(\mu)\right). \end{aligned}$$

□

4.6.3 Simulation Considerations

Since our need to simulate non-central Wishart random variables will arise in the coming sections we will address it right now. Given that the simulation in the non-central case is linked to the simulation in the central case, we start by giving the answer in this case.

4.6.3.a The Central Case

Several authors have searched efficient ways to simulate central Wishart random variables. Indeed, at least in the case of an integer number of degrees of freedom β, a naive way to simulate such a random variable would be to simulate β vectorial Gaussian variables with dimension n, which could become highly time consuming for big matrix. The idea, developed notably by Mahalanobis et al. in [40], Wisjman in [46] and Kshirsagar in [34], is to use the Bartlett decomposition of a Wishart matrix and realize that it has a rather simple form. They all manage to get the following theorem, but for their proof to be valid, β has to be an integer. We will provide a proof valid in all cases in Appendix A

Theorem 3. *Let Σ be a symmetric positive definite matrix, let β be an integer strictly greater than n and let $A \in M_n(\mathbb{R})$ defined as follows*

$$\forall\ i \in [\![1, n]\!],\ A^{ii} = \chi_{\beta-i+1}$$

$$\forall\ 1 \leq j < i \leq n,\ A^{ij} = N^{ij},$$

where the $\chi_{\beta-i+1}$ denote the square roots of independent χ^2 random variables with $\beta-i+1$ degrees of freedom, the N^{ij} are independent ordinary Gaussian random variables (also independent of the χ^2) and where all other entries of A are zero.

Then the random matrix $\sqrt{\Sigma}AA^T\sqrt{\Sigma}$ has the law of $\mathcal{W}_\beta(0, \Sigma)$.

Proof. See Appendix A. □

Remarks

- If β is an integer, it is easy to define central Wishart random variables even if $\beta \leq n-1$ thanks to the formula (18). For those variables, the method above does not work, but the Gaussian interpretation allows us to simulate them.
- For the simulation of the chi-square law, see appendix B.

4.6.3.b The non-central Case

An easy way to simulate non-central Wishart random variables has been proposed by Gleser in [25]. It is based on the fact that a non-central Wishart random matrix can be written as the sum of a non-central Wishart with n degrees of freedom and a central Wishart with $\beta - n$ degrees of freedom. There are no problems here as long as β is an integer according to the remark above, but if it is not the case, then we need to have $\beta > 2n-1$ for the method to work. Unfortunately, we were not able to find an easy way to simulate non-central Wishart matrix in all cases. That being said, let us present explicitly the method.

Proposition 8. *Let β satisfying the conditions above and let M and Σ be symmetric positive matrix. Let $(m_i)_{1\leq i\leq n}$ be the i^{th} column of $\sqrt{M}$. Let now the $(Z_i)_{1\leq i\leq n}$ be independent vectorial Gaussian variables $\sim \mathcal{N}(0, I_n)$ and V be a central Wishart random variable $\sim \mathcal{W}_{\beta-n}(0, I_n)$ independent of the $(Z_i)_{1\leq i\leq n}$. We have*

$$\sum_{i=1}^{n} \left(m_i + \sqrt{\Sigma}Z_i\right)\left(m_i + \sqrt{\Sigma}Z_i\right)^T + \sqrt{\Sigma}V\sqrt{\Sigma} \sim \mathcal{W}_\beta(M, \Sigma). \quad (23)$$

Proof. Since it is clear that

$$\sum_{i=1}^{n} m_i m_i^T = M,$$

we have that

$$\sum_{i=1}^{n} \left(m_i + \sqrt{\Sigma} Z_i\right) \left(m_i + \sqrt{\Sigma} Z_i\right)^T \sim \mathcal{W}_n(M, \Sigma).$$

Then, using the Laplace transform of a Wishart random variable, it is easy to show that

$$\sum_{i=1}^{n} \left(m_i + \sqrt{\Sigma} Z_i\right) \left(m_i + \sqrt{\Sigma} Z_i\right)^T + \sqrt{\Sigma} V \sqrt{\Sigma} \sim \mathcal{W}_\beta(M, \Sigma).$$

which ends the proof. □

Remark Of course, it is possible in the formula above to replace $\sqrt{\Sigma}$ by, for example, a Cholesky factorization of the symmetric positive matrix Σ, which computation is less time-consuming (see [26] or [32] for details about Cholesky factorizations).

4.6.4 Relation to Wishart Processes

By comparing (7) and (19) we have the following theorem

Theorem 4. *If the process $(V_t)_{t\geq 0}$ has the dynamic (1) then conditionally on V_t, V_{t+h} has the distribution $\mathcal{W}_\beta\left(\Delta(h)V_t\Delta(h)^T, \Sigma(h)\right)$ where*

$$\Delta(h) = e^{hM}$$

$$\Sigma(h) = \int_0^h e^{sM} Q^T Q e^{sM^T} ds$$

As an immediate consequence of theorem 4 and propositions 6 and 7, we have now access to the conditional moments of the Wishart process.

Proposition 9. *If the process $(V_t)_{t\geq 0}$ has the dynamic (1) then we have $\forall (i,j,k,l) \in [\![1,n]\!]^4$*

$$\begin{aligned}
\mathbb{E}(V_{t+h}^{ij} V_{t+h}^{kl} | V_t) &= \mu_t(h)^{ij}\mu_t(h)^{kl} + \mu_t(h)^{il}\Sigma(h)^{kj} + \mu_t(h)^{ik}\Sigma(h)^{jl} \\
&+ \mu_t(h)^{jl}\Sigma(h)^{ik} + \mu_t(h)^{kj}\Sigma(h)^{il} + \Sigma(h)^{ij}\Sigma(h)^{kl}\beta^2 \\
&+ \left(\Sigma(h)^{il}\Sigma(h)^{kj} + \Sigma(h)^{ik}\Sigma(h)^{jl} + \mu_t(h)^{ij}\Sigma(h)^{kl} + \mu_t(h)^{kl}\Sigma(h)^{ij}\right)\beta
\end{aligned}$$

$$\beta \mathbb{V}ar\left[V_{t+h}|V_t\right] = \left(\mu_t(h) + \beta\Sigma(h)\right)^2 - \mu_t(h)^2 + \left(\mu_t(h) + \beta\Sigma(h)\right)\mathrm{Tr}\left(\mu_t(h) + \beta\Sigma(h)\right)$$
$$- \mu_t(h)\,\mathrm{Tr}(\mu_t(h))$$

$$\mathbb{E}\left[V_{t+h}|V_t\right] = \mu_t(h) + \beta\Sigma(h),$$

where

$$\mu_t(h) = e^{hM} V_t e^{hM^T}$$

$$\Sigma(h) = \int_0^h e^{sM} Q^T Q e^{sM^T} ds$$

$$\mathbb{V}ar\left[V_{t+h}|V_t\right] = \mathbb{E}\left[\left(V_{t+h} - \mathbb{E}\left[V_{t+h}|V_t\right]\right)^2 |V_t\right]$$

5 Option valuation in the Wishart Model: the characteristic function method

Following previous works by Carr, Madan, Lewis and Lee in particular (see [11], [47], [36] or even [42]), it is now well known that the price of a European option can be easily calculated via Fourier integrals, provided that the characteristic function of the log-spot price has a closed form.

5.1 The Characteristic Function

Just as the Heston model or the Double Heston model, the Wishart model belongs to the larger class of affine models, for which the computation of the characteristic function is rather straightforward. Let us first note $X_t = \log S_t$ and:

$$\Psi(k, t, \tau) = \mathbb{E}\left[e^{ikX_{t+\tau}}|X_t, V_t\right] \tag{24}$$

The dynamic of X_t is given by:

$$dX_t = \left(r - \frac{\mathrm{Tr}(V_t)}{2}\right) dt + \mathrm{Tr}\left(\sqrt{V_t} dZ_t\right). \tag{25}$$

According to the Feynman-Kac theorem, we know that the function Ψ is the solution of the following PDE:

$$\begin{cases} \dfrac{\partial f}{\partial \tau} &= \mathcal{L}_{X,V} f \\ f(k, t, \tau) &= e^{ikX_t}, \end{cases}$$

where $\mathcal{L}_{X,V}$ is the infinitesimal generator of the diffusion of the process (X_t, V_t).

The following proposition of [21] gives the infinitesimal generator

Proposition 10. *The infinitesimal generator of (X_t, V_t) is given by*

$$\mathcal{L}_{X,V} = \left(r - \frac{1}{2}\operatorname{Tr}(V)\right)\frac{\partial}{\partial x} + \frac{1}{2}\operatorname{Tr}(V)\frac{\partial^2}{\partial x^2} + \mathcal{L}^V + 2\operatorname{Tr}\left[VR^TQD\right]\frac{\partial}{\partial x}, \tag{26}$$

where $\mathcal{L}^V$ and D are given in proposition 3.

Proof. See Appendix A □

It follows that (26) can be rewritten explicitly

$$\begin{aligned}\frac{\partial\Psi(k,t,\tau)}{\partial\tau} &= \left(r - \frac{1}{2}\operatorname{Tr}(V)\right)\frac{\partial\Psi(k,t,\tau)}{\partial x} + \frac{1}{2}\operatorname{Tr}(V)\frac{\partial^2\Psi(k,t,\tau)}{\partial x^2}\\ &+ \mathcal{L}^V\Psi(k,t,\tau) + 2\operatorname{Tr}\left[VR^TQD\right]\frac{\partial\Psi(k,t,\tau)}{\partial x}.\end{aligned} \tag{27}$$

Since we have already seen that the characteristic function of the Wishart process is exponentially affine, we guess that the function $\Psi(k,t,x)$ will be the exponential of an affine combination of X and the terms of the Wishart matrix. That is to say

$$\Psi(k,t,\tau) = e^{\operatorname{Tr}(A(\tau)V_t) + b(\tau)X_t + c(\tau)}, \tag{28}$$

where

$$\begin{aligned}&A : \mathbb{R}^+ \to M_n(\mathbb{C}),\ A(0) = 0\\ &b : \mathbb{R}^+ \to \mathbb{C},\ b(0) = ik\\ &c : \mathbb{R}^+ \to \mathbb{C},\ c(0) = 0.\end{aligned}$$

Then replacing (28) in (27) leads to

$$\begin{aligned}\operatorname{Tr}\left(\frac{\partial A(\tau)}{\partial\tau}V\right) + \frac{\partial b(\tau)}{\partial\tau}X + \frac{\partial c(\tau)}{\partial\tau} &= \left(r - \frac{1}{2}\operatorname{Tr}(V)\right)b(\tau) + \frac{1}{2}\operatorname{Tr}(V)b^2(\tau)\\ &+ 2\operatorname{Tr}\left[VR^TQA(\tau)\right]b(\tau) + 2VA(\tau)Q^TQA(\tau)\\ &+ \operatorname{Tr}\left[\left(\beta Q^TQ + MV + VM^T\right)A(\tau)\right].\end{aligned}$$

By identifying the coefficients of X and V, we finally obtain the following system

$$\begin{aligned}
\mathrm{Tr}\left(\frac{\partial A(\tau)}{\partial \tau}\right) &= \mathrm{Tr}\left[A(\tau)M + M^T A(\tau) + 2A(\tau)Q^T QA(\tau)\right] \\
&+ \mathrm{Tr}\left[2R^T QA(\tau)b(\tau) + \frac{b(\tau)(b(\tau)-1)}{2} I_n\right] \qquad (29) \\
\frac{\partial c(\tau)}{\partial \tau} &= \mathrm{Tr}\left[\beta Q^T QA(\tau)\right] + b(\tau)r \qquad (30) \\
\frac{\partial b(\tau)}{\partial \tau} &= 0. \qquad (31)
\end{aligned}$$

The complete resolution has been made by Da Fonseca et al. in [21]. They give the following solution

Proposition 11. *The characteristic function of the log-price in the Wishart model is given by*

$$\mathbb{E}\left[e^{ikX_{t+\tau}}|X_t, V_t\right] = e^{\mathrm{Tr}(A(\tau)V_t) + b(\tau)X_t + c(\tau)},$$

with

$$\begin{aligned}
A(\tau) &= A_{22}(\tau)^{-1}A_{21}(\tau) \\
c(\tau) &= -\frac{\beta}{2}\left(\log\left[\det\left(A_{22}(\tau)\right)\right] + \mathrm{Tr}\left[M^T + 2ikR^T Q\right]\tau\right) + ikr\tau \\
b(\tau) &= ik,
\end{aligned}$$

where

$$\begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} = \exp\ \tau \begin{pmatrix} M & -2Q^T Q \\ -\frac{k(k+i)}{2} I_n & -\left(M^T + 2ikR^T Q\right) \end{pmatrix}$$

Proof. See Appendix A □

Remarks

- The formulation above is very interesting for the purpose of computation, since we have avoided to deal with numerical integration.

- The formula giving $c(\tau)$ involves complex logarithms. The readers familiar with the Heston model know that if precautions are no taken, those logarithms can lead to non-continuous characteristic functions and therefore badly priced derivatives. Even though there are ways to avoid this issue with the Heston and Double Heston models (see [38] and [23] for example), we did not manage to solve the problem without having to keep track of the values of the complex logarithm and switch the branch used in order to maintain continuity. It is rather easy when using a numerical integration technique where the points of evaluation are in increasing order, that is the reason why we decided to use Gauss-Legendre integration (see [1]) when computing European options prices with the inverse Fourier Transform method. As can be seen on the graphic below, the continuity problems are a real issue, even though it seems that they only arise when considering long maturities.

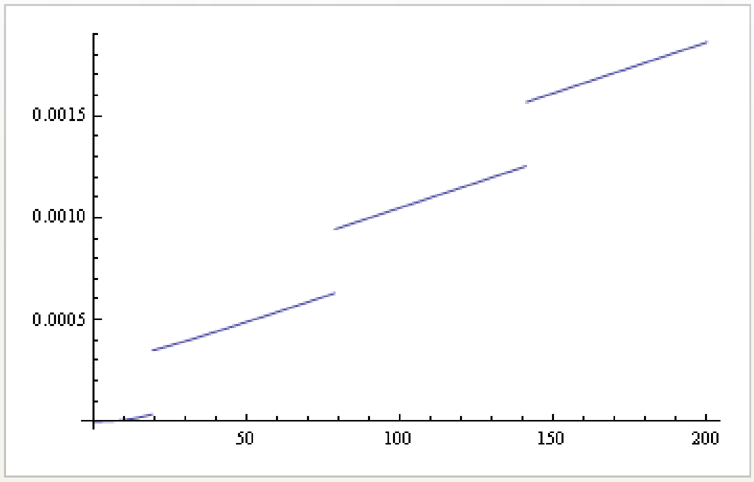


Figure 1: Imaginary part of $c(\tau)$

5.2 The Call Option Price

We know, following Lewis and Lee ([47] and [36]), that the price of a call option of strike K and maturity T is given by:

$$C(K,T) = S_0 - \frac{1}{\pi}\sqrt{K}e^{-\frac{rT}{2}} \int_0^{\infty} \mathcal{R}e\left[e^{iuk}\Psi\left(u - \frac{i}{2}\right)\right] \frac{du}{u^2 + \frac{1}{4}} \tag{32}$$

where $k = rT - \log K$

Given that we have obtained a closed-form formula for the characteristic function involved in the integration, the only thing remaining in order to numerically evaluate the option price is to compute the integral.

Our formula is different from that proposed by Da Fonseca et al. [21]. Readers which are familiar with the characteristic function method are already aware of this recurrent issue. The original call formula requires a Fourier inversion technique which often proves very time consuming, while ours, which is exactly that of Lewis [36], will only use numerical integration routines which are faster. Thus, we gain not only on computational time but also on numerical precision.

5.3 Improving the closed formula: the control variate method

The Call Option Case When we consider options which are deep-in-the-money, the vega is usually very small. Therefore it is necessary to have a very efficient pricing formula, if one wants to be able to calibrate the model to the market. For those extreme strikes, the current pricing formula does not allow to calibrate to the market implied volatility, since it calculates the option price in the money with a too large error. Therefore, we tried to improve it by using a control variate.
The idea here is to use the option price when the volatility matrix is deterministic (i.e $Q = 0$). In this case we're in a simple Black-Scholes framework, and the price of the option can be computed as:

$$\widetilde{C}(K,T) = S_0\mathcal{N}(\delta_1) - Ke^{-rT}\mathcal{N}(\delta_0) \tag{33}$$

where

$$\delta_0 = \frac{1}{\sigma_T\sqrt{T}}\log\left(\frac{S_0}{Ke^{-rT}}\right) - \frac{1}{2}\sigma_T\sqrt{T}$$

$$\delta_1 = \delta_0 + \sigma_T\sqrt{T}$$

$$\sigma_T = \sqrt{\frac{1}{T}\int_0^T \mathrm{Tr}\left(V_t\right)dt}.$$

We can get a closed-form solution for σ_T by noticing that we now have

$$dV_t = \left(MV_t + V_tM^T\right)dt,$$

which is easily solved by

$$V_t = e^{tM}V_0e^{tM^T}.$$

Thus we have

$$\sigma_T^2 = \frac{1}{T}\operatorname{Tr}\left(\int_0^T V_t dt\right)$$
$$= \frac{1}{T}\operatorname{Tr}\left(\int_0^T e^{tM} V_0 e^{tM^T} dt\right).$$

In the (very) particular case where $M \in \widetilde{S}_n^-$ the integration is easy and leads to

$$\int_0^T \operatorname{Tr}\left(e^{tM} V_0 e^{tM^T}\right) dt = \operatorname{Tr}\left(\int_0^T e^{2tM} dt V_0\right)$$
$$= \operatorname{Tr}\left[V_0 \left(2M\right)^{-1}\left(e^{2TM} - I_n\right)\right].$$

In the general case the formula above is no longer applicable, but using techniques from [37] we also have

$$\int_0^T e^{tM} V_0 e^{tM^T} dt = F_1(T)^T G(T),$$

where

$$\begin{pmatrix} F_0(T) & G(T) \\ 0 & F_1(T) \end{pmatrix} = \exp\ T \begin{pmatrix} -M & V_0 \\ 0 & M^T \end{pmatrix}$$

Therefore in all case we have a closed-form for σ_T.

Besides, we can apply the characteristic option formula. Putting $Q = 0$ in (27) we obtain the following PDE for the characteristic function $\widetilde{\Psi}(k, t, \tau)$ in this case:

$$\frac{\partial \widetilde{\Psi}(k,t,\tau)}{\partial \tau} = \left(r - \frac{1}{2}\operatorname{Tr}(V)\right)\frac{\partial \widetilde{\Psi}(k,t,\tau)}{\partial x} + \frac{1}{2}\operatorname{Tr}(V)\frac{\partial^2 \widetilde{\Psi}(k,t,\tau)}{\partial x^2} + \mathcal{L}^V \widetilde{\Psi}(k,t,\tau).$$

Then we assume again that

$$\widetilde{\Psi}(k,t,\tau) = e^{\operatorname{Tr}\left(\widetilde{A}(\tau)V_t\right) + \widetilde{b}(\tau)X_t + \widetilde{c}(\tau)},$$

Their resolution is identical to the one already done and leads to

$$\widetilde{A}(\tau) = A_{22}(\tau)^{-1} A_{21}(\tau)$$
$$\widetilde{c}(\tau) = 0$$
$$\widetilde{b}(\tau) = ik,$$

where

$$\begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} = \exp\ \tau \begin{pmatrix} M & 0 \\ -\frac{k(k+i)}{2} I_n & -M^T \end{pmatrix}$$

Therefore, the price of the option in this deterministic case can also be given by:

$$\widetilde{C}(K,T) = S_0 - \frac{1}{\pi}\sqrt{S_0 K} e^{-\frac{rT}{2}} \int_0^\infty \mathcal{R}e \left[e^{iuk} \widetilde{\Psi} \left(u - \frac{i}{2} \right) \right] \frac{du}{u^2 + \frac{1}{4}} \tag{34}$$

Thus we have a new formulation for the call option price in the Wishart Model, using control variate:

$$C(K,T) = S_0 \mathcal{N}(\delta_1) - K e^{-rT} \mathcal{N}(\delta_0) - \frac{1}{\pi}\sqrt{S_0 K} e^{-\frac{rT}{2}} \int_0^\infty \mathcal{R}e \left[e^{iuk} \left(\Psi - \widetilde{\Psi} \right) \left(u - \frac{i}{2} \right) \right] \frac{du}{u^2 + \frac{1}{4}}$$

The presence of $\widetilde{\Psi}$ in the integral reduces the integrand, preventing the numerical evaluation with the Gauss-Legendre quadrature method to create large errors with extreme strikes.

6 Efficient discretization

6.1 The Price Process

6.1.1 The predictor-corrector Euler scheme

Derivation of the scheme Rather than diffusing the logarithm of the spot $\ln S_t$, it may be interesting to consider the logarithm of the discounted spot $\ln\left(e^{-rt} S_t\right)$. By applying Itō's formula, we remark that the $r_t dt$ terms cancel out. For the discounted asset price, we choose to apply the Predictor-Corrector scheme. This discretization scheme is known to provide better results than the Euler scheme at a low additional effort.
The idea is to start from an exact expression of the logarithm of the discounted price. By Itō's formula we have:

$$\begin{aligned} d\left(\ln\left(e^{-rt} S_t\right)\right) &= -r dt + \frac{dS_t}{S_t} - \frac{1}{2 S_t^2} \langle dS, dS \rangle_t \\ &= -\frac{\mathrm{Tr}(V_t)}{2} dt + \sqrt{V_t} \left(dW_t R + dB_t \sqrt{I_n - R^T R} \right). \end{aligned}$$

Then a simple integration on an interval $[t, t + \Delta t]$ leads to:

$$e^{-r(t+\Delta t)}S_{t+\Delta t} = e^{-rt}S_t - \frac{1}{2}\int_t^{t+\Delta t} \mathrm{Tr}\,(V_s)\,ds + \mathrm{Tr}\left(\int_t^{t+\Delta t} \sqrt{V_s}dW_sR\right)$$

$$+\,\mathrm{Tr}\left(\int_t^{t+\Delta t} \sqrt{V_s}dB_s\sqrt{I_n - R^TR}\right) \tag{35}$$

We now have to approximate the integrals in the above formula. The idea of the predictor corrector scheme is first to handle the time integral in a centered manner, that is to say:

$$\int_t^{t+\Delta t} \mathrm{Tr}(V_s)ds \approx \frac{\Delta t}{2}\,\mathrm{Tr}\,(V_{t+\Delta t} + V_t)\,.$$

Then, given that B is independent of V, the stochastic integral with respect to B in the formula above is (conditionally) Gaussian with zero mean and variance

$$\mathrm{Tr}\left(\int_t^{t+\Delta t} V_sds\left(I_n - R^TR\right)\right),$$

which can be approximated as previously.

Finally, we have to deal with the stochastic integral against the Brownian W. Unfortunately, we cannot simulate it by integrating the dynamic of the variance process as in the Heston and Double Heston models. For the time being we will then use the usual Euler approximation

$$\mathrm{Tr}\left(\int_t^{t+\Delta t} \sqrt{V_s}dW_sR\right) \approx \sqrt{\Delta t}\,\mathrm{Tr}\left(\sqrt{V_t}GR\right),$$

where G is a matrix which elements are independent Gaussian random variables.

Finally, we obtain the following scheme

$$\log\left(\frac{e^{-r(t+\Delta t)}\widehat{S}_{t+\Delta t}}{e^{-rt}\widehat{S}_t}\right) = \sqrt{\frac{\Delta t}{2}}\sqrt{\mathrm{Tr}\left[\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right)(I_n - R^TR)\right]}Z$$

$$+\sqrt{\Delta t}\,\mathrm{Tr}\left[\sqrt{\widehat{V}_t}GR\right] - \frac{\Delta t}{4}\,\mathrm{Tr}\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right), \tag{36}$$

where Z is a Gaussian random variable independent of G.

Martingale Correction As noted by Andersen in [5] and Andersen and Piterbarg in [6] the discretized scheme for the discounted price may not always be arbitrage free. That is to say that even though the continuous-time process of discounted price is a martingale, the scheme defined by (36) is not. Even though the practical relevance of this is often minor, it is interesting in our Wishart framework to see how the results of Andersen may be generalized. Therefore, let us examine whether it is possible to modify the scheme (36) to enforce that

$$\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}\middle|\,\widehat{S}_t\right] = \widehat{S}_t. \tag{37}$$

By iterated conditional expectations and by taking the exponential in (36), we have:

$$\begin{aligned}
\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}|\widehat{S}_t\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}\middle|\,\widehat{S}_t, \widehat{V}_{t+\Delta t}\right]\right] \\
&= \widehat{S}_t e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_{t+\Delta t}\right)}\times\right. \\
&\left.\mathbb{E}_t\left[e^{\sqrt{\Delta t}\operatorname{Tr}\left(\sqrt{\widehat{V}_t}GR\right)+\sqrt{\frac{\Delta t}{2}}\sqrt{\operatorname{Tr}\left(\left(\widehat{V}_t+\widehat{V}_{t+\Delta t}\right)(I_n-R^TR)\right)}Z}\middle|\,\widehat{V}_{t+\Delta t}\right]\right] \\
&= \widehat{S}_t e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_{t+\Delta t}\right)}\mathbb{E}_t\left[e^{\sqrt{\Delta t}\operatorname{Tr}\left(\sqrt{\widehat{V}_t}GR\right)}\middle|\,\widehat{V}_{t+\Delta t}\right]\right. \\
&\left.\mathbb{E}_t\left[e^{\sqrt{\frac{\Delta t}{2}}\sqrt{\operatorname{Tr}\left(\left(\widehat{V}_t+\widehat{V}_{t+\Delta t}\right)(I_n-R^TR)\right)}Z}\middle|\,\widehat{V}_{t+\Delta t}\right]\right] \\
&= \widehat{S}_t e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_{t+\Delta t}\right)+\frac{\Delta t}{2}\operatorname{Tr}\left(R^TR\widehat{V}_t\right)}\times\right. \\
&\left.e^{\frac{\Delta t}{4}\operatorname{Tr}\left(\left(\widehat{V}_t+\widehat{V}_{t+\Delta t}\right)(I_n-R^TR)\right)}\middle|\,\widehat{V}_t\right] \\
&= \widehat{S}_t e^{\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_{t+\Delta t}\right)}\middle|\,\widehat{V}_t\right],
\end{aligned}$$

where we used standard results on Gaussian random variables and the fact that G and Z are independent.

Let us then note $M = \mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_{t+\Delta t}\right)}\middle|\,\widehat{V}_t\right]$.

If we were to modify the scheme (36) as follows

$$
\begin{aligned}
\log\left(\frac{e^{-r(t+\Delta t)}\widehat{S}_{t+\Delta t}}{e^{-rt}\widehat{S}_t}\right) &= -\log M - \frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\left(I_n + R^T R\right) + \widehat{V}_{t+\Delta t}\right) \\
&+ \sqrt{\frac{\Delta t}{2}}\sqrt{\operatorname{Tr}\left[\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right)\left(I_n - R^T R\right)\right]}Z \\
&+ \sqrt{\Delta t}\operatorname{Tr}\left[\sqrt{\widehat{V}_t}GR\right], \qquad (38)
\end{aligned}
$$

then the martingale condition would be clearly verified.

Of course, for everything we just said to be meaningful, the quantity M has to be finite. In most of the variance discretization that we will consider (the exception being the full truncation Euler-Scheme), the variance process remains in the cone of positive matrix (even if we only have $\beta > n-1$). Therefore in those cases, since the matrix $R^T R$ is also positive, we have (see proposition 12)

$$
\operatorname{Tr}\left(R^T R V_t\right) \geq 0,
$$

and thus the quantity which expectation is taken in the definition of M is positive and bounded by 1, hence $M < \infty$. The martingale correction is then always possible, as long as the variance process remains a positive matrix.

6.1.2 Ornstein-Uhlenbeck Scheme

6.1.2.a When β is an integer

Derivation of the scheme In the particular case where β is an integer, we have already seen in proposition 1 that a Wishart process could be written as a sum of vectorial Ornstein-Uhlenbeck processes which dynamic is given by :

$$
dX_{k,t} = MX_{k,t}dt + Q^T dW_{k,t},
$$

where $\{W_{k,t}, t \geq 0, 1 \leq k \leq \beta\}$ are independent vectorial Brownian motions and where we have

$$
\sqrt{V_t}dW_t = \sum_{k=1}^{\beta} X_{k,t}dW_{k,t}^T.
$$

Therefore, we can modify the previous Euler Scheme with the idea to diffuse the processes $X_{k,t}$ instead of V_t, so that we would not have to deal with matrix square roots anymore. A standard discretization for the dynamic of the $X_{k,t}$ can be written

$$\widehat{X}_{k,t+\Delta t} = \widehat{X}_{k,t} + \Delta t M X_{k,t} + \sqrt{\Delta t} Q^T \epsilon_{k,t+\Delta t}, \tag{39}$$

where the $(\epsilon_{k,t+\Delta t})_{1 \leq k \leq \beta}$ are standard vectorial Gaussian random variables, independent of the corresponding Gaussians at the other times of discretization.

We now have another approximation for the stochastic integral against the matrix Brownian motion W_t

$$\int_t^{t+\Delta t} \sqrt{V_s} dW_s = \sum_{k=1}^{\beta} \int_t^{t+\Delta t} X_{k,s} dW_{k,s}^T$$

$$\approx \sqrt{\Delta t} \sum_{k=1}^{\beta} X_{k,t} \epsilon_{k,t+\Delta t}^T.$$

Hence the following scheme, which already integrates the discretization of the variance process

$$\log\left(\frac{e^{-r(t+\Delta t)}\widehat{S}_{t+\Delta t}}{e^{-rt}\widehat{S}_t}\right) = \sqrt{\Delta t}\,\mathrm{Tr}\left[\sum_{k=1}^{\beta} \widehat{X}_{k,t} \epsilon_{k,t+\Delta t}^T R\right] - \frac{\Delta t}{4}\,\mathrm{Tr}\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right)$$

$$+ \sqrt{\frac{\Delta t}{2}} \sqrt{\mathrm{Tr}\left[\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right)(I_n - R^T R)\right]} Z$$

$$\widehat{X}_{k,t+\Delta t} = \widehat{X}_{k,t} + \Delta t M X_{k,t} + \sqrt{\Delta t} Q^T \epsilon_{k,t+\Delta t}$$

$$\widehat{V}_t = \sum_{k=1}^{\beta} \widehat{X}_{k,t} \widehat{X}_{k,t}^T, \tag{40}$$

with the same notations as in the previous section.

Remains the initialization of the scheme. In this regard, since $V_0 \in S_n^+$, it can be diagonalized in an orthonormal basis, that is to say

$$V_0 = \sum_{k=1}^{n} \lambda_k \phi_k \phi_k^T,$$

where the λ_k are the eigenvalues of V_0 and the ϕ_k the corresponding eigenvectors.

Then it suffices to set

$$X_{0,k} = \mathbb{1}_{k \leq n} \sqrt{\lambda_k} \phi_k.$$

Martingale Correction The calculations are quite similar to those of the previous Euler scheme and lead to

$$
\begin{aligned}
\mathbb{E}\left[e^{-r\Delta t}\widehat{S}_{t+\Delta t}\middle|\,\widehat{S}_t\right] &= \widehat{S}_t e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_{t+\Delta t}\right)}\times\right.\\
&\mathbb{E}\left[e^{\sqrt{\Delta t}\operatorname{Tr}\left[\sum_{k=1}^{\beta}\widehat{X}_{k,t}\epsilon_k^T R\right]}\middle|\,\widehat{S}_t,\left(\widehat{X}_{k,t+\Delta t}\right)_{1\le k\le\beta}\right]\times\\
&\left.\mathbb{E}\left[e^{\sqrt{\frac{\Delta t}{2}}\sqrt{\operatorname{Tr}\left(\left(\widehat{V}_t+\widehat{V}_{t+\Delta t}\right)\left(I_n-R^TR\right)\right)}Z}\middle|\,\widehat{S}_t,\left(\widehat{X}_{k,t+\Delta t}\right)_{1\le k\le\beta}\right]\right]\\
&= \widehat{S}_t e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(\widehat{V}_{t+\Delta t}\right)+\frac{\Delta t}{2}\operatorname{Tr}\left(R^TR\widehat{V}_t\right)}\times\right.\\
&\left.e^{\frac{\Delta t}{4}\operatorname{Tr}\left(\left(\widehat{V}_t+\widehat{V}_{t+\Delta t}\right)\left(I_n-R^TR\right)\right)}\middle|\,\left(\widehat{X}_{k,t+\Delta t}\right)_{1\le k\le\beta}\right]\\
&= \widehat{S}_t e^{\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_t\right)}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_{t+\Delta t}\right)}\middle|\,\left(\widehat{X}_{k,t}\right)_{1\le k\le\beta}\right],
\end{aligned}
$$

where it is important to remark that the natural filtration of the $\left\{\left(\widehat{X}_{k,t}\right)_{1\le k\le\beta}, t\ge 0\right\}$ contains the natural filtration of $\left\{\widehat{V}_t,\ t\ge 0\right\}$.

Therefore, the choice that we made with the discretization scheme (39) allows us to get back on our feet and have the same martingale correction for the two schemes that we have considered so far.

Besides, in this case since we know the dynamic of the discretized variance process, we can finish the calculations. First of all, thanks to (39) we know that conditional on the $\left(\widehat{X}_{k,t}\right)_{1\le k\le\beta}$, we have

$$
\widehat{X}_{k,t+\Delta t}\sim\mathcal{N}_n\left(\left(I_n+\Delta tM\right)\widehat{X}_{k,t},\Delta tQ^TQ\right). \tag{41}
$$

Therefore, conditional on the $\left(\widehat{X}_{k,t}\right)_{1\le k\le\beta}$ (and thus on $\widehat{V}_t$), $\widehat{V}_{t+\Delta t}$ is distributed as a non-central Wishart distribution

$$
V_{t+\Delta t}\sim\mathcal{W}_\beta\left(\left(I_n+\Delta tM\right)\widehat{V}_t,\Delta tQ^TQ\right),
$$

so much so that we have with proposition 5

$$
\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^TR\widehat{V}_{t+\Delta t}\right)}\middle|\,\left(\widehat{X}_{t+\Delta t}\right)_{1\le k\le\beta}\right]=\frac{e^{-\frac{\Delta t}{4}\operatorname{Tr}\left[R^TR\left(I_n+\frac{\Delta t^2}{2}Q^TQR^TR\right)^{-1}\left(I_n+\Delta tM\right)\widehat{V}_t\right]}}{\left[\det\left(I_n+\frac{\Delta t^2}{2}Q^TQR^TR\right)\right]^{\beta/2}},
$$

which in turn leads to the following martingale corrected scheme

$$
\begin{aligned}
\log\left(\frac{e^{-r(t+\Delta t)}\widehat{S}_{t+\Delta t}}{e^{-rt}\widehat{S}_t}\right) &= \sqrt{\Delta t}\,\mathrm{Tr}\left[\sum_{k=1}^{\beta}\widehat{X}_{k,t}\epsilon_k^T R\right] - \frac{\beta}{2}\log\left[\det\left(I_n + \frac{\Delta t^2}{2}Q^TQR^TR\right)\right] \\
&+ \frac{\Delta t}{4}\,\mathrm{Tr}\left[R^TR\left(I_n + \frac{\Delta t^2}{2}Q^TQR^TR\right)^{-1}(I_n+\Delta t M)\,\widehat{V}_t\right] \\
&+ \sqrt{\frac{\Delta t}{2}}\sqrt{\mathrm{Tr}\left[\left(\widehat{V}_t + \widehat{V}_{t+\Delta t}\right)(I_n - R^TR)\right]}Z \\
&- \frac{\Delta t}{4}\,\mathrm{Tr}\left[\widehat{V}_t\left(I_n + R^TR\right) + \widehat{V}_{t+\Delta t}\right]. \qquad (42)
\end{aligned}
$$

6.1.2.b A more general case: $\beta \geq n+1$

When the parameter β is no longer an integer, the previous discretization scheme fails to work. The idea is then to change the probability measure under which we are working, so that the new dynamic of V_t would be characterized by an integer $\widetilde{\beta}$. Fortunately, equivalent of the celebrated Girsanov formula have been found in [16] as a generalization of results on Bessel and square Bessel processes. We have the following theorem

Theorem 5. *Let $\beta \geq n+1$, let V_t be the unique symmetric positive definite solution of (1), let $\nu > 0$ and let $\mathbb{Q}^*$ be the probability measure defined by its Radon-Nikodym density*

$$
\left.\frac{d\mathbb{Q}^*}{d\mathbb{Q}}\right|_{\mathcal{F}_T} = e^{-\nu\int_0^T \mathrm{Tr}\left(V_s^{-1/2}dW_sQ\right) - \frac{\nu^2}{2}\int_0^T \mathrm{Tr}\left(V_s^{-1}Q^TQ\right)ds}. \qquad (43)
$$

Then the matrix valued stochastic process W_t^* defined by

$$
W_t^* = W_t + \nu\int_0^t V_s^{-1/2}Q^T ds,
$$

is a matrix Brownian motion under the probability $\mathbb{Q}^*$, and we have

$$
dV_t = \left((\beta - 2\nu)\,Q^TQ + MV_t + V_tM^T\right)dt + \sqrt{V_t}dW_t^*Q + Q^T\left(dW_t^*\right)^T\sqrt{V_t}.
$$

Therefore if we were to choose $\nu = \frac{1}{2}(\beta - [\beta])$ then $\beta - 2\nu$ would exactly be the integer part of β and our aim would be achieved.

Proof. It suffices to apply the usual Girsanov formula to the matrix valued predictable process $H_t = \nu V_t^{-1/2}$, which is correctly defined since $\beta \geq n+1$ and therefore V_t is always positive definite. $\square$

Thus, in order to price, say, a European option with payoff $f(S_T)$, we can write its risk-neutral price

$$\begin{aligned} P_f &= \mathbb{E}\left[e^{-rT} f(S_T)\right] \\ &= \mathbb{E}^{\mathbb{Q}^*}\left[e^{-rT} e^{\nu \int_0^T \mathrm{Tr}\left(V_s^{-1/2} dW_s Q\right) + \frac{\nu^2}{2} \int_0^T \mathrm{Tr}\left(V_s^{-1} Q^T Q\right) ds} f(S_T)\right]. \end{aligned}$$

Then at each discretization step we can use the previous scheme in the integer case to simulate the price process under the probability $\mathbb{Q}^*$ (with of course the correction of the dynamic due to the probability measure change, namely we have to add a drift term $-\nu \,\mathrm{Tr}\left(Q^T R\right) dt$ in the dynamic of $d \log S_t$), and a simpler Euler scheme to approximate the Radon-Nykodym density. In this regard, it seems interesting to try and simplify the expression (43). But if we use the results of proposition 2, we can notice that

$$\int_0^T \mathrm{Tr}\left[V_t^{-1/2} dW_t Q\right] = \frac{1}{2} \log\left(\frac{\det V_T}{\det V_0}\right) - \frac{1}{2}\left((\beta - n - 1) \int_0^T \mathrm{Tr}\left[V_t^{-1} Q^T Q\right] dt + 2T \,\mathrm{Tr}\,(M)\right).$$

Therefore, we have

$$\left.\frac{d\mathbb{Q}}{d\mathbb{Q}^*}\right|_{\mathcal{F}_T} = \left(\frac{\det V_T}{\det V_0}\right)^{\frac{\nu}{2}} e^{-\nu T \,\mathrm{Tr}(M) + \frac{\nu}{2}(\nu - \beta + n + 1) \int_0^T \mathrm{Tr}\left[V_s^{-1} Q^T Q\right] ds}. \tag{44}$$

The integral in the formula above can then easily approximated at each step by the quantity

$$\int_t^{t+\Delta t} \mathrm{Tr}\left[V_s^{-1} Q^T Q\right] ds \approx \frac{\Delta t}{2} \mathrm{Tr}\left[\left(V_t^{-1} + V_{t+\Delta t}^{-1}\right) Q^T Q\right].$$

Of course, this method can be extended to any path-dependent option. Unfortunately, we did not find a solution to implement this scheme in the case $n-1 < \beta < n+1$ and $\beta \neq n$, since all the properties above require that V_t remains a positive definite matrix. In the next sections, we will explore other simulations schemes of the variance process which will be used in conjunction with the predictor-corrector Euler scheme and which could overcome this flaw of the Ornstein-Uhlenbeck scheme.

6.1.2.c Consistence of the Scheme

It is important to check whether the scheme above allows us to recover the original dynamic of the variance process, one way or another. Since we have discretized the dynamic of the $\left(\widehat{X}_{k,t}\right)_{1\leq k\leq\beta}$, we can write, limiting ourself to the first order on Δt

$$\begin{aligned}
\widehat{V}_{t+\Delta t} &= \sum_{k=1}^{\beta} \widehat{X}_{k,t+\Delta t}\widehat{X}_{k,t+\Delta t}^{T} \\
&= \sum_{k=1}^{\beta} \left(\widehat{X}_{k,t} + \Delta t M X_{k,t} + \sqrt{\Delta t} Q^{T} \epsilon_{k,t+\Delta t}\right) \left(\widehat{X}_{k,t}^{T} + \Delta t X_{k,t}^{T} M^{T} + \sqrt{\Delta t}\epsilon_{k,t+\Delta t}^{T} Q\right) \\
&= \widehat{V}_{t} + \Delta t \left(\widehat{V}_{t} M^{T} + M\widehat{V}_{t}\right) + \Delta t Q^{T} \left(\sum_{k-1}^{\beta} \epsilon_{k,t+\Delta t}\epsilon_{k,t+\Delta t}^{T}\right) Q \\
&+ \sqrt{\Delta t}\left(\sum_{k=1}^{\beta} X_{k,t}\epsilon_{k,t+\Delta t}^{T} Q + Q^{T}\sum_{k=1}^{\beta} \epsilon_{k,t+\Delta t} X_{k,t}^{T}\right) + o(\Delta t).
\end{aligned}$$

Then, according to a previous approximation, we can consider that

$$\sqrt{\Delta t}\sum_{k=1}^{\beta} X_{k,t}\epsilon_{k,t+\Delta t}^{T} \approx \sqrt{\widehat{V}_{t}}\Delta W_{t+\Delta t},$$

where W is the matrix Brownian motion of (1) and $\Delta W_{t+\Delta t} := W_{t+\Delta t} - W_{t}$.

Finally, we get (almost) the basic Euler approximation of the dynamic of V_t, with the exception of the term $\Delta t Q^{T}\sum_{k-1}^{\beta} \epsilon_{k,t+\Delta t}\epsilon_{k,t+\Delta t}^{T} Q$, but its expectation corresponds to the correct term, since it equals $\Delta t\beta Q^{T} Q$.

6.2 Variance Process

6.2.1 Full-Truncation Euler Scheme

For the variance process, the first option is to use a slightly modified naive Euler scheme. The most naive discretization scheme here would be to write

$$\widehat{V}_{t+\Delta t} = \widehat{V}_{t} + \left(\beta Q^{T} Q + M\widehat{V}_{t} + \widehat{V}_{t} M^{T}\right)\Delta t + \sqrt{\widehat{V}_{t}} G Q\sqrt{\Delta t} + Q^{T} G^{T}\sqrt{\widehat{V}_{t}}\sqrt{\Delta t}, \qquad (45)$$

where G is the same matrix of independent standard Gaussian random variables as defined in the previous section.

The fatal flaw of this scheme is that the approximated process $\widehat{V}$ is no longer bound to remain in the cone of symmetric positive matrix. For example if one looked at the diagonal terms of $\widehat{V}$ (which are necessarily positive for $\widehat{V}$ to be positive), it is clear from properties of Gaussian random variables that they can become strictly negative with a strictly positive probability. We then face the same problems that arise in the Heston model (see e.g. [13] and [17]), that is to say that the square root of the variance may be no longer defined. In order to solve this problem, several fixes have been proposed in the literature, and the one which seems to produce the smallest bias is known as the "full truncation" scheme. The idea behind is to allow the process to become negative. It can be written as

$$\widehat{V}_{t+\Delta t} = \widehat{V}_t + \left(\beta Q^T Q + M\left(\widehat{V}_t\right)^+ + \left(\widehat{V}_t\right)^+ M^T\right)\Delta t + \sqrt{\left(\widehat{V}_t\right)^+} G Q \sqrt{\Delta t}$$
$$+ Q^T G^T \sqrt{\left(\widehat{V}_t\right)^+}\sqrt{\Delta t}, \tag{46}$$

where the positive part of a symmetric matrix is defined as follows

Definition If $A \in S_n(\mathbb{R})$, thanks to the spectral theorem we can write

$$A = PDiag\left(\lambda_1, ..., \lambda_n\right) P^T.$$

Then

$$A^+ = PDiag\left(\left(\lambda_1\right)^+, ..., \left(\lambda_n\right)^+\right) P^T.$$

The purpose of this scheme is mainly to be a comparison basis for the other schemes that we will derive in this paper.

6.2.2 Extension of the QE scheme to Wishart processes

In [5], Andersen derived his so-called Quadratic-Exponential scheme for the Heston model, which has been proved to outperform by far most of existing schemes. We extended his results for the Double Heston model in [23], so much so that a generalization of this scheme to the case of Wishart processes is also bound to perform well. In this section we will try to derive such a scheme. Before beginning, we will start by recalling a certain number of results of matrix theory which will prove useful later on.

6.2.2.a Important results of Matrix Theory

We will now give a catalog of results some of which are generalizations of existing ones.

Proposition 12. *Let $A, B \in S_n^+(\mathbb{R})$, let $\lambda_{max}(A)$ and $\lambda_{max}(B)$ be the greatest eigenvalues of A and B respectively. Then, all the eigenvalues of AB are in the compact $[0, \lambda_{max}(A)\lambda_{max}(B)]$.*

Proof. We know that for every $x \in \mathbb{R}^n$ we have

$$0 \leq \langle x, Bx \rangle \leq \lambda_{max}(B) \langle x, x \rangle ,$$

where $<,>$ denotes the usual scalar product on $\mathbb{R}^n$.

Therefore, we have for all $y \in \mathbb{R}^n$

$$0 \leq \left\langle y, \sqrt{A}B\sqrt{A}y \right\rangle \leq \lambda_{max}(B) \langle y, Ay \rangle \leq \lambda_{max}(A)\lambda_{max}(B) \langle y, y \rangle ,$$

which in turn means that $\sqrt{A}B\sqrt{A}$ has all its eigenvalues in $[0, \lambda_{max}(A)\lambda_{max}(B)]$.

To conclude it suffices to remark that for every matrix M and N, MN and NM have the same eigenvalues, so that here $\sqrt{A}B\sqrt{A}$ and AB also have the same eigenvalues. □

Proposition 13. *Let $A, B \in \widetilde{S}_n^+(\mathbb{R})$ and let $\alpha \in \mathbb{R}_+^*$. Then we have*

(i) if $\alpha < \frac{\lambda_{min}(A)}{\lambda_{max}(B)}$ then the matrix $A - \alpha B$ is symmetric definite positive.

(ii) if $\alpha = \frac{1}{\lambda_{max}(BA^{-1})}$ then the matrix $A - \alpha B$ is symmetric positive with zero determinant.

Before proving this proposition, we will need the following two lemmas

Lemma 1.

(i) Let $A \in S_n(\mathbb{R})$ and $B \in S_n^+(\mathbb{R})$. Then we have

$$\lambda_{min}(A)\mathrm{Tr}\,(B) \leq \mathrm{Tr}\,(AB) \leq \lambda_{max}(A)\,\mathrm{Tr}\,(B) .$$

(ii) Let A and $B \in S_n(\mathbb{R})$, then we have

$$\lambda_{min}(A) + \lambda_{min}(B) \leq \lambda_{min}(A + B) \leq \lambda_{min}(A) + \lambda_{max}(B).$$

Lemma 2. *Let $A, B \in \widetilde{S}_n^+(\mathbb{R})$, let I be a subset of $\{1, ..., n\}$ and A_{II} and B_{II} the corresponding principal minors. Then we have*

$$\lambda_{max}(BA^{-1}) \geq \lambda_{max}(B_{II}A_{II}^{-1}).$$

Proof of lemma 1.

(i) Since A is symmetric we know from the spectral theorem that there exist an orthogonal matrix P and a diagonal matrix D such that $A = PDP^T$, and where the entries of D are the eigenvalues of A. Since the trace operator is similarity-invariant, if we note $\widetilde{B} = PBP^T$, we have

$$\text{Tr}(AB) = \sum_{i=i}^{n} D^{ii}\widetilde{B}^{ii}.$$

Since B is a positive matrix, its diagonal elements are all positive, and therefore we easily get

$$\lambda_{min}(A)\text{Tr}\,(B) \leq \text{Tr}\,(AB) \leq \lambda_{max}\,(A)\,\text{Tr}\,(B)\,.$$

(ii) This is a simple consequence of the Courant-Fischer theorem for the eigenvalues of a symmetric matrix (see [26] chapter 8 for more details).

□

Proof of lemma 2. The case where the subset I has the form $I = [\![1, k]\!]$ has been proven by Wang et al. in [45]. For our generalized result let us first write the following block-matrix structure for A and B

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{12}^T & A_{II} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix}$$

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{12}^T & B_{II} & B_{23} \\ B_{13}^T & B_{23}^T & B_{33} \end{pmatrix}.$$

We then know that there exists a permutation matrix P, such that $P^{-1} = P^T$ and

$$PAP^T = \begin{pmatrix} A_{II} & A_{12}^T & A_{23} \\ A_{12} & A_{11} & A_{13} \\ A_{23}^T & A_{13}^T & A_{33} \end{pmatrix}$$

$$PBP^T = \begin{pmatrix} B_{II} & B_{12}^T & B_{23} \\ B_{12} & B_{11} & B_{13} \\ B_{23}^T & B_{13}^T & B_{33} \end{pmatrix}.$$

Therefore we have easily $PBA^{-1}P^T = \left(PBP^T\right)\left(PAP^T\right)^{-1}$ and we can apply the result of [45], which ends the proof since the eigenvalues of the matrix are not changed by the permutation transformation. □

We can now prove the proposition

Proof of proposition 13.

(i) Using lemma 1 we know that the smallest eigenvalue of $A - \alpha B$ verifies

$$\lambda_{min}(A - \alpha B) \geq \lambda_{min}(A) - \alpha\lambda_{max}(B).$$

Therefore if $\alpha < \frac{\lambda_{min}(A)}{\lambda_{max}(B)}$, the smallest eigenvalue of $A - \alpha B$ will be strictly positive, and since it is also clearly symmetric, it is positive-definite.

(ii) We first have

$$\det(A - \alpha B) = \det(A)\det\left(I_n - \alpha BA^{-1}\right).$$

Since $\det(A) > 0$ and that the eigenvalues of $I_n - \alpha BA^{-1}$ are the $1 - \alpha\lambda(BA^{-1})$, it is clear that if we choose $\alpha = \frac{1}{\lambda_{max}(BA^{-1})}$ then the determinant above will be zero.

Then, we know that a symmetric matrix is positive if and only if the determinant of its principal minors are all positive, that is to say

$\forall I \subset [\![1, n]\!]$

$$\det\left(A_{II} - \frac{1}{\lambda_{max}(BA^{-1})}B_{II}\right) \geq 0$$

$$\Leftrightarrow \det\left(I_{\#I} - \frac{1}{\lambda_{max}(BA^{-1})}B_{II}A_{II}^{-1}\right) \geq 0.$$

But with lemma 2 we know that for all I, $\lambda_{max}(BA^{-1}) \geq \lambda_{max}(B_{II}A_{II}^{-1})$. Thus for all I we have

$$\lambda_{min}\left(I_{\#I} - \frac{1}{\lambda_{max}(BA^{-1})}B_{II}A_{II}^{-1}\right) = 1 - \frac{\lambda_{max}(B_{II}A_{II}^{-1})}{\lambda_{max}(BA^{-1})} \geq 0,$$

which ends the proof.

□

We will now study a matrix equation which will be involved in the derivation of our scheme

Proposition 14. *Let M be a symmetric positive matrix of order n, let $a, b \in \mathbb{R}_+^*$ and let us consider the following matrix equation*

$$aX^2 + \mathrm{Tr}(X)X = bM. \tag{47}$$

Then (47) has a unique symmetric positive solution which is given by

$$X = \sqrt{\frac{b}{a}M + \frac{\nu^2}{4a^2}I_n} - \frac{\nu}{2a}I_n,$$

where we have

- *If $n = 2$*

$$\nu = \sqrt{\frac{b}{(1+a)(a+2)}\left((1+a)\,\mathrm{Tr}(M) + \sqrt{\mathrm{Tr}^2(M) + 4a(a+2)\det(M)}\right)}$$

- *If $n > 2$, ν is the unique solution of the following equation*

$$-\nu\left(1 + \frac{n}{2a}\right) + \mathrm{Tr}\left[\sqrt{\frac{b}{a}M + \frac{\nu^2}{4a^2}I_n}\right] = 0.$$

Proof. See appendix A. □

Remark When $n > 2$ the equation above has to be numerically solved. However, given the form of the function, as can be seen in the following graph, and since its first derivative can easily be computed, it is clear that a simple Newton-Raphson type algorithm will do the job.

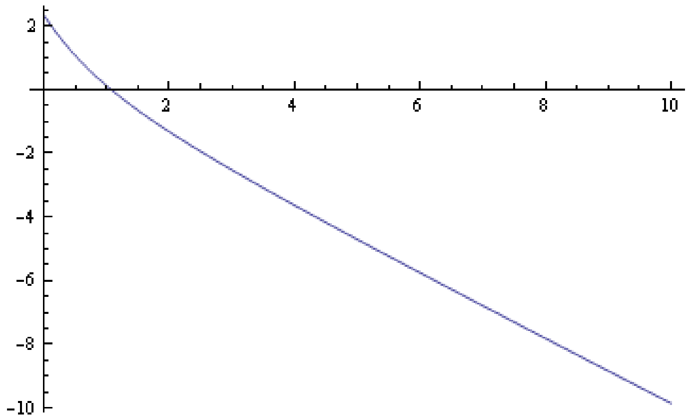


Figure 2: Representation of the equation to solve

6.2.2.b The Wishart QE-Scheme

First Moment-Matching The first idea here is to approximate the conditional distribution of $V_{t+\Delta t}$ by a non-central Wishart random variable $\mathcal{W}_k(M, \Sigma)$ where M and Σ are respectively symmetric positive and symmetric positive definite and are chosen so that the first two "moments" would be identical, and the choice of k will be explained later on. Here comes the first problem. Indeed, if we were to choose as "second" moment of the distribution the matrix $\mathbb{E}\left[(V_{t+\Delta t} - \mathbb{E}_t\left[V_{t+\Delta t}\right]) \otimes (V_{t+\Delta t} - \mathbb{E}_t\left[V_{t+\Delta t}\right])\right]$, which entries are given by proposition 9, we would end up with a system of equations completely inextricable. In order to be able to have closed-form solutions, we have therefore decided to match the "variance" matrix instead, since, thanks to the Alam and Mitra formula, it has a rather pleasant expression. If we note

$$\mathbb{E}_W(\Delta t) = \mathbb{E}\left[\, V_{t+\Delta t} |\, V_t\right]$$

$$\mathbb{V}_W(\Delta t) = \mathbb{E}\left[\, V_{t+\Delta t}^2 - \mathbb{E}\left[\, V_{t+\Delta t} |\, V_t\right] |\, V_t\right],$$

were the exact expressions can be found in proposition 9, then the moment-matching system can be written

$$\mathbb{E}_W(\Delta t) = M + k\Sigma$$

$$k\mathbb{V}_W(\Delta t) = (M + n\Sigma)^2 + (M + n\Sigma)\operatorname{Tr}\left[M + n\Sigma\right] - M^2 - M\operatorname{Tr}\left[M\right],$$

that is to say

$$\Sigma = \frac{1}{k}\left(\mathbb{E}_W(\Delta t) - M\right)$$

$$M^2 + M\operatorname{Tr}\left[M\right] = \mathbb{E}_W^2(\Delta t) + \mathbb{E}_W(\Delta t)\operatorname{Tr}\left[\mathbb{E}_W(\Delta t)\right] - k\mathbb{V}_W(\Delta t).$$

Let us note

$$P = \mathbb{E}_W^2(\Delta t) + \mathbb{E}_W(\Delta t)\operatorname{Tr}\left[\mathbb{E}_W(\Delta t)\right] - k\mathbb{V}_W(\Delta t). \tag{48}$$

According to proposition 14, it is necessary that the matrix P is positive for the system to have a solution. But we have, with the notations of proposition 9

$$P = \left(\mathbb{E}_W^2(\Delta t) + \mathbb{E}_W(\Delta t)\operatorname{Tr}\left[\mathbb{E}_W(\Delta t)\right]\right)\left(1 - \frac{k}{\beta}\right) + \mu_t(\Delta t)^2 + \mu_t(\Delta t)\operatorname{Tr}\left[\mu_t(\Delta t)\right].$$

Since by definition the matrix $\mu_t(\Delta t)$ and $\mathbb{E}_W^2(\Delta t) + \mathbb{E}_W(\Delta t)\operatorname{Tr}\left[\mathbb{E}_W(\Delta t)\right]$ are symmetric positive, it is clear that the matrix P above is going to be positive if $k \leq \beta$. Therefore we make the following choice for k, based also on the case for which we know how to simulate a non-central Wishart random variable

- If $n - 1 < \beta < n$ then $k = n$.
- If $n < \beta \leq 2n - 1$ and $\beta \notin \mathbb{N}$ then $k = [\beta]$.
- If $\beta \geq n$ and $\beta \in \mathbb{N}$ or $\beta > 2n - 1$ then $k = \beta$.

In the first case the moment-matching may fail. We have to check whether the matrix P is positive or not, and if not we will have to use the second moment-matching of the next paragraph. In the third case however, the moment matching is trivial, it suffices to set $M = \mu_t(\Delta t)$ and $\Sigma = \Sigma(\Delta t)$ and to use proposition 8 for the simulation. Finally in the second case, the matrix P is positive and the matrix M is given by proposition 14

$$M = \sqrt{P + \frac{\nu^2}{4} I_n} - \frac{\nu}{2} I_n, \tag{49}$$

where we have

- If $n = 2$

$$\nu = \sqrt{\frac{1}{6}\left(2\operatorname{Tr}(P) + \sqrt{\operatorname{Tr}^2(P) + 12\det(P)}\right)}$$

- If $n > 2$, ν is the unique solution of the following equation

$$-\nu\left(1 + \frac{n}{2}\right) + \operatorname{Tr}\left[\sqrt{P + \frac{\nu^2}{4} I_n}\right] = 0.$$

We then have

$$\Sigma = \frac{1}{k}\left(\mathbb{E}_W(\Delta t) - M\right). \tag{50}$$

Unfortunately we have no way to know if the matrix Σ above is going to be definite positive. If that is the case, then the moment matching is over. Otherwise, that means that the variance process has a non-zero probability to degenerate and that the approximation by a non-central Wishart with a definite positive matrix of covariance is no longer pertinent. Hence the second moment-matching.

Second Moment-Matching Let B, W_1, W_2 be independent random variables which laws are respectively the Bernouilli law $\mathcal{B}(c)$, the non-central Wishart $\mathcal{W}_k(R, S)$ and the non-central Wishart $\mathcal{W}_k(V, W)$, where $0 < c < 1$, $R, S \in S_n^+(\mathbb{R})$ with $\det S = 0$ and $V, W \in \widetilde{S}_n^+(\mathbb{R})$. We approximate the conditional variance process by

$$V_{t+\Delta t} \approx BW_1 + (1 - B)W_2.$$

In other words, we now allow the approximated process to become degenerate with a strictly positive probability (namely c). Besides, in order to reduce the number of unknown variables in our system, we impose the following relations

$$V = \alpha W$$

$$R = \alpha S,$$

where the positive parameter α will be one of the key to make the following system of moment-matching always solvable.

$$\mathbb{E}_W(\Delta t) = c\,(R + kS) + (1 - c)\,(V + kW)$$

$$\mathbb{V}_W(\Delta t) + \mathbb{E}_W^2(\Delta t) = c\,(R + kS)^2 + (1 - c)\,(V + kW)^2$$

$$+\, c\frac{(R + kS)^2 + (R + kS)\operatorname{Tr}[R + kS] - R^2 - R\operatorname{Tr}[R]}{k}$$

$$+\, (1 - c)\frac{(V + kW)^2 + (V + kW)\operatorname{Tr}[V + kW] - V^2 - V\operatorname{Tr}[V]}{k},$$

which gives after simplification

$$\mathbb{E}_W(\Delta t) = (\alpha + k)\,(cS + (1 - c)W)$$

$$\mathbb{V}_W(\Delta t) + \mathbb{E}_W^2(\Delta t) = (\alpha + k)^2 \left(cS^2 + (1 - c)W^2\right)$$

$$+\, (2\alpha + k)\left(c\left(S^2 + S\operatorname{Tr}[S]\right) + (1 - c)\left(W^2 + W\operatorname{Tr}[W]\right)\right),$$

and finally

$$S = \frac{\mathbb{E}_W(\Delta t)}{(\alpha+k)c} - \frac{1-c}{c}W$$

$$(1+a)\widetilde{W}^2 + a\widetilde{W}\,\mathrm{Tr}[\widetilde{W}] = \frac{c}{1-c}G,$$

where

$$a = \frac{2\alpha+k}{(\alpha+k)^2}$$

$$\widetilde{W} = (\alpha+k)W - \mathbb{E}_W(\Delta t)$$

$$G = \mathbb{V}_W(\Delta t) - a\left(\mathbb{E}^2_W(\Delta t) + \mathbb{E}_W(\Delta t)\,\mathrm{Tr}\left[\mathbb{E}^2_W(\Delta t)\right]\right).$$

Let us note

$$G := X - aY.$$

As well as with the first moment-matching, the matrix G above needs to be definite positive for the system to have a solution. But by definition the matrix X and Y above are positive definite. Indeed, with the notations of proposition 9, $\Sigma(\Delta t)$ is definite positive because $Q \in GL_n(\mathbb{R})$ and therefore $e^{\Delta t M}Q^TQe^{\Delta t M^T}$ is definite positive for all $t \geq 0$. Thus, since $\mu_t(\Delta t)$ is also positive, $\mathbb{E}_W(\Delta t)$ is definite positive, which clearly shows that X and Y are definite positive too. Therefore, we can now apply proposition 13, from which we get that if

$$a < \frac{\lambda_{min}(X)}{\lambda_{max}(Y)}, \tag{51}$$

then G is definite positive.

But (51) is equivalent to

$$\frac{2\alpha+k}{(\alpha+k)^2} < \frac{\lambda_{min}(X)}{\lambda_{max}(Y)}$$

$$\text{i.e. } \frac{\lambda_{min}(X)}{\lambda_{max}(Y)}\alpha^2 + 2\left(k\frac{\lambda_{min}(X)}{\lambda_{max}(Y)} - 1\right)\alpha + k\left(k\frac{\lambda_{min}(X)}{\lambda_{max}(Y)} - 1\right) > 0.$$

The discriminant of the second order polynomial above is given by

$$\Delta = 4\left(1 - k\frac{\lambda_{min}(X)}{\lambda_{max}(Y)}\right).$$

Therefore, if $k > \frac{\lambda_{max}(Y)}{\lambda_{min}(X)}$, G is always definite positive and we can take for α any strictly positive number. Otherwise, we only need to take α so that

$$\alpha > \frac{\lambda_{max}(Y) - k\lambda_{min}(X) + \sqrt{\lambda_{max}(Y)\left(\lambda_{max}(Y) - k\lambda_{min}(X)\right)}}{\lambda_{min}(X)}. \tag{52}$$

This problem being solved, the solution for the matrix S can be obtained thanks to proposition 14

$$W = \frac{1}{\alpha + k}\left[\mathbb{E}_W(\Delta t) + \sqrt{\frac{c}{1-c}}\left(\sqrt{\frac{G}{1+a} + \frac{\nu^2 a^2}{4(1+a)^2} I_n} - \frac{\nu a}{2(1+a)} I_n\right)\right], \tag{53}$$

where we have

- If $n = 2$

$$\nu = \frac{1}{\sqrt{(1+2a)(1+3a)}}\sqrt{(1+2a)\,\mathrm{Tr}(G) + \sqrt{a^2\,\mathrm{Tr}^2(G) + 4(1+a)(1+3a)\det(G)}}.$$

- If $n > 2$, ν is the unique solution of the following equation

$$-\nu\left(1 + \frac{na}{2(1+a)}\right) + \mathrm{Tr}\left[\sqrt{\frac{G}{1+a} + \frac{\nu^2 a^2}{4(1+a)^2} I_n}\right] = 0.$$

We notice immediately that W is a symmetric definite positive matrix, since it is the sum of a symmetric definite positive matrix (i.e. $\frac{\mathbb{E}_W(\Delta t)}{\alpha+k}$) and a symmetric positive matrix (the rest of the expression).

Moving on to S we obtain easily

$$S = \frac{1}{\alpha + k}\left[\mathbb{E}_W(\Delta t) - \sqrt{\frac{1-c}{c}}\left(\sqrt{\frac{G}{1+a} + \frac{\nu^2 a^2}{4(1+a)^2} I_n} - \frac{\nu a}{2(1+a)} I_n\right)\right]. \tag{54}$$

With this form it is unclear whether S is going to be positive or not. Besides, we also need to have $\det(S) = 0$. Nonetheless, if we note

$$S := \frac{1}{\alpha + k}\left(A - \sqrt{\frac{1-c}{c}} B\right),$$

we know thanks to proposition 13 that if we take

$$\sqrt{\frac{1-c}{c}} = \frac{1}{\lambda_{max}(BA^{-1})}$$

$$\text{i.e. } c = \frac{\lambda_{max}^2(BA^{-1})}{1+\lambda_{max}^2(BA^{-1})}, \tag{55}$$

then the matrix S will be positive with zero determinant.

In the end, we were able to solve the initial system of moment-matching under all constraints. We now give a summary of the algorithm

Algorithm 1.

- *Step 1*
  - *If* $n-1 < \beta < n$
    1. *Check if the matrix* P *in (48) is definite positive. If not go to step 2.*
    2. *Define* M *as in (49).*
    3. *Check if the matrix* Σ *in (50) is definite positive. If not go to step 2.*
    4. *Simulate* $v \sim \mathcal{W}_n(M, \Sigma)$ *with proposition 8.*
    5. *Set* $\widehat{V}_{t+\Delta t} = v$.
  - *If* $n < \beta \leq 2n-1$ *and* $\beta \notin \mathbb{N}$
    1. *Define* M *as in (49).*
    2. *Check if the matrix* Σ *in (50) is definite positive. If not go to step 2.*
    3. *Simulate* $v \sim \mathcal{W}_{[\beta]}(M, \Sigma)$ *with proposition 8.*
    4. *Set* $\widehat{V}_{t+\Delta t} = v$.
  - *If* $\beta \geq n$ *and* $\beta \in \mathbb{N}$ *or* $\beta > 2n-1$
    1. *Define* $\mu_t(\Delta t)$ *and* $\Sigma(\Delta t)$ *as in proposition 9.*
    2. *Simulate* $v \sim \mathcal{W}_\beta(\mu_t(\Delta t), \Sigma(\Delta t))$ *with proposition 8.*
    3. *Set* $\widehat{V}_{t+\Delta t} = v$.

- *Step 2*
  1. *Define* k *according to* $k = n\mathbb{1}_{n-1<\beta<n} + [\beta]\mathbb{1}_{n<\beta\leq 2n-1 \text{ and } \beta\notin\mathbb{N}}$.
  2. *Define* α, W, S *and* c *according to (52), (53), (54) and (55) respectively.*
  3. *Simulate* $U \sim \mathcal{U}([0,1])$.
     - *If* $U < c$ *simulate* $v \sim \mathcal{W}_k(\alpha S, S)$ *with proposition 8.*
     - *If* $U \geq c$ *simulate* $v \sim \mathcal{W}_k(\alpha W, W)$ *with proposition 8.*
  4. *Set* $\widehat{V}_{t+\Delta t} = v$.

Remark We can easily make a little simplification in the definition of α in (52). Indeed it is easy to show that the function $x \rightarrow \frac{x-k\lambda_{min}(X)+\sqrt{x(x-k\lambda_{min}(X))}}{\lambda_{min}(X)}$ is increasing in x and therefore we can replace the value of $\lambda_{max}(Y)$ by any larger constant. Then we only have to use the well-known Gersgorin's circles (see e.g. the reference [26]) and replace $\lambda_{max}(Y)$ by $\max_{1\leq i\leq n}\left(Y^{ii}+\sum_{j\neq i}\left|Y^{ij}\right|\right)$. Therefore we have one less eigenvalue to compute.

Martingale Correction Since the approximated law are, in both moment-matchings, linear combinations of independent non-central Wishart random variables (conditional on the previous time step), it is easy to compute the Laplace transform of the discretized scheme. Indeed, we have

- First Moment-Matching

$$\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^T R\widehat{V}_{t+\Delta t}\right)}|\widehat{V}_t\right]=\frac{e^{-\frac{\Delta t}{4}\operatorname{Tr}\left[R^T R\left(I_n+\frac{\Delta t}{2}\Sigma R^T R\right)^{-1}M\right]}}{\left[\det\left(I_n+\frac{\Delta t}{2}\Sigma R^T R\right)\right]^{k/2}}.$$

- Second Moment-Matching

$$\begin{aligned}\mathbb{E}\left[e^{-\frac{\Delta t}{4}\operatorname{Tr}\left(R^T R\widehat{V}_{t+\Delta t}\right)}|\widehat{V}_t\right]&=c\frac{e^{-\alpha\frac{\Delta t}{4}\operatorname{Tr}\left[R^T R\left(I_n+\frac{\Delta t}{2}SR^T R\right)^{-1}S\right]}}{\left[\det\left(I_n+\frac{\Delta t}{2}SR^T R\right)\right]^{k/2}}\\&+(1-c)\frac{e^{-\alpha\frac{\Delta t}{4}\operatorname{Tr}\left[R^T R\left(I_n+\frac{\Delta t}{2}WR^T R\right)^{-1}W\right]}}{\left[\det\left(I_n+\frac{\Delta t}{2}WR^T R\right)\right]^{k/2}},\end{aligned}$$

  from which we get the martingale corrected scheme using (38).

Consistency of the Scheme A formal analysis of the convergence properties for the Wishart QE scheme proposed in this paper is difficult and complicated by the fact that the X process may not have any high-order moments. As such, the usual examination of (weak) convergence of expectations of polynomials of X is not always meaningful. While we could, in principle, undertake an examination of the convergence of expectations on selected slow-growing payouts (e.g. call options), the technicalities of such an analysis are considerable and we skip it (see [39] for examples of this type of analysis). Instead, as pointed out by Andersen in [5] we focus on a simpler concept, namely that of weak consistency. As shown in [33] p. 328, there is a strong link between weak consistency and weak convergence. We will show that our scheme is almost weakly consistent, with the exception of the covariance between the discretized asset prices and the discretized variance.

Proposition 15. *The Wishart QE scheme verifies that conditional on $\widehat{S}_t$ and $\widehat{V}_t$, we have*

$$\mathbb{V}ar_t\left[\frac{\widehat{V}_{t+\Delta t}-\widehat{V}_t}{\sqrt{\Delta t}}\right] \underset{\Delta t\to 0}{\to} \widehat{V}_tQ^TQ + Q^TQ\widehat{V}_t + Q^TQ\,\mathrm{Tr}[\widehat{V}_t] + \widehat{V}_t\,\mathrm{Tr}[Q^TQ]$$

$$\mathbb{E}_t\left[\frac{\widehat{V}_{t+\Delta t}-\widehat{V}_t}{\Delta t}\right] \underset{\Delta t\to 0}{\to} M\widehat{V}_t + \widehat{V}_tM^T + \beta Q^TQ$$

$$\mathbb{V}ar_t\left[\frac{1}{\sqrt{\Delta t}}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right] \underset{\Delta t\to 0}{\to} \mathrm{Tr}[\widehat{V}_t]$$

$$\mathbb{E}_t\left[\frac{1}{\Delta t}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right] \underset{\Delta t\to 0}{\to} r - \frac{1}{2}\mathrm{Tr}[\widehat{V}_t].$$

Proof. Since the random variable Z is independent of $\widehat{V}_{t+\Delta t}$, we have thanks to proposition 9 (since the Wishart QE scheme is made so that the first two moments of the variance process are matched)

$$\begin{aligned}\mathbb{E}\left[\frac{1}{\Delta t}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right] &= r - \frac{1}{4}\mathrm{Tr}\left[\widehat{V}_t + \mu_t(\Delta t) + \beta\Sigma(\Delta t)\right] \\ &\underset{\Delta t\to 0}{\to} r - \frac{1}{2}\mathrm{Tr}[\widehat{V}_t].\end{aligned}$$

Then we also have

$$\begin{aligned}\mathbb{V}ar_t\left[\frac{1}{\sqrt{\Delta t}}\log\frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t}\right] &= \frac{1}{2}\mathrm{Tr}\left[\left(\widehat{V}_t + \mathbb{E}_t\left[\widehat{V}_{t+\Delta t}\middle|\widehat{V}_t\right]\right)\left(I_n - R^TR\right)\right] \\ &+ \mathbb{E}\left[\mathrm{Tr}^2\left(\sqrt{\widehat{V}_t}GR\right)\middle|\widehat{V}_t\right] + O(\Delta t^{1/2}) \\ &= \frac{1}{2}\mathrm{Tr}\left[\left(\widehat{V}_t + \mu_t(\Delta t) + \beta\Sigma(\Delta t)\right)\left(I_n - R^TR\right)\right] \\ &+ \mathrm{Tr}[\widehat{V}_tR^TR] + O(\Delta t^{1/2}) \\ &\underset{\Delta t\to 0}{\to} \mathrm{Tr}[\widehat{V}_t].\end{aligned}$$

Finally, since we know the first two moments of the discretized process, the two other results follow easily as soon as we notice that

$$\frac{e^{\Delta t M}\widehat{V}_t e^{\Delta t M^T} - \widehat{V}_t}{\Delta t} \underset{\Delta t \to 0}{\rightarrow} M\widehat{V}_t + \widehat{V}_t M^T$$

$$\frac{\Sigma(\Delta t)}{\Delta t} \underset{\Delta t \to 0}{\rightarrow} Q^T Q.$$

□

The results above are interesting, since they show that our scheme is consistent with the dynamic of S_t and V_t since we can easily show (after some calculations) that

$$\begin{aligned}
\mathrm{Drift}(d\log S_t) &= r dt - \frac{1}{2}\mathrm{Tr}[\widehat{V}_t] \\
\langle d\log S_t, d\log S_t \rangle &= \mathrm{Tr}[\widehat{V}_t] dt \\
\mathrm{Drift}(dV_t) &= \left(MV_t + V_t M^T + \beta Q^T Q\right) dt \\
\langle dV_t, dV_t \rangle &= \left(V_t Q^T Q + Q^T Q V_t + V_t \,\mathrm{Tr}[Q^T Q] + Q^T Q\, \mathrm{Tr}[V_t]\right) dt,
\end{aligned}$$

where the definition of $\langle dV_t, dV_t \rangle$ when V is a matrix valued process is recalled in the proof of theorem 2.

Unfortunately, when it comes to the covariance between the asset and the variance process, we can show that

$$\langle d\log S_t I_n, dV_t \rangle = \left(V_t R^T Q + Q^T R V_t\right) dt,$$

whereas, if we assume that

$$\mathbb{E}\left[\widehat{V}_{t+\Delta t}\,\mathrm{Tr}[\widehat{V}_{t+\Delta t}] \middle| \widehat{V}_t\right] \underset{\Delta t \to 0}{\rightarrow} \widehat{V}_t\,\mathrm{Tr}[\widehat{V}_t],$$

which we know to be exact only when $\beta \in \mathbb{N}\backslash\{0, 1, ..., n\}$ or $\beta > 2n - 1$, we get that conditional on $\widehat{V}_t$

$$\mathbb{C}ov\left[\frac{1}{\sqrt{\Delta t}} \log \frac{\widehat{S}_{t+\Delta t}}{\widehat{S}_t} I_n, \frac{\widehat{V}_{t+\Delta t} - \widehat{V}_t}{\sqrt{\Delta t}}\right] \underset{\Delta t \to 0}{\rightarrow} 0.$$

Therefore our scheme is not completely consistent, but it was to be expected since, unlike with Heston or Double Heston model, the dynamic of the variance process does not intervene at all in the discretization scheme for the price process.

7 Control variate

Control variate is one of the most effective tool for improving Monte Carlo convergence. The core idea of control variate is to exploit the error made on estimates of known quantities to correct estimates of unknown quantities. Let $(X_i)_{1\leq i\leq n}$ and $(Y_i)_{1\leq i\leq n}$ be the simulations of two payoffs X and Y. Furthermore, let assume that $\mathbb{E}[X]$ is known. The standard Monte-Carlo estimators of $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{56}$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i. \tag{57}$$

Both estimators are unbiased. Therefore, for any fixed β, we can construct a second unbiased estimator for $\mathbb{E}(Y)$:

$$\bar{Y}(\beta) = \bar{Y} + \beta(\mathbb{E}[X] - \bar{X}). \tag{58}$$

This estimator is unbiased because

$$\begin{aligned}
\mathbb{E}[\bar{Y}(\beta)] &= \mathbb{E}[\bar{Y} + \beta(\mathbb{E}[X] - \bar{X})] \\
&= \mathbb{E}[\bar{Y}] + \beta(\mathbb{E}[\mathbb{E}[X]] - \mathbb{E}[\bar{X}])] \\
&= \mathbb{E}[\bar{Y}] + \beta(\mathbb{E}[X] - \mathbb{E}[X])] \\
&= \mathbb{E}[\bar{Y}].
\end{aligned}$$

Now, in order to apply control variate, we still have to choose appropriately the β coefficient. For the moment, our new estimator matches the expectation of Y. A desirable property is to match also the variance of Y.

$$\begin{aligned}
\mathbb{V}ar[\bar{Y}(\beta)] &= \mathbb{V}ar[\bar{Y} + \beta(\mathbb{E}[X] - \bar{X})] \\
&= \sigma_Y^2 - 2\beta\sigma_X\sigma_Y\rho_{XY} + \beta^2\sigma_X^2,
\end{aligned} \tag{59}$$

where $\sigma_X^2 = \mathbb{V}ar[X]$, $\sigma_Y^2 = \mathbb{V}ar[Y]$ and $\langle X, Y\rangle = \rho_{XY}$.

The control variate estimator has smaller variance than the standard estimator, which corresponds to the case $\beta = 0$ if

$$\beta^2\sigma_X^2 - 2\beta\sigma_X\sigma_Y\rho_{XY} < 0.$$

The optimal coefficient β minimizes the variance (59) and is given by

$$\beta^* = \frac{\sigma_Y}{\sigma_X}\rho_{XY} = \frac{\mathbb{C}ov[X,Y]}{\mathbb{V}ar[X,Y]}. \tag{60}$$

The effectiveness of control variates depends strongly on the choice of the reference payoff. It can indeed be shown that control variate works better when X and Y are strongly correlated, which is quite intuitive. For a rigorous proof and further insights on control variate, Chapter 4 of Glasserman's book [24] is an excellent reference.

In our numerical examples, since we are pricing a call using Monte-Carlo, a good reference is the spot value. The call price and the spot value are indeed strongly correlated.

8 Numerical results

8.1 Test cases

We consider a call option of strike K with maturities 6 months (case I) and 3 years (case II). The other parameters values are given in appendix C. They have been chosen to match roughly market conditions which are not restrictive. In order to evaluate the call option price using a specific discretization scheme, we use the so-called Monte-Carlo method, with or without the control-variate techniques described above. All the numerical results are given in the appendix. Besides, we use two values of the parameter β, an integer one and a non-integer one, in order to highlight what seems to be two different behaviors of our schemes.

8.2 Conclusions of the Numerical Tests

8.2.1 Numerical results

The Integer Case First of all, it is rather evident when considering our results that the Monte-Carlo simulations for the Euler and Ornstein schemes really benefit from the use of the control variate method. Indeed, the gain is often quite spectacular, ranging between a factor 2 and a factor 10. However, for the QE scheme, the results are more mitigated, since the mean of the relative errors for all the simulations are roughly the same with or without control variate (namely around 3%).

When it comes to the efficiency of the schemes, a quick examination of the results in the appendix shows without any doubt that in this case the Euler and Ornstein schemes completely outperform the QE scheme. Indeed, even with the rather low number of simulations that we have made, the Monte-Carlo simulations for those two schemes almost always give a price with a bias non-significantly different from zero, and the mean of the relative error for the Euler scheme and Ornstein scheme are both below 1%, whereas the QE scheme is above 3% as mentionned previously. Let us also note that even though the QE scheme performs rather poorly whith far out-of-the-money options ($K = 1.3S_0$), it seems to behave better with options at-the-money with short maturities (case I), where it is often at least twice as precise as the two other schemes.

Now when comparing the Euler and Ornstein schemes, with all our simulations we came to the conclusions that even though almost all the biases obtained were not significantly different from zero, the Euler scheme performed better than the Ornstein scheme 60% of the time. Indeed, the mean of the relative errors is 0.67% for the Euler scheme, whereas it is 0.80% for the Ornstein scheme.

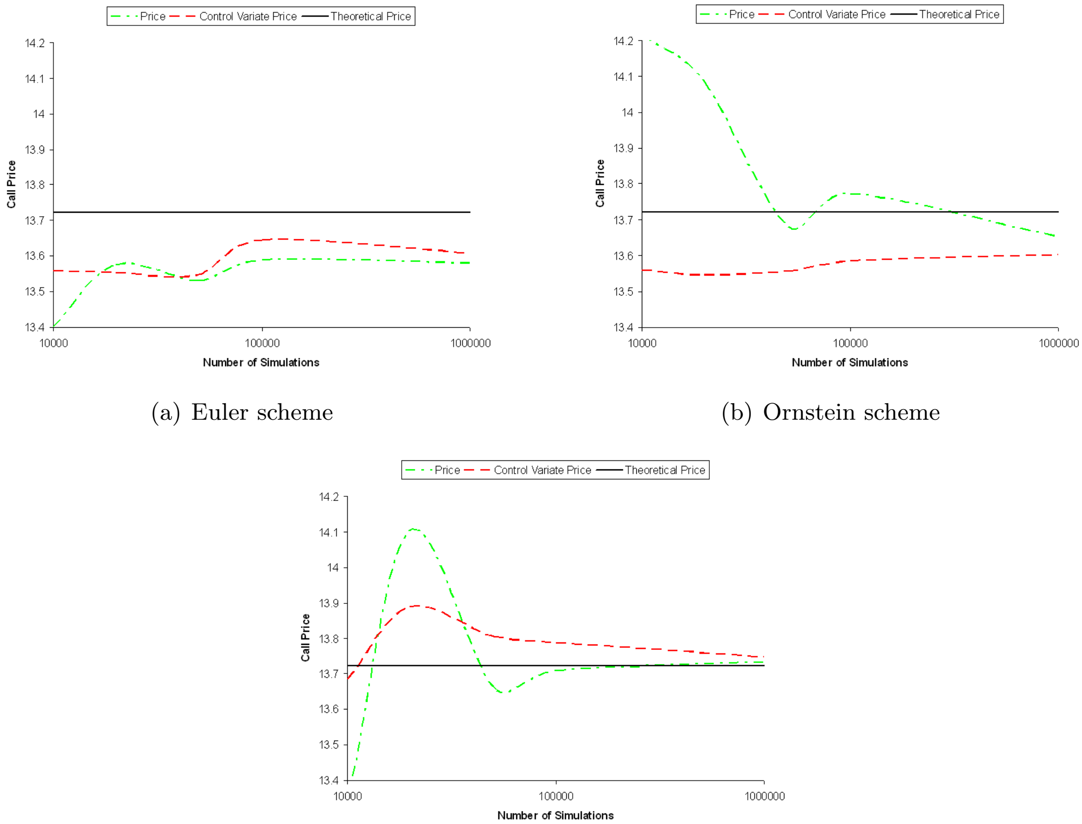
The Non-Integer Case In all the Monte-Carlo simulations concerning the Wishart model that we found in the literature, the parameter β was always chosen integer. In order to test the behavior of our schemes in the most comprehensive way possible, we also include simulations with a non-integer β, and the results that we obtained were quite surprising. First of all, the conclusion concerning the use of the control variate method is the same: especially with the Ornstein and QE scheme, it does not always improve the precision of the price obtained, even though it does on average.

Then, the Euler scheme, which had the best precision in the previous case, is now outperformed by the two other schemes, which are roughly 1.5 times as precise. Now when comparing the Ornstein and QE scheme, it is difficult to find a winner. The results are very close with control variate (an average error of 2.3% for the Ornstein scheme and 2.4% for the QE scheme), and the QE scheme has the upper hand without control variate (3.3% against 4%). When we compare the two cases I and II, it appears that the QE scheme performs better with short maturities, and the Ornstein scheme with long maturities.

Finally, it is also clear that the performance of all schemes (with the notable exception of the QE one) has dramatically decreased with a non-integer β. It might be linked to the rate of convergence of the Ornstein scheme (which is bound to be slower in this case since this scheme was originally meant to deal only with integer β) and the fact that the full-truncated Euler scheme is known to sometimes engender wide biases. Even though it would be very interesting to study further this modification of behavior from a theoretical point of view, we postpone this analysis which would likely be rather tedious.

8.2.2 Convergence

The Integer Case We represent below the behavior of our schemes in terms of convergence with respect to the number of simulations used (figure 3) and the number of steps (figure 4).

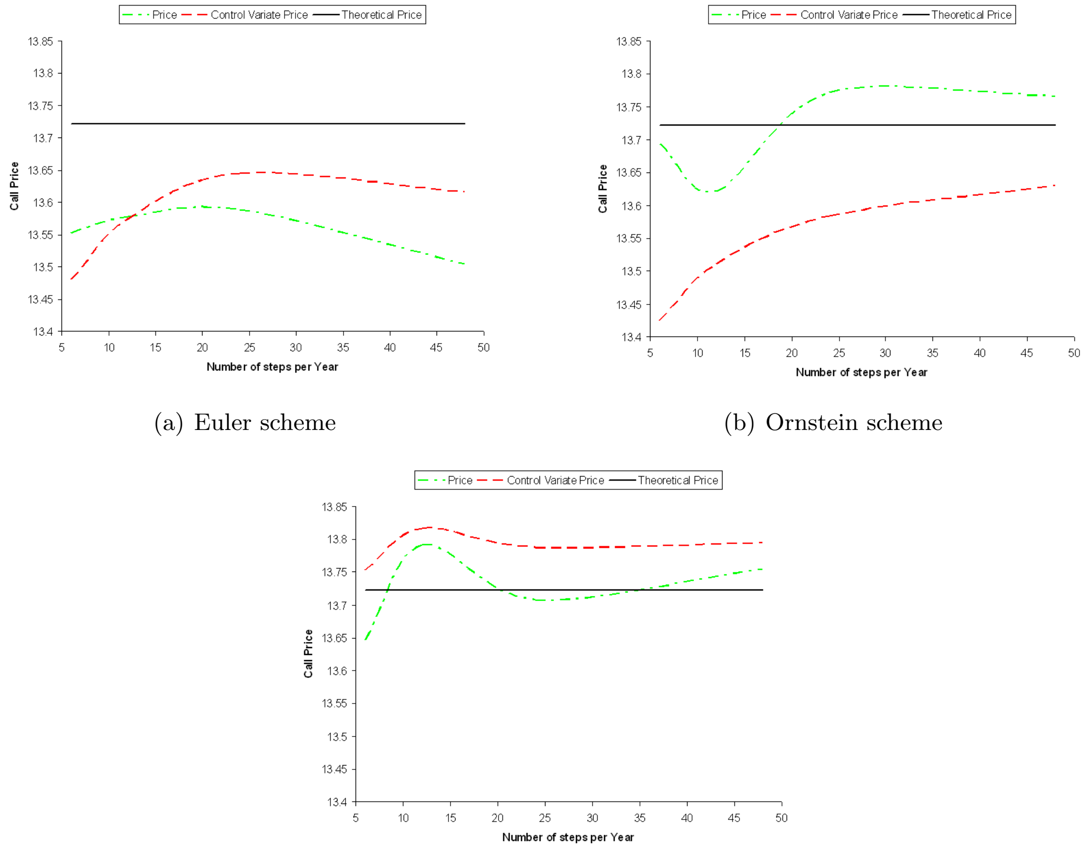


(a) Euler scheme

(b) Ornstein scheme

(c) QE scheme

Figure 3: Convergence of the schemes with the number of simulations



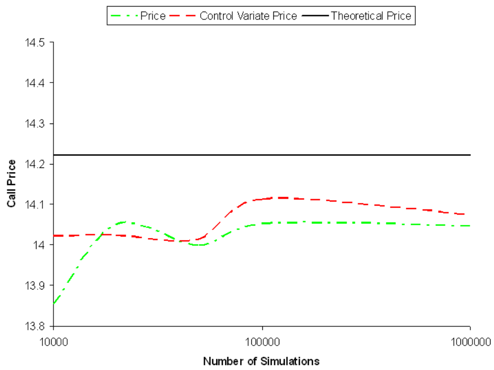
(a) Euler scheme

(b) Ornstein scheme

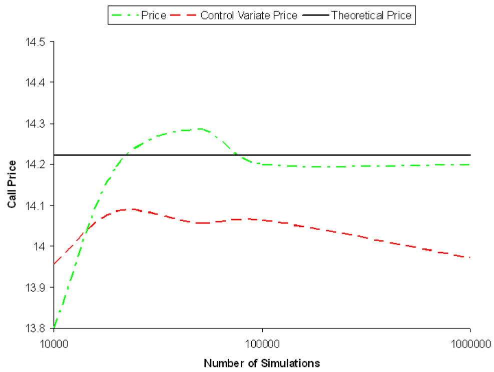
(c) QE scheme

Figure 4: Convergence of the schemes with the number of steps

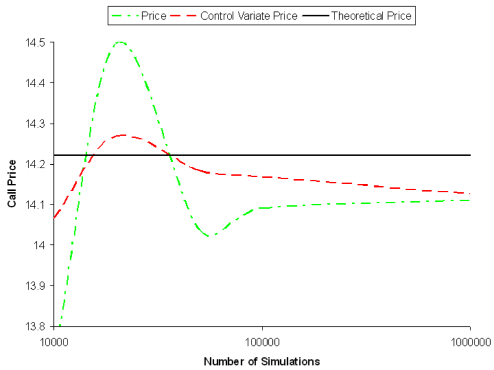
The Non-Integer Case We represent below the behavior of our schemes in terms of convergence with respect to the number of simulations used (figure 5) and the number of steps (figure 6).



(a) Euler scheme

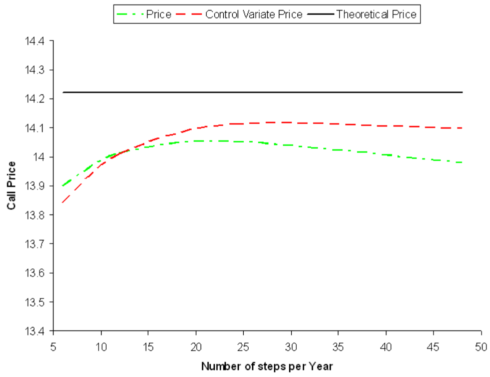


(b) Ornstein scheme

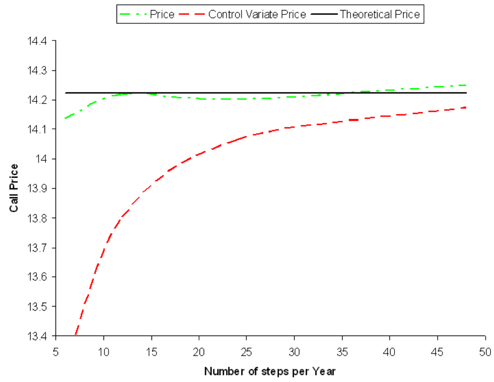


(c) QE scheme

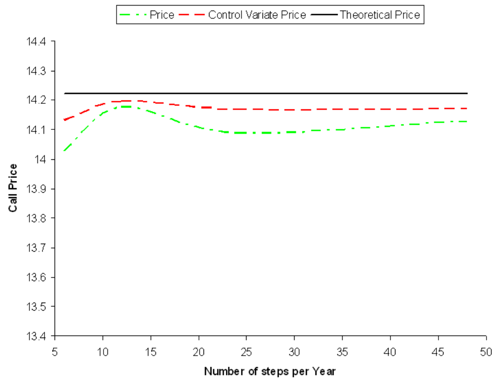
Figure 5: Convergence of the schemes with the number of simulations



(a) Euler scheme



(b) Ornstein scheme



(c) QE scheme

Figure 6: Convergence of the schemes with the number of steps

Our conclusions remain the same as in the previous section concerning the hierarchy of the schemes. The convergence of the Ornstein and QE schemes with the number of steps in the non-integer case (i.e. when Δt goes to 0) appears clearly faster than that of the Euler scheme, the situation being reversed in the integer case. Nonetheless, more tests with a higher number of simulations and steps may be required in order to have a better understanding of the behavior of those three schemes.

8.2.2.1 Computation time Of course the results above would not be comprehensive without mentioning the computational time of all the schemes. In case I we have the following results for 10000 simulations and 24 steps:

	Euler	Ornstein	Ornstein Non-Integer	QE-M
Time	3.297	2.391	4.593	13.468

Table 1: Computational Times for the different schemes (s)

In the integer case, the Ornstein scheme is by far the fastest, making it our preferred choice, all the more since when the dimension of the matrix used increase, the Euler scheme will have to diagonalize bigger matrix at each steps of the simulation, which will become very rapidly extremely time-consuming. In the non-integer case, the QE scheme is approximately 3 times slower than the Ornstein scheme (which is also the difference between the Euler and QE scheme in the Heston or Double Heston models, see [23]), but it can be more precise depending on the parameters. Besides, the QE scheme can deal with values of β in $]n-1, n+1[$ whereas the Ornstein cannot. Therefore, depending on the values of the parameters we recommend the use of either the QE scheme or the Ornstein scheme.

9 Further work

From what we have seen so far, it is clear that our QE scheme can be improved, either by changing the approximative laws or by bringing out faster and more efficient ways of simulating non-central Wishart processes with any number of degrees of freedom. Another promising road would be to try and adapt the schemes proposed by Alfonsi in [4] for the Heston model and their versions for the Double Heston model proposed by the authors in [23]. Indeed, they have been proven to be the most efficient schemes available for those models, in terms of precision and computation time. The work has actually been started by Ahdida in an upcoming PhD thesis [2], with, so far, promising results. Finally, we think that for improving the efficiency of the schemes, we will need to understand the origin of the difference of their behavior depending on whether β is an integer or not.

10 Conclusion

In this paper we have addressed the simulation and discretization issues of the Wishart model introduced by Da Fonseca et al in [20]. As far as we know, this question has not been addressed in the literature previously, despite a growing interest in the class of Wishart models for modelization in Finance. Indeed, they are the natural generalization to a multidimensional framework of the widely used Heston model, and thus remain on one hand analytically tractable, and on the other hand allow a far better capture of the volatility term structure (see [12] for an empirical analysis of the performance of multifactor stochastic volatility models). We have proposed in this paper three different methods for simulating the process, each of them having its own limitations. The full truncation Euler scheme performs surprisingly well in the integer case when compared to the results in Heston or Double Heston models for example. This could be linked to the fact the condition $\beta \geq n+1$ that is imposed in the definition of the Wishart model corresponds exactly when $n=1$ to the well known condition for which the CIR process never reaches 0, for which it is known that the truncated Euler schemes are efficient. However, it suffers from the fact that its computational time increase drastically when the dimension of the problem increases, and it performs poorly when β is no longer an integer. Concerning the Ornstein scheme, even though it is originally defined for integer values of β, a simple change of probability measure allows us to extend it to all values of $\beta \geq n+1$, but it cannot be used when $\beta \in]n-1, n+1[\setminus\{n\}$. Nonetheless, it is easy to implement and performs quite well even when $\beta \notin \mathbb{N}$, and also remains one of the fastest of our schemes. Finally, our first attempt to generalize the QE scheme of Andersen to the Wishart model has given us promising results. Even though it does not seem to always perform well when β is an integer, it is slightly better than the Ornstein scheme when $\beta \notin \mathbb{N}$, and more importantly, can be used as soon as we have $\beta > n-1$. Of course, it is three times slower than the Ornstein scheme, but we are convinced that a careful implementation and some modifications of the scheme could lead to important improvements in this regard.

Acknowledgments

Both authors are very grateful to Aurélien Alfonsi and Abdelkodousse Ahdida for pointing out several very interesting references and for their insightful remarks.

References

[1] Phillip Abbot. Tricks of the trade : Legendre-Gauss quadrature. *Mathematica Journal*, 9:689–691, 2005.

[2] Abdelkodousse Ahdida. Phd thesis Ecole des Ponts. *in preparation*.

[3] K. Alam and A. Mitra. On estimated the scale and noncentrality matrices of a Wishart distribution. *Sankha Series B*, 52:133–143, 1990.

[4] Aurélien Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Mathematics of Computation*, 2008.

[5] Leif Andersen. Efficient simulation of the Heston stochastic volatility model. *Working Paper*, December 2005.

[6] Leif Andersen and Vladimir Piterbarg. Efficient simulation of the Heston stochastic volatility model. *Working Paper*, July 2005.

[7] George A. Anderson. An asymptotic expansion for the noncentral Wishart distribution. *The Annals of Mathematical Statistics*, 41(5):1700–1707, 1970.

[8] T. W. Anderson. The non-central Wishart distribution and certain problems of multivariate statistics. *The Annals of Mathematical Statistics*, 17(4):409–431, 1946.

[9] Marie-France Bru. Processus de Wishart. *Compte-rendu de l'Académie des Sciences de Paris*, 308(Série I):29–32, 1989.

[10] Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4:725–743, 1991.

[11] Peter Carr and Dilip B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61:73, 1999.

[12] Peter Christoffersen, Steven Heston, and Kris Jacobs. The shape and term structure of the index option smirks : why multifactor stochastic volatility models work so well. *Working paper,McGill University*, February 2009.

[13] John Cox, Jonathan Ingersoll, and Stephen Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[14] D.Duffie, D.Filipovic, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

[15] Armido R. DiDonato and Alfred H. Morris. Computation of the incomplete Gamma function ratios and their inverse. *ACM Transactions on Mathematical Software*, 12(4), December 1986.

[16] Catherine Donati-Martin, Yan Doumerc, Hiroyuki Matsumoto, and Marc Yor. Some properties of the Wishart processes and a matrix extension of the Hartman-Watson laws. *Publ. RIMS, Kyoto Univ.*, 2004.

[17] Daniel Dufresne. The integrated square-root process. *Working Paper, University of Montréal*, 2001.

[18] José Da Fonseca, Martino Grasselli, and Florian Ielpo. Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. *Research report, ESILV*, November 2007.

[19] José Da Fonseca, Martino Grasselli, and Florian Ielpo. Hedging (co)variance risk with variance swaps. *Working paper, ESILV*, November 2008.

[20] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. A multifactor volatility Heston model. *Working paper, ESILV*, March 2006.

[21] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. Option pricing when correlations are stochastic : an analytical framework. *Research report, ESILV*, September 2006.

[22] Gerhard Freiling. A survey of nonsymmetric Riccati equations. *Linear algebra and its applications*, 351(2):243–270, 2002.

[23] Pierre Gauthier and Dylan Possamaï. Efficient simulation of the double Heston model. *Working Paper*, July 2009. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1434853.

[24] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, 2004.

[25] Leon Jay Gleser. A canonical representation for the noncentral Wishart distribution useful for simulation. *Journal of the American Statistical Association*, 71(355):690–695, September 1976.

[26] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 3rd Edition, 1996.

[27] Christian Gouriéroux. Continuous time Wishart process for stochastic risk. *CREF*, October 2006.

[28] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6:327–343, 1993.

[29] Nicholas J. Higham. Computing the polar decomposition, with applications. *Working Paper, University of Manchester*, 1986.

[30] Nicholas J. Higham. Stable iterations for the matrix square root. *Numerical Algorithms*, 15:227–242, 1997.

[31] Nicholas J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM Journal on Matrix Analysis and Applications*, 26(4):1179–1193, 2005.

[32] Nicholas J. Higham. Cholesky factorization. *Working Paper, University of Manchester*, November 2008.

[33] P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations.* 3rd ed. Springer Verlag, 1999.

[34] A.M. Kshirsagar. Bartlett decomposition and Wishart distribution. *The Annals of Mathematical Statistics*, 30(1):239–241, March 1959.

[35] G. Letac and H. Massam. A tutorial on non central Wishart distributions. *Technical Paper, Toulouse University*, October 2004.

[36] Alan L. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. *Working paper, Envision Financial Systems*, August 2001.

[37] Charles Van Loan. Computing integrals involving the matrix exponential. *Journal of Automatic Control*, 23(3):395–404, June 1978.

[38] Richard Lord and Christian Kahl. Complex logarithms in Heston-like models. *SSRN Working Paper Series*, February 2008.

[39] Roger Lord, Remmert Koekkoek, and Dick van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Working Paper, Tinbergen Institute*, May 2006.

[40] P.C Mahalabonis, R.C. Bose, and S.N. Roy. Normalization of statistical variates and the use of rectangular co-ordinates in the theory of sampling distributions. *Sankhyā*, 3:1–40, 1937.

[41] Beresford N. Parlett. *The Symmetric Eigenvalue Problem.* Society for Industrial Mathematics, Berlin, classics edition edition, 1987.

[42] Arthur Sepp. Fourier transform for option pricing under affine jump-diffusions: An overview. *Working paper, Tartu University*, September 2003.

[43] J. Y. Tourneret, A. Ferrari, and G. Letac. The noncentral Wishart distribution: Properties and application to speckle imaging. *Workshop on Stat. Signal Proc*, pages 1856–1860, 2005.

[44] Pelsser Van Haastrecht, Lord and Schrager. Pricing long-maturity equity and fx derivatives with stochastic interest rates and stochastic volatility. *SSRN*, 2008.

[45] Bo-Ying Wang, Bo-Yan Xi, and Fuzhen Zhang. Some inequalities for sum and product of positive semidefinite matrices. *Linear Algebra and its Applications*, 293:39–49, 1999.

[46] R.A. Wisjman. Random orthogonal transformations. *The Annals of Mathematical Statistics*, 28:415–422, 1957.

[47] Roger W.Lee. Option pricing by transform methods : Extensions, unification and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

A Technical Proofs

Proof of theorem 2. We first assume that we are in the relevant case for us, that is to say that almost surely the eigenvalues of the process $\{\xi_t\}_{t\geq 0}$ do not collide. We then introduce the following notations: $U_t^T \xi_t U_t = R_t$ and if $\{H_t\}_{t\geq 0}$ and $\{K_t\}_{t\geq 0}$ are two semi-martingales matrix-valued processes then

$$\langle dH, dK \rangle_t = \left(\sum_{k=1}^{n} d \left\langle H^{ik}, K^{kj} \right\rangle_t \right)_{1\leq i,j\leq n}.$$

Thanks to this notation we can write that

$$d\left(H_t K_t\right) = dH_t K_t + H_t dK_t + \langle dH, dK \rangle_t.$$

Besides, for $A \in O_n$ we can notice that

$$\left\langle dHA^T, AdK \right\rangle_t = \langle dH, dK \rangle_t.$$

Then, using Itō's lemma and the fact that $U_t \in O_n$, we have

$$0 = dU_t^T U_t + U_t^T dU_t + \left\langle dU^T, dU \right\rangle_t \tag{61}$$

$$\begin{aligned} dR_t &= dU_t^T \xi_t U_t + U_t^T d\xi_t U_t + U_t^T \xi_t dU_t + \left\langle dU^T, d\xi U \right\rangle_t \\ &+ \left\langle U^T d\xi, dU \right\rangle_t + \left\langle dU^T, \xi dU \right\rangle_t \\ &= dU_t^T U_t R_t + U_t^T d\xi_t U_t + R_t U_t^T dU_t + \left\langle dU^T U, U^T d\xi U \right\rangle_t \\ &+ \left\langle U^T d\xi U, U^T dU \right\rangle_t + \left\langle dU^T U, RU^T dU \right\rangle_t. \end{aligned} \tag{62}$$

Using notations already introduced, we then have from (62)

$$d\lambda_t^i = 2\left(dU_t^T U_t\right)^{ii} \lambda_t^i + dM_t^{ii} + dY_t^{ii} + 2\left\langle dM, U^T dU \right\rangle_t^{ii} + \left\langle dU^T U, RU^T dU \right\rangle_t^{ii}, \tag{63}$$

and for $i \neq j$

$$0 = \left(dU_t^T U_t\right)^{ij} \lambda_t^j + \left(U_t^T dU_t\right)^{ij} \lambda_t^i + dM_t^{ij} + dX_t^{ij}, \tag{64}$$

where X is a process with finite variation.

Using (61) and (64) we then get, since M is symmetric

$$\Gamma_t^{ij}dt = \langle dM^{ij}, dM^{ij}\rangle = \left\langle \left(dU^TU\right)^{ij}\lambda^j + \left(U^TdU\right)^{ij}\lambda^i, \left(dU^TU\right)^{ij}\lambda^j + \left(U^TdU\right)^{ij}\lambda^i\right\rangle_t$$

$$= \left(\lambda_t^i - \lambda_t^j\right)^2 \left\langle \left(dU^TU\right)^{ij}, \left(U^TdU\right)^{ji}\right\rangle_t, \tag{65}$$

and

$$\Gamma_t^{ij}dt = \langle dM^{ij}, dM^{ij}\rangle = \left\langle dM^{ij}, -\left(dU^TU\right)^{ij}\lambda^j - \left(U^TdU\right)^{ij}\lambda^i\right\rangle_t$$

$$= \left(\lambda_t^i - \lambda_t^j\right)\left\langle dM^{ij}, \left(U^TdU\right)^{ji}\right\rangle_t. \tag{66}$$

Therefore we have

$$\left\langle dU^TU, RU^TdU\right\rangle_t^{ii} = \left\langle dU^T, dU\right\rangle_t^{ii}\lambda_t^i + \sum_{j\neq i}\left(\lambda_t^i - \lambda_t^j\right)\left\langle \left(dU^TU\right)^{ij}, \left(U^TdU\right)^{ji}\right\rangle$$

$$= -2\left(dU_t^TU_t\right)^{ii} - \sum_{j\neq i}\frac{\Gamma_t^{ij}dt}{\lambda_t^i - \lambda_t^j},$$

and

$$\left\langle dM, U^TdU\right\rangle_t^{ii} = \sum_{j\neq i}\frac{\Gamma_t^{ij}dt}{\lambda_t^i - \lambda_t^j},$$

which ends the proof. □

Proof of proposition 3. Let $f : S_n \to \mathbb{R}$ be a two times differentiable function. Thanks to Itō's lemma, we have:

$$df(V_t) = \sum_{i,j=1}^{n}\frac{\partial f}{\partial V^{ij}}(V_t)\,dV_t^{ij} + \frac{1}{2}\sum_{i,j,k,l=1}^{n}\frac{\partial^2 f}{\partial V^{ij}\partial V^{kl}}(V_t)\,d\left\langle V^{ij}, V^{kl}\right\rangle_t. \tag{67}$$

Then from (1), we get:

$$dV_t^{ij} = \left(\beta Q^TQ + MY_t + Y_tM^T\right)^{ij}dt + \sum_{p,q=1}^{n}\left(\sqrt{V_t}\right)^{ip}(dW_t)^{pq}Q^{qj} + Q^{qi}(dW_t)^{pq}\left(\sqrt{V_t}\right)^{jp}.$$

Thus

$$d\left\langle V^{ij},V^{kl}\right\rangle_t = \sum_{p,q=1}^{n}\left[\left(\sqrt{V_t}\right)^{ip}Q^{qj}+Q^{qi}\left(\sqrt{V_t}\right)^{jp}\right]\left[\left(\sqrt{V_t}\right)^{kp}Q^{ql}+Q^{qk}\left(\sqrt{V_t}\right)^{lp}\right]$$

$$= 4\sum_{p,q=1}^{n}\left(\sqrt{V_t}\right)^{ip}Q^{qj}\left(\sqrt{V_t}\right)^{kp}Q^{ql}$$

$$= 4V_t^{ik}\left(Q^TQ\right)^{jl}$$

Replacing those quantities in (67) leads then easily to the matrix formulation (5). □

Proof of proposition 10. Let $f:\mathbb{R}\times S_n^+\to\mathbb{R}$ be a two times differentiable function.

It is clear that the first two terms in (26) correspond to the infinitesimal generator of the process X_t and that the third one corresponds to the infinitesimal generator of the Wishart process. Therefore, the only non-trivial term comes from crossed-differentiation in Itō's formula. That is to say the term :

$$\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x\partial V^{ij}}\left(X_t,V_t\right)d\left\langle X,V^{ij}\right\rangle_t.$$

But we have

$$d\left\langle X,V^{ij}\right\rangle_t = \sum_{k,l,p=1}^{n}\left(\sqrt{V}\right)^{ik}\left(\sqrt{V}\right)^{pk}R^{lp}Q^{lj}+\sum_{k,l,p=1}^{n}\left(\sqrt{V}\right)^{jk}\left(\sqrt{V}\right)^{pk}R^{lp}Q^{li}$$

$$= \left(VR^TQ\right)^{ij}+\left(VR^TQ\right)^{ji}.$$

Therefore

$$\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x\partial\partial V^{ij}}\left(X_t,V_t\right)d\left\langle X,V^{ij}\right\rangle_t = 2\operatorname{Tr}\left(VR^TQD\right)\frac{\partial}{\partial x}.$$

□

Proof of proposition 11. From the system already obtained, it is clear that we have $b(\tau)=ik$. Then using known properties of the Trace operator, the equation for $A(\tau)$ reduces to

$$\frac{\partial A(\tau)}{\partial\tau} = A(\tau)M+\left(M^T+2ikR^TQ\right)A(\tau)+2A(\tau)Q^TQA(\tau)-\frac{k(k+i)}{2}I_n. \quad (68)$$

It is once again a non-symmetric matrix Riccati equation which can be solved by the same linearization method that we already used. Namely, put

$$A(\tau) = F(\tau)^{-1}G(\tau), \tag{69}$$

for $F(h) \in GL_n(\mathbb{C})$, $G(h) \in M_n(\mathbb{C})$, $F(0) = I_n$ and $G(0) = 0$.

With these new variables the equation (68) becomes after identification of the terms containing $A(\tau)$

$$\begin{aligned}
\frac{dG}{d\tau}(\tau) &= -\frac{k(k+i)}{2}F(\tau) + G(\tau)M \\
\frac{dF}{d\tau}(\tau) &= -F(\tau)\left(M^T + 2ikR^TQ\right) - 2G(\tau)Q^TQ.
\end{aligned}$$

The system is solved by

$$\begin{pmatrix} G(\tau) & F(\tau) \end{pmatrix} = \begin{pmatrix} 0 & I_n \end{pmatrix} \exp\ \tau \begin{pmatrix} M & -2Q^TQ \\ -\frac{k(k+i)}{2}I_n & -\left(M^T + 2ikR^TQ\right) \end{pmatrix}. \tag{70}$$

Since $A(0) = 0$ the result for $A(\tau)$ follows.

For $c(\tau)$ we use the same trick as the one used when computing the characteristic function of the Wishart process. We have

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

and plugging into (69) and using the properties of the trace we deduce

$$\frac{dc}{d\tau}(\tau) = -\frac{\beta}{2}\operatorname{Tr}\left(F(\tau)^{-1}\frac{dF}{d\tau}(\tau) + M^T + 2ikR^TQ\right) + ikr.$$

And we can easily integrate to get:

$$c(\tau) = -\frac{\beta}{2}\left(\tau \operatorname{Tr}\left(M^T + 2ikR^TQ\right) + \operatorname{Tr}\left(\log F(\tau)\right)\right) + ikr\tau.$$

The logarithm of the complex matrix $F(\tau)$ is well-defined since by definition $F(\tau) \in GL_n(\mathbb{C})$, and using properties of the matrix logarithm we get:

$$c(\tau) = -\frac{\beta}{2}\left(\tau \operatorname{Tr}\left(M^T + 2ikR^TQ\right) + \log\left(\det F(\tau)\right)\right) + ikr\tau,$$

which ends the proof.

□

Proof of proposition 14. If there exists a positive solution X to (47), then its trace is positive. Let us note $\nu = \mathrm{Tr}(X)$. We have by simple calculations

$$\left(X + \frac{\nu}{2a} I_n\right)^2 = \frac{b}{a} M + \frac{\nu^2}{4a^2} I_n,$$

and therefore since $\frac{b}{a}M + \frac{\nu^2}{4a^2} I_n$ is positive as a sum of two positive matrix

$$X + \frac{\nu}{2a} I_n = \sqrt{\frac{b}{a} M + \frac{\nu^2}{4a^2} I_n}. \tag{71}$$

If we now take the trace, we have

$$-\nu\left(1 + \frac{n}{2a}\right) + \mathrm{Tr}\left[\sqrt{\frac{b}{a} M + \frac{\nu^2}{4a^2} I_n}\right] = 0,$$

which means that the trace of the matrix X depends only on the matrix M.

Let us now note

$$f(x) = -x\left(1 + \frac{n}{2a}\right) + \mathrm{Tr}\left[\sqrt{\frac{b}{a} M + \frac{x^2}{4a^2} I_n}\right].$$

The second derivative of f is simple to compute and one can easily see that the function f is convex since M is positive. Besides, we have $f(0) > 0$ and $f(x) \underset{x \to +\infty}{\to} -\infty$. Therefore the equation has a unique positive solution.

The result for X can then be obtained thanks to (71).

In the particular case where $n = 2$, it is possible to give explicitly the solution. Indeed, we first have to remark the following two identities which are true only in this case

$$X\,\mathrm{Tr}(X) = X^2 + \det(X) I_2$$

$$\mathrm{Tr}(X^2) = \mathrm{Tr}^2(X) - 2\det(X).$$

Therefore, by taking the trace of (47) we get

$$\mathrm{Tr}^2(X)\left(1 + \frac{1}{a}\right) = 2\det(X) + \frac{b}{a}\,\mathrm{Tr}(M), \tag{72}$$

and by taking the determinant of (47)

$$\det(X)^2 + \operatorname{Tr}^2(X)\left(1+\frac{1}{a}\right)\frac{\det(X)}{a} - \frac{b^2}{a^2}\det(M) = 0$$

$$\text{i.e.} \quad \left(1+\frac{2}{a}\right)\det(X)^2 + \frac{b}{a^2}\operatorname{Tr}(M)\det(X) - \frac{b^2}{a^2}\det(M) = 0.$$

The unique positive solution is given by

$$\det(X) = \frac{b}{2a(a+2)}\left(-\operatorname{Tr}(M) + \sqrt{\operatorname{Tr}^2(M) + 4a(a+2)\det(M)}\right),$$

which, thanks to (72), leads to

$$\operatorname{Tr}(X) = \sqrt{\frac{b}{(a+1)(a+2)}}\sqrt{(a+1)\operatorname{Tr}(M) + \sqrt{\operatorname{Tr}^2(M) + 4a(a+2)\det(M)}},$$

which ends the proof. □

Proof of Theorem 3. First of all, we can remark that if $X \sim \mathcal{W}_\beta(0, I_n)$ and if $\Sigma \in S_n^+((R))$, then $\sqrt{\Sigma}X\sqrt{\Sigma} \sim \mathcal{W}_\beta(0,\Sigma)$. Indeed, by computing the Laplace transform, we have for a symmetric matrix S so that $I_n + 2\Sigma S \in \widetilde{S}_n^+$

$$\begin{aligned}
\mathbb{E}\left[e^{-\operatorname{Tr}\left(S\sqrt{\Sigma}X\sqrt{\Sigma}\right)}\right] &= \mathbb{E}\left[e^{-\operatorname{Tr}\left(\sqrt{\Sigma}S\sqrt{\Sigma}X\right)}\right] \\
&= \frac{1}{\left[\det\left(I_n + 2\sqrt{\Sigma}S\sqrt{\Sigma}\right)\right]^{\beta/2}}.
\end{aligned}$$

Then, since the matrix I_n and $2\sqrt{\Sigma}S\sqrt{\Sigma}$ commute, they can be codiagonalized, and the eigenvalues of their sum are the $1+\lambda$, where the λ are eigenvalues of $2\sqrt{\Sigma}S\sqrt{\Sigma}$. Since $2\sqrt{\Sigma}S\sqrt{\Sigma}$ and $2\Sigma S$ have the same eigenvalues, we finally have

$$\mathbb{E}\left[e^{-\operatorname{Tr}\left(S\sqrt{\Sigma}X\sqrt{\Sigma}\right)}\right] = \frac{1}{\left[\det\left(I_n + 2\Sigma S\right)\right]^{\beta/2}},$$

which proves the first result.

Therefore, it suffices to prove that $AA^T \sim \mathcal{W}_\beta(0, I_n)$. We will do it using recursion. Let S be a symmetric matrix so that $I_n + 2S \in \widetilde{S}_n^+$ and let us first define the following matrix sequences

$$S_n = S$$

$$\forall\ 1 \le i \le n-2,\ S_{n-i} = \begin{pmatrix} S^{ii} & g_{n-i}^T \\ g_{n-i} & S_{n-i-1} \end{pmatrix}$$

$$S_1 = S^{nn},$$

and

$$A_n = A$$

$$\forall\ 1 \le i \le n-2,\ A_{n-i} = \begin{pmatrix} A^{ii} & 0 \\ v_{n-i} & A_{n-i-1} \end{pmatrix}$$

$$A_1 = A^{nn}.$$

We then have

$$\begin{aligned} \operatorname{Tr}\left[SAA^T\right] &= \operatorname{Tr}\left[S_n A_n A_n^T\right] \\ &= \operatorname{Tr}\left[S_{n-1}A_{n-1}A_{n-1}^T\right] + S^{11}(A^{11})^2 + 2A^{11}g_{n-1}^T v_{n-1} + v_{n-1}^T S_{n-1} v_{n-1}. \end{aligned}$$

Therefore by simple iteration and using the independence of the entries of A, we get

$$\begin{aligned} \mathbb{E}\left[e^{-\operatorname{Tr}\left(SAA^T\right)}\right] &= \mathbb{E}\left[e^{-S^{nn}(A^{nn})^2}\right] \prod_{i=1}^{n-1} \mathbb{E}\left[e^{-S^{ii}(A^{ii})^2 + 2A^{ii}g_{n-i}^T v_{n-i} + v_{n-i}^T S_{n-i} v_{n-i}}\right] \\ &= (1+2S^{nn})^{-\frac{\beta-n+1}{2}} \prod_{i=1}^{n-1} \mathbb{E}\left[e^{-S^{ii}(A^{ii})^2 + 2A^{ii}g_{n-i}^T v_{n-i} + v_{n-i}^T S_{n-i} v_{n-i}}\right]. \end{aligned}$$

Since the law of the entries of A is known, we have

$$\mathbb{E}\left[e^{-S^{ii}(A^{ii})^2+2A^{ii}g_{n-i}^T v_{n-i}+v_{n-i}^T S_{n-i}v_{n-i}}\right]$$

$$= \frac{1}{(2\pi)^{\frac{n-i}{2}}}\int_{\mathbb{R}^{n-i}}\int_0^{+\infty} e^{-\frac{\|v\|^2}{2}-v^T S_{n-i}v-2\sqrt{x}v^T g_{n-i}}\frac{e^{-xS^{ii}-\frac{x}{2}}x^{\frac{\beta-i-1}{2}}}{\Gamma\left(\frac{\beta-i+1}{2}\right)2^{\frac{\beta-i+1}{2}}}dxdv$$

$$= \frac{1}{(2\pi)^{\frac{n-i}{2}}\left[\det\left(I_{n-i}+2S_{n-i}\right)\right]^{1/2}}\int_{\mathbb{R}^{n-i}}\int_0^{+\infty} e^{-\frac{\|u\|^2}{2}-2\sqrt{x}u^T(I_{n-i}+2S_{n-i})^{-1/2}g_{n-i}}\times$$

$$\frac{e^{-\frac{x}{2}\left(1-2S^{ii}\right)}x^{\frac{\beta-i-1}{2}}}{\Gamma\left(\frac{\beta-i+1}{2}\right)2^{\frac{\beta-i+1}{2}}}dxdu$$

$$= \frac{1}{\left[\det\left(I_{n-i}+2S_{n-i}\right)\right]^{1/2}}\int_0^{+\infty}\frac{e^{-\frac{x}{2}\left(1-2S^{ii}\right)+2xg_{n-i}^T(I_{n-i}+2S_{n-i})^{-1}g_{n-i}}x^{\frac{\beta-i-1}{2}}}{\Gamma\left(\frac{\beta-i+1}{2}\right)2^{\frac{\beta-i+1}{2}}}dx$$

$$= \frac{1}{\left[\det\left(I_{n-i}+2S_{n-i}\right)\right]^{1/2}\left(1+2S^{ii}-4g_{n-i}^T\left(I_{n-i}+2S_{n-i}\right)^{-1}g_{n-i}\right)^{\frac{\beta-i+1}{2}}}.$$

But we can also easily show, using known results on the Schur's complement, that

$$1+2S^{ii}-4g_{n-i}^T\left(I_{n-i}+2S_{n-i}\right)^{-1}g_{n-i} = \frac{\det\left(I_{n-i+1}+2S_{n-i+1}\right)}{\det\left(I_{n-i}+2S_{n-i}\right)},$$

where all the quantities above are defined since $I_n+2S\in\widetilde{S}_n^+$, which means that all the leading minors of the matrix I_n+2S are definite positive.

Then, we get

$$\prod_{i=1}^{n-1}\left[1+2S^{ii}-4g_{n-i}^T\left(I_{n-i}+2S_{n-i}\right)^{-1}g_{n-i}\right]^{\frac{\beta-i+1}{2}} = \frac{\left(1+2S^{nn}\right)^{\frac{n-\beta}{2}-1}\left[\det(I_n+2S)\right]^{\frac{\beta+1}{2}}}{\prod_{i=1}^{n-1}\left[\det\left(I_{n-i+1}-2S_{n-i+1}\right)\right]^{1/2}},$$

and thus

$$\mathbb{E}\left[e^{-\operatorname{Tr}\left(SAA^T\right)}\right] = \frac{\left(1+2S^{nn}\right)^{\frac{1}{2}}}{\left[\det(I_n+2S)\right]^{\frac{\beta+1}{2}}}\left[\prod_{i=1}^{n-1}\frac{\det(I_{n-i+1}+2S_{n-i+1})}{\det(I_{n-i}+2S_{n-i})}\right]^{1/2}$$

$$= \frac{1}{\left[\det\left(I_n+2S\right)\right]^{\beta/2}},$$

which ends the proof.

□

B Numerical Considerations

B.1 On the Matrix Exponential

There are many methods to compute the exponential of a given matrix. In all our simulations we have chosen to use an algorithm of Higham (see [31]) which is actually a slight modification of the so called method of Scaling and Squaring coupled with the use of Padé's approximation. Roughly, the exponential of the matrix A is computed thanks to the following approximation

$$e^A \approx R_{pq}(A),$$

where $(p,q) \in \mathbb{N}^* \times \mathbb{N}^*$ and

$$R_{pq}(A) = D_{pq}(A)^{-1} N_{pq}(A)$$

$$N_{pq}(A) = \sum_{j=0}^{p} \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!} A^j$$

$$D_{pq}(A) = \sum_{j=0}^{q} \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!} A^j.$$

Then, since the approximation is good only when the norm of the matrix is near zero, the algorithm choose an integer h and compute, instead of $R_{pq}(A)$, $\left[R_{pq}(2^{-h}A)\right]^{2^h}$.

Besides, when the dimension of the matrix is 2, the exponential has a simple closed form which can be proven (it suffices to calculate the eigenvalues of A and to diagonalize or trigonalize it) to be

Proposition 16. *Let $A \in M_2(\mathbb{C})$, and let $\Delta = \sqrt{(A^{11}-A^{22})^2 + 4A^{12}A^{21}}$. We have*

- *If $\Delta \neq 0$*

$$e^A = \frac{e^{(A^{11}+A^{22})/2}}{\Delta} \begin{pmatrix} \Delta \mathrm{ch}(\Delta/2) + (A^{11}-A^{22})\mathrm{sh}(\Delta/2) & 2A^{12}\mathrm{sh}(\Delta/2) \\ 2A^{21}\mathrm{sh}(\Delta/2) & \Delta \mathrm{ch}(\Delta/2) - (A^{11}-A^{22})\mathrm{sh}(\Delta/2) \end{pmatrix}.$$

- *If $\Delta = 0$*

$$e^A = e^{(A^{11}+A^{22})/2} \begin{pmatrix} 1 + \frac{1}{2}(A^{11}-A^{22}) & A^{12} \\ A^{21} & 1 - \frac{1}{2}(A^{11}-A^{22}) \end{pmatrix}.$$

B.2 On the Matrix Square Root

For all our calculations, we need a fast and accurate way to compute the matrix square-root of symmetric definite or semi-definite matrix A. The simplest algorithm for these kind of computations all use iterations procedures. One of the most efficient has been proposed by Higham in [29] and [30] and can be summed up as follows

Algorithm 2. *1. Compute the Cholesky factorization of A, $A = R^T R$.*

2. Compute $U = X_\infty$ from the iteration

$$X_0 = R$$

$$\forall k \geq 1,\ X_{k+1} = \frac{1}{2}\left(X_k + \left(X_k^{-1}\right)^T\right).$$

3. $\sqrt{A} = U^T R$.

Besides the convergence of the iterations can be readily accelerated by the incorporation of scaling parameters (see [29] for more details).

B.3 On the Matrix Smallest and Largest Eigenvalues

In our generalization of the QE algorithm to Wishart processes, we need to be able to compute the smallest and largest eigenvalues of a given (semi)definite matrix A. Several algorithms are known for this purpose, among which the most efficient seem to be the power iteration algorithm, the inverse power iteration algorithm and the Rayleigh Quotient algorithm. We do not detail here those well-known algorithms and refer instead to [26] and [41].

However, in order to avoid a lot of computations, we decided in our implementations to use the fact that in dimensions 2 and 3 the largest and smallest eigenvalues of a semi-definite matrix can easily be calculated at a small computational cost. Indeed, we have

Proposition 17. *Let $A \in S_n^+(\mathbb{R})$. Then we have*

1. If $n = 2$

$$\lambda_{min} = \frac{\mathrm{Tr}^2(A) - \sqrt{\mathrm{Tr}^2(A) - 4\det(A)}}{2}$$

$$\lambda_{max} = \frac{\mathrm{Tr}^2(A) + \sqrt{\mathrm{Tr}^2(A) - 4\det(A)}}{2}$$

2. *If* $n = 3$ *Let* $a = \frac{1}{2}(\mathrm{Tr}(A^2) - \mathrm{Tr}^2(A))$, $b = \mathrm{Tr}(A)$ *and* $c = -\det(A)$. *Define* $m = 2a^3 - 9ab + 27c$, $k = a^2 - 3b$ *and* $g = m^2 - 4k^3$. *Consider now the complex number*

$$\left(\frac{m + \sqrt{g}}{2}\right)^{1/3} := \alpha + i\beta,$$

where $(\alpha, \beta) \in \mathbb{R}^2$.

Then we have

$$\lambda_{min} = \frac{1}{3}\min\left(-a - 2\alpha, -a + \alpha - \sqrt{3}\beta, -a + \alpha + \sqrt{3}\beta\right)$$

$$\lambda_{max} = \frac{1}{3}\max\left(-a - 2\alpha, -a + \alpha - \sqrt{3}\beta, -a + \alpha + \sqrt{3}\beta\right)$$

Proof. We only need to compute the characteristic polynomial of the matrix A in both cases and to solve it. For $n = 2$ the resolution is trivial and for $n = 3$ it suffices to use the well-known formulas due to Tartaglia. □

B.4 On the Simulation of the Central Chi-Square law

According to theorem 3, in order to simulate central-Wishart random variables, we have to simulate central chi-square random variables. In order to do so, we prefer avoiding the method based on acceptance/rejection algorithms and prefer using the fact that the inverse of the cumulative distribution function of the chi-square law can be easily approximated (see [15]).

C Numerical Results

All our numerical results have been performed with the following parameters.

Parameter	**Value**
Spot	61.90
Short rate	3%
Times Steps per Year	24
β	4 or 4.8

Table 2: Parameters used for Monte-Carlo simulation

V_0	Q	M	R
$\begin{pmatrix} 50\% & 0 \\ 0 & 50\% \end{pmatrix}$	$\begin{pmatrix} 35\% & 0 \\ 0 & 40\% \end{pmatrix}$	$\begin{pmatrix} -0.6 & 0 \\ 0 & -0.4 \end{pmatrix}$	$\begin{pmatrix} -0.5 & 0 \\ 0 & -0.5 \end{pmatrix}$

Table 3: Parameters used for Monte-Carlo simulation

For each discretization scheme, we have computed call prices for various Monte Carlo simulations number (ranging from 10000 to 1000000), strikes (70%, 100% and 130%) and maturities (6 months and 3 years). In each table, we provide the price obtained by Monte Carlo simulation, the corresponding standard deviation, the corresponding relative error, the price obtained by Control Variate, the corresponding standard deviation and the corresponding relative error.
The error is computed with respect to the theoretical price given by formula 35. Table 4 lists the numerical prices for our different test cases.

Strikes	**Term**	$\beta = 4$	$\beta = 4.8$
100%	6M	13.7223	14.2210
100%	3	36.3354	38.5084
70%	6M	23.6987	24.0157
70%	3	41.2395	42.9492
130%	6M	7.6089	8.1435
130%	3	32.4443	34.9568

Table 4: Theoretical call prices

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
10000	100%	6M	13.40	0.25	2.35%	13.56	0.10	1.20%
20000	100%	6M	13.58	0.18	1.06%	13.55	0.07	1.24%
50000	100%	6M	13.53	0.18	1.39%	13.55	0.08	1.27%
100000	100%	6M	13.59	0.08	0.97%	13.65	0.03	0.56%
1000000	100%	6M	13.58	0.03	1.03%	13.61	0.01	0.83%
10000	70%	6M	23.38	0.30	1.35%	23.58	0.06	0.49%
20000	70%	6M	23.57	0.22	0.55%	23.54	0.05	0.68%
50000	70%	6M	23.51	0.23	0.78%	23.53	0.09	0.70%
100000	70%	6M	23.54	0.10	0.65%	23.62	0.02	0.35%
1000000	70%	6M	23.56	0.03	0.59%	23.60	0.01	0.43%
10000	130%	6M	7.33	0.19	3.68%	7.44	0.11	2.21%
20000	130%	6M	7.52	0.14	1.18%	7.50	0.08	1.40%
50000	130%	6M	7.45	0.12	2.12%	7.46	0.07	1.97%
100000	130%	6M	7.50	0.06	1.41%	7.54	0.03	0.88%
1000000	130%	6M	7.49	0.02	1.52%	7.51	0.01	1.25%
10000	100%	3Y	35.82	1.70	1.42%	36.37	0.20	0.08%
20000	100%	3Y	35.06	0.96	3.52%	36.37	0.14	0.09%
50000	100%	3Y	36.11	0.64	0.62%	36.37	0.09	0.10%
100000	100%	3Y	37.31	0.56	2.67%	36.46	0.06	0.34%
1000000	100%	3Y	36.37	0.15	0.09%	36.32	0.02	0.05%
10000	70%	3Y	40.73	1.72	1.23%	41.29	0.15	0.12%
20000	70%	3Y	39.92	0.99	3.21%	41.27	0.10	0.07%
50000	70%	3Y	41.00	0.65	0.59%	41.27	0.06	0.07%
100000	70%	3Y	42.19	0.57	2.29%	41.32	0.05	0.20%
1000000	70%	3Y	41.28	0.16	0.09%	41.22	0.01	0.04%
10000	130%	3Y	31.96	1.67	1.49%	32.49	0.25	0.16%
20000	130%	3Y	31.19	0.94	3.86%	32.47	0.17	0.07%
50000	130%	3Y	32.24	0.62	0.62%	32.50	0.11	0.17%
100000	130%	3Y	33.42	0.55	3.02%	32.59	0.08	0.45%
1000000	130%	3Y	32.47	0.15	0.09%	32.42	0.02	0.06%

Table 5: Call prices for Euler discretization scheme in the integer case

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
10000	100%	6M	13.85	0.26	2.59%	14.02	0.10	1.39%
20000	100%	6M	14.05	0.19	1.20%	14.02	0.07	1.39%
50000	100%	6M	14.00	0.19	1.57%	14.01	0.09	1.45%
100000	100%	6M	14.05	0.08	1.18%	14.11	0.03	0.76%
1000000	100%	6M	14.05	0.03	1.23%	14.08	0.01	1.02%
10000	70%	6M	23.66	0.31	1.47%	23.88	0.07	0.58%
20000	70%	6M	23.86	0.23	0.63%	23.83	0.05	0.77%
50000	70%	6M	23.81	0.24	0.88%	23.83	0.10	0.79%
100000	70%	6M	23.84	0.10	0.74%	23.91	0.02	0.43%
1000000	70%	6M	23.85	0.03	0.67%	23.89	0.01	0.52%
10000	130%	6M	7.82	0.21	4.02%	7.94	0.11	2.52%
20000	130%	6M	8.03	0.15	1.43%	8.01	0.08	1.67%
50000	130%	6M	7.95	0.13	2.38%	7.96	0.07	2.23%
100000	130%	6M	8.00	0.07	1.76%	8.04	0.04	1.23%
1000000	130%	6M	7.99	0.02	1.86%	8.01	0.01	1.59%
10000	100%	3Y	37.85	2.09	6.98%	38.53	0.21	5.56%
20000	100%	3Y	37.02	1.16	8.97%	38.54	0.14	5.56%
50000	100%	3Y	38.19	0.77	6.23%	38.54	0.09	5.54%
100000	100%	3Y	39.73	0.72	3.12%	38.62	0.07	5.32%
1000000	100%	3Y	38.53	0.19	5.56%	38.49	0.02	5.69%
10000	70%	3Y	42.30	2.11	5.16%	42.98	0.15	3.86%
20000	70%	3Y	41.44	1.18	7.06%	42.99	0.10	3.91%
50000	70%	3Y	42.61	0.78	4.54%	42.97	0.07	3.92%
100000	70%	3Y	44.16	0.73	1.78%	43.03	0.05	3.79%
1000000	70%	3Y	42.98	0.19	3.89%	42.93	0.01	4.02%
10000	130%	3Y	34.34	2.07	8.57%	35.01	0.26	7.04%
20000	130%	3Y	33.49	1.14	10.8%	34.98	0.17	7.13%
50000	130%	3Y	34.67	0.76	7.76%	35.02	0.11	7.03%
100000	130%	3Y	36.18	0.71	4.39%	35.08	0.08	6.77%
1000000	130%	3Y	34.98	0.18	7.10%	34.94	0.02	7.25%

Table 6: Call prices for Euler discretization scheme in the non-integer case

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
10000	100%	6M	14.22	0.26	3.61%	13.56	0.10	1.18%
20000	100%	6M	14.08	0.18	2.64%	13.55	0.07	1.29%
50000	100%	6M	13.68	0.11	0.29%	13.56	0.04	1.21%
100000	100%	6M	13.77	0.08	0.36%	13.58	0.03	1.00%
1000000	100%	6M	13.65	0.03	0.50%	13.60	0.01	0.87%
10000	70%	6M	24.34	0.32	2.70%	23.51	0.06	0.80%
20000	70%	6M	24.28	0.22	2.45%	23.60	0.05	0.43%
50000	70%	6M	23.71	0.14	0.03%	23.55	0.03	0.64%
100000	70%	6M	23.81	0.10	0.45%	23.57	0.02	0.55%
1000000	70%	6M	23.65	0.03	0.19%	23.59	0.01	0.46%
10000	130%	6M	7.93	0.21	4.23%	7.46	0.11	2.02%
20000	130%	6M	7.87	0.15	3.42%	7.48	0.08	1.69%
50000	130%	6M	7.59	0.09	0.31%	7.50	0.05	1.50%
100000	130%	6M	7.67	0.06	0.75%	7.53	0.03	1.02%
1000000	130%	6M	7.57	0.02	0.52%	7.53	0.01	1.00%
10000	100%	3Y	37.70	1.35	3.76%	36.32	0.19	0.03%
20000	100%	3Y	39.30	1.12	8.15%	36.47	0.14	0.36%
50000	100%	3Y	37.90	0.72	4.32%	36.36	0.09	0.07%
100000	100%	3Y	37.50	0.50	3.19%	36.34	0.06	0.00%
1000000	100%	3Y	36.41	0.15	0.19%	36.45	0.02	0.33%
10000	70%	3Y	42.70	1.38	3.54%	41.28	0.14	0.10%
20000	70%	3Y	44.25	1.14	7.30%	41.35	0.10	0.28%
50000	70%	3Y	42.85	0.73	3.90%	41.27	0.07	0.07%
100000	70%	3Y	42.43	0.51	2.89%	41.25	0.05	0.02%
1000000	70%	3Y	41.27	0.15	0.08%	41.32	0.01	0.20%
10000	130%	3Y	33.67	1.31	3.78%	32.34	0.23	0.33%
20000	130%	3Y	35.35	1.10	8.96%	32.59	0.17	0.45%
50000	130%	3Y	33.95	0.71	4.64%	32.44	0.11	0.01%
100000	130%	3Y	33.59	0.49	3.52%	32.46	0.08	0.03%
1000000	130%	3Y	32.55	0.15	0.31%	32.59	0.02	0.46%

Table 7: Call prices for Ornstein discretization scheme in the integer case

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
10000	100%	6M	13.80	0.25	2.97%	13.96	0.13	1.86%
20000	100%	6M	14.19	0.18	0.19%	14.09	0.09	0.96%
50000	100%	6M	14.28	0.12	0.45%	14.06	0.06	1.17%
100000	100%	6M	14.20	0.08	0.14%	14.06	0.05	1.11%
1000000	100%	6M	14.20	0.03	0.16%	13.97	0.01	1.76%
10000	70%	6M	23.62	0.30	1.67%	23.82	0.11	0.82%
20000	70%	6M	24.01	0.22	0.01%	23.87	0.08	0.59%
50000	70%	6M	24.17	0.14	0.64%	23.88	0.05	0.57%
100000	70%	6M	24.07	0.11	0.22%	23.89	0.04	0.51%
1000000	70%	6M	24.08	0.03	0.26%	23.79	0.01	0.95%
10000	130%	6M	7.82	0.20	3.92%	7.94	0.12	2.53%
20000	130%	6M	8.07	0.14	0.86%	7.99	0.09	1.83%
50000	130%	6M	8.20	0.09	0.67%	8.03	0.06	1.40%
100000	130%	6M	8.11	0.07	0.40%	8.01	0.04	1.63%
1000000	130%	6M	8.11	0.02	0.46%	7.94	0.01	2.49%
10000	100%	3Y	40.47	1.80	5.10%	37.65	0.25	2.23%
20000	100%	3Y	38.95	1.11	1.15%	37.99	0.17	1.35%
50000	100%	3Y	40.80	0.84	5.94%	37.96	0.11	1.42%
100000	100%	3Y	40.83	0.86	6.03%	37.90	0.08	1.57%
1000000	100%	3Y	40.21	0.20	4.42%	37.95	0.03	1.45%
10000	70%	3Y	45.12	1.84	5.06%	42.24	0.20	1.65%
20000	70%	3Y	43.50	1.14	1.27%	42.51	0.13	1.02%
50000	70%	3Y	45.39	0.86	5.68%	42.49	0.09	1.06%
100000	70%	3Y	45.40	0.87	5.72%	42.45	0.06	1.17%
1000000	70%	3Y	44.79	0.21	4.28%	42.49	0.02	1.07%
10000	130%	3Y	36.76	1.77	5.16%	34.00	0.30	2.73%
20000	130%	3Y	35.30	1.09	0.98%	34.37	0.20	1.69%
50000	130%	3Y	37.13	0.83	6.22%	34.35	0.13	1.73%
100000	130%	3Y	37.18	0.86	6.35%	34.28	0.10	1.94%
1000000	130%	3Y	36.55	0.20	4.55%	34.33	0.03	1.80%

Table 8: Call prices for Ornstein discretization scheme in the non-integer case

MC sims	**Strike**	**Term**	**MC**	**Std-Dev**	**Error**	**CV**	**Std-Dev**	**Error**
10000	100%	6M	13.34	0.29	2.79%	13.69	0.10	0.27%
20000	100%	6M	14.10	0.21	2.78%	13.89	0.07	1.22%
50000	100%	6M	13.66	0.13	0.44%	13.81	0.05	0.61%
100000	100%	6M	13.71	0.09	0.10%	13.79	0.03	0.48%
1000000	100%	6M	13.73	0.03	0.08%	13.75	0.01	0.19%
10000	70%	6M	22.90	0.34	3.39%	23.32	0.06	1.61%
20000	70%	6M	23.70	0.25	0.02%	23.44	0.04	1.08%
50000	70%	6M	23.23	0.15	1.97%	23.41	0.03	1.23%
100000	70%	6M	23.27	0.11	1.83%	23.36	0.02	1.42%
1000000	70%	6M	23.34	0.03	1.52%	23.36	0.01	1.44%
10000	130%	6M	7.81	0.24	2.66%	8.08	0.12	6.16%
20000	130%	6M	8.40	0.18	10.4%	8.23	0.09	8.16%
50000	130%	6M	8.02	0.11	5.35%	8.13	0.05	6.80%
100000	130%	6M	8.10	0.08	6.48%	8.16	0.04	7.30%
1000000	130%	6M	8.09	0.02	6.32%	8.10	0.01	6.48%
10000	100%	3Y	36.25	2.57	0.23%	37.61	0.20	3.50%
20000	100%	3Y	38.46	1.85	5.83%	37.69	0.14	3.74%
50000	100%	3Y	36.87	1.00	1.46%	37.53	0.09	3.30%
100000	100%	3Y	37.44	0.82	3.05%	37.33	0.06	2.74%
1000000	100%	3Y	37.25	0.30	2.53%	37.33	0.02	2.74%
10000	70%	3Y	40.47	2.59	1.86%	41.84	0.15	1.46%
20000	70%	3Y	42.75	1.86	3.66%	41.98	0.10	1.80%
50000	70%	3Y	41.17	1.01	0.16%	41.85	0.07	1.48%
100000	70%	3Y	41.83	0.83	1.42%	41.71	0.05	1.15%
1000000	70%	3Y	41.63	0.31	0.94%	41.70	0.02	1.13%
10000	130%	3Y	32.97	2.55	1.63%	34.32	0.25	5.77%
20000	130%	3Y	35.10	1.83	8.19%	34.35	0.18	5.86%
50000	130%	3Y	33.53	0.99	3.34%	34.19	0.11	5.38%
100000	130%	3Y	34.05	0.82	4.94%	33.94	0.08	4.60%
1000000	130%	3Y	33.87	0.30	4.39%	33.94	0.03	4.62%

Table 9: Call prices for QE discretization scheme in the integer case

MC sims	Strike	Term	MC	Std-Dev	Error	CV	Std-Dev	Error
10000	100%	6M	13.72	0.30	6.20%	14.07	0.10	3.77%
20000	100%	6M	14.50	0.22	0.82%	14.27	0.07	2.33%
50000	100%	6M	14.04	0.14	3.93%	14.18	0.05	2.92%
100000	100%	6M	14.09	0.10	3.61%	14.17	0.03	3.04%
1000000	100%	6M	14.11	0.03	3.43%	14.13	0.01	3.33%
10000	70%	6M	23.14	0.35	4.67%	23.57	0.06	2.91%
20000	70%	6M	23.97	0.25	1.30%	23.69	0.05	2.39%
50000	70%	6M	23.48	0.16	3.26%	23.66	0.03	2.53%
100000	70%	6M	23.52	0.11	3.12%	23.61	0.02	2.72%
1000000	70%	6M	23.59	0.04	2.82%	23.61	0.01	2.74%
10000	130%	6M	8.20	0.25	4.08%	8.47	0.12	0.81%
20000	130%	6M	8.79	0.19	3.11%	8.62	0.09	1.06%
50000	130%	6M	8.40	0.11	1.57%	8.51	0.05	0.21%
100000	130%	6M	8.49	0.08	0.51%	8.55	0.04	0.25%
1000000	130%	6M	8.48	0.03	0.66%	8.49	0.01	0.51%
10000	100%	3Y	37.96	2.99	5.86%	39.36	0.20	2.34%
20000	100%	3Y	40.04	2.08	0.14%	39.39	0.14	2.11%
50000	100%	3Y	38.48	1.16	4.27%	39.26	0.09	2.53%
100000	100%	3Y	39.18	0.97	2.77%	39.06	0.07	3.06%
1000000	100%	3Y	38.96	0.37	3.26%	39.06	0.02	3.06%
10000	70%	3Y	41.86	3.01	5.77%	43.26	0.15	2.58%
20000	70%	3Y	44.02	2.09	0.47%	43.37	0.11	2.25%
50000	70%	3Y	42.45	1.17	4.13%	43.24	0.07	2.56%
100000	70%	3Y	43.23	0.97	2.61%	43.10	0.05	2.88%
1000000	70%	3Y	42.99	0.37	3.08%	43.10	0.02	2.90%
10000	130%	3Y	34.89	2.97	5.68%	36.27	0.25	1.83%
20000	130%	3Y	36.93	2.07	0.41%	36.29	0.18	1.74%
50000	130%	3Y	35.36	1.15	4.08%	36.14	0.11	2.19%
100000	130%	3Y	36.02	0.96	2.60%	35.90	0.08	2.92%
1000000	130%	3Y	35.80	0.37	3.12%	35.91	0.03	2.90%

Table 10: Call prices for QE discretization scheme in the non-integer case