ORIGINAL PAPER

On the application of Wishart process to the pricing of equity derivatives: the multi-asset case

Gaetano La Bua[1] · Daniele Marazzina[1]

Abstract

Given the inherent complexity of financial markets, a wide area of research in the field of mathematical finance is devoted to develop accurate models for the pricing of contingent claims. Focusing on the stochastic volatility approach (i.e. we assume to describe asset volatility as an additional stochastic process), it appears desirable to introduce reliable dynamics in order to take into account the presence of several assets involved in the definition of multi-asset payoffs. In this article we deal with the multi asset Wishart Affine Stochastic Correlation model, that makes use of Wishart process to describe the stochastic variance covariance matrix of assets return. The resulting parametrization turns out to be a genuine multi-asset extension of the Heston model: each asset is exactly described by a single instance of the Heston dynamics while the joint behaviour is enriched by cross-assets and cross-variances stochastic correlation, all wrapped in an affine modeling. In this framework, we propose a fast and accurate calibration procedure, and two Monte Carlo simulation schemes.

Keywords Wishart process · Calibration · Monte Carlo · Multi assets

JEL Classification C02 · C63 · G12 · G13

1 Introduction

A Wishart process is a matrix-valued continuous time stochastic process with a marginal Wishart distribution, i.e., a generalization to multiple dimensions of the chi-squared distribution, or, in the case of non-integer degrees of freedom, of the gamma distribution. While the introduction of Wishart-based Stochastic Volatility models in

✉ Daniele Marazzina
daniele.marazzina@polimi.it

Gaetano La Bua
gaetano.labua@polimi.it

[1] Department of Mathematics, Politecnico di Milano, 32 p.zza Leonardo da Vinci, 20133 Milan, Italy

⚇ Springer

finance is fully motivated by the need to describe the multidimensional structure of asset variances (see La Bua and Marazzina (2019) and references therein), in this article our goal is to exploit the Wishart process in a multi assets framework. More precisely, we deal with the Wishart Affine Stochastic Correlation model (WASC), introduced in Da Fonseca et al. (2007) with the purpose of reproducing well-known multi-asset stylized facts in a tractable way. WASC model makes use of Wishart process to describe the stochastic variance covariance matrix of asset returns.

In our analysis we focus on the so-called Wishart processes introduced in Bru (1991) as a matrix generalization of square-root processes. A remarkable feature is that the analytical tractability is fully preserved since these processes belong to the class of affine processes. Given the strict connection with the well-known CIR processes, Wishart processes have been used to define multi-factor (Da Fonseca et al. 2008; La Bua and Marazzina 2019) and multi-asset (Da Fonseca et al. 2007) extensions of the classic Heston model, which is one of the most known and used models in finance, see for example (Goudenege et al. 2019; Yolcu-Okur et al. 2018).

Despite the analytical tractability (La Bua and Marazzina 2019), the implementation of Wishart-based models poses non-trivial challenges from a numerical point of view. In this article we extend the results presented in La Bua and Marazzina (2019) for the single asset Wishart Multidimensional Stochastic Volatility model (WMSV) to the WASC one. More precisely, we deal with the model calibration problem, presenting an innovative and efficient methodology to calibrate WASC parameters to market data. The algorithm exploits the close link existing between the Heston model (Heston 1993) and marginal WASC dynamics. Considering the single-asset case, in La Bua and Marazzina (2019) we show that, for an appropriate choice of parameters, both Heston (Heston 1993) and the Bi-Heston (Christoffersen et al. 2009) models may provide a reliable approximation of WMSV. In this article, for the multi-asset case, we extend this approximating technique making use of the distributional law of diagonal elements of Wishart process to connect the WASC calibration problem to the Heston one. We provide the analytical form of the gradient of calibration problem objective function with respect to Wishart-based parameters allowing for a further reduction in the computational burden.

Additionally, we propose model approximations that permit us to introduce two numerical schemes for Monte Carlo simulations. It is well known that a standard discretization (e.g. Euler scheme) is unfeasible in the Wishart-based framework, since we also need to take into account the evolution of non diagonal elements of the Wishart matrix $\Sigma(t)$ to determine the dependence structure and satisfy the positive semi-definiteness constraint for the Wishart process. Therefore, as a first algorithm, we propose an adapted version of the scalar full truncated Euler. Secondly, we extend the Gaussian variable approximation scheme presented in La Bua and Marazzina (2019) to the WASC model. Extensive numerical results to compare the two simulation schemes are provided.

The article is organized as follows. In Sect. 2 we present the Wishart process and its basic properties, while in Sect. 3 we deal with the WASC model. Finally the calibration procedure is described in Sect. 4, while the Monte Carlo simulation schemes are presented in Sect. 5. In both cases, numerical results are presented.

2 Definition of Wishart process and basic properties

In this section, we introduce the Wishart process. We refer to La Bua and Marazzina (2019) for details.

Definition 1 (*Wishart process*) Let $W(t)$ be a $d \times d$ Brownian motion (i.e. a matrix of $d \times d$ independent scalar Brownian motions) and $\mathcal{S}_d^+(\mathbb{R})$ the set of real $d \times d$ positive semidefinite matrices. We define the Wishart process as the solution on $\mathcal{S}_d^+(\mathbb{R})$ of the following stochastic differential equation (SDE):

$$d\Sigma(t) = (\Omega\Omega^\top + M\Sigma(t) + \Sigma(t)M^\top)dt + \sqrt{\Sigma(t)}\,dW(t)\,Q + Q^\top\,dW^\top(t)\,\sqrt{\Sigma(t)},$$
$$\Sigma(0) = \Sigma_0 \in \mathcal{S}_d^+(\mathbb{R}) \tag{1}$$

with $\Omega, Q, M \in \mathcal{M}_d(\mathbb{R})$ (the set of real $d \times d$ square matrices).

As in La Bua and Marazzina (2019), in order to embed mean-reversion and stationarity, we consider matrix M to have only eigenvalues with negative real part. Moreover, we relate the deterministic part of the drift in (1), $\Omega\Omega^\top$, to the expected long-term value of the process, Σ_∞, by the equation

$$-\Omega\Omega^\top = M\Sigma_\infty + \Sigma_\infty M^\top. \tag{2}$$

As shown in La Bua and Marazzina (2019), if we set $d = 1$ in (1), we end up with a scalar CIR process defined by the SDE

$$dv(t) = \kappa(\theta - v(t))dt + \eta\sqrt{v(t)}dw_v(t), \quad v(0) = v_0, \tag{3}$$

with κ, θ, and η strictly positive parameters, $v_0 \geq 0$ and $w_v(t)$ a scalar Brownian motion. Therefore Wishart processes can be considered as multidimensional extension of a scalar CIR process. Moreover, a Wishart process entails a non-trivial dependence structure among its elements, since we have

$$d\left[\Sigma_{ij}(t), \Sigma_{kl}(t)\right] = \left(\Sigma_{ik}(t)Q_{jl}^* + \Sigma_{il}(t)Q_{jk}^* + \Sigma_{jk}(t)Q_{il}^* + \Sigma_{jl}(t)Q_{ik}^*\right)dt, \tag{4}$$

where the notation $[\cdot, \cdot]$ refers as usual to the quadratic covariation of two stochastic processes, $\Sigma_{ij}(t)$ indicates the element in the i-th row and $j-$th column of $\Sigma(t)$, and $Q^* = Q^\top Q$.

Existence and uniqueness conditions of the solution of the Wishart process SDE are given by the following results.

Proposition 1 (Proposition 2.1 in La Bua and Marazzina (2019)) *Let $X(t)$ be a generic affine process with continuous trajectories defined in $\mathcal{S}_d^+(\mathbb{R})$ by the following SDE*

$$X(t) = X(0) + \int_0^t (D_X + \mathcal{L}[X(s)])\,ds$$
$$+ \int_0^t \left(\sqrt{X(s)}\,dW(s)\,C_X + C_X^\top\,dW^\top(s)\,\sqrt{X(s)}\right), \tag{5}$$

where $X(0)$, $D_X \in \mathcal{S}_d^+(\mathbb{R})$, $C_X \in \mathcal{M}_d(\mathbb{R})$, $\mathcal{L} : \mathcal{S}_d^+(\mathbb{R}) \to \mathcal{S}_d^+(\mathbb{R})$ is a linear transformation. Such process admits a unique weak solution in $\mathcal{S}_d^+(\mathbb{R})$ if

(a) $D_X - (d-1)C_X^\top C_X \in \mathcal{S}_d^+(\mathbb{R})$,
(b) $\forall\, P_1, P_2 \in \mathcal{S}_d^+(\mathbb{R})$ s.t. $\mathrm{Tr}\,[P_1 P_2] = 0 \;\Rightarrow\; \mathrm{Tr}\,[\mathcal{L}(P_1)P_2] \geq 0$, where $\mathrm{Tr}\,[\cdot]$ is the trace of a square matrix (i.e., the sum of the elements on the main diagonal).

If $X(0)$ is in the set of real positive definite matrices $\mathcal{S}_d^{++}(\mathbb{R})$ and condition a) is replaced by the stronger requirement

(c) $D_X - (d+1)C_X^\top C_X \in \mathcal{S}_d^+(\mathbb{R})$,

then there exist a unique strong solution to (5) in $\mathcal{S}_d^{++}(\mathbb{R})$.

As stated in La Bua and Marazzina (2019), we can obtain the Wishart SDE (1) from (5) by setting $D_X = \Omega\Omega^\top$, $C_X = Q$, and $\mathcal{L}[P_0] = MP_0 + P_0 M^\top$. Moreover, if we assume a more restrictive parametrization for the deterministic part of the drift

$$\Omega\Omega^\top = \beta Q^\top Q, \tag{6}$$

conditions (a) and (c) of Proposition 1 are satisfied as soon as

$$\beta \geq d - 1, \tag{7}$$
$$\beta \geq d + 1, \tag{8}$$

respectively, where the real positive parameter β plays the role of Feller's condition in the univariate case. Additionally if condition a) is not met the whole process is not well defined. For the rest of the paper we consider a Wishart process defined by (1) and (6) as usually done in financial literature. Notice that a significant constraint has thus to be imposed on parameter β. In the case $d = 2$, for example, we must require $\beta \geq 1$. As shown in La Bua and Marazzina (2019), this condition is not usually met when we perform a straight calibration of Wishart-based pricing models to market prices of plain vanilla options, while it does not seem to be a real limitation, according to our knowledge, when a maximum likelihood or a moments estimation is considered (Alfonsi et al. 2016; Boloorforoosh et al. 2020; Da Fonseca et al. 2014; Gourieroux and Sufana 2010). Given our interest in exploiting the model for pricing purposes, we rely to the calibration on plain vanilla options, and therefore to the implied volatility surface, analyzing the impact of Conditions (7, 8) in our numerical results. In La Bua and Marazzina (2020) we show how this impact can be smoothed adding a local volatility component, i.e., exploiting a stochastic-local volatility hybrid model.

2.1 Distribution of Wishart process and related results

In this section we deal with the analogy between Wishart and CIR processes. In fact, as shown in La Bua and Marazzina (2019), exploiting the affine nature of the Wishart process, we have that its characteristic function is an exponential affine transformation of the initial state as shown in the following Proposition:

Proposition 2 (Proposition 2.2 in La Bua and Marazzina (2019)) *Let Λ be a real symmetric $d \times d$ matrix, $t \geq 0$ and $T - t = \tau > 0$. The (conditional) characteristic function of the Wishart process defined by (1) and (6) is*

$$
\begin{aligned}
\phi_\Sigma(\Lambda, \tau) &= \mathbb{E}\left[\exp\left(\iota \operatorname{Tr}\left[\Lambda \Sigma(T)\right]\right) \mid \Sigma(t)\right] \\
&= \exp\left(\operatorname{Tr}\left[A_\Sigma(\Lambda, \tau)\Sigma(t)\right] + b_\Sigma(\Lambda, \tau)\right)
\end{aligned}
\tag{9}
$$

where ι is the imaginary unit (i.e. $\iota = \sqrt{-1}$) and matrix $A_\Sigma(\Lambda, \tau)$ and scalar function $b_\Sigma(\Lambda, \tau)$ are such that

$$
\begin{aligned}
\operatorname{Tr}\left[A_\Sigma(\Lambda, \tau)\Sigma(t)\right] &= \operatorname{Tr}\left[\iota\Lambda\left(\mathbb{I}_d - 2\iota\Theta(\tau)\Lambda\right)^{-1}\Gamma(\tau)\right], \\
b_\Sigma(\Lambda, \tau) &= -\frac{\beta}{2}\operatorname{Tr}\left[\log\left(\left(\mathbb{I}_d - 2\iota\Theta(\tau)\Lambda\right)\exp\left(\tau M^\top\right)\right) - \tau M\right],
\end{aligned}
$$

$\mathbb{I}_d$ being the d-dimensional identity matrix. The additional matrix functions appearing in the above equations are given by

$$
\Gamma(\tau) = \exp(\tau M)\Sigma(t)\exp(\tau M^\top),
$$

$$
\Theta(\tau) = \int_0^\tau \exp\left(uM\right) Q^\top Q \exp\left(uM^\top\right) du.
$$

As a consequence of the analytical tractability of Wishart process and of the knowledge of its characteristic function, we are able to present an additional result regarding the distribution of diagonal elements of the Wishart process (a similar result for the trace of the Wishart process has been obtained in La Bua and Marazzina 2019, Corollary 2.3).

Corollary 1 (Distribution of elements on the main diagonal of Wishart process) *Let $\Sigma_i(t) = \Sigma_{ii}(t)$ be the i-th element on the main diagonal of $\Sigma(t)$ and $F_{\chi^2}(x; \nu, \delta)$ the cumulative distribution function of a non-central chi-square random variable with ν degrees of freedom and non-centrality parameter δ. Then, for a fixed $T > t$, we have:*

$$
\Pr\left[\Sigma_i(T) \leq \upsilon | \Sigma(t)\right] = F_{\chi^2}\left(\frac{\upsilon}{\vartheta_i}, \beta, \delta_i\right)
\tag{10}
$$

where $\delta_i = \gamma_i / \vartheta_i$ with $\Gamma(\tau) = (\gamma_{ij})_{1 \leq i, j \leq d}$ and $\Theta(\tau) = (\vartheta_{ij})_{1 \leq i, j \leq d}$. For ease of notation, we also set $\gamma_i = \gamma_{ii}$ and $\vartheta_i = \vartheta_{ii}$.

Proof We use the fact that $\exp\left(\text{Tr}\left[\log(G)\right]\right) = \det[G]$ for any matrix invertible matrix[1] G to write

$$\exp\left(b_{\Sigma}(\Lambda, \tau)\right) = \det\left[(\mathbb{I}_d - 2i\Theta(\tau)\Lambda)\exp\left(\tau M^{\top}\right)\right]^{-\frac{\beta}{2}} \exp\left(\frac{\beta}{2}\text{Tr}\left[M\right]\tau\right)$$

$$= \det\left[\mathbb{I}_d - 2i\Theta(\tau)\Lambda\right]^{-\frac{\beta}{2}},$$

and (9) becomes

$$\phi_{\Sigma}(\Lambda, \tau) = \det\left[\mathbb{I}_d - 2\iota\Theta(\tau)\Lambda\right]^{-\frac{\beta}{2}} \exp\left(\text{Tr}\left[\iota\Lambda\left(\mathbb{I}_d - 2\iota\Theta(\tau)\Lambda\right)^{-1}\Gamma(\tau)\right]\right). \quad (11)$$

Let λ be a real variable and $e_i^d = (\mathbb{1}_{k=\ell=i})_{1\leq k,\ell\leq d}$, then by setting $\Lambda_i = \lambda e_i^d$, the characteristic function of $\Sigma_i(T)$ is

$$\phi_{\Sigma_i}(\lambda, \tau) = \phi_{\Sigma}\left(\Lambda_i, \tau\right) = (1 - 2\iota\lambda\vartheta_i)^{-\frac{\beta}{2}} \exp\left(\frac{\iota\lambda\gamma_i}{1 - 2\iota\lambda\vartheta_i}\right) \quad (12)$$

from which (10) follows by the definition of non-central chi-square distribution. $\square$

An alternative way to obtain (10) is to recognize that (11) is the characteristic function associated to the non-central Wishart distribution with scale matrix $\Theta(\tau)$ and non-centrality matrix $\Gamma(\tau)$ and apply results in Kourouklis and Moschopoulos (1985). For the sake of completeness, we point out that an analogous claim is shown in Da Fonseca and Grasselli (2011). Our formulation, however, gives a direct interpretation of parameters involved in the distribution of $\Sigma_i(T)$ in terms of matrices describing the Wishart process. As we will see, this turns out to be particularly useful for computational purposes.

An important consequence of (10) is that we can define an exact mapping between $\Sigma_i(T)$ and a CIR process:

Proposition 3 (CIR process mapping $\Sigma_i(T)$) *Let $v(t)$ be a CIR process defined by* (3). *For a fixed $T > t$, it holds that (conditionally on $v(t)$ and $\Sigma(t)$ respectively) $v(T)$ and $\Sigma_i(T)$ share the same distribution provided that*

$$v(t) = \Sigma_i(t), \quad (13)$$

$$\kappa = -\frac{1}{t}\log\left(\frac{\gamma_i}{v(t)}\right), \quad (14)$$

$$\eta = 2\sqrt{\frac{\vartheta_i\kappa}{\left(1 - e^{-\kappa t}\right)}}, \quad (15)$$

$$\theta = \frac{\beta\eta^2}{4\kappa}, \quad (16)$$

---

[1] Here we exploit the invertibility of $(\mathbb{I}_d - 2i\Theta(\tau)\Lambda)\exp\left(\tau M^{\top}\right)$ which is given by the invertibility of both $\mathbb{I}_d - 2i\Theta(\tau)\Lambda$ (see the proof of Proposition 2.2 in La Bua and Marazzina (2019)) and $\exp\left(\tau M^{\top}\right)$ (by definition of matrix exponential).

where γ_i and ϑ_i have been introduced in Corollary 1.

Proof The correspondence of the distributions relies on the properties of the CIR process. We refer to Cox et al. (1985) for details. □

3 The wishart affine stochastic correlation model

With the purpose of reproducing well-known multi-asset stylized facts in a tractable way, in Da Fonseca et al. (2007) the authors introduce the Wishart Affine Stochastic Correlation model (WASC) that makes use of Wishart process to describe the stochastic variance covariance matrix of asset returns. The model proposes the following joint dynamics for a vector of forward asset prices:

$$d\mathbf{f}(t) = \text{diag}\left[\mathbf{f}(t)\right]\sqrt{\Sigma(t)}\,d\mathbf{b}(t), \quad \mathbf{f}(0) \in \mathbb{R}^d_+, \tag{17}$$

where $\text{diag}\left[\cdot\right]$ is the operator that transforms a d-dimensional column vector into a $d \times d$ diagonal matrix. In (17) $\mathbf{b}(t)$ is a d-vector Brownian motion such that

$$\mathbf{b}(t) = \sqrt{1 - \mathbf{r}^\top \mathbf{r}}\,\mathbf{z}(t) + W(t)\,\mathbf{r} \tag{18}$$

with $\mathbf{z}(t)$ another d-vector Brownian motion independent on $W(t)$ and $\mathbf{r} \in [-1, 1]^d$ such that $\mathbf{r}^\top \mathbf{r} \leq 1$. Here $\mathbf{r}$ can be interpreted as the vector of coefficients meant to drive the linear correlation between the shocks on asset returns and shocks on variance-covariance matrix $\Sigma(t)$. The choice of the correlation structure (18) represents the major improvement with respect to the model in Gourieroux and Sufana (2004) and aims at accommodating realistic single asset volatility skews still preserving the affinity of the model. Remarkably, the resulting WASC dynamics (17) allows for stochastic correlation among asset returns in a tractable framework where each asset is enriched with a stochastic volatility behavior consistent with the effects observed on plain vanilla markets. Let the i-th forward asset price at time t be denoted by $f_i(t) = f_i(0)e^{y_i(t)}$, $t \geq 0$. The peculiarities of the model can be fully appreciated by referring to the individual, or scalar, dynamics of asset returns y_i:

$$dy_i(t) = -\frac{1}{2}\sum_{j=1}^d s_{ij}^2(t)dt + \sum_{j=1}^d s_{ij}(t)db_j(t)$$

$$= -\frac{1}{2}\Sigma_i(t) + \sum_{j=1}^d s_{ij}(t)db_j(t), \quad i = 1, ..., d \tag{19}$$

where $\widehat{S}(t) = \widehat{S}(t)^\top = \left(s_{ij}\right)_{1 \leq i, j \leq d}$ is the unique positive semi-definite square root of $\Sigma(t)$. We also use the notation $\Sigma_i(t) = \Sigma_{ii}(t)$ to denote the i-th diagonal element of Wishart process. By straightforward computations, from (19) we can compute the quadratic covariation of two given assets:

$$d\left[y_k(t), y_\ell(t)\right] = d\left[\sum_{j=1}^{d} s_{kj}(t) db_j(t), \sum_{j=1}^{d} s_{\ell j}(t) db_j(t)\right] = \Sigma_{k\ell}(t) dt \quad (20)$$

that highlights the role of Wishart process, used to describe the stochastic evolution of the asset returns variance covariance matrix. Furthermore, we can explicitly define the cross-asset correlation matrix $C_{\mathbf{y}}(t)$ as

$$C_{\mathbf{y}}(t) = \left(\rho_{ij}(t)\right)_{1 \le i, j \le d} = \left(\frac{\Sigma_{ij}(t)}{\sqrt{\Sigma_i(t)\Sigma_j(t)}}\right)_{1 \le i, j \le d}. \quad (21)$$

By exploiting the properties of Wishart process, it can be shown that $C_{\mathbf{y}}(t)$ is a well-defined correlation matrix (i.e. $C_{\mathbf{y}}(t)$ is positive semi-definite and each $\rho_{ij}(t) \in [-1, 1]$) as soon as condition (7) is satisfied and provided that $\Sigma_{ij}(t) \neq 0$ for $i, j = 1, ..., d$. Indeed, if $\beta \ge d - 1$, $\Sigma(t)$ is positive semi-definite and $C_{\mathbf{y}}(t)$ admits the decomposition

$$C_{\mathbf{y}}(t) = D^{-1}\Sigma(t)D^{-1} = D^{-1}\widehat{S}(t)\widehat{S}(t)D^{-1} = LL^\top,$$

where $D = \sqrt{\text{diag}\left[\Sigma(t)\right]}$. To show that each element $\rho_{ij}(t)$ is bounded in $[-1, 1]$ we use the following theorem that applies for Wishart distributed matrices:

Theorem 1 (Theorem 2.4.2 in Kollo and von Rosen (2006)) *Let $X_{\mathcal{W}} \in \mathcal{S}_d^+(\mathbb{R}) \sim \mathcal{W}_d(\beta, \Theta, \Gamma)$, i.e. $X_{\mathcal{W}}$ is a $d \times d$ symmetric matrix that follows a non-central Wishart distribution with degrees of freedom β, scale Θ and non-centrality matrix Γ. Then for a $n \times d$ matrix B, we have that $Y = BX_{\mathcal{W}}B^\top \sim \mathcal{W}_n(\beta, B\Theta B^\top, B\Gamma B^\top)$.*

If we set $X_{\mathcal{W}} = \Sigma(t)$ and $B = \left[\mathbf{e}_i^d, \mathbf{e}_j^d\right]^\top$ for some admissible i and j (with $\mathbf{e}_i^d$ the i-th element of the standard basis of $\mathbb{R}^d$), the previous theorem shows that the resulting matrix $B\Sigma(t)B^\top = \begin{bmatrix} \Sigma_i(t) & \Sigma_{ij}(t) \\ \Sigma_{ij}(t) & \Sigma_j(t) \end{bmatrix}$ is a well-defined 2×2 Wishart process and then it holds that $\Sigma_{ij}^2(t) \le \Sigma_i \Sigma_j$. This proves that each element $\rho_{ij}(t)$ defined in (21) is bounded in $[-1, 1]$.

By combining (19) and (20) with Proposition 3, and fixing a time horizon T, we can represent the (T-specific) WASC dynamics as the following $2d$ system of scalar SDEs (for $i = 1, ..., d$)

$$dy_i(t) = -\frac{1}{2}\Sigma_i(t)dt + \sqrt{\Sigma_i(t)}dw_i^y(t), \quad (22)$$

$$d\Sigma_i(t) = \kappa_i(\theta_i - \Sigma_i(t))dt + \eta_i\sqrt{\Sigma_i(t)}dw_i^\Sigma(t), \quad (23)$$

where the parameters κ_i, θ_i and η_i are given, respectively, in (14), (15) and (16). Here the correlation structure among Brownian motions $\mathbf{w} = \left[w_1^\Sigma, w_2^\Sigma, ..., w_d^\Sigma, w_1^y, w_2^y, ..., w_d^y\right]^\top$ is described by means of the stochastic block matrix

$$C(t) = \begin{bmatrix} C_{\Sigma}(t) & C_{\Sigma\mathbf{y}}^{\top}(t) \\ C_{\Sigma\mathbf{y}}(t) & C_{\mathbf{y}}(t) \end{bmatrix}, \tag{24}$$

where the submatrices, other than $C_{\mathbf{y}}$ already introduced in (21), will be described in the following. Interestingly, from (22) and (23) we have that, for any $T > 0$, the scalar dynamics of each asset is consistent with a standard Heston model driven by the i-th diagonal element of $\Sigma(t)$. This assures that the behaviour induced by WASC in terms of reconstructed implied volatility surfaces is in line with the documented findings of traditional one factor stochastic volatility models. Consistently with Heston model, the asset specific returns-volatility correlation is constant: following Da Fonseca et al. (2007), we have

$$\text{Corr}_t\left[dy_i(t), d\Sigma_i(t)\right] = \rho_i dt = \frac{\text{Tr}\left[Q R_i\right]}{\sqrt{Q_{ii}^*}} dt, \tag{25}$$

where R_i is the matrix with $\mathbf{r}$ on the i-th row and zero elsewhere. We can even generalize the previous result by explicitly deriving the correlation between the i-th log-asset and a generic diagonal element of $\Sigma(t)$: we, indeed, have that (as shown in Da Fonseca et al. (2007)) the covariation between asset returns and volatility terms is given by

$$d\left[y_i(t), \Sigma_j(t)\right] = 2\Sigma_{ij}(t)\,\text{Tr}\left[R_j Q\right] dt. \tag{26}$$

By combining (26) and (20) and exploiting (4), we get the generic element of matrix $C_{\Sigma\mathbf{y}}(t)$ as

$$\begin{aligned}
\text{Corr}_t\left[dy_i(t), d\Sigma_j(t)\right] &= \frac{2\Sigma_{ij}(t)\,\text{Tr}\left[R_j Q\right]}{\sqrt{\Sigma_i(t)}\sqrt{4\Sigma_j(t)Q_{jj}^*}} dt \\
&= \frac{\text{Tr}\left[R_j Q\right]}{\sqrt{Q_{jj}^*}} \frac{\Sigma_{ij}(t)}{\sqrt{\Sigma_i(t)}\sqrt{\Sigma_j(t)}} dt = \rho_j \rho_{ij}(t) dt. \quad (27)
\end{aligned}$$

In the last equality of (27), we introduce a new representation of such correlations that can be seen, quite fascinatingly, as the product of the (constant) *proper*, or scalar, j-th asset-volatility correlation and the cross-asset correlation between i-th and j-th assets. This result highlights the peculiar dependence structure inherent in the WASC model and could help in gaining more insights on parameters impact on correlation surfaces in the spirit of the study carried out in Da Fonseca et al. (2007).

Further, using (4), it follows that the elements of $C_{\Sigma}(t)$ have the form

$$\text{Corr}_t\left[d\Sigma_i(t), d\Sigma_j(t)\right] = \frac{4\Sigma_{ij}(t)Q_{ij}^*}{4\sqrt{\Sigma_i(t)\Sigma_j(t)}\sqrt{Q_{ii}^* Q_{jj}^*}} dt = \frac{Q_{ij}^*}{\sqrt{Q_{ii}^* Q_{jj}^*}} \rho_{ij}(t) dt. \tag{28}$$

From (21), (27) and (28), we have that the stochastic evolution of (24) is fully described by processes $\rho_{ij}(t)$. Let us consider, for example, the case $d = 2$: matrix $C(t)$ then reads as

$$C(t) = \begin{bmatrix} 1 & q_{12}\rho_{12}(t) & \rho_1 & \rho_1\rho_{12}(t) \\ q_{12}\rho_{12}(t) & 1 & \rho_2\rho_{12}(t) & \rho_2 \\ \rho_1 & \rho_2\rho_{12}(t) & 1 & \rho_{12}(t) \\ \rho_1\rho_{12}(t) & \rho_2 & \rho_{12}(t) & 1 \end{bmatrix}$$

where $q_{ij} = \dfrac{Q_{ij}^*}{\sqrt{Q_{ii}^* Q_{jj}^*}}$. This representation highlights the peculiar dependence structure induced in the WASC model and could provide useful insights on the role of Q and $\mathbf{r}$ in determining the relation among state variables.

3.1 A restricted version of the model

In this section we consider a restricted, more intuitive, specification of WASC model: we assume matrix M to be diagonal and with negative entries. This setting leads to a very interesting dynamics for the diagonal elements of Wishart process. From Proposition 3, if M is diagonal, direct computation gives

$$\kappa_i = -2M_{ii}, \quad \eta_i = 2\sqrt{Q_{ii}^*},$$

for all $T > 0$. An immediate consequence is that the asset instantaneous variances are now described by time-independent CIR processes, in the sense that the parameters involved are no longer function of the time horizon considered.

In our opinion the resulting parametrization turns out to be the most genuine multi-asset extension of the Heston model: each asset is exactly described by a single instance of the Heston dynamics while the joint behaviour is enriched by cross-assets and cross-variances stochastic correlation, all wrapped in an affine framework. As far as we know, there are no alternative settings that can reach a comparable degree of flexibility. The exact Heston representation of asset dynamics also helps in understanding the role and the impact of WASC parameters, that in the general formulation appear somehow unclear. In particular, it is worthwhile to point out that the pricing of single-asset European claims is only affected by the corresponding diagonal element of Σ_0. To see this, it suffices to notice that for $\boldsymbol{\lambda} = \lambda_i \mathbf{e}_i^d$, the matrix $A_{\mathbf{y}}(\tau)$ in (31) has a non-null i-th diagonal element and zeros elsewhere. This peculiarity is in line with the asymptotic analysis provided in Da Fonseca and Grasselli (2011) where the i-th implied volatility approximation for short time to maturity is found to be[2]

$$\sigma_{imp,i}^2 = \Sigma_{ii} + (r_1 Q_{1i} + r_2 Q_{2i})m_f + \frac{1}{2}\frac{4\left(Q_{1i}^2 + Q_{2i}^2\right) - 7(r_1 Q_{1i} + r_2 Q_{2i})^2}{6\Sigma_{ii}}m_f^2,$$

with $m_f = \log\left(\frac{K}{f_i(0)}\right)$ denoting the log-forward moneyness. Consequently, the off-diagonal entries of Σ_0 can be used to match multi-asset stylized facts without compromising the shape of individual volatility surfaces. This represents an additional degree of freedom that in the general WASC model we would not have. A

---

[2] Without loss of generality only the case $d = 2$ is considered.

possible calibration strategy could be to set the off-diagonal entries of Σ_0 in order to match a predefined initial cross-asset correlation matrix. Alternatively, provided that liquid multi-asset derivatives are traded, we could try to fit the implied correlation market evidences. In order to further develop this point, we now study the impact of Σ_{12} on the price of Best-Of put options which payoff is

$$(K - \max\,[S_1(T),\,S_2(T)])^+\,, \tag{29}$$

$S_i(t)$ being the asset price at time t, $t \geq 0$. Let $\boldsymbol{\pi}_{WA}$ be the following set of WASC parameters:

$$\beta = 1.1, \ \ \Sigma_0 = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix}, \ \ M = \begin{bmatrix} -0.7 & 0 \\ 0 & -1.2 \end{bmatrix}, \ \ Q = \begin{bmatrix} 0.3 & 0.3 \\ 0.2 & 0.3 \end{bmatrix},$$

$$\mathbf{r} = \begin{bmatrix} -0.6 \\ -0.1 \end{bmatrix},$$

that are meant to describe realistic market scenarios and allowing for a well defined Wishart process. Figure 1 shows the implied correlation profiles corresponding to different values of Σ_{12}, where the implied correlation is defined to be the value of parameter ρ such that the WASC price equals the one obtained in a two-assets Black-Scholes setting, i.e.

$$P_{Best-Of}^{BS}(\sigma_{imp,1}^{WA}(K, T), \sigma_{imp,2}^{WA}(K, T), \rho, K, T) = P_{Best-Of}^{WA}(\boldsymbol{\pi}_{WA}, K, T). \tag{30}$$

Here $\sigma_{imp,i}^{WA}(K, T)$ is the Black-Scholes implied volatility corresponding to the option written on the i-th asset with strike K and maturity T whose price is computed with WASC model. By exploiting the affinity of the model, Best-Of put options are priced by numerically computing a bi-dimensional inverse Fourier transform. We refer the reader to Da Fonseca et al. (2007) where this pricing methodology is developed for Best-Of contracts. From the numerical results, it is evident that the off-diagonal Wishart element plays a significant role in modelling the implied correlation skew: we observe an increase in implied correlation levels for higher values of $\Sigma_{1,2}$. This, in turn, induces an increase in option prices consistently with the fact that Best-Of put options are long correlation products that benefit from lower assets returns dispersion.

3.2 WASC Characteristic function

The chosen correlation structure (18) assures the affinity of WASC model. This means, once more, that we can express the (joint) characteristic function of the asset returns vector $\mathbf{y}(T)$ as an exponential affine transformation of state variables $\mathbf{y}(t)$ and $\Sigma(t)$ as recalled in the following Proposition:

Proposition 4 *(Joint characteristic function of log-prices in WASC) Let the log-forward prices vector $\mathbf{y}(t)$ be described by (19) and $\boldsymbol{\lambda}$ be an auxiliary vector-valued variable $\boldsymbol{\lambda} = [\lambda_1, ..., \lambda_d]^\top$. Then for $T > t$, the WASC (conditional) characteristic function of $\mathbf{y}(T)$ admits the following closed formula representation*
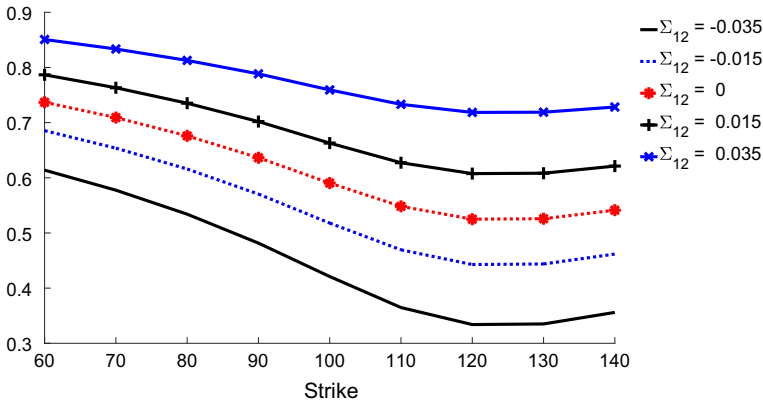
Fig. 1 Best-Of put options implied correlation skew for different values of Σ_{12}. Other parameters used: $f_1(0) = f_2(0) = 100$, $T = 1$ and $r = 0\%$

$$\phi_{\mathbf{y}}^{WA}(\boldsymbol{\lambda}, \tau) = \mathbb{E}\left[\exp\left(\iota\langle\boldsymbol{\lambda}, \mathbf{y}(T)\rangle\right) | \mathbf{y}(t)\right]$$
$$= \exp\left(\iota\langle\boldsymbol{\lambda}, \mathbf{y}(t)\rangle + \text{Tr}\left[A_{\mathbf{y}}(\tau)\Sigma(t)\right] + b_{\mathbf{y}}(\tau)\right), \quad (31)$$

with the deterministic matrix $A_{\mathbf{y}}(\tau)$ and the scalar function $b_{\mathbf{y}}(\tau)$ given by

$$A_{\mathbf{y}}(\tau) = A_{22}(\tau)^{-1} A_{21}(\tau),$$
$$b_{\mathbf{y}}(\tau) = -\frac{\beta}{2}\text{Tr}\left[\log(A_{22}(\tau)) + \tau(M + \iota\lambda Q^{\top}R^{\top})\right],$$

and

$$\begin{bmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{bmatrix} = \exp\left(\tau\begin{bmatrix} M + \iota Q^{\top}\mathbf{r}\boldsymbol{\lambda}^{\top} & -2Q^{\top}Q \\ -\frac{1}{2}\left(\boldsymbol{\lambda}\boldsymbol{\lambda}^{\top} + \iota\,\text{diag}[\boldsymbol{\lambda}]\right) & -(M + \iota Q^{\top}\mathbf{r}\boldsymbol{\lambda}^{\top})^{\top} \end{bmatrix}\right).$$

Proof See Da Fonseca et al. (2007). □

Thanks to Proposition 4, we are able to price both plain vanilla and multi-asset options (if transform-based techniques are applicable[3]) in a comprehensive framework. In particular, we price options on the *i*-th asset as a basket option with degenerate weights vector $\mathbf{e}_i^d$, such that $\boldsymbol{\lambda} = \lambda_i \mathbf{e}_i^d$. Furthermore, the knowledge of the joint characteristic function of asset returns vector allows to make use of bounds techniques as those developed in Caldana et al. (2016) for basket options. Despite the analytical tractability, several numerical issues arise when we try to calibrate WASC model to market data by exploiting (31). Not only, indeed, we have to evaluate functions of matrix argument for each computation of the characteristic function (as in the WMSV case), but, even worse, we are required to perform *d* different plain vanilla pricing (one for each asset)

---

[3] We refer here to those cases in which transformed payoff functions are available and evaluation of (inverse) transforms are numerically feasible.

for a single parameters set. This is due to the lack of liquid multi-asset derivatives that force us to calibrate model parameters to the individual market implied volatility surfaces. As reported in Da Fonseca and Grasselli (2011), such a naive algorithm can take up to 15 minutes in the simplest case $d = 2$. This is, clearly, not feasible for real market applications. Therefore in the next section we propose an accurate and fast calibration procedure.

4 A new calibration procedure

In this section we present an innovative and efficient methodology to calibrate WASC parameters that exploits the close link existing between Heston model and marginal WASC dynamics. The proposed algorithm is firstly tested in a simplified framework and then applied to market data. The results obtained also highlight the impact of parameter β on model accuracy in reproducing market volatility smiles.

Let us consider a WASC parameters set $\boldsymbol{\pi}_{WA}$ (with cardinality N_{WA}) and fix a maturity T. For the generic i-th asset described by (19) we can define a function g_i^{H-WA} that maps WASC parameters to those of a scalar Heston dynamics. In other words, we set $g_i^{H-WA} : \mathbb{R}^{N_{WASC}} \times \mathbb{R}_{>0} \to \mathbb{R}^5$ such that $g_i^{H-WA}(\boldsymbol{\pi}_{WA}, T) = \boldsymbol{\pi}_{i,H} = [v_{0,i}, \kappa_i, \theta_i, \eta_i, \rho_i]^\top$ as defined in (13)–(16) along with the assets-volatility correlation (25). Consequently, for calibration purposes, we can replace the cumbersome WASC characteristic function with the simpler Heston one (see La Bua and Marazzina 2019, Appendix A.1). Furthermore, we can compute analytically the gradient of the objective function with respect to WASC parameters. In Appendix A we show how to compute explicitly the matrix $J_{i,T}^{H-WA}(\boldsymbol{\pi}_{WA}) = \nabla g_i^{H-WA} \in \mathbb{R}^{N_{WA} \times 5}$, i.e. the Jacobian matrix of function h_i^{WA} with elements:

$$j_{q,r}^{H-WA} = \frac{\partial g_{i,r}^{H-WA}(\boldsymbol{\pi}_{WA}, T)}{\partial \boldsymbol{\pi}_{WA,q}}. \tag{32}$$

Then, the Jacobian matrix of $\tilde{\mathbf{r}}_{i,T}(\boldsymbol{\pi}_{WA})$ (the residuals vector composed of options with maturity T written on the i-th asset) can be written as

$$J_{i,T}^{WA}(\boldsymbol{\pi}_{WA}) = J_{i,T}^{H-WA}(\boldsymbol{\pi}_{WA}) J_i^H(g_i^{H-WA}(\boldsymbol{\pi}_{WA}, T)) \tag{33}$$

where the second matrix in the right-hand side of (33) is known thanks to Cui et al. (2017) (see also La Bua and Marazzina 2019, Appendix A.2). To obtain the overall Jacobian matrix J^{WA}, we simply need to compute (33) for each maturity and asset taken into account and aggregate the resulting matrices. Notice the relationship between (La Bua and Marazzina 2019, Equation (43)) and Equation (32). Finally, similarly to La Bua and Marazzina (2019), the gradient of the calibration problem objective function is given by $\nabla f_{obj} = J^{WA}\tilde{\mathbf{r}}(\boldsymbol{\pi}_{WA})$.

The calibration algorithm so defined avoids the computation of WASC characteristic function and significantly reduces the issues due to the possible presence of multiple

minima. Given that we rely on the law identity in Proposition 3 rather than on some approximation, the routine does not require any further step.

4.1 A simplified calibration exercise

The accuracy of the proposed algorithm is illustrated by considering the following numerical experiment: let us suppose that a fictitious two-assets market is perfectly described by the WASC parameters reported in the first column of Table 1. Even if simplified, the data outline realistic market environments: they represent the calibrated parameters set (truncated at the first significant decimal digit) found in Da Fonseca and Grasselli (2011) for the couple of indices EuroStoxx50-DAX. We construct a full implied volatility surface for each asset assuming to have options with maturities $T = [0.25, 0.5, 1, 3]$ and 41 equally spaced strikes ranging from 0.5 to 1.5 (initial asset values are set for simplicity equal to 1). Each of the resulting surfaces consists of 164 options. The goal is to implement the proposed algorithm in order to find a suitable parameters set that reproduces the supposed market data. Hopefully, we expect the calibrated parameters to be reasonably close to the original ones (accuracy) and to experience a limited dependency on the initial guess (robustness). For this test we set the starting values of the optimization routine as shown in the second column of Table 1. The choice is meant to assess the robustness of the algorithm in the case in which the initial guess is very far from the optimal set. Indeed, not only the discrepancy is mixed - some values are overestimated, others underestimated - but the distance between initial guess and optimal values is substantial: the smallest gap, defined as percentage difference, is equal to 35.29%. The mistaken initialization of the problem and the high dimensionality of the parameters space make the calibration task more challenging and could potentially lead to suboptimal outcomes. Notwithstanding, the proposed algorithm is able to produce results very close to the original values: the norm of the errors between true prices and calibrated ones is 2.2069×10^{-7}. Most remarkably, the procedure takes only 3.56 seconds using Matlab Mex files, on a laptop PC with an Intel Core i7 CPU and 8 GB RAM. By considering parallelization and porting to more efficient languages we can obtain a further speedup. It is worthwhile also noting that in realistic applications, the calibration problem is somehow facilitated thanks to the availability of previous optimal sets that act as efficient guesses. In the lights of all these evidences, we believe that the proposed methodology represents a highly efficient tool for the calibration of WASC model. This is particularly true if we intend to increase the number of assets involved with the subsequent growth of dimensionality.

4.2 Calibration to market data

We now want to validate the procedure with realistic market data. Despite the general applicability of the algorithm, we focus our attention on the restricted specification of the model introduced in Sect. 3.1. With this in mind, we select a basket of market quoted instruments composed of 201 European call options written on EuroStoxx50 index and 182 on DAX index. The set of derivatives on the DAX is the same set

Table 1 Results for the calibration exercise described in Sect. 4.1

Parameter	True values	Initial guess	Calibrated values
β	0.7	1	0.6951
$\Sigma_{11}(0)$	0.04	0.1	0.0402
$\Sigma_{12}(0)$	0.03	0	0.0349
$\Sigma_{22}(0)$	0.04	0.1	0.0399
M_{11}	-0.7	-2	-0.6791
M_{12}	-0.3	-1	-0.2476
M_{21}	-0.05	-1	-0.0836
M_{22}	-1.2	-2	-1.1417
Q_{11}	0.3	0.1	0.2756
Q_{12}	0.3	0	0.3201
Q_{21}	0.2	0	0.2187
Q_{22}	0.3	0.1	0.2976
r_1	-0.6	0	-0.6490
r_2	-0.1	0	-0.0912

used in La Bua and Marazzina (2019). We further set, for simplicity, interest rates and dividends to zero. Thanks to the efficiency of the new calibration algorithm, we are able to calibrate model parameters in less than 3 seconds. The outputs of the optimization routine are shown in the leftmost column of Table 2. Given that, as illustrated above, the off-diagonal element of Σ_0 does not impact the pricing of univariate call options, we set its value such that the initial correlation among the two indices equal the one-year historical one (that is found to be 0.9715).[4]

The most interesting result is that β is lower than 1. This is coherent with the evidences in Da Fonseca and Grasselli (2011) where similar results are found. Figure 2 shows the calibrated model implied volatility skews for the two indices with respect to maturities of one month, one year and three years. The model succeeds in reproducing the shape of market volatility surfaces but the mispricing is not negligible for short term far-from-the-money options. This is particularly true for the EuroStoxx50 index as highlighted from the fact that the error in volatility terms is roughly 3 times higher than the error made for the DAX.

Additionally, we can compare the evidences from the calibration of the WASC model against the Wishart Multidimensional Stochastic Volatility model (WMSV, La Bua and Marazzina 2019), as well as the Heston (1993) and Christoffersen et al. (2009) models, calibrated to the same basket of DAX options. Table 3 shows the calibrated initial variance of asset returns along with the Mean Squared Error with respect to both price and implied volatility for the four models. Consistently, the estimates of initial variance are in strict agreement: all the models agree on the initial volatility. Moreover, w.r.t. the accuracy of the calibration, the two multi-factor models, i.e., the WMSV and the Bi-Heston, tend to perform quite similarly (although errors for WMSV are slightly smaller) and substantially outperform the simpler Heston and

---

[4] We simply consider the historical value as a target correlation level defined by model user.

Table 2 Calibration on February, 3 2016 with the WASC over a full set of EuroStoxx50 and DAX indices European call options

Parameter	$\beta \geq 0$	$\beta \geq 1$	$\beta \geq 3$
β	0.8577	1.0389	3.0110
$\Sigma_{11}(0)$	0.0697	0.0554	0.0556
$\Sigma_{12}(0)$	0.0765	0.0654	0.0643
$\Sigma_{22}(0)$	0.0890	0.0817	0.0788
M_{11}	-1.9763	-0.4626	-0.6740
M_{22}	-1.1605	-0.7311	-0.8590
Q_{11}	0.4422	0.1977	0.1259
Q_{12}	0.1448	0.0893	0.0920
Q_{21}	0.1075	0.0004	0.0440
Q_{22}	0.3843	0.2787	0.1529
r_1	-0.5145	-0.4888	-0.4706
r_2	-0.5247	-0.5262	-0.8490
Error Price	2.67E-06	3.58E-06	7.66E-06
Error Price	2.97E-06	3.16E-06	6.11E-06
Error Vol	2.62E-04	3.25E-04	5.20E-04
Error Vol	9.88E-05	1.53E-04	8.33E-04
Time (s)	2.81	0.88	0.92

As in Da Fonseca and Grasselli (2011) Error Price stands for the Mean Squared Error (MSE) in price normalized by the forward price and Error Vol denotes the MSE in implied volatility. For each column, the first Error Price value refers to the EuroStoxx50 index while the second one to the DAX index. The same applies for the Error Vol values. In all cases we set $\Sigma_{12}(0)$ so that the initial value of cross-asset correlation is equal to 0.9715 (the one year historical correlation computed with daily market data)

Table 3 Comparison of calibration outputs on DAX index

	WMSV	Bi-Heston	Heston	WASC
Initial Variance	$\mathrm{Tr}\,[\Sigma_0]$	$v_{0,1} + v_{0,2}$	Σ_{12}	v_0
Calibrated Value	0.0866	0.0868	0.0842	0.0890
Error Price	2.45E-06	2.83E-06	2.95E-06	2.97E-06
Error Vol	2.85E-05	3.56E-05	1.07E-04	9.88E-05

The Initial Variance row specifies the initial variance of asset returns in the corresponding model

WASC dynamics. In particular, by comparing the error of WMSV and WASC models, the outperformance of the former is clearly evident. This is not surprisingly since we contrast a multi-factor volatility setting (WMSV) with the WASC single-asset dynamics that, as developed in Sect. 3, is equivalent to 1-factor parametrization. It is important to remark, however, that the models are meant to address rather different tasks (i.e. single-asset and multi-asset modelling).

Moving back to the multi-asset calibration in Table 2, some fix is required in order to enforce the existence and uniqueness condition for matrix process $\Sigma(t)$. First of
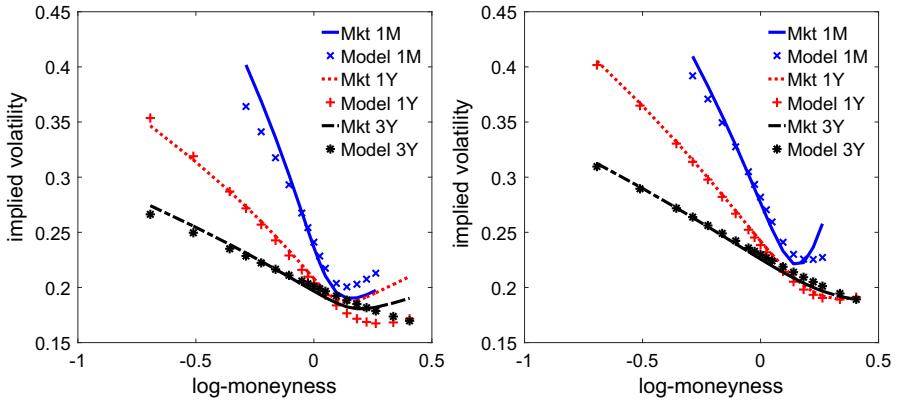
Fig. 2 Calibration results for WASC. Resulting value for parameter β is 0.8577. Comparison with market implied volatility for EuroStoxx50 (left) and DAX (right) indices for selected tenors
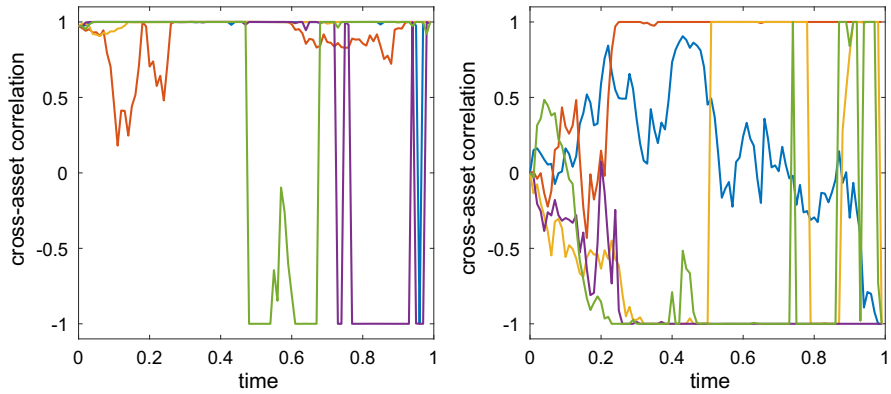

Fig. 3 Simulated trajectories of cross-asset correlation $\rho_{12}(t)$ for $t \in [0, 1]$ generated with calibrated parameters obtained imposing $\beta \geq 1$. Left panel: $\Sigma_{12} = 0.0654$. Right panel: $\Sigma_{12} = 0$
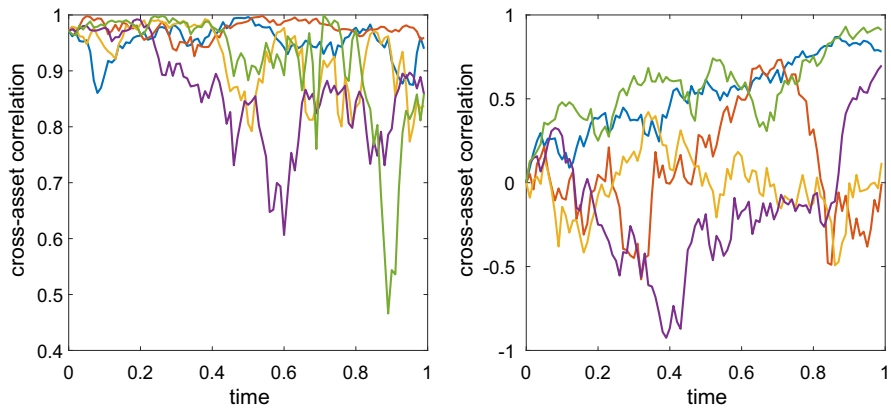

Fig. 4 Simulated trajectories of cross-asset correlation $\rho_{12}(t)$ for $t \in [0, 1]$ generated with calibrated parameters obtained imposing $\beta \geq 3$. Left panel: $\Sigma_{12} = 0.0643$. Right panel: $\Sigma_{12} = 0$

all, we tackle the calibration problem imposing $\beta \geq 1$. Results are exhibited in the second column of Table 2. Even if the loss in accuracy seems to be somehow limited (the error measure are just slightly higher than in the unconstrained setting), a relevant issue arises: in Fig. 3 we report simulated trajectories of WASC cross-asset correlation obtained with the resulting parameters set. The fact that β satisfies condition (7) effectively ensures $\rho_{12}(t)$ to lie in the range $[-1, 1]$. However, the fact that the parameter is just slightly above the threshold ($\beta = 1$) makes the boundary of $\mathcal{S}_2^+(\mathbb{R})$ very likely to be attained. Very often, then the absolute value of correlation is stuck at 1. We can also experience sudden changes in correlation from $+1$ to -1 (or viceversa) in a very restricted time frame (even on a daily basis). In order to study the dependence of the observed phenomenon on the initial value of correlation, in the rightmost panel of Fig. 3 we also consider the case $\Sigma_{12} = 0$ that produces a similar erratic correlation dynamics.

Given the intent to apply WASC model to describe the joint behaviour of asset prices, this represents a major issue. To tackle the problem, we decide to enforce the positive definiteness condition for $\Sigma(t)$, given by (8), that in our setting equals to set $\beta \geq 3$. Calibrated parameters are collected in the rightmost column of Table 2. The corresponding cross-asset correlation dynamics is depicted in Fig. 4: the trajectories are now much more meaningful. Further, as a consequence of the fact that $\Sigma(t)$ is defined on the interior of $\mathcal{S}_2^+(\mathbb{R})$, $\rho_{1,2}(t)$ is bounded in $(-1, +1)$. Nonetheless, the stronger condition enforced has a severe impact on the ability of the model to reproduce single-asset market evidences. Reconstructed volatility skews are shown in Fig. 5. Significant discrepancies now emerge for far-from-the-money options. In particular, in the very short-end of the volatility term structure the error with respect to market volatilities can be as high as 11.87% (in-the-money options on EuroStoxx50) and 10.80% (out-of-the-money options on DAX). Disappointingly, we face a non trivial trade-off between plain vanilla pricing accuracy and realistic modellization of cross-asset correlation. A possible solution to mitigate the problem could be to set β equal to some value in the range $(1, 3]$. This alternative, however, would require to couple the plain vanilla analysis with adequate market evidences on multi-asset derivatives.

5 Simulation schemes for the WASC

This section is devoted to present simulation algorithms specifically devised for WASC model. As far as we know, indeed, there are no previous attempts in literature to deal with the discretization of prices trajectories (17). In particular, our task is to develop an efficient, yet accurate, scheme to discretize the system of SDEs (22)–(23). It is evident that a standard discretization (e.g. via Euler scheme) is unfeasible, since we also need to take into account the evolution of non diagonal elements of $\Sigma(t)$ to determine the dependence structure and satisfy the positive semi-definiteness constraint for the Wishart process.
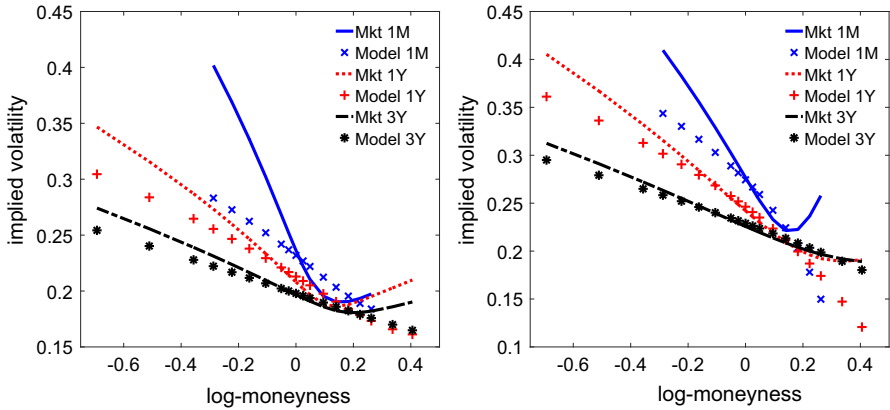
Fig. 5 Calibration results for WASC. Resulting value for parameter β is 3.011. Comparison with market implied volatility for EuroStoxx50 (left) and DAX (right) indices for selected tenors

As a first algorithm, we implement an adapted version of the scalar full truncated Euler (TE) scheme that reads as

$$\widehat{\mathbf{y}}(t + \Delta) = -\frac{1}{2} \operatorname{Vec}\left[\widehat{\Sigma}^+(t)\right] \Delta + \sqrt{\widehat{\Sigma}^+(t)}\left(\sqrt{1 - \mathbf{r}^\top \mathbf{r}}\ \widetilde{\mathbf{z}} + \widetilde{W}\ \mathbf{r}\right)\sqrt{\Delta}$$

$$\widehat{\Sigma}(t + \Delta) = \widehat{\Sigma}(t) + \left(\beta Q^\top Q + M\widehat{\Sigma}^+(t) + \widehat{\Sigma}^+(t)M^\top\right)\Delta + \sqrt{\widehat{\Sigma}^+(t)}\widetilde{W}Q\sqrt{\Delta}$$

$$+ Q^\top \widetilde{W}^\top \sqrt{\widehat{\Sigma}^+(t)}\sqrt{\Delta}$$

with $\widetilde{\mathbf{z}}$ and $\widetilde{W}$, respectively, d-dimensional vector and square matrix of independent standard gaussian random variables. Here $\operatorname{Vec}[\cdot]$ is the operator that extracts the elements on the main diagonal of a square matrix into a column vector, while $\widehat{\Sigma}^+$ is the positive part of matrix $\widehat{\Sigma}$.

As a second algorithm, we deal with the Wishart process sampling scheme developed in Ahdida and Alfonsi (2013): that is, we consider as given an entire discretized path of $\Sigma(t)$ over the time grid $0 = t_0 < t_1 < ... < t_{M_T} = T$ with time step Δ. In this way we are only left with the problem of sampling the log-prices trajectories. The most challenging task here is to embed the correlation structure (24) in the discretization of (22). In standard cases, we would compute the Cholesky decomposition of matrix $C(t)$ such that

$$\left[w_1^\Sigma, w_2^\Sigma, ..., w_d^\Sigma, w_1^y, w_2^y, ..., w_d^y\right]^\top$$
$$= L_C(t)\left[w_1^*, w_2^*, ..., w_d^*, w_{d+1}^*, w_{d+2}^*, ..., w_{2d}^*\right]^\top, \tag{34}$$

where $L_C(t) = \left(\ell_{i,j}(t)\right)_{1 \leq i \leq d, 1 \leq j \leq i}$ is the lower triangular matrix that satisfies $C(t) = L_C(t)L_C^\top(t)$ and $\mathbf{w}^* = \left[w_1^*, w_2^*, ..., w_d^*, w_{d+1}^*, w_{d+2}^*, ..., w_{2d}^*\right]^\top$ is a vector of independent Brownian motions. By exploiting (34), we can rewrite (22) as

$$dy_i(t) = -\frac{1}{2}\Sigma_i(t)dt + \sqrt{\Sigma_i(t)} \sum_{j=1}^{d+i} \ell_{d+i,j}(t)dw_j^*(t), \tag{35}$$

that can be discretized by generating $2d$ independent gaussian random variables. Unfortunately, in our setting this is not readily doable as a consequence of the mechanics of Wishart sampling algorithm. In other words, we do not have a direct "access" to the discretized paths of $w_1^\Sigma, w_2^\Sigma, ..., w_d^\Sigma$.

The simple idea underlying the new simulation scheme is to exploit the auxiliary scalar dynamics of $\Sigma_i(t)$ to get an approximation of $w_i^\Sigma(t + \Delta) - w_i^\Sigma(t)$. Let $\hat{\Sigma}(t)$ and $\hat{\Sigma}(t + \Delta)$ be the realizations of the trajectory of Wishart process for two adjacent points on the time grid computed by means of the exact scheme in Ahdida and Alfonsi (2013). The discretized version of (1), that reads

$$\hat{\Sigma}(t + \Delta) - \hat{\Sigma}(t) \approx \kappa_i \left(\theta_i - \hat{\Sigma}(t)\right)\Delta + \eta_i \sqrt{\hat{\Sigma}(t)}\widetilde{w}_i, \tag{36}$$

can now be used to approximate the gaussian variable $\widetilde{w}_i$. Let $\widetilde{w}_{\Sigma_i}$ be the result of (36), for a sufficiently small time interval, $\widetilde{\mathbf{w}}_\Sigma = \left[\widetilde{w}_{\Sigma_1}, \widetilde{w}_{\Sigma_2}, ..., \widetilde{w}_{\Sigma_d}\right]^\top$ represents an approximation of a vector of gaussian variables with correlation matrix $\hat{C}_\Sigma(t)$ (i.e. the realization at time t of matrix C_Σ). Further, let $\hat{L}_\Sigma(t)$ be the lower triangular matrix obtained from the Cholesky decomposition of $\hat{C}_\Sigma(t)$, then

$$\widetilde{\mathbf{w}}_\Sigma^* = \hat{L}_\Sigma^{-1}(t)\widetilde{\mathbf{w}}_\Sigma$$

is composed of d approximated independent gaussian random variables. By sampling an additional random vector $\hat{\mathbf{w}}_\mathbf{y}^*$ from $\mathcal{N}\left(\mathbf{0}_d, \Delta\mathbb{I}_d\right)$ (the d-variate gaussian distribution) and setting

$$\hat{\mathbf{w}}^* = \begin{bmatrix} \widetilde{\mathbf{w}}_\Sigma^* \\ \hat{\mathbf{w}}_\mathbf{y}^* \end{bmatrix}$$

we can finally approximate (35) as

$$\hat{y}_i(t + \Delta) = \hat{y}_i(t) - \frac{1}{2}\hat{\Sigma}_i(t)\Delta + \sqrt{\hat{\Sigma}_i(t)} \sum_{j=1}^{d+i} \hat{\ell}_{d+i,j}(t)\hat{w}_j^*(t) \tag{37}$$

where $\hat{L}_C(t)$ results from the factorization of $\hat{C}(t)$. If $\hat{C}(t)$ turns out not to be positive definite, we take its positive part, $\hat{C}^+(t)$, (defined as the matrix obtained from the spectral decomposition of $\hat{C}(t)$ with negative eigenvalues replaced by zeros) and apply the extended Cholesky decomposition described in Golub and Van Loan (2012). The complete algorithm is exhibited in Algorithm 1, and we refer to it as Gaussian Variables Approximation (GVA) scheme for WASC model. Notice that this algorithm can be considered as an extension of the GVA algorithm for the WMSV case developed in La Bua and Marazzina (2019).

Algorithm 1 The *Gaussian Variables Approximation* (GVA) scheme for the WASC

1: **for** each asset i, $i = 1, ..., d$ **do**
2:     Given a time horizon T, compute parameters κ_i, θ_i, η_i and ρ_i as given in Proposition 3
3: **end for**
4: **for** each simulation trial n, $n = 1, ..., N$ **do**
5:     Initialize $\widehat{\Sigma}(0) = \Sigma_0$
6:     Initialize $\widehat{\mathbf{y}}(0) = \log(\mathbf{f}(0))$
7:     **for** each time-step t_m, $m = 0, ..., M_T - 1$ **do**
8:         Compute correlation matrix $\widehat{C}(t_m)$ as given by (24)
9:         Compute $\hat{L}_C(t)$ from Cholesky decomposition of $\hat{C}(t_m)$ or, in alternative, from the extended
            Cholesky decomposition of $\hat{C}^+(t)$
10:          Sample the Wishart process for a time step of Δ with initial state $\widehat{\Sigma}(t_m)$
            using the scheme in Ahdida and Alfonsi (2013)
11:          Approximate the d gaussian random variables $\widetilde{\mathbf{w}}$ by means of (36)
12:          Compute $\widetilde{\mathbf{w}}_{\Sigma}^* = \hat{L}_{\Sigma}^{-1}(t)\widetilde{\mathbf{w}}_{\Sigma}$
13:          Sample $\hat{\mathbf{w}}_{\mathbf{y}}^*$ from $\mathcal{N}(\mathbf{0}_d, \Delta \mathbb{I}_d)$
14:          Discretize (35) by means of (37)
15:     **end for**
16: **end for**

5.1 Numerical results

Even if these new schemes would apply to the general specification of the model, here we suppose to deal with the reduced model presented in Sect. 3.1 (i.e. we consider matrix M to be diagonal). In particular, we develop an extensive numerical investigation based on the parameters set calibrated to market data enforcing the condition $\beta \geq 3$ and shown in the rightmost column of Table 2. Considering 5×10^5 simulation paths, we price European call options written on any of the 2 assets with maturity $T = 1$ and moneyness in the range $\{70\%, 100\%, 130\%\}$. For the sake of simplicity, we assume interest rates and dividends equal to zero and $f_1(0) = f_2(0) = 100$. The asterisk in the following tables means that the corresponding reference value lies outside of the 95% confidence interval.

In the context of plain vanilla options, results in Tables 4, 5, 6, 7 show that both schemes allow for accurate price estimates as the size of the time step is sufficiently small. More in detail, the TE scheme is found to outperform the GVA scheme when the time grid is coarse (10, 20 and 50 steps per year), while for smaller mesh widths the two approaches tend to perform similarly. With this setting and taking into account option prices for both assets, the absolute mean percentage error is respectively equal to 0.232% for the GVA scheme and to 0.275% for the TE scheme. It is worthwhile to remark, though, that the GVA scheme systematically requires a finer time discretization to produce reliable estimates (true prices lying in the 95% confidence interval) compared to the simpler TE scheme. This is due to the fact that the approximation exploited in (36) seems to be adequate only for very small time intervals. From a computational point of view, the TE scheme greatly outperforms the GVA scheme with the latter that results $110\% - 130\%$ slower than the former as exhibited in Table 8. When implementing the GVA scheme, indeed, at each time step we are asked to perform the Cholesky factorization of the $2d$-dimensional matrix $\widehat{C}(t)$ with a considerable increase in the computational burden.

Table 4 Plain vanilla option written on Asset 1, TE scheme

Strike	Reference Value	No. of time steps	MC estimates	Confidence Interval (95%)	Error
K = 70	30.8462	10	30.8737	30.8198–30.9276	−0.09%
		20	30.8160	30.7625–30.8696	0.10%
		50	30.8513	30.7978–30.9048	−0.02%
		100	30.8318	30.7783–30.8853	0.05%
		200	30.8249	30.7715–30.8783	0.07%
K = 100	8.3780	10	8.4584∗	8.4256–8.4912	−0.96%
		20	8.4073	8.3749–8.4398	−0.35%
		50	8.3973	8.3649–8.4296	−0.23%
		100	8.3668	8.3344–8.3992	0.13%
		200	8.3705	8.3381–8.4028	0.09%
K = 130	0.5467	10	0.5869∗	0.5786 − 0.5951	−7.34%
		20	0.5696∗	0.5616–0.5775	−4.17%
		50	0.5573∗	0.5495–0.5652	−1.94%
		100	0.5470	0.5391–0.5549	−0.05%
		200	0.5523	0.5444–0.5602	−1.01%

Table 5 Plain vanilla option written on Asset 1, GVA scheme

Strike	Reference Value	No. of time steps	MC estimates	Confidence Interval (95%)	Error
K = 70	30.8462	10	30.3028*	30.2511–30.3545	1.76%
		20	30.5859*	30.5333–30.6384	0.84%
		50	30.7412*	30.6881–30.7943	0.34%
		100	30.8328	30.7796–30.8860	0.04%
		200	30.8521	30.7987–30.9055	− 0.02%
K = 100	8.3780	10	7.8458*	7.8151–7.8766	6.35%
		20	8.1059*	8.0744–8.1375	3.25%
		50	8.2812*	8.2491–8.3132	1.16%
		100	8.3399*	8.3078–8.3720	0.45%
		200	8.3799	8.3476–8.4122	− 0.02%
K = 130	0.5467	10	0.4424*	0.4355–0.4494	19.08%
		20	0.4895*	0.4822–0.4968	10.47%
		50	0.5240*	0.5164–0.5316	4.16%
		100	0.5368*	0.5291–0.5445	1.82%
		200	0.5454	0.5375–0.5532	0.25%

Table 6 Plain vanilla option written on Asset 2, TE scheme

Strike	Reference Value	No. of time steps	MC estimates	Confidence Interval (95%)	Error
$K = 70$	31.5397	10	31.5774	31.5190–31.6359	−0.12%
		20	31.4685*	31.4104–31.5266	0.23%
		50	31.5378	31.4798–31.5958	0.01%
		100	31.5116	31.4537–31.5695	0.09%
		200	31.5149	31.457–31.5728	0.08%
$K = 100$	9.6689	10	9.7750*	9.7402–9.8097	−1.10%
		20	9.6683	9.6339–9.7027	0.01%
		50	9.6837	9.6493–9.7181	−0.15%
		100	9.6558	9.6215–9.6901	0.14%
		200	9.6603	9.6260–9.6946	0.09%
$K = 130$	0.6072	10	0.6643*	0.6570–0.6717	−9.41%
		20	0.6234*	0.6164–0.6303	−2.66%
		50	0.6173*	0.6105–0.6242	−1.67%
		100	0.6117	0.6049–0.6185	−0.74%
		200	0.6091	0.6023 − 0.6158	− 0.31%

Table 7 Plain vanilla option written on Asset 2, GVA scheme

Strike	Reference Value	No. of time steps	MC estimates	Confidence Interval (95%)	Error
$K = 70$	31.5397	10	30.7237*	30.6679–30.7794	2.59%
		20	31.1076*	31.0509–31.1644	1.37%
		50	31.3530*	31.2956–31.4104	0.59%
		100	31.5186	31.461–31.5762	0.07%
		200	31.5372	31.4794–31.5950	0.01%
$K = 100$	9.6689	10	8.8934*	8.8612–8.9256	8.02%
		20	9.2535*	9.2203–9.2868	4.30%
		50	9.4986*	9.4648–9.5325	1.76%
		100	9.6168*	9.5827–9.6508	0.54%
		200	9.6496	9.6154–9.6838	0.20%
$K = 130$	0.6072	10	0.4325*	0.4270–0.4381	28.76%
		20	0.5132*	0.5071–0.5193	15.48%
		50	0.5688*	0.5623–0.5753	6.32%
		100	0.5863*	0.5797–0.5929	3.44%
		200	0.6018	0.5951–0.6085	0.89%

Table 8 Computational time as function of the number of time steps. The number of simulated paths is fixed to 5×10^5

Time steps	10	20	50	100	200
TE (s)	5.48	10.74	25.84	53.56	104.08
GVA (s)	11.66	23.48	59.68	119.74	244.54

Results refer to the implementation on a laptop PC with an Intel Core i7 CPU and 8 GB RAM via Matlab MEX files

Table 9 Asian option with monthly monitoring

	TE		GVA	
M_s	Value	Confidence Interval (95%)	Value	Confidence Interval (95%)
5×10^5 simulations				
1	11.6163	11.5731–11.6595	10.6334	10.5937–10.6732
5	11.5223	11.4799–11.5646	11.2828	11.2412–11.3244
10	11.4814	11.4392–11.5236	11.3440	11.3022–11.3859
15	11.4659	11.4238–11.5079	11.4115	11.3695–11.4534
20	11.4441	11.4020–11.4861	11.4318	11.3898–11.4739
5×10^6 simulations				
1	11.6414	11.6277–11.6551	10.6233	10.6108–10.6359
5	11.4852	11.4718–11.4985	11.2837	11.2705–11.2969
10	11.4711	11.4578–11.4845	11.3663	11.3530–11.3795
15	11.4545	11.4412–11.4678	11.4281	11.4104–11.4373
20	11.4456	11.4323–11.4589	11.4481	11.4306–11.4669

The time step is $\Delta = \Delta_M / M_s = \frac{1}{12 M_s}$

We now compare the two simulation schemes pricing an Asian option with payoff in $T = 1$ equal to

$$\left(\frac{1}{12} \sum_{i=1}^{12} f_1(0) e^{y_1(i \Delta_M)} + \frac{1}{12} \sum_{i=1}^{12} f_2(0) e^{y_2(i \Delta_M)} - K \right)^+ ,$$

with $\Delta_M = \frac{1}{12}$, i.e., we are considering an Asian option with monthly monitoring, and $K = 200$ (all the other parameters as above). In Table 9 we compare the two simulation schemes: results confirm that the TE scheme outperforms the GVA when the time grid is coarse.

Nonetheless, the GVA scheme proposed embeds the inherent advantage to implement the exact sampling of Wishart process thanks to the algorithm in Ahdida and Alfonsi (2013). This feature is of great importance in all the cases in which we need to estimate the (conditional) moments of the distribution of the elements of $\Sigma(t)$. The ability to consistently deal with the discretization of Wishart process is, indeed, the main reason that led us to develop the new scheme. To better highlight this advantage, in Table 10 we exploit the two simulation algorithms to compute the absolute value of the real part of the joint characteristic function

Table 10 Absolute value of the real part of the joint characteristic function (38): 5×10^5 simulations, parameters as in Eq.(39)

Reference Value	No. of time steps	TE Confidence Interval	GVA Confidence Interval
$a = 0$			
0.9835	10	0.9828–0.9832 *	0.9834–0.9835
	20	0.9833–0.9834 *	0.9834–0.9835
	50	0.9834–0.9835	0.9834–0.9835
	100	0.9834–0.9835	0.9834–0.9835
	200	0.9835–0.9836	0.9834–0.9835
$a = 0.1$			
0.8204	10	0.8209–0.8212 *	0.8206–0.8209 *
	20	0.8206–0.8209 *	0.8203–0.8206
	50	0.8204–0.8207	0.8203–0.8206
	100	0.8203–0.8206	0.8202–0.8205
	200	0.8203–0.8206	0.8203–0.8206
$a = 0.5$			
0.7561	10	0.7549–0.7546 *	0.7546–0.7543 *
	20	0.7555–0.7552 *	0.7554–0.7551 *
	50	0.7560–0.7557 *	0.7559–0.7556 *
	100	0.7561–0.7558	0.7562–0.7558
	200	0.7560–0.7559	0.7561–0.7558
$a = 1$			
0.0065	10	0.0040–0.0051 *	0.0009–0.0019 *
	20	0.0049–0.0060 *	0.0030–0.0041 *
	50	0.0056–0.0067	0.0047–0.0058 *
	100	0.0057–0.0067	0.0055–0.0065
	200	0.0057–0.0068	0.0056–0.0066

$$|\text{Re}\phi(\Lambda_X, \Lambda_V, T)| = \left|\text{Re}\mathbb{E}\left[\exp\left(\iota \langle \Lambda_X, X(T)\rangle + \iota \text{Tr}\left[\Lambda_V \Sigma(T)\right]\right)\right]\right|, \quad (38)$$

setting

$$X(T) = \begin{bmatrix} \log(f_1(T)) \\ \log(f_2(T)) \end{bmatrix}, \quad \Lambda_X = \begin{bmatrix} 0 \\ a \end{bmatrix}, \quad \Lambda_V = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \quad (39)$$

varying the parameter a. Notice that a closed form solution for this characteristic function is provided in Da Fonseca and Grasselli (2011). For large values of a, e.g., $a = 1$, the TE scheme outperforms the GVA one, exactly as in the option pricing cases above described. However, the opposite happens for small values of a, i.e., when the contribution of the simulation of the Wishart process is more important w.r.t. the log-asset value. Moreover, when $a = 0$, i.e., we only need to simulate the Wishart process

$\Sigma(T)$, the GVA scheme performance is independent on the number of the time steps, since it is an exact simulation scheme, as shown in Ahdida and Alfonsi (2013).

6 Concluding remarks

The matrix structure of Wishart-based stochastic volatility models provides a remarkable degree of flexibility in describing the evolution of asset(s) volatility. Realistic implementations, though, require the development of specific numerical techniques in order to deal with the inherent level of complexity. In this article we have shown, leveraging on a thorough analysis of distributional properties of Wishart process, some possible solutions intended to make this class of model more suitable for real market applications. Accordingly, we hope that our contribution will increase the interest of researchers and practitioners towards matrix-variate stochastic volatility dynamics.

Appendix A Jacobian matrix of WASC-Heston parameters mapping

Let $g_i^{H-WA}(\pi_{WA}, T)$ be the mapping presented in Sect. 4 between the WASC and Heston model parameters corresponding to the i-th basket component. Here we show the elements of matrix $J_{i,T}^{H-WA}(\pi_{WA})$ defined by (32). For the sake of simplicity, and without loss of generality, we consider $d = 2$. By direct computation the following holds for $i, j, k = 1, 2$:

$$\frac{\partial v_{0,i}}{\partial \Sigma_{j,k}} = \mathbb{1}_{i=j=k}, \quad \frac{\partial \kappa_i}{\partial \Sigma_{j,k}} = \frac{1}{t}\left(-\frac{1}{\gamma_i}\frac{\partial \gamma_i}{\partial \Sigma_{j,k}} + \frac{1}{\Sigma_{j,k}}\mathbb{1}_{i=j=k}\right),$$

$$\frac{\partial \kappa_i}{\partial M_{j,k}} = -\frac{1}{\gamma_i t}\frac{\partial \gamma_i}{\partial M_{j,k}}, \quad \frac{\partial \eta_i}{\partial \Sigma_{j,k}} = \frac{2\vartheta_i\left(1 - (1+\kappa_i t)e^{-\kappa_i t}\right)}{\eta_i\left(1 - e^{-\kappa_i t}\right)}\frac{\partial \kappa_i}{\partial \Sigma_{j,k}},$$

$$\frac{\partial \eta_i}{\partial M_{j,k}} = \frac{2\kappa_i}{\eta_i\left(1 - e^{-\kappa_i t}\right)}\frac{\partial \vartheta_i}{\partial M_{j,k}} + \frac{1 - (1+\kappa_i t)e^{-\kappa_i t}}{1 - e^{-\kappa_i t}}\vartheta_i\frac{\partial \kappa_i}{\partial M_{j,k}},$$

$$\frac{\partial \eta_i}{\partial Q_{j,k}} = \frac{2\kappa_i}{\eta_i\left(1 - e^{-\kappa_i t}\right)}\frac{\partial \vartheta_i}{\partial Q_{j,k}},$$

$$\frac{\partial \theta_i}{\partial \Sigma_{j,k}} = \frac{\beta \eta_i}{4\kappa_i^2}\left(2\kappa_i\frac{\partial \eta_i}{\partial \Sigma_{j,k}} - \eta_i\frac{\partial \kappa_i}{\partial \Sigma_{j,k}}\right), \quad \frac{\partial \theta_i}{\partial \beta} = \frac{\eta_i^2}{4\kappa_i},$$

$$\frac{\partial \theta_i}{\partial M_{j,k}} = \frac{\beta \eta_i}{4\kappa_i^2}\left(2\kappa_i \frac{\partial \eta_i}{\partial M_{j,k}} - \eta_i \frac{\partial \kappa_i}{\partial M_{j,k}}\right), \qquad \frac{\partial \theta_i}{\partial Q_{j,k}} = \frac{\beta \eta_i}{2\kappa_i}\frac{\partial \eta_i}{\partial Q_{j,k}},$$

$$\frac{\partial \rho_1}{\partial Q_{1,1}} = \frac{Q_{2,1}(Q_{2,1}r_1 - Q_{1,1}r_2)}{(Q_{1,1}^2 + Q_{2,1}^2)^{3/2}}, \qquad \frac{\partial \rho_1}{\partial Q_{2,1}} = \frac{Q_{1,1}(Q_{1,1}r_2 - Q_{2,1}r_1)}{(Q_{1,1}^2 + Q_{2,1}^2)^{3/2}},$$

$$\frac{\partial \rho_2}{\partial Q_{1,2}} = \frac{Q_{2,2}(Q_{2,2}r_1 - Q_{1,2}r_2)}{(Q_{1,2}^2 + Q_{2,2}^2)^{3/2}}, \qquad \frac{\partial \rho_2}{\partial Q_{2,2}} = \frac{Q_{1,2}(Q_{1,2}r_2 - Q_{2,2}r_1)}{(Q_{1,2}^2 + Q_{2,2}^2)^{3/2}},$$

$$\frac{\partial \rho_i}{\partial r_j} = \frac{Q_{j,i}}{\sqrt{Q_{1,i}^2 + Q_{2,i}^2}}.$$

The unspecified elements of J^{H-WA} are equal to zero.

References

Ahdida A, Alfonsi A (2013) Exact and high-order discretization schemes for Wishart processes and their affine extensions. Ann Appl Probab 23(3):1025–1073

Alfonsi A, Kebaier A, Rey C (2016) Maximum likelihood estimation for Wishart processes. Stochast Process Appl 126(11):3243–3282

Boloorforoosh A, Christoffersen P, Fournier M, Gouriéroux C (2020) Beta risk in the cross-section of equities. Rev Financial Stud 33(9):4318–4366

Bru MF (1991) Wishart processes. J Theor Probab 4(4):725–751

Caldana R, Fusai G, Gnoatto A, Grasselli M (2016) General closed-form basket option pricing bounds. Quantitat Finance 16(4):535–554

Christoffersen P, Heston S, Jacobs K (2009) The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. Manage Sci 55(12):1914–1932

Cox JC, Ingersoll JE, Ross SA (1985) A theory of the term structure of interest rates. Econ J Econ Soc 53(2):385–407

Cui Y, del Baño Rollin S, Germano G (2017) Full and fast calibration of the Heston stochastic volatility model. Eur J Oper Res 263(3):625–638

Da Fonseca J, Grasselli M (2011) Riding on the smiles. Quantitat Finance 11(11):1609–1632

Da Fonseca J, Grasselli M, Ielpo F (2014) Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. Stud Nonlinear Dyn Econ 18(3):253–289

Da Fonseca J, Grasselli M, Tebaldi C (2007) Option pricing when correlations are stochastic: an analytical framework. Rev Derivat Res 10(2):151–180

Da Fonseca J, Grasselli M, Tebaldi C (2008) A multifactor volatility Heston model. Quantitat Finance 8(6):591–604

Golub GH, Van Loan CF (2012) Matrix computations, vol 3. JHU Press, Baltimore

Goudenege L, Molent A, Zanette A (2019) Pricing and hedging GMWB in the Heston and in the Black-Scholes with stochastic interest rate models. Comput Manage Sci 16(1–2):217–248

Gourieroux C, Sufana R (2004) Derivative pricing with multivariate stochastic volatility: Application to credit risk. Les Cahiers du CREF of HEC Montréal Working Paper No. CREF, pp 04–09

Gourieroux C, Sufana R (2010) Derivative pricing with Wishart multivariate stochastic volatility. J Business Econ Stat 28(3):438–451

Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financial Stud 6(2):327–343

Kollo T, von Rosen D (2006) Advanced multivariate statistics with matrices, vol 579. Springer, Berlin

Kourouklis S, Moschopoulos PG (1985) On the distribution of the trace of a noncentral Wishart matrix. Metron 43(2):85–92

La Bua G, Marazzina D (2019) Calibration and advanced simulation schemes for the Wishart stochastic volatility model. Quantitat Finance 19(6):997–1016

La Bua, G, Marazzina D (2020) A new class of multidimensional Wishart-based hybrid models. Submitted - Available at SSRN

Yolcu-Okur Y, Sayer T, Yilmaz B, Inkaya BA (2018) Computation of the delta of European options under stochastic volatility models. Comput Manage Sci 15(2):213–237

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.