

A Parsimonious Multi-Asset Heston Model: Calibration and Derivative Pricing

Georgi Dimitroff* Stefan Lorenz* Alexander Szimayer**

Abstract

We propose a parsimonious multi-asset Heston model, give a straightforward Monte Carlo simulation algorithm for derivative pricing, and provide an easy-to-implement calibration algorithm. The multi-asset Heston is consistently nested, i.e., each single-asset sub-model of our multi-asset model is a classical Heston model itself. As a notable feature the correlation structure between two single-assets sub-models is determined by the asset return correlation only. This specification is parsimonious and sound since the asset cross correlations are observable, in contrast to correlations of latent variables such as volatilities. The model is especially customized to market situations where no cross-asset options are liquidly traded and it is not possible to back out cross-asset correlation parameters. For this situation we suggest a hybrid calibration approach for identifying the model parameters consistent with option price data and asset price data.

Key Words: Heston model, multi-asset, option pricing, calibration, correlation

JEL Classification: G13

Mathematics Subject Classification (2000): 60G44, 60H30, 62M10

This version: September 22, 2009

*Fraunhofer ITWM, Kaiserslautern

**Faculty of Economics and Law, University of Bonn

1 Introduction

For pricing and managing medium-term and long-term equity options the model of Heston [1993] is widely applied in practice. Generalizations of Black-Scholes, other than Heston, have been put forward to capture volatility smile and skew observed in option price data (see, e.g., Bates [1996] for a Heston extension allowing for jumps, Hagan et al. [2002] for the popular SABR model, and in general Fouque et al. [2000] for an overview of diffusion-based stochastic volatility models and Cont and Tankov [2008] for a wide range of Lévy process models). However, the Heston model still appeals because of its richness and simplicity. The Heston model caters for volatility smile and skew, and its parameters have a direct economic interpretation. The latter is advantageous when the range of liquidly traded options does not fully cover the maturity of the option which is to be priced. Then calibration is rather a matter of extrapolation, and the appropriateness of the “extrapolated” Heston model can be assessed by inspecting the transparent model parameters (e.g., the mean level of the volatility), see Mikhailov and Nögel [2003] and Gatheral [2006]. The practitioner’s preference for Heston when pricing medium-term and long-term options is further supported by the empirical findings of Bakshi et al. [2000] where they conclude: “Modeling stochastic volatility is perhaps of true first-order importance.”

Recently, in the finance industry the need emerged for adequate pricing models for valuing medium-term and long-term options which are written on two or more assets. The need stems from medium-term and long-term options embedded in financial products in retail banking, insurance, and the pension fund business covering products such as principal protected notes, minimal return guarantees as part of unit-linked life-insurance products, and portfolio insurance. In this paper we put forward a model to price and manage medium-term and long-term multi-asset options.

We propose a parsimonious multi-asset Heston model, give a straightforward Monte Carlo simulation algorithm for derivative pricing, and provide an easy-to-implement calibration algorithm. The multi-asset Heston is consistently nested, i.e., each single-asset sub-model of our multi-asset model is a classical Heston model itself. This enables the

calibration of the single-asset parameters from single-asset option price data. More precisely, for each asset and its corresponding volatility, the parameters can be calibrated in the same fashion as done for the single-asset Heston. As a notable feature the correlation structure between two single-assets sub-models is determined by the asset return correlation only. This specification is parsimonious and sound since the asset cross correlations are observable, in contrast to correlations of latent variables such as volatilities.

The model is especially customized to market situations where no cross-asset options are liquidly traded and it is not possible to back out cross-asset correlation parameters. For this situation we suggest a hybrid approach for identifying the model parameters consistent with option price data and asset price data. In particular, we combine information from option price data for calibrating the implied parameters of each single-asset sub-model and information from asset price data for calibrating the empirical cross-asset dependence parameters. Our calibration procedure is based on the two-asset case, and can be extended to the full multi-asset model by considering all two-asset sub-models. Thus the stability of the method for d assets is identical to the simple two-asset case, and the computer run times grow with $(d - 1)d/2$ (the number of two-dimensional sub-models). For the two-asset Heston, there are four missing cross correlations. The cross-correlation asset-asset is determined and the remaining three cross-correlations (asset-vol, vol-asset, and vol-vol) are expressed as functions of the other parameters. Paths are then generated using Monte-Carlo simulation. These paths are observed at the same frequency as the data on which the empirical correlation is based. From the simulated observations the simulated empirical correlation is computed. Using the bisection methodology, the model cross-correlation asset-asset is then adjusted such that the simulated empirical correlation and the empirical correlation coincide. In case the pairwise cross-asset correlations result in a matrix that is not a correlation matrix we investigate three different regularization procedures.

The multi-asset Heston we propose is parsimonious and well manageable. A related and more complex extension of Heston is provided by Da Fonseca et al. [2008]. They use a multifactor Wishart process to model the asset volatility replacing the univariate square-root process in the single-asset Heston framework, see also Gouriéroux [2006]. Da Fonseca

et al. [2007] extended the Wishart-based Heston extension to a multi-asset version. The dependence structure we utilize is also investigated by Kahl and Günther [2005]. However, it is interesting to note that their Theorem 4.1 proves to be dispensable since the matrix they study for conditions for being a correlation matrix is shown to be a true correlation matrix by construction in the subsequent modeling section.

The paper is organized as follows. Section 2 introduces the model setup, and studies the correlations under certain transformations. The calibration procedure is laid out in Section 3. In Section 4 Monte Carlo simulation and the hybrid calibration algorithm are discussed. Section 5 investigates calibration from asset price data if only historical data is available. Section 6 concludes.

2 Model

In this section, the stochastic volatility model of Heston [1993] is reviewed. Subsequently, a parsimonious multi-asset extension is developed and studied for its properties.

2.1 The Single-Asset Heston Model

In the framework of Heston [1993], the risk-neutral dynamics of the asset price process $S(t)$ is given by

$$dS(t) = (r(t) - q(t)) S_t dt + \sqrt{\nu(t)} S_t dW(t), \quad S(0) > 0, \quad (1)$$

where $r(t)$ is risk-free rate, $q(t)$ is the dividend rate, and $W(t)$ is a Wiener process. The risk neutral dynamics of the volatility $\sqrt{\nu(t)}$ is given by

$$d\nu(t) = \kappa (\bar{\nu} - \nu(t)) dt + \eta \sqrt{\nu(t)} [\rho dW(t) + \sqrt{1 - \rho^2} d\widetilde{W}(t)], \quad \nu(0) > 0, \quad (2)$$

where \widetilde{W} is another Wiener process independent of W . Thus, $\nu(t)$ describes a mean-reverting process with reversion rate $\kappa > 0$, mean level $\bar{\nu}$, “volatility” $\eta > 0$ and initial value $\nu(0)$. The correlation parameter ρ allows for modeling possible leverage effects. The parameters $\rho, \kappa, \bar{\nu}, \nu(0), \eta$ are assumed to be given, e.g., calibrated from plain vanilla option price data.

2.2 A Parsimonious Multi-Asset Heston Model

Heston can be extended to a multi-asset model in different ways. For example, the Heston model can be extended by generalizing the real-valued variance rate process to a matrix valued version. Wishart processes represent the matrix analogue of the square-root mean-reverting process and are used for volatility modeling by Gouriéroux [2006].

We propose a parsimonious multi-asset extension of the single-asset Heston with properties: (1) each single-asset sub-model forms a traditional Heston model; and (2) parameters are single-asset Heston parameters and asset-asset cross-correlations. These properties are desirable since (1) enables using established calibration procedures for the single-asset sub-models parameters, and (2) is parsimonious in the sense that $d(d-1)/2$ asset-asset cross-correlations are required for our d -asset Heston model.

Now consider a system of d assets with price and volatility processes $(S_1(t), \nu_1(t)), \dots, (S_d(t), \nu_d(t))$. For $i = 1, \dots, d$ we can write (1) and (2) in the vectorized form

$$\begin{aligned} \begin{pmatrix} dS_i(t) \\ d\nu_i^2(t) \end{pmatrix} &= \begin{pmatrix} S_i(t)(r(t) - q_i(t)) \\ \kappa_i (\bar{\nu}_i - \nu_i(t)) \end{pmatrix} dt \\ &+ \begin{pmatrix} S_i(t) \sqrt{\nu_i(t)} & 0 \\ 0 & \eta_i \sqrt{\nu_i(t)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \rho_i & \sqrt{1 - \rho_i^2} \end{pmatrix} \begin{pmatrix} dW_i(t) \\ d\widetilde{W}_i(t) \end{pmatrix}, \end{aligned} \quad (3)$$

where $W_i(t)$ and $\widetilde{W}_i(t)$ are independent Wiener processes. The parameter describing the bivariate process are collected in $\theta_i = (\rho_i, \kappa_i, \bar{\nu}_i, \nu_i^2(0), \eta_i)$.

The model is completely described except for the dependence structure. So far, the correlations asset-volatility are given for $i = 1, \dots, d$, but the cross-correlations asset-asset of $W_i(t)$ and $W_j(t)$, asset-volatility of $W_i(t)$ and $\widetilde{W}_j(t)$, and volatility-volatility of $\widetilde{W}_i(t)$ and $\widetilde{W}_j(t)$, ($i \neq j$), are not yet specified.

Define the d -dimensional Wiener processes $W(t) = (W_1(t), \dots, W_d(t))$ and $\widetilde{W}(t) = (\widetilde{W}_1(t), \dots, \widetilde{W}_d(t))$. Our assumptions on the dependence structure are summarized below.

Assumption 2.1. *The Wiener processes $W(t)$ and $\widetilde{W}(t)$ satisfy:*

- (1) $W(t)$ has correlation matrix $\Sigma^S = (\rho_{i,j})_{1 \leq i, j \leq d}$, i.e., $dW_i(t) dW_j(t) = \rho_{i,j} dt$;

(2) $\widetilde{W}(t)$ has correlation matrix I_d , i.e., $d\widetilde{W}_i(t) d\widetilde{W}_j(t) = \delta_{i,j} dt$,¹

(3) $W(t)$ and $\widetilde{W}(t)$ are independent.

The dependence assumption on $(W(t), \widetilde{W}(t))$ is permissible. To see this, observe that the correlation matrix of $(W(t), \widetilde{W}(t))$ is

$$\Sigma^{(W, \widetilde{W})} = \begin{pmatrix} \Sigma^S & 0 \\ 0 & I_d \end{pmatrix}.$$

Provided that Σ^S is a correlation matrix, which holds by assumption, the matrix $\Sigma^{(W, \widetilde{W})}$ is also a correlation matrix.

Assumption (1) allows all possible correlation matrices for the asset prices. In contrast, assumptions (2) and (3) are restrictions stating that the dependence of the volatilities is carried via the correlations of the corresponding asset prices $\rho_{i,j}$, which is then transferred to each volatility by the corresponding Heston parameter asset-volatility correlation ρ_i and ρ_j .

2.2.1 Correlation Structure

The model specification (3) and Assumption 2.1 determine the correlation structure of the vector process $(S_1(t), \nu_1(t), \dots, S_d(t), \nu_d(t))$. The instantaneous correlation structure is stated in the following Proposition.

Proposition 2.2. *Let $(S_1(t), \nu_1(t), \dots, S_d(t), \nu_d(t))$ be the $2d$ -dimensional process defined in (3) and Assumption 2.1 holds. Then for $1 \leq i, j \leq d$ the instantaneous covariances of the components are*

$$(1) \quad dS_i(t) dS_j(t) = \rho_{i,j} S_i(t) S_j(t) \sqrt{\nu_i(t) \nu_j(t)} dt;$$

$$(2) \quad dS_i(t) d\nu_j(t) = \rho_{i,j} \rho_j S_i(t) \eta_j \sqrt{\nu_i(t) \nu_j(t)} dt;$$

$$(3) \quad d\nu_i(t) d\nu_j(t) = \begin{cases} \rho_{i,j} \rho_i \rho_j \eta_i \eta_j \sqrt{\nu_i(t) \nu_j(t)} dt, & \text{for } i \neq j, \\ \eta_i \eta_j \sqrt{\nu_i(t) \nu_j(t)} dt, & \text{for } i = j. \end{cases}$$

¹ $\delta_{i,j}$ is Kronecker's delta, i.e. $\delta_{i,j} = 1$ if $i = j$, and $\delta_{i,j} = 0$ if $i \neq j$.

Thus the instantaneous correlations of the components are

$$(1) \frac{dS_i(t) dS_j(t)}{\sqrt{(dS_i(t))^2 (dS_j(t))^2}} = \rho_{i,j};$$

$$(2) \frac{dS_i(t) d\nu_j(t)}{\sqrt{(dS_i(t))^2 (d\nu_j(t))^2}} = \rho_{i,j} \rho_j;$$

$$(3) \frac{d\nu_i(t) d\nu_j(t)}{\sqrt{(d\nu_i(t))^2 (d\nu_j(t))^2}} = \begin{cases} \rho_{i,j} \rho_i \rho_j, & \text{for } i \neq j, \\ 1, & \text{for } i = j. \end{cases}$$

Proof. The proof utilizes Assumption (2.1) and is a straightforward computation along the lines of the model (3). Observe that we use the "infinitesimal correlation" notation

$$dS_i(t)dS_j(t) := \frac{d}{dt} \langle S_i, S_j \rangle_t$$

□

Remark 2.3. *The dependence structure we utilize is also investigated by Kahl and Günther [2005]. However, it is interesting to note that their Theorem 4.1 proves to be dispensable since the matrix they study for conditions for being a correlation matrix is shown to be a true correlation matrix by construction in our setup.*

2.3 Empirical Correlations

Our model is designed for markets with little cross-asset option price data complicating the calibration of potential correlation parameters. Assuming that the single-asset Heston parameters are calibrated, the free parameters for correlating the single-asset Heston models are contained in the matrix Σ^S . Thus $(d-1)d/2$ additional parameters have to be calibrated to correlate the single-asset models.

In the absence of sufficient cross-asset option price data, we rely on historical time series from which we extract empirical cross-asset correlations $\widehat{\Sigma}^{\text{emp}}$ of the asset prices. The key idea for calibration is to adjust the model cross-correlations in Σ^S according to the observed empirical correlations Σ^{emp} . It is important to note that Σ^S is the correlation of infinitesimal noise carried by the Wiener process $W(t)$, whereas $\widehat{\Sigma}^{\text{emp}}$ is the correlation of observed discrete-time asset price data. This observation is crucial since the Σ^S is not

influenced by an equivalent measure transformations, while the cross-asset correlations may and typically will change. Therefore in one of the estimation setups considered below we will estimate the Σ^S using $\mathbb{E}^P \Sigma^{\text{emp}}$ under the objective measure. The switch to the physical measure utilizes the Heston assumption for the shape of the market price of volatility risk.

Let us now formulate what is exactly meant by the empirical correlations $\widehat{\Sigma}^{\text{emp}}$. For observation times $(t_k)_{k=0,\dots,K}$ we are given asset price data $(S_1(t_k), \dots, S_d(t_k))_{k=0,\dots,K}$. The asset prices are suitably transformed to returns, log-returns or first differences, i.e.: $X_i = (X_i(t_k))_{k=1,\dots,K}$ with $X_i = F(S_i)$.

$\widehat{\Sigma}_K^{\text{emp}}(\Sigma)$ is then the empirical correlation matrix of these $(X_i)_{i=1,\dots,d}$ using the K -step time series. In particular, the elements of $\widehat{\Sigma}_K^{\text{emp}}(\Sigma) = (\widehat{\rho}_{i,j:K}^{\text{emp}}(\Sigma))_{1 \leq i,j \leq d}$ are defined by

$$\widehat{\rho}_{i,j:K}^{\text{emp}}(\Sigma) = \frac{\widehat{\nu}_{i,j:K}^{\text{emp}}(\Sigma)}{\sqrt{\widehat{\nu}_{i,i:K}^{\text{emp}}(\Sigma) \widehat{\nu}_{j,j:K}^{\text{emp}}(\Sigma)}}, \quad (4)$$

where

$$\widehat{\nu}_{i,j:K}^{\text{emp}}(\Sigma) = \frac{1}{K-1} \sum_{k=1,\dots,K} (X_i(t_k) - \bar{X}_i)(X_j(t_k) - \bar{X}_j), \quad \text{and} \quad \bar{X}_i = \frac{1}{K} \sum_{k=1,\dots,K} X_i(t_k). \quad (5)$$

The transformation F applied commonly to asset price data is to form log returns. Other alternatives are possible (see Example 1), though sensible transforms are closely related to scaled first differences.

Example 1. *Transformations typically used include*

$$(1) \text{ first differences: } F(S(t_0), \dots, S(t_K)) = (S(t_1) - S(t_0), \dots, S(t_K) - S(t_{K-1}));$$

$$(2) \text{ returns: } F(S(t_0), \dots, S(t_K)) = \left(\frac{S(t_1) - S(t_0)}{S(t_0)}, \dots, \frac{S(t_K) - S(t_{K-1})}{S(t_{K-1})} \right);$$

$$(3) \text{ log-returns: } F(S(t_0), \dots, S(t_K)) = (\log(S(t_1)/S(t_0)), \dots, \log(S(t_K)/S(t_{K-1}))).$$

In the following we will concentrate on the log-returns as the most widely used transforms of the time series. However, we stress that the results will hold also for the returns and the first differences.

2.3.1 Properties

As already mentioned above we focus on the log-returns and for a discretized asset price $S(t_k)$ in the following we consider

$$X(t_k) := \log \frac{S(t_k)}{S(t_{k-1})}.$$

Theorem 2.4. *Let $(S_i(t), \nu_i(t))_{i=1, \dots, d}$ be the processes defined in (3) and Assumption 2.1 holds. Let $\pi_n = (t_{k,n})_{k=1, \dots, K_n(T)}$ be partitions of $[0, T]$, with $\text{mesh}(\pi_n) = \max_{i=1, \dots, K_n} |t_{i,n} - t_{i-1,n}| \rightarrow 0$, for $n \rightarrow \infty$. Then for all $1 \leq i, j \leq d$ we have*

- (1) $\lim_{T \rightarrow 0} \lim_{n \rightarrow \infty} \widehat{\rho}_{i,j:K_n(T)}^{\text{emp}}(\Sigma) = \rho_{i,j}$ in probability,
- (2) $\lim_{T \rightarrow \infty} \lim_{n \rightarrow \infty} \widehat{\rho}_{i,j:K_n(T)}^{\text{emp}}(\Sigma) = \rho_{i,j} \frac{\mathbb{E} \sqrt{\nu_{ij}(\infty)}}{\sqrt{\mathbb{E} \nu_i(\infty) \mathbb{E} \nu_j(\infty)}} = \rho_{i,j} \frac{\mathbb{E} \sqrt{\nu_{ij}(\infty)}}{\sqrt{\nu_i \nu_j}}$ in probability, where $\nu_i(\infty)$, $\nu_j(\infty)$ and $\nu_{ij}(\infty)$ denote random variables having the stationary distributions of the processes $\nu_i(t)$, $\nu_j(t)$ and $\nu_i(t) \cdot \nu_j(t)$ respectively.

Proof. A straightforward transformation of the expression for $\widehat{\rho}_{i,j:K_n(T)}^{\text{emp}}(\Sigma)$ yields

$$\widehat{\rho}_{i,j:K_n(T)}^{\text{emp}}(\Sigma) = \frac{\sum_{k=1}^{K_n(T)} X_i(t_k) X_j(t_k) - K_n(T) \bar{X}_i \bar{X}_j}{\sqrt{\sum_{k=1}^{K_n(T)} X_i^2(t_k) - K_n(T) \bar{X}_i^2} \sqrt{\sum_{k=1}^{K_n(T)} X_j^2(t_k) - K_n(T) \bar{X}_j^2}}. \quad (6)$$

Following the model (3) we have

$$S_i(t) = S_i(0) \exp \left[\int_0^t \sqrt{\nu_i(s)} dW_i(s) + \int_0^t (r(s) - q_i(s)) ds - \frac{1}{2} \int_0^t \nu_i(s) ds \right]$$

and the very definition of quadratic variation implies that almost surely the following convergence holds in probability:

$$\begin{aligned} \sum_{k=1}^{K_n(T)} X_i(t_k) X_j(t_k) &= \sum_{k=1}^{K_n(T)} (\log S_i(t_k) - \log S_i(t_{k-1})) (\log S_j(t_k) - \log S_j(t_{k-1})) \\ &\xrightarrow{n \rightarrow \infty} \langle \log S_i, \log S_j \rangle_T = \rho_{i,j} \int_0^t \sqrt{\nu_i(s) \nu_j(s)} ds. \end{aligned} \quad (7)$$

On the other hand for all i, j we have

$$K_n(T) \bar{X}_i \bar{X}_j = \frac{1}{K_n(T)} \log \frac{S_i(T)}{S_i(0)} \log \frac{S_j(T)}{S_j(0)} \xrightarrow{n \rightarrow \infty} 0, \quad \mathbb{P} - \text{a.s.} \quad (8)$$

Combining (6), (7) and (8) we obtain

$$\lim_{n \rightarrow \infty} \widehat{\rho}_{i,j;K_n(T)}^{\text{emp}}(\Sigma) = \rho_{i,j} \frac{\int_0^T \sqrt{\nu_i(s)\nu_j(s)} ds}{\sqrt{\int_0^T \nu_i(s) ds} \sqrt{\int_0^T \nu_j(s) ds}} \quad \text{in probability.} \quad (9)$$

Letting $T \rightarrow 0$ in (9) and recalling the continuity of ν_i we obtain the first claim in the theorem.

For the second claim we need the ergodicity of the processes ν_i , ν_j and (ν_i, ν_j) . This is well-known for the one dimensional CIR process. In the Appendix we provide the ergodicity of (ν_i, ν_j) as we were not able to find a reference for this fact in the existing literature. Clearly, the two dimensional result in particular implies the ergodicity of the one dimensional marginals. For the proof we use an approach developed and presented in Has'minskii [1980]. Once the ergodicity of the above processes is verified, we can apply the ergodic theorem to obtain the second almost sure convergence:

$$\lim_{n \rightarrow \infty} \widehat{\rho}_{i,j;K_n(T)}^{\text{emp}}(\Sigma) = \rho_{i,j} \frac{\frac{1}{T} \int_0^T \sqrt{\nu_i(s)\nu_j(s)} ds}{\sqrt{\frac{1}{T} \int_0^T \nu_i(s) ds} \sqrt{\frac{1}{T} \int_0^T \nu_j(s) ds}} \xrightarrow{T \rightarrow \infty} \rho_{i,j} \frac{\mathbb{E} \sqrt{\nu_i(\infty)\nu_j(\infty)}}{\sqrt{\mathbb{E} \nu_i(\infty)} \sqrt{\mathbb{E} \nu_j(\infty)}}.$$

Clearly the expectation of the invariant distribution of $\nu_i(t)$ is the mean reversion level $\bar{\nu}_i$ for all $i = 1, \dots, d$. This completes the proof. \square

Remark 2.5.

- (1) *The first part of Theorem 2.4 states that calculating correlations from high-frequency data observed over a rather short period renders $\widehat{\Sigma}^{\text{emp}}(\Sigma) = \Sigma$. Thus $\Sigma = \Sigma^{\text{emp}}$, and the correlation matrix is therefore not adjusted.*
- (2) *Note, that the first part of Theorem 2.4 is a theoretical statement. High-frequency data -if available at all- is typically subject to other effects, e.g., caused by the market micro-structure. These effects are making the observations noisy and are potentially distorting the correlation estimates. Therefore, we adopt the second statement as our approach for estimating the correlation structure.*

Remark 2.6.

- (1) *The Cauchy-Schwarz inequality immediately gives $|\widehat{\rho}_{i,j;n}^{\text{emp}}(\Sigma)| \leq |\rho_{i,j}|$ asymptotically for big n .*

(2) The empirical correlation is approximately linear in ρ with a positive slope strictly smaller than 1. Varying the model correlation $\rho_{i,j}$ by $\Delta\rho$, we can write

$$\widehat{\rho}_{i,j;n}^{\text{emp}}(\Sigma + \Delta\rho) \approx \widehat{\rho}_{i,j;n}^{\text{emp}}(\Sigma) + \Delta\rho \frac{\mathbb{E}\sqrt{\nu_i(s)\nu_j(\infty)}}{\sqrt{\mathbb{E}\nu_i(\infty)}\sqrt{\mathbb{E}\nu_j(\infty)}}.$$

where the distribution of $(\nu_1(\infty), \dots, \nu_d(\infty))$ depends on Σ .

(3) The range of the possible empirical correlations is contained in $(-1, 1)$, except for the case if two single-asset Heston sub-models are perfectly correlated in the asset price and the variance rate process, which we want to exclude as it introduces singularity into the model.

3 Calibration and simulation of the multi-asset Heston model

We assume that the parameters $\theta_i = (\rho_i, \kappa_i, \bar{\nu}_i, \nu_i(0), \eta_i)$ describing the risk-neutral dynamics of the single-asset Heston sub-models are given. They can be obtained by single-asset Heston calibration routine, which is rather standard. It remains to calibrate the correlation structure between the one-dimensional models. The ideal situation would be if we had a liquid market of correlation sensitive cross-asset derivatives which can be used for calibration. Unfortunately, the existence of such markets is rather an exception and therefore we focus on the alternative to use the historical asset time series to determine the correlations. This is done by calculating the historical empirical cross-asset correlations $\widehat{\Sigma}^{\text{emp}}$ using (4) and extracting Σ out of them. While the empirical cross-asset correlations $\widehat{\Sigma}^{\text{emp}}$ depend on the underlying measure the correlation structure Σ of the driving Brownian noise is invariant under equivalent measure change, so it must be the same. As the time series evolve according to the objective measure and therefore the obtained cross-asset empirical correlations are with respect to the physical measure we also need to determine the mean-reversion level of all volatility processes under this measure.

The calibration procedure capitalizes on the fact that $\widehat{\Sigma}^{\text{emp}} \rightarrow \mathbb{E}^P \widehat{\Sigma}^{\text{emp}}(\Sigma^{\text{true}})$ where $\mathbb{E}^P \widehat{\Sigma}^{\text{emp}}(\Sigma)$ denotes the expectation under the objective measure P of the empirical corre-

lations given the true correlation structure Σ^{true} between the driving Brownian motions. Therefore determining the unknown correlation parameters Σ amounts to solving the problem

$$\min_{\Sigma \in \text{Cor}(d)} \left\| \mathbb{E}^P \widehat{\Sigma}^{\text{emp}}(\Sigma) - \widehat{\Sigma}^{\text{emp}} \right\|, \quad (10)$$

where $\text{Cor}(d)$ is the space of $d \times d$ -dimensional correlation matrices, i.e., $d \times d$ -matrices over \mathbb{R} that are symmetric, positive semi-definite and unit elements on the main diagonal, and $\|\cdot\|$ is a suitable matrix-norm.

In some situations it is sensible to assume that the historical cross-asset empirical correlations are actually empirical correlations under the risk neutral dynamics. This would be the case if we assume that the market price of risk is zero. Essentially (7) states that, if we have rather high frequency data e.g. daily returns, this amounts to the assumption that the market price of volatility risk is zero, as the drift of the asset prices under the objective measure becomes negligible when we build the empirical correlations. In this setting the calibration procedure boils down to solving

$$\min_{\Sigma \in \text{Cor}(d)} \left\| \mathbb{E}^Q \widehat{\Sigma}^{\text{emp}}(\Sigma) - \widehat{\Sigma}^{\text{emp}} \right\|, \quad (11)$$

where \mathbb{E}^Q denotes expectation with respect to the equivalent martingale measure Q .

The input of the algorithm in both settings are the parameters describing the one-dimensional risk neutral dynamics and the historical time series of the assets.

It is sometimes even relevant to assume that the expected cross-asset empirical correlations $\Sigma^{\text{emp}} = \mathbb{E}^Q \widehat{\Sigma}^{\text{emp}}$ under the equivalent martingale measure are provided and our task is to calibrate the infinitesimal correlations Σ directly under the relevant pricing measure. In this case we have as an input the expected empirical correlations and do not need the asset time series.

Here calibrating the yet unknown parameters contained in the correlation matrix Σ^S is done by solving the problem

$$\min_{\Sigma \in \text{Cor}(d)} \left\| \mathbb{E}^Q \widehat{\Sigma}^{\text{emp}}(\Sigma) - \Sigma^{\text{emp}} \right\|. \quad (12)$$

We first start by describing the algorithm in the setup where the cross-asset empirical correlations $\widehat{\Sigma}^{\text{emp}}$ or even the expected cross-asset empirical correlations Σ^{emp} under the

equivalent martingale measure are provided and we aim to determine Σ . That is we solve (11) and (12). We call the algorithm *calibration with correlation adjustment*.

The calibration in the first setup, where we extract the empirical correlations from the historical data and then solve (10) is provided in section 3.2.

3.1 Correlation adjustment algorithm

If we consider (12) the input is a $d \times d$ -dimensional matrix Σ^{emp} , given by the user, which does not necessarily have to be positive semi-definite. It only has to be symmetric with diagonal entries of value 1. This is then taken as the expected empirical correlation matrix under the equivalent martingale measure. The procedure for (11) is the same, with the only difference that we take $\widehat{\Sigma}^{\text{emp}}$ as a proxy for Σ^{emp} . Therefore in the following we will refer to (12) meaning also (11). Throughout this section all expectations are with respect to the risk neutral measure Q .

From section 2.2 we know, that we do not need to solve a d -dimensional problem, but $(d-1)d/2$ 2-dimensional problems. For these 2-dimensional problems, we propose an algorithm that calibrates the model and returns the desired correlation matrix Σ .

First, the two-asset model and its calibration is investigated. With given Heston parameters θ_1 and θ_2 for the two models, we have $\mathbb{E}^Q \widehat{\rho}_{1,2}^{\text{emp}}(\rho_{1,2}) = f_{(\theta_1, \theta_2)}(\rho_{1,2})$, for a strictly increasing and continuous function $f_{(\theta_1, \theta_2)} : [-1, 1] \mapsto \mathbb{R}$, parameterized by θ_1 and θ_2 . Given that $\rho_{1,2}^{\text{emp}}$ is contained in the range $\{f_{(\theta_1, \theta_2)}(x) : |x| \leq 1\}$, the model can be uniquely calibrated by numerically solving the equation $\mathbb{E}^Q \widehat{\rho}_{1,2}^{\text{emp}}(\rho_{1,2}) = \rho_{1,2}^{\text{emp}}$.²

²Remember that here $\rho_{1,2}^{\text{emp}}$ is provided by the user.

3.1.1 Two-Asset Calibration

Setting $d = 2$, we can expand (3) to the four-dimensional system:

$$\begin{aligned} \begin{pmatrix} dS_1(t) \\ d\nu_1^2(t) \\ dS_2(t) \\ d\nu_2^2(t) \end{pmatrix} &= \begin{pmatrix} S_1(t)(r(t) - q(t)) \\ \kappa_1 (\bar{\nu}_1^2 - \nu_1^2(t)) \\ S_2(t)(r(t) - q(t)) \\ \kappa_2 (\bar{\nu}_2^2 - \nu_2^2(t)) \end{pmatrix} dt \\ &+ \begin{pmatrix} S_1(t) \nu_1(t) \\ \eta_1 \sqrt{\nu_1^2(t)} \\ S_2(t) \nu_2(t) \\ \eta_2 \sqrt{\nu_2^2(t)} \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 & 0 \\ \rho_1 & \sqrt{1 - \rho_1^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \rho_2 & \sqrt{1 - \rho_2^2} \end{pmatrix} \begin{pmatrix} dW_1(t) \\ d\widetilde{W}_1(t) \\ dW_2(t) \\ d\widetilde{W}_2(t) \end{pmatrix}, \end{aligned} \quad (13)$$

where $W_1(t)$ and $W_2(t)$ are correlated with $\rho_{1,2}$, and the pairs $W_1(t)$ and $\widetilde{W}_1(t)$, $W_2(t)$ and $\widetilde{W}_2(t)$ and $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent.

For $|\rho_{1,2}| < 1$ we apply the standard Gram-Schmidt-type orthogonalization procedure and define $\bar{W}_1(t) = W_1(t)$ and $\bar{W}_2(t) = (\rho_{1,2} W_1(t) - W_2(t))/\sqrt{1 - \rho_{1,2}^2}$.³ Then $(\bar{W}_1(t), \widetilde{W}_1(t), \bar{W}_2(t), \widetilde{W}_2(t))$ is a Wiener process with uncorrelated components. The latter part of (13) then becomes

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ \rho_1 & \sqrt{1 - \rho_1^2} & 0 & 0 \\ \rho_{1,2} & 0 & \sqrt{1 - \rho_{1,2}^2} & 0 \\ \rho_2 \rho_{1,2} & 0 & \rho_2 \sqrt{1 - \rho_{1,2}^2} & \sqrt{1 - \rho_2^2} \end{pmatrix}}_{=: L} \begin{pmatrix} d\bar{W}_1(t) \\ d\widetilde{W}_1(t) \\ d\bar{W}_2(t) \\ d\widetilde{W}_2(t) \end{pmatrix}.$$

The correlation matrix is then given by

$$C = L L^\top = \begin{pmatrix} 1 & \rho_1 & \rho_{1,2} & \rho_{1,2} \rho_2 \\ \rho_1 & 1 & \rho_{1,2} \rho_1 & \rho_{1,2} \rho_1 \rho_2 \\ \rho_{1,2} & \rho_{1,2} \rho_1 & 1 & \rho_2 \\ \rho_{1,2} \rho_2 & \rho_{1,2} \rho_1 \rho_2 & \rho_2 & 1 \end{pmatrix}. \quad (14)$$

³Note that for $|\rho_{1,2}| = 1$ we have the linear dependence $W_1(t) = \rho_{1,2} W_2(t)$, and $W_2(t)$ and its transform $\bar{W}_2(t)$ drops effectively out of the system.

We see that the cross-correlation over the two distinct asset-volatility sub-models is specified by the asset-volatility correlations ρ_1 and ρ_2 of the single-asset Heston models, and the asset-asset cross-correlation $\rho_{1,2}$, i.e.

$$C^{(1,2)} = \rho_{1,2} \begin{pmatrix} 1 & \rho_2 \\ \rho_1 & \rho_1 \rho_2 \end{pmatrix}.$$

Given the single-asset parameters, the dependence structure can be expressed exclusively by a single parameter, the asset-asset cross-correlation $\rho_{1,2}$.

Next, we have to numerically find the values ρ which solve the problem

$$\min_{|\rho| \leq 1} |\mathbb{E}^Q \widehat{\rho}_{1,2}^{\text{emp}}(\rho) - \rho_{1,2}^{\text{emp}}|.$$

This can simply be done by some line search method such as bisectioning. Such a bisection procedure adapted to the given problem is stated here:

Input: Lower bound $\underline{\rho} = -1$, upper bound $\bar{\rho} = 1$,

Algorithm:

- (1) Set $\rho_{1,2}^{\text{empL}} = \underline{\rho}$ and $\rho_{1,2}^{\text{empU}} = \bar{\rho}$
- (2) Define $\underline{\mathbb{E}} := \mathbb{E}^Q(\widehat{\rho}_{1,2}^{\text{emp}}(\rho_{1,2}^{\text{empL}}))$ and $\bar{\mathbb{E}} := \mathbb{E}^Q(\widehat{\rho}_{1,2}^{\text{emp}}(\rho_{1,2}^{\text{empU}}))$
- (3) Update parameters:

$$\rho_{1,2}^{\text{empNew}} := 0.5 \cdot (\rho_{1,2}^{\text{empL}} + \rho_{1,2}^{\text{empU}})$$

$$\mathbb{E}^* := \mathbb{E}^Q(\widehat{\rho}_{1,2}^{\text{emp}}(\rho_{1,2}^{\text{empNew}})).$$

- (4) If $|\mathbb{E}^* - \rho_{1,2}^{\text{emp}}| < \epsilon \rightarrow$ **STOP** and output optimal parameter $\rho = \rho_{1,2}^{\text{empNew}}$.
 If $\mathbb{E}^* > \rho_{1,2}^{\text{emp}} \rightarrow \bar{\rho} := \rho_{1,2}^{\text{empNew}}$ and **GOTO** step 1.
 If $\mathbb{E}^* < \rho_{1,2}^{\text{emp}} \rightarrow \underline{\rho} := \rho_{1,2}^{\text{empNew}}$ and **GOTO** step 1.

Output: Optimal parameter $\rho_{1,2}^{\text{empNew}}$.

Throughout the whole procedure the calculation of $\mathbb{E}^Q \widehat{\rho}_{1,2}^{\text{emp}}(\rho)$ is needed. This can be done by simple Monte-Carlo simulation. Even if at first glance it seems computationally very exhaustive to perform Monte-Carlo simulation at every step one should observe that the $\widehat{\rho}_{1,2}^{\text{emp}}(\rho)$ is actually a ratio of time averages and therefore the ergodic theorem states that it will converge to the corresponding ratio of expected values for large T . Therefore for calculating the expectation it is enough to take a fairly small number of Monte-Carlo runs. In our tests a value of 10 already provided good results.

3.1.2 Multi-Asset Calibration

In the general multi-asset setting, the two-asset procedure can be carried out for all asset pairs. The correlation matrix has two kinds of blocks

$$C^{(i)} = \begin{pmatrix} 1 & \rho_i \\ \rho_i & 1 \end{pmatrix}, \quad \text{and} \quad C^{(i,j)} = \rho_{i,j} \begin{pmatrix} 1 & \rho_j \\ \rho_i & \rho_i \rho_j \end{pmatrix}. \quad (15)$$

Using these blocks, the whole correlation matrix can be represented. For example C in (14) can be expressed as

$$C = L L^\top = \begin{pmatrix} C^{(1)} & C^{(1,2)} \\ C^{(2,1)} & C^{(2)} \end{pmatrix}. \quad (16)$$

By this block representation, we can rearrange the single blocks, such that for any two assets i and k we obtain a matrix

$$C_{i,k} = L_{i,k} L_{i,k}^\top = \begin{pmatrix} C^{(i)} & C^{(i,k)} \\ C^{(k,i)} & C^{(k)} \end{pmatrix}. \quad (17)$$

After we have repeated the bisection procedure for all 2-dimensional problems as described in the previous section, we get a candidate solution Σ^* of (12). To validate this candidate Σ^* , we must check whether Σ^* is positive semi-definite or not. If the test succeeds, the calibration is finished. Otherwise, some regularization has to be done (see section 3.3).

3.2 Calibrating the correlations using historical time series

As already mentioned, in case we do not have liquid correlations dependent derivatives and we are not provided with the expected cross-asset empirical correlations under the pricing measure, we are left with the only option to determine the correlations historically. In this case we need to solve the optimization problem (10) involving expectations under the objective measure. We will treat only the transition to the physical measure as the rest is exactly as in the correlation adjustment algorithm described in Section 3.1, with the only difference that the expectations are taken with respect to the objective measure.

3.2.1 Algorithm

The input is the d -dimensional historical time series $(S_1(t_k), \dots, S_d(t_k))_{k=0, \dots, K}$ of the assets in interest. They are then transformed into log returns $X_i(t_k) := \log \frac{X_i(t_k)}{X_i(t_{k-1})}$ and the historical empirical correlations $\widehat{\rho}_{i,j:K}^{\text{emp}}$ are defined as in (4).

The next step is to determine the dynamics of the model under the objective measure as in the correlation adjustment algorithm we should evaluate expectations with respect to this measure. To do so we utilize Heston's assumption that the market price of volatility risk has the form $\lambda_\nu \cdot \sqrt{\nu}$ and we aim at estimating the historical mean reversion level of the volatilities of all single-asset models and the corresponding market prices of risk parameters λ_S . However, as we need the objective measure only to evaluate expectations of empirical cross-asset correlations and according to (7) the drift term of the log returns becomes negligible compared to the martingale part we can actually safely assume that $\lambda_S = 0$ or take a robust proxy for it.

In the next step we calibrate the correlation matrix Σ using the historical empirical correlation by the correlation adjustment method presented in Section 3.1. The only difference is that the simulations are being performed under the physical measure rather than the equivalent martingale measure.

Obviously the regularization of the matrix obtained by the successive two-asset calibrations in order to end up with a valid correlation matrix does not depend on the measure under which we perform the calibration and is exactly as in Section 3.3.

3.2.2 Dynamics under the physical measure

We utilize Heston's assumption in Heston [1993] that the market price of volatility risk has the form $\frac{\lambda_\nu}{\eta} \cdot \sqrt{\nu}$ (we omit the subscripts indicating the assets as we only treat the one dimensional Heston in this section). Denoting the market price of risk parameter by λ_S the dynamics under the objective measure is governed by

$$\begin{aligned} dS(t) &= [(r(t) - q(t)) + \lambda_S \nu(t)] S_t dt + \sqrt{\nu(t)} S_t dW^P(t) \\ d\nu(t) &= \underbrace{[\kappa(\bar{\nu} - \nu(t)) + \lambda_\nu \nu(t)]}_{=: \kappa^*(\bar{\nu}^* - \nu(t))} dt + \eta \sqrt{\nu(t)} [\rho dW^P(t) + \sqrt{1 - \rho^2} d\widetilde{W}^P(t)], \end{aligned} \quad (18)$$

where

$$\begin{aligned} W^P(t) &= W(t) - \lambda_S \int_0^t \sqrt{\nu(u)} du \\ \widetilde{W}^P(t) &= \widetilde{W}(t) - \frac{1}{\eta \sqrt{1 - \rho^2}} (\lambda_\nu - \rho \eta \lambda_S) \int_0^t \sqrt{\nu(u)} du \end{aligned} \quad (19)$$

are independent Brownian motions under P . The Heston parameters θ^* under P are related to the risk neutral ones θ via the following relations

$$\rho = \rho^*, \quad \eta = \eta^*, \quad \kappa = \kappa^* + \lambda_\nu, \quad \bar{\nu} = \frac{\kappa^* \bar{\nu}^*}{\kappa^* + \lambda_\nu}. \quad (20)$$

We immediately conclude that $\kappa^* \cdot \bar{\nu}^* = \kappa \cdot \bar{\nu}$ and therefore in order to obtain the dynamics of the volatility process under the objective measure it is enough to estimate the mean reversion level $\bar{\nu}^*$ under this measure.

We chose the obvious estimator based on the following well-known convergence.

Proposition 3.1. *Let $(S(t), \nu(t))$ be the processes defined in (1) and (2). Let $\pi_n = (t_{k,n})_{k=1, \dots, K_n(T)}$ be partitions of $[0, T]$, with $\text{mesh}(\pi_n) = \max_{i=1, \dots, K_n(T)} |t_{i,n} - t_{i-1,n}| \rightarrow 0$, for $n \rightarrow \infty$. Then for the log returns $X(t_k) := \log \frac{S(t_k)}{S(t_{k-1})}$*

$$\lim_{T \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{T} \sum_{k=1}^{K_n(T)} X^2(t_k) = \bar{\nu}^* \quad \text{in probability}$$

Proof. We know that almost surely

$$\sum_{k=1}^{K_n(T)} X^2(t_k) = \sum_{k=1}^{K_n(T)} (\log S(t_k) - \log S(t_{k-1}))^2 \xrightarrow[n \rightarrow \infty]{} \langle \log S \rangle_T = \int_0^T \nu(s) ds$$

and the ergodic theorem implies

$$\frac{1}{T} \int_0^T \nu(s) ds \xrightarrow{T \rightarrow \infty} \bar{\nu} \quad \text{a.s.}$$

Combining the above facts proves the statement. \square

A straightforward estimator for the mean reversion level under the physical measure is obtained by stipulating that for large T and every $i = 1, \dots, d$ we have $\sum_{k=1}^{K_n(T)} X_i^2(t_k) \approx T\bar{\nu}_i^*$, i.e. we set

$$\widehat{\nu}_i^* := \frac{1}{T} \sum_{k=1}^K X_i^2(t_k) \quad \text{for all } i = 1, \dots, d. \quad (21)$$

As we already noted we need the physical measure only for the purpose of evaluating expectations of empirical cross-asset correlations of the form

$$\widehat{\rho}_{i,j:K_n(T)}^{\text{emp}}(\Sigma) = \frac{\sum_{k=1}^{K_n(T)} X_i(t_k) X_j(t_k) - K_n(T) \bar{X}_i \bar{X}_j}{\sqrt{\sum_{k=1}^{K_n(T)} X_i^2(t_k) - K_n(T) \bar{X}_i^2} \sqrt{\sum_{k=1}^{K_n(T)} X_j^2(t_k) - K_n(T) \bar{X}_j^2}},$$

where the X_i and X_j denote the log returns in a "short" time interval, say time interval of length Δ . In the above expression the drift terms of the log returns have the magnitude of Δ^2 and are certainly dominated by the volatility terms involving $\nu(t)$ and having the magnitude of Δ . Letting $\Delta \rightarrow 0$ completely eliminates the drift terms from the empirical correlation expressions (see (7)). Therefore for our purposes in case we have data of reasonably high frequency, say daily observations, the drifts of the assets under the objective measure are of negligible importance compared to the parameters governing the dynamics of $\nu(t)$. Having all this in mind we could either assume $\lambda_S = 0$ or use an obvious proxy for it, e.g.

$$\lambda_S = \frac{1}{2} + \frac{\log \frac{S_T}{S_0} - \int_0^T (r(u) - q(u)) du}{\bar{\nu}^* \cdot T}, \quad (22)$$

where $\int_0^T (r(u) - q(u)) du$ is the accumulated interest rate minus the paid dividends associated with the corresponding asset time series. The intuition for (22) comes from the fact that

$$\mathbb{E}^P \log \frac{S_T}{S_0} = \int_0^T [(r(u) - q(u)) + (\lambda_S - \frac{1}{2}) \mathbb{E}^P \nu(u)] du.$$

We could also use different proxies, e.g. the integrated volatility proxy proposed in Aït-Sahalia and Kimmel [2007]. It is worth mentioning that the latter reference presents an efficient method for estimating the parameters of the dynamics of stochastic volatility models under the objective measure. As for our purposes it is not crucial to have very good estimator for the market price of risk λ_S for sake of simplicity we do not follow the estimation procedure of Aït Sahalia and Kimmel.

3.3 Generate a valid correlation matrix

We investigated three algorithms for transforming Σ^* into a valid correlation matrix.

3.3.1 Regularization by Jäckel [2002]

One way out of this problem is a regularization proposed by Jäckel [2002]. The following algorithm briefly demonstrates how this regularization works:

Input: Model correlation matrix Σ^* (not necessarily positive definite).

Algorithm: (1) Do an eigenvalue decomposition of $\Sigma^* \rightarrow \Sigma^* = S\Lambda S^T$, with $\Lambda = \text{diag}(\lambda_i)$. Here, $\lambda_i, i = 1, \dots, d$ represent the eigenvalues of Σ^* .

(2) Define the diagonal matrix Λ' with $\lambda'_i = \begin{cases} \lambda_i & \lambda_i \geq 0 \\ 0 & \lambda_i < 0 \end{cases}$

(3) Create the diagonal matrix T with $t_i := [\sum_m s_{im}^2 \lambda'_m]^{-1}$.

(4) Define $B := \sqrt{T}S\sqrt{\Lambda'}$.

(5) $\widehat{\Sigma}^* := BB^T$

Output: Positive definite correlation matrix $\widehat{\Sigma}^*$.

3.3.2 Regularization by Mishra [2004]

Another way out of the problem of non-positive definite correlation matrices is a regularization proposed by Mishra [2004]. The following algorithm briefly demonstrates how this regularization works:

Input: Model correlation matrix Σ^* (not necessarily positive definite).

Algorithm: (1) Do an eigenvalue decomposition of $\Sigma^* \rightarrow \Sigma^* = SAS^T$, with $\Lambda = \text{diag}(\lambda_i)$. Here, $\lambda_i, i = 1, \dots, d$ represents the eigenvalues of Σ^* .

$$(2) \text{ Define the diagonal matrix } \Lambda' \text{ with } \lambda'_i = \begin{cases} \lambda_i & \lambda_i \geq 0 \\ 0 & \lambda_i < 0 \end{cases}$$

(3) Generate d uniformly $(0, 1)$ -distributed random numbers and add them to the diagonal elements of Λ . Normalize Λ such that its trace is equal to d .

(4) By random walk methods of optimization find best possible Λ such that $\text{trace}(\Lambda) = m$ and $\widehat{\Sigma}^* = SAS^T$ has positive determinant and is positive definite and closest (wrt. maximum norm) to Σ^* .

(5) IF all diagonal entries of $\widehat{\Sigma}^*$ are approximately 1, depending on a tolerance level THEN STOP, ELSE set diagonal entries to 1 and GOTO step 1.

Output: Positive definite correlation matrix $\widehat{\Sigma}^*$.

Remark 3.2. *By construction, both algorithms are well defined, as they both deliver a positive definite correlation matrix, if they stop. Obviously, the first two steps of Jaeckel's and Mishra's algorithms are equal. In the proceeding steps, Jaeckel uses a rather deterministic way to create the positive definite correlation matrix, while Mishra uses a stochastic algorithm. Unfortunately, there is also a chance of non-convergence for the Mishra Algorithm in a case that the intermediate matrix $\widehat{\Sigma}^*$ in the intermediate step 4 is near-singular.*

3.3.3 Regularization by convex combination

If the input matrix Σ^{emp} is already positive definite, but the model matrix Σ^* is not, there is another easy way of creating a positive definite correlation matrix as a convex combination of the original input matrix and the model matrix. The following algorithm briefly demonstrates how this regularization works.

Input: Correlation matrix Σ (positive definite) and model correlation matrix Σ^* (not necessarily positive definite).

Algorithm: By standard line search algorithms find smallest $\lambda \in [0, 1]$ such that $\widehat{\Sigma}^* := \lambda\Sigma + (1 - \lambda)\Sigma^*$ is positive definite.

Output: Positive definite correlation matrix $\widehat{\Sigma}^*$

Remark 3.3. *By construction, the matrix $\widehat{\Sigma}^*$ is symmetric and positive definite and its diagonal contains only values 1, and is therefore a valid correlation matrix.*

3.3.4 Comparison of the different methods

We tested the regularization procedure of all three methods for one example. As the convex combination only works if the empirical input matrix is already a correlation matrix, we constructed the example such that this is fulfilled. We chose 10 assets with correlation

$$\rho_{ij} = \begin{cases} 0.8 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (23)$$

except for one asset pair (asset 1 and asset 8), for which we defined a correlation of 0.56. This matrix is a valid correlation matrix. The calibrated matrix Σ^* is in this setting

$$\begin{pmatrix} 100.00\% & 84.61\% & 84.61\% & 83.87\% & 84.61\% & 84.61\% & 84.61\% & 84.61\% & 61.60\% & 84.61\% \\ 84.61\% & 100.00\% & 83.13\% & 83.13\% & 83.87\% & 83.87\% & 83.13\% & 83.87\% & 83.87\% & 83.13\% \\ 84.61\% & 83.13\% & 100.00\% & 83.13\% & 83.87\% & 83.87\% & 83.87\% & 83.87\% & 83.87\% & 83.13\% \\ 83.87\% & 83.13\% & 83.13\% & 100.00\% & 83.13\% & 83.13\% & 83.13\% & 83.87\% & 83.87\% & 83.13\% \\ 84.61\% & 83.87\% & 83.87\% & 83.13\% & 100.00\% & 84.61\% & 84.61\% & 85.35\% & 85.35\% & 85.35\% \\ 84.61\% & 83.87\% & 83.87\% & 83.13\% & 84.61\% & 100.00\% & 84.61\% & 85.35\% & 84.61\% & 84.61\% \\ 84.61\% & 83.13\% & 83.87\% & 83.13\% & 84.61\% & 84.61\% & 100.00\% & 84.61\% & 84.61\% & 84.61\% \\ 84.61\% & 83.87\% & 83.87\% & 83.87\% & 85.35\% & 85.35\% & 84.61\% & 100.00\% & 84.61\% & 84.61\% \\ 61.60\% & 83.87\% & 83.87\% & 83.87\% & 85.35\% & 84.61\% & 84.61\% & 84.61\% & 100.00\% & 84.61\% \\ 84.61\% & 83.13\% & 83.13\% & 83.13\% & 85.35\% & 84.61\% & 84.61\% & 84.61\% & 84.61\% & 100.00\% \end{pmatrix}$$

This matrix has negative eigenvalues and therefore it is not positive semi-definite. After trying the different regularizations we can compute the error in terms of some matrix norm. According to different norms, the errors can be seen in Table 1. For the chosen setting, Jäckel clearly performs best. Mishra also produces good results. The convex combination method returns the worst results.

To do further investigations, we also calculated an error matrix $\Sigma_{\text{error}} := |\widehat{\Sigma}^* - \Sigma^*|$. Figures 1 to 3 graphically show the distribution of these errors in the error matrices for

Norm	Jckel	Mishra	Convex
Euclidean norm	5.66%	8.27%	37.63%
maximum norm	1.03%	1.15%	5.25%
row-sum norm	8.50%	10.00%	39.13%

Table 1: Errors of different methods

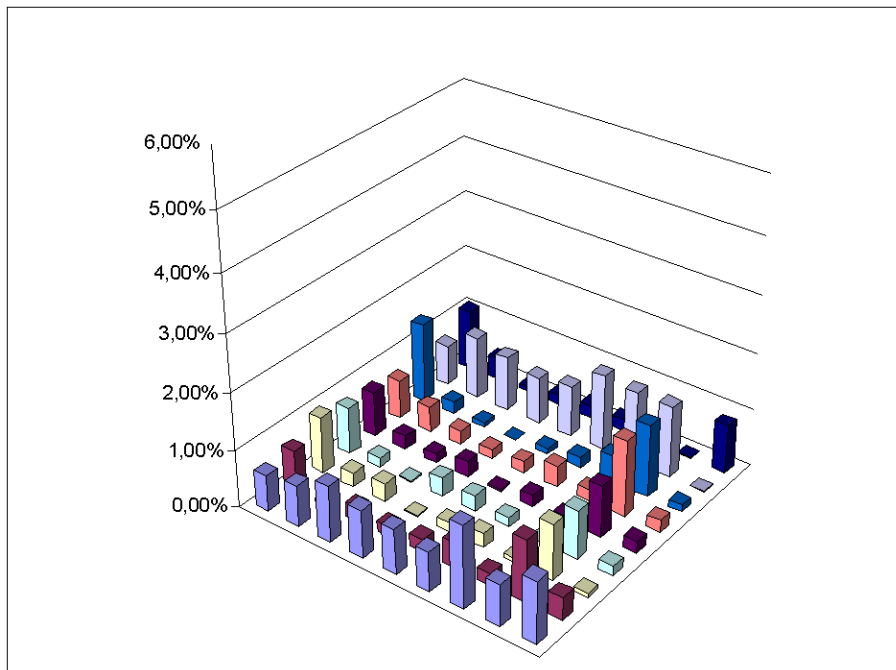


Figure 1: Jäckel regularization

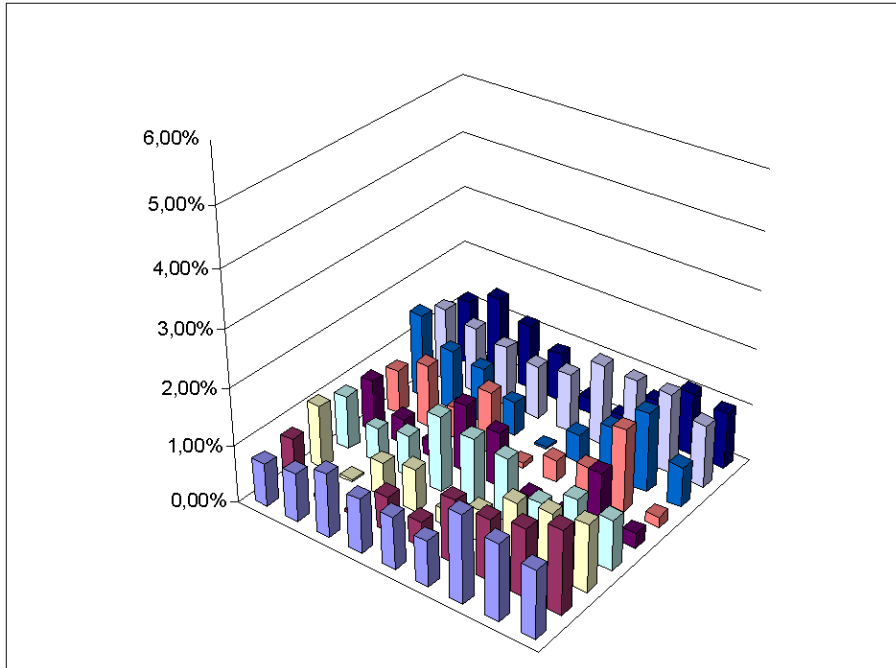


Figure 2: Mishra regularization

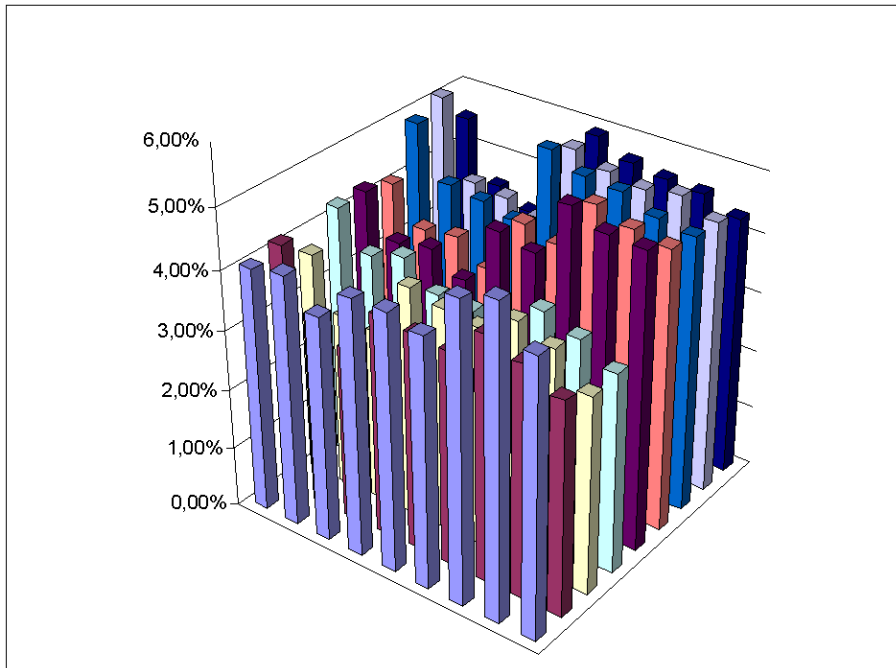


Figure 3: Convex combination regularization

the different methods. To illustrate the differences of the three methods, we applied the same scale to every graph. In figure 1 one clearly sees, that the error is concentrated vertically and horizontally from the critical value, which was the correlation of asset 1 and asset 9. Using Mishra's method, the errors in figure 2 are distributed more equally all over the error matrix, however, the maximum error is worse than for Jäckel. What we can already derive from the numbers in Table 1 gets even more obvious when we take a look at figure 3 for the case of convex combination. This method almost completely fails and there are large errors all over the whole matrix. So, in this case the regularization by Jäckel clearly does best.

3.4 Simulation Algorithm

Once calibrated, the multi-asset Heston model developed here can be simulated in a rather standard manner using the full truncated Euler method for example. We outline this method for sake of completeness. For the time step $t \mapsto t + \Delta t$, the full truncation adjustment of Lord et al. [2008] of the Euler scheme for single-asset Heston is given by

$$\begin{pmatrix} \log \widehat{S}(t + \Delta t) \\ \widehat{v}(t + \Delta t) \end{pmatrix} = \begin{pmatrix} \log \widehat{S}(t) + (r(t) - q(t)) \Delta t + \sqrt{\widehat{v}^+(t)} \Delta t \varepsilon^S \\ \widehat{v}(t) + \kappa(\theta - \widehat{v}^+(t)) \Delta t + \eta \sqrt{\widehat{v}^+(t)} \Delta t \varepsilon^\nu \end{pmatrix} \quad (24)$$

where $(\varepsilon^S, \varepsilon^\nu)$ is a bivariate standard normal random variable with zero mean and unit variances and correlation ρ . The issue of the volatility becoming negative when using the standard Euler scheme can be addressed using other truncation schemes from Lord et al. [2008] or the Andersen approach (see Andersen [2007]).

For the multi-asset Heston the correlation structure set out in Assumption 2.1 suggests the following algorithm:

- simulate a normal random vector $\varepsilon^S = (\varepsilon_1^S, \dots, \varepsilon_d^S)$ with covariance matrix Σ^S by:
 - (a) calculate the Cholesky square root L of $\Sigma^S = (\rho_{i,j})_{1 \leq i, j \leq d}$, i.e. L is a lower triangular matrix with $L L^\top = \Sigma^S$;
 - (b) generate d independent standard normally distributed random numbers z_1, \dots, z_d , and set $z = (z_1, \dots, z_d)^\top$;
 - (c) compute $\varepsilon^S = Lz$.

- simulate a normal random vector $\varepsilon^\nu = (\varepsilon_1^\nu, \dots, \varepsilon_d^\nu)$ by:
 - (a) generate d independent standard normally distributed random numbers $\tilde{z}_1, \dots, \tilde{z}_d$, and set $\tilde{z} = (\tilde{z}_1, \dots, \tilde{z}_d)^\top$;
 - (b) calculate $\varepsilon_i^\nu = \rho_i \varepsilon_i^S + \sqrt{1 - \rho_i^2} \tilde{z}_i$, for $i = 1, \dots, d$.

The algorithm follows essentially the construction implicit in Assumption 2.1. One can verify that the correlation matrix of $(\varepsilon^S, \varepsilon^\nu)$ corresponds to the one given in Proposition 2.2.

4 Conclusion

In this paper we presented a parsimonious multi-asset Heston model. All single-asset sub-models follow the well-known Heston dynamics and their parameters are calibrated as usual on implied market volatilities. Our focus was on the calibration of the correlation structure between the single-asset marginals in the absence of sufficient liquid cross-asset option price data. We presented two general calibration setups corresponding to relevant practical situations: (1) when the empirical cross-asset correlations in the risk neutral world are given by the user and we need to calibrate the correlations between the driving Brownian motions or (2) when they have to be estimated from the historical time series. The theoretical background, including the ergodicity of the multidimensional CIR process, for the proposed estimators are worked out.

The presented model is parsimonious in the sense that $d(d-1)/2$ asset-asset cross-correlations are required for a d -asset Heston model. This correlation matrix can be used to simulate multi-asset dynamics using the Heston model.

We also tested three methods for obtaining a valid correlation matrix, in case the calibrated cross-asset correlations do not form such a matrix.

A Ergodicity of the two dimensional CIR process

Theorem A.1. *Let $\nu := (\nu_1, \nu_2)$ be a two dimensional CIR process with parameters $\kappa_1, \kappa_2, \eta_1, \eta_2, \bar{\nu}_1, \bar{\nu}_2 > 0$ and $\rho \in (-1, 1)$, i.e. it is a continuous Markov process with*

infinitesimal generator

$$\mathcal{A} = \kappa_1(\bar{\nu}_1 - x)\partial_x + \kappa_2(\bar{\nu}_2 - y)\partial_y + \frac{1}{2}\eta_1^2 x\partial_{xx} + \frac{1}{2}\eta_2^2 y\partial_{yy} + \rho\eta_1\eta_2\sqrt{xy}\partial_{xy}.$$

ν has a stationary probability distribution $\mu_\infty(dx dy)$ on \mathbb{R}_+^2 , and moreover, for every μ -integrable $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ we have for μ -almost all (x, y)

$$\mathbb{P}_{(x,y)} \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(\nu_1(t), \nu_2(t)) dt = \int f(u, v) \mu(du dv) \right] = 1.$$

Remark A.2. We actually need the above result for the function $f(x, y) = \sqrt{xy}$. The fact that f is μ -integrable is a simple consequence of the Hölder inequality. Indeed

$$\int \sqrt{xy} \mu(dx dy) \leq \sqrt{\int x \mu(dx dy) \int y \mu(dx dy)} = \sqrt{\bar{\nu}_1 \cdot \bar{\nu}_2}.$$

Proof. The statement of the theorem is a special case of Theorem 5.1, Chapter IV of Has'minskii. However, we need to verify that the following condition holds: there is a bounded domain D in \mathbb{R}_+^2 having the following properties:

- (1) The smallest eigenvalue of the diffusion matrix of the generator \mathcal{A} is bounded away from zero on D .
- (2) For every $(x, y) \in \mathbb{R}_+^2 \setminus D$ the mean time τ at which a trajectory $(\nu_1^x(t), \nu_2^y(t))$ started at (x, y) reaches U is finite, and even $\sup_{(x,y) \in K} \tau < \infty$ for every compact set $K \subset \mathbb{R}_+^2 \setminus D$.

We first present a rather concise solution in the case where both one dimensional CIR processes do not reach the origin (i.e. the stability condition $2\kappa_i \bar{\nu}_i \geq \eta_i^2$ is satisfied). In this case the state space is $E := (0, \infty)^2$ as $P(\nu_i(t) > 0 \forall t) = 1$, $i = 1, 2$.

The first condition will obviously be satisfied if the closure of D is contained in E . Indeed, the eigenvalues are given by

$$\lambda_{1,2} = \frac{\bar{\eta}_1^2 x + \bar{\eta}_2^2 y \pm \sqrt{(\bar{\eta}_1^2 x + \bar{\eta}_2^2 y)^2 + 4(\rho^2 - 1)\bar{\eta}_1^2 \bar{\eta}_2^2 xy}}{2},$$

which can be uniformly bounded away from zero on any compact set contained in E as long as $\rho^2 < 1$.

The second condition per definition actually states that the process η has to be D -recurrent (see Chapter III, Section 7 in Has'minskii [1980]). According to Theorem 7.1 in Chapter III in the same source, a sufficient condition for the D -recurrence is the existence of a nonnegative (Lyapunov) function $V(s, x)$ on $R_+ \times D^C$, twice continuously differentiable in the space variable and continuously differentiable in the time variable and such that

$$\mathcal{A}V \leq -\delta \text{ for some positive } \delta > 0 \text{ on } R_+ \times D^C.$$

As ν is a time homogeneous diffusion we consider Lyapunov functions depending only on the state variable. We take

$$V(x, y) = ax \ln x + by \ln y + c : E \rightarrow \mathbb{R},$$

where $a, b, c > 0$. V is a C^2 function on E . Since the mapping $x \mapsto x \ln x$ is bounded from below on E for every set of parameters $a, b > 0$ we can achieve $V(x, y) > 0$ on E by choosing c big enough. Now observe that

$$\begin{aligned} \mathcal{A}V(x, y) &= \kappa_1(\bar{\nu}_1 - x)a(1 + \ln x) + \kappa_2(\bar{\nu}_2 - y)b(1 + \ln y) + \frac{1}{2}(\eta_1^2 a + \eta_2^2 b) \\ &= -\kappa_1 a(x + x \ln x) + \kappa_1 a \bar{\nu}_1 \ln x - \kappa_2 b(y + y \ln y) + \kappa_2 b \bar{\nu}_2 \ln y \\ &\quad + \frac{1}{2}(\eta_1^2 a + \eta_2^2 b) + \kappa_1 a \bar{\nu}_1 + \kappa_2 b \bar{\nu}_2. \end{aligned} \tag{25}$$

Since of course $x \ln x$ and $y \ln y$ outperform the other terms for large x and y , the set M on which $\mathcal{A}V$ is bigger than say -1 is a bounded subset of E . Furthermore, letting $x \rightarrow 0$ we have $\mathcal{A}V(x, y) \rightarrow -\infty$ for all $y > 0$ and analogously for $y \rightarrow 0$ we get $\mathcal{A}V(x, y) \rightarrow -\infty$ for arbitrary $x > 0$. Thus we conclude that the closure of M is still contained in E . Indeed, assume e.g. $(x_0, 0)$ is in the closure of M . Then we would have a sequence $M \ni (x_n, y_n) \rightarrow (x_0, 0)$ and we would get $-\infty = \lim_{n \rightarrow \infty} \mathcal{A}V(x_n, y_n) \geq -1$. Summarizing, M is a bounded subset of E , such that the closure of M is also in E and we have that $\mathcal{A}V \leq -1$ on M^C . The nonnegativity of V on the whole E is achieved by setting $c > 0$ big enough. This verifies the conditions (1) and (2) stated in the beginning of the proof. This proves the statement theorem in case the stability condition $P(\nu_i(t) > 0 \forall t) = 1, i = 1, 2$ is satisfied.

We will now sketch the proof in the general case. We need to set $E = [0, \infty)^2$ as the one dimensional CIR process may reach the zero boundary. For small enough $\epsilon > 0$ define function $g^\epsilon : [0, \infty)$ to be the quadratic truncation of $x \ln x$, i.e.

$$g(x) = \begin{cases} x \ln x & \text{if } x > \epsilon \\ a_\epsilon x^2 + b_\epsilon x + c_\epsilon & x \in [0, \epsilon] \end{cases},$$

where the coefficients a_ϵ , b_ϵ and c_ϵ are chosen in such a manner that the function $g(x)$ is C^2 . We now consider the Lyapunov function

$$V^\epsilon(x, y) = ag^\epsilon(x) + bg^\epsilon(y) + c : E \rightarrow \mathbb{R}$$

with $a, b, c > 0$. The nonnegativity can be achieved as above by choosing c large enough. An analogous leading term argument implies that the set M on which $\mathcal{A}V^\epsilon(x, y)$ is bigger than -1 is bounded. A rather technical calculation shows that by choosing ϵ small enough we can make $\mathcal{A}V^\epsilon(x, y) \leq -2$ on the boundary of E . Clearly, there is no surprise as we approximate the case, which we have worked out above. This, however, proves that actually the closure of M is contained in the interior of E , as otherwise we would have for example a sequence $M \ni (x_n, y_n) \rightarrow (x_0, 0)$ and we would get $-2 \geq \lim_{n \rightarrow \infty} \mathcal{A}V^\epsilon(x_n, y_n) \geq -1$. Therefore the set M satisfies the requirements of the conditions (1) and (2) in the beginning of the proof and we are done.

An alternative way to prove this would be to take g^ϵ to be the linear rather than quadratic truncation of $x \ln x$. This would actually also formally suffice even if the resulting V^ϵ is only a C^1 rather than C^2 function. The reason that it still works is that the proof of Theorem 8.1 in Chapter III in Has'minskii which we invoke relies on Itô's formula and it still holds for C^1 functions with existing and continuous second derivatives everywhere but on a countable exception set and having bounded second derivatives around the singularities.

□

References

- Yacine Aït-Sahalia and Robert Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83:413–452, 2007.
- L.B.G. Andersen. Efficient simulation of the heston stochastic volatility model. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=946405, 2007.
- G. Bakshi, C. Cao, and Z. Chen. Pricing and hedging long-term options. *Journal of Econometrics*, 94(1–2):277–318, 2000.
- D.S. Bates. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *Review of financial studies*, 1(1):69–107, 1996.
- R. Cont and P. Tankov. *Financial modelling with jump processes (2nd Ed.)*. Chapman & Hall, 2008.
- J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180, 2007.
- J. Da Fonseca, M. Grasselli, and C. Tebaldi. A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604, 2008.
- J.P. Fouque, G. Papanicolaou, and K.R. Sircar. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, 2000.
- J. Gatheral. *The volatility surface: a practitioner’s guide*. Wiley Finance, 2006.
- C. Gouriéroux. Continuous time Wishart process for stochastic risk. *Econometric Reviews*, 25(2):177–217, 2006.
- P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward. Managing smile risk. *Wilmott*, September:84–108, 2002.
- R.Z. Has’minskii. *Stochastic stability of differential equations, vol. 7 of Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.

- S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, pages 327–343, 1993.
- P. Jäckel. *Monte Carlo methods in finance*. John Wiley & Sons, Ltd. West Sussex, 2002.
- C. Kahl and M. Günther. Complete the correlation matrix. <http://www.math.uni-wuppertal.de/~kahl/publi.html>, 2005.
- R. Lord, R. Koekkoek, and D.J.C. Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Journal of Quantitative Finance*, 2008.
- S. Mikhailov and U. Nögel. Heston’s stochastic volatility model: implementation, calibration, and some extensions. *Wilmott*, July:75–79, 2003.
- S.K. Mishra. Optimal solution of the nearest correlation matrix problem by minimization of the maximum norm. <http://mpa.ub.uni-muenchen.de/1783/>, 2004.