Markovian projection to a Displaced Volatility Heston model

A. Antonov, M. Arneguy and N. Audet *

March 12, 2008

Abstract

Markovian Projection is an optimal approximation of a complex underlying process with a simpler one, keeping essential properties of the initial process. The Heston process, as the Markovian Projection target, is an example [4]. In this article, we generalize the results of Markovian Projection onto a Heston model to a wider class of approximating models, a Heston model with displaced volatility. As an important application, we derive an effective approximation for FX/EQ options for the Heston model, coupled with correlated Gaussian interest rates. The main technical result is an option evaluation for correlated Heston/Lognormal processes. Unlike the case of exactly solvable (affine) zero correlation or its uncorrelated displacement generalization, considered by Andreasen [2], non-trivial correlations destroy affine structure and exact solvability. Using the powerful technique of Markovian Projection onto a Heston model with displaced volatility, we produce an effective approximation and present its numerical confirmation.

1 Introduction

The growing sophistication and interpenetration of financial markets requires complex hybrid models dealing with numerous underlyings. Equity baskets in possibly different currencies, and in the presence of non-trivial interest rates, appear more and more frequently in financial institutions. Stochastic volatility models having observable dynamics, and handling implied volatility skew and smile, are standard participants of such hybrids. An FX-rate or equity following a Heston model, with Gaussian interest rates, is one such example. Modern computational facilities permit hybrid pricing using Monte Carlo methods, but the model calibration still requires effective analytical approximations.

Recently, a powerful method of Markovian Projection (MP) was introduced in the financial mathematics by Piterbarg [9]. Different challenging calibration problems, such as Cross-Currency Gaussian and LIBOR Market Model equipped with a CEV exchange rate [9] and [3], Heston Basket Index options [4], and others, were solved with the help of this method. In this article, we generalize the technique of MP to a Heston model, and illustrate the theory with an important example of the calibration of FX-options for a cross-currency Heston model, correlated with Gaussian interest rates.

The MP theory is based on the Gyöngy result [6]. Its early applications in financial mathematics were proposed by Dupire [5]. For European option pricing, it is often possible to come up with a single underlying process, S, usually complex and non-Markovian, but with an explicitly known SDE. The option

*Numerix LLC 150 East 42nd Street, 15th Floor, New York, NY 10017 www.numerix.com

price with a strike K can be expressed in a standard manner, $\mathbb{E}[(S(t)-K)^+]$. Suppose the underlying process, in its martingale measure, satisfies the following general SDE,

$$dS(t) = \lambda(t) \cdot dW(t),$$

for some fixed initial value. Here and below, $W(t)$ is an F -component vector of independent Brownian motions, $\lambda(t)$ is an F-component stochastic volatility process, and the dot denotes a scalar product. According to Gyöngy, under some technical assumptions, there exists a local-volatility model,

$$dS^*(t) = \lambda^*(t, S^*(t)) \cdot dW(t), \quad S^*(0) = S(0),$$

such that all one-dimensional marginal distributions of $S(t)$ and $S^*(t)$ coincide. To ensure this match, the effective local volatility should satisfy a condition,

$$|\lambda^*(t, S^*)|^2 = \mathbb{E}[|\lambda(t)|^2 | S(t) = S^*].$$

Non-trivial calculus of the above conditional expectation can be avoided by postulating a simpler form of the local volatility. In [9] and [3], the authors considered a linear dependence on the underlying, leading to a Displaced-Diffusion model,

$$\lambda^*(t, S^*) = (1 + b(t)\Delta S^*(t))\, s(t),$$

where a difference between current underlying value and its initial value was denoted as $\Delta S^*(t) = S^*(t) - S^*(0)$. Thus, reducing the local volatility function space to a linear one, it is possible to calculate an optimal shift $b(t)$ and volatility $s(t)$. The mimicking displaced-diffusion model S^* gives explicit formulas for European options.

Unlike the full-space local volatility model, an effective displaced diffusion cannot reproduce one-dimensional marginals of the original process exactly. Furthermore, while a displaced diffusion is sufficient to recover the skewness of implied volatilities, it is obviously inappropriate if the initial model also has a smile.

In order to incorporate a smile to the mimicking model S^*, one can enlarge a space of basis functions from linear to U-shaped ones, or choose a stochastic volatility model. In [4], the authors proposed the second way whereby the underlying is mimicked with a shifted Heston process with time-dependent coefficients,

$$\begin{aligned} dS^*(t) &= (1 + \beta(t)\, \Delta S^*(t)) \sqrt{z(t)}\, \sigma_H(t) \cdot dW(t), \\ dz(t) &= \alpha(t)\, (1 - z(t))\, dt + \sqrt{z(t)}\, \sigma_z(t) \cdot dW(t), \quad z(0) = 1, \end{aligned}$$

where $\beta(t)$ is a shift parameter controlling the model skew, $\sigma_H(t)$ is a vector volatility, and $z(t)$ is a stochastic volatility or, more precisely, a stochastic *multiplier* with initial unit value. The process $z(t)$ is defined by its mean-reversion $\alpha(t)$ and volatility vector $\sigma_z(t)$. This Heston model form, proposed by [1] in the context of interest rates, is equivalent to that initially introduced by Heston [7]. We prefer to work with the Andersen-Andreasen setup above, and give a map to the initial Heston form in Appendix A, which also contains option pricing details.

Markovian Projection to Heston models can approximate both skew and smile properties of implied volatilities. An important application to Heston basket index options was considered in [4].

In this paper, we enlarge the mimicking model space to affine bridges between the Heston model and the displaced diffusion. Having more degrees of freedom, the bridge model can deliver a more accurate approximation. The new MP target has the following SDE,

$$\begin{aligned} dS^*(t) &= (1+\beta(t)\,\Delta S^*(t))\left(\sqrt{z(t)}\,\sigma_H(t)+\sigma_D(t)\right)\cdot dW(t), \\ dz(t) &= \alpha(t)\,(1-z(t))\,dt+\sqrt{z(t)}\,\sigma_z(t)\cdot dW(t), \quad z(0)=1. \end{aligned}$$

A new vector parameter $\sigma_D(t)$, volatility displacement, should be perpendicular to both vector $\sigma_H(t)$ and vector $\sigma_z(t)$, which guarantees affine properties and exact solvability. Below, we will refer to the bridge model as a displaced volatility Heston model (Heston DV).

The model first appeared in [2] in the context of an FX-rate with stochastic Gaussian interest rates. A correlation between interest rates and FX was organized via the volatility displacement. The Heston driving factors were uncorrelated with the interest rates to preserve affine properties and FX-option solvability.

As an important application of MP to the Heston DV model, we consider FX-option pricing of the Heston model, coupled with correlated Gaussian interest rates. Unlike Andreasen [2], we consider a *standard* Heston model for the FX-rate, *arbitrarily correlated* with the interest rates. The exact FX-option pricing is impossible in this case, but the MP can be elaborated to attain excellent accuracy.

This paper is organized as follows. In Section 2, we recall the MP key formulas in the case of a target Heston model. We present new results in MP to Heston models with displaced volatility in Section 3. Section 4 is devoted to FX-option pricing approximations for Gaussian interest rates, with numerical results in Section 5.

2 Projection to the Heston model

In this section, we recall key facts of Markovian Projection to the Heston model. An effective approximation of a general non-Markovian process with skew and smile,

$$dS(t)=\lambda(t)\cdot dW(t), \tag{1}$$

was derived in [4]. The mimicking process was a Heston model,

$$\begin{aligned} dS^*(t) &= (1+\beta(t)\,\Delta S^*(t))\sqrt{z(t)}\,\sigma_H(t)\cdot dW(t), \quad S^*(0)=S(0), \\ dz(t) &= \theta(t)\,(1-z(t))\,dt+\sqrt{z(t)}\,\sigma_z(t)\cdot dW(t), \quad z(0)=1, \end{aligned} \tag{2}$$

with coefficients calculated by Markovian projection. Here and below, we denote $\Delta S^*(t)=S^*(t)-S^*(0)$. Function $\beta(t)$ is a shift parameter controlling the model skew, $\sigma_H(t)$ is the underlying vector volatility, and $z(t)$ is a stochastic volatility factor with initial unit value. The process $z(t)$ is defined by its mean-reversion $\alpha(t)$ and volatility vector $\sigma_z(t)$. Below, we always suppose that $S^*(0)=S(0)$ and $z(0)=1$, and omit these expressions from the formulas for better legibility. In Appendix A, one can find option pricing details for model (2).

The MP to the Heston model is often done in two steps. First, we estimate the shift parameter $\beta(t)$. Several recommendations on its calculation are proposed in [4], although, in certain cases, one can simply set the shift according to a form of the SDE (1). Having shift parameter $\beta(t)$ fixed, we calculate the effective Heston parameters. Define a (displaced-diffusion) volatility $\Lambda(t)$ and variance $V(t)$ as

$$\Lambda(t) = \frac{\lambda(t)}{1 + \beta(t)\,\Delta S(t)} \tag{3}$$

and

$$V(t) = |\Lambda(t)|^2 = \frac{|\lambda(t)|^2}{(1 + \beta(t)\,\Delta S(t))^2}. \tag{4}$$

Suppose that an SDE for the process $V(t)$ is known,

$$\begin{aligned} dS(t) &= (1 + \Delta S(t)\beta(t))\,\Lambda(t)\cdot dW(t), && (5)\\ dV(t) &= \mu_V(t)\,dt + \sigma_V(t)\cdot dW(t), && (6) \end{aligned}$$

for some stochastic drift $\mu_V(t)$ and volatility $\sigma_V(t)$. Then, the effective Heston parameters can be expressed as averages of underlying processes,

$$\begin{aligned} |\sigma_H(t)|^2 &= \mathbb{E}[V(t)], && (7)\\ \theta(t) &= (\ln \mathbb{E}[V(t)])' - \frac{1}{2}(\ln \mathrm{Var}[V(t)])' + \frac{\mathbb{E}[|\sigma_V(t)|^2]}{2\,\mathrm{Var}[V(t)]}, && (8)\\ |\sigma_z(t)|^2 &= \frac{\mathbb{E}[V(t)|\sigma_V(t)|^2]}{\mathbb{E}[V^2(t)]\mathbb{E}[V(t)]}, && (9)\\ \rho(t) &= \frac{\mathbb{E}[V(t)\Lambda(t)\cdot\sigma_V(t)]}{\sqrt{\mathbb{E}[V^2(t)]\mathbb{E}[V(t)|\sigma_V(t)|^2]}}. && (10) \end{aligned}$$

Although the expectations in the above formulas look complicated, in practice, one can come up with their suitable approximations, which often require less precision than standard moment calculations. On the other hand, it is important to preserve general properties like skews, smiles, underlying supports, etc. For example, if the variance process $V(t)$ can have an appreciable positive floor, its MP to the Heston model would likely give a poor quality, as far as the effective variance process $z(t)\,|\sigma_H(t)|^2$ spans from zero to infinity. In the next section, we will elaborate on the target process having a positive floor for stochastic variance.

3 Markovian projection to a Heston model with displaced volatility

Sometimes the approximation quality given by MP to the Heston model is not sufficient due to a floor on the stochastic variance $V(t)$. In such cases, we take a wider model, still conserving its affine properties, a *Shifted Heston model with displaced volatility* (Heston DV),

$$\begin{aligned} dS^*(t) &= (1 + \beta(t)\,\Delta S^*(t))(\sqrt{z(t)}\,\sigma_H(t) + \sigma_D(t))\cdot dW(t),\\ dz(t) &= \theta(t)\,(1 - z(t))\,dt + \sqrt{z(t)}\,\sigma_z(t)\cdot dW(t), && (11) \end{aligned}$$

where the volatility displacement vector $\sigma_D(t)$ is perpendicular to both rate volatility vector $\sigma_H(t)$ and volatility of volatility vector $\sigma_z(t)$. The Heston DV calibration procedure is similar to that of the Heston model, its analytical option pricing details can be found in Appendix B.

As in cases of a Heston model as the projection target, we use the 2D version of Gyöngy's lemma stating that the initial process pair $\{S(t), V(t)\}$,

$$\begin{aligned} dS(t) &= (1 + \beta(t)\,\Delta S(t))\;\Lambda(t) \cdot dW(t), \\ dV(t) &= \mu_V(t)\,dt + \sigma_V(t) \cdot dW(t), \end{aligned}$$

can be mimicked by a Markovian pair $\{S^*(t), V^*(t)\}$,

$$\begin{aligned} dS^*(t) &= (1 + \beta(t)\,\Delta S^*(t))\;\sigma^*_S(t; S^*(t), V^*(t))\;\cdot dW(t), \\ dV^*(t) &= \mu^*_V(t; S^*(t), V^*(t))\,dt + \sigma^*_V(t; S^*(t), V^*(t))\;\cdot dW(t), \end{aligned}$$

provided that the effective drift, μ^*_V, and volatilities, σ^*_S and σ^*_V, are calculated using the initial process conditional expectations,

$$\mu^*_V(t; s, v) = \mathbb{E}[\mu_V(t) \mid S(t) = s, V(t) = v], \tag{12}$$

$$|\sigma^*_S(t; s, v)|^2 = \mathbb{E}[|\Lambda(t)|^2 \mid S(t) = s, V(t) = v] = v, \tag{13}$$

$$|\sigma^*_V(t; s, v)|^2 = \mathbb{E}[|\sigma_V(t)|^2 \mid S(t) = s, V(t) = v], \tag{14}$$

$$\sigma^*_S(t; s, v)\cdot\sigma^*_V(t; s, v) = \mathbb{E}[\Lambda(t)\cdot\sigma_V(t) \mid S(t) = s, V(t) = v]. \tag{15}$$

We look for a mimicking pair $\{S^*(t), V^*(t)\}$ related to the Heston DV model (11), after identifying variance

$$V^*(t) = z(t)\,|\sigma_H(t)|^2 + V_D(t),$$

where we have denoted the displaced variance as $V_D(t) = |\sigma_D(t)|^2$. This leads to the following SDEs for the pair $\{S^*(t), V^*(t)\}$,

$$\begin{aligned} dS^*(t) &= (1 + \beta(t)\,\Delta S^*(t)) \left(\frac{\sqrt{V^*(t) - V_D(t)}}{|\sigma_H(t)|}\,\sigma_H(t) + \sigma_D(t) \right) \cdot dW(t), \\ dV^*(t) &= \Big((V^*(t) - V_D(t)) \left(\left(\ln|\sigma_H(t)|^2\right)' - \theta(t) \right) + V'_D(t) + \theta(t)\,|\sigma_H(t)|^2 \Big)\,dt \\ &+ \;|\sigma_H(t)|\,\sqrt{V^*(t) - V_D(t)}\,\sigma_z(t)\cdot dW(t), \end{aligned}$$

where $\Delta S^*(t) = S^*(t) - S^*(0)$, and prime $'$ denotes the derivative of a deterministic function of time. Thus, the Heston DV model specifies an *ansatz* for the conditional expectations (12)–(15) given by[1]

$$\begin{aligned} \mu^*_V(t; s, v) &= (v - V_D(t)) \left(\left(\ln|\sigma_H(t)|^2\right)' - \theta(t) \right) + V'_D(t) + \theta(t)\,|\sigma_H(t)|^2, \\ |\sigma^*_V(t; s, v)|^2 &= (v - V_D(t))\;|\sigma_H(t)|^2\,|\sigma_z(t)|^2, \\ \sigma^*_S(t; s, v)\cdot\sigma^*_V(t; s, v) &= (v - V_D(t))\;\sigma_z(t)\cdot\sigma_H(t). \end{aligned}$$

[1]Eq. (13) is trivially satisfied by the choice of the volatility process and need not be considered.

Due to the L^2-measure minimizing property of the conditional expectation, our task reduces to a joint minimization of the following criteria for all times t,

$$\chi_1^2(t) = \mathbb{E}\left[\left(\mu_V(t) - \left((V(t)-V_D(t))\left(\left(\ln|\sigma_H(t)|^2\right)' - \theta(t)\right) + V_D'(t) + \theta(t)\,|\sigma_H(t)|^2\right)\right)^2\right], \quad (16)$$

$$\chi_2^2(t) = \mathbb{E}\left[\left(|\sigma_V(t)|^2 - |\sigma_H(t)|^2\,|\sigma_z(t)|^2\,(V(t)-V_D(t))\right)^2\right], \quad (17)$$

$$\chi_3^2(t) = \mathbb{E}\left[(\Lambda(t)\cdot\sigma_V(t) - \sigma_z(t)\cdot\sigma_H(t)(V(t)-V_D(t)))^2\right]. \quad (18)$$

For any fixed variance displacement $V_D(t)$, we have the following solution for the unknown functions $|\sigma_H(t)|$, $|\sigma_z(t)|$, $\theta(t)$, and $\rho(t)$,

$$|\sigma_H(t)|^2 = \mathbb{E}[V(t)] - V_D(t), \quad (19)$$

$$\theta(t) = \left(\ln|\sigma_H(t)|^2\right)' - \frac{1}{2}\left(\ln \mathrm{Var}[V(t)]\right)' + \frac{\mathbb{E}[|\sigma_V(t)|^2]}{2\,\mathrm{Var}[V(t)]}, \quad (20)$$

$$|\sigma_z(t)|^2 = \frac{\mathbb{E}[(V(t)-V_D(t))|\sigma_V(t)|^2]}{\mathbb{E}[(V(t)-V_D(t))^2]\mathbb{E}[V(t)-V_D(t)]}, \quad (21)$$

$$\rho(t) = \frac{\mathbb{E}[(V(t)-V_D(t))\Lambda(t)\cdot\sigma_V(t)]}{\sqrt{\mathbb{E}[(V(t)-V_D(t))^2]\mathbb{E}[(V(t)-V_D(t))|\sigma_V(t)|^2]}}, \quad (22)$$

where $\rho(t)$ is the effective Heston correlation, $\rho(t) = \frac{\sigma_H(t)\cdot\sigma_z(t)}{|\sigma_H(t)|\,|\sigma_z(t)|}$.

As shown in Appendix C, the first χ_1^2 criterion minimization gives the optimal volatility module $|\sigma_H(t)|$ (19) and the mean-reversion $\theta(t)$ (20) as a function of underlying averages and a displacement $V_D(t)$. The other optimal parameters are calculated to minimize the criteria χ_2^2 and χ_3^2. More precisely, a product of squares of volatilities, $M_1(t) = |\sigma_H(t)|^2\,|\sigma_z(t)|^2$, and their dot-product, $M_2(t) = \sigma_z(t)\cdot\sigma_H(t)$, can be obtained by minimizing χ_2^2 and χ_3^2, for fixed variance displacement $V_D(t)$. Namely, setting a derivative $\frac{\partial\chi_2^2}{\partial M_1}$ to zero, one has

$$M_1(t) = |\sigma_H(t)|^2\,|\sigma_z(t)|^2 = \frac{\mathbb{E}[(V(t)-V_D(t))|\sigma_V(t)|^2]}{\mathbb{E}[(V(t)-V_D(t))^2]}. \quad (23)$$

Similarly, optimal condition $\frac{\partial\chi_3^2}{\partial M_2} = 0$ leads to

$$M_2(t) = \sigma_z(t)\cdot\sigma_H(t) = \frac{\mathbb{E}[(V(t)-V_D(t))\Lambda(t)\cdot\sigma_V(t)]}{\mathbb{E}[(V(t)-V_D(t))^2]}. \quad (24)$$

Then, optimal volatility of volatility module $|\sigma_z(t)|$ (21) and correlation $\rho(t)$ (22) can be restored from the already-calculated volatility $|\sigma_H(t)|$ (19), the product of squares of volatilities $|\sigma_H(t)|^2\,|\sigma_z(t)|^2$, and volatilities dot-product $\sigma_z(t)\cdot\sigma_H(t) = |\sigma_H(t)|\,|\sigma_z(t)|\,\rho(t)$.

Now, let us evaluate the optimal variance displacement which was, up to now, a fixed parameter. Substituting the optimal multipliers M_1 and M_2 into criteria χ_2^2 and χ_3^2, respectively, we derive the optimization defects as functions of the displacement,

$$D_2(t) = \mathbb{E}\left[|\sigma_V(t)|^4\right] - \frac{\left(\mathbb{E}[(V(t)-V_D(t))|\sigma_V(t)|^2]\right)^2}{\mathbb{E}[(V(t)-V_D(t))^2]}, \tag{25}$$

$$D_3(t) = \mathbb{E}\left[(\Lambda(t)\cdot\sigma_V(t))^2\right] - \frac{\left(\mathbb{E}[(V(t)-V_D(t))\Lambda(t)\cdot\sigma_V(t)]\right)^2}{\mathbb{E}[(V(t)-V_D(t))^2]}. \tag{26}$$

Minimizing the defects leads, in general, to different values of the displacement $V_D(t)$. In other words, optimal displacement for $D_2(t)$ is not necessarily optimal for $D_3(t)$, and vice-versa. Thus, we should reconcile both defects' optimizations. This can be done in practical geometrical terms. Namely, we are looking for the displacement $V_D(t)$ such that a "vector" $V(t) - V_D(t)$ remains as "parallel" as possible to two "vectors" $|\sigma_V(t)|^2$ and $\Lambda(t)\cdot\sigma_V(t)$. Unknown coefficients $|\sigma_H(t)|^2\,|\sigma_z(t)|^2$ and $\sigma_z(t)\cdot\sigma_H(t)$ can be considered as optimum scaling of $V(t) - V_D(t)$, minimizing distance with the "vector" $|\sigma_V(t)|^2$ and the "vector" $\Lambda(t)\cdot\sigma_V(t)$, respectively.

In order to optimize the variance displacement, we should find a "main direction," a special linear combination of target "vectors" $|\sigma_V(t)|^2$ and $\Lambda(t)\cdot\sigma_V(t)$. Namely, we construct their "dot product" matrix

$$\begin{pmatrix} \mathbb{E}[|\sigma_V(t)|^4] & \mathbb{E}[|\sigma_V(t)|^2\,\Lambda(t)\cdot\sigma_V(t)] \\ \mathbb{E}[|\sigma_V(t)|^2\,\Lambda(t)\cdot\sigma_V(t)] & \mathbb{E}[(\Lambda(t)\cdot\sigma_V)^2] \end{pmatrix}$$

and compute eigenvector $\{e_1(t), e_2(t)\}$ corresponding to the biggest eigenvalue. The main direction vector

$$X_D(t) = e_1(t)\,|\sigma_V(t)|^2 + e_2(t)\,\Lambda(t)\cdot\sigma_V(t)$$

will form a new criterion for optimal volatility displacement,

$$\chi_D^2(t) = \mathbb{E}\left[\left(X_D(t) - M(t)(V(t) - V_D(t))\right)^2\right],$$

for a multiplier $M(t)$. Minimization of $\chi_D^2(t)$ over the multiplier $M(t)$ and $V_D(t)$ leads to the optimal volatility displacement,

$$V_D(t) = \mathbb{E}[V(t)] - \frac{\mathbb{E}[X_D(t)]\mathrm{Var}[V(t)]}{\mathbb{E}[X_D(t)V(t)] - \mathbb{E}[X_D(t)]\mathbb{E}[V(t)]}. \tag{27}$$

4 FX/EQ options in the presence of Gaussian correlated rates

In this section, we consider an important example of hybrid FX/EQ models following the Heston process, coupled with Gaussian interest rates. For non-zero correlations between the Heston evolution and the interest rates, a standard approach to FX-option analytical pricing fails. Indeed, the zero correlations preserve the affine model structure and make possible analytical calculation of the characteristic function. Thus, for non-zero correlation cases when the standard affine technique does not work, one should provide an approximation. Below, we will apply the elaborated MP formalism to the Heston DV model.

Consider the Hull-White model for domestic and foreign interest rates. Take, for simplicity, the one-factor case[2] for domestic and foreign short rates (denoted $r_1(t)$ and $r_2(t)$). In the respective risk-neutral measures, the Hull-White SDEs look like

$$dr_i(t) = (\zeta_i(t) - r_i(t)\,a_i(t))\,dt + \sigma_i(t)\cdot dW_i(t), \tag{28}$$

[2]Multi-factor cases can be treated in the same manner.

where subscript index $i = 1, 2$ corresponds to a market number (domestic/foreign) and parameters a_i and σ_i are time-dependent mean-reversion and volatility. Parameter ζ_i is the usual yield curve adaptor, and W_1 (respectively W_2) is a risk-neutral domestic (respectively, risk-neutral foreign) vector Brownian motion, mutually uncorrelated as usual. For further applications, we introduce a zero bond $P_i(t,T)$ maturing at T that satisfies standard log-normal evolution

$$\frac{dP_i(t,T)}{P_i(t,T)} = r_i(t)\,dt - \Sigma_i(t,T)\cdot dW_i(t),$$

where zero bond volatilities $\Sigma_i(t,T)$ are expressed as

$$\Sigma_i(t,T) = \sigma_i(t)\int_t^T d\tau e^{-\int_t^\tau a_i(s)\,ds}.$$

Let $X(t)$ be the FX-rate between domestic and foreign markets following a Heston process. Then, the hybrid model under the domestic Brownian motion $W^* = W_1$ will have SDEs

$$\begin{aligned}
dr_1(t) &= (\zeta_1(t) - r_1(t)\,a_1(t))\,dt + \sigma_1(t)\cdot dW^*(t), && (29)\\
dr_2(t) &= (\zeta_2(t) - r_2(t)\,a_2(t) - \sqrt{z(t)}\,\lambda(t)\cdot\sigma_2(t))\,dt + \sigma_2(t)\cdot dW^*(t), && (30)\\
dX(t) &= X(t)(r_1(t) - r_2(t))\,dt + X(t)\sqrt{z(t)}\,\lambda(t)\ \cdot dW^*(t), && (31)\\
dz(t) &= \alpha(t)\,(1 - z(t))\,dt + \sqrt{z(t)}\,\gamma(t)\ \cdot dW^*(t), \quad z(0) = 1. && (32)
\end{aligned}$$

Correlations in the model are introduced via vector volatilities.

Before proceeding with the MP approximation, we briefly recall Andreasen [2] affine modifications of the setup (29-32). His approach results in a transfer of correlations between the interest rates and the Heston driving factors into a volatility displacement. Namely, Andreasen replaces the standard Heston evolution (31) by its displaced version,

$$dX(t) = X(t)(r_1(t) - r_2(t))\,dt + X(t)\left(\sqrt{z(t)}\,\lambda(t) + \nu(t)\right)\ \cdot dW^*(t),$$

where the displacement vector $\nu(t)$ is the only rate component *correlated* with the interest rates; the other Heston driving factors, i.e., vectors $\lambda(t)$ and $\gamma(t)$, are *uncorrelated* with the displacement vector and the interest rates. This correlation transfer makes the model affine and exactly solvable. In spite of the Andreasen model calibration's simplicity, we prefer the initial model setup (29-32) as being standard with the familiar parameter meanings. Finally, we give two other minor advantages in favor of the standard setup. The Andreasen scheme cannot correlate interest rates with the stochastic volatility. Also, a presence of the IR correlated displacement $\nu(t)$ in the FX-rate diffusion term can eventually lead to extreme parameters (volatility of volatility and Heston correlation) for big values of $|\nu(t)|$, or to tiny *effective* correlations between the FX and the interest rates for small values of $|\nu(t)|$.

A European option price with strike K and maturity T under risk neutral-measure reads

$$\mathbb{E}\left[\frac{(X(T) - K)^+}{N_1(T)}\right],$$

where the domestic savings account is $N_1(t) = \exp\left(\int_0^t ds\, r_1(s)\right)$. As usual, we can represent it in a more convenient domestic T-forward measure associated with domestic zero bond $P_1(t,T)$ maturing at T,

$$\mathbb{E}\left[\frac{(X(T) - K)^+}{N_1(T)}\right] = P_1(0,T)\,\mathbb{E}_T\left[(X(T) - K)^+\right].$$

Introducing the forward FX-rate $S(t)$, a martingale process under the domestic T-forward measure,

$$S(t) = \frac{X(t)\, P_2(t,T)}{P_1(t,T)},$$

we can further simplify the options price

$$\mathbb{E}\left[\frac{(X(T)-K)^+}{N_1(T)}\right] = P(0,T)\,\mathbb{E}_T\left[(S(T)-K)^+\right].$$

Given domestic and foreign zero bond volatilities $\Sigma_i(t,T)$, one can easily calculate the SDE for the forward FX-rate,

$$dS(t) = S(t)\left(\sqrt{z_T(t)}\,\lambda(t) + \Sigma_1(t,T) - \Sigma_2(t,T)\right)\cdot dW_T,$$

where W_T is T-forward Brownian motion, related to risk-neutral motion as $dW_T(t) = dW^*(t) + \Sigma_1(t,T)\,dt$, and z_T is the stochastic variance in the T-forward measure satisfying SDE

$$dz_T(t) = \left(\alpha(t)\,(1-z_T(t)) - \Sigma_1(t,T)\cdot\gamma(t)\,\sqrt{z_T(t)}\right)\,dt + \sqrt{z_T(t)}\,\gamma(t)\,\cdot dW_T(t),\quad z_T(0)=1. \tag{33}$$

Denoting a vector volatility displacement $\eta(t)\equiv\Sigma_1(t,T)-\Sigma_2(t,T)$ and a multiplier $\varepsilon(t)\equiv-\Sigma_1(t,T)\cdot\gamma(t)$, we approach evaluation of average $\mathbb{E}\left[(S(T)-K)^+\right]$ for the forward process[3]

$$dS(t) = S(t)\left(\sqrt{z_T(t)}\,\lambda(t) + \eta(t)\right)\cdot dW \tag{34}$$

with

$$dz_T(t) = \left(\alpha(t)\,(1-z_T(t)) + \varepsilon(t)\,\sqrt{z_T(t)}\right)\,dt + \sqrt{z_T(t)}\,\gamma(t)\,\cdot dW(t),\quad z_T(0)=1. \tag{35}$$

If the displacement volatility vector $\eta(t)$ were not correlated with $\lambda(t)$ and $\gamma(t)$, leading, in addition, to $\varepsilon(t)=0$, the option could be easily evaluated by its affine properties [2]. For non-trivial correlation, an exact solution does not exist, thus, some approximation should be done. We apply the elaborated above technique of the MP to the Heston DV model. For this, we fix the shift parameter β equal to 1, in order to reproduce the initial process (34) log-normality. The process Λ from (3) can be identified with

$$\Lambda = \sqrt{z_T(t)}\,\lambda(t) + \eta(t), \tag{36}$$

which gives the model stochastic variance,

$$V(t) = |\Lambda(t)|^2 = z_T(t)|\lambda(t)|^2 + 2\sqrt{z_T(t)}\lambda(t)\cdot\eta(t) + |\eta(t)|^2. \tag{37}$$

Finally, the volatility term σ_V of the stochastic variance, $dV = \cdots + \sigma_V\cdot dW$, can be obtained using Ito's formula,

$$\sigma_V(t) = \Lambda(t)\cdot\lambda(t)\;\gamma(t). \tag{38}$$

Substituting the above $\Lambda(t)$, $V(t)$, and $\sigma_V(t)$ into (19-22) and (27), one can derive the optimal Heston DV parameters via the shifted CIR process moments, $\mathbb{E}[z_T(t)]$, $\mathbb{E}[\sqrt{z_T(t)}]$, $\mathbb{E}[z_T^{\frac{3}{2}}(t)]$, and $\mathbb{E}[z_T^2(t)]$. Computational details can be found in Appendix D.

[3] We omit subscript T from average operator and Brownian motion.

5 Numerical results

As an important example, we consider a hybrid equity-interest rate setup with *correlated* driving factors. It is a special case of the general exchange rate hybrid (29-32) when the foreign currency rates become deterministic, $\sigma_2 = 0$, and have a dividend yield sense. In the numerical experiments, we use the following Heston (31-32) time-independent parameters[4],

	Heston info
spot, $X(0)$	100 %
vol of rate, $\lvert\lambda\rvert$	25 %
correlation, ρ	-40 %
SV mean-reversion, α	25 %
vol-of-vol, $\lvert\gamma\rvert$	250 %

where the Heston correlation is that of two volatility vectors, $\rho = \frac{\lambda \cdot \gamma}{|\lambda|\,|\gamma|}$.
The domestic interest rates (29) are set as follows:

	domestic HW info
yield	5 %
vol, $\lvert\sigma_1\rvert$	1 %
mean-reversion, a_1	5 %

The foreign interest rates are deterministic with 2% yield, or, in equity terms, the continuous dividend yield is 2%.

The structure in the factor space is characterized by the following values of the correlations between the Brownian motions that drive the underlying short rate, FX- or equity-rate, and the stochastic volatility,

$$
\begin{array}{lcll}
\text{Corr between IR and FX-rate} & \rightarrow & \dfrac{\lambda \cdot \sigma_1}{|\lambda|\,|\sigma_1|} & = \quad 30\%, \\
\text{Corr between IR and SV} & \rightarrow & \dfrac{\gamma \cdot \sigma_1}{|\gamma|\,|\sigma_1|} & = \quad 15\%,
\end{array}
$$

which results in the following correlation matrix between the three underlying processes, Heston FX-rate $X(t)$, stochastic volatility multiplier $z(t)$, and HW domestic short rate $r_1(t)$:

	FX	SV	IR
FX	1	-0.4	0.3
SV	-0.4	1	0.15
IR	0.3	0.15	1

Below, we present the results table, where we compare Black implied volatilities of European option prices for a large set of maturities and strikes. The strikes are presented in percentage of forward values, 100% strike corresponds to ATM options. Target volatility values ("Sim vol") are calculated using 50,000 low-discrepancy Monte Carlo paths, their standard deviation ("*std.dev.*") is estimated for 50 independent runs. We present two different approximation techniques. The first one ("Heston DV") is the MP to the displaced volatility generalization of the Heston model with optimal coefficients (19-22) and optimal

[4] Recall that the MP optimal formulas are valid for general time-dependent setups.

variance displacement (27), the second one ("Heston") is the standard Heston MP (7-10), corresponding to *zero* volatility displacement.

Maturity (*year*)	Strike	Sim vol (*std.dev.*)	MP analytic vol		MP analytic vol error	
			Heston DV	Heston	Heston DV	Heston
1	86.07	24.45 (0.06)	24.49	24.50	**0.04**	0.04
1	92.77	22.25 (0.05)	22.27	22.26	**0.02**	0.02
1	100.00	20.36 (0.05)	20.32	20.30	**-0.04**	-0.06
1	107.79	19.42 (0.05)	19.34	19.32	**-0.08**	-0.10
1	116.18	19.67 (0.06)	19.64	19.63	**-0.03**	-0.04
3	77.12	22.61 (0.08)	22.65	22.65	**0.03**	0.04
3	87.82	20.05 (0.08)	20.05	20.03	**0.01**	-0.02
3	100.00	17.95 (0.09)	17.91	17.80	**-0.04**	-0.15
3	113.87	17.23 (0.13)	17.14	17.02	**-0.09**	-0.21
3	129.67	18.02 (0.18)	17.92	17.88	**-0.09**	-0.14
5	71.50	21.89 (0.06)	21.94	21.95	**0.06**	0.06
5	84.56	19.43 (0.05)	19.45	19.37	**0.02**	-0.06
5	100.00	17.49 (0.06)	17.44	17.21	**-0.05**	-0.28
5	118.26	16.83 (0.08)	16.72	16.46	**-0.11**	-0.37
5	139.85	17.55 (0.12)	17.42	17.30	**-0.13**	-0.25
10	62.23	21.55 (0.07)	21.61	21.57	**0.06**	0.02
10	78.89	19.52 (0.07)	19.51	19.26	**0.00**	-0.26
10	100.00	18.01 (0.08)	17.91	17.33	**-0.10**	-0.69
10	126.77	17.41 (0.11)	17.22	16.53	**-0.19**	-0.88
10	160.70	17.75 (0.16)	17.51	17.08	**-0.24**	-0.67
20	51.13	22.28 (0.06)	22.32	22.08	**0.03**	-0.20
20	71.50	20.91 (0.06)	20.86	20.19	**-0.05**	-0.73
20	100.00	19.94 (0.06)	19.77	18.61	**-0.17**	-1.33
20	139.85	19.44 (0.09)	19.16	17.77	**-0.27**	-1.67
20	195.58	19.40 (0.13)	19.05	17.86	**-0.35**	-1.54

The results show an excellent approximation quality for the MP to the displaced Heston for all maturities. On the other hand, the MP to the standard Heston does not match well the simulation values for large time-horizons. The reason is that the variance process $V(t) = \left|\sqrt{z_T(t)}\,\lambda(t) + \eta(t)\right|^2$ of the underlying forward rate (34) has a positive floor which cannot be taken into account by the standard Heston MP, but can be successfully handled by its displaced generalization.

For visualization, we present graphs for a selected 10-year maturity, including Black volatility values and corresponding errors, compared with the Monte Carlo standard deviation.

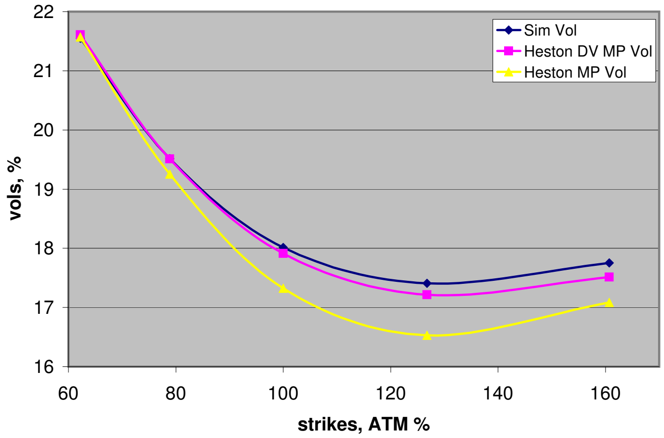

Figure 1: Heston/HW option implied volatilities for 10Y maturity

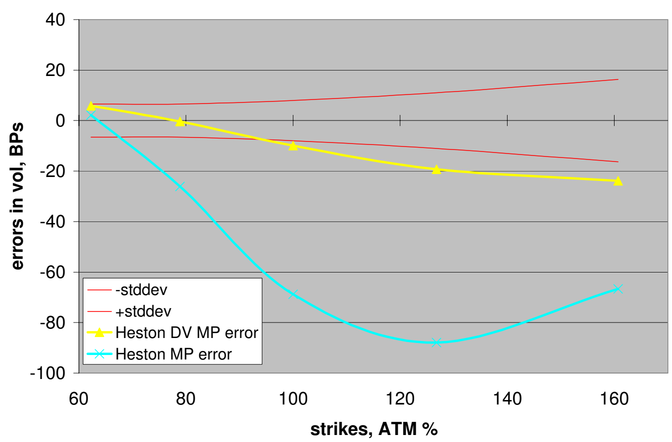

Figure 2: Heston/HW option implied volatility *errors* for 10Y maturity

6 Conclusion

In this article, we generalized the Markovian Projection technique to the displaced volatility Heston model. As an important application, we derived the effective approximation for FX/EQ options for the Heston model, coupled with correlated Gaussian interest rates. The presented numerical experiments show the excellent approximation quality. Once again, we demonstrated the efficiency of the Markovian Projection technique in solving challenging approximation problems.

The authors are indebted to Serguei Mechkov for discussions and numerical implementation help as well as to their colleagues at NumeriX, especially, to Gregory Whitten for supporting this work and Patti Harris for the excellent editing. AA is grateful to Vladimir Piterbarg, Jesper Andreasen, and Dominique Bang for stimulating discussions.

A Different forms of the Heston model and its option pricing

In his famous paper [7], Heston proposed an exactly solvable stochastic volatility model[5],

$$\begin{aligned} dS_t &= S_t\sqrt{v_t}dU, && (39)\\ dv_t &= \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dV, && (40)\\ \langle dU dV\rangle &= \rho dt. && (41)\end{aligned}$$

Andersen and Andreasen [1] defined a version of the Heston model suitable for time-dependent coefficients[6],

$$\begin{aligned} dS_t &= S_t\sqrt{z_t}\lambda_t\, dU, && (42)\\ dz_t &= a_t(1 - z_t)dt + \sqrt{z_t}\gamma_t\, dV, \qquad \text{for } z_0 = 1, && (43)\\ \langle dU dV\rangle &= \rho_t dt. && (44)\end{aligned}$$

A map between the initial version (39) and the Andersen-Andreasen version (42), $v_t = \lambda_t^2\, z_t$, is achieved if

$$\begin{aligned} a_t &= \frac{\kappa\theta}{(v_0 - \theta)\, e^{-\kappa\, t} + \theta}, && (45)\\ \lambda_t &= \sqrt{(v_0 - \theta)\, e^{-\kappa\, t} + \theta}, && (46)\\ \gamma_t &= \frac{\xi}{\sqrt{(v_0 - \theta)\, e^{-\kappa\, t} + \theta}}. && (47)\end{aligned}$$

Intuitive derivation of this result is based on the observation that $\mathbb{E}[z_t] = 1$, which gives

$$\lambda_t^2 = \mathbb{E}[v_t] \equiv m_t,$$

and $a_t = \kappa\theta/m_t$ and $\gamma_t = \xi/\sqrt{m_t}$. Here, the variance v_t average satisfies $m_t' = \kappa(\theta - v_t)$, resulting in

$$m_t = (v_0 - \theta)\, e^{-\kappa\, t} + \theta.$$

[5] Here, we omit a drift depending on rates as far as it can be easily eliminated by scaling.

[6] To preserve similarities with the Heston version, we adopt here correlations via Brownian motions, equivalent to our initial volatility vector correlations (2). Also, we denote time-dependence as a subscript.

The advantage of the Heston model with time-dependent parameters is its higher number of degrees of freedom, resulting in the possibility of calibration to multiple options per exercise. To fit our vector volatility form (2) with a shift parameter, it is sufficient to identify $|\sigma_H(t)| = \lambda_t$, $|\sigma_z(t)| = \gamma_t$, $\sigma_H(t)\cdot\sigma_z(t) = \rho_t\,\lambda_t\,\gamma_t$, and $\beta(t) = 1$. The inverse map handles volatilities accordingly, but can require supplementary work for time-dependent shifts $\beta(t)$. Indeed, to transform (2) into a log-normal shape, one should apply the shift averaging technique [8], $\beta(t) \to \bar{\beta}$, and consider $1 + \bar{\beta}\,\Delta S(t)$ as a new log-normal variable. Below, we recall a technique of analytical calculation of option prices for the log-normal case, having in mind its generalization to a skewed case. We use the Laplace transform, instead of the equivalent Fourier one initially introduced by Heston [7].

Our goal is to calculate an option price

$$\mathbb{E}[(S_t - K)^+]$$

via Laplace transform of the characteristic function, which is known analytically. Introduce a process y_t,

$$y_t = \ln\left(\frac{S_t}{S_0}\right),$$

satisfying

$$dy = -\frac{1}{2}\,z_t\lambda_t^2\,dt + \sqrt{z_t}\lambda_t\,dU. \tag{48}$$

Then, due to affine properties of the model, one can write the ODE for the process y_t characteristic function,

$$\phi_H(T,\xi) = E\left[e^{\xi y_T}\right], \tag{49}$$

which can be used for the price calculation via Laplace transform. Indeed,

$$(S_0\,e^{y_t} - K)^+ = \frac{K}{2\pi i}\int_{C^+}\frac{d\xi}{\xi(\xi-1)}\exp\left(\xi\,y_t - \xi\,\ln\left(\frac{K}{S_0}\right)\right), \tag{50}$$

where contour C^+ is parallel to the imaginary axis and passes to the right of the integrand poles (0 and 1). Then, the option price can be represented in terms of a moment generating function (MGF) $\mathcal{K}(\xi) \equiv \ln\phi(T,\xi)$, i.e.,

$$E\left[(S_0\,e^{y_t} - K)^+\right] = \frac{K}{2\pi i}\int_{C^+}\frac{d\xi}{\xi(\xi-1)}\exp\left(\mathcal{K}(\xi) - \xi\,\ln\left(\frac{K}{S_0}\right)\right). \tag{51}$$

The standard approach is to compute this integral numerically. As we will see below, it is possible for certain cases to reduce a complex integration to a real one. We take the integral along the line $\xi = \frac{1}{2} + i\omega$ with real ω, taking into account the contribution of the pole at $\xi = 1$,

$$E\left[(S_0\,e^{y_t} - K)^+\right] = S_0 - \frac{K}{2\pi}\int_{-\infty}^{\infty}\frac{d\omega}{\omega^2+\frac{1}{4}}\exp\left(\mathcal{K}\left(\frac{1}{2}+i\omega\right) - (i\omega+\frac{1}{2})\ln\left(\frac{K}{S_0}\right)\right). \tag{52}$$

If the values of $\mathcal{K}\left(\frac{1}{2}+i\omega\right)$ are real, one can replace

$$\exp\left(-(i\omega+\frac{1}{2})\ln\left(\frac{K}{S_0}\right)\right) \to \left(\frac{K}{S_0}\right)^{-\frac{1}{2}}\cos\left(\omega\ln\left(\frac{K}{S_0}\right)\right).$$

In order to calculate the characteristic function $\phi(T, \xi)$, we consider a conditional expectation

$$u(t; z, y) = E\left[e^{\xi y_T} \mid z_t = z, y_t = y\right],$$

which satisfies PDE (zero drift condition)

$$u_t - \tfrac{1}{2}\lambda_t^2\, z\, u_y + \tfrac{1}{2}\lambda_t^2\, z\, u_{yy} + a_t(1-z)\, u_z + \tfrac{1}{2}\gamma_t^2\, z\, u_{zz} + \lambda_t\gamma_t\rho_t\, z\, u_{yz} = 0. \tag{53}$$

Looking for the solution in affine form,

$$u(t; z, y) = e^{A_t + B_t\, z + C_t\, y},$$

and imposing final conditions $A(T) = B(T) = 0$ and $C(T) = \xi$, we have

$$\begin{aligned}
C_t &= \xi, \\
B_t' &= -\frac{1}{2}\gamma_t^2 B_t^2 + (a_t - \lambda_t\gamma_t\rho_t\xi)\, B_t - \lambda_t^2 \frac{\xi^2 - \xi}{2}, \\
A_t' &= -a_t\, B_t.
\end{aligned} \tag{54}$$

These ODEs can be resolved numerically or, for step-constant parameters, iteratively using a Riccati solution.

Coming back to the above remark about the possibility of real instead of complex integration, we should mention that, for zero correlation, the MGF $\mathcal{K}\left(\frac{1}{2} + i\omega\right)$ is real, and we can do a real integration; although, for the non-zero correlation case, a complex integration is necessary.

Finally, to calculate the European option price, one performs numerical integration (52), finding $\mathcal{K}\left(\frac{1}{2} + i\omega\right)$ for each value of ω solving the ODEs (54).

B European option pricing of the Heston DV model

In this Appendix, we generalize the above results of European option pricing of the standard Heston model to the Heston model with the orthogonal displaced volatility. In reality, it is sufficient to make a few modifications of the above formulas. Indeed, one can rewrite the vector Heston DV definition (11) using in a form similar to (42),

$$dS_t = S_t\left(\sqrt{z_t}\lambda_t\, dU + \alpha_t\, dZ\right), \tag{55}$$

$$dz_t = a_t(1 - z_t)dt + \sqrt{z_t}\gamma_t\, dV, \tag{56}$$

$$\langle dU dV\rangle = \rho_t dt, \quad \langle dU dZ\rangle = \langle dV dZ\rangle = 0. \tag{57}$$

As in the previous Appendix, we apply a shift averaging procedure and eliminate the shift parameter. Again, introducing logarithm process

$$\tilde{y}_t = \ln\left(\frac{S_t}{S_0}\right)$$

satisfying

$$d\tilde{y}_t = -\frac{1}{2}\, z_t\lambda_t^2\, dt + \sqrt{z_t}\lambda_t\, dW \quad - \frac{1}{2}\alpha_t^2\, dt + \alpha_t\, dZ, \tag{58}$$

one notices that it differs from the standard Heston one (48) by the two last terms, which are uncorrelated with the rest. Thus, the Heston DV characteristic function

$$\phi_{HDV}(T,\xi) = E\left[e^{\xi \tilde{y}_T}\right] \tag{59}$$

can be obtained from the standard Heston one (49) by a simple multiplication[7],

$$\phi_{HDV}(T,\xi) = \phi_H(T,\xi)\, e^{\frac{1}{2}\, v_T^{(\alpha)}\, \xi\,(\xi-1)}, \tag{60}$$

where $v_T^{(\alpha)} = \int_0^T dt\, \alpha_t^2$ is a realized variance of the Gaussian component. Having calculated the Heston DV characteristic function, one can calculate the option price using a numerical integration from (52) and the MGF $\mathcal{K}_{HDV}(\xi) \equiv \ln \phi_{HDV}(T,\xi)$.

C Details of the MP onto the Heston DV model

In this Appendix, we give details on minimization of the criterion (16)

$$\chi_1^2 = \mathbb{E}\left[\left(\mu_V - \left((\ln|\sigma_H|^2)' - \theta\right)(V - V_D) - V_D' - |\sigma_H|^2\theta\right)^2\right].$$

Here, for better legibility, we omit the time argument t, supposing time-dependence implicitly with prime sign $'$ denoting the derivative of a deterministic function of time.

As in Section 3, we fix the displacement variance V_D before proceeding to the minimization. Denoting $A = (\ln|\sigma_H|^2)' - \theta$ and $B = V_D' + |\sigma_H|^2\theta$, one can rewrite

$$\chi_1^2 = \mathbb{E}\left[(\mu_V - A(V - V_D) - B)^2\right].$$

Solving the first optimal equation $\frac{\partial \chi_1^2}{\partial B} = 0$, we get

$$\mathbb{E}[\mu_V] - A\,(\mathbb{E}[V] - V_D) - B = 0. \tag{61}$$

Substituting $\mathbb{E}[\mu_V] = \mathbb{E}[V]'$ and expressions for A and B into (61) gives

$$\mathbb{E}[V(t)]' - V_D' - \left((\ln|\sigma_H|^2)' - \theta\right)(\mathbb{E}[V] - V_D) - |\sigma_H|^2\theta = 0.$$

It is easy to check that solution $|\sigma_H|^2 = \mathbb{E}[V] - V_D$ satisfies the condition. This proves the first optimal condition for the effective volatility (19).

Proceed now to the second optimal equation, $\frac{\partial \chi_1^2}{\partial A} = 0$, resulting in

$$\mathbb{E}\left[(V - V_D)\,(\mu_V - A(V - V_D) - B)\right] = 0.$$

Subtracting from it the first condition (61) multiplied by average $\mathbb{E}[V - V_D]$, we have

$$A = \frac{\mathbb{E}[(V - V_D)\,\mu_V] - \mathbb{E}[(V - V_D)]\,\mathbb{E}[\mu_V]}{\mathbb{E}[(V - V_D)^2] - \left(\mathbb{E}[(V - V_D)]\right)^2} = \frac{\mathbb{E}[V\,\mu_V] - \mathbb{E}[V]\,\mathbb{E}[\mu_V]}{\mathrm{Var}[V]}. \tag{62}$$

[7] See also [2].

Averages in the last numerator, $\mathbb{E}[V\,\mu_V] - \mathbb{E}[V]\,\mathbb{E}[\mu_V]$, can be simplified using Ito's lemma applied to a square of the process V, satisfying the SDE $dV = \mu_V dt + \sigma_V \cdot dW$, see (5),

$$dV^2 = (2\,V\,\mu_V + |\sigma_V|^2)\,dt + 2\,V\,\sigma_V \cdot dW(t).$$

Applying the expectation operator to the above SDE and reusing the average $\mathbb{E}[\mu_V] = \mathbb{E}[V]'$, one can obtain a practical form of the numerator,

$$\mathbb{E}[V\,\mu_V] - \mathbb{E}[V]\,\mathbb{E}[\mu_V] = \frac{1}{2}\,(\mathrm{Var}[V])' - \frac{1}{2}\,\mathbb{E}[|\sigma_V|^2].$$

Substituting it into the condition (62) for $A = (\ln|\sigma_H|^2)' - \theta$, we restore the optimal expression for the mean-reversion (20),

$$\theta = (\ln|\sigma_H|^2)' - \frac{1}{2}\,(\ln\mathrm{Var}[V])' + \frac{1}{2}\,\frac{\mathbb{E}[|\sigma_V|^2]}{\mathrm{Var}[V]}.$$

D FX/EQ options in the presence of Gaussian correlated rates: computational details

To reduce formulas, we introduce the two processes,

$$Y_\lambda(t) = \Lambda(t)\cdot\lambda(t) \quad \text{and} \quad Y_\gamma(t) = \Lambda(t)\cdot\gamma(t).$$

Then, the processes $V(t)$ and $\sigma_V(t)$ transform to

$$V(t) = \frac{Y_\lambda^2(t)}{|\lambda(t)|^2} + |\eta(t)|^2\,D(t), \tag{63}$$

$$\sigma_V(t) = Y_\lambda(t)\;\gamma(t), \tag{64}$$

where decorrelation parameter $D(t) \equiv 1 - \frac{(\lambda(t)\cdot\eta(t))^2}{|\lambda(t)|^2\,|\eta(t)|^2}$.

All the averages underlying efficient Heston DV coefficients (19-22) and (27) can be expressed via averages of $Y_\lambda(t)$ and $Y_\gamma(t)$. After elementary algebra, we have[8]

$$\mathbb{E}[V] = \frac{\mathbb{E}[Y_\lambda^2]}{|\lambda|^2} + |\eta|^2\,D, \tag{65}$$

$$Var[V] = \frac{\mathbb{E}[Y_\lambda^4] - (\mathbb{E}[Y_\lambda^2])^2}{|\lambda|^4}, \tag{66}$$

$$\mathbb{E}[|\sigma_V|^2] = |\gamma|^2\;\mathbb{E}[Y_\lambda^2], \tag{67}$$

$$\mathbb{E}[V\,|\sigma_V|^2] = |\gamma|^2\,\left(\frac{\mathbb{E}[Y_\lambda^4]}{|\lambda|^2} + |\eta|^2\,D\;\mathbb{E}[Y_\lambda^2]\right), \tag{68}$$

$$\mathbb{E}[\Lambda\cdot\sigma_V] = \mathbb{E}[Y_\lambda\,Y_\gamma], \tag{69}$$

$$\mathbb{E}[V\,\Lambda\cdot\sigma_V] = \frac{\mathbb{E}[Y_\lambda^3 Y_\gamma]}{|\lambda|^2} + |\eta|^2\,D\;\mathbb{E}[Y_\lambda\,Y_\gamma]. \tag{70}$$

Averages $\mathbb{E}[Y_\lambda^2]$, $\mathbb{E}[Y_\lambda^4]$, $\mathbb{E}[Y_\lambda Y_\gamma]$, and $\mathbb{E}[Y_\lambda^3 Y_\gamma]$ can be expressed via the CIR process moments, $\mathbb{E}[z_T(t)]$, $\mathbb{E}[\sqrt{z_T(t)}]$, $\mathbb{E}[z_T^{\frac{3}{2}}(t)]$, and $\mathbb{E}[z_T^2(t)]$.

[8] For better legibility, we omit time dependence.

E Calculations of the shifted CIR moments

The shifted CIR process z_T,

$$dz_T(t) = \Big(\alpha(t)\,(1 - z_T(t)) + \varepsilon\,\sqrt{z_T(t)}\Big)\;dt + \sqrt{z_T(t)}\,\gamma(t)\;\cdot dW(t), \quad z_T(0) = 1, \tag{71}$$

is not affine anymore. Thus, one should make an approximation in order to calculate its moments. This can be done using Markovian Projection. Indeed, we will look for the *affine* approximating process

$$d\tilde{z}_T(t) = (A(t) - B(t)\tilde{z}_T(t))\;dt + \sqrt{\tilde{z}_T(t)}\,\gamma(t)\;\cdot dW(t), \quad \tilde{z}_T(0) = 1, \tag{72}$$

keeping the same diffusion term and linearizing the drift. To calculate unknown functions $A(t)$ and $B(t)$, one should minimize $\chi^2(t)$ criterion

$$\chi^2(t) = \mathbb{E}\left[\Big(\alpha(t)\,(1 - z_T(t)) + \varepsilon\,\sqrt{z_T(t)} - A(t) + B(t)z_T(t)\Big)^2\right],$$

for any time t. Denoting corrections $A_1(t) = A(t) - \alpha(t)$ and $B_1(t) = B(t) - \alpha(t)$, we have

$$\chi^2(t) = \mathbb{E}\left[\Big(\varepsilon\,\sqrt{z_T(t)} - A_1(t) + B_1(t)z_T(t)\Big)^2\right].$$

Setting $\chi^2(t)$ derivatives over $A_1(t)$ and $B_1(t)$ to zero for each t, one can obtain the optimal corrections,

$$A_1(t) \;=\; \varepsilon(t)\,\mathbb{E}[\sqrt{z_T(t)}] + B_1(t)\,\mathbb{E}[z_T(t)], \tag{73}$$

$$B_1(t) \;=\; -\varepsilon(t)\,\frac{\mathbb{E}[z_T^{\frac{3}{2}}(t)] - \mathbb{E}[\sqrt{z_T(t)}]\,\mathbb{E}[z_T(t)]]}{\mathrm{Var}[z_T(t)]}. \tag{74}$$

In general, the shift ε magnitude is quite small, 10-20%. The corrections $A_1(t)$ and $B_1(t)$ have the same order in magnitude. In order to calculate their values in a leading order, it is sufficient to evaluate averages in the r.h.s. of (73-74) in zero-th order in ε or, in other words, set ε to zero in (71), transforming it to the standard CIR model, $z_T \to z$.

The first and second moments are available analytically for the standard CIR model. To calculate its non-entire moments, one can use the explicitly available PDF function corresponding the non-centered χ^2 distribution[9]. The density is available as a series expansion. Thus, to calculate the average $\mathbb{E}[\sqrt{z(t)}]$ or $\mathbb{E}[z^{\frac{3}{2}}(t)]$, it is sufficient to integrate terms of the density series with $\sqrt{z}$ or $z^{\frac{3}{2}}$, respectively, and make summations. Each term in the resulting sum is expressed using Gamma-functions, which can be efficiently computed using standard algorithms. And, due to fast convergence of the obtained series, the proposed way of calculation should be numerically efficient.

Note at the end that the affine model (72) can be transformed to a product of the classical CIR process and a deterministic function,

$$\tilde{z}_T(t) = M(t)\,\tilde{z}(t),$$

[9]Strictly speaking, this is valid for the CIR model with *time-independent* coefficients. For our time-dependent case, one can use the averaging technique [8] to transform the initial model to approximate a time-independent one.

with the multiplier

$$M(t) = 1 + e^{-\int_0^t dsB(s)} \int_0^t d\tau \left(A(\tau) - B(\tau)\right) e^{\int_0^\tau dsB(s)},$$

and the CIR process

$$d\tilde{z}(t) = \frac{A(t)}{M(t)} \left(1 - \tilde{z}(t)\right) \, dt + \sqrt{\tilde{z}(t)} \, \frac{\gamma(t)}{\sqrt{M(t)}} \cdot dW(t), \quad \tilde{z}(0) = 1.$$

References

[1] L. Andersen and J. Andreasen (2002) “Volatile volatilities”, *Risk*, Dec, 163–168.

[2] J. Andreasen (2006) “Closed Form Pricing of FX Options under Stochastic Rates and Volatility”, Global Derivatives Conference, ICBI.

[3] A. Antonov and T. Misirpashaev (2006) “Markovian Projection onto a Displaced Diffusion: Generic Formulas with Applications,” available at SSRN: http://ssrn.com/abstract=937860.

[4] A. Antonov, T. Misirpashaev and V. Piterbarg (2007) “Markovian Projection Onto a Heston Model,” available at SSRN: http://ssrn.com/abstract=997001.

[5] B. Dupire (1997) “A Unified Theory of Volatility,” Banque Paribas working paper, reprinted in “Derivatives Pricing: the Classic Collection,” edited by P. Carr, Risk Books, London, 2004.

[6] I. Gyöngy (1986) “Mimicking the One-Dimensional Marginal Distributions of Processes Having an Ito Differential”, *Probability Theory and Related Fields*, 71, 501–516.

[7] S. Heston (1993) “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”, *The Review of Financial Studies*, Vol. 6, No. 2.

[8] V. Piterbarg (2005) “Time to Smile,” *Risk*, May, 71–75.

[9] V. Piterbarg (2007) “Markovian Projection for Volatility Calibration,” *Risk*, April, 84–89.