SIAM J. FINANCIAL MATH.
Vol. 2, pp. 255–286

On the Heston Model with Stochastic Interest Rates*

Lech A. Grzelak† and Cornelis W. Oosterlee‡

Abstract. We discuss the Heston model [*Rev. Financ. Stud.*, 6 (1993), pp. 327–343] with stochastic interest rates driven by Hull–White (HW) [*J. Derivatives*, 4 (1996), pp. 26–36] or Cox–Ingersoll–Ross (CIR) [*Econometrica*, 53 (1985), pp. 385–407] processes. Two projection techniques to derive affine approximations of the original hybrid models are presented. In these approximations we can prescribe a nonzero correlation structure between all underlying processes. The affine approximate models admit pricing basic derivative products by Fourier techniques [P. P. Carr and D. B. Madan, *J. Comput. Finance*, 2 (1999), pp. 61–73, F. Fang and C. W. Oosterlee, *SIAM J. Sci. Comput.*, 31 (2008), pp. 826–848] and can therefore be used for fast calibration of the hybrid model.

Key words. equity-interest rate hybrid models, stochastic volatility, Heston–Hull–White and Heston–Cox–Ingersoll–Ross processes, approximation by affine diffusion process

AMS subject classifications. 91G60, 91G30, 91G20

DOI. 10.1137/090756119

1. Introduction.

Modelling derivative products in finance usually starts with the specification of a system of stochastic differential equations (SDEs) that correspond to state variables like stock, interest rate, and volatility. By correlating the SDEs from the different asset classes, one can define so-called hybrid models and use them for pricing multiasset derivatives. Even if each of these SDEs yields a closed-form solution, a nonzero correlation structure between the processes may cause difficulties for modeling and product pricing. Typically, a closed-form solution of the hybrid models is not known, and numerical approximation by means of Monte Carlo (MC) simulation or discretization of the corresponding partial differential equations (PDEs) has to be employed for model evaluation and derivative pricing. The speed of pricing European products is, however, crucial, especially for the calibration. Several theoretically attractive SDE models that cannot fulfill the speed requirements are not used in practice.

The aim of this paper is to define hybrid SDE models that fit in the class of affine diffusion processes (AD), as in Duffie, Pan, and Singleton [13]. For processes within this class a closed-form solution of the characteristic function (ChF) exists. Suppose we have given a system of SDEs, i.e.,

$$\mathrm{d}\mathbf{X}(t) = \mu(\mathbf{X}(t))\mathrm{d}t + \sigma(\mathbf{X}(t))\mathrm{d}\mathbf{W}(t). \tag{1.1}$$

*Received by the editors April 15, 2009; accepted for publication (in revised form) January 12, 2011; published electronically March 15, 2011.
http://www.siam.org/journals/sifin/2/75611.html

†Delft University of Technology, Delft Institute of Applied Mathematics, Mekelweg 4, 2628 CD, Delft, The Netherlands, and Derivatives Research and Validation Group, Rabobank, Jaarbeursplein 22, 3521 AP, Utrecht, The Netherlands (L.A.Grzelak@tudelft.nl).

‡CWI – Centrum Wiskunde & Informatica, Amsterdam, The Netherlands, and Delft University of Technology, Delft Institute of Applied Mathematics Mekelweg 4, 2628 CD, Delft, The Netherlands (C.W.Oosterlee@cwi.nl).

This system (1.1) is said to be of affine form if

$$\mu(\mathbf{X}(t)) = a_0 + a_1\mathbf{X}(t) \quad \text{for any } (a_0, a_1) \in \mathbb{R}^n \times \mathbb{R}^{n\times n}, \tag{1.2}$$

$$(\sigma(\mathbf{X}(t))\sigma(\mathbf{X}(t))^{\mathrm{T}})_{i,j} = (c_0)_{ij} + (c_1)_{ij}^{\mathrm{T}}\mathbf{X}(t), \quad \text{with } (c_0, c_1) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n\times n}, \tag{1.3}$$

$$r(\mathbf{X}(t)) = r_0 + r_1^{\mathrm{T}}\mathbf{X}(t) \quad \text{for } (r_0, r_1) \in \mathbb{R} \times \mathbb{R}^n, \tag{1.4}$$

for $i, j = 1, \ldots, n$, with $r(\mathbf{X}(t))$ being an interest rate component. Then, the discounted ChF is of the following form [13]:

$$\phi(\mathbf{u}, \mathbf{X}(t), t, T) = \mathbb{E}^{\mathbb{Q}}\left(\exp\left(-\int_t^T r(s)\mathrm{d}s + i\mathbf{u}^{\mathrm{T}}\mathbf{X}(T)\right)\Big|\mathcal{F}(t)\right) = \mathrm{e}^{A(\mathbf{u},\tau)+\mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)\mathbf{X}(t)},$$

where the expectation is taken under the risk-neutral measure, $\mathbb{Q}$. For a time lag, $\tau := T - t$, the coefficients $A(\mathbf{u}, \tau)$ and $\mathbf{B}^{\mathrm{T}}(\mathbf{u}, \tau)$ have to satisfy the following complex-valued ordinary differential equations (ODEs):

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{B}(\mathbf{u},\tau) = & -r_1 + a_1^{\mathrm{T}}\mathbf{B}(\mathbf{u},\tau) + & \dfrac{1}{2}\mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)c_1\mathbf{B}(\mathbf{u},\tau), \\ \dfrac{\mathrm{d}}{\mathrm{d}\tau}A(\mathbf{u},\tau) = & -r_0 + \mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)a_0 + & \dfrac{1}{2}\mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)c_0\mathbf{B}(\mathbf{u},\tau), \end{cases} \tag{1.5}$$

with $a_i, c_i, r_i,\ i = 0, 1$, as in (1.2), (1.3), and (1.4).

In this article we focus our attention specifically on a hybrid model which combines the equity and interest rate asset classes. Brigo and Mercurio [8] have shown that the assumption of constant interest rates in the classical Black–Scholes model [7] can be generalized, and by including the stochastic interest rate process of Hull and White [23], one is still able to obtain a closed-form solution for European-style option prices. Originally, the Black–Scholes–Hull–White model in [8] was not dedicated to pricing hybrid products but to increasing the accuracy for long-maturity options. The model is not, however, able to describe any smile and skew shapes present in the equity markets.

In [37] a hybrid model was presented which could provide the skew pattern for the equity and included a stochastic (but uncorrelated) interest rate process. Generalizations were presented in [19, 3], where the Heston stochastic volatility model [22] was used. Some form of correlation was indirectly modeled by including additional terms in the SDEs (this approach is discussed in some detail in section 3.1.1).

In [20, 21] the Heston stochastic volatility model was replaced by the Schöbel–Zhu (SZ) model [35], while the interest rate was still driven by a Hull–White (HW) process (SZHW model). In this model a full matrix of correlations can be directly imposed on the driving Brownian motions. The model is well defined under the class of AD processes, but since the SZHW model is based on a Vašiček-type process [36] for the stochastic volatility, the volatilities can become negative.

A different approach to modelling equity-interest rate hybrids was presented by Benhamou, Rivoira, and Gruz [6], extending the local volatility framework of Dupire [15] and Derman and Kani [12] and incorporating stochastic interest rates.

Here, we investigate the Heston–Hull–White (HHW) and the Heston–Cox–Ingersoll–Ross (HCIR) hybrid models and propose approximations so that we can obtain their ChFs. The

framework presented is relatively easy to understand and implement. It is inspired by the techniques in [19, 2].

Our approximations do not require several preliminary calculations of expectations like in the case of Markovian projection methods [4, 5]. The resulting option pricing method benefits greatly from the speed of ChF evaluations.

The interest rate models studied here cannot generate *interest rate* implied volatility smiles or skews. They can therefore mainly be used for long-term equity options and for "not too complicated" equity-interest rate hybrid products. As described in [24], for accurate modeling of hybrid derivatives, it is necessary to be able to describe a nonzero correlation between equity and interest rate. This is possible in the approximations presented here.

This paper is organized as follows. In section 2 we discuss the full-scale Heston hybrid models with stochastic interest rate processes. Section 3 presents a deterministic approximation of the HHW hybrid model, together with the corresponding ChF, and section 4 gives the ChF based on another stochastic approximation of that hybrid model. In section 5 we deal with the HCIR model. In section 6 the calibration based on the approximations of the full-scale hybrid models is applied. Section 7 offers concluding remarks. Details of proofs and tests are in the appendices, where, in particular, in Appendix C we compare the performance of the approximations developed with the Markovian projection method studied in [4, 5].

2. Heston hybrid models with stochastic interest rate.

With state vector $\mathbf{X}(t) = [S(t), v(t)]^{\mathrm{T}}$, under the risk-neutral pricing measure, the Heston stochastic volatility model [22], which is our point of departure, is specified by the following system of SDEs:

$$\begin{cases} \mathrm{d}S(t)/S(t) = & r\mathrm{d}t + & \sqrt{v(t)}\mathrm{d}W_x(t), & S(0) > 0, \\ \mathrm{d}v(t) = & \kappa(\bar{v} - v(t))\mathrm{d}t + & \gamma\sqrt{v(t)}\mathrm{d}W_v(t), & v(0) > 0, \end{cases} \tag{2.1}$$

with $r > 0$ a constant interest rate, correlation $\mathrm{d}W_x(t)\mathrm{d}W_v(t) = \rho_{x,v}\mathrm{d}t$, and $|\rho_{x,v}| < 1$. The variance process, $v(t)$, of the stock, $S(t)$, is a mean reverting square-root process, in which $\kappa > 0$ determines the speed of adjustment of the volatility towards its theoretical mean $\bar{v} > 0$, and $\gamma > 0$ is the second-order volatility, i.e., the volatility of the volatility.

As already indicated in [22], the model given in (2.1) is not in the class of affine processes, whereas under the log-transform for the stock, $x(t) = \log S(t)$, it is. Then, the discounted ChF is given by

$$\phi_{\mathrm{H}}(u, \mathbf{X}(t), \tau) = \exp\left(A(u,\tau) + B(u,\tau)x(t) + C(u,\tau)v(t)\right), \tag{2.2}$$

where the functions $A(u,\tau)$, $B(u,\tau)$, and $C(u,\tau)$ are known in closed form (see [22]).

The ChF is explicit, but its inverse also has to be found for pricing purposes. Because of the form of the ChF, we cannot get its inverse analytically, and a numerical method for integration has to be used; see, for example, [10, 16, 29, 30] for Fourier methods.

2.1. Full-scale hybrid models.

A constant interest rate, r, may be insufficient for pricing interest rate sensitive products. Therefore, we extend our state vector with an additional stochastic quantity, i.e., $\mathbf{X}(t) = [S(t), v(t), r(t)]^{\mathrm{T}}$. This model corresponds to a *hybrid stochastic volatility equity model with a stochastic interest rate* process, $r(t)$. In particular, we

add to the Heston model the HW interest rate [23] or the square-root CIR process [11]. The extended model can be presented in the following way:

$$\text{(2.3)}\qquad \begin{cases} \mathrm{d}S(t)/S(t) = & r(t)\mathrm{d}t + & \sqrt{v(t)}\mathrm{d}W_x(t), & S(0)>0, \\ \mathrm{d}v(t) = & \kappa(\bar{v}-v(t))\mathrm{d}t + & \gamma\sqrt{v(t)}\mathrm{d}W_v(t), & v(0)>0, \\ \mathrm{d}r(t) = & \lambda(\theta(t)-r(t))\mathrm{d}t + & \eta r^p(t)\mathrm{d}W_r(t), & r(0)>0, \end{cases}$$

where exponent $p=0$ in (2.3) represents the HHW model and for $p=\frac{1}{2}$ it becomes the HCIR model. For both models the correlations are given by $\mathrm{d}W_x(t)\mathrm{d}W_v(t)=\rho_{x,v}\mathrm{d}t$, $\mathrm{d}W_x(t)\mathrm{d}W_r(t)=\rho_{x,r}\mathrm{d}t$, and $\mathrm{d}W_v(t)\mathrm{d}W_r(t)=\rho_{v,r}\mathrm{d}t$, and κ, γ, and $\bar{v}$ are as in (2.1); $\lambda>0$ determines the speed of mean reversion for the interest rate process; $\theta(t)$ is the interest rate term structure; and η controls the volatility of the interest rate. We note that the interest rate process in (2.3) for $p=\frac{1}{2}$ is of the same form as the variance process $v(t)$.

System (2.3) is not in the affine form, not even with $x(t)=\log S(t)$. In particular, the symmetric instantaneous covariance matrix is given by

$$\text{(2.4)}\qquad \sigma(\mathbf{X}(t))\sigma(\mathbf{X}(t))^{\mathrm{T}} = \begin{bmatrix} v(t) & \rho_{x,v}\gamma v(t) & \rho_{x,r}\eta r^p(t)\sqrt{v(t)} \\ * & \gamma^2 v(t) & \rho_{r,v}\gamma\eta r^p(t)\sqrt{v(t)} \\ * & * & \eta^2 r^{2p}(t) \end{bmatrix}_{(3\times 3)}.$$

Setting the correlation $\rho_{r,v}$ to zero would still not make the system affine. Matrix (2.4) is of the linear form with respect to state vector $[x(t)=\log S(t), v(t), r(t)]^{\mathrm{T}}$ if two correlations, $\rho_{r,v}$ and $\rho_{x,r}$, are set to zero.[1] Models with two correlations equal to zero are covered in [31].

Since for pricing equity-interest rate products a nonzero correlation between stock and interest rate is crucial (see, for example, [24]), approximations to the Heston hybrid models need to be formulated, so that correlations can be imposed. Variants are discussed in the sections to follow. These approximate models are evaluated with the help of the Cholesky decomposition of a correlation matrix.

We can decompose a given general symmetric correlation matrix, $\mathbf{C}$, denoted by

$$\text{(2.5)}\qquad \mathbf{C} = \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ * & 1 & \rho_3 \\ * & * & 1 \end{bmatrix},$$

as $\mathbf{C}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$, where $\mathbf{L}$ is a lower triangular matrix, with

$$\text{(2.6)}\qquad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_1 & \sqrt{1-\rho_1^2} & 0 \\ \rho_2 & \frac{\rho_3-\rho_2\rho_1}{\sqrt{1-\rho_1^2}} & \sqrt{1-\rho_2^2-\left(\frac{\rho_3-\rho_2\rho_1}{\sqrt{1-\rho_1^2}}\right)^2} \end{bmatrix}.$$

We then rewrite the system of SDEs in terms of the independent Brownian motions, $\mathrm{d}\widetilde{\mathbf{W}}(t)$, with the help of the lower triangular matrix $\mathbf{L}$.

[1] Here we assume positive parameters.

Since our main objective is to derive a closed-form ChF while assuming a nonzero correlation between the equity process, $S(t)$, and the interest rate, $r(t)$, we first assume that the Brownian motions for the interest rate $r(t)$ and the variance $v(t)$ are not correlated (the case of a full correlation structure is discussed in detail in Appendix B).

By exchanging the order of the state variables $\mathbf{X}(t) = [S(t), v(t), r(t)]^{\mathrm{T}}$ to $\mathbf{X}^*(t) = [r(t), v(t), S(t)]^{\mathrm{T}}$, the HHW and HCIR models in (2.3) then have $\rho_1 \equiv \rho_{r,v} = 0$, $\rho_2 \equiv \rho_{x,r} \neq 0$, and $\rho_3 \equiv \rho_{x,v} \neq 0$ in (2.5) and read as

$$\begin{bmatrix} \mathrm{d}r(t) \\ \mathrm{d}v(t) \\ \dfrac{\mathrm{d}S(t)}{S(t)} \end{bmatrix} = \begin{bmatrix} \lambda(\theta(t) - r(t)) \\ \kappa(\bar{v} - v(t)) \\ r(t) \end{bmatrix} \mathrm{d}t + \sigma(\mathbf{X}^*(t)) \begin{bmatrix} \mathrm{d}\widetilde{W}_r(t) \\ \mathrm{d}\widetilde{W}_v(t) \\ \mathrm{d}\widetilde{W}_x(t) \end{bmatrix}, \tag{2.7}$$

with

$$\sigma(\mathbf{X}^*(t)) = \begin{bmatrix} \eta r^p(t) & 0 & 0 \\ 0 & \gamma\sqrt{v(t)} & 0 \\ \rho_{x,r}\sqrt{v(t)} & \rho_{x,v}\sqrt{v(t)} & \sqrt{1-\rho_{x,v}^2-\rho_{x,r}^2}\sqrt{v(t)} \end{bmatrix}. \tag{2.8}$$

2.2. Reformulated Heston hybrid models. In the previous section we have seen that for the HHW and HCIR models with a full matrix of correlations given in (2.3), the affinity relations [13] are not satisfied, so that the ChF cannot be obtained by standard techniques.

In order to obtain a well-defined Heston hybrid model with an *indirectly imposed correlation*, $\rho_{x,r}$, we propose the following system of SDEs:

$$\mathrm{d}S(t)/S(t) = r(t)\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t) + \Omega(t)r^p(t)\mathrm{d}W_r(t) + \Delta\sqrt{v(t)}\mathrm{d}W_v(t), S(0) > 0, \tag{2.9}$$

with

$$\begin{aligned} \mathrm{d}v(t) &= \kappa(\bar{v} - v(t))\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v(t), \quad v(0) > 0, \\ \mathrm{d}r(t) &= \lambda(\theta(t) - r(t))\mathrm{d}t + \eta r^p(t)\mathrm{d}W_r(t), \quad r(0) > 0, \end{aligned} \tag{2.10}$$

where

$$\mathrm{d}W_x(t)\mathrm{d}W_v(t) = \hat{\rho}_{x,v}\mathrm{d}t, \quad \mathrm{d}W_x(t)\mathrm{d}W_r(t) = 0, \quad \mathrm{d}W_v(t)\mathrm{d}W_r(t) = 0, \tag{2.11}$$

where $p = 0$ for HHW and $p = \frac{1}{2}$ for HCIR. We have included a function,[2] $\Omega(t)$, and a constant parameter, Δ. Note that we still assume independence between the instantaneous short rate, $r(t)$, and the variance process, $v(t)$, i.e., $\hat{\rho}_{r,v} = 0$.

By exchanging the order of the state variables to $\mathbf{X}^*(t) = [r(t), v(t), S(t)]^{\mathrm{T}}$, system (2.9) is given, in terms of the independent Brownian motions, by

$$\begin{bmatrix} \mathrm{d}r(t) \\ \mathrm{d}v(t) \\ \dfrac{\mathrm{d}S(t)}{S(t)} \end{bmatrix} = \begin{bmatrix} \lambda(\theta(t) - r(t)) \\ \kappa(\bar{v} - v(t)) \\ r(t) \end{bmatrix} \mathrm{d}t + \hat{\sigma}(\mathbf{X}^*(t)) \begin{bmatrix} \mathrm{d}\widetilde{W}_r(t) \\ \mathrm{d}\widetilde{W}_v(t) \\ \mathrm{d}\widetilde{W}_x(t) \end{bmatrix}, \tag{2.12}$$

[2]This under certain conditions can also be stochastic.

where

$$\hat{\sigma}(\mathbf{X}^*(t)) = \begin{bmatrix} \eta r^p(t) & 0 & 0 \\ 0 & \gamma\sqrt{v(t)} & 0 \\ \Omega(t) r^p(t) & \sqrt{v(t)}\,(\hat{\rho}_{x,v} + \Delta) & \sqrt{v(t)}\sqrt{1-\hat{\rho}_{x,v}^2} \end{bmatrix}. \tag{2.13}$$

In the following lemma we show that the model (2.9) is equivalent to the full-scale HHW model in (2.3), with nonzero correlation $\rho_{x,r}$.

Lemma 2.1. *Model* (2.9) *satisfies the system in* (2.3) *with nonzero correlation* $\rho_{x,r}$ *for*

$$\Omega(t) = \rho_{x,r}\frac{\sqrt{v(t)}}{r^p(t)}, \quad \hat{\rho}_{x,v}^2 = \rho_{x,v}^2 + \rho_{x,r}^2, \quad \Delta = \rho_{x,v} - \hat{\rho}_{x,v}, \tag{2.14}$$

where correlation $\hat{\rho}_{x,v}$ *is as in model* (2.9) *and* $\rho_{x,v}$ *as in model* (2.3).

Proof. We presented the two models (2.3) and (2.9) in terms of the independent Brownian motions, (2.7) and (2.12), respectively. By matching the appropriate coefficients in (2.7) and (2.12), we find that the following relations should hold:

$$\begin{cases} \Omega(t) r^p(t) S(t) = & \rho_{x,r}\sqrt{v(t)}S(t), \\ \sqrt{1-\hat{\rho}_{x,v}^2}\sqrt{v(t)}S(t) = & \sqrt{1-\rho_{x,v}^2-\rho_{x,r}^2}\sqrt{v(t)}S(t), \\ (\hat{\rho}_{x,v}+\Delta)\sqrt{v(t)}S(t) = & \rho_{x,v}\sqrt{v(t)}S(t). \end{cases} \tag{2.15}$$

By simplifying (2.15) the proof is finished. ∎

If results (2.14) were directly included in the main system (2.9), the affinity property of the system would be lost. So, in order to satisfy the affinity constraints, appropriate *approximations* need to be introduced.

2.3. Log-transform. Before going into the details of the approximations of the HHW and HCIR models, let us first find the dynamics for the log-transform of the reformulated Heston hybrid models. By applying Itô's lemma, model (2.9) in log-equity space, $x(t) = \log S(t)$, with a constant parameter, Δ, and a function, $\Omega(t)$, is given by

$$\begin{aligned} \mathrm{d}x(t) &= \left[r(t) - \frac{1}{2}\left(\Omega^2(t) r^{2p}(t) + v(t)\left(1 + \Delta^2 + 2\hat{\rho}_{x,v}\Delta\right)\right)\right]\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t) \\ &\quad + \Omega(t) r^p(t)\mathrm{d}W_r(t) + \Delta\sqrt{v(t)}\mathrm{d}W_v(t) \\ &= \left(r(t) - \frac{1}{2}v(t)\right)\mathrm{d}t + \sqrt{v(t)}\mathrm{d}W_x(t) + \Omega(t) r^p(t)\mathrm{d}W_r(t) + \Delta\sqrt{v(t)}\mathrm{d}W_v(t), \end{aligned}$$

because of (2.14).

For a given state vector $\mathbf{X}^*(t) = [r(t), v(t), x(t)]^{\mathrm{T}}$, the symmetric instantaneous covariance matrix (1.3) is given by

$$\mathbf{\Sigma} := \begin{bmatrix} \eta^2 r^{2p}(t) & 0 & \eta\Omega(t) r^{2p}(t) \\ * & \gamma^2 v(t) & \gamma v(t)\,(\hat{\rho}_{x,v} + \Delta) \\ * & * & \Omega^2(t) r^{2p}(t) + v(t)\left(1 + \Delta^2 + 2\hat{\rho}_{x,v}\Delta\right) \end{bmatrix}. \tag{2.16}$$

As we consider two cases for parameter $p = \{0, 1/2\}$, the affinity issue appears in only one term of matrix (2.16), namely, in element $(1,3)$:

$$\mathbf{\Sigma}_{(1,3)} = \eta\Omega(t)r^{2p}(t) = \eta\rho_{x,r}\sqrt{v(t)}r^p(t) = \begin{cases} \eta\rho_{x,r}\sqrt{v(t)} & \text{for HHW,} \\ \eta\rho_{x,r}\sqrt{v(t)}\sqrt{r(t)} & \text{for HCIR.} \end{cases} \tag{2.17}$$

Although term $\mathbf{\Sigma}_{(3,3)}$ does not seems to be of the affine form, by (2.14), it equals $\mathbf{\Sigma}_{(3,3)} = v(t)$, and therefore it is linear in the state variables.

Remark 1. We see that, in order to make either the HHW or the HCIR model affine, one does not necessarily need to approximate function $\Omega(t)$; only the nonaffine terms in the corresponding instantaneous covariance matrix[3] need be approximated. By approximation of the nonaffine covariance term, $\mathbf{\Sigma}_{(1,3)}$, the corresponding pricing PDE also changes. The Kolmogorov backward equation for the log-stock price (see, for example, [33]) is now given by

$$\begin{aligned} 0 = &\frac{\partial\phi}{\partial t} + \left(r - \frac{1}{2}v\right)\frac{\partial\phi}{\partial x} + \kappa(\bar{v} - v)\frac{\partial\phi}{\partial v} + \lambda(\theta(t) - r)\frac{\partial\phi}{\partial r} + \frac{1}{2}v\frac{\partial^2\phi}{\partial x^2} \\ &+ \frac{1}{2}\gamma^2 v\frac{\partial^2\phi}{\partial v^2} + \frac{1}{2}\eta^2 r^{2p}\frac{\partial^2\phi}{\partial r^2} + \rho_{x,v}\gamma v\frac{\partial^2\phi}{\partial x\partial v} + \mathbf{\Sigma}_{(1,3)}\frac{\partial^2\phi}{\partial x\partial r} - r\phi, \end{aligned} \tag{2.18}$$

subject to terminal condition $\phi(u, \mathbf{X}(T), T, T) = \exp(iux(T))$.

The derivations in section 2.3 show that system (2.9) is nothing but a reformulation of the original HHW system under the conditions in (2.14). It is therefore sufficient to linearize the nonaffine terms in the covariance matrix to determine an affine approximation of the full-scale model. In the sections to follow we discuss two possible approximations for $\mathbf{\Sigma}_{(1,3)}$.

3. Deterministic approximation for hybrid models.

In order to make the Heston hybrid model affine, we provide a first approximation for the expressions in (2.17) in section 3.1. The corresponding ChF is derived in section 3.2.

3.1. Deterministic approach: The H1-HW model.

The first approach to finding an approximation for the term $\mathbf{\Sigma}_{(1,3)} = \eta\rho_{x,r}\sqrt{v(t)}r^p(t)$ in matrix (2.16) is to replace it by its expectation; i.e.,

$$\mathbf{\Sigma}_{(1,3)} \approx \eta\rho_{x,r}\mathbb{E}\left(r^p(t)\sqrt{v(t)}\right) \overset{\perp\!\!\!\perp}{=} \eta\rho_{x,r}\mathbb{E}(r^p(t))\mathbb{E}(\sqrt{v(t)}), \tag{3.1}$$

assuming independence between $r(t)$ and $v(t)$.

The approximation for $\mathbf{\Sigma}_{(1,3)}$ in (3.1) consists of two expectations: one with respect to $\sqrt{v(t)}$ and another with respect to $r^p(t)$. $\mathbb{E}(r^p(t)) = 1$ for $p = 0$, and it is $\mathbb{E}(\sqrt{r(t)})$ for $p = 1/2$. Since the processes for $v(t)$ and $r(t)$ are then of the same type, the approximations are analogous. By taking the expectations of the stochastic variables, the model becomes of affine form, so that we can obtain the corresponding ChF.

In Lemma 3.1 the closed-form expressions for the expectation and variance of $\sqrt{v(t)}$ (a CIR-type process) are presented.

[3]The drifts and the interest rate are already in the affine form, presented in (1.2) and (1.4).

Lemma 3.1 (expectation and variance for CIR-type process). *For a given time $t > 0$ the expectation and variance of $\sqrt{v(t)}$, where $v(t)$ is a CIR-type process (2.1), are given by*

$$\mathbb{E}(\sqrt{v(t)}) = \sqrt{2c(t)}\mathrm{e}^{-\lambda(t)/2}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{\lambda(t)}{2}\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma(\frac{d}{2}+k)} \tag{3.2}$$

and

$$\mathbb{V}ar\left(\sqrt{v(t)}\right) = c(t)(d+\lambda(t)) - 2c(t)\mathrm{e}^{-\lambda(t)}\left(\sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{\lambda(t)}{2}\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma\left(\frac{d}{2}+k\right)}\right)^2, \tag{3.3}$$

where

$$c(t) = \frac{1}{4\kappa}\gamma^2(1-\mathrm{e}^{-\kappa t}),\quad d = \frac{4\kappa\bar{v}}{\gamma^2},\quad \lambda(t) = \frac{4\kappa v(0)\mathrm{e}^{-\kappa t}}{\gamma^2(1-\mathrm{e}^{-\kappa t})}, \tag{3.4}$$

with $\Gamma(k)$ being the gamma function defined by

$$\Gamma(k) = \int_0^{\infty} t^{k-1}\mathrm{e}^{-t}\mathrm{d}t.$$

Proof. By [14] one can find the closed-form expression for the expectation $\mathbb{E}(\sqrt{v(t)})$, which by the principle of Kummer [28] can be simplified. ∎

The analytic expression for the expectation, either of $\sqrt{v(t)}$ or $\sqrt{r(t)}$ in (3.1), is involved and requires rather expensive numerical operations.

In order to find a first-order approximation, we can apply the so-called *delta method* (see, for example, [1, 32]), which states that a function $\varphi(X)$ can be approximated by a first-order Taylor expansion at $\mathbb{E}(X)$, for a given random variable, X, with expectation, $\mathbb{E}(X)$, and variance, $\mathbb{V}\mathrm{ar}(X)$, assuming that for $\varphi(X)$ its first derivative with respect to X exists and is sufficiently smooth.

The result below provides details of the approximation.

Result 3.2. *The expectation, $\mathbb{E}(\sqrt{v(t)})$, with stochastic process $v(t)$ given by (2.3), can be approximated by*

$$\mathbb{E}(\sqrt{v(t)}) \approx \sqrt{c(t)(\lambda(t)-1) + c(t)d + \frac{c(t)d}{2(d+\lambda(t))}} =: \Lambda(t), \tag{3.5}$$

with $c(t)$, d, and $\lambda(t)$ given in Lemma 3.1, and where κ, $\bar{v}$, γ, and $v(0)$ are the parameters given in (2.3).

In order to find the approximation in Result 3.2, we can use the delta method as follows. Assuming the function φ to be sufficiently smooth and the first two moments of X to exist, we obtain by first-order Taylor expansion

$$\varphi(X) \approx \varphi(\mathbb{E}X) + (X-\mathbb{E}X)\frac{\partial\varphi}{\partial X}(\mathbb{E}X). \tag{3.6}$$

Since the variance of $\varphi(X)$ can be approximated by the variance of the right-hand side of (3.6), we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}(\varphi(X)) &\approx \mathbb{V}\mathrm{ar}\left(\varphi(\mathbb{E}X) + (X - \mathbb{E}X)\frac{\partial\varphi}{\partial X}(\mathbb{E}X)\right)\\ &= \left(\frac{\partial\varphi}{\partial X}(\mathbb{E}X)\right)^2 \mathbb{V}\mathrm{ar}X. \end{aligned} \tag{3.7}$$

Now, by using this result for function $\varphi(v(t)) = \sqrt{v(t)}$, we find

$$\mathbb{V}\mathrm{ar}(\sqrt{v(t)}) \approx \left(\frac{1}{2}\frac{1}{\sqrt{\mathbb{E}(v(t))}}\right)^2 \mathbb{V}\mathrm{ar}(v(t)) = \frac{1}{4}\frac{\mathbb{V}\mathrm{ar}(v(t))}{\mathbb{E}(v(t))}. \tag{3.8}$$

However, from the definition of the variance, we also have

$$\mathbb{V}\mathrm{ar}(\sqrt{v(t)}) = \mathbb{E}(v(t)) - \left(\mathbb{E}(\sqrt{v(t)})\right)^2, \tag{3.9}$$

and by combining (3.8) and (3.9) we obtain the following approximation:

$$\mathbb{E}(\sqrt{v(t)}) \approx \sqrt{\mathbb{E}(v(t)) - \frac{1}{4}\frac{\mathbb{V}\mathrm{ar}(v(t))}{\mathbb{E}(v(t))}}. \tag{3.10}$$

Since $v(t)$ is a square-root process, as in (2.9), we have

$$v(t) = v(0)\mathrm{e}^{-\kappa t} + \bar{v}(1 - \mathrm{e}^{-\kappa t}) + \gamma\int_0^t \mathrm{e}^{\kappa(s-t)}\sqrt{v(s)}\mathrm{d}W_v(s). \tag{3.11}$$

The expectation reads $\mathbb{E}(v(t)) = c(t)(d + \lambda(t))$, and for the variance we get $\mathbb{V}\mathrm{ar}(v(t)) = c^2(t)(2d + 4\lambda(t))$, with $c(t)$, d, and $\lambda(t)$ as given in (3.4).

Now, by substituting these expressions in (3.10), the result is proved. Since Result 3.2 provides an explicit approximation for $\mathbf{\Sigma}_{(1,3)}$ in (3.1) in terms of a deterministic function for $\mathbb{E}(\sqrt{v(t)})$, we are, in principle, able to derive the corresponding ChF.

We show here for which parameters the expression under the square root in approximation (3.5) is nonnegative.

Consider the following inequality:

$$c(t)(\lambda(t) - 1) + c(t)d + \frac{c(t)d}{2(d + \lambda(t))} \geq 0. \tag{3.12}$$

Division by $c(t) > 0$ gives

$$\frac{2\left(\lambda(t) + d\right)\left(d + \lambda(t)\right) + d}{2(d + \lambda(t))} \geq 1. \tag{3.13}$$

So,

$$2\left(\lambda(t) + d\right)^2 - 2(\lambda(t) + d) + d \geq 0. \tag{3.14}$$

By setting $y = \lambda(t) + d$ we find $2y^2 - 2y + d \geq 0$. The parabola is nonnegative for the discriminant $4 - 4 \cdot 2 \cdot d \leq 0$, so that the expression in (3.5) is nonnegative for $d \geq \frac{1}{2}$ (i.e., $2d \geq 1$). With $d = 4\kappa\bar{v}/\gamma^2$ we can compare the inequality obtained to the Feller condition. If the Feller condition is satisfied, the expression under the square root is certainly well defined. If $8\kappa\bar{v}/\gamma^2 \geq 1$ but the Feller condition is not satisfied, the approximation is also valid. If the expression under the square root in (3.5) becomes negative, we suggest using the expressions in Lemma 3.1 instead.

Remark 2. We assume that the first-order linear terms around the parameter values in the Taylor expansion give an accurate representation. However, this may not work satisfactorily for "flat" density functions, like those from a uniform distribution. In order to increase the accuracy, higher-order terms can be included in the expansion [1]. More discussion on the conditions for the delta method to perform well can be found in [32].

The approximation for $\mathbb{E}(\sqrt{v(t)})$ in (3.5) is still nontrivial and may cause difficulties when deriving the corresponding ChFs. In order to find the coefficients of the ChF, a routine for numerically solving the corresponding ODEs has to be incorporated. Numerical integration, however, slows down the option pricing engine and would make the SDE model less attractive. As we aim to find a closed-form expression for the ChF, we further simplify $\Lambda(t)$ in (3.5). Expectation $\mathbb{E}(\sqrt{v(t)})$ can be further approximated by a function of the following form:

$$\mathbb{E}(\sqrt{v(t)}) \approx a + b\mathrm{e}^{-ct} =: \widetilde{\Lambda}(t), \tag{3.15}$$

with a, b, and c constant. Appropriate values for a, b, and c in (3.15) can be obtained via an optimization problem of the form $\min_{a,b,c} ||\Lambda(t) - \widetilde{\Lambda}(t)||_n$, where $||\cdot||_n$ is any nth norm.

We propose here, instead of a numerical approximation for these coefficients, a simple analytic expression in Result 3.3.

Result 3.3. *By matching functions $\Lambda(t)$ and $\widetilde{\Lambda}(t)$ for $t \to +\infty$, $t \to 0$, and $t = 1$, we find*

$$\begin{aligned} \lim_{t\to+\infty} \Lambda(t) &= \sqrt{\bar{v} - \frac{\gamma^2}{8\kappa}} = a = \lim_{t\to+\infty} \widetilde{\Lambda}(t), \\ \lim_{t\to 0} \Lambda(t) &= \sqrt{v(0)} = a + b = \lim_{t\to 0} \widetilde{\Lambda}(t), \\ \lim_{t\to 1} \Lambda(t) &= \Lambda(1) = a + b\mathrm{e}^{-c} = \lim_{t\to 1} \widetilde{\Lambda}(t). \end{aligned} \tag{3.16}$$

The values a, b, and c can now be estimated by

$$a = \sqrt{\bar{v} - \frac{\gamma^2}{8\kappa}}, \quad b = \sqrt{v(0)} - a, \quad c = -\log\left(b^{-1}(\Lambda(1) - a)\right), \tag{3.17}$$

where $\Lambda(t)$ is given by (3.5).

The approximation given in Result 3.3 may give difficulties for $\bar{v} < \gamma^2/8\kappa$ in (3.17) (the expression under the square root then becomes negative). We recognize that this expression is well defined, as the expression under the square root in the function $\Lambda(t)$ in Result 3.2 is positive.

In order to measure the quality of approximation (3.17) to $\mathbb{E}(\sqrt{v(t)})$ in (3.2), we perform a numerical experiment (see the results in Figure 1). For randomly chosen sets of parameters the approximation (3.17) resembles $\mathbb{E}(\sqrt{v(t)})$ in (3.2) very closely.

We call the resulting model the H1-HW model (HHW model-1).

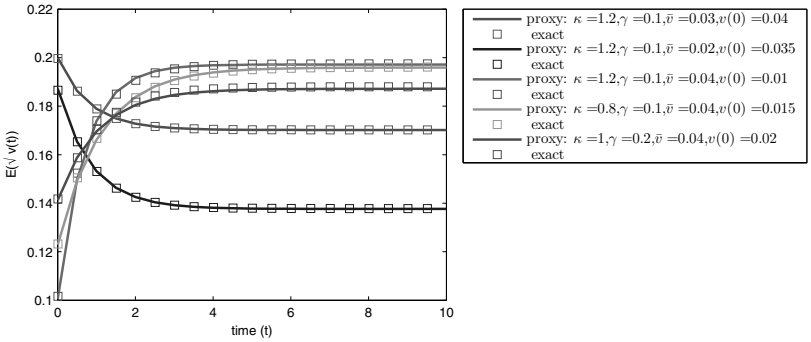

Figure 1. *The quality of the approximation* $\mathbb{E}(\sqrt{v(t)}) \approx a+b\mathrm{e}^{-ct}$ *(continuous line) versus the exact solution given in* (3.2) *(squares) for* 5 *random* κ, γ, $\bar{v}$, *and* $v(0)$.

3.1.1. The case $\Delta = 0$ and $\Omega(t) \equiv$ const. With $\Delta = 0$ in systems (2.9) and (2.12), the model resembles the one in [19, 3]. There, a constant parameter $\bar{\Omega} = \Omega(t)$ was prescribed, and an instantaneous correlation was imposed *indirectly*.

The following lemma, however, shows that this model with $\Delta = 0$ resembles the full-scale HHW and HCIR models only for correlation $\rho_{x,r} = 0$.

Lemma 3.4. *The hybrid models* (2.9) *with* $\Delta = 0$ *are full-scale HHW and HCIR models, in the sense of system* (2.3)*, only if the instantaneous correlation between the stock and the interest rate processes in system* (2.3) *equals zero, i.e.,* $\rho_{x,r} = 0$.

Proof. The proof is analogous to the proof of Lemma 2.1. We see from the equalities in (2.14) that system (2.7) resembles system (2.12) with $\Delta = 0$ only if

$$\bar{\Omega} = \rho_{x,r}\frac{\sqrt{v(t)}}{r^p(t)}, \quad \hat{\rho}_{x,v} = \rho_{x,v}, \quad \hat{\rho}^2_{x,v} = \rho^2_{x,v} + \rho^2_{x,r}. \tag{3.18}$$

The equations (3.18) hold only for $\rho_{x,r} = 0$. So, the models with $\Delta = 0$ are not full-scale HHW and HCIR models with nonzero correlation $\rho_{x,r}$. ∎

Although the model with $\Delta = 0$ is not a properly defined Heston hybrid model, one can still proceed with the analysis. Parameter $\bar{\Omega}$ was derived based on the following equality (see [19]), using the definition of the instantaneous correlation:

$$\hat{\rho}_{x,r} = \frac{\mathbb{E}\left(\mathrm{d}S(t)\mathrm{d}r(t)\right) - \mathbb{E}(\mathrm{d}S(t))\mathbb{E}(\mathrm{d}r(t))}{\sqrt{v(t)S^2(t)\,\mathrm{d}t + \bar{\Omega}^2 r^{2p}(t)S^2(t)\,\mathrm{d}t}\sqrt{\eta^2 r^{2p}(t)\,\mathrm{d}t}} = \frac{\bar{\Omega}r^p(t)}{\sqrt{v(t) + \bar{\Omega}^2 r^{2p}(t)}}. \tag{3.19}$$

To deal with the affinity issue a constant approximation for $\bar{\Omega}$ was proposed, given by

$$\bar{\Omega} \approx \frac{\hat{\rho}_{x,r}}{\sqrt{1-\hat{\rho}^2_{x,r}}}\mathbb{E}\left(\frac{1}{T}\int_0^T v(t)\mathrm{d}t\right)^{\frac{1}{2}} \Big/ \mathbb{E}\left(\frac{1}{T}\int_0^T r(t)\mathrm{d}t\right)^p. \tag{3.20}$$

By choosing $\bar{\Omega} = 0$ the model collapses to the well-known HHW model ($p = 0$) or the HCIR model ($p = \frac{1}{2}$) with zero correlation $\rho_{x,r}$.

In Figure 2 we present the behavior of the instantaneous correlation between the equity and the interest rates. We see that for time-dependent $\Omega(t)$, as defined in Lemma 2.1, the instantaneous correlations are stable and oscillate around the exact value, chosen to be $\rho_{x,r} = 60\%$, whereas for the model with $\Omega(t) = \bar{\Omega}$ a different correlation pattern is observed. For the latter model, initially the correlation is significantly higher than 60%, and it decreases in time. These results show that a constant $\bar{\Omega}$ in the model with $\Delta = 0$ may give an average correlation close to the exact value, although the instantaneous correlation is not stable in time.

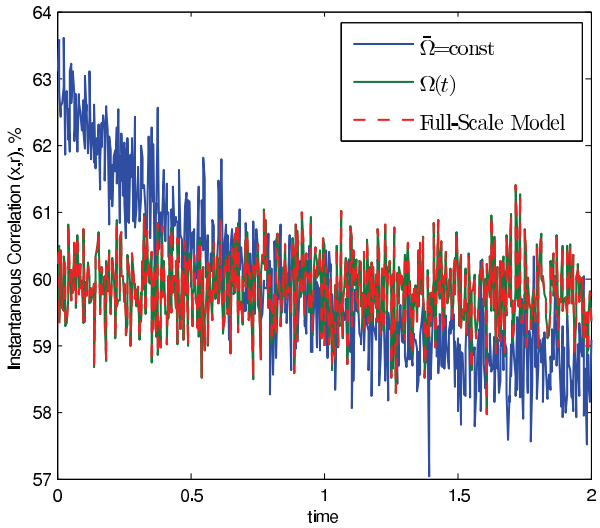

Figure 2. *The instantaneous correlations for different models. The blue line represents the model with $\Delta = 0$ with constant $\bar{\Omega}$, the dotted-red line corresponds to the full-scale HHW model, and the green line corresponds to the model with time-dependent $\Omega(t)$. Maturity is chosen to $\tau = 2$ years.*

The assumptions of constant $\bar{\Omega}$ and $\Delta = 0$ also have an impact on the corresponding pricing PDE. With the Feynman–Kac theorem the corresponding PDE is given by

$$0 = \frac{\partial \phi}{\partial t} + \left[r - \frac{1}{2}\left(v + r^{2p}\bar{\Omega}^2\right)\right]\frac{\partial \phi}{\partial x} + \kappa(\bar{v} - v)\frac{\partial \phi}{\partial v} + \lambda(\theta(t) - r)\frac{\partial \phi}{\partial r} + \frac{1}{2}\left(v + r^{2p}\bar{\Omega}^2\right)\frac{\partial^2 \phi}{\partial x^2}$$

$$+\frac{1}{2}\gamma^2 v\frac{\partial^2 \phi}{\partial v^2} + \frac{1}{2}\eta^2 r^{2p}\frac{\partial^2 \phi}{\partial r^2} + \hat{\rho}_{x,v}\gamma v\frac{\partial^2 \phi}{\partial x \partial v} + \eta\bar{\Omega}r^{2p}\frac{\partial^2 \phi}{\partial x \partial r} - r\phi, \tag{3.21}$$

with the same terminal condition as for (2.18). The assumption of constant $\bar{\Omega}$ and $\Delta = 0$ gives rise to additional terms in the convection and diffusion parts of PDE (3.21).

By means of a numerical experiment, we check the accuracy of the model with $\Delta = 0$ and determine whether the model approximates the full-scale HHW hybrid model sufficiently well. We consider here the following set of parameters: $S(0) = 1$, $\kappa = 2$, $v(0) = \bar{v} = 0.05$, $\gamma = 0.1$, $\lambda = 1.2$, $r(0) = \theta = 0.05$, $\eta = 0.01$, and correlation $\rho_{x,v} = -40\%$. In the simulation we choose two different values for correlation $\rho_{x,r} = \{30\%, 50\%\}$.

We compare the following three models: the full-scale HHW model (with MC simulation), the model with $\Delta = 0$, and our approximation for $\mathbf{\Sigma}_{(1,3)}$ in PDE (2.18) with the projection according to (3.1).

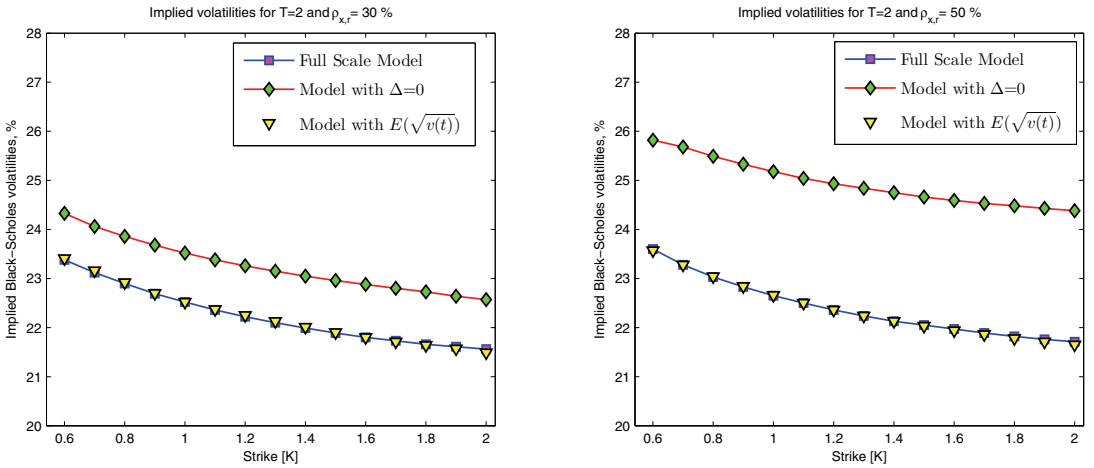

Figure 3. *The implied Black–Scholes volatilities for the full-scale Heston model and two approximations: deterministic approach (model (2.18) with (3.1)), and model with* $\Delta = 0$ *(model (3.21)).*

In Figure 3 the implied volatilities obtained are compared. The model with $\Delta = 0$ in (3.21) does not provide a satisfactory fit to the full-scale HHW model, whereas the implied volatilities obtained with the deterministic hybrid approximation compare very well (they essentially overlap) with the full-scale reference results; see Figure 3. The *volatility compensator* Δ, as defined in Lemma 2.1, cannot be neglected when approximating the full-scale HHW model, as was stated in Lemma 3.4.

3.2. ChF for the H1-HW model. We derive a ChF for the HHW hybrid model given in (2.18). For $p = 0$, the nonaffine term, $\mathbf{\Sigma}_{(1,3)}$, in matrix (2.18) equals $\mathbf{\Sigma}_{(1,3)} = \eta\rho_{x,r}\sqrt{v(t)}$ and will be approximated by $\mathbf{\Sigma}_{(1,3)} \approx \eta\rho_{x,r}\mathbb{E}(\sqrt{v(t)})$.

We assume here that the term structure for the interest rate $\theta(t)$ is constant: $\theta(t) = \theta$. A generalization can be found in [8].

According to [13], the discounted ChF for the H1-HW model is of the following form:

$$\phi_{\text{H1-HW}}(u, \mathbf{X}(t), \tau) = \exp\left(A(u,\tau) + B(u,\tau)x(t) + C(u,\tau)r(t) + D(u,\tau)v(t)\right), \tag{3.22}$$

with final conditions $A(u,0) = 0$, $B(u,0) = iu$, $C(u,0) = 0$, $D(u,0) = 0$, and $\tau := T - t$.

The ChF for the H1-HW model can be derived in closed form, with the help of the following lemmas.

Lemma 3.5 (ODEs related to the H1-HW model). *The functions* $B(u,\tau) =: B(\tau)$, $C(u,\tau) =: C(\tau)$, $D(u,\tau) =: D(\tau)$, *and* $A(u,\tau) =: A(\tau)$ *for* $u \in \mathbb{R}$ *and* $\tau \geq 0$ *in (3.22) for the H1-HW model satisfy the following system of ODEs:*

$$\begin{aligned}
&B'(\tau) = 0, \quad B(u,0) = iu,\\
&C'(\tau) = -1 - \lambda C(\tau) + B(\tau), \quad C(u,0) = 0,\\
&D'(\tau) = B(\tau)(B(\tau)-1)/2 + (\gamma\rho_{x,v}B(\tau) - \kappa)\,D(\tau) + \gamma^2 D^2(\tau)/2, \quad D(u,0) = 0,\\
&A'(\tau) = \lambda\theta C(\tau) + \kappa\bar{v}D(\tau) + \eta^2 C^2(\tau)/2 + \eta\rho_{x,r}\mathbb{E}(\sqrt{v(t)})B(\tau)C(\tau), \quad A(u,0) = 0,
\end{aligned}$$

with $\tau = T - t$, and where κ, λ, and θ and η, $\rho_{x,r}$, and $\rho_{x,v}$ correspond to the parameters in the HHW model (2.3).

Proof. The proof can be found in section A.1. ∎

The following lemma gives the closed-form solution for the functions $B(u,\tau)$, $C(u,\tau)$, $D(u,\tau)$, and $A(u,\tau)$ in (3.22).

Lemma 3.6 (ChF for the H1-HW model). *The solution of the ODE system in Lemma 3.5 is given by*

$$B(u,\tau) = iu, \tag{3.23}$$

$$C(u,\tau) = (iu-1)\lambda^{-1}(1-\mathrm{e}^{-\lambda\tau}), \tag{3.24}$$

$$D(u,\tau) = \frac{1-\mathrm{e}^{-D_1\tau}}{\gamma^2\left(1-g\mathrm{e}^{-D_1\tau}\right)}\left(\kappa-\gamma\rho_{x,v}iu-D_1\right), \tag{3.25}$$

$$A(u,\tau) = \lambda\theta I_1(\tau) + \kappa\bar{v}I_2(\tau) + \frac{1}{2}\eta^2 I_3(\tau) + \eta\rho_{x,r}I_4(\tau), \tag{3.26}$$

with $D_1 = \sqrt{(\gamma\rho_{x,v}iu-\kappa)^2 - \gamma^2 iu(iu-1)}$, and where $g = \frac{\kappa-\gamma\rho_{x,v}iu-D_1}{\kappa-\gamma\rho_{x,v}iu+D_1}$, κ, θ, λ, and γ are as in (2.10).

The integrals $I_1(\tau)$, $I_2(\tau)$, and $I_3(\tau)$ admit an analytic solution, and $I_4(\tau)$ admits a semianalytic solution:

$$\begin{aligned}
I_1(\tau) &= \frac{1}{\lambda}(iu-1)\left(\tau + \frac{1}{\lambda}(\mathrm{e}^{-\lambda\tau}-1)\right),\\
I_2(\tau) &= \frac{\tau}{\gamma^2}\left(\kappa-\gamma\rho_{x,v}iu-D_1\right) - \frac{2}{\gamma^2}\log\left(\frac{1-g\mathrm{e}^{-D_1\tau}}{1-g}\right),\\
I_3(\tau) &= \frac{1}{2\lambda^3}(i+u)^2\left(3+\mathrm{e}^{-2\lambda\tau}-4\mathrm{e}^{-\lambda\tau}-2\lambda\tau\right),\\
I_4(\tau) &= iu\int_0^\tau \mathbb{E}(\sqrt{v(T-s)})C(u,s)\mathrm{d}s\\
&= -\frac{1}{\lambda}(iu+u^2)\int_0^\tau \mathbb{E}(\sqrt{v(T-s)})\left(1-\mathrm{e}^{-\lambda s}\right)\mathrm{d}s.
\end{aligned}$$

Proof. The proof can be found in section A.2. ∎

Note that by taking $\mathbb{E}(\sqrt{v(T-s)}) \approx a + b\mathrm{e}^{-c(T-s)}$, with a, b, and c as given in (3.15), we obtain a closed-form expression:

$$I_4(\tau) = -\frac{1}{\lambda}(iu+u^2)\left[\frac{b}{c}\left(\mathrm{e}^{-ct}-\mathrm{e}^{-cT}\right) + a\tau + \frac{a}{\lambda}\left(\mathrm{e}^{-\lambda\tau}-1\right) + \frac{b}{c-\lambda}\mathrm{e}^{-cT}\left(1-e^{-\tau(\lambda-c)}\right)\right].$$

In Appendix B we present the generalization to a *full matrix of nonzero correlations* between the processes.

4. Stochastic approximation for hybrid models. In the previous section a rather straightforward way to approximate the nonaffine elements in the instantaneous covariance matrix was presented. Here, we model those elements alternatively by stochastic processes and call the resulting approximate model H2-HW (HHW model-2).

4.1. Stochastic approach: The H2-HW model. In the result below an approximation for finite time t and a nonzero centrality parameter is presented.

Result 4.1 (normal approximation for $\sqrt{v(t)}$, for $0 < t < \infty$). *For any time $t < \infty$, the square root of $v(t)$ in (2.9) can be approximated by*

$$\sqrt{v(t)} \approx \mathcal{N}\left(\sqrt{c(t)(\lambda(t)-1)+c(t)d+\frac{c(t)d}{2(d+\lambda(t))}}, c(t)-\frac{c(t)d}{2(d+\lambda(t))}\right), \tag{4.1}$$

with $c(t)$, d, and $\lambda(t)$ from (3.4). Moreover, for a fixed value of x in the cumulative distribution function $F_{\sqrt{v(t)}}(x)$ and a fixed value for parameter d, the error is of order $\mathcal{O}(\lambda^2(t))$ for $\lambda(t) \to 0$ and $\mathcal{O}(\lambda(t)^{-\frac{1}{2}})$ for $\lambda(t) \to \infty$.

To show the validity of the approximations presented above, we follow Patnaik in [34], who found that an accurate approximation for the noncentral chi-square distribution, $\chi^2_d(\lambda(t))$, can be obtained by an approximation with a centralized chi-square distribution, i.e.,

$$\chi^2(d,\lambda(t)) \approx a(t)\chi^2(f(t)), \tag{4.2}$$

with $a(t)$ and $f(t)$ in (4.2) chosen so that the first two moments match, i.e.,

$$a(t)=\frac{d+2\lambda(t)}{d+\lambda(t)}, \quad f(t)=d+\frac{\lambda^2(t)}{d+2\lambda(t)}. \tag{4.3}$$

It was shown in [11, 9] that, for a given time $t > 0$, $v(t)$ is distributed as $c(t)$ times a noncentral chi-squared random variable, $\chi^2(d,\lambda(t))$, with degrees of freedom parameter d and noncentrality parameter $\lambda(t)$, i.e., $v(t) = c(t)\chi^2\left(d,\lambda(t)\right),\ t > 0$. By combining this with (4.2) we have

$$\sqrt{v(t)} \approx \sqrt{c(t)}\sqrt{a(t)\chi^2(f(t))}. \tag{4.4}$$

Now, we use a result by Fisher [17] that for a given central chi-square random variable, $\chi^2(d)$, the expression $\sqrt{2\chi^2(d)}$ is approximately normally distributed with mean $\sqrt{2d-1}$ and unit variance, i.e.,

$$F_{\chi^2(d)}(x) \approx \Phi\left(\sqrt{2x}-\sqrt{2d-1}\right), \tag{4.5}$$

which implies

$$\sqrt{v(t)} \approx \mathcal{N}\left(\sqrt{\left(f(t)-\frac{1}{2}\right)c(t)a(t)}, \frac{1}{2}c(t)a(t)\right). \tag{4.6}$$

The order of this approximation can be found in [26].

Remark 3. Also in [34] it was indicated that the normal approximation resembles the noncentral chi-square distribution very well for either a large number of degrees of freedom, d, or a large noncentrality, $\lambda(t)$. For $t \to 0$, the noncentrality parameter, $\lambda(t)$, tends to infinity. Therefore, accurate approximations are expected.

In the case of long maturities, the noncentrality parameter converges to 0, which may give an inaccurate approximation. In this case, satisfactory results depend on the size of the degrees of freedom parameter d. It is clear that d in (3.4) is directly related to the Feller condition. In practical applications, however, $2\kappa\bar{v}$ is often smaller than γ^2. In the numerical experiments to follow, we will study the impact of *not* satisfying the Feller condition.

In Result 4.1 we have shown that $\sqrt{v(t)}$ can be well approximated by a normally distributed random variable. As the application of Itô's lemma to find the dynamics for $\sqrt{v(t)}$ is not allowed (the square-root process is not twice differentiable at the origin [25]), we construct here a stochastic process, $\xi(t)$, so that *equality in distribution* holds, i.e., $\xi(t) \overset{\mathrm{d}}{\approx} \sqrt{v(t)}$. Since a normal random variable is completely described by the first two moments, we need to ensure that $\mathbb{E}(\xi(t)) = \mathbb{E}(\sqrt{v(t)})$ and $\mathbb{V}\mathrm{ar}(\xi(t)) = \mathbb{V}\mathrm{ar}(\sqrt{v(t)})$. For this purpose we propose the following dynamics:

$$\mathrm{d}\xi(t) = \mu^{\xi}(t)\mathrm{d}t + \psi^{\xi}(t)\mathrm{d}W_v(t), \quad \xi(0) = \sqrt{v(0)}, \tag{4.7}$$

with some deterministic, time-dependent functions $\mu^{\xi}(t)$ and $\psi^{\xi}(t)$, determined so that the first two moments match. By moment matching, the unknown functions $\mu^{\xi}(t)$ and $\psi^{\xi}(t)$ in (4.7) read as

$$\mu^{\xi}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(\sqrt{v(t)}), \quad \psi^{\xi}(t) = \sqrt{\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{V}\mathrm{ar}(\sqrt{v(t)})}. \tag{4.8}$$

Using the results from Lemma 3.1, the expectation, $\mathbb{E}(\sqrt{v(t)})$, and the variance, $\mathbb{V}\mathrm{ar}(\sqrt{v(t)})$, can be derived:

$$\begin{aligned}\mu^{\xi}(t) = \frac{1}{2\sqrt{2}}\frac{\Gamma\left(\frac{1+d}{2}\right)}{\sqrt{c(t)}}\Bigg[{}_1\widetilde{F}_1\left(-\frac{1}{2},\frac{d}{2},-\frac{\lambda(t)}{2}\right)\frac{1}{2}\gamma^2\mathrm{e}^{-\kappa t} \\ + {}_1\widetilde{F}_1\left(\frac{1}{2},\frac{2+d}{2},-\frac{\lambda(t)}{2}\right)\frac{v(0)\kappa}{1-\mathrm{e}^{\kappa t}}\Bigg],\end{aligned}$$

$$\psi^{\xi}(t) = \left(\kappa(\bar{v}-v(0))\mathrm{e}^{-\kappa t} - 2\mathbb{E}(\sqrt{v(t)})\,\mu^{\xi}(t)\right)^{\frac{1}{2}}. \tag{4.9}$$

Here, $\mathbb{E}(\sqrt{v(t)})$ and d, $c(t)$, and $\lambda(t)$ are as in (3.2), and the regularized hypergeometric function ${}_1\widetilde{F}_1(a;b;z) =: {}_1F_1(a;b;z)/\Gamma(b)$.

The expressions for $\mu^{\xi}(t)$ and $\psi^{\xi}(t)$ in (4.9) are exact. However, since those expressions are not cheap to compute, one can find suitable approximations based on the results in Result 3.2, which are, however, not guaranteed to be well defined for all sets of parameters.

Since the approximate hybrid models are to be used for the calibration to European-style options (with one terminal payment), we do not need pathwise equality between processes $\xi(t)$ and $\sqrt{v(t)}$; only equality in terminal distribution is needed.

Remark 4. In section 4.1 we projected $\sqrt{v(t)}$ onto a normal process, $\xi(t)$. As is common with approximations by normal processes (a nonnegative random variable is projected onto another variable $\in \mathbb{R}$), this approximation comes with an error (as we indicated in Result 4.1). During stress testing, examples of which are presented in section 4.3 and in Appendix C, we did

not encounter any problems with this approximation. Typically, the stochastic approximation is somewhat more accurate than the deterministic approach (which is not based on a normal approximation).[4]

4.2. ChF for the H2-HW model. We now use the (stochastic) approximation for the term $\mathbf{\Sigma}_{(1,3)}$, with the process $\mathrm{d}\xi(t)$ given by (4.7), and the time-dependent functions $\mu^{\xi}(t)$ and $\psi^{\xi}(t)$ as in (4.9).

This approximation gives rise to an extension of the three-dimensional space variable $\mathbf{X}(t) = [S(t), v(t), r(t)]^{\mathrm{T}}$ to a four-dimensional space $\widetilde{\mathbf{X}}(t) = [S(t), v(t), r(t), \xi(t)]^{\mathrm{T}}$, with the following system of SDEs:

$$\left\{\begin{array}{rrrl} \mathrm{d}S(t)/S(t) = & r(t)\mathrm{d}t + & \sqrt{v(t)}\mathrm{d}W_x(t), & S(0) > 0, \\ \mathrm{d}v(t) = & \kappa(\bar{v} - v(t))\mathrm{d}t + & \gamma\sqrt{v(t)}\mathrm{d}W_v(t), & v(0) > 0, \\ \mathrm{d}r(t) = & \lambda(\theta(t) - r(t))\mathrm{d}t + & \eta\mathrm{d}W_r(t), & r(0) > 0, \\ \mathrm{d}\xi(t) = & \mu^{\xi}(t)\mathrm{d}t + & \psi^{\xi}(t)\mathrm{d}W_v(t), & \xi(0) = \sqrt{v(0)}, \end{array}\right. \tag{4.10}$$

where

$$\left\{\begin{array}{rcl} \mathrm{d}W_x(t)\mathrm{d}W_v(t) & = & \rho_{x,v}\mathrm{d}t, \\ \mathrm{d}W_x(t)\mathrm{d}W_r(t) & = & \rho_{x,r}\mathrm{d}t, \\ \mathrm{d}W_v(t)\mathrm{d}W_r(t) & = & 0, \end{array}\right. \tag{4.11}$$

with $\sqrt{v(t)} \approx \xi(t)$ and $\mu^{\xi}(t)$,$\psi^{\xi}(t)$ as defined in (4.9).

By taking the log-transform, $x(t) = \log S(t)$, in the model above all the drift terms are linear, and the symmetric instantaneous covariance matrix, with $\xi(t) \approx \sqrt{v(t)}$, is given by

$$\widetilde{\mathbf{\Sigma}} = \begin{bmatrix} v(t) & \gamma\rho_{x,v}v(t) & \rho_{x,r}\eta\xi(t) & \rho_{x,v}\psi^{\xi}(t)\xi(t) \\ * & \gamma^2 v(t) & 0 & \gamma\psi^{\xi}(t)\xi(t) \\ * & * & \eta^2 & 0 \\ * & * & * & \left(\psi^{\xi}(t)\right)^2 \end{bmatrix}, \tag{4.12}$$

which, since $\psi^{\xi}(t)$ is a deterministic time-dependent function, is now affine.

Since the system of SDEs (4.10) is affine, we derive the corresponding ChF:

$$\phi_{\text{H2-HW}}(u, \mathbf{X}(t), \tau) = \exp\left(A(u,\tau) + B(u,\tau)x(t) + C(u,\tau)r(t) + D(u,\tau)v(t) + E(u,\tau)\xi(t)\right), \tag{4.13}$$

with final conditions $\phi_{\text{H2-HW}}(u, \mathbf{X}(T), 0) = \exp(iux(T))$ and $\xi(t) \approx \sqrt{v(t)}$.

The functions $A(u,\tau)$, $B(u,\tau)$, $C(u,\tau)$, $D(u,\tau)$, and $E(u,\tau)$ satisfy the complex-valued ODEs given by the following lemma.

Lemma 4.2 (ODEs related to the H2-HW model). *The functions* $B(u,\tau) =: B(\tau)$, $C(u,\tau) := C(\tau)$, $D(u,\tau) =: D(\tau)$, $E(u,\tau) =: E(\tau)$, *and* $A(u,\tau) =: A(\tau)$ *for* $u \in \mathbb{R}$

[4]The method by Antonov from [4, 5] is also not based on normal approximations.

and $\tau = T - t > 0$ *in* (4.13) *satisfy*

$$
\begin{aligned}
&B'(\tau) = 0, \quad B(u,0) = iu,\\
&C'(\tau) = -1 + B(\tau) - \lambda C(\tau), \quad C(u,0) = 0,\\
&D'(\tau) = (B(\tau)-1)\,B(\tau)/2 + (\gamma\rho_{x,v}B(\tau) - \kappa)\,D(\tau) + \gamma^2 D^2(\tau)/2, \quad D(u,0) = 0,\\
&E'(\tau) = \rho_{x,r}\eta B(\tau)C(\tau) + \psi^{\xi}(t)\rho_{x,v}B(\tau)E(\tau) + \gamma\psi^{\xi}(t)D(\tau)E(\tau), \quad E(u,0) = 0,\\
&A'(\tau) = \kappa\bar{v}D(\tau) + \lambda\theta C(\tau) + \mu^{\xi}(t)E(\tau) + \eta^2 C^2(\tau)/2 + \left(\psi^{\xi}(t)\right)^2 E^2(\tau)/2, \quad A(u,0) = 0,
\end{aligned}
$$

with $\mu^{\xi}(t)$, $\psi^{\xi}(t)$ *as given in* (4.9).

Proof. The proof is very similar to the proof of Lemma 3.5. ∎

Solutions to the ODEs for $B(u,\tau)$, $C(u,\tau)$, and $D(u,\tau)$ can be found in Lemma 3.6, where the deterministic linearization was applied.

Note that the remaining two functions, $E(u,\tau)$ and $A(u,\tau)$, contain the rather complicated functions $\mu^{\xi}(t)$ and $\psi^{\xi}(t)$. We leave these equations to be solved numerically by a basic ODE routine.

A detailed analysis of the properties of the ChF will be studied in a followup article with theoretical research.

4.3. Numerical experiment. Here we check the performance of the deterministic (section 3.2) and the stochastic (section 4.2) approximations to the full-scale HHW model in terms of differences in implied volatilities. The HHW benchmark prices were obtained by MC simulation, as in [2].

In Table 1 we present the errors for Black–Scholes implied volatilities, $\epsilon(\rho_{x,r})$, for different correlations between the stock, $S(t)$, and the short rate, $r(t)$, and different strikes. We show results for a maturity of 10 years, $\tau = 10$, and for parameters that do not satisfy the Feller condition.[5]

Both approximations give very similar, highly accurate results for low correlations, $\rho_{x,r}$. This is different for high values of $\rho_{x,r}$. The deterministic approach generates somewhat more bias for high strikes, whereas the stochastic approach is essentially bias-free. The errors presented in Table 1 depend on the size of the volatility parameter of the interest rate process, η. For very low volatility, the two approximations provide a similar level of accuracy. As the volatility of the short rate process increases, a higher accuracy is expected for the stochastic approximation.

Calibration results will be presented in section 6. The performance of the methods developed is also presented in Appendix C, where our schemes are compared to the Markovian projection method [5].

5. HCIR hybrid model. We also present the ChF for an HCIR hybrid model, $p = 1/2$ in (2.3), which is more involved than the HW-based hybrid models. In the HCIR model the nonaffine term is given in (2.17). Again we use two approximations to obtain the ChF. In the first model, H1-CIR, we use the deterministic setup, and for the second model, H2-CIR, we determine the stochastic approximation.

[5]For short maturities, $\tau < 10$, and for model parameters for which the Feller condition is satisfied, we did not find any significant differences between the two approximations and the full-scale model.

Table 1

The implied volatilities and errors for the deterministic approximation (Approx 1) from (2.18) with approximation (3.2) and the stochastic approximation (Approx 2) from section 4.1 of the HHW model compared to the MC simulation performed with $20T$ steps and 100.000 paths. The error is defined as a difference between reference implied volatilities and the approximations. The parameters were chosen as $\kappa = 0.3$, $\gamma = 0.6$, $v(0) = \bar{v} = 0.05$, $\lambda = 0.01$, $r(0) = \theta = 0.02$, $\eta = 0.01$, and $S(0) = 100$ and the correlations as $\rho_{x,v} = -30\%$ and $\rho_{x,r} \in \{20\%, 60\%\}$. Numbers in brackets indicate standard deviations.

$\rho_{x,r}$	Strike	MC Imp. vol. [%]	Approx 1	Approx 2	Err. 1	Err. 2
20%	40	26.26 (0.22)	25.87	25.99	0.39	0.27
	80	20.07 (0.22)	20.03	20.02	0.04	0.05
	100	18.43 (0.24)	18.55	18.36	−0.12	0.07
	120	17.51 (0.20)	17.74	17.42	−0.23	0.09
	180	17.40 (0.22)	17.55	17.36	−0.15	0.04
60%	40	26.27 (0.14)	26.21	26.61	0.06	−0.34
	80	20.59 (0.11)	21.00	20.91	−0.41	−0.32
	100	19.11 (0.10)	19.84	19.22	−0.72	−0.10
	120	18.31 (0.10)	19.21	18.18	−0.90	0.13
	180	18.25 (0.11)	18.92	18.34	−0.67	−0.09

5.1. ChF for the H1-CIR model. The dynamics for the stock, $S(t)$, in the HCIR model read as

$$\begin{cases} \mathrm{d}S(t)/S(t) = & r(t)\mathrm{d}t + & \sqrt{v(t)}\mathrm{d}W_x(t), & S(0) > 0, \\ \mathrm{d}v(t) = & \kappa(\bar{v} - v(t))\mathrm{d}t + & \gamma\sqrt{v(t)}\mathrm{d}W_v(t), & v(0) > 0, \\ \mathrm{d}r(t) = & \lambda(\theta(t) - r(t))\mathrm{d}t + & \eta\sqrt{r(t)}\mathrm{d}W_r(t), & r(0) > 0, \end{cases} \tag{5.1}$$

with $\mathrm{d}W_x(t)\mathrm{d}W_v(t) = \rho_{x,v}\mathrm{d}t$, $\mathrm{d}W_x(t)\mathrm{d}W_r(t) = \rho_{x,r}\mathrm{d}t$, and $\mathrm{d}W_v(t)\mathrm{d}W_r(t) = 0$.

Here, we assume that the nonaffine term in the pricing PDE (2.18), $\mathbf{\Sigma}_{(1,3)}$, in (2.17) can be approximated as

$$\mathbf{\Sigma}_{(1,3)} \approx \eta\rho_{x,r}\mathbb{E}\left(\sqrt{r(t)}\sqrt{v(t)}\right) \overset{\perp\!\!\!\perp}{=} \eta\rho_{x,r}\mathbb{E}(\sqrt{r(t)})\mathbb{E}(\sqrt{v(t)}). \tag{5.2}$$

Since the processes involved are of the same type, the expectations in (5.2) can be determined as presented in section 3.1. For the log-stock, $x(t) = \log S(t)$, the ChF and the corresponding Riccati ODEs are defined as below:

$$\phi_{\text{H1-CIR}}(u, \mathbf{X}(t), \tau) = \exp\left(A(u,\tau) + B(u,\tau)x(t) + C(u,\tau)r(t) + D(u,\tau)v(t)\right). \tag{5.3}$$

Lemma 5.1 (ODEs related to the H1-CIR model). *The functions $B(u,\tau) =: B(\tau)$, $C(u,\tau) =: C(\tau)$, $D(u,\tau) =: D(\tau)$, and $A(u,\tau) =: A(\tau)$ for $u \in \mathbb{R}$ and $\tau > 0$ in (5.3) satisfy*

$$\begin{aligned} &B'(\tau) = 0, \quad B(u,0) = iu, \\ &C'(\tau) = -1 + B(\tau) - \lambda C(\tau) + \eta^2 C^2(\tau)/2, \quad C(u,0) = 0, \\ &D'(\tau) = (B(\tau) - 1)\, B(\tau)/2 + (\gamma\rho_{x,v}B(\tau) - \kappa)\, D(\tau) + \gamma^2 D^2(\tau)/2, \quad D(u,0) = 0, \\ &A'(\tau) = \kappa\bar{v}D(\tau) + \lambda\theta C(\tau) + \eta\rho_{x,r}\mathbb{E}(\sqrt{v(t)})\mathbb{E}(\sqrt{r(t)})B(\tau)C(\tau), \quad A(u,0) = 0, \end{aligned} \tag{5.4}$$

with $\tau = T - t$, $\mathbb{E}(\sqrt{v(t)})$, *and* $\mathbb{E}(\sqrt{r(t)})$ *from Lemma* 3.1.

Proof. The proof is very similar to the proof of Lemma 3.5 in section A.1. ∎

Lemma 5.2 (solutions for the ChF coefficients of the H1-CIR model). *The solutions for the ODEs for* $B(u,\tau)$, $C(u,\tau)$, $D(u,\tau)$, *and* $A(u,\tau)$, *defined in Lemma* 5.1, *are given by*

$$B(u,\tau) = iu, \tag{5.5}$$

$$C(u,\tau) = \frac{1-\mathrm{e}^{-D_1\tau}}{\eta^2\left(1-G_1\mathrm{e}^{-D_1\tau}\right)}\left(\lambda - D_1\right), \tag{5.6}$$

$$D(u,\tau) = \frac{1-\mathrm{e}^{-D_2\tau}}{\gamma^2\left(1-G_2\mathrm{e}^{-D_2\tau}\right)}\left(\kappa - \gamma\rho_{x,v}iu - D_2\right), \tag{5.7}$$

and

$$A(u,\tau) = \int_0^\tau \left(\kappa\bar{v}D(u,s) + \lambda\theta C(u,s) + \rho_{x,r}\eta iu\mathbb{E}(\sqrt{v(T-s)})\mathbb{E}(\sqrt{r(T-s)})C(u,s)\right)\mathrm{d}s,$$

with $D_1 = \sqrt{\lambda^2 + 2\eta^2(1-iu)}$, $D_2 = \sqrt{\left(\gamma\rho_{x,v}iu - \kappa\right)^2 - (iu-1)iu\gamma^2}$, $G_1 = \frac{\lambda - D_1}{\lambda + D_1}$, *and* $G_2 = \frac{\kappa - \gamma\rho_{x,v}iu - D_2}{\kappa - \gamma\rho_{x,v}iu + D_2}$.

Proof. The proof is very similar to the proof of Lemma 3.6 in section A.2. ∎

The integral for $A(u,\tau)$ in Lemma 5.2 can be determined analytically only for constant approximations of the two expectations involved.

5.2. ChF for the H2-CIR model. As before, we aim to find an approximation of the instantaneous covariance matrix for which the affinity of the approximation model is obtained, but now with the stochastic approximation.

$\mathbf{\Sigma}_{(1,3)}$ now consists of two stochastic components, $\sqrt{v(t)}$ and $\sqrt{r(t)}$. We approximate both and obtain

$$\mathbf{\Sigma}_{(1,3)} \approx \widetilde{\mathbf{\Sigma}}_{(1,3)} = \rho_{x,r}\eta\xi(t)R(t), \quad R(t) \approx \sqrt{r(t)}, \quad \xi(t) \approx \sqrt{v(t)}. \tag{5.8}$$

This form, based on the product of two random variables, is not affine. To linearize (5.8) we need to specify the joint dynamics, $\mathrm{d}(\sqrt{v(t)}\sqrt{r(t)})$. If we assume that the dynamics for $\mathrm{d}(\sqrt{v(t)})$ and $\mathrm{d}(\sqrt{r(t)})$ can be approximated by normally distributed processes, we find, by Itô's lemma, that the dynamics of $z(t) = \xi(t)R(t)$ are given by

$$\mathrm{d}z(t) = \left(\mu^R(t)\xi(t) + \mu^\xi(t)R(t)\right)\mathrm{d}t + \psi^\xi(t)R(t)\mathrm{d}W_v(t) + \psi^R(t)\xi(t)\mathrm{d}W_r(t). \tag{5.9}$$

With three additional variables, $\xi(t)$, $R(t)$, and $z(t)$, the state vector $\mathbf{X}(t)$, with log-stock process $x(t) = \log S(t)$, is expanded to $\mathbf{X}(t) = [x(t), v(t), r(t), \xi(t), R(t), z(t)]^\mathrm{T}$, with the following corresponding system of SDEs:

$$\begin{cases} \mathrm{d}x(t) = & \left(r(t) - \frac{1}{2}v(t)\right)\mathrm{d}t + & \sqrt{v(t)}\mathrm{d}W_x(t), & x(0) = \log(S(0)), \\ \mathrm{d}v(t) = & \kappa(\bar{v} - v(t))\mathrm{d}t + & \gamma\sqrt{v(t)}\mathrm{d}W_v(t), & v(0) > 0, \\ \mathrm{d}r(t) = & \lambda(\theta(t) - r(t))\mathrm{d}t + & \eta\sqrt{r(t)}\mathrm{d}W_r(t), & r(0) > 0, \end{cases} \tag{5.10}$$

with the linearizing variables $\xi(t)$, $R(t)$, and $z(t)$ given by

(5.11)
$$\begin{aligned}
\mathrm{d}\xi(t) &= \mu^{\xi}(t)\mathrm{d}t + \psi^{\xi}(t)\mathrm{d}W_v(t), & \xi(0) &= \sqrt{v(0)},\\
\mathrm{d}R(t) &= \mu^{R}(t)\mathrm{d}t + \psi^{R}(t)\mathrm{d}W_r(t), & R(0) &= \sqrt{r(0)},\\
\mathrm{d}z(t) &= \left(\mu^{R}(t)\xi(t) + \mu^{\xi}(t)R(t)\right)\mathrm{d}t + \psi^{\xi}(t)\sqrt{r(t)}\mathrm{d}W_v(t) + \psi^{R}(t)\sqrt{v(t)}\mathrm{d}W_r(t),
\end{aligned}$$

with $z(0) = \sqrt{r(0)}\sqrt{v(0)}$, $\xi(t) \approx \sqrt{v(t)}$, $R(t) \approx \sqrt{r(t)}$, $z(t) \approx \sqrt{v(t)}\sqrt{r(t)}$, and the other parameters as in (2.3). The symmetric instantaneous covariance matrix reads as

$$\widetilde{\mathbf{\Sigma}}^{*} = \begin{bmatrix}
v(t) & \rho_{x,v}\gamma v(t) & \rho_{x,r}\eta z(t) & \rho_{x,v}\psi^{\xi}(t)\xi(t) & \rho_{x,r}\psi^{R}(t)\xi(t) & s_1(t)\\
* & \gamma^2 v(t) & 0 & \psi^{\xi}(t)\gamma\xi(t) & 0 & \gamma\psi^{\xi}(t)z(t)\\
* & * & \eta^2 r(t) & 0 & \psi^{R}(t)\eta R(t) & \eta\psi^{R}(t)z(t)\\
* & * & * & (\psi^{\xi}(t))^2 & 0 & (\psi^{\xi}(t))^2 R(t)\\
* & * & * & * & (\psi^{R}(t))^2 & (\psi^{R}(t))^2\xi(t)\\
* & * & * & * & * & s_2(t)
\end{bmatrix}, \tag{5.12}$$

with $s_1(t) = \rho_{x,v}\psi^{\xi}(t)z(t) + \rho_{x,r}\psi^{R}(t)v(t)$ and $s_2(t) = (\psi^{\xi}(t))^2 r(t) + (\psi^{R}(t))^2 v(t)$.

Since $\psi^{\xi}(t)$ and $\psi^{R}(t)$ are deterministic time-dependent functions, the approximate H2-CIR model is now affine, and we can derive the corresponding ChF:

$$\begin{aligned}
\phi_{\text{H2-CIR}}(u, \mathbf{X}(t), \tau) = \exp\big(&A(u,\tau) + B(u,\tau)x(t) + C(u,\tau)r(t) + D(u,\tau)v(t)\\
&+ E(u,\tau)\xi(t) + F(u,\tau)R(t) + G(u,\tau)z(t)\big),
\end{aligned} \tag{5.13}$$

with $\xi(t) = \sqrt{v(t)}$, $R(t) = \sqrt{r(t)}$, and $z(t) = \sqrt{v(t)}\sqrt{r(t)}$, and where the functions $A(u,\tau)$, $B(u,\tau)$, $C(u,\tau)$, $D(u,\tau)$, $E(u,\tau)$, $F(u,\tau)$, and $G(u,\tau)$ satisfy the ODEs given by the following lemma.

Lemma 5.3 (ODEs related to the H2-CIR model). *The functions* $B(u,\tau) =: B(\tau)$, $C(u,\tau) =: C(\tau)$, $D(u,\tau) =: D(\tau)$, $E(u,\tau) =: E(\tau)$, $F(u,\tau) =: F(\tau)$, $G(u,\tau) =: G(\tau)$, *and* $A(u,\tau) =: A(\tau)$ *for* $u \in \mathbb{R}$ *and* $\tau > 0$ *in* (5.13) *satisfy*

$$\begin{aligned}
B'(\tau) &= 0,\\
C'(\tau) &= -1 + B(\tau) - \lambda C(\tau) + \eta^2 C^2(\tau)/2 + (\psi^{\xi}(t))^2 G^2(\tau)/2,\\
F'(\tau) &= \mu^{\xi}(t)G(\tau) + \psi^{R}(t)\eta C(\tau)G(\tau) + (\psi^{\xi}(t))^2 E(\tau)G(\tau),\\
G'(\tau) &= \eta\rho_{x,r}B(\tau)C(\tau) + \rho_{x,v}\psi^{\xi}(t)B(\tau)G(\tau) + \gamma\psi^{\xi}(t)D(\tau)G(\tau) + \eta\psi^{R}(t)C(\tau)G(\tau)
\end{aligned}$$

and

$$\begin{aligned}
D'(\tau) &= B(\tau)\left(B(\tau) - 1\right)/2 - \kappa D(\tau) + \gamma\rho_{x,v}B(\tau)D(\tau) + \gamma^2 D^2(\tau)/2\\
&\quad + \rho_{x,r}\psi^{R}(t)B(\tau)G(\tau) + (\psi^{R}(t))^2 G^2(t)/2,\\
E'(\tau) &= \mu^{R}(t)G(\tau) + \psi^{\xi}(t)\rho_{x,v}B(\tau)E(\tau) + \gamma\psi^{\xi}(t)D(\tau)E(\tau)\\
&\quad + \rho_{x,r}\psi^{R}(t)B(\tau)F(\tau) + (\psi^{R}(t))^2 F(\tau)G(\tau),\\
A'(\tau) &= \kappa\bar{v}D(\tau) + \lambda\theta C(\tau) + \mu^{\xi}(t)E(\tau) + \mu^{R}(t)F(\tau)\\
&\quad + (\psi^{\xi}(t))^2 E^2(\tau)/2 + (\psi^{R}(\tau))^2 F^2(\tau)/2,
\end{aligned}$$

with the final conditions $B(u,0) = iu$, $C(u,0) = 0$, $D(u,0) = 0$, $E(u,0) = 0$, $F(u,0) = 0$, $G(u,0) = 0$, *and* $A(u,0) = 0$. *Parameters* $\mu^{\xi}(t)$, $\mu^{R}(t)$, $\psi^{\xi}(t)$, $\psi^{R}(t)$ *are specified in* (4.9), *and the remaining parameters are in* (5.10).

Proof. The proof is very similar to the proof of Lemma 3.5 in section A.1. ∎

The system of the ODEs given in Lemma 5.3 is difficult to solve analytically. To find the solution we have used an explicit Runge–Kutta method [18, 27], ode45 from the MATLAB package. Numerical results are presented in the next subsection.

The extension of the H2-CIR model to the case of a full matrix of correlations is a trivial exercise.

5.3. Numerical experiment. We compare the performance of the approximations H1-CIR and H2-CIR with the full-scale HCIR model. As in the case of the HHW models, we have chosen here $T = 10$, and the model parameters are chosen so that the Feller condition does not hold. The results, presented in Table 2, are very satisfactory. Both approximation models, H1-CIR and H2-CIR, provide an error, $\epsilon(\rho_{x,r})$, for a call option within the confidence bounds. For higher correlation $\rho_{x,r}$, the error grows, but it is still small.

Table 2

The implied volatilities and errors for the deterministic approximation (Approx 1*) from* (2.18) *with approximation* (3.2) *and the stochastic approximation (Approx* 2*) from section* 4.1 *of the HCIR model compared to the MC simulation performed with* $20\,T$ *steps and* 100.000 *paths. The error is defined as a difference between the reference implied volatilities and the approximation. The parameters were chosen as* $\kappa = 0.3$, $\gamma = 0.6$, $v(0) = \bar{v} = 0.05$, $\lambda = 0.01$, $r(0) = \theta = 0.02$, $\eta = 0.01$, *and* $S(0) = 100$ *and the correlations as* $\rho_{x,v} = -30\%$ *and* $\rho_{x,r} \in \{20\%, 60\%\}$. *Numbers in brackets indicate standard deviations.*

$\rho_{x,r}$	Strike	MC Imp. vol. [%]	Approx 1	Approx 2	Err. 1	Err. 2
20%	40	25.66 (0.17)	25.68	25.74	−0.02	−0.08
	80	19.17 (0.15)	19.21	19.25	−0.04	−0.08
	100	17.10 (0.18)	17.19	17.09	−0.09	−0.01
	120	15.77 (0.17)	15.90	15.85	−0.14	−0.08
	180	15.84 (0.18)	15.90	15.86	−0.06	−0.02
60%	40	24.95 (0.14)	25.72	25.79	−0.77	−0.84
	80	18.93 (0.12)	19.32	19.32	−0.39	−0.39
	100	16.92 (0.13)	17.37	17.08	−0.44	−0.15
	120	15.60 (0.13)	16.17	15.93	−0.57	−0.32
	180	15.57 (0.14)	16.10	15.98	−0.53	−0.41

We also present the time needed for obtaining the plain vanilla option prices, with the ChFs H2-HW (section 4.2) and H2-CIR (section 5.2) based on the numerical solution for the system of Riccati ODEs. Table 3 shows that, although the ODEs in Lemma 5.3 need to be solved numerically, the time for obtaining European option prices, by the COS pricing method [16], is often less than 0.1 seconds. The pricing of the options by means of the COS method, a method based on Fourier cosine series expansions, was performed with a fixed number of 250 terms, which guaranteed highly accurate option prices (up to machine precision).

The tolerance for the ODE solves, by ode45 from MATLAB, is varied in the experiments shown in the table.

Table 3

Time in seconds for pricing a call option based on an explicit Runge–Kutta method combined with the COS method [16].

Model	Accuracy	Maturity				
		$\tau = 0.5$	$\tau = 1$	$\tau = 2$	$\tau = 5$	$\tau = 10$
H2-HW	10^{-2}	4.37e-2	4.80e-2	6.41e-2	7.49e-2	8.10e-2
	10^{-5}	5.32e-2	5.82e-2	8.05e-2	9.74e-2	1.21e-1
H2-CIR	10^{-2}	7.78e-2	7.80e-2	8.38e-2	8.48e-2	8.90e-2
	10^{-5}	8.33e-2	8.97e-2	1.05e-1	1.34e-1	1.62e-1

6. Calibration of the Heston hybrid models. Here, we evaluate the performance of the approximations H1-HW and H2-HW for the HHW hybrid model in a calibration setting.

We reference call option prices, based on synthetic data that is representative for the skew and smile patterns observed in real-life applications. For all models the simulation was performed with an a priori defined speed of mean reversion for the variance process, $\kappa = 0.3$ (which is set small on purpose). The calibration is performed here with constant correlation $\rho_{x,r} = 20\%$. In practice, this correlation can be obtained from historical data.

The calibration procedure is performed in two stages. First, the parameters for the short rate process are determined (independent of the equity part). In the second stage, the calibrated $r(t)$ is included in the Heston model, and the remaining parameters are determined. The parameters for the interest rate part are found to be $\lambda_{\text{HW}} = 0.501$, $\eta_{\text{HW}} = 0.005$, and $r(0) = 0.04$.

First, we also perform, as a benchmark, the calibration of the pure Heston model with constant interest rate; see Table 4. SSE stands stands for the "sum-squared error." We calibrate the models for different maturities, τ.

Table 4

Calibration results for the Heston stochastic volatility model with deterministic interest rate. The mean reversion parameter is $\kappa = 0.3$.

Model	γ	$\bar{v}$	$\rho_{x,v}$	$v(0)$	r	SSE
Heston ($\tau = 0.5$)	0.5992	0.0823	-58.32%	0.0407	0.04	4.9063E-4
Heston ($\tau = 10$)	0.6019	0.0828	-48.49%	0.0411	0.04	1.2182E-4

In Table 5 the calibration results for the HHW approximations, H1-HW and H2-HW, are presented. For both models a highly satisfactory fit is obtained, with a slightly better performance of the stochastic approximation H2-HW. For $\rho_{x,r} = 20\%$ the calibration procedure gives roughly the same sets of parameters for both models. When comparing the calibration results for HHW with those for the pure Heston model, we see that the inclusion of stochastic interest rates in the model results in a lower vol-vol parameter, γ, and a more negative correlation, $\rho_{x,v}$. The lower value of parameter γ can be explained by the additional volatility which comes from the interest rate process.

In Figure 4 the corresponding implied volatilities, for the full-scale model, for a short and a long maturity time ($\tau = 0.5y$ and $\tau = 10y$) are presented. The left-hand sides of the figure

Table 5

Calibration results for the H1-HW model from section 3.2, and the H2-HW model from section 4.2, with $\kappa = 0.3$ *and correlation* $\rho_{x,r} = 20\%$.

Model	τ	γ	$\bar{v}$	$\rho_{x,v}$	$v(0)$	SSE
H1-HW	$\tau = 0.5$	0.5840	0.0822	-60.06%	0.0407	4.4581E-4
	$\tau = 10$	0.4921	0.0826	-61.50%	0.0418	3.2912E-4
H2-HW	$\tau = 0.5$	0.5879	0.0930	-60.10%	0.0398	4.9677E-4
	$\tau = 10$	0.4884	0.0820	-60.72%	0.0421	8.5934E-5

present the implied volatilities and their errors for H1-HW and H2-HW. The related implied volatilities of the full-scale HHW model, with the parameters from H1-HW and H2-HW, are shown in the right-hand sides of the figure.

Both hybrid models perform very well. For long maturities a higher accuracy for the hybrid models compared to the plain Heston model can be observed.

7. Concluding remarks.

In this article we have presented the extension of the Heston stochastic volatility equity model by stochastic interest rates. We have focused our attention on two hybrid models, the HHW and the HCIR models.

By approximations of the nonaffine terms in the corresponding instantaneous covariance matrix, we placed the approximation hybrid models in the framework of AD processes. The approximations in the models have been validated by comparing the implied volatilities to the full-scale hybrid models.

The approximations in the HHW and the HCIR models lead to highly efficient determination of the corresponding ChFs. The more sophisticated approximation is based on a transformation of the three-dimensional HCIR model to a six-dimensional representation.

The deterministic and the stochastic approaches for approximating the instantaneous covariance matrix of the hybrid model provide highly satisfactory approximations for prices for European options.

Appendix A. Proofs and details.

A.1. Proof of Lemma 3.5.

Proof. For a given state vector $X(t) = [x(t), r(t), v(t)]^{\mathrm{T}}$ and $\phi := \phi(u, X(t), t, T)$, we find the system of the ODEs satisfying the following pricing PDE:

$$\begin{aligned} 0 = &\frac{\partial \phi}{\partial t} + \left(r - \frac{1}{2}v\right)\frac{\partial \phi}{\partial x} + \kappa(\bar{v} - v)\frac{\partial \phi}{\partial v} + \lambda(\theta(t) - r)\frac{\partial \phi}{\partial r} + \frac{1}{2}v\frac{\partial^2 \phi}{\partial x^2} \\ &+ \frac{1}{2}\gamma^2 v\frac{\partial^2 \phi}{\partial v^2} + \frac{1}{2}\eta^2\frac{\partial^2 \phi}{\partial r^2} + \rho_{x,v}\gamma v\frac{\partial^2 \phi}{\partial x \partial v} + \eta\rho_{x,r}\mathbb{E}(\sqrt{v(t)})\frac{\partial^2 \phi}{\partial x \partial r} - r\phi, \end{aligned} \tag{A.1}$$

subject to terminal condition $\phi(u, \mathbf{X}(T), T, T) = \exp(iux(T))$.

Since the PDE in (A.1) is affine, its solution is of the following form:

$$\phi := \phi(u, X(t), t, T) = \exp\left(A(u,t,T) + B(u,t,T)x(t) + C(u,t,T)r(t) + D(u,t,T)v(t)\right).$$

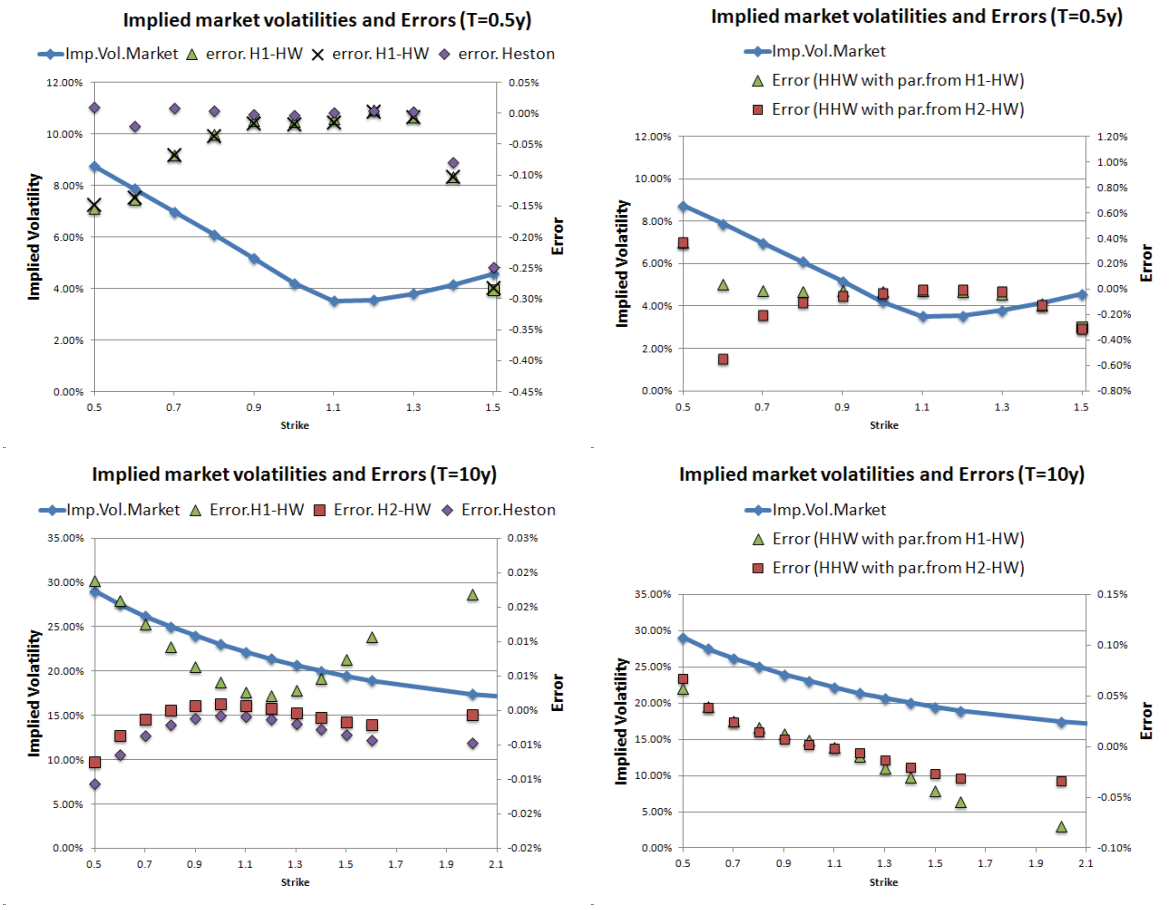

Figure 4. *For $\tau = 0.5$ and $\tau = 10$, $\rho_{x,r} = 20\%$, the implied Black–Scholes volatilities for Heston hybrid models are compared to the pure Heston model and a reference implied volatility curve. The left-hand graphs present the implied volatilities and errors for H1-HW and H2-HW. The implied volatilities for the full-scale HHW model, with the parameters from H1-HW and H2-HW, are in the right-hand figures.*

By setting $A := A(u,t,T)$, $B := B(u,t,T)$, $C := C(u,t,T)$, and $D := D(u,t,T)$, we find the following partial derivatives:

$$\frac{\partial \phi}{\partial t} = \phi\left(\frac{\partial A}{\partial t} + x(t)\frac{\partial B}{\partial t} + r(t)\frac{\partial C}{\partial t} + v(t)\frac{\partial D}{\partial t}\right), \tag{A.2}$$

$$\frac{\partial \phi}{\partial x} = B\phi, \quad \frac{\partial^2 \phi}{\partial x^2} = B^2\phi, \quad \frac{\partial^2 \phi}{\partial x \partial v} = BD\phi, \quad \frac{\partial^2 \phi}{\partial x \partial r} = BC\phi, \tag{A.3}$$

$$\frac{\partial \phi}{\partial r} = C\phi, \quad \frac{\partial^2 \phi}{\partial r^2} = C^2\phi, \tag{A.4}$$

$$\frac{\partial \phi}{\partial v} = D\phi, \quad \frac{\partial^2 \phi}{\partial v^2} = D^2\phi. \tag{A.5}$$

By substitution, PDE (A.1) becomes

$$0 = \frac{\partial A}{\partial t} + x\frac{\partial B}{\partial t} + r\frac{\partial C}{\partial t} + v\frac{\partial D}{\partial t} + \left(r - \frac{1}{2}v\right)B + \kappa(\bar{v} - v)D + \lambda(\theta(t) - r)C$$
$$+ \frac{1}{2}vB^2 + \frac{1}{2}\gamma^2 vD^2 + \frac{1}{2}\eta^2 C^2 + \rho_{x,v}\gamma vBD + \eta\rho_{x,r}\mathbb{E}(\sqrt{v(t)})BC - r. \tag{A.6}$$

Now, by collecting the terms for $x(t)$, $r(t)$, and $v(t)$, we find the following set of ODEs:

$$\frac{\partial B}{\partial t} = 0, \tag{A.7}$$
$$\frac{\partial C}{\partial t} = -B + \lambda C + 1, \tag{A.8}$$
$$\frac{\partial D}{\partial t} = \frac{1}{2}B + \kappa D - \frac{1}{2}\gamma^2 D^2 - \rho_{x,v}\gamma BD - \frac{1}{2}B^2, \tag{A.9}$$
$$\frac{\partial A}{\partial t} = -\kappa\bar{v}D - \lambda\theta C - \frac{1}{2}\eta^2 C^2 - \rho_{x,r}\eta\mathbb{E}(\sqrt{v(t)})BC. \tag{A.10}$$

By setting $\tau = T - t$, the proof is finished. ■

A.2. Proof of Lemma 3.6.

Proof. Obviously, due to the final condition, $B(u,0) = iu$, we have $B(u,\tau) = iu$. For the second ODE, by multiplying both sides with $\mathrm{e}^{\lambda\tau}$, we get

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\mathrm{e}^{\lambda\tau}C(u,\tau)\right) = (iu - 1)\mathrm{e}^{\lambda\tau}; \tag{A.11}$$

by integrating both sides and using the final condition, $C(u,0) = 0$, we find

$$C(u,\tau) = (iu - 1)\lambda^{-1}\left(1 - \mathrm{e}^{-\lambda\tau}\right).$$

By setting $a = -\frac{1}{2}(u^2 + iu)$, $b = \gamma\rho_{x,v}iu - \kappa$, and $c = \frac{1}{2}\gamma^2$, the ODEs for $D(u,\tau)$ and $I_2(\tau)$ are given by the following Riccati-type equation:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}D(u,\tau) = a + bD(u,\tau) + cD^2(u,\tau), \quad D(u,0) = 0, \tag{A.12}$$
$$I_2(\tau) = \kappa\bar{v}\int_0^\tau D(u,s)\mathrm{d}s. \tag{A.13}$$

Equations (A.12) and (A.13) are of the same form as those in [22]. Their solutions are given by

$$D(u,\tau) = \frac{-b - D_1}{2c(1 - G\mathrm{e}^{-D_1\tau})}(1 - \mathrm{e}^{-D_1\tau}), \tag{A.14}$$
$$I_2(\tau) = \frac{1}{2c}\left((-b - D_1)\tau - 2\log\left(\frac{1 - G\mathrm{e}^{-D_1\tau}}{1 - G}\right)\right), \tag{A.15}$$

with $D_1 = \sqrt{b^2 - 4ac}$, $G = \frac{-b-D_1}{-b+D_1}$.

The evaluation of the integrals $I_1(\tau)$, $I_3(\tau)$, and $I_4(\tau)$ is straightforward. The proof is finished by appropriate substitutions. ∎

Appendix B. Hybrid model with full matrix of correlations. Similar to the approximation of the nonaffine terms in the instantaneous covariance matrix of the Heston hybrid model presented in section 3.1, we discuss here the inclusion of the additional correlation, $\rho_{r,v}$, between the interest rate, $r(t)$, and the stochastic variance, $v(t)$. We call the resulting model the HHW hybrid model-3 and denote it by H3-HW. For the state vector $\mathbf{X}(t) = [x(t), v(t), r(t)]^{\mathrm{T}}$, the H3-HW model has the following symmetric instantaneous covariance matrix:

$$\mathbf{\Sigma} := \sigma(\mathbf{X}(t))\sigma(\mathbf{X}(t))^{\mathrm{T}} = \begin{bmatrix} v(t) & \rho_{x,v}\gamma v(t) & \rho_{x,r}\eta\sqrt{v(t)} \\ * & \gamma^2 v(t) & \rho_{r,v}\gamma\eta\sqrt{v(t)} \\ * & * & \eta^2 \end{bmatrix}_{(3\times 3)}. \tag{B.1}$$

The affinity issue arises in two terms of matrix (B.1), namely, in elements $(1,3)$ and $(2,3)$:

$$\mathbf{\Sigma}_{(1,3)} = \rho_{x,r}\eta\sqrt{v(t)}, \quad \mathbf{\Sigma}_{(2,3)} = \rho_{r,v}\gamma\eta\sqrt{v(t)}.$$

For completeness, we also present the associated Kolmogorov backward equation, which is now given by

$$\begin{aligned} 0 = {} & \frac{\partial\phi}{\partial t} + \left(r - \frac{1}{2}v\right)\frac{\partial\phi}{\partial x} + \kappa(\bar{v} - v)\frac{\partial\phi}{\partial v} + \lambda(\theta(t) - r)\frac{\partial\phi}{\partial r} + \frac{1}{2}v\frac{\partial^2\phi}{\partial x^2} + \frac{1}{2}\gamma^2 v\frac{\partial^2\phi}{\partial v^2} \\ & + \frac{1}{2}\eta^2\frac{\partial^2\phi}{\partial r^2} + \rho_{x,v}\gamma v\frac{\partial^2\phi}{\partial x\partial v} + \mathbf{\Sigma}_{(1,3)}\frac{\partial^2\phi}{\partial x\partial r} + \mathbf{\Sigma}_{(2,3)}\frac{\partial^2\phi}{\partial r\partial v} - r\phi, \end{aligned} \tag{B.2}$$

with the final condition equal to

$$\phi(u, \mathbf{X}(T), T, T) = \exp(iux(T)).$$

With $\rho_{r,v} = 0$ the H3-HW model with a full matrix of correlations collapses to the setup in section 3.1.

As before, we can use the deterministic approximations $\mathbf{\Sigma}_{(1,3)} \approx \rho_{x,r}\eta\mathbb{E}(\sqrt{v(t)})$ and $\mathbf{\Sigma}_{(2,3)} \approx \rho_{r,v}\gamma\eta\mathbb{E}(\sqrt{v(t)})$, for which Result 3.3 can be used.

The representations of the HHW model in (2.9) and the model in (2.3) with $\rho_{r,v} \neq 0$ for $p = 0$ are closely related. The following lemma specifies the relation in terms of the coefficients of the corresponding ChF.

Lemma B.1 (ChF for the H3-HW model with a full matrix of correlations). *The discounted ChF for the H3-HW model is of the following form:*

$$\phi_{H3\text{-}HW}(u, \mathbf{X}(t), \tau) = \exp\left(\hat{A}(u,\tau) + \hat{B}(u,\tau)x(t) + \hat{C}(u,\tau)r(t) + \hat{D}(u,\tau)v(t)\right),$$

with the functions $\hat{A}(u,\tau)$, $\hat{B}(u,\tau)$, $\hat{C}(u,\tau)$, and $\hat{D}(u,\tau)$ given by

$$\hat{B}(u,\tau) = B(u,\tau), \quad \hat{C}(u,\tau) = C(u,\tau), \quad \hat{D}(u,\tau) = D(u,\tau), \tag{B.3}$$

with $B(u,\tau)$ *in* (3.23), $C(u,\tau)$ *in* (3.24), *and* $D(u,\tau)$ *in* (3.25). *For* $\hat{A}(u,\tau)$ *we have*

$$\hat{A}(u,\tau) = A(u,\tau) + \rho_{r,v}\gamma\eta \int_0^\tau \mathbb{E}(\sqrt{v(T-s)})\hat{C}(u,s)\hat{D}(u,s)\mathrm{d}s, \tag{B.4}$$

where $A(u,\tau)$ *is given in* (3.26).

The accuracy of the HHW approximations with a full matrix of correlations is discussed in Appendix C.

Appendix C. Comparison to Markov projection method. In this appendix we compare our results to the Markovian projection (MP) method [5]. We check the results of three different approximation schemes: the MP method, Approx 1, i.e., the approximation with $\sqrt{v(t)} \approx \mathbb{E}(\sqrt{v(t)})$ (section 3.1), and Approx 2, i.e., the method with $\sqrt{v(t)} \approx \mathcal{N}(\cdot)$ (section 4.1).

In the experiment, taken directly from [4], we price an equity option with continuous dividend. The model parameters for the HHW model are given by $\kappa = 0.25$, $\bar{v} = v(0) = 0.0625$, $\gamma = 0.625$, $\lambda = 0.05$, and $\eta = 0.01$, and a zero-coupon bond is given by $P(0,T) = \mathrm{e}^{-0.05T}$ and a continuous dividend of 2%. A full matrix of correlations, as in [4], is given by

$$\mathbf{C} = \begin{bmatrix} 1 & \rho_{x,v} & \rho_{x,r} \\ \rho_{x,v} & 1 & \rho_{v,r} \\ \rho_{x,r} & \rho_{v,r} & 1 \end{bmatrix} = \begin{bmatrix} 100\% & -40\% & 30\% \\ -40\% & 100\% & 15\% \\ 30\% & 15\% & 100\% \end{bmatrix}. \tag{C.1}$$

The MC reference for the implied volatilities, the corresponding standard deviations, and the results for the MP method are all taken from [4].

In order to incorporate a continuous dividend in the equity model, one can model foreign-exchange (FX), in which the volatility of the foreign interest rates is set to zero. In such a setup, the forward, $F(t)$, is defined as

$$F(t) = S(t)\frac{P_f(t,T)}{P_d(t,T)} \quad \text{and} \quad F(0) = S(0)\frac{\mathrm{e}^{-0.02T}}{\mathrm{e}^{-0.05T}},$$

where $P_f(t,T)$ and $P_d(t,T)$ are the foreign and domestic zero-coupon bonds, respectively, paying €1 at the maturity T. By switching from the spot risk-neutral measure, $\mathbb{Q}$, to the T-forward measure, $\mathbb{Q}^T$, discounting will be decoupled from taking the expectation, i.e.,

$$\mathbb{E}\left(\frac{1}{B(T)}\max(S(T)-K,0)|\mathcal{F}(0)\right) = P_d(0,T)\mathbb{E}^T\left(\max(F(T)-K,0)|\mathcal{F}(0)\right).$$

Moreover, the forward, $F(t)$, is a martingale with dynamics given by

$$\begin{aligned} \mathrm{d}F(t)/F(t) &= \sqrt{v(t)}\mathrm{d}W_x^T(t) - \eta B_r(t,T)\mathrm{d}W_r^T(t), \\ \mathrm{d}v(t) &= \left(\kappa(\bar{v}-v(t)) + \gamma\rho_{v,r}\eta B_r(t,T)\sqrt{v(t)}\right)\mathrm{d}t + \gamma\sqrt{v(t)}\mathrm{d}W_v^T(t), \end{aligned} \tag{C.2}$$

where $B_r(t,T) = \frac{1}{\lambda}\left(\mathrm{e}^{-\lambda(T-t)}-1\right)$, and the full correlation structure given in (C.1).

Table 6

The error for a deterministic (Approx 1) and a stochastic approximation (Approx 2) of the HHW model compared to the MP method. The MP and MC results with the corresponding standard deviations were taken from [4]. The error is defined as a difference between the reference implied volatilities and the approximation.

T	Strike	Imp. vol [%]	MP	Approx 1	Approx 2	Err. MP	Err. 1	Err. 2
	86.07	24.45	24.49	24.48	24.48	−0.04	−0.03	−0.03
	92.77	22.25	22.27	22.27	22.25	−0.02	−0.02	0.00
1y	100.00	20.36	20.32	20.35	20.30	0.04	0.01	0.06
	107.79	19.42	19.34	19.38	19.34	0.08	0.04	0.08
	116.18	19.67	19.64	19.62	19.64	0.03	0.05	0.03
	77.12	22.61	22.65	22.61	22.63	−0.04	0.00	−0.02
	87.82	20.05	20.05	20.09	20.06	0.00	−0.04	−0.01
3y	100.00	17.95	17.91	18.09	17.90	0.04	−0.14	0.05
	113.87	17.23	17.14	17.32	17.15	0.09	−0.09	0.08
	129.67	18.02	17.92	17.93	18.00	0.10	0.09	0.02
	71.50	21.89	21.94	21.90	21.95	−0.05	−0.01	−0.06
	84.56	19.43	19.45	19.52	19.48	−0.02	−0.09	−0.05
5y	100.00	17.49	17.44	17.71	17.45	0.05	−0.22	0.04
	118.26	16.83	16.72	17.01	16.76	0.11	−0.18	0.07
	139.85	17.55	17.42	17.49	17.57	0.13	0.06	−0.02
	62.23	21.55	21.61	21.57	21.68	−0.06	−0.02	−0.13
	78.89	19.52	19.51	19.67	19.61	0.01	−0.15	−0.09
10y	100.00	18.01	17.91	18.31	17.97	0.10	−0.30	−0.04
	126.77	17.41	17.22	17.67	17.30	0.19	−0.26	0.11
	160.70	17.75	17.51	17.79	17.78	0.24	−0.04	−0.03
	51.13	22.28	22.32	22.37	22.47	−0.04	−0.09	−0.19
	71.50	20.91	20.86	21.14	21.03	0.05	−0.23	−0.12
20y	100.00	19.94	19.77	20.27	19.91	0.17	−0.33	0.03
	139.85	19.44	19.16	19.77	19.32	0.28	−0.33	0.12
	195.58	19.40	19.05	19.63	19.39	0.35	−0.23	0.01

Under the log-transform, $x(t) = \log F(t)$, the Kolmogorov backward PDE reads as

$$
\begin{aligned}
-\frac{\partial \phi}{\partial t} &= \kappa(\bar{v} - v)\frac{\partial \phi}{\partial v} + \left(\frac{1}{2}v - \rho_{x,r}\eta B_r(t,T)\sqrt{v} - \frac{1}{2}\eta^2 B_r^2(t,T)\right)\left(\frac{\partial^2 \phi}{\partial x^2} - \frac{\partial \phi}{\partial x}\right) \\
&\quad + \left(\rho_{x,v}\gamma v - \rho_{v,r}\gamma\eta\sqrt{v}B_r(t,T)\right)\frac{\partial^2 \phi}{\partial x \partial v} + \frac{1}{2}\gamma^2 v\frac{\partial^2 \phi}{\partial v^2} + \rho_{v,r}\gamma\eta\sqrt{v}\frac{\partial \phi}{\partial v},
\end{aligned} \tag{C.3}
$$

with the final condition $\phi(u, \mathbf{X}(T), T, T) = \mathrm{e}^{iux(T)}$.

We linearize PDE (C.3) in two ways: by the deterministic approach described in section 3.1 and Appendix B, and by the stochastic approach as in section 4.1. Both approximations result in affine approximations of PDE (C.3).[6]

The results of the experiments performed, presented in Table 6, show a highly satisfactory accuracy of the HHW approximations introduced in this paper. When comparing to the MP method, we see that the MP method is more accurate for low strike values, whereas our proxies perform favorably for larger strike values, especially when large maturities are considered.

[6]Since the moments of the square-root process under the T-forward measure are difficult to find, we first project $\sqrt{v(t)}$ on a normal process, under measure $\mathbb{Q}$, and then change measures.

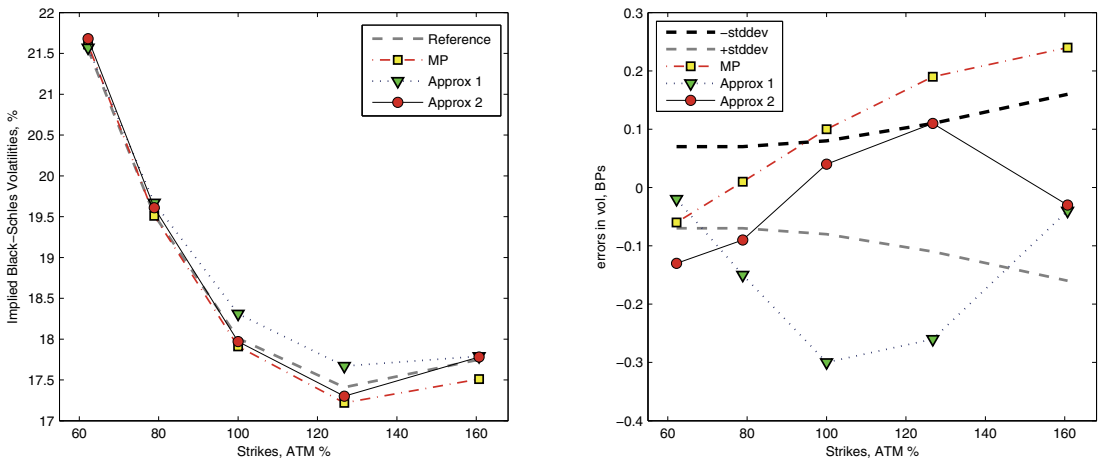

Figure 5. *Left: Implied volatilities for a maturity of* 10 *years. Right: The error for the different approximations. (MP stands for Markovian projection, Approx* 1 *is the deterministic approach, and Approx* 2 *corresponds to the approximation with* $\sqrt{v(t)} \approx \mathcal{N}(\cdot)$.*)*

In Figure 5 the error results for $T = 10$ are presented. In this experiment, the stochastic approximation, Approx 2, performed somewhat better than the deterministic approach, Approx 1.

In the case of the deterministic approach, pricing of European options is done in a split second (the corresponding ChF is analytic when the Feller condition is satisfied; one integration step is required otherwise). In the case of the stochastic approach a numerical routine for solving the ODEs is employed. This, however, can also be done highly efficiently, as already presented in Table 3.

Acknowledgments. The authors would like to thank the anonymous referees for valuable suggestions. Moreover, the authors thank Natalia Borovykh and Sacha van Weeren from Rabobank International for fruitful discussions and helpful comments.

REFERENCES

[1] S. Amstrup, L. MacDonald, and B. Manly, *Handbook of Capture-Recapture Analysis*, Princeton University Press, Princeton, NJ, 2006.

[2] L. Andersen, *Simple and efficient simulation of the Heston stochastic volatility model*, J. Comput. Finance, 11 (2008), pp. 1–42.

[3] J. Andreasen, *Closed Form Pricing of FX Options under Stochastic Rates and Volatility*, presentation at Global Derivatives Conference, Paris, 2006.

[4] A. Antonov, *Effective Approximation of FX/EQ Options for the Hybrid Models: Heston and Correlated Gaussian Interest Rates*, presentation at MathFinance 2007; available online from http://conference.mathfinance.com/2008/papers/antonov/slides.

[5] A. Antonov, M. Arneguy, and N. Audet, *Markovian Projection to a Displaced Volatility Heston Model*, SSRN working paper, 2008; available online from http://ssrn.com/abstract=1106223.

[6] E. Benhamou, A. Rivoira, and A. Gruz, *Stochastic Interest Rates for Local Volatility Hybrid Models*, SSRN working paper, 2008; available online from http://ssrn.com/abstract=1107711.

[7] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Polit. Econ., 81 (1973), pp. 637–654.
[8] D. Brigo and F. Mercurio, *Interest Rate Models—Theory and Practice: With Smile, Inflation and Credit*, 2nd ed., Springer Finance, Springer-Verlag, Berlin, 2007.
[9] M. Broadie and Y. Yamamoto, *Application of the fast Gauss transform to option pricing*, Management Sci., 49 (2003), pp. 1071–1088.
[10] P. P. Carr and D. B. Madan, *Option valuation using the fast Fourier transform*, J. Comput. Finance, 2 (1999), pp. 61–73.
[11] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica, 53 (1985), pp. 385–407.
[12] E. Derman and I. Kani, *Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility*, Int. J. Theor. Appl. Finance, 1 (1998), pp. 61–110.
[13] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, 68 (2000), pp. 1343–1376.
[14] D. Dufresne, *The Integrated Square-Root Process*, University of Montreal working paper, 2001; available online from http://repository.unimelb.edu.au/10187/1413.
[15] B. Dupire, *Pricing with a smile*, Risk, 7 (1994), pp. 18–20.
[16] F. Fang and C. W. Oosterlee, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM J. Sci. Comput., 31 (2008), pp. 826–848.
[17] R. A. Fisher, *On the interpretation of χ^2 from contingency tables and calculations of P*, J. R. Statist. Soc., 85 (1922), pp. 87–94.
[18] G. Forsythe, M. Malcolm, and C. Moler, *Computer Methods for Mathematical Computations*, Prentice–Hall, Englewood Cliffs, NJ, 1977.
[19] A. Giese, *On the Pricing of Auto-callable Equity Securities in the Presence of Stochastic Volatility and Stochastic Interest Rates*, presentation at MathFinance Workshop: Derivatives and Risk Management in Theory and Practice, Frankfurt, 2006; available online from http://www.mathfinance.de/workshop/2006/papers/giese/slides.pdf.
[20] L. A. Grzelak, C. W. Oosterlee, and S. van Weeren, *Extension of stochastic volatility equity models with Hull-White interest rate process*, Quant. Finance, 2009, pp. 1469–7696.
[21] A. van Haastrecht, R. Lord, A. Pelsser, and D. Schrager, *Pricing long-maturity equity and FX derivatives with stochastic interest rates and stochastic volatility*, Insur. Math. Econ., 45 (2009), pp. 436–448.
[22] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Financ. Stud., 6 (1993), pp. 327–343.
[23] J. Hull and A. White, *Using Hull-White interest rate trees*, J. Derivatives, 4 (1996), pp. 26–36.
[24] C. Hunter, *Hybrid derivatives*, in The Euromoney Derivatives and Risk Management Handbook, Euromoney Institutional Investor, London, 2005.
[25] P. Jäckel, *Stochastic volatility models: Past, present and future*, in The Best of Wilmott I: Incorporating the Quantitative Finance Review, John Wiley and Sons, Chichester, 2004, pp. 379–390.
[26] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Vol. 2, 2nd ed., Wiley, New York, 1994.
[27] D. Kahaner, C. Moler, and S. Nash, *Numerical Methods and Software*, Prentice–Hall, Englewood Cliffs, NJ, 1989.
[28] E. E. Kummer, *Über die Hypergeometrische Reihe $F(a;b;x)$*, J. Reine Angew. Math., 15 (1936), pp. 39–83.
[29] R. Lee, *Option pricing by transform methods: Extensions, unification, and error control*, J. Comput. Finance, 7 (2004), pp. 51–86.
[30] A. Lewis, *Option Valuation under Stochastic Volatility*, Finance Press, Newport Beach, CA, 2001.
[31] M. Muskulus, K. in't Hout, J. Bierkens, A. P. C. van der Ploeg, J. in't Panhuis, F. Fang, B. Janssens, and C. W. Oosterlee, *The ING problem—a problem from financial industry; three papers on the Heston-Hull-White model*, in Proceedings of the 58th European Study Group Mathematics with Industry, Utrecht, The Netherlands, 2007.
[32] G. W. Oehlert, *A note on the delta method*, Amer. Statist., 46 (1992), pp. 27–29.
[33] B. Øksendal, *Stochastic Differential Equations*, 5th ed., Springer-Verlag, Berlin, 2000.

[34] P. B. Patnaik, *The non-central χ^2 and F-distributions and their applications*, Biometrika, 36 (1949), pp. 202–232.
[35] R. Schöbel and J. Zhu, *Stochastic volatility with an Ornstein-Uhlenbeck process: An extension*, Eur. Financ. Rev., 3 (1999), pp. 23–46.
[36] O. A. Vašiček, *An equilibrium characterization of the term structure*, J. Financ. Econ., 5 (1977), pp. 177–188.
[37] J. Zhu, *Modular Pricing of Options*, Springer-Verlag, Berlin, 2000.