DELFT UNIVERSITY OF TECHNOLOGY

REPORT 09-07

EFFICIENT OPTION PRICING WITH MULTI-FACTOR EQUITY-INTEREST RATE HYBRID MODELS

LECH A. GRZELAK, CORNELIS W. OOSTERLEE
& SACHA VAN WEEREN

ISSN 1389-6520

Reports of the Department of Applied Mathematical Analysis

Delft 2009

EFFICIENT OPTION PRICING WITH MULTI-FACTOR EQUITY-INTEREST RATE HYBRID MODELS*

LECH A. GRZELAK,[a,b,†] CORNELIS W. OOSTERLEE[a,c]
& SACHA VAN WEEREN[b]

[a] DELFT INSTITUTE OF APPLIED MATHEMATICS, DELFT UNIVERSITY OF TECHNOLOGY,
Mekelweg 4, 2628 CD, Delft, the Netherlands

[b] DERIVATIVES RESEARCH AND VALIDATION GROUP, RABOBANK,
Jaarbeursplein 22, 3521 AP, Utrecht, the Netherlands

[c] CWI - NATIONAL RESEARCH INSTITUTE FOR MATHEMATICS AND COMPUTER SCIENCE,
Science Park 123, 1098 XG Amsterdam, the Netherlands

first version: March 20, 2009
this version: July 15, 2009

Abstract

In this article we discuss multi-factor equity-interest rate hybrid models with a full matrix of correlations. We assume the equity part to be modeled by the Heston model [Heston-1993] with as a short rate process either a Gaussian two-factor model [Brigo,Mercurio-2007] or a stochastic volatility short rate process of Heston type [Heidari, *et al.*-2007]. We develop an approximation for the discounted characteristic function. Our approximation scheme is based on the observation that $\sqrt{\sigma_t}$, with σ_t a stochastic quantity of CIR type [Cox, *et al.*-1985], can be well approximated by a normal distribution. Our approximate hybrid fits almost perfectly to the original model in terms of implied Black-Scholes [Black,Scholes-1973] volatilities for European options. Since fast integration techniques allow us to get European style option prices for a whole strip of strikes in a split second, the hybrid approximation can be directly used for model calibration.

Key words: Hybrid, Heston-Gaussian Two-Factor Model (H-G2++), Heston-Stochastic Volatility Interest Rate (H-H2++), Affine Diffusion, Stochastic Volatility.

1 Introduction

Pricing modern contracts involving multiple asset classes requires well-developed pricing models from quantitative analysts. Among them the hybrid models, that include features from different asset classes are of interest.

In this article we analyze hybrids between two particular asset classes: the equity and the interest rates (IR). It is already well-known [Brigo,Mercurio-2007] that these models can be used for pricing specific hybrid products or for accurate pricing of long-term equity options. In either case, any hybrid model needs to be calibrated to some simple, European-style, liquid contracts. Although multi-dimensional hybrids can be relatively easily defined, real use of the models is guaranteed if the hybrid model is properly defined for each asset class (i.e., fit to implied volatility structures etc.), and if there is a non-zero correlation structure among the processes from the different asset classes. Furthermore, highly efficient pricing of fundamental contracts needs to be available. In this article we aim at models that satisfy these requirements.

*The authors would like to thank Natalia Borovykh from Rabobank International and Stefan Singor from Delft University of Technology for fruitful discussions and helpful comments.

†Corresponding author. E-mail address: L.A.Grzelak@tudelft.nl.

It is well-known that the Black-Scholes [Black,Scholes-1973] model may lead to significant option miss-pricing. Although stochastic and local volatility models [Heston-1993; Schöbel,Zhu-1999; Dupire-2008; Derman,Kani-1998] provide an explanation for the so-called leverage effect, also known as the market implied volatility *smile/skew*, these models are not always satisfactory for pricing exotic hybrid stock and interest rate products. A basic generalization to time-dependent deterministic interest rates does not allow the calibration of the interest rate part to any interest rate products, except to the current yield curve. Extending the model with a correlated (multi-factor) stochastic interest rate (short rate) process increases the flexibility of the model. On the other hand it also increases the model's complexity.

Pricing long-maturity options with equity-interest rate hybrid models is common practice in the market. The basic hybrid model of Black-Scholes for the equity and Hull-White [Hull,White-1996] for interest rate was presented in [Brigo,Mercurio-2007]. In [Grzelak, *et al.*,-2008] a stochastic volatility equity hybrid model with a full matrix of correlations (Schöbel-Zhu-Hull-White) was presented. Since the model was based on a normally distributed volatility process it was limited in implied volatility shapes. The Heston-Hull-White hybrid model was then presented in [Grzelak,Oosterlee-2009]. In the same article the interest rate process of Cox-Ingersoll-Ross [Cox, *et al.*-1985] was analyzed. Our current paper is a generalization of that model, since we extend the 1D interest rate process to a multi-factor process.

In this paper we construct two equity-interest rate hybrid models. We consider hybrids with the equity part to be driven by the Heston model [Heston-1993] while for the short rate process we assume either a Gaussian two-factor model [Hull-2006] or a Heston stochastic volatility short rate process [Heidari, *et al.*-2007].

The Hull-White two-factor interest rate model [Hull-2008] provides a rich pattern for the term structure movements and recovers a *humped* volatility structure observed in the market. However, it can not model a stable interest rate smile. Unfortunately, the model also allows the rates to become negative which is an obvious disadvantage. Therefore, as an alternative hybrid model, we also consider the interest rates to be driven by a square-root stochastic volatility model in this paper.

By approximations we place these hybrid models in the affine diffusion framework for which the corresponding characteristic function can be obtained. This facilitates the use of Fourier-based algorithms [Carr,Madan-1999; Fang,Oosterlee-2008] for efficient pricing of plain vanilla contracts.

This article is divided in several parts. In Section 2 we define the main hybrid models. In the follow-up section we discuss, in detail, approximations for the square root processes appearing. These approximations are later used in Section 4 when redefining the hybrid model. This is the heart of our article. In Section 5 we derive the characteristic functions of the introduced hybrid models. Section 6 is dedicated to the numerical experiments of pricing plain vanilla options under the hybrid models. In an appendix we show how to price swaptions and bonds and calibrate with a stochastic volatility short rate model.

2 Hybrid with Multi-Factor Short Rate Process

Suppose we have given two asset classes defined by the vectors $\mathbf{X}_t^{\bar{n}\times 1}$, $\bar{n}\in\mathbb{N}^+$ for the equity and for the interest rates $\mathbf{R}_t^{\bar{m}\times 1}$, $\bar{m}\in\mathbb{N}^+$. One can take high-dimensional processes involving stochastic volatility. In general, we can define the following system of governing stochastic differential equations (SDEs):

$$\begin{cases} \mathrm{d}\mathbf{X}_t = & a(\mathbf{X}_t,\mathbf{R}_t)\mathrm{d}t + & b(\mathbf{X}_t)\mathrm{d}\mathbf{W}_t^{\mathbf{X}}, \\ \mathrm{d}\mathbf{R}_t = & c(\mathbf{R}_t)\mathrm{d}t + & d(\mathbf{R}_t)\mathrm{d}\mathbf{W}_t^{\mathbf{R}}, \\ \mathbf{Z}_t\mathbf{Z}_t^{\mathrm{T}} = & \mathbf{C}^{\mathbf{H}}\mathrm{d}t, & \end{cases} \tag{2.1}$$

where $\mathbf{H_t}=[\mathbf{X}_t,\mathbf{R}_t]$, $\mathbf{Z}_t=[\mathrm{d}\mathbf{W}_t^X,\mathrm{d}\mathbf{W}_t^R]^{\mathrm{T}}$, $\mathbf{C}^{\mathbf{H}}$ is a $(\bar{n}+\bar{m})\times(\bar{n}+\bar{m})$ matrix which represents the instantaneous correlation between the Brownian motions. Moreover, since the noises $\mathrm{d}\mathbf{W}_t$ are assumed to be multi-dimensional, correlation within the asset classes is allowed, i.e.: $\mathbf{C}^{\mathbf{X}}=(\mathrm{d}\mathbf{W}_t^{\mathbf{X}})(\mathrm{d}\mathbf{W}_t^{\mathbf{X}})^{\mathrm{T}}$, $\mathbf{C}^{\mathbf{Y}}=(\mathrm{d}\mathbf{W}_t^{\mathbf{R}})(\mathrm{d}\mathbf{W}_t^{\mathbf{R}})^{\mathrm{T}}$.

Since the Heston model in [Heston-1993] is sufficiently complex for explaining the smile-shaped implied volatilities in equity, we take this model as a benchmark for the equity part. In particular, the

model for the state vector $\mathbf{X}_t = [x_t = \log S_t, \sigma_t]^{\mathrm{T}}$ is described by the following system of SDEs:

$$\begin{cases} \mathrm{d}x_t = & \left(r_t - \frac{1}{2}\sigma_t\right)\mathrm{d}t + & \sqrt{\sigma_t}\mathrm{d}W_t^x, & x_0 > 0, \\ \mathrm{d}\sigma_t = & \epsilon\left(\bar{\sigma} - \sigma_t\right)\mathrm{d}t + & \omega\sqrt{\sigma_t}\mathrm{d}W_t^\sigma, & \sigma_0 > 0, \\ \mathbf{C}_{1,2}^{\mathbf{X}} = & \rho_{x,\sigma}\mathrm{d}t, & & \end{cases} \tag{2.2}$$

with the speed of mean reversion $\epsilon > 0$; $\bar{\sigma} > 0$ is a long-term mean of stochastic volatility and $\omega > 0$ specifies the volatility of the stochastic volatility process. Note that term $\sigma_t/2$ in the x_t process results from Itô's formula (see for example [Øksendal-2000]) when deriving the dynamics for $\log S_t$.

For the interest rate process we consider multi-factor short rate processes. In the general setup, for a given state vector $\mathbf{R}_t = [r_t, v_t]^{\mathrm{T}}$ we consider the following system of SDEs:

$$\begin{cases} \mathrm{d}r_t = & \kappa(\theta_t + (1-p)v_t - r_t)\mathrm{d}t + & \eta\sqrt{v_t}^p\mathrm{d}W_t^r, & r_0 > 0, \\ \mathrm{d}v_t = & \lambda(\bar{v}p - v_t)\mathrm{d}t + & \gamma\sqrt{v_t}^p\mathrm{d}W_t^v, & \\ \mathbf{C}_{1,2}^{\mathbf{R}} = & \rho_{r,v}\mathrm{d}t, & & \end{cases} \tag{2.3}$$

with

$$\begin{cases} v_0 = 0 & \text{for } p = 0, \\ v_0 > 0 & \text{for } p = 1, \end{cases} \tag{2.4}$$

where $p = \{0, 1\}$, $\kappa > 0$, $\lambda > 0$ are the mean reversion parameters, $\eta > 0$ determines the volatility magnitude, while $\gamma > 0$ controls the volatility of the volatility process in interest rate. In the system above the coefficients $\theta_t > 0, \forall t$ and $\bar{v} > 0$ stand for long term interest rate (which usually is calibrated to the current yield curve) and the long term volatility level, respectively.

In model (2.3) parameter p needs some additional comments. Parameter p can lead to different models: for $p = 0$ the model becomes a Gaussian two-factor model [Brigo,Mercurio-2007] (G2++) (also known as a two-factor Hull-White model):

$$\begin{cases} \mathrm{d}r_t = & \kappa(\theta_t + v_t - r_t)\mathrm{d}t + & \eta\mathrm{d}W_t^r, & r_0 > 0, \\ \mathrm{d}v_t = & -\lambda v_t\mathrm{d}t + & \gamma\mathrm{d}W_t^v, & v_0 = 0, \\ \mathbf{C}_{1,2}^{\mathbf{R}} = & \rho_{r,v}\mathrm{d}t. & & \end{cases} \tag{2.5}$$

The G2++ model provides a satisfactory fit to the At-The-Money (ATM) humped structures of the volatility of the instantaneous forward rates. Moreover, simple model construction (multivariate normal distribution) provides closed form solutions for caps and swaptions, allowing fast calibration. On the other hand, since the model is assumed to be normal, the interest rates can become negative. In order to improve, an extended model can be applied. In this article, as the alternative model we choose the recently developed stochastic volatility short rate model of Heston type [Heidari, *et al.*-2007] (H2++). Therefore, by taking $p = 1$ the short rate process r_t is driven by the following system of SDEs:

$$\begin{cases} \mathrm{d}r_t = & \kappa(\theta_t - r_t)\mathrm{d}t + & \eta\sqrt{v_t}\mathrm{d}W_t^r, & r_0 > 0, \\ \mathrm{d}v_t = & \lambda(\bar{v} - v_t)\mathrm{d}t + & \gamma\sqrt{v_t}\mathrm{d}W_t^v, & v_0 > 0, \\ \mathbf{C}_{1,2}^{\mathbf{R}} = & \rho_{r,v}\mathrm{d}t, & & \end{cases} \tag{2.6}$$

By taking the equity model $\mathbf{X}_t$ as introduced in (2.2) and the interest rate part $\mathbf{R}_t$ from (2.3) we define a hybrid model $\mathbf{H}_t^p = [\mathbf{X}_t, \mathbf{Y}_t]^{\mathrm{T}}$ with the following instantaneous correlation structure:

$$\mathbf{C}^{\mathbf{H}} = \left(\begin{array}{cc|cc} 1 & \rho_{x,\sigma} & \rho_{x,r} & \rho_{x,v} \\ * & 1 & \rho_{\sigma,r} & \rho_{\sigma,v} \\ \hline * & * & 1 & \rho_{r,v} \\ * & * & * & 1 \end{array}\right)_{4\times 4}. \tag{2.7}$$

Model $\mathbf{H}_t^p$ for $p = 0$ is the Heston-Gaussian two-factor hybrid model (H-G2++) and for $p = 1$ it becomes the Heston-Heston two-factor hybrid model (H-H2++). Note that the equity and the interest

rate asset classes are linked by correlations in the right-upper and left-lower diagonal blocks of matrix $\mathbf{C^H}$. Our main objective is the preservation of the correlation, $\rho_{x,r}$, between the log-equity and the interest rate. However, we also show that additional correlations (degrees of freedom) may be used in the calibration to better fit to the market quotes.

Assuming $\mathbf{V} := \mathbf{V}(t, x_t, \sigma_t, r_t, v_t)$ to represent the value of a European claim, with the help of the arbitrage free pricing theorem and the use of Itô's formula, we can derive the corresponding pricing Partial Differential Equation (PDE) [Gatheral-2006]:

$$\begin{aligned} 0 \quad = \quad & (r - 1/2\sigma)\,\mathbf{V}_x + \epsilon\,(\bar{\sigma} - \sigma)\,\mathbf{V}_\sigma + \kappa(\theta(t) + (1-p)v - r)\mathbf{V}_r + \lambda\,(\bar{v}p - v)\,\mathbf{V}_v \\ & + \frac{1}{2}\sigma\mathbf{V}_{xx} + \frac{1}{2}\omega^2\sigma\mathbf{V}_{\sigma\sigma} + \frac{1}{2}\eta^2 v^p\mathbf{V}_{rr} + \frac{1}{2}\gamma^2 v^p\mathbf{V}_{vv} + \rho_{x,\sigma}\omega\sigma\mathbf{V}_{x\sigma} \\ & + \rho_{x,r}\eta\sqrt{\sigma}\sqrt{v}^p\mathbf{V}_{xr} + \rho_{x,v}\gamma\sqrt{\sigma}\sqrt{v}^p\mathbf{V}_{xv} + \rho_{\sigma,r}\eta\omega\sqrt{\sigma}\sqrt{v}^p\mathbf{V}_{\sigma r} \\ & + \rho_{\sigma,v}\gamma\omega\sqrt{\sigma}\sqrt{v}^p\mathbf{V}_{\sigma v} + \rho_{r,v}\gamma\eta v^p\mathbf{V}_{rv} - r\mathbf{V} + \mathbf{V}_t, \end{aligned} \tag{2.8}$$

where subscripts denote partial derivatives and with state variables $x := x_t = \log S_t$, $\sigma := \sigma_t$, $r := r_t$, $v := v_t$, and specific boundary and final conditions (for details on boundary conditions for similar problems, see, for example, [Duffie-2006] pp.241).

The solution of the 4D convection-diffusion PDE above can be approximated by means of standard numerical techniques, like finite differences (see for example [Morton-2005]). This may however result in substantial CPU time needed for the model evaluation. An alternative is to use the Feynman-Kac theorem and reformulate the problem to an integral equation related to the discounted expected payoff. Before doing this, let us derive the associated instantaneous covariance matrix $\Sigma_{\mathbf{H}}\Sigma_{\mathbf{H}}^{\mathrm{T}}$ of hybrid model (2.1) with (2.2) and (2.3):

$$\mathbf{S} := \mathbf{\Sigma_H}\mathbf{\Sigma_H^T} = \left(\begin{array}{cc|cc} \sigma_t & \rho_{x,\sigma}\omega\sigma_t & \rho_{x,r}\eta\sqrt{\sigma_t}\sqrt{v_t}^p & \rho_{x,v}\gamma\sqrt{\sigma_t}\sqrt{v_t}^p \\ * & \omega^2\sigma_t & \rho_{\sigma,r}\eta\omega\sqrt{\sigma_t}\sqrt{v_t}^p & \rho_{\sigma,v}\omega\gamma\sqrt{\sigma_t}\sqrt{v_t}^p \\ \hline * & * & \eta^2 v_t^p & \rho_{r,v}\eta\gamma v_t^p \\ * & * & * & \gamma^2 v_t^p \end{array}\right)_{\mathbf{4\times 4}}. \tag{2.9}$$

In the next section we discuss in detail the problems associated with the non-affine coefficients in matrix (2.9).

3 The Square Root Process

Let us have a closer look at the elements in the matrix (2.9). There are similarities between the non-affine elements. Each of them involves the term $\sqrt{\sigma_t}\sqrt{v_t}^p$. Moreover, for both cases the volatility processes σ_t and v_t are of square-Bessel mean-reverting CIR type [Cox, *et al.*-1985]. These two processes are guaranteed to be positive if the Feller conditions [Feller-1971], i.e., $2\epsilon\bar{\sigma} \geq \omega^2$ for σ_t and $2\lambda\bar{v} \geq \gamma^2$ for v_t are satisfied. Since the processes σ_t and v_t are of the same type, we continue the analysis of σ_t; the results for v_t are analogous.

It is shown in [Cox, *et al.*-1985; Broadie,Kaya-2006] that, for a given time $t > 0$, σ_t is distributed as $c(t)$ times a non-central chi-squared random variable, $\chi^2(d, \lambda(t))$, with d the "degrees of freedom" parameter and non-centrality parameter $\lambda(t)$, i.e.:

$$\sigma_t = c(t)\chi^2\left(d, \lambda(t)\right), \quad t > 0, \tag{3.1}$$

with

$$c(t) = \frac{1}{4\epsilon}\omega^2(1 - \mathrm{e}^{-\epsilon t}), \quad d = \frac{4\epsilon\bar{\sigma}}{\omega^2}, \quad \lambda(t) = \frac{4\epsilon\sigma_0\mathrm{e}^{-\epsilon t}}{\omega^2(1 - \mathrm{e}^{-\epsilon t})}. \tag{3.2}$$

So, the corresponding cumulative distribution function (CDF) can be expressed as:

$$F_{\sigma_t}(x) = \mathbb{P}(\sigma_t \leq x|\sigma_0) = \mathbb{P}\left(\chi^2\left(d, \lambda(t)\right) \leq x/c(t)|\sigma_0\right) = F_{\chi^2(d,\lambda(t))}\left(x/c(t)\right), \tag{3.3}$$

where:

$$F_{\chi^2(d,\lambda(t))}(y) = \sum_{k=0}^{\infty}\exp\left(-\frac{\lambda(t)}{2}\right)\frac{\left(\frac{\lambda(t)}{2}\right)^k}{k!}\frac{\Gamma\left(k + \frac{d}{2}, \frac{y}{2}\right)}{\Gamma\left(k + \frac{d}{2}\right)}, \tag{3.4}$$

with

$$\Gamma(a,z) = \int_0^z t^{a-1}\mathrm{e}^{-t}\mathrm{d}t, \quad \Gamma(z) = \int_0^\infty t^{z-1}\mathrm{e}^{-t}\mathrm{d}t. \tag{3.5}$$

Further, the corresponding density function (see for example [Moser-2007]) reads:

$$f_{\chi^2(d,\lambda(t))}(y) = \frac{1}{2}\mathrm{e}^{-\frac{1}{2}(y+\lambda(t))}\left(\frac{y}{\lambda(t)}\right)^{\frac{1}{2}\left(\frac{d}{2}-1\right)}\mathcal{B}_{\frac{d}{2}-1}(\sqrt{\lambda(t)y}), \tag{3.6}$$

with

$$\mathcal{B}_a(z) = \left(\frac{z}{2}\right)^a \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(a+k+1)}, \tag{3.7}$$

which is a modified Bessel function of the first kind (see for example [Abramowitz-1972; Gradshteyn,Ryzhik-1996]).

From above, the density for σ_t can be expressed as:

$$f_{\sigma_t}(x) \overset{def}{=} \frac{\mathrm{d}}{\mathrm{d}x}F_{\sigma_t}(x) = \frac{\mathrm{d}}{\mathrm{d}x}F_{\chi^2(d,\lambda(t))}(x/c(t)) = \frac{1}{c(t)}f_{\chi(d,\lambda(t))}\left(x/c(t)\right). \tag{3.8}$$

By using the properties of the non-central chi-square distribution the mean and variance of the process σ_t are known explicitly:

$$\begin{aligned} \mathbb{E}(\sigma_t|\sigma_0) = &\quad c(t)(d+\lambda(t)), \\ \mathbb{V}\mathrm{ar}(\sigma_t|\sigma_0) = &\quad c^2(t)(2d+4\lambda(t)). \end{aligned} \tag{3.9}$$

In the lemma below we derive the corresponding moments for $\sqrt{\sigma_t}$.

Lemma 3.1 (Expectation and variance for $\sqrt{\sigma_t}$)**.** *For a given time $t > 0$ the expectation and variance of $\sqrt{\sigma_t}$ where σ_t has a non-central chi-square distribution function with CDF in (3.4) are given by:*

$$\mathbb{E}(\sqrt{\sigma_t}|\sigma_0) = \sqrt{2c(t)}\mathrm{e}^{-\lambda(t)/2}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\lambda(t)/2\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma(\frac{d}{2}+k)}, \tag{3.10}$$

and

$$\mathbb{V}\mathrm{ar}\left(\sqrt{\sigma_t}|\sigma_0\right) = c(t)(d+\lambda(t)) - 2c(t)\mathrm{e}^{-\lambda(t)}\left(\sum_{k=0}^{\infty}\frac{1}{k!}\left(\lambda(t)/2\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma\left(\frac{d}{2}+k\right)}\right)^2. \tag{3.11}$$

Proof. First of all by [Dufresne-2001] we have that:

$$\begin{aligned} \mathbb{E}(\sqrt{\sigma_t}|\sigma_0) &\overset{def}{=} \int_0^\infty \frac{\sqrt{x}}{c(t)}f_{\chi^2(d,\lambda(t))}\left(\frac{x}{c(t)}\right)\mathrm{d}x \\ &= \sqrt{2c(t)}\frac{\Gamma\left(\frac{1+d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}{}_1F_1\left(-\frac{1}{2},\frac{d}{2},-\frac{\lambda(t)}{2}\right), \end{aligned} \tag{3.12}$$

where ${}_1F_1(a;b;z)$ is a confluent hyper-geometric function, which is also known as Kummer's function [Kummer-1936] of the first kind, given by:

$${}_1F_1(a;b;z) = \sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{z^k}{k!}, \tag{3.13}$$

with $(a)_k$ and $(b)_k$ being Pochhammer symbols of the form:

$$(a)_k = \frac{\Gamma(a+k)}{\Gamma(a)} = a(a+1)\cdot\cdots\cdot(a+k-1). \tag{3.14}$$

Now, using the principle of Kummer (see [Koepf-1998] pp.42) we find:

$$ {}_1F_1\left(-\frac{1}{2},\frac{d}{2},-\frac{\lambda(t)}{2}\right)=\mathrm{e}^{-\lambda(t)/2}\,{}_1F_1\left(\frac{1+d}{2},\frac{d}{2},\frac{\lambda(t)}{2}\right) \tag{3.15} $$

Therefore, by (3.14) and (3.15), Equation (3.12) reads:

$$ \begin{aligned} \mathbb{E}(\sqrt{\sigma_t}|\sigma_0) &= \sqrt{2c(t)}\mathrm{e}^{-\lambda(t)/2}\frac{\Gamma\left(\frac{1+d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\,{}_1F_1\left(\frac{1+d}{2},\frac{d}{2},\frac{\lambda(t)}{2}\right) && (3.16)\\ &= \sqrt{2c(t)}\mathrm{e}^{-\lambda(t)/2}\frac{\Gamma\left(\frac{1+d}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\lambda(t)/2\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma\left(\frac{1+d}{2}\right)}\frac{\Gamma\left(\frac{d}{2}\right)}{\Gamma\left(\frac{d}{2}+k\right)} && (3.17)\\ &= \sqrt{2c(t)}\mathrm{e}^{-\lambda(t)/2}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\lambda(t)/2\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma\left(\frac{d}{2}+k\right)}. && (3.18) \end{aligned} $$

In order to calculate the variance we have:

$$ \mathbb{V}\mathrm{ar}\left(\sqrt{\sigma_t}|\sigma_0\right) \overset{def}{=} \mathbb{E}(\sigma_t|\sigma_0)-\left(\mathbb{E}(\sqrt{\sigma_t})|\sigma_0\right)^2. \tag{3.19} $$

Now, by combining (3.9) and (3.19) with the already derived $\mathbb{E}(\sqrt{\sigma_t}|\sigma_0)$ we obtain:

$$ \mathbb{V}\mathrm{ar}\left(\sqrt{\sigma_t}|\sigma_0\right)=c(t)(d+\lambda(t))-2c(t)\mathrm{e}^{-\lambda(t)}\left(\sum_{k=0}^{\infty}\frac{1}{k!}\left(\lambda(t)/2\right)^k\frac{\Gamma\left(\frac{1+d}{2}+k\right)}{\Gamma\left(\frac{d}{2}+k\right)}\right)^2, \tag{3.20} $$

which concludes the proof. □

3.1 Approximations for the Square Root Process

We develop here the analysis of the distribution of $\sqrt{\sigma_t}$. As a start, in the first lemma we show that for $t\to\infty$ the random variable $\sqrt{\sigma_t}$, with σ_t a square root process as in (2.2), can be approximated by a normally distributed random variable.

Lemma 3.2 (Normal approximation for $\sqrt{\sigma_t}$ for $t\to\infty$)**.** *For $t\to\infty$, we find:*

$$ \sqrt{\sigma_\infty}:=\lim_{t\to\infty}\sqrt{\sigma_t}\approx\mathcal{N}\left(\sqrt{\bar{\sigma}-\frac{\omega^2}{8\epsilon}},\frac{\omega^2}{8\epsilon}\right), \tag{3.21} $$

with $\bar{\sigma}$, ω and ϵ as in (2.2), and $\mathcal{N}(\mu,\sigma)$ a normally distributed random variable with expectation μ and variance σ.

Proof. It was already shown that the CDF for σ_t is given by:

$$ F_{\sigma_t}(x)=F_{\chi^2(d,\lambda(t))}\left(x/c(t)\right), \tag{3.22} $$

where d, $\lambda(t)$ and $c(t)$ are given in (3.2). For $t\to\infty$ we find the limits:

$$ \lim_{t\to\infty}\lambda(t)=0,\ \text{and}\ \lim_{t\to\infty}c(t)=\frac{\omega^2}{4\epsilon}. \tag{3.23} $$

For $t\to\infty$ the non-centrality parameter converges to 0, so, by using the properties of the non-central chi-square distribution, σ_t converges to the standard chi-square distribution, i.e.,:

$$ \lim_{t\to\infty}\sigma_t= c(\infty)\chi^2(d,\lambda(\infty))\overset{def}{=}\frac{\omega^2}{4\epsilon}\chi^2(d), \tag{3.24} $$

where $\chi^2(d)$ is a chi-square random variable with d degrees of freedom.

We use the result by Fisher [Fisher-1922] 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.:

$$F_{\chi^2(d)}(x) \approx \Phi\left(\sqrt{2x} - \sqrt{2d-1}\right). \tag{3.25}$$

From (3.24) we have:

$$\chi^2(d) = \frac{4\epsilon}{\omega^2}\sigma_\infty. \tag{3.26}$$

It follows that:

$$\sqrt{\frac{8\epsilon}{\omega^2}\sigma_\infty} \approx \mathcal{N}\left(\sqrt{\frac{8\epsilon\bar{\sigma}}{\omega^2}} - 1, 1\right).$$

So, finally we conclude that

$$\sqrt{\sigma_\infty} \approx \mathcal{N}\left(\sqrt{\bar{\sigma} - \frac{\omega^2}{8\epsilon}}, \frac{\omega^2}{8\epsilon}\right), \tag{3.27}$$

which finishes the proof. □

The lemma provides us with an intuition about the behavior of the distribution of $\sqrt{\sigma_t}$, however, since for finite time t the non-centrality parameter $\lambda(t)$ does not converge to zero another approximation for finite time t is necessary. In the lemma below this approximation, for finite time t and a non-zero centrality parameter, is presented.

Lemma 3.3 (Normal Approximation for $\sqrt{\sigma_t}$ for $0 < t < \infty$)**.** *For any time, $t < \infty$, the square root of σ_t in (2.2) can be approximated by*

$$\sqrt{\sigma_t} \approx \mathcal{N}\left(\sqrt{c(t)(\lambda(t)-1) + c(t)d + \frac{c(t)d}{2(d+\lambda(t))}}, c(t) - \frac{c(t)d}{2(d+\lambda(t))}\right), \tag{3.28}$$

with $c(t)$, d and $\lambda(t)$ from (3.2). Moreover, for a fixed value of x in the cumulative distribution function $F_{\sqrt{\sigma(t)}}(x)$, 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 $\lambda(t) \to \infty$.

Proof. As given in [Patnaik-1949] an accurate approximation for the non-central chi-square distribution, $\chi^2_d(\lambda(t))$, can be obtained by an approximation with a centralized chi-square distribution, i.e.:

$$\chi^2(d, \lambda(t)) \approx a(t)\chi^2(f(t)), \tag{3.29}$$

with $a(t)$ and $f(t)$ in (3.29) chosen so that the first two moments match, i.e.:

$$a(t) = \frac{d+2\lambda(t)}{d+\lambda(t)}, \quad f(t) = d + \frac{\lambda(t)^2}{d+2\lambda(t)}. \tag{3.30}$$

From (3.1) and (3.29) we have:

$$\sqrt{\sigma_t} \approx \sqrt{c(t)}\sqrt{a(t)\chi^2(f(t))}. \tag{3.31}$$

Therefore, by (3.25),

$$\sqrt{\sigma_t} \approx \mathcal{N}\left(\sqrt{\left(f(t) - \frac{1}{2}\right)c(t)a(t)}, \frac{1}{2}c(t)a(t)\right). \tag{3.32}$$

The order of this approximation can be found in [Johnson, *et al.*-1994]. □

Remark 1 (Delta method and moments approximations)**.** *The mean, $\mathbb{E}(\sqrt{\sigma_t})$, and variance, $\mathbb{V}\mathrm{ar}(\sqrt{\sigma_t})$, provided in Lemma 3.3 are equivalent with first first order approximations obtained by the delta method, discussed in [Grzelak,Oosterlee-2009].*

Remark 2 (Remark on Accuracy)**.** *As already indicated in [Patnaik-1949] the normal approximation resembles the non-central chi-square distribution very well for either a large number of degrees of freedom, d, or a large non-centrality $\lambda(t)$. As already mentioned, for $t \to 0$, the non-centrality parameter, $\lambda(t)$, tends to infinity. Therefore, accurate approximations are expected. In the case of large maturities, the non-centrality parameter converges to 0, which may give an inaccurate approximation. In that case, satisfactory results depend on the size of the degrees of freedom parameter d. It is clear that d in (3.2) is directly related to the Feller condition which by an inequality ensures the process to be positive. In practical applications however $2\epsilon\bar{\sigma}$ is often lower than ω^2. In Section 6 we check, by a numerical experiment, the impact of not satisfying the Feller condition on our approximations.*

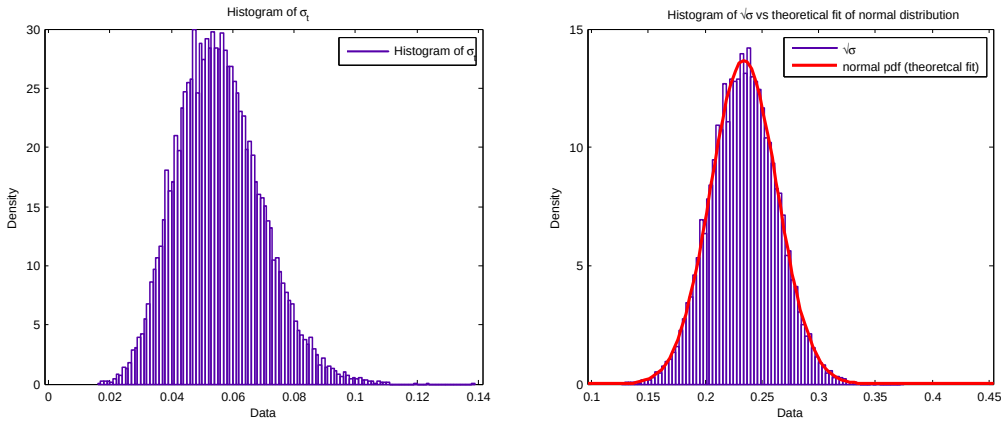

Figure 3.1: For maturity T $= 0.1$; LEFT: a histogram for σ_t, RIGHT: a histogram for $\sqrt{\sigma_t}$ and the theoretical fit of normal distribution. The Monte Carlo simulation was performed with 20.000 paths with 500 steps for $\epsilon = 1.2$, $\omega = 0.2$, $\bar{\sigma} = 0.1$, $\sigma_0 = 0.05$.

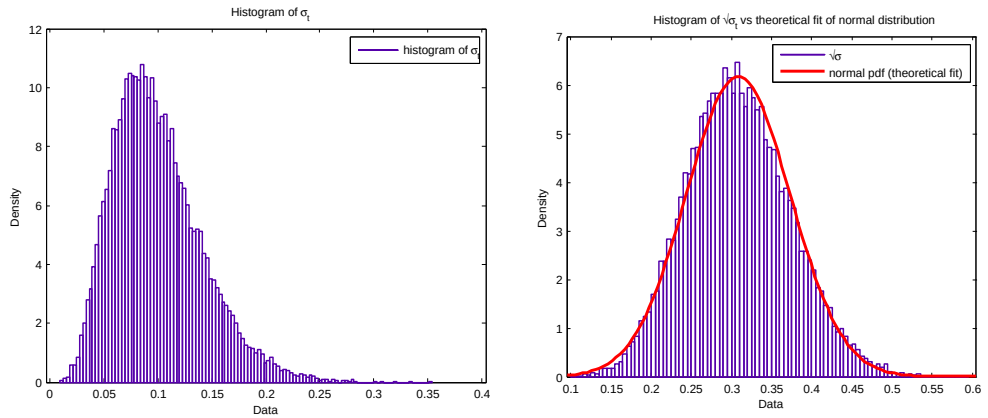

Figure 3.2: For maturity T $= 5$; LEFT: a histogram for σ_t, RIGHT: a histogram for $\sqrt{\sigma_t}$ and the theoretical fit of normal distribution. The Monte Carlo setup was chosen the same as in Figure 3.1.

Figures 3.1 and 3.2 show that, for the non-central chi-square distributed σ_t given by (2.2), $\sqrt{\sigma_t}$ indeed resembles a normal distribution very well for finite maturities, T $= 0.1$ and T $= 5$. In Table 3.1 the accuracy results of the approximations of the first two moments are presented.

Table 3.1: The expectation and variance of $\sqrt{\sigma_T}$ for T $= 0.1$ and T $= 5$ obtained from Monte Carlo simulation (20.000 paths and 500 steps), approximation given in Lemma 3.3 and exact representation from Lemma 3.1. For the exact representation we have truncated the sum at $\widetilde{n} = 100$.

Maturity	$\mathbb{E}(\cdot)$, $\mathbb{V}\mathrm{ar}(\cdot)$	Monte Carlo	Proxy (Lem. 3.3)	Exact (Lem. 3.1)
T $= 0.1$	$\mathbb{E}(\sqrt{\sigma_t})$	0.23602	0.23411	0.23410
	$\mathbb{V}\mathrm{ar}(\sqrt{\sigma_t})$	8.76694E-4	8.46600E-4	8.50367E-4
T $= 5$	$\mathbb{E}(\sqrt{\sigma_t})$	0.30896	0.30938	0.30952
	$\mathbb{V}\mathrm{ar}(\sqrt{\sigma_t})$	0.00432	0.00416	0.00407

3.2 Approximation of the Dynamics $\mathrm{d}\sqrt{\sigma_t}$

Now, let us have a closer look at the dynamics of process $\sqrt{\sigma}_t$ from another perspective. By using Itô's formula one can simply get:

$$\mathrm{d}\sqrt{\sigma_t} = \left(-\frac{\omega^2}{8\sqrt{\sigma_t}} + \frac{\epsilon}{2}\left(\frac{\bar{\sigma}}{\sqrt{\sigma_t}} - \sqrt{\sigma_t}\right)\right)\mathrm{d}t + \frac{\omega}{2}\mathrm{d}W_t^{\sigma}, \quad \sqrt{\sigma_0} > 0. \tag{3.33}$$

The dynamics of the process $\sqrt{\sigma_t}$ presented above include a constant volatility coefficient, $\omega/2$, whereas the drift is of a non-affine form. In the previous subsection, in Lemma 3.3, we have shown that for any finite time t, $\sqrt{\sigma_t}$ may be simply approximated by a normal distribution. Instead of using the approximations for expectation and variance given in Lemma 3.3 we can use the exact moment estimates, i.e.:

$$\sqrt{\sigma_t} \approx \mathcal{N}\left(\mathbb{E}(\sqrt{\sigma_t}|\sigma_0), \mathbb{V}\mathrm{ar}(\sqrt{\sigma_t}|\sigma_0)\right), \tag{3.34}$$

with exact expectation $\mathbb{E}(\sqrt{\sigma_t}|\sigma_0)$ and variance $\mathbb{V}\mathrm{ar}(\sqrt{\sigma_t}|\sigma_0)$ as given in Lemma 3.1.

The distribution for $\sqrt{\sigma_t}$, is then used to construct an approximation for the dynamics of $\sqrt{\sigma_t}$ by matching the corresponding moments. The lemma below provides the necessary coefficients.

Lemma 3.4 (Construction of the dynamics of $\sqrt{\sigma_t}$)**.** *For a given stochastic volatility process, σ_t, from (2.2) the process u_t modeled by (3.35) has the same first two moments as $\sqrt{\sigma_t}$, if:*

$$\mathrm{d}u_t = \mu_t^u \mathrm{d}t + \psi_t^u \mathrm{d}W_t^{\sigma}, \qquad u_0 = \sqrt{\sigma_0} > 0, \tag{3.35}$$

with deterministic time-dependent drift, μ_t^u, and volatility, ψ_t^u, given by:

$$\begin{aligned}\mu_t^u = \frac{1}{2\sqrt{2}}\frac{\Gamma\left(\frac{1+d}{2}\right)}{\sqrt{c(t)}}\Bigg({}_1\widetilde{F}_1\left(-\frac{1}{2}, \frac{d}{2}, -\frac{\lambda(t)}{2}\right)\frac{1}{2}\omega^2 \mathrm{e}^{-t\epsilon} \\ - {}_1\widetilde{F}_1\left(\frac{1}{2}, \frac{2+d}{2}, -\frac{\lambda(t)}{2}\right)\frac{4c(t)\epsilon^2\sigma_0\mathrm{e}^{\epsilon t}}{\left(\mathrm{e}^{\epsilon t}-1\right)^2\omega^2}\Bigg),\end{aligned} \tag{3.36}$$

and:

$$\psi_t^u = \left(\frac{1}{4}\omega^2\mathrm{e}^{-t\epsilon}(d+\lambda(t)) - \frac{4c(t)\epsilon^2\sigma_0\mathrm{e}^{\epsilon t}}{\left(\mathrm{e}^{\epsilon t}-1\right)^2\omega^2} - 2\mathbb{E}(\sqrt{\sigma_t}|\sqrt{\sigma_0})\,\mu_t^u\right)^{\frac{1}{2}}.$$

Here, $\mathbb{E}(\sqrt{\sigma_t}|\sigma_0)$ is given by Lemma 3.1, d, $c(t)$ and $\lambda(t)$ are as in (3.2) and the regularized hypergeometric function ${}_1\widetilde{F}_1(a;b;z) =: {}_1F_1(a;b;z)/\Gamma(b)$.

Proof. By integrating Equation (3.35) we obtain:

$$u_t = u_0 + \int_0^t \mu_s^u \mathrm{d}s + \int_0^t \psi_s^u \mathrm{d}W_s^{\sigma}, \qquad u_0 = \sqrt{\sigma_0} > 0. \tag{3.37}$$

By taking the conditional expectation and variance from both sides we get:

$$\begin{cases} \mathbb{E}(u_t|u_0) = & u_0 + \int_0^t \mu_s^u \mathrm{d}s, \\ \mathbb{V}\mathrm{ar}(u_t|u_0) = & \int_0^t \mathbb{E}\left(\psi_s^u\right)^2 \mathrm{d}s. \end{cases} \tag{3.38}$$

Now, by matching these equations with $\mathbb{E}(\sqrt{\sigma_t}|\sqrt{\sigma_0})$ and $\mathbb{V}\mathrm{ar}(\sqrt{\sigma_t}|\sqrt{\sigma_0})$, respectively, the following equalities need to be satisfied:

$$\begin{cases} \mu_t^u = & \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(\sqrt{\sigma_t}|\sqrt{\sigma_0}), \\ \psi_t^u = & \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{V}\mathrm{ar}\left(\sqrt{\sigma_t}|\sqrt{\sigma_0}\right)\right)^{\frac{1}{2}}. \end{cases} \tag{3.39}$$

With the expressions in the RHS of (3.39), determined in Lemma 3.1, the proof is finished after some basic algebra. □

In Lemma 3.4 we have derived the functional form of the coefficients μ_t^u and ψ_t^u needed for the construction of process $\mathrm{d}u_t$ which serves as an approximation for process $\mathrm{d}\sqrt{\sigma_t}$. Both coefficients are determined in terms of the hyper-geometric function ${}_1F_1(\cdot,\cdot,\cdot)$ which is given by an infinite sum (see Equation (3.13)). Since, we are interested in finite expressions, a truncation of the infinite sum is necessary. Numerically, we give an indication of the number of terms needed for an accurate function evaluation. Figure 6.2 shows that a high accuracy is already obtained with the first few terms in the summation. For both functions we observe a tendency to faster convergence for higher maturities. In Remark 3 we indicate how to deal with an increasing error for $t \to 0$.

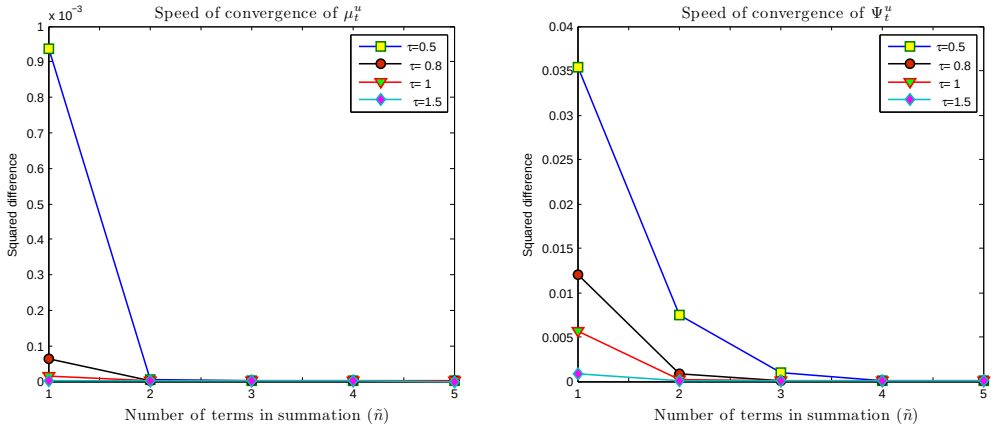

Figure 3.3: Speed of convergence related to ${}_1F_1$ for different maturities. LEFT: μ_t^u, RIGHT: ψ_t^u. The parameters were chosen as $\epsilon = 1.5$, $\omega = 0.4$, $\bar{\sigma} = 0.2$ and $\sigma_0 = 0.05$.

Remark 3. *Since for $t \to 0$, $\lambda(t) \to \infty$ and $c(t) \to 0$ some numerical instabilities in calculating the moments in Lemma 3.1 may arise. Therefore for $t \to 0$ we advise the use of the estimates provided in Lemma 3.3.*

4 Construction of the Affine Hybrid Model

For both hybrid models, the H-G2++ model with $p = 0$ and H-H2++ for $p = 1$, the instantaneous covariance matrices in (2.9) are not affine in all terms on the right-upper block. One can immediately see that the affinity problem disappears for $\rho_{x,r} = 0$, $\rho_{x,v} = 0$, $\rho_{\sigma,r} = 0$ and $\rho_{\sigma,v} = 0$. This, however,

introduces independence between the asset classes. In order to stay in the affine class while assuming the correlations between the assets to be non-zero some approximations need to be introduced.

In this section we use the approximation introduced for process $\mathrm{d}\sqrt{\sigma_t}$ to find a proxy for the hybrid models under consideration, so that we obtain *affine hybrid multi-factor models*. In matrix (2.9) the terms that cause affinity problems are all of the same form. For example, element $(1,3)$ from matrix $\mathbf{S}$ is given by:

$$\mathbf{S}_{(1,3)} = \rho_{x,r}\eta\sqrt{\sigma_t}\sqrt{v_t}^p,$$

where $\rho_{x,r} \neq 0$, $\eta > 0$, $p = \{0,1\}$ and both processes, σ_t and v_t, are CIR-type processes given by (2.2). Now, based on the results from the previous section, let us define an approximation for $\mathbf{S}_{(1,3)}$:

$$\mathbf{S}_{(1,3)} \approx \widetilde{\mathbf{S}}_{(1,3)} = \rho_{x,r}\eta u_t^\sigma \left(u_t^v\right)^p, \tag{4.1}$$

where the processes u_t^σ, u_t^v are given by:

$$\begin{cases} \mathrm{d}u_t^\sigma = & \mu_t^\sigma \mathrm{d}t + & \psi_t^\sigma \mathrm{d}W_t^\sigma, & u_0^\sigma = \sqrt{\sigma_0} > 0, \\ \mathrm{d}u_t^v = & \mu_t^v \mathrm{d}t + & \psi_t^v \mathrm{d}W_t^v, & u_0^v = \sqrt{v_0} > 0, \\ \rho_{\sigma,v}\mathrm{d}t = & \mathrm{d}W_t^\sigma \mathrm{d}W_t^v, & & \end{cases} \tag{4.2}$$

with μ_t^σ, μ_t^v as in Equation (3.36), ψ_t^σ, and ψ_t^v as in Equation (3.37). Depending on the values of parameter p we can distinguish the following cases:

$$\begin{cases} p = 0: & \widetilde{\mathbf{S}}_{(1,3)} = & \rho_{x,r}\eta u_t^\sigma, \\ p = 1: & \widetilde{\mathbf{S}}_{(1,3)} = & \rho_{x,r}\eta u_t^\sigma u_t^v. \end{cases} \tag{4.3}$$

For the simplest case, $p = 0$, $\widetilde{\mathbf{S}}_{(1,3)}$ is already in the affine form, however, for $p = 1$ it is not. In order to *repair* this for the case of $p = 1$, we introduce an additional variable

$$z_t := u_t^\sigma u_t^v,$$

for which we find the following dynamics:

$$\mathrm{d}z_t = \left(\mu_t^\sigma u_t^v + \mu_t^v u_t^\sigma + \psi_t^\sigma \psi_t^v \rho_{\sigma,v}\right)\mathrm{d}t + u_t^v \psi_t^\sigma \mathrm{d}W_t^\sigma + u_t^\sigma \psi_t^v \mathrm{d}W_t^v. \tag{4.4}$$

Process (4.4) gives rise to additional affinity problems, but since we have defined the approximations $u_t^\sigma \approx \sqrt{\sigma_t}$ and $u_t^v \approx \sqrt{v_t}$, we can simply write:

$$\mathrm{d}z_t = \left(\mu_t^\sigma u_t^v + \mu_t^v u_t^\sigma + \psi_t^\sigma \psi_t^v \rho_{\sigma,v}\right)\mathrm{d}t + \sqrt{v_t}\psi_t^\sigma \mathrm{d}W_t^\sigma + \sqrt{\sigma_t}\psi_t^v \mathrm{d}W_t^v. \tag{4.5}$$

Now, by combining Equations (2.2), (2.3) with (4.2) and (4.5) we end up with the approximate hybrid model of our choice:

$$\text{Hybrid Model} \begin{cases} \text{Equity} & \begin{cases} \mathrm{d}x_t = & \left(r_t - \frac{1}{2}\sigma_t\right)\mathrm{d}t + & \sqrt{\sigma_t}\mathrm{d}W_t^x, \\ \mathrm{d}\sigma_t = & \epsilon\left(\bar{\sigma} - \sigma_t\right)\mathrm{d}t + & \omega\sqrt{\sigma_t}\mathrm{d}W_t^\sigma, \end{cases} \\ \text{Short Rate} & \begin{cases} \mathrm{d}r_t = & \kappa(\theta_t + (1-p)v_t - r_t)\mathrm{d}t + & \eta\sqrt{v_t}^p \mathrm{d}W_t^r, \\ \mathrm{d}v_t = & \lambda(\bar{v}p - v_t)\mathrm{d}t + & \gamma\sqrt{v_t}^p \mathrm{d}W_t^v, \end{cases} \\ \text{Affinity correction} & \begin{cases} \mathrm{d}u_t^\sigma = & \mu_t^\sigma \mathrm{d}t + & \psi_t^\sigma \mathrm{d}W_t^\sigma, \\ \mathrm{d}u_t^v = & \mu_t^v \mathrm{d}t + & \psi_t^v \mathrm{d}W_t^v, \\ \mathrm{d}z_t = & \Delta_t \mathrm{d}t + & \sqrt{v_t}\psi_t^\sigma \mathrm{d}W_t^\sigma + \sqrt{\sigma_t}\psi_t^v \mathrm{d}W_t^v, \end{cases} \end{cases} \tag{4.6}$$

with Δ_t given by:

$$\Delta_t = \mu_t^\sigma u_t^v + \mu_t^v u_t^\sigma + \psi_t^\sigma \psi_t^v \rho_{\sigma,v},$$

and

$$\begin{cases} u_t^\sigma \approx & \sqrt{\sigma_t}, & u_0^\sigma = \sqrt{\sigma_0} > 0, \\ u_t^v \approx & \sqrt{v_t}, & u_0^v = \sqrt{v_0} > 0, \\ z_t = & u_t^\sigma u_t^v \approx \sqrt{\sigma_t}\sqrt{v_t}, & z_0 = \sqrt{\sigma_0}\sqrt{v_0} > 0. \end{cases} \tag{4.7}$$

Here $x_0 > 0$ and $r_0 > 0$, and deterministic time-dependent functions μ_t^σ, μ_t^v, ψ_t^σ, ψ_t^v are given by Lemma 3.4 with the remaining parameters detailed in Section 2 and a full matrix of correlations given by (2.7). Since we have defined some additional processes which correspond to the affinity correction, our original 4D state vector, $\mathbf{X}_t = [x_t, \sigma_t, r_t, v_t]^{\mathrm{T}}$ now turns into a 5D vector $\widetilde{\mathbf{H}}_t^0 := \widetilde{\mathbf{H}}_t^{p=0} = [x_t, \sigma_t, r_t, v_t, u_t^\sigma]^{\mathrm{T}}$ and a 7D vector $\widetilde{\mathbf{H}}_t^1 := \widetilde{\mathbf{H}}_t^{p=1} = [x_t, \sigma_t, r_t, v_t, u_t^\sigma, u_t^v, z_t]^{\mathrm{T}}$.

Let us define a state vector $\mathbf{J}_t(p) = [x_t, \sigma_t, r_t, v_t, u_t^\sigma]^{\mathrm{T}}$. Note that, $\mathbf{J}_t(0) = \widetilde{\mathbf{H}}_t^0$. For model $\mathbf{J}_t(p)$ we find the following instantaneous covariance matrix:

$$\boldsymbol{\Sigma}_{\mathbf{J}}\boldsymbol{\Sigma}_{\mathbf{J}}^{\mathrm{T}}(p) = \left(\begin{array}{cc|cc|c} \sigma_t & \rho_{x,\sigma}\omega\sigma_t & \rho_{x,r}\eta\sqrt{\sigma_t}\sqrt{v_t}^p & \rho_{x,v}\gamma\sqrt{\sigma_t}\sqrt{v_t}^p & \rho_{x,\sigma}\psi_t^\sigma\sqrt{\sigma_t} \\ * & \omega^2\sigma_t & \rho_{\sigma,r}\omega\eta\sqrt{\sigma_t}\sqrt{v_t}^p & \rho_{\sigma,v}\omega\gamma\sqrt{\sigma_t}\sqrt{v_t}^p & \omega\psi_t^\sigma\sqrt{\sigma_t} \\ \hline * & * & \eta^2 v_t^p & \rho_{r,v}\eta\gamma v_t^p & \rho_{r,\sigma}\eta\psi_t^\sigma\sqrt{v_t}^p \\ * & * & * & \gamma^2 v_t^p & \rho_{\sigma,v}\gamma\psi_t^\sigma\sqrt{v_t}^p \\ * & * & * & * & (\psi_t^\sigma)^2 \end{array}\right)_{5\times 5}. \tag{4.8}$$

We can also find an instantaneous covariance matrix for $\widetilde{\mathbf{H}}_t^{p=1}$:

$$\boldsymbol{\Sigma}_{\widetilde{\mathbf{H}}_t^1}\boldsymbol{\Sigma}_{\widetilde{\mathbf{H}}_t^1}^{\mathrm{T}} = \left(\begin{array}{cc} \boldsymbol{\Sigma}_{\mathbf{J}}\boldsymbol{\Sigma}_{\mathbf{J}}^{\mathrm{T}}(1) & \mathbf{D}_t \\ \mathbf{D}_t^{\mathrm{T}} & \mathbf{A}_t \end{array}\right)_{7\times 7}, \tag{4.9}$$

with

$$\mathbf{D}_t = \left(\begin{array}{cc} \rho_{x,v}\psi_t^v\sqrt{\sigma_t} & \rho_{x,\sigma}\psi_t^\sigma\sqrt{\sigma_t v_t} + \rho_{x,v}\psi_t^v\sigma_t \\ \rho_{\sigma,v}\omega\psi_t^v\sqrt{\sigma_t} & \omega\psi_t^\sigma\sqrt{\sigma_t v_t} + \rho_{\sigma,v}\omega\psi_t^v\sigma_t \\ \rho_{r,v}\eta\psi_t^v\sqrt{v_t}^p & \rho_{r,\sigma}\eta\psi_t^\sigma v_t + \rho_{r,v}\eta\psi_t^v\sqrt{v_t\sigma_t} \\ \gamma\psi_t^v\sqrt{v_t}^p & \rho_{\sigma,v}\gamma\psi_t^\sigma v_t + \gamma\psi_t^v\sqrt{v_t\sigma_t} \\ \rho_{\sigma,v}\psi_t^v\psi_t^\sigma & (\psi_t^\sigma)^2\sqrt{v_t} + \rho_{\sigma,v}\psi_t^\sigma\psi_t^v\sqrt{\sigma_t} \end{array}\right)_{5\times 2}. \tag{4.10}$$

and

$$\mathbf{A}_t = \left(\begin{array}{cc} (\psi_t^v)^2 & \rho_{\sigma,v}\psi_t^v\psi_t^\sigma\sqrt{v_t} + (\psi_t^v)^2\sqrt{\sigma_t} \\ * & v_t(\psi_t^\sigma)^2 + \sigma_t(\psi_t^v)^2 + 2\rho_{\sigma,v}\psi_t^\sigma\psi_t^v\sqrt{\sigma_t}\sqrt{v_t} \end{array}\right)_{2\times 2}. \tag{4.11}$$

Both matrices, (4.8) and (4.9), are not in the affine form, but with the *affinity correction* variables u_t^σ, u_t^v and z_t given in (4.7), we can write:

$$\boldsymbol{\Sigma}_{\widetilde{\mathbf{H}}_t^0}\boldsymbol{\Sigma}_{\widetilde{\mathbf{H}}_t^0}^{\mathrm{T}} = \left(\begin{array}{cc|cc|c} \sigma_t & \rho_{x,\sigma}\omega\sigma_t & \rho_{x,r}\eta u_t^\sigma & \rho_{x,v}\gamma u_t^\sigma & \rho_{x,\sigma}\psi_t^\sigma u_t^\sigma \\ * & \omega^2\sigma_t & \rho_{\sigma,r}\omega\eta u_t^\sigma & \rho_{\sigma,v}\omega\gamma u_t^\sigma & \omega\psi_t^\sigma u_t^\sigma \\ \hline * & * & \eta^2 & \rho_{r,v}\eta\gamma & \rho_{r,\sigma}\eta\psi_t^\sigma \\ * & * & * & \gamma^2 & \rho_{\sigma,v}\gamma\psi_t^\sigma \\ * & * & * & * & (\psi_t^\sigma)^2 \end{array}\right)_{5\times 5}, \tag{4.12}$$

and for $p = 1$ we obtain:

$$\boldsymbol{\Sigma}_{\mathbf{J}}\boldsymbol{\Sigma}_{\mathbf{J}}^{\mathrm{T}}(1) = \left(\begin{array}{cc|cc|c} \sigma_t & \rho_{x,\sigma}\omega\sigma_t & \rho_{x,r}\eta z_t & \rho_{x,v}\gamma z_t & \rho_{x,\sigma}\psi_t^\sigma u_t^\sigma \\ * & \omega^2\sigma_t & \rho_{\sigma,r}\omega\eta z_t & \rho_{\sigma,v}\omega\gamma z_t & \omega\psi_t^\sigma u_t^\sigma \\ \hline * & * & \eta^2 v_t & \rho_{r,v}\eta\gamma v_t & \rho_{r,\sigma}\eta\psi_t^\sigma u_t^v \\ * & * & * & \gamma^2 v_t & \rho_{\sigma,v}\gamma\psi_t^\sigma u_t^v \\ * & * & * & * & (\psi_t^\sigma)^2 \end{array}\right)_{5\times 5}, \tag{4.13}$$

with

$$\mathbf{D}_t = \left(\begin{array}{cc} \rho_{x,v}\psi_t^v u_t^\sigma & \rho_{x,\sigma}\psi_t^\sigma z_t + \rho_{x,v}\psi_t^v\sigma_t \\ \rho_{\sigma,v}\omega\psi_t^v u_t^\sigma & \omega\psi_t^\sigma z_t + \rho_{\sigma,v}\omega\psi_t^v\sigma_t \\ \rho_{r,v}\eta\psi_t^v u_t^v & \rho_{r,\sigma}\eta\psi_t^\sigma v_t + \rho_{r,v}\eta\psi_t^v z_t \\ \gamma\psi_t^v u_t^v & \rho_{\sigma,v}\gamma\psi_t^\sigma v_t + \gamma\psi_t^v z_t \\ \rho_{\sigma,v}\psi_t^v\psi_t^\sigma & (\psi_t^\sigma)^2 u_t^v + \rho_{\sigma,v}\psi_t^\sigma\psi_t^v u_t^\sigma \end{array}\right)_{5\times 2}, \tag{4.14}$$

and

$$\mathbf{A}_t = \left(\begin{array}{cc} (\psi_t^v)^2 & \rho_{\sigma,v}\psi_t^v\psi_t^\sigma u_t^v + (\psi_t^v)^2 u_t^\sigma \\ * & v_t(\psi_t^\sigma)^2 + \sigma_t(\psi_t^v)^2 + 2\rho_{\sigma,v}\psi_t^\sigma\psi_t^v z_t \end{array}\right)_{2\times 2}. \tag{4.15}$$

Since the functions μ_t^σ, μ_t^v, ψ_t^σ and ψ_t^v are given deterministic functions of time, the drifts (4.6) are of a linear form, and the covariance matrices $\mathbf{\Sigma}_{\widetilde{\mathbf{H}}_t^0}\mathbf{\Sigma}_{\widetilde{\mathbf{H}}_t^0}^{\mathrm{T}}$ and $\mathbf{\Sigma}_{\widetilde{\mathbf{H}}_t^1}\mathbf{\Sigma}_{\widetilde{\mathbf{H}}_t^1}^{\mathrm{T}}$ do not consist of any non-affine functions of the state variables. So, our modified multi-factor hybrid model (4.6) has an affine structure. Therefore, we are able to derive the corresponding characteristic functions which form the basis for an approximation of the solution of the pricing PDE given by (2.8).

5 Approximation of the Characteristic Function

As shown in [Duffie-2006] the relevance of the Fourier transform for PDEs lies in the transformation to a problem which can be solved efficiently. For this one needs to derive the characteristic function (CF). As already described in [Lee-2004] once the CF is found, a wide class of options can be efficiently and accurately priced, with the help of fast numerical routines.

The focus of this section is on the derivation of the CF for the affine versions of the two hybrid models introduced previously, H-G2++ and H-H2++.

5.1 Characteristic Function for the Heston-G2++ Hybrid Model

Model (4.6), with $p = 0$, is the Heston-Hull-White two-factor model:

$$\begin{cases} \mathrm{d}x_t = & (r_t - 1/2\sigma_t)\,\mathrm{d}t + & \sqrt{\sigma_t}\mathrm{d}W_t^x, & x_0 > 0, \\ \mathrm{d}\sigma_t = & \epsilon(\bar{\sigma} - \sigma_t)\mathrm{d}t + & \omega\sqrt{\sigma_t}\mathrm{d}W_t^\sigma, & \sigma_0 > 0, \\ \mathrm{d}r_t = & \kappa(\theta_t + v_t - r_t)\mathrm{d}t + & \eta\mathrm{d}W_t^r, & r_0 > 0, \\ \mathrm{d}v_t = & -\lambda v_t\mathrm{d}t + & \gamma\mathrm{d}W_t^v, & v_0 = 0, \end{cases} \tag{5.1}$$

with the additional state variable for the affinity of the system:

$$(\mathrm{d}\sqrt{\sigma_t} \approx)\ \mathrm{d}u_t^\sigma = \quad \mu_t^\sigma\mathrm{d}t + \quad \psi_t^\sigma\mathrm{d}W_t^\sigma, \quad u_0 = \sqrt{\sigma_0} > 0, \tag{5.2}$$

and all the processes correlated as given in (2.7).

By using the standard technique we obtain the pricing PDE for the approximate H-G2++ model:

$$\begin{aligned} 0 \quad = \quad & (r - 1/2\sigma)\,\mathbf{V}_x + \epsilon(\bar{\sigma} - \sigma)\mathbf{V}_\sigma + \kappa\,(\theta(t) - r + v)\,\mathbf{V}_r - \lambda v\mathbf{V}_v + \mu^\sigma(t)\mathbf{V}_u + \frac{1}{2}\sigma\mathbf{V}_{xx} \\ & + \frac{1}{2}\omega^2\sigma\mathbf{V}_{\sigma\sigma} + \frac{1}{2}\eta^2\mathbf{V}_{rr} + \frac{1}{2}\gamma^2\mathbf{V}_{vv} + \frac{1}{2}(\psi^\sigma(t))^2\mathbf{V}_{uu} + \rho_{x,\sigma}\omega\sigma\mathbf{V}_{x,\sigma} + \rho_{x,r}\eta u\mathbf{V}_{x,r} \\ & + \rho_{x,v}\gamma u\mathbf{V}_{x,v} + \rho_{x,\sigma}\psi^\sigma(t)u\mathbf{V}_{x,u} + \rho_{\sigma,r}\eta\omega u\mathbf{V}_{\sigma,r} + \rho_{\sigma,v}\gamma\omega u\mathbf{V}_{\sigma,v} + \psi^\sigma(t)\omega u\mathbf{V}_{\sigma,u} \\ & + \rho_{r,v}\gamma\eta\mathbf{V}_{r,v} + \rho_{r,\sigma}\eta\psi^\sigma(t)\mathbf{V}_{r,u} + \rho_{\sigma,v}\gamma\psi^\sigma(t)\mathbf{V}_{v,u} - r\mathbf{V} + \mathbf{V}_t, \end{aligned}$$

with state variables $x := x_t$, $\sigma := \sigma_t$, $r := r_t$, $v := v_t$ and $u := u_t^\sigma$ and time-dependent functions $\theta(t)$, $\mu^\sigma(t)$ and $\psi^\sigma(t)$. The PDE given above is linear in its state variables. The discounted characteristic function for the state vector, $\widetilde{\mathbf{H}}_t^0 = [x_t, \sigma_t, r_t, v_t, u_t^\sigma]^{\mathrm{T}}$, is known, under the risk neutral measure, as:

$$\phi_{\text{H-G2++}}(\mathbf{u}, \widetilde{\mathbf{H}}_t^0, \tau) = \mathbb{E}^{\mathbb{Q}}\left(\mathrm{e}^{-\int_t^T r_s\mathrm{d}s}\mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\widetilde{\mathbf{H}}_T^0}|\mathcal{F}_0\right) = \mathrm{e}^{A(\mathbf{u},\tau)+\mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)\widetilde{\mathbf{H}}_t^0}, \tag{5.3}$$

with the initial condition:

$$\phi_{\text{H-G2++}}(\mathbf{u}, \widetilde{\mathbf{H}}_T^0, 0) = \mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\widetilde{\mathbf{H}}_T^0}, \tag{5.4}$$

with $\tau = T - t$. Here, $A(\mathbf{u}, \tau)$ and $\mathbf{B}(\mathbf{u}, \tau)$ satisfy the following Riccati ordinary differential equations (see [Duffie, *et al.*-2000]):

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{B}(\mathbf{u}, \tau) = & -r_1 + a_1^T\mathbf{B} + & \dfrac{1}{2}\mathbf{B}^T c_1\mathbf{B}, \\ \dfrac{\mathrm{d}}{\mathrm{d}\tau}A(\mathbf{u}, \tau) = & -r_0 + \mathbf{B}^T a_0 + & \dfrac{1}{2}\mathbf{B}^T c_0\mathbf{B}, \end{cases} \tag{5.5}$$

with $a_i, c_i, r_i,\ i = 0, 1$, given by a linear decomposition:

$$\begin{aligned}
\mu(\widetilde{\mathbf{H}}_t^0) &= a_0 + a_1 \widetilde{\mathbf{H}}_t^0, \text{ for any } (a_0, a_1) \in \mathbb{R}^l \times \mathbb{R}^{l\times l}, && (5.6)\\
\sigma(\widetilde{\mathbf{H}}_t^0)\sigma(\widetilde{\mathbf{H}}_t^0)^{\mathrm{T}} &= (c_0)_{ij} + (c_1)_{ij}^{\mathrm{T}} \widetilde{\mathbf{H}}_t^0, \text{ for arbitrary } (c_0, c_1) \in \mathbb{R}^{l\times l} \times \mathbb{R}^{l\times l\times l}, && (5.7)\\
r(\widetilde{\mathbf{H}}_t^0) &= r_0 + r_1^{\mathrm{T}} \widetilde{\mathbf{H}}_t^0, \text{ for } (r_0, r_1) \in \mathbb{R} \times \mathbb{R}^l, && (5.8)
\end{aligned}$$

where l indicates the dimension of the state vector $\widetilde{\mathbf{H}}_t^0$. Now, we set $\mathbf{u} = [u, 0, 0, 0, 0]^{\mathrm{T}}$ and finite maturity $\tau > 0$. The following two lemmas provide us with the coefficients needed for evaluating the characteristic function.

Lemma 5.1 (CF coefficients for the Heston-G2++ hybrid model)**.** *The coefficients of the characteristic function for the H-G++ model in (5.1) are the solutions of the following Riccati ODEs:*

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau} B_x(u,\tau) &= 0,\\
\frac{\mathrm{d}}{\mathrm{d}\tau} B_\sigma(u,\tau) &= \frac{1}{2}\left(B_x - 1\right) B_x + \left(\rho_{x,\sigma}\omega B_x - \epsilon\right) B_\sigma + \frac{1}{2}\omega^2 B_\sigma^2,\\
\frac{\mathrm{d}}{\mathrm{d}\tau} B_r(u,\tau) &= -1 - \kappa B_r + B_x,\\
\frac{\mathrm{d}}{\mathrm{d}\tau} B_v(u,\tau) &= \kappa B_r - \lambda B_v,\\
\frac{\mathrm{d}}{\mathrm{d}\tau} B_u(u,\tau) &= \rho_{x,r}\eta B_x B_r + \rho_{\sigma,r}\omega\eta B_\sigma B_r + \rho_{x,v}\gamma B_x B_v + \rho_{\sigma,v}\omega\gamma B_\sigma B_v\\
&\quad + \rho_{x,\sigma}\psi_\tau^\sigma B_x B_u + \omega\psi_\tau^\sigma B_\sigma B_u,\\
\frac{\mathrm{d}}{\mathrm{d}\tau} A(u,\tau) &= \epsilon\bar{\sigma} B_\sigma + \kappa\theta_\tau B_r + \mu_\tau^\sigma B_u + \frac{1}{2}\eta^2 B_r^2 + \frac{1}{2}\gamma^2 B_v^2 + \frac{1}{2}(\psi_t^\sigma)^2 B_u^2\\
&\quad + \rho_{r,v}\eta\gamma B_r B_v + \rho_{r,\sigma}\eta\psi_\tau^\sigma B_r B_u + \rho_{\sigma,v}\gamma\psi_\tau^\sigma B_v B_u.
\end{aligned}$$

Lemma 5.2 provides the analytic solutions to the complex-valued ODEs $\mathbf{B}(u,\tau)$ for the variables x_t, σ_t, r_t, and v_t.

Lemma 5.2 (Solutions to the CF coefficients of the H-G2++)**.** *The solutions to $B_x(u,\tau)$, $B_\sigma(u,\tau)$, $B_r(u,\tau)$ and $B_v(u,\tau)$, defined in Lemma 5.1, are given by:*

$$\begin{aligned}
B_x(u,\tau) &= iu, && (5.9)\\
B_\sigma(u,\tau) &= \frac{1 - \mathrm{e}^{-D\tau}}{\omega^2\left(1 - g\mathrm{e}^{-D\tau}\right)}\left(\epsilon - \rho_{x,\sigma}\omega iu - D\right), && (5.10)\\
B_r(u,\tau) &= (iu - 1)\kappa^{-1}\left(1 - \mathrm{e}^{-\kappa\tau}\right), && (5.11)\\
B_v(u,\tau) &= \frac{i\mathrm{e}^{-\kappa\tau}(i+u)}{(\kappa-\lambda)\lambda}\left(\mathrm{e}^{\kappa\tau}(\kappa-\lambda) + \lambda - \kappa\mathrm{e}^{(\kappa-\lambda)\tau}\right), && (5.12)
\end{aligned}$$

with $D = \sqrt{\left(\rho_{x,\sigma}\omega iu - \epsilon\right)^2 + \omega^2\left(u^2 + iu\right)}$ *and* $g = \dfrac{-\rho_{x,\sigma}\omega iu + \epsilon - D}{-\rho_{x,\sigma}\omega iu + \epsilon + D}$.
Since $\kappa = \lambda$, the expression for $B_v(u,\tau)$ diverges, and we need to use the following limit function:

$$\lim_{\kappa\to\lambda} B_v(u,\tau) = \frac{1}{\lambda} i\mathrm{e}^{-\lambda\tau}(i+u)\left(\mathrm{e}^{\lambda\tau} - \lambda\tau - 1\right). \tag{5.13}$$

The remaining two functions, $B_u(u,\tau)$ and $A(u,\tau)$, involve the time-dependent ψ_t^σ and μ_t^σ from Lemma 3.4 and therefore need to be solved numerically.

5.2 Characteristic Function for the H-H2++ Hybrid Model

More involved than H-G2++ is the H-H2++ model from (4.6), with $p = 1$. As we have shown before, the model is of the affine form if three additional state variables are included in the system,

leading to the 7-dimensional system $\widetilde{\mathbf{H}}_t^1 = [x_t, \sigma_t, r_t, v_t, u_t^\sigma, u_t^v, z_t]^{\mathrm{T}}$. Although the model is of high-dimensionality, we show that one can still find the corresponding CF with numerical techniques. Following the notation from Section 4 the following form for the discounted CF is found:

$$\phi_{\text{H-H2++}}(\mathbf{u}, \widetilde{\mathbf{H}}_t^1, \tau) = \mathbb{E}^{\mathbb{Q}}\left(\mathrm{e}^{-\int_t^T r_s \mathrm{d}s} \mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\widetilde{\mathbf{H}}_T^1} | \mathcal{F}_0\right) = \mathrm{e}^{A(\mathbf{u},\tau)+\mathbf{B}^{\mathrm{T}}(\mathbf{u},\tau)\widetilde{\mathbf{H}}_t^1}, \tag{5.14}$$

with the initial condition:

$$\phi_{\text{H-H2++}}(\mathbf{u}, \widetilde{\mathbf{H}}_T^1, 0) = \mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\widetilde{\mathbf{H}}_T^1}, \tag{5.15}$$

The governing ODEs for the functions of the CF are presented in the following lemma.

Lemma 5.3 (ODEs for the approximate Heston-H2++ hybrid model)**.** *The functions* $B_x(u,\tau)$, $B_r(u,\tau)$,$B_\sigma(u,\tau)$,$B_v(u,\tau)$, $B_{u^\sigma}(u,\tau)$, $B_{u^v}(u,\tau)$, $B_z(u,\tau)$ *and* $A(u,\tau)$ *for* $u \in \mathbb{R}$ *satisfy the following system of ODEs:*

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}B_\sigma(u,\tau) &= B_x\left(\frac{1}{2}\left(B_x-1\right)+\psi_\tau^v\rho_{x,v}B_z\right)+\omega B_\sigma\left(\rho_{x,\sigma}B_x-\frac{\epsilon}{\omega}+\frac{1}{2}\omega B_\sigma+\psi_\tau^v\rho_{\sigma,v}B_z\right),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}B_v(u,\tau) &= \gamma B_v\left(-\frac{\lambda}{\gamma}+\eta\rho_{r,v}B_r+\frac{1}{2}\gamma B_v+\psi_\tau^\sigma\rho_{\sigma,v}B_z\right)+\eta B_r\left(\frac{1}{2}\eta B_r+\psi_\tau^\sigma\rho_{r,\sigma}B_z\right),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}B_{u^\sigma}(u,\tau) &= B_z\left(\frac{1}{2}(\psi_t^v)^2B_z+\mu_\tau^\sigma+\psi_\tau^\sigma B_{u^\sigma}\left(\rho_{x,\sigma}+\psi_\tau^v\rho_{\sigma,v}\right)+(\psi_\tau^v)^2B_v\right)+\psi_\tau^v\rho_{x,v}B_xB_{u^v}\\
&\quad+\omega B_\sigma\left(\psi_\tau^\sigma B_{u^\sigma}+\psi_\tau^v\rho_{\sigma,v}B_{u^v}\right),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}B_{u^v}(u,\tau) &= B_z\left(\mu_\tau^\sigma+(\psi_\tau^\sigma)^2B_{u^\sigma}+\psi_\tau^\sigma\psi_\tau^v\rho_{\sigma,v}B_v+\frac{1}{2}(\psi_\tau^\sigma)^2B_z\right)+\psi_\tau^vB_{u^v}\left(\eta\rho_{r,v}B_r+\gamma B_v\right)\\
&\quad+\psi_\tau^\sigma B_{u^\sigma}\left(\eta\rho_{r,\sigma}B_r+\gamma\rho_{\sigma,v}B_v\right),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}B_z(u,\tau) &= B_z\big(\psi_\tau^\sigma\left(\rho_{x,\sigma}B_x+\omega B_\sigma\right)+\psi_\tau^v\big(\eta\rho_{r,v}B_r+\gamma B_v+\psi_\tau^\sigma\rho_{\sigma,v}B_z\big)\big)+\eta\rho_{x,r}B_rB_x\\
&\quad+\omega\eta\rho_{\sigma,r}B_rB_\sigma+\gamma B_v\left(\rho_{x,v}B_x+\omega\rho_{\sigma,v}B_\sigma\right),\\
\frac{\mathrm{d}}{\mathrm{d}\tau}A(u,\tau) &= \epsilon\bar{\sigma}B_\sigma+\kappa\theta_\tau B_r+\lambda\bar{v}B_v+B_{u^\sigma}\left(\mu_\tau^\sigma+\frac{1}{2}(\psi_\tau^\sigma)^2B_{u^\sigma}\right)+\psi_\tau^\sigma\psi_\tau^v\rho_{\sigma,v}B_z\\
&\quad+B_{u^v}\left(\mu_t^v+\frac{1}{2}(\psi_\tau^v)^2B_{u^v}+\psi_\tau^\sigma\psi_\tau^v\rho_{\sigma,v}B_{v^\sigma}\right),
\end{aligned}$$

where $B_x(u,\tau) = iu$, *and* $B_r(u,\tau)$ *is as in the H-G2++ model (see Equation (5.11)).*

The ODEs appearing in Lemma 5.3 cannot easily be solved analytically, except for $B_x(u,\tau)$ and $B_r(u,\tau)$.

In order to calculate $B_i(u,\tau)$ and $A(u,\tau)$ for a specific u and given τ, these Riccati ODEs need to be solved numerically. Crucial in determining the CF in (5.14) is a fast computation of these coefficients. Commonly available explicit Runge-Kutta methods, without any additional features, are highly efficient for the simultaneous computation of all $B_i(u,\tau)$, and $A(u,\tau)$ on a given numerical grid for u. We employ the standard Matlab routine *ode45* for this purpose.

The same routine is used to determine the remaining two coefficients, $A(u,\tau)$ and $B_u(u,\tau)$, from Lemma 5.1.

5.3 Numerical Method

For obtaining the European option prices at the highest efficiency, we use the COS pricing method from [Fang,Oosterlee-2008], which is based on the availability of the characteristic function. The method employs a Fourier cosine expansion of the density function.

From the general risk-neutral pricing formula the price of any claim $V(T,S_T)$ defined in terms of the underlying stock process S_T can be written as:

$$\Pi(t,S_t) = \mathbb{E}^{\mathbb{Q}}\left(\mathrm{e}^{-\int_t^T r_s\mathrm{d}s}V(T,S_T)|\mathcal{F}_t\right) = \int_{\mathbb{R}} V(T,y)\widehat{f}_Y(y|x)\mathrm{d}y, \tag{5.16}$$

where $\widehat{f}_Y(y|x) = \int_{\mathbb{R}} \mathrm{e}^z f_{Y,Z}(y,z|x)\mathrm{d}z$, with $z = -\int_t^T r_s \mathrm{d}s$. Note that we use the joint distribution $f_{Y,Z}(\cdot)$ since S_t and r_t are correlated.

Assuming fast decay of the density function, we use the following approximation:

$$\Pi(t,x) \approx \int_{\delta_1}^{\delta_2} V(T,y)\widehat{f}_Y(y|x)\mathrm{d}y, \tag{5.17}$$

choosing $x = \log(S_t/K)$, $y = \log(S_T/K)$, and $\delta_1 < \delta_2$. Now, in order to recover the density $\widehat{f}_Y(y|x)$ one can use the following Fourier-cosine expansion, based on the availability of the characteristic function:

$$\widehat{f}_Y(y|x) \approx \sum_{n=0}^{N} \frac{2\omega_n}{\delta_2-\delta_1} \Re\left(\phi\left(\frac{n\pi}{\delta_2-\delta_1}\right)\exp\left(-\frac{n\pi i\delta_1}{\delta_2-\delta_1}\right)\right)\cos\left(n\pi\frac{y-\delta_1}{\delta_2-\delta_1}\right), \tag{5.18}$$

with $\omega_0 = 1/2$ and $\omega_n = 1$, $n \in \mathbb{N}^+$. Using this density expansion, we can replace the probability function $\widehat{f}_Y(y|x)$ in Equation (5.16), i.e.:

$$\Pi(t,x) \approx \sum_{n=0}^{N} \omega_n \Re\left(\phi\left(\frac{n\pi}{\delta_2-\delta_1}\right)\exp\left(-\frac{n\pi i\delta_1}{\delta_2-\delta_1}\right)\right)\Gamma_n^{\delta_1,\delta_2}, \tag{5.19}$$

where the coefficients $\Gamma_n^{\delta_1,\delta_2}$ are known analytically for European options, see [Fang,Oosterlee-2008] for details.

The expansion in (5.19) exhibits an exponential convergence in the number of terms. Typically small values of N, the number of terms, are needed. Moreover, a whole vector of strikes can be priced simultaneously. A proper choice for the range of integration in (5.1) is a guarantee for fast convergence with only a few terms in the Fourier-cosine expansion. In [Fang,Oosterlee-2008], the integration range was based on some insight in the behavior of the probability density function. There, one choice was to set $\delta_1 = -L\sqrt{\tau}$ and $\delta_2 = L\sqrt{\tau}$, with $L = 8$. We also use this integration range here.

6 Numerical Experiments

Some numerical experiments are presented in this section. We compare the plain vanilla option prices obtained by the approximate affine hybrid models (computed by the COS method), with numerical results from the full-scale two-factor hybrid models, computed by the Monte Carlo approach.

In the table below we first present the time needed for pricing the European option with the two affine hybrid models, H-G2++ and H-H2++. In both cases we price a whole strip of 50 strikes $K = \{0.1, 0.2, \ldots, 5\}$. The table shows timing results and sum-squared-errors (SSE); the total time used consists of the time needed for the integration of the Riccati ODEs, as well as the time used by the COS method for pricing the options. For both models the reference prices are calculated with the same method, using a large number (2^{12}) of terms in the expansion.

The tolerance for the explicit Runge-Kutta method for the Riccati ODEs was set to 10^{-10}. The models were evaluated with the following parameters:

$$\epsilon = 1.2, \bar{\sigma} = 0.1, \omega = 0.05, \kappa = 1.5, \theta = 0.05, \eta = 0.1, \lambda = 1.5, \gamma = 0.1,$$

and a full matrix of correlations:

$$\rho_{x,\sigma} = -0.4, \rho_{x,r} = 0.4, \rho_{x,v} = -0.6, \rho_{\sigma,r} = 0.1, \rho_{\sigma,v} = 0.2, \rho_{r,v} = 0.3,$$

and initial values:

$$S_0 = 1, \sigma_0 = 0.05, r_0 = 0.03,$$

For H-G2++ we set $v_0 = 0$ and for H-H2++ $v_0 = 0.05$.

Highly satisfactory results are found in Table 6.1, obtained on a standard PC[1]. Hundred terms in the expansion are sufficient regarding the accuracy.

[1] Intel(R)Core(TM)2CPU, 6400 @ 2.13GHz, 1GB of RAM.

Table 6.1: Total CPU time and sum-squared-error (SSE) of European calls for a whole set of strikes.

characteristic			number of expansion terms (N)			
model	**maturity**	**des.**	40	50	100	200
H-G2++	$\tau = 1y$	SSE	34.8956	1.06618	5.756E-11	1.011E-16
		time [s]	0.0619s	0.0734s	0.0843s	0.0931s
	$\tau = 10y$	SSE	4.856	2.352E-5	1.357E-11	1.263E-11
		time [s]	0.0925s	0.0950s	0.1051s	0.1152s
H-H2++	$\tau = 1$	SSE	4.505E2	16.671	2.746E-8	1.045E-17
		time [s]	0.402s	0.450s	0.512s	0.612s
	$\tau = 10y$	SSE	1.02E2	2.231E-3	8.479E-16	1.453E-16
		time [s]	0.723s	0.817s	0.932s	1.023s

In a next experiment we check the accuracy of the approximate hybrid model. The experiment is as follows. First of all, we generate European call prices with the original hybrid model by Monte Carlo simulation. Secondly, we compare, in terms of implied volatilities, with the same results from the approximate affine hybrid models obtained by the COS method. We consider two cases, one where the parameters appearing satisfy the Feller condition for the stock and a second experiment where they do not satisfy this condition. For both tests we have performed the Monte Carlo simulation with the scheme proposed in [Andersen-2006].

Experiment 6.1 (Feller's condition satisfied i.e.: $2\epsilon\bar{\sigma} > \omega^2$)**.** *We have performed numerical experiments with both models, i.e., H-G2++ and H-H2++. However, as we have obtained almost identical results, we only present the results obtained with H-G2++. The parameters are chosen as:*

$$\epsilon = 1.2, \bar{\sigma} = 0.1, \omega = 0.05, \kappa = 1.5, \theta = 0.05, \eta = 0.1, \lambda = 1.5, \gamma = 0.1,$$

$$\rho_{x,\sigma} = -0.4, \rho_{x,r} = 0.4, \rho_{x,v} = -0.6, \rho_{\sigma,r} = 0.1, \rho_{\sigma,v} = 0.2, \rho_{r,v} = 0.3.$$

The initial conditions set are: $S_0 = 1$, $r_0 = 0.03$, $v_0 = 0$. *With the specified parameters the Feller's condition for the stock*[2] *is satisfied, i.e.:* $\epsilon\bar{\sigma} > \omega^2$*:* $0.24 > 0.0025$. *In the experiment we choose three maturities* $\tau = 1$, $\tau = 5$ *and* $\tau = 10$. *For the model H-G2++ Figure 6.1(a) shows an almost perfect correspondence between the volatilities from the Monte Carlo method (for the full-scale hybrid model) and the COS method (for the approximate affine hybrid model).*

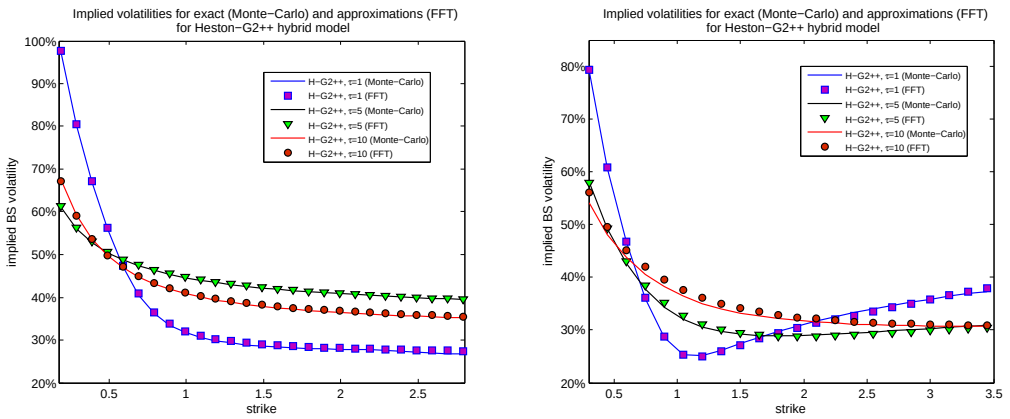

Figure 6.1: Implied volatilities for the equity, for different maturities. The full-scale hybrid model (Monte Carlo, with 50.000 paths and 1000 time steps) vs the approximate affine hybrid model (Fourier-cosine expansion method) (a) Experiment 6.1, (b) Experiment 6.2.

[2]We discuss here the Feller's condition for the stock, however, one may also check the Feller condition for the short-rate process in the H-H2++ model

Experiment 6.2 (Feller's condition not satisfied i.e.: $2\epsilon\bar{\sigma} < \omega^2$)**.** *In practice there are many cases in which the Feller condition is not satisfied. Therefore we check the performance of the approximate affine hybrid model in such a setup. In this experiment we therefore choose* $\epsilon = 0.8$, $\bar{\sigma} = 0.05$ *and* $\omega = 0.5$ *and remaining parameters as in Experiment 6.1. Clearly, the Feller condition does not hold in this case, as* $0.08 \not\geq 0.25$*. Therefore, the probability of hitting zero is positive. Figure 6.1(b) shows however that our approximate hybrid model also performs very satisfactory in this test. In the experiment the error increases with increasing maturity. This is due to the fact that the noncentrality parameter,* $\lambda(t)$*, converges to zero for large maturities and because the degrees of freedom parameter, d, is small. The results are in accordance with the discussion in Remark 2.*

Experiment 6.3 (Impact of correlation, $\rho_{x,r}$, on implied volatility)**.** *In this experiment we investigate the effect of the correlation,* $\rho_{x,r}$*, on the implied ATM volatilities. In Figure 6.2 the results obtained are presented. From the graph we see that the correlations have significant impact on the volatilities. For both models, H-G2++ and H-H2++, we noticed the following relation: the higher the correlation the higher is the volatility level. This effect is stronger for the H-G2++ model.*

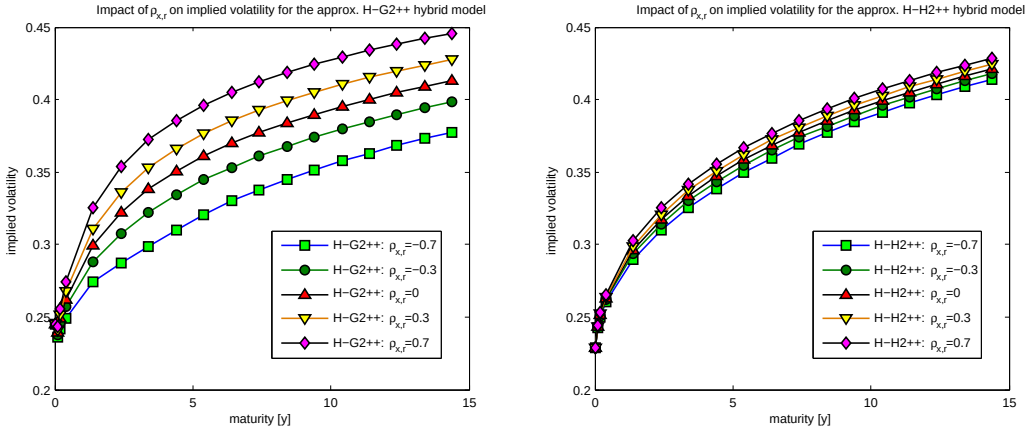

Figure 6.2: The impact of the correlation, $\rho_{x,r}$, between the log-stock x_t and short rate r_t, on the equity implied volatility. The model parameters were set as: $\epsilon = 1.2$, $\bar{\sigma} = 0.05$, $\omega = 0.1$, $\kappa = 1.2$, $\theta = 0.05$, $\eta = 0.1$, $\lambda = 1.2$, $\gamma = 0.1$, $\rho_{x,\sigma} = -0.5$ and the remaining correlations equal to 0.

7 Conclusions and Final Remarks

In this article we have constructed an approximation for the characteristic function of the multi-factor equity interest hybrid models, Heston-Gaussian two-factor (H-G2++) and Heston-stochastic volatility short rate (H-H2++). The governing stochastic processes are placed in the class of affine diffusion processes. Our approximations allow us to define a full matrix of correlations between the processes. By a straightforward implementation we obtain the plain vanilla prices (for a whole strip of strikes) in less than a second. The numerical experiments confirm a highly accurate approximation of the prices of plain vanilla products compared to the full hybrid model. These hybrid models can therefore be used for calibration purposes.

The approximation of the hybrid model is based on replacing the "square-root of square-root" process by a normal distribution. The experiments show a very satisfactory approximation even if the Feller condition, which is often ignored in practice, does not hold.

References

[Abramowitz-1972] M. ABRAMOWITZ, I. A. STEGUN, Modified Bassel Functions $\mathbf{I}$ and $\mathbf{K}$, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover,* 374-377, 1972.

[Andersen-2006] L. ANDERSEN, Efficient Simulation of the Heston Stochastic Volatility Model, SSRN working paper, 2009. Available at SSRN: `http://ssrn.com/abstract=946405`.

[Black,Scholes-1973] F. BLACK, M. SCHOLES, The Pricing of Options and Corporate Liabilities. *J. Political Economy,* 81: 637–654, 1973.

[Brigo,Mercurio-2007] D. BRIGO, F. MERCURIO, *Interest Rate Models- Theory and Practice: With Smile, Inflation and Credit*. Springer Finance, second edition, 2007.

[Broadie,Kaya-2006] M. BROADIE, Ö. KAYA, Exact Simulation of Stochastic Volatility and other Affine Jump Diffusion Processes. *Operations Research,* 54: 217–231, 2006.

[Carr,Madan-1999] P. P. CARR, D. B. MADAN, Option Valuation Using the Fast Fourier Transform. *J. Comp. Finance,* 2:61-73, 1999.

[Cox, *et al.*-1985] J. C. COX, J. E. INGERSOLL, S. A. ROSS, A theory of the term structure of interest rates. *Econometrica* 53: 385-407, 1985.

[Dufresne-2001] P. COLLIN-DUFRESNE, R. S. GOLDSTEIN, Pricing Swaptions Within an Affine Framework. *J. Derivatives,* 10(1): 9–26, Fall 2002.

[Derman,Kani-1998] E. DERMAN, I. KANI, Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility. *Int. J. Theoretical Appl. Finance,* 1: 61–110, 1998.

[Duffie, *et al.*-2000] D. DUFFIE, J. PAN, K. SINGLETON, Transform analysis and asset pricing for affine jump-diffusions. *Econometrica,* 68: 1343–1376, 2000.

[Duffie-2006] D. J. DUFFIE, *F*inite Difference Methods in Financial Engineering. A partial Differential Equation Approach. John Wiley & Sons, Ltd, 2006.

[Dufresne-2001] D. DUFRESNE, The Integrated Square-Root Process. Working paper, University of Montreal, 2001.

[Dupire-2008] B. DUPIRE, Pricing with a Smile. *Risk,* 7: 18–20, 1994.

[Fang,Oosterlee-2008] F. FANG, C. W. OOSTERLEE, A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions. *SIAM J. Sci. Comput.*, 31: 826, 2008.

[Feller-1971] W. FELLER, An Introduction fo Probability Theory and Its Applications, Volume 2 (2nd Ed.) Wiley, Chicester.

[Fisher-1922] R. A. FISHER, On the Interpretation of χ^2 from Contingency Tables and Calculations of P. *J. R. Statist. Soc.,* 85: 87-94, 1922.

[Gatheral-2006] J. GATHERAL, *The Volatility Surface. A Practitioner's Guide.* John Wiley & Sons, Ltd, first edition, 2006.

[Geman-1995] H. GEMAN, N. EL KAROUI, J. C. ROCHET, Changes of Numéraire, Changes of Probability Measures and Pricing of Options. *J. Appl. Prob.* 32: 443–458, 1995.

[Gradshteyn,Ryzhik-1996] I. S. GRADSHTEYN, I. M. RYZHIK, *Table of Integrals, Series, and Products*, 5th ed., A. Jeffrey, Ed. Academic Press, San Diego, 1996.

[Grzelak, *et al.*,-2008] L. A. GRZELAK, C. W. OOSTERLEE, S.VAN WEEREN, Extension of Stochastic Volatility Equity models with Hull-White Interest Rate Process. *Quant. Finance, to appear.* Available at: `http://ssrn.com/abstract=1344959`.

[Grzelak,Oosterlee-2009] L. A.GRZELAK, C. W. OOSTERLEE, On the Heston model with stochastic interest rates. SSRN working paper, 2009. Available at SSRN: `http://ssrn.com/abstract=1382902`.

[Heston-1993] S. L. HESTON, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Rev. Finan. Stud.,* 2(6): 327–343, 1993.

[Heidari, *et al.*-2007] M. HEIDARI, A. HIRSA, D. B. MADAN, Pricing of Swaptions in Affine Term Structures with Stochastic Volatility. Advances in Mathematical Finance. Birkhäuser 2007.

[Hull-2006] J. HULL, Interest Rate Derivatives: Models of the Short Rate. *Option, Futures, and Other Derivatives,* 6: 657-658, 2006.

[Hull-2008] J. HULL, *Options, Futures, and Other Derivatives, Seventh Edition.* Prentice Hall, 2008.

[Hull,White-1996] J. HULL, A. WHITE, Using Hull-White Interest Rate Trees, *J. Derivatives,* 4: 26–36, 1996.

[Jamshidian-1989] F. JAMSHIDIAN, An Exact Bond Option Pricing Formula. *J. Finance,* 44: 204–209, 1989.

[Johnson, *et al.*-1994] N. L. JOHNSON, N. L. KOTZ, N. BALAKRISHNAN, *Continuous Univariate Distributions*, Volume 2, Second Edition, Wiley, New York, 1994.

[Kummer-1936] E. E. KUMMER, Über die hypergeometrische Reihe $F(a;b;x)$. *J. reine angew. Math.,* 15: 39–83, 1936.

[Koepf-1998] W. KOEPF, Hypergeometric Summation: An Algorithmic Approach to Summation and Special

Function Identities. Braunschweig, Germany: Vieweg, 1998.
[Lee-2004] R. LEE, Option Pricing by Transform Methods: Extensions, Unification, and Error Control. *J. Comp. Finance,*, 7(3): 51–86, 2004.
[Morton-2005] K. W. MORTON, D. F. MAYERS, *Numerical Solution of Partial Differential Equations, An Introduction.* Cambridge University Press, 2005.
[Moser-2007] S. M. MOSER, Some Expectations of a Non-Central Chi-Square Distribution with an Even Number of Degrees of Freedom, *TENCON 2007 - 2007 IEEE Region 10 Conference*, Oct. 30 2007-Nov. 2 2007.
[Øksendal-2000] B. ØKSENDAL, *Stochastic Differential Equations*, Fifth Ed., Springer Verlag, 2000.
[Patnaik-1949] P. B. PATNAIK, The Non-Central χ^2 and F-Distributions and Their Applications. *Biometrika,* 36: 202-232, 1949.
[Schöbel,Zhu-1999] R. SCHÖBEL, J. ZHU, Stochasic Volatility with an Ornstein-Uhlenbeck Process: An extension. *Europ. Fin. Review*, 3:23–46, 1999.
[Schrager,Pelsser-2002] D. F. SCHRAGER, A. J. PELSSER, Pricing Swaptions and Coupon Bonds Options in Affine Term Structure Models. *Math. Finance,* 16(4): 427–446, 2002.

A The Hybrid Model Calibration

In this appendix we outline a calibration procedure for the equity-interest rate hybrid models, and we give details on the calibration of the interest rate part. Since the calibration of multi-factor models in general is a very difficult task, we calibrate the hybrid model in the following three steps:

Algorithm 1 Hybrid model calibration procedure

1. Assume the interest rate asset classes to be independent of the equity asset class.
2. Calibrate the interest rate model to the market data available: Bonds, caplets, swaptions etc.
3. Take the equity-interest rate hybrid model (with previously calibrated interest rate parameters and unknown coefficients for the equity), and calibrate the remaining parameters. In the case of unknown correlations either:
3.1→ estimate them from historical data, or
3.2→ calibrate them with other hybrid parameters, or
3.3→ estimate some correlations from historical data, and find others determine via the calibration procedure.

The fundamental interest rate derivatives are bonds and options on the interest rate swaps (commonly called swaptions). These interest rate contracts correspond to stocks and plain vanilla options in the equity world. Typically, in the case of swaptions the underlying interest rate swap, can be seen as a contract for exchanging the coupon-bearing bond with a floating-rate note (for details see for example [Brigo,Mercurio-2007]). Since the swaptions are basic products in the interest rate world, it is crucial that new models can be efficiently calibrated to these products. In this section we discuss how efficient pricing of swaptions under affine processes can be done. Since for the G2++ model analytic formulas are available, we proceed with the H2++ model [Heidari, *et al.*-2007].

The class of affine processes is well-studied for many years (see, for example, [Dufresne-2001; Schrager,Pelsser-2002]). In this section we follow the results of Heidari [Heidari, *et al.*-2007] where the characteristic function for the H2++ short rate model was derived. We start with the prices for zero-coupon bonds, $P(t,T)$. We use the stochastic volatility short rate process H2++ of the following form:

$$\begin{cases} \mathrm{d}r_t = & \kappa(\theta_t - r_t)\mathrm{d}t + & \eta\sqrt{v_t}\mathrm{d}W_t^r, \\ \mathrm{d}v_t = & \lambda(\bar{v} - v_t)\mathrm{d}t + & \gamma\sqrt{v_t}\mathrm{d}W_t^v, \\ \mathrm{d}W_t^r\mathrm{d}W_t^v = & & \rho_{r,v}\mathrm{d}t, \end{cases} \tag{A.1}$$

where $\kappa > 0$, $\eta > 0$, $\lambda > 0$, $\bar{v} > 0$, $\gamma > 0$ and $|\rho_{r,v}| \leq 1$, as in (2.3).

Since we deal with a two-factor system, $\mathbf{R}_t = [r_t, v_t]^{\mathrm{T}}$, the corresponding characteristic function is also 2D, i.e., $\mathbf{u} = [u_r, u_v]^{\mathrm{T}}$. From the definition of a characteristic function we simply get:

$$\phi_{\text{H2++}}(\mathbf{u}, \mathbf{R}_t, \tau) \quad \stackrel{def}{=} \quad \mathbb{E}^{\mathbb{Q}}\left(\mathrm{e}^{-\int_t^T r_s \mathrm{d}s}\mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\mathbf{R}_T}|\mathcal{F}_t\right) = \mathrm{e}^{A(\mathbf{u},\tau)+B_r(\mathbf{u},\tau)r_t+B_v(\mathbf{u},\tau)v_t}, \tag{A.2}$$

where $\tau = T - t$, and r_t, v_t are the processes in (A.1). It is easy to find that $A(\mathbf{u}, \tau)$, $B_r(\mathbf{u}, \tau)$ and $B_v(\mathbf{u}, \tau)$ should satisfy the following system of ODEs:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}\tau} B_r(\mathbf{u}, \tau) = & -1 - \kappa B_r, \\ \frac{\mathrm{d}}{\mathrm{d}\tau} B_v(\mathbf{u}, \tau) = & \left(-\lambda + \frac{1}{2}\gamma^2 B_v + \eta\gamma\rho_{r,v} B_r\right) B_v + \frac{1}{2}\eta^2 B_r^2, \\ \frac{\mathrm{d}}{\mathrm{d}\tau} A(\mathbf{u}, \tau) = & \kappa\theta B_r + \lambda\bar{v} B_v, \end{cases} \tag{A.3}$$

with $\tau = T - t$ and the boundary condition $\mathbf{B}(\mathbf{u}, 0) = i\mathbf{u}$.

The second and third ODE in (A.3) are not easily solved analytically, but for the first equation we obtain:

$$B_r(\mathbf{u}, \tau) = \frac{1}{\kappa}\left(\mathrm{e}^{-\kappa\tau} + i\mathbf{u}\kappa\mathrm{e}^{\kappa\tau} - 1\right).$$

Now, if we take $\mathbf{u} = [0, 0]$ we get the price for a zero-coupon bond, i.e.,:

$$P(t, T) \overset{def}{=} \mathbb{E}^{\mathbb{Q}}\left(\mathrm{e}^{-\int_t^T r_s \mathrm{d}s} | \mathcal{F}_t\right) = \mathrm{e}^{A(0,\tau) + B_r(0,\tau) r_t + B_v(0,\tau) v_t}. \tag{A.4}$$

The zero-coupon bonds can easily be priced via Formula (A.4) and numerical integration for solving (A.3). Therefore, we derive an approximation for the log-swap rate characteristic function. Take $T_0 > 0$ as the start time of the interest rate swap, and $T > T_0$ the end time of the swap, and t such that $t < T_0 < T$. Suppose that fixed payments occur at times T_k, $n < k < M$ where $T_M = T$. For simplicity, we assume $T_n = T_0$ and take $\tau_k = T_k - T_{k-1}$.

By definition the forward swap rate at time t is given by:

$$\mathrm{Swap}_t^{T_0,T} = \frac{P(t, T_0) - P(t, T)}{\sum_{j=n+1}^{M} \tau_j P(t, T_j)} := S_t^{T_0,T}. \tag{A.5}$$

The swaption price can be expressed as the risk-neutral expectation of the discounted payoff, i.e.:

$$\Pi\left(t, S_t^{T_0,T}\right) = B_t \mathbb{E}^{\mathbb{Q}}\left(\frac{C_{T_0}^{T_0,T}}{B_{T_0}} \max\left\{S_{T_0}^{T_0,T} - K, 0\right\} | \mathcal{F}_t\right), \tag{A.6}$$

with

$$C_t^{T_0,T} = \sum_{j=n+1}^{M} \tau_j P(t, T_j). \tag{A.7}$$

Now, by changing the risk neutral measure $\mathbb{Q}$ to the swap rate measure $\mathbb{Q}^{\Pi}$ (see for example [Jamshidian-1989; Geman-1995]) induced by numéraire $C_t^{T_0,T}$ we have:

$$\mathrm{d}\mathbb{Q} = \frac{B_{T_0}}{B_t} \frac{C_t^{T_0 T}}{C_{T_0}^{T_0,T}} \mathrm{d}\mathbb{Q}^{\Pi}. \tag{A.8}$$

Therefore:

$$\Pi\left(t, S_t^{T_0,T}\right) = \mathbb{E}^{\Pi}\left(C_t^{T_0,T} \max\left\{S_{T_0}^{T_0,T} - K, 0\right\} | \mathcal{F}_t\right), \tag{A.9}$$

and since $C_t^{T_0,T}$ is already determined at time t we have:

$$\Pi\left(t, S_t^{T_0,T}\right) = \sum_{j=n+1}^{M} \tau_j P(t, T_j) \mathbb{E}^{\Pi}\left(\max\left\{S_{T_0}^{T_0,T} - K, 0\right\} | \mathcal{F}_t\right). \tag{A.10}$$

The price of a swaption at time t is expressed as an expectation under the swap measure multiplied by a sum of zero-coupon bonds. Since the swap-rate $S_{T_0}^{T_0,T}$ is a random quantity we want to find its dynamics. The next lemma provides the necessary results:

Lemma A.1 (Dynamics of Swaprate $S_{T_0}^{T_0,T}$)**.** *The swap rate given by Equation (A.5) has the following dynamics:*

$$\frac{\mathrm{d}S_t^{T_0,T}}{S_t^{T_0,T}} = \frac{\mathrm{d}P(t,T_0)-\mathrm{d}P(t,T)}{P(t,T_0)-P(t,T)} - \frac{\sum_{j=n+1}^{M}\tau_j \mathrm{d}P(t,T_j)}{\sum_{j=n+1}^{M}\tau_j P(t,T_j)}. \tag{A.11}$$

The stochastic differential equation in Lemma A.11 depends on the dynamics of the appropriate zero-coupon bonds. Since the swap rate under its measure does not contain $\mathrm{d}t$-terms, the dynamics for the zero-coupon bond also have to be drift-less. From the previous derivations we found that a zero-coupon bond reads:

$$P(t,T) = \mathrm{e}^{A(0,\tau)+B_r(0,\tau)r_t+B_v(0,\tau)v_t}. \tag{A.12}$$

Therefore, the dynamics without the $\mathrm{d}t$-terms are given by:

$$\frac{\mathrm{d}P(t,T)}{P(t,T)} = B_r(0,\tau)\mathrm{d}r_t + B_v(0,\tau)\mathrm{d}v_t. \tag{A.13}$$

By *freezing* the zero-coupon bonds, see [Schrager,Pelsser-2002], at initial time, $t=0$, we obtain the following approximation:

$$\frac{\mathrm{d}S_t^{T_0,T}}{S_t^{T_0,T}} \approx \frac{P(0,T_0)}{P(0,T_0)-P(0,T)}\left(B_r(0,T_0-t)\mathrm{d}r_t + B_v(0,T_0-t)\mathrm{d}v_t\right) \tag{A.14}$$

$$- \frac{P(0,T)}{P(0,T_0)-P(0,T)}\left(B_r(0,T-t)\mathrm{d}r_t + B_v(0,T-t)\mathrm{d}v_t\right) \tag{A.15}$$

$$- \frac{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}\left(B_r(0,T_j-t)\mathrm{d}r_t + B_v(0,T_j-t)\mathrm{d}v_t\right). \tag{A.16}$$

With,

$$\Gamma_t^r = \frac{P(0,T_0)B_r(0,T_0-t)}{P(0,T_0)-P(0,T)} - \frac{P(0,T)B_r(0,T-t)}{P(0,T_0)-P(0,T)} - \frac{\sum_{j=n+1}^{M}\tau_j P(0,T_j)B_r(0,T_j-t)}{\sum_{j=n+1}^{M}\tau_j P(0,T_j)},$$

and,

$$\Gamma_t^v = \frac{P(0,T_0)B_v(0,T_0-t)}{P(0,T_0)-P(0,T)} - \frac{P(0,T)B_v(0,T-t)}{P(0,T_0)-P(0,T)} - \frac{\sum_{j=n+1}^{M}\tau_j P(0,T_j)B_v(0,T_j-t)}{\sum_{j=n+1}^{M}\tau_j P(0,T_j)},$$

we end up with the following approximation for the dynamics of the swap rate:

$$\frac{\mathrm{d}S_t^{T_0,T}}{S_t^{T_0,T}} \approx \Gamma_t^r \mathrm{d}r_t + \Gamma_t^v \mathrm{d}v_t. \tag{A.17}$$

A.0.1 Swaprate dynamics under swap measure

The swap rate dynamics are given in terms of the dynamics of the underlying processes, r_t, v_t, and some time-dependent functions, Γ_t^1 and Γ_t^2. Previously, we found that in order to price swaption (Equation(A.10)) under the swap measure $\mathbb{Q}^{\Pi}$ one needs to find the dynamics of $\mathrm{d}r_t$ and $\mathrm{d}v_t$ in that measure first.

Using Equation (A.8) we can write:

$$\mathrm{d}\mathbb{Q}^{T_0,T} = \Lambda_t \mathrm{d}\mathbb{Q}, \text{ with } \Lambda_t = \frac{C_t^{T_0,T}}{B_t}\frac{B_0}{C_0^{T_0,T}}. \tag{A.18}$$

From the Girsanov Theorem we know that Λ_t has to be a martingale under the risk-neutral measure $\mathbb{Q}$. In order to extract the Girsanov kernel, we need the diffusion term from the dynamics for Λ_t:

$$\begin{aligned} \mathrm{d}\Lambda_t &= \frac{B_0}{C_0^{T_0,T}} \mathrm{d}\left(\frac{C_t^{T_0,T}}{B_t}\right), && \text{(A.19)} \\ &= (\dots)\mathrm{d}t - \frac{B_0}{C_0^{T_0,T}} \frac{C_t^{T_0,T}}{B_t^2} \mathrm{d}B_t + \frac{B_0}{C_0^{T_0,T}} \frac{1}{B_t} \mathrm{d}C_t^{T_0,T}. && \text{(A.20)} \end{aligned}$$

Recall that the money savings account, B_t, is given by:

$$\mathrm{d}B_t = r_t \exp\left(\int_0^t r_s \mathrm{d}s\right) \mathrm{d}t, \tag{A.21}$$

which gives:

$$\begin{aligned} \mathrm{d}\Lambda_t &= (\dots)\mathrm{d}t + \frac{B_0}{C_0^{T_0,T}} \frac{1}{B_t} \mathrm{d}C_t^{T_0,T}, && \text{(A.22)} \\ &= (\dots)\mathrm{d}t + \frac{B_0}{C_0^{T_0,T}} \frac{1}{B_t} \sum_{j=n+1}^{M} \tau_j \mathrm{d}P(t,T_j). && \text{(A.23)} \end{aligned}$$

With Equation (A.13), we obtain:

$$\mathrm{d}\Lambda_t = (\dots)\mathrm{d}t + \frac{B_0}{C_0^{T_0,T}} \frac{1}{B_t} \sum_{j=n+1}^{M} \tau_j P(t,T_j) \left(B_r(0,T_j-t)\mathrm{d}r_t + B_v(0,T_j-t)\mathrm{d}v_t\right) \tag{A.24}$$

This simply reads:

$$\begin{aligned} \frac{\mathrm{d}\Lambda_t}{\Lambda_t} &\approx (\dots)\mathrm{d}t + \frac{1}{C_t^{T_0,T}} \sum_{j=n+1}^{M} \tau_j P(t,T_j) \left(B_r(0,T_j-t)\mathrm{d}r_t + B_v(0,T_j-t)\mathrm{d}v_t\right), && \text{(A.25)} \\ &\approx (\dots)\mathrm{d}t + \frac{1}{C_0^{T_0,T}} \sum_{j=n+1}^{M} \tau_j P(0,T_j) \left(B_r(0,T_j-t)\mathrm{d}r_t + B_v(0,T_j-t)\mathrm{d}v_t\right), && \text{(A.26)} \end{aligned}$$

Finally, in terms of the independent Brownian motions, $\mathrm{d}\widetilde{W}_t^r$ and $\mathrm{d}\widetilde{W}_t^v$, the process $\mathrm{d}\Lambda_t$ can be expressed as:

$$\begin{aligned} \frac{\mathrm{d}\Lambda_t}{\Lambda_t} = (\dots)\mathrm{d}t + \frac{1}{C_0^{T_0,T}} \Bigg(& \sum_{j=n+1}^{M} \tau_j P(0,T_j) \left(B_r(0,T_j-t)\eta\sqrt{v_t} + B_v(0,T_j-t)\gamma\rho_{r,v}\sqrt{v_t}\right) \mathrm{d}\widetilde{W}_t^r \\ & + \sum_{j=n+1}^{M} \tau_j P(0,T_j) \left(\gamma\sqrt{v_t}\sqrt{1-\rho_{r,v}^2} B_v(0,T_j-t)\right) \mathrm{d}\widetilde{W}_t^v \Bigg). \end{aligned}$$

Using the Girsanov transform:

$$\begin{cases} \mathrm{d}\widetilde{W}_t^r = \phi_t^r \mathrm{d}t + \mathrm{d}\widetilde{W}_t^{r,\Pi}, \\ \mathrm{d}\widetilde{W}_t^v = \phi_t^v \mathrm{d}t + \mathrm{d}\widetilde{W}_t^{v,\Pi}, \end{cases} \tag{A.27}$$

one can simply find that the dynamics of r_t and v_t under the swap measure are given by:

$$\begin{cases} \mathrm{d}r_t = \kappa(\theta - r_t + \zeta_t^r v_t)\mathrm{d}t + \eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi}, \\ \mathrm{d}v_t = \widetilde{\lambda}\left(\widetilde{v} - v_t\right)\mathrm{d}t + \gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi} \\ \rho_{r,v}\mathrm{d}t = \mathrm{d}W_t^{r,\Pi}\mathrm{d}W_t^{v,\Pi}, \end{cases} \tag{A.28}$$

where

$$\zeta_t^r = \frac{1}{\kappa}\frac{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}\left(B_r(0,T_j-t)\eta^2 + B_v(0,T_j-t)\gamma\eta\rho_{r,v}\right),$$

with

$$\widetilde{\lambda} = \lambda - \zeta_t^v,$$
$$\widetilde{v} = \frac{\bar{v}}{1-\zeta_t^v/\lambda},$$

and,

$$\zeta_t^v = \frac{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}{\sum_{j=n+1}^{M}\tau_j P(0,T_j)}\left(B_r(0,T_j-t)\rho_{r,v}\eta\gamma + \gamma^2 B_v(0,T_j-t)\right).$$

We know that the swap rate, $S_t^{T_0,T}$, is a martingale under the swap measure, so it does not include $\mathrm{d}t$-terms. Therefore, we have:

$$\frac{\mathrm{d}S_t^{T_0,T}}{S_t^{T_0,T}} \approx \Gamma_t^r\left(\eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi}\right) + \Gamma_t^v\left(\gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi}\right). \tag{A.29}$$

In order to obtain an affine process for the swap rate, we take the log-transform, i.e., $\log S_t^{T_0,T}$, for which we find the following dynamics:

$$\begin{aligned}\mathrm{d}\log S_t^{T_0,T} &= \frac{1}{S_t^{T_0,T}}(\mathrm{d}S_t^{T_0,T}) - \frac{1}{2}\frac{1}{(S_t^{T_0,T})^2}\left(\mathrm{d}S_t^{T_0,T}\right)^2,\\ &= \Gamma_t^r\eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi} + \Gamma_t^v\gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi} - \frac{1}{2}\left(\Gamma_t^r\eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi} + \Gamma_t^v\gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi}\right)^2.\end{aligned}$$

So, under the swap measure in the log-space for the state vector, $\mathbf{Y}_t = [\pi_t^{T_0,T} := \log S_t^{T_0,T}, r_t, v_t]$, we need to solve the following system of SDEs:

$$\begin{cases}\mathrm{d}\pi_t^{T_0,T} = & \Omega_t v_t\mathrm{d}t + & \Gamma_t^r\eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi} + & \Gamma_t^v\gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi},\\ \mathrm{d}r_t = & \kappa(\theta - r_t + \zeta_t^r v_t)\mathrm{d}t + & \eta\sqrt{v_t}\mathrm{d}W_t^{r,\Pi}, & \\ \mathrm{d}v_t = & \widetilde{\lambda}(\widetilde{v} - v_t)\mathrm{d}t + & \gamma\sqrt{v_t}\mathrm{d}W_t^{v,\Pi}, & \\ \rho_{r,v}\mathrm{d}t = & \mathrm{d}W_t^{r,\Pi}\mathrm{d}W_t^{v,\Pi}, & & \end{cases} \tag{A.30}$$

with $\Omega_t = -\frac{1}{2}(\Gamma_t^r)^2\eta^2 - \frac{1}{2}(\Gamma_t^v)^2\gamma^2 - (\Gamma_t^r)(\Gamma_t^v)\eta\gamma\rho_{r,v}$.

Since we deal with a 3D system (A.30) of affine SDEs for pricing the log-swap rate, the corresponding characteristic function is of the following form:

$$\phi_\pi(\mathbf{u},\mathbf{V}_t,\tilde{\tau}) = \mathrm{e}^{A(\mathbf{u},\tilde{\tau}) + B_\pi(\mathbf{u},\tilde{\tau})\pi_t + B_r(\mathbf{u},\tilde{\tau})r_t + B_v(\mathbf{u},\tilde{\tau})v_t}, \tag{A.31}$$

with $\tilde{\tau} = T_0 - t$ and $\pi_t = \log S_t^{T_0,T}$, as in (A.6), and the boundary condition,

$$\phi_\pi(\mathbf{u},\mathbf{V}_0,0) = \mathrm{e}^{i\mathbf{u}^{\mathrm{T}}\mathbf{V}_{T_0}}.$$

For the vector $\mathbf{u} = [u,0,0]^{\mathrm{T}}$ we obtain the following set of ODEs for the characteristic function: $B_\pi(u,\tilde{\tau}) = iu$, $B_r(u,\tilde{\tau}) = 0$,

$$\frac{\mathrm{d}}{\mathrm{d}\tilde{\tau}}B_v(u,\tilde{\tau}) = \Omega_t B_\pi - \lambda B_v + \frac{1}{2}\left(a_1(\tilde{\tau})B_\pi^2 + \gamma^2 B_v^2\right) + a_2(\tilde{\tau})B_\pi B_v, \tag{A.32}$$

and

$$A(u,\tilde{\tau}) = \int_0^{\tilde{\tau}}\widetilde{\lambda}\widetilde{v}B_v(u,s)\mathrm{d}s, \tag{A.33}$$

with: $a_1(\tilde{\tau}) = \left((\psi_{\tilde{\tau}}^1)^2\eta^2 + (\psi_{\tilde{\tau}}^2)^2\gamma^2 + 2\psi_{\tilde{\tau}}^1\psi_{\tilde{\tau}}^2\gamma\eta\rho_{r,v}\right)$ and $a_2(\tilde{\tau}) = \left(\psi_{\tilde{\tau}}^1\eta\gamma\rho_{r,v} + \psi_{\tilde{\tau}}^2\gamma^2\right)$.