To appear in *Quantitative Finance*, Vol. 00, No. 00, Month 20XX, 1–31

A Novel Monte Carlo Approach to Hybrid Local Volatility Models

Anthonie W. van der Stoep*† ‡ §, Lech A. Grzelak§¶ and Cornelis W. Oosterlee‡§

†Pricing Model Validation (PMV), Rabobank, Utrecht, the Netherlands
‡Centrum Wiskunde & Informatica (CWI) – National Research Institute for Mathematics and Computer Science, Amsterdam, the Netherlands
§Delft Institute of Applied Mathematics (DIAM), TU Delft, Delft, the Netherlands
¶Quantitative Analytics, ING, Amsterdam, the Netherlands

(*v2.0 released September 2016*)

We present in a Monte Carlo simulation framework a novel approach for the evaluation of hybrid local volatility (Dupire 1994, Derman and Kani 1998) models. In particular, we consider the stochastic local volatility model – see e.g. Lipton *et al.* (2014), Piterbarg (2007), Tataru and Fisher (2010), Lipton (2002) – and the local volatility model incorporating stochastic interest rates – see e.g. Atlan (2006), Piterbarg (2006), Deelstra and Rayée (2012), Ren *et al.* (2007). For both model classes a particular (conditional) expectation needs to be evaluated, which cannot be extracted from the market and is expensive to compute. We establish accurate and 'cheap to evaluate' approximations for the expectations by means of the stochastic collocation method (Babuška *et al.* 2007, Xiu and Hesthaven 2005, Beck *et al.* 2012, Nobile *et al.* 2008, Sankaran and Marsden 2011), which was recently applied in the financial context (Grzelak *et al.* 2014, Grzelak and Oosterlee 2017), combined with standard regression techniques. Monte Carlo pricing experiments confirm that our method is highly accurate and fast.

Keywords: Local volatility, Monte Carlo, hybrid, stochastic volatility, stochastic local volatility, stochastic interest rates, stochastic collocation, regression, SABR, Heston, Hull-White.

JEL Classification: C15; C63

1. Introduction

In this article we propose for two types of *hybrid local volatility models* a novel, highly efficient Monte Carlo simulation method. We consider stochastic local volatility (SLV) models and the local volatility model incorporating stochastic interest rates. These hybrid models, by construction, can be calibrated perfectly to the plain vanilla market, while (partially) inheriting particular desirable features from their 'pure' stochastic volatility counterparts or including stochastic interest rates, which may yield an enhancement in the pricing of long-dated FX and equity-linked structures. Although this makes these models attractive to the financial industry, their evaluation is not trivial. A particular (conditional) expectation needs to be established, which cannot be extracted from the market quotes. The *stochastic collocation method* (Babuška *et al.* 2007, Xiu and Hesthaven 2005, Beck *et al.* 2012, Nobile *et al.* 2008, Sankaran and Marsden 2011), which was recently applied in the financial context (Grzelak and Oosterlee 2017, Grzelak *et al.* 2014), allows us to determine for both types of hybrid local volatility models the expectation in a way that is highly accurate and

*Corresponding author. Email: Anton.van.der.Stoep@rabobank.com

fast.

In a nutshell, the concept of stochastic collocation is to approximate a problem variable of interest, which is expensive to compute, as a function of a more convenient 'cheap to compute' random variable. The function is a polynomial expansion, which is established by determining basis functions and computing interpolation points, based on the input distribution (corresponding to the 'cheap' variable). An increase in the order of this polynomial yields exponential convergence.

The paper is organized as follows. In the follow-up section we describe the main characteristics of the traditional local volatility model and the Heston and SABR stochastic volatility models. We proceed with a brief overview of evaluation approaches for stochastic local volatility models and the local volatility model incorporating stochastic interest rates in Sections 1.2 and 1.3, respectively. The basics of stochastic collocation are discussed in Section 1.4. In Section 2 we present and numerically test the stochastic collocation based approach for the evaluation of stochastic local volatility models. We subsequently apply this method to the local volatility model enhanced by stochastic interest rates in Section 3. Section 4 concludes.

1.1. *Local volatility & stochastic volatility models*

In the financial industry, the local volatility (LV) model introduced by Dupire (Dupire 1994) and Derman & Kani (Derman and Kani 1998) has widely been employed for managing smile and skew risk. As it is well-known, by its construction, the LV model can be calibrated *perfectly* to any set of arbitrage-free European-type option prices. Although this makes the model practically appealing, it has certain undesirable features; it e.g. exhibits a *flattening* of the forward smile, which may lead to a significant mispricing of forward volatility sensitive contracts like forward-starting and path-dependent options (Van der Stoep *et al.* 2014, Rebonato 1999, Clark 2011). Also, it does not accurately predict the smile movement when the value of the underlying changes, resulting in possibly unstable hedges (Hagan *et al.* 2002, Baker *et al.* 2004, Johnson and Lee 2003). Further, the model assumes deterministic interest rates, which may imply mispricing of long-term interest rate sensitive hybrids, like the *Power-Reverse Dual-Currency (PRDC)* notes in the FX market, see e.g. Piterbarg (2006), Deelstra and Rayée (2012), Bloch and Nakashima (2008).

Alternative models to overcome these issues are the well-established Heston (Heston 1993) and SABR (Hagan *et al.* 2002) models, in which the volatility is assumed to follow a stochastic process. The Heston model predicts that the forward smile has a shape that is similar to the implied volatility smile observed today, which is more in line with market observations and yields a more accurate pricing of forward volatility sensitive products (Baker *et al.* 2004, Gatheral 2006). Further, the SABR model is, besides for the availability of a closed-form formula for the implied volatility ('Hagan's formula'), well-known for the prediction that the smile 'follows' the underlying. In fact, one of the main reasons for Hagan et al. (Hagan *et al.* 2002) to introduce the SABR model is the typically inaccurate smile movement with respect to the underlying, predicted by the local volatility model. When considering long-term products like PRDCs, interest rates can no longer be assumed deterministic – introducing a short-rate process in the model, like Vasicek, Black & Karasinkski (Black and Karasinski 1991) and Hull & White (Hull and White 1993) may enhance the pricing results.

In this article we consider the SABR and Heston models, enhanced by a non-parametric local volatility component, which we refer to as the *SABR-Local Volatility*[1] (SABR-LV) and *Heston-Stochastic Local Volatility* (Heston-SLV or H-SLV) models. Also, we study the local volatility model incorporating stochastic interest rates governed by Hull-White dynamics, the so-called *Local Volatility-Hull White* (LV-HW) model. The SABR-LV and Heston-SLV models, compared to the traditional local volatility model, typically yield a more stable hedging performance and a more

[1]Note that, in fact, the 'pure' SABR model is already a stochastic local volatility model with a *parametric* local volatility component.

accurate pricing of forward volatility sensitive products. Enriching the local volatility model with (Hull-White) stochastic interest rates enhances the pricing of long-dated FX and equity-linked structures.

We evaluate the SABR-LV, Heston-SLV and LV-HW models based on a method that combines stochastic collocation and standard regression techniques. In the follow-up Sections 1.2 and 1.3 an overview of other approaches for the evaluation of hybrid local volatility models is provided. The basics of stochastic collocation are discussed in Section 1.4.

1.2. *Stochastic local volatility models*

The class of stochastic local volatility models was developed by Jex et al. (Jex *et al.* 1999) and Lipton (Lipton 2002, Lipton and McGhee 2002)[1], amongst others. As e.g. pointed out in Clark (2011) and Lipton *et al.* (2014), for the pricing of FX options SLV models are typically used. Also in the Bloomberg note Tataru and Fisher (2010) stochastic local volatility models are presented in an FX context.

In an SLV framework a conditional expectation of the form $\mathrm{E}\left[\psi^2(V(t))\middle| S(t) = K\right]$ needs to be established, see e.g. Dupire (1996). The exact form of $\psi(\cdot)$ depends on the choice of stochastic volatility model, e.g. for the Heston model $\psi(x) = \sqrt{x}$. The conditional expectation cannot be extracted from the market. A common evaluation approach consists of solving the Kolmogorov forward equation (Deelstra and Rayée 2012, Ren *et al.* 2007, Clark 2011) forward one step at a time, recovering simultaneously the conditional expectation and the complete stochastic local volatility component. In this iterative procedure the joint density of $S(\cdot)$ and $V(\cdot)$ is solved for all time-points. PDE-discretization techniques are common practice in the financial industry in a hybrid local volatility context. As an alternative to the standard ADI methods, Lipton et al. (Lipton *et al.* 2014), in a Quadratic Local Stochastic Volatility (QLSV) framework, introduce a Galerkin-Ritz inspired method for solving a system of PDEs and demonstrate that it is efficient. Another approach to handle the problem of computational burden is presented in Tian *et al.* (2015), who employ GPUs to accelerate the computations.

The *Markovian projection* technique has also been applied in an SLV context (Henry-Labordère 2009, Piterbarg 2007). Although this method is generally applicable, it involves several conditional expectations that typically need to be approximated. Moreover, this technique does not preserve marginal distributions of order higher than one. This may result in a significant mismatch in prices of contracts depending on stock values at multiple times, such as American and barrier options, implied by the original and projected models.

Other attempts for solving the SLV model are presented in Tataru and Fisher (2010), where a Levenberg-Marquardt optimization technique for a non-linear Fokker-Planck equation is applied and in Deelstra and Rayée (2012), where zero correlation is assumed between the volatility process and the underlying asset, yielding an efficient simulation of the extended Schöbel-Zhu model (Schöbel and Zhu 1999). In a more general framework, based on the theory of generalized Wiener functionals, see e.g. Watanabe (1987), An and Li (An and Li 2015) provide closed-form expansions for evaluating a general conditional expectation that involves marginal distributions which are generated by stochastic differential equations. In Lorig *et al.* (2015) for a general class of stochastic local volatility models a family of asymptotic expansions for European-style option prices and implied volatilities is derived. Further, in Pascucci and Mazzon (2015) the authors derive an asymptotic expansion for forward-starting options in a multi-factor local-stochastic volatility model, which results in explicit approximation formulas for the forward implied volatility. In Van der Stoep *et al.* (2014) we introduced in a Monte Carlo setting a non-parametric method for the evaluation of the problematic conditional expectation, which relies on splitting the Monte Carlo realizations in bins.

[1] In Lipton and McGhee (2002) Lipton and McGhee present a more general form of stochastic local volatility models including jumps.

A similar idea was presented in e.g. Guyon and Henry-Labordère (2012), based on kernel estimators in an interacting particle system. All the above-mentioned numerical techniques are found to be relatively costly, or limited in applicability.

1.3. *Local volatility model with stochastic interest rates*

Regarding local volatility in a stochastic interest rates (SIR) framework, the literature is less rich. In Atlan (2006) Gyöngy's (Gyöngy 1986) mimicking techniques are used to incorporate stochastic interest rates in a local volatility framework. More generally, he shows how Gyöngy's theorem can be used to relate any continuous stochastic volatility model with stochastic interest rates to local volatility with deterministic interest rates.

In Piterbarg (2006) Piterbarg states that the slope of the FX volatility is a major factor affecting the values of PRDC swaps. He therefore comes up with a skew-enabled model, namely the local volatility model with domestic and foreign interest rates following Hull-White dynamics, which serves as an extension to the traditional three-factor log-normal model (without skew). For the stability of the calibration a CEV specification for the local volatility function is chosen. This yields an essentially instantaneous calibration procedure which is based on a Markovian representation technique[1] of the dynamics of the forward FX rate and skew averaging techniques. The calibration basically 'captures' mainly the slope of the implied volatility.

Further, in Ren *et al.* (2007) an expression is derived for the local volatility in a stochastic interest rates framework, consisting of the particular expectation $\mathrm{E}^T\left[r(T)\mathbb{1}_{S(T)>K}\right]$, which they compute by iteratively solving the corresponding Kolmogorov forward equation forward in time. Benhamou et al. (Benhamou *et al.* 2008) specify the bias between the local volatility with and without stochastic interest rates. By means of numerical experiments they illustrate the importance of this bias which, in line with intuition, gets larger for longer maturity. In another paper, based on his work on perturbation methods for local volatility models, Benhamou (Benhamou *et al.* 2012) presents and numerically tests the expansions approximating the prices of European options in a local volatility model with stochastic interest rates. In numerical experiments, similar as in Piterbarg (2006), a CEV diffusion for the spot is chosen. In Deelstra and Rayée (2012) the authors present, in an FX context with stochastic interest rates, four methods to compute the local volatility function for different strikes and time-points. Although this article provides a clear overview of the ways the local volatility component can be computed, no concrete calibration or pricing experiments are included. Last, in Guyon and Henry-Labordère (2012) the authors evaluate the Ho-Lee/Dupire hybrid model by an approach that is based on McKean's particle method.

1.4. *Stochastic collocation basics*

In this section we briefly discuss the basics of the stochastic collocation method. The original idea of stochastic collocation is to project uncertainty onto a probability space with known properties and conditions (Babuška *et al.* 2007, Xiu and Hesthaven 2005). In particular, we approximate a variable of interest, Y, which is expensive to compute, by a function of a more convenient 'cheap to evaluate' random variable X.

Collocation methods have been studied and employed in various disciplines for uncertainty quantification, see e.g. Xiu and Hesthaven (2005), Ganapathysubramanian and Zabaras (2007), Witteveen and Iaccarino (2012). In collocation methods the target is to satisfy governing differential equations at a discrete set of points, in the corresponding probability space. Two of the main approaches of high-order stochastic collocation methods are the Lagrange interpolation approach, see e.g. Xiu and Hesthaven (2005), and the pseudo-spectral generalized polynomial chaos approach from e.g. Xiu (2007).

[1] In Piterbarg (2007) Piterbarg formalizes this procedure as the *Markovian projection* method.

We explain the stochastic collocation method in a sampling setting. Suppose we wish to sample values y_n from the distribution of Y. This is typically established by first drawing samples u_n from a standard uniform distribution $U \overset{\mathrm{d}}{=} \mathcal{U}([0,1])$ and subsequently applying the inversions $y_n = F_Y^{-1}(u_n)$. However, this sampling approach is not preferred in the case that the inversion $F_Y^{-1}(\cdot)$ is *expensive* – the sampling is not performed in an efficient way. By the stochastic collocation method this issue is overcome.

The stochastic collocation technique relies on the fact that $F_Y(Y) \overset{\mathrm{d}}{=} U \overset{\mathrm{d}}{=} F_X(X)$, for an arbitrary random variable X, with $U \overset{\mathrm{d}}{=} \mathcal{U}([0,1])$, i.e. the CDFs of Y and X (*not* Y and X themselves) are equal in distribution. In a sampling setting, with y_n and x_n denoting samples from the distributions of Y and X, respectively, the target is to find a function $g(\cdot)$ such that

$$F_Y(g(x_n)) = F_X(x_n), \;\; y_n = g(x_n). \tag{1}$$

When the function $g(\cdot)$ is determined, sampling from Y can be performed by sampling from X, without performing the expensive inversion $F_Y^{-1}(\cdot)$, which is needed when sampling in the traditional way. Trivially, equation (1) implies $g(\cdot) = F_Y^{-1}(F_X(\cdot))$. However, the task is to find a function which does *not* require many expensive inversions $F_Y^{-1}(\cdot)$. This can be achieved in a polynomial chaos expansion framework, where a sample y_n is approximated in terms of Lagrange basis polynomials $\ell(\cdot)$ evaluated at a sample of X, x_n, as

$$y_n = g(x_n) \approx g_N(x_n) := \sum_{i=1}^{N} y_i \ell_i(x_n), \;\; \ell_i(x_n) := \prod_{j=1, j\neq i}^{N} \frac{x_n - x_j}{x_i - x_j},$$

where x_i and x_j are so-called *collocation points*, y_i is the exact 'expensive' evaluation at the collocation point x_i, i.e. $y_i = F_Y^{-1}(F_X(x_i))$. Choosing the interpolation polynomial in the Lagrange form is well-accepted in the field of Uncertainty Quantification (when the stochastic collocation method is applied), see e.g. Sankaran and Marsden (2011). By a change of basis it can be written in terms of monomials, $g_N(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_{N-1} x_n^{N-1}$, where the coefficients a_0, a_1, …, a_{N-1} are obtained by solving a linear system $\mathbf{Va} = \mathbf{y}$, with matrix $\mathbf{V}$ denoting the Vandermonde matrix (see e.g. Grzelak *et al.* (2014) for more details).

Once the function $g_N(\cdot)$ is established by N expensive inversions $F_Y^{-1}(\cdot)$, we are able to generate any number of samples y_n without significant additional cost. The collocation points x_i can basically be chosen arbitrarily, however we choose them in an *optimal way*, i.e. based on the zeros of an orthogonal polynomial.

In this article we choose X to be standard normally distributed. This implies that the *optimal* collocation points are the zeros of the Hermite polynomials (abscissas of the Gauss-Hermite quadrature)(Abramowitz and Stegun 1972, Grzelak *et al.* 2014). It turns out that choosing X to be normally distributed works highly satisfactory; the method is accurate and efficient, as the inversion of a normal distribution is 'cheap'. We therefore do not consider other distributions for X, which may yield a method with similar accuracy, but more expensive inversions. What is more, the Cameron-Martin Theorem (Cameron and Martin 1947) states that polynomial chaos approximations based on the normal distribution converge to any distribution.

In the following sections we employ the stochastic collocation method for the efficient Monte Carlo evaluation of stochastic local volatility models and the local volatility model incorporating stochastic interest rates.

2. Stochastic Local Volatility Models

In this section we discuss stochastic local volatility models. In Jex *et al.* (1999), Lipton (2002), Lipton and McGhee (2002), amongst others, the class of stochastic local volatility models was

introduced, which combine properties of the traditional local volatility model (Dupire 1994, Derman *et al.* 1995) and stochastic volatility models, like the SABR model (Hagan *et al.* 2002) and the Heston model (Heston 1993). According to Clark (Clark 2011) neither the 'sticky-delta' property of stochastic volatility models nor the 'sticky-strike' characteristic corresponding to the local volatility model is in line with the actual smile behaviour in FX markets; the reality is somewhere between the two and therefore typically a stochastic local volatility model is used. In line with this, as pointed out in Lipton *et al.* (2014), SLV models are *de facto* standard for pricing FX options.

Assuming constant deterministic interest rate r, no dividends and instantaneous correlation $\rho_{s,v}$, the general SLV model dynamics under the risk-neutral Q-measure[1] read

$$\mathrm{d}S(t)/S(t) = r\mathrm{d}t + \sigma(t,S(t))\psi(S(t),V(t))\mathrm{d}W_s^{\mathrm{Q}}(t),\ S(0)=S_0 \tag{2}$$

$$\mathrm{d}V(t) = a_v(t,V(t))\mathrm{d}t + b_v(t,V(t))\mathrm{d}W_v^{\mathrm{Q}}(t),\ V(0)=V_0 \tag{3}$$

with $\mathrm{d}W_s^{\mathrm{Q}}(t)\mathrm{d}W_v^{\mathrm{Q}}(t) = \rho_{s,v}\mathrm{d}t$ and[2] $\sigma^2(t,K) = \sigma^2_{\mathrm{LV}}(t,K)/\mathrm{E}^{\mathrm{Q}}\left[\psi^2(S(t),V(t))\middle|\, S(t)=K\right]$, where $\sigma^2_{\mathrm{LV}}(t,K)$ denotes Dupire's local volatility component (Dupire 1994):

$$\sigma^2_{\mathrm{LV}}(t,K) = \frac{\frac{\partial C(t,K)}{\partial t} + rK\frac{\partial C(t,K)}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C(t,K)}{\partial K^2}}.$$

For notation purposes, we suppress the Q-superscript from this point on. By choosing $\psi(S(t),V(t)) = \sqrt{V(t)}S^{\beta-1}(t)$, $a_v(t,V(t)) = \gamma^2 V(t)$ and $b_v(t,V(t)) = 2\gamma V(t)$ we obtain the *SABR-LV model*[3]:

$$\mathrm{d}S(t)/S(t) = r\mathrm{d}t + \sigma(t,S(t))\sqrt{V(t)}S^{\beta-1}(t)\mathrm{d}W_s(t),\ S(0)=S_0 \tag{4}$$

$$\mathrm{d}V(t) = \gamma^2 V(t)\mathrm{d}t + 2\gamma V(t)\mathrm{d}W_v(t),\ V(0)=V_0 \tag{5}$$

with $\mathrm{d}W_s(t)\mathrm{d}W_v(t) = \rho_{s,v}\mathrm{d}t$. In the SABR-LV model the local volatility component is specified by

$$\sigma^2(t,K) = \frac{\sigma^2_{\mathrm{LV}}(t,K)}{K^{2\beta-2}\mathrm{E}\left[V(t)\middle|\, S(t)=K\right]}. \tag{6}$$

We use a standard Euler discretization scheme to simulate the SABR-LV model.

The choices $\psi(S(t),V(t)) = \sqrt{V(t)}$, $a_v(t,V(t)) = \kappa(\bar{V}-V(t))$ and $b_v(t,V(t)) = \gamma\sqrt{V(t)}$ provide us with the *Heston-SLV (H-SLV) model*:

$$\mathrm{d}S(t)/S(t) = r\mathrm{d}t + \sigma(t,S(t))\sqrt{V(t)}\mathrm{d}W_s(t),\ S(0)=S_0 \tag{7}$$

$$\mathrm{d}V(t) = \kappa\left(\bar{V}-V(t)\right)\mathrm{d}t + \gamma\sqrt{V(t)}\mathrm{d}W_v(t),\ V(0)=V_0 \tag{8}$$

---

[1] Note that the general SLV model as described by equations (2)-(3) is an incomplete market model, which implies that a unique risk-neutral pricing measure does not exist, see e.g. Fouque *et al.* (2000).

[2] A derivation of the local volatility component, consisting of Dupire's local volatility and a conditional expectation, can be found in e.g. Gatheral (2006), Van der Stoep *et al.* (2014).

[3] To prevent double use of the σ-notation we write the variance dynamics instead of the more common volatility dynamics. The traditional SABR model dynamics are given by the following two SDEs (Hagan *et al.* 2002, Rebonato *et al.* 2011):

$$\mathrm{d}F^T(t) = \sigma(t)(F^T(t))^{\beta}\mathrm{d}W_s^T(t),\ \ \mathrm{d}\sigma(t) = \gamma\sigma(t)\mathrm{d}W_v^T(t),$$

with $F^T(\cdot)$ denoting the forward corresponding to expiry T and $\mathrm{d}W_s^T(t)\mathrm{d}W_v^T(t) = \rho_{s,v}\mathrm{d}t$.

with $\mathrm{d}W_s(t)\mathrm{d}W_v(t) = \rho_{s,v}\mathrm{d}t$ and

$$\sigma^2(t,K) = \frac{\sigma^2_{\mathrm{LV}}(t,K)}{\mathrm{E}\left[V(t)|\,S(t)=K\right]}. \tag{9}$$

Similar dynamics are presented in the paper of Jex et al. (Jex *et al.* 1999). In order to simulate the Heston-SLV model, we use an *adapted version* of the *Quadratic Exponential (QE) scheme* introduced in Andersen (2008), which we derive in Van der Stoep *et al.* (2014). The difference between the original and the adapted version lies in the fact that only the latter incorporates the local volatility component. Let $i = 0, 1, \ldots, N$ and $j = 1, 2, \ldots, M$ indicate the time-step and path, respectively. Defining $\Delta := T/N$ as the time-step size, the discretization scheme for $V(\cdot)$ and $X(\cdot) := \log(S(\cdot))$ reads

$$v_{i+1,j} \sim c(\Delta)\chi^2(d, \lambda(t_i, v_{i,j})),\ v_{0,j} = V_0 \tag{10}$$

$$x_{i+1,j} = x_{i,j} + r\Delta - \frac{1}{2}\widehat{\sigma}^2(t_i, x_{i,j})v_{i,j}\Delta + \frac{\rho_{s,v}}{\gamma}\widehat{\sigma}(t_i, x_{i,j})\left(v_{i+1,j} - \kappa\bar{v}\Delta + v_{i,j}c_1\right) \tag{11}$$

$$+\rho_1\sqrt{\widehat{\sigma}^2(t_i, x_{i,j})v_{i,j}\Delta}Z,\ x_{0,j} = \log(S_0) \tag{12}$$

with $Z \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$, $\rho_1 := (1-\rho^2_{s,v})^{1/2}$, $c_1 := \kappa\Delta - 1$ and

$$c(\Delta) := \frac{\gamma^2}{4\kappa}(1-\mathrm{e}^{-\kappa\Delta}),\ d := \frac{4\kappa\bar{V}}{\gamma^2},\ \lambda(t, V(t)) := \frac{4\kappa\mathrm{e}^{-\kappa\Delta}}{\gamma^2(1-\mathrm{e}^{-\kappa\Delta})}V(t), \tag{13}$$

where $\chi^2(d, \lambda(t, V(t)))$ represents a noncentral chi-squared distribution with d degrees of freedom and non-centrality parameter $\lambda(t, V(t))$. Further, the local volatility component reads

$$\widehat{\sigma}^2(t_i, x_{i,j}) \overset{\mathrm{def}}{=} \sigma^2(t_i, e^{x_{i,j}}) = \frac{\sigma^2_{\mathrm{LV}}(t_i, s_{i,j})}{\mathrm{E}\left[V(t_i)|S(t_i)=s_{i,j}\right]}. \tag{14}$$

Numerical comparisons between the Euler and the *original* QE scheme have been provided in the literature (Andersen 2008). In Van der Stoep *et al.* (2014) we numerically demonstrate that the *adapted* QE scheme outperforms the standard Euler scheme: it yields a higher accuracy and a faster convergence to the reference for a decaying time-step size.

Equation (14) makes clear that in a Monte Carlo simulation framework, for both the SABR-LV model and the Heston-SLV model, we need to evaluate the conditional expectation *for each path, at each time-step*. A closed-form representation does not exist, as the joint distribution of $S(\cdot)$ and $V(\cdot)$ is unknown. We require the evaluation to be *efficient* and *accurate* – if it is not, the error introduced accumulates in the simulation and the results are biased. The principle of stochastic collocation (Babuška *et al.* 2007, Xiu and Hesthaven 2005), discussed in Section 1.4, allows for an evaluation that satisfies both requirements.

2.1. *Establishing* $\mathrm{E}\left[\boldsymbol{V(t)|S(t)=K}\right]$

In this section we evaluate the conditional expectation of interest $\mathrm{E}\left[V(t)|S(t)=K\right]$, which is present in both the SABR-LV model (4)-(5) and the Heston-SLV model (7)-(8). Our approach essentially consists of two projection steps. We first project $V(\cdot)$ and $S(\cdot)$ on standard normal random variables, where, by means of stochastic collocation, $\mathrm{E}[V(t)|S(t)=K]$ is decomposed into a *series of conditional expectations*. Secondly, similar as in e.g. Longstaff and Schwartz (2001), Jain and Oosterlee (2015), each of these conditional expectations, which are expressed in terms of

standard normal random variables, is approximated by a projection on a set of basis functions and applying standard regression techniques.

We start by projecting $S(\cdot)$ at a given fixed time t, on a standard normal random variable $X \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$ via the function $g(\cdot)$, defined by

$$g(\cdot) := F^{-1}_{S(t)}(F_X(\cdot)), \tag{15}$$

which ensures

$$S(t) \overset{\mathrm{d}}{=} g(X)$$

and, moreover, for elements $S(t) = s$ and $X = x$:

$$s = g(x). \tag{16}$$

In a similar way we project $V(t)$ on a standard normal random variable $Z \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$:

$$V(t) \overset{\mathrm{d}}{=} h(Z), \;\; h(\cdot) := F^{-1}_{V(t)}(F_Z(\cdot)), \tag{17}$$

which also yields for elements $V(t) = v$ and $Z = z$:

$$v = h(z). \tag{18}$$

By the element-wise equalities (16) and (18), the conditional expectation can be written in terms of X and Z:

$$\mathrm{E}\left[V(t) \middle| S(t) = K\right] = \mathrm{E}\left[h(Z) \middle| g(X) = K\right]. \tag{19}$$

The joint distribution of X and Z is not analytically known. Although X and Z are both normally distributed, the joint distribution of X and Z is *not* bivariate normal[1] – only the reverse holds in general. We therefore cannot evaluate the right-hand side of (19) analytically and we proceed by determining an approximation for it. This is established by approximating the function $h(\cdot)$ by a polynomial $h_{N_V}(\cdot)$ with degree $N_V - 1$, which is obtained by the stochastic collocation method with N_V collocation points. In particular, given the collocation points z_i, that are *a priori* known[2], we compute the corresponding *exact* evaluations of $V(t)$:

$$v_i = h(z_i) = F^{-1}_{V(t)}\left(F_Z\left(z_i\right)\right), \; i = 1, 2, \ldots, N_V. \tag{20}$$

Next, we apply Lagrangian[3] interpolation through the v_i-values. For an arbitrary value $V(t) = v$, it holds that

$$v = h_{N_V}(z) + \epsilon_1(z) := \sum_{i=1}^{N_V} v_i \ell_i\left(z\right) + \epsilon_1(z), \;\; \ell_i(z) := \prod_{k=1, k\neq i}^{N_V} \frac{z - z_k}{z_i - z_k}, \tag{21}$$

[1] A well-known test for multi-variate normality is Mardia's, see Mardia (1974), which is based on multivariate extensions of skewness and kurtosis measures.

[2] We choose the collocation points in an *optimal* way, namely as the zeros of the Hermite polynomials (abscissas of the Gauss-Hermite quadrature) (Abramowitz and Stegun 1972, Grzelak *et al.* 2014).

[3] Choosing the interpolation polynomial in the Lagrange form is well-accepted in the field of uncertainty quantification (when the stochastic collocation method is applied), see e.g. Sankaran and Marsden (2011).

where $\epsilon_1(z)$ denotes the interpolation error corresponding to the particular argument z. By a change of basis we can write the Lagrange polynomial in terms of monomials:

$$h_{N_V}(z) = a_0 + a_1 z + \ldots + a_{N_V-1} z^{N_V-1}, \tag{22}$$

where the coefficients a_0, a_1, …, a_{N_V-1} are obtained by solving a linear system involving a Vandermonde matrix, see Grzelak *et al.* (2014) for more details. Given (22), we approximate the conditional expectation on the right-hand side of (19) as follows:

$$\begin{aligned}
&\mathrm{E}\left[h(Z)\middle|\, g(X) = K\right] \\
&\quad = \mathrm{E}\left[h_{N_V}(Z) + \epsilon_1(Z)\middle|\, g(X) = K\right] \\
&\quad = \mathrm{E}\left[h_{N_V}(Z)\middle|\, g(X) = K\right] + \mathrm{E}\left[\epsilon_1(Z)\middle|\, g(X) = K\right] \\
&\quad = a_0 + a_1 \mathrm{E}\left[Z\middle|\, X = g^{-1}(K)\right] + \ldots + a_{N_V-1}\mathrm{E}\left[Z^{N_V-1}\middle|\, X = g^{-1}(K)\right] + \bar{\epsilon}_1(K),
\end{aligned} \tag{23}$$

with $\bar{\epsilon}_1(K) := \mathrm{E}\left[\epsilon_1(Z)\middle|\, g(X) = K\right]$. The inversions of functions $g(\cdot)$ and $h(\cdot)$, defined in (15) and (17), respectively, are cheap, as both merely consist of (1) the inversion of a standard normal random variable and (2) the evaluation of $F_{S(t)}(\cdot)$ or $F_{V(t)}(\cdot)$. As CDFs are strictly monotonic, the inversions of $g(\cdot)$ and $h(\cdot)$ provide a *bijective* mapping between the original probability space and the new space.

As we mentioned, the joint distribution of X and Z is not analytically known. To approximate the conditional expectations in (23), we assume that we can approximate the conditional expectation $\mathrm{E}\left[Z^p\middle|\, X = x\right]$ in terms of functions of x, the *basis functions* $\psi_{kp}(\cdot)$, $k = 1, 2, \ldots, n$, $p = 1, 2, \ldots, N_V - 1$:

$$\mathrm{E}\left[Z^p\middle|\, X = x\right] = \sum_{k=1}^{n} b_{kp}\psi_{kp}(x) + \epsilon_{2p}. \tag{24}$$

Equation (24) is motivated rigorously by assuming that the conditional expectations in (23) are elements of the L^2-space of square integrable functions. As the L^2-space is a Hilbert space, it possesses a countable orthonormal basis and the conditional expectations (which are deterministic functions) can be expressed as a linear combination of the elements of this basis. A similar idea is used in Longstaff and Schwartz (2001) in the context of valuing American options by simulation, where the value of continuing with the option is expressed as a conditional expectation. We approximate the conditional expectation in (24) by using the first n orthogonal polynomials[1] $\{1,\ x,\ x^2, \ldots\ x^{n-1}\}$ and, similar as in e.g. Longstaff and Schwartz (2001), Jain and Oosterlee (2015), we apply OLS regression to compute the corresponding coefficients, which yields

$$\mathrm{E}\left[Z^p\middle|\, X = g^{-1}(K)\right] = \widehat{\mathrm{E}}\left[Z^p\middle|\, X = g^{-1}(K)\right] + \widehat{\epsilon}_{2p}, \tag{25}$$

with $\widehat{\mathrm{E}}\left[Z^p\middle|\, X = g^{-1}(K)\right] = \widehat{\beta}_{0p} + \widehat{\beta}_{1p} g^{-1}(K) + \widehat{\beta}_{2p}\left(g^{-1}(K)\right)^2 + \ldots + \widehat{\beta}_{n-1,p}\left(g^{-1}(K)\right)^{n-1}$. Combining this result with (23) yields

$$\mathrm{E}\left[V(t)\middle|\, S(t) = K\right] = a_0 + \sum_{p=1}^{N_V-1} a_p \sum_{k=0}^{n-1} \widehat{\beta}_{kp}\left(g^{-1}(K)\right)^k + \bar{\epsilon}_1(K) + \bar{\epsilon}_2, \tag{26}$$

[1] Other, more complex types of basis functions we may use are the Laguerre, Hermite, Legendre, Chebyshev, Gegenbauer and Jacobi polynomials, see e.g. Chapter 22 of Abramowitz and Stegun (1972). In this article we do not consider these basis functions, as the set of simple polynomials $\{1,\ x,\ \ldots, x^{n-1}\}$ already yields highly satisfactory results, see the numerical experiments in Section 2.3.

with $g(\cdot)$ defined in (15), $\overline{\epsilon}_1(K) := \mathrm{E}\left[\epsilon_1(Z)|\, X = g^{-1}(K)\right]$ and $\overline{\epsilon}_2 := \sum_{p=1}^{N_V-1} \widehat{\epsilon}_{2p}$.

A brief analysis of the errors $\overline{\epsilon}_1(\cdot)$ and $\overline{\epsilon}_2$ can be found in Appendix A. In Babuška *et al.* (2007), in an elliptic PDE framework, a rigorous convergence analysis of the stochastic collocation method is provided, where exponential convergence with respect to the number of "Gauss points" is proven. Our numerical experiments in Section 2.3.1 for a base case are in line with this.

Computation of the approximation in (26) is efficient, as it only requires N_V inversions of $F_{V(t)}(\cdot)$, see (20). Determining x and z and the OLS estimates $\widehat{\beta}_{0p},\ \widehat{\beta}_{1p}, \ldots, \widehat{\beta}_{n-1,p}$ does not involve significant computational cost. In the following, we refer to the approach presented in this section as the 'stochastic collocation – regression' or 'SC–R' approach.

In a Monte Carlo simulation framework, we apply the SC-R approach as described in Algorithm 1. In this algorithm $i = 1, 2, \ldots, N$ denotes the time-step and $j = 1, 2, \ldots, M$ indicates the path. For simulating the SABR-LV model we use a standard Euler discretization scheme, whereas for the Heston-SLV model we apply the adapted version of the QE scheme, see equations (10)-(14). After the Monte Carlo simulation described in Algorithm 1 we price European call options based on the obtained values for $S(\cdot)$ at the time of maturity. This results in the model implied volatility values $\overline{\sigma}_{\text{model}}$ displayed in Figures 2.4 and 2.5 and the errors reported in Tables 2 and 3. In the Monte Carlo simulation, it may be necessary to apply one or more of the enhancements we describe in the follow-up section.

for *each time-step* $t_i,\ i = 1, 2, \ldots, N$ **do**
1. Generate M pairs $(s_{i,j}, v_{i,j}),\ j = 1, 2, \ldots, M$ by going forward one time-step in the Euler scheme (SABR-LV model) or the adapted QE scheme (Heston-SLV model).
2. Compute $\mathrm{E}\left[V(t_i)|\, S(t_i) = s_{i,j}\right]$ using the SC-R approach, see equation (26).
3. Establish the local volatility component $\sigma^2(t_i, s_{i,j})$ by equation (6) for the SABR-LV model or equation (9) for the Heston-SLV model – use its value in step 1.

end

4. Price European call options based on the obtained values for $S(\cdot)$ at the time to maturity.

Algorithm 1: Pricing European call options by a Monte Carlo simulation of the SABR-LV and Heston-SLV models, incorporating the SC-R approach (Section 2.1).

Remark 2.1 *In a stochastic local volatility framework, directly applying OLS regression may yield reasonable results as well. However, as we describe in Remark 3.2 of Van der Stoep et al. (2014), non-negativity of the conditional expectation cannot be guaranteed for cases where the Feller condition is violated and improvements must be made. Further, by applying stochastic collocation we can use the analytical expression of the CDF of* $V(\cdot)$ *in order to obtain values for the coefficients* $a_0,\ a_1, \ldots$. *Moreover, in the context of the local volatility model with stochastic interest rates, see Section 3, by projecting* $S(\cdot)$ *on a standard normal random variable we can employ the analytical expression for moments of a truncated standard normal random variable, see Result 3.1.*

2.2. *Enhancements*

In this section we discuss three adaptations to the stochastic collocation – regression method which may enhance the results.

First, we observe that at the boundaries of the X-domain (recall $X := g^{-1}(K) = F_X^{-1}(F_{S(t)}(K))$) the performance of the regression deteriorates due to the presence of a small number of observations, which may yield a significant increase of $\overline{\epsilon}_2$. We therefore set for $K \leq s_{\min}$: $\mathrm{E}\left[V(t)|\, S(t) = K\right] = \mathrm{E}\left[V(t)|\, S(t) \leq s_{\min}\right]$ and for $K \geq s_{\max}$: $\mathrm{E}\left[V(t)|\, S(t) = K\right] = \mathrm{E}\left[V(t)|\, S(t) \geq s_{\max}\right]$, where $s_{\min}$ and $s_{\max}$ are percentiles of the $S(t)$-distribution, i.e. $s_{\min} = F_{S(t)}^{-1}(p_{s,\min})$ and $s_{\max} = F_{S(t)}^{-1}(p_{s,\max})$. Here $0 \leq p_{s,\min} < p_{s,\max} \leq 1$ denote fractions of the total number of Monte Carlo realizations. *In all pricing experiments in Section 2.3.2 we apply this adaptation and choose* $p_{s,\min} = 0.1$ *and* $p_{s,\max} = 0.9$.

The approximation of the expectation in (26) is not guaranteed to be positive. This may be problematic in the case that a significant part of the variance realizations is close to zero, e.g. if the Feller condition in the Heston-SLV model is strongly violated. In this case we may split the conditional expectation in two parts in the following way:

$$\begin{aligned}\mathrm{E}\left[V(t)|\, S(t)=K\right] = \mathrm{E}\left[V(t)|\, S(t)=K, V(t)\leq v^*\right]\mathrm{Q}[V(t)\leq v^*] \\ +\mathrm{E}\left[V(t)|\, S(t)=K, V(t)> v^*\right]\left(1-\mathrm{Q}[V(t)\leq v^*]\right). \end{aligned} \tag{27}$$

The first conditional expectation we approximate by $\mathrm{E}\left[V(t)|\, V(t)\leq v^*\right]$, the second conditional expectation is approximated by the stochastic collocation – regression approach. We can choose v^* to be a fixed value, or based on a fixed percentile p_v^*, i.e. $v^* = F^{-1}_{V(t)}(p_v^*)$. We prefer the latter, as in this case at each time-step in the Monte Carlo simulation we naturally control the fraction of the total number of observations on which we apply the stochastic collocation – regression approach. So we obtain

$$\begin{aligned}\mathrm{E}\left[V(t)|\, S(t)=K\right] &= \left(\mathrm{E}\left[V(t)|\, V(t)\leq v^*\right]+\epsilon_3\right)F_{V(t)}(v^*) \\ &\quad + \left(a_0 + \sum_{p=1}^{N_V-1} a_p \sum_{k=0}^{n-1} \widehat{\beta}_{kp}\left(g^{-1}(K)\right)^k + \bar{\epsilon}_1(K) + \bar{\epsilon}_2\right)\left(1-F_{V(t)}(v^*)\right) \\ &:= \mathcal{V}(K) + \epsilon, \end{aligned} \tag{28}$$

where ϵ denotes the approximation error. By means of this adaptation we leave out the smallest variance realizations when applying the stochastic collocation – regression approach, which makes it less likely that the corresponding SC–R approximation yields negativity. To this approximation we moreover add the positive term $\left(\mathrm{E}\left[V(t)|\, V(t)\leq v^*\right]\right)F_{V(t)}(v^*)$.

Although the former adaptation guarantees non-negativity for $\mathcal{V}(K)$, in extreme cases the fraction of $V(\cdot)$-realizations close to zero is substantial, and we would need to choose a relatively large value for p_v^* to ensure non-negativity of $\mathcal{V}(K)$. This would make the approximation for the conditional expectation inaccurate, as in this case it is for a large part determined by the naive approximation $\mathrm{E}\left[V(t)|\, V(t)\leq v^*\right]F_{V(t)}(v^*)$. Therefore, in the case that the approximation $\mathcal{V}(K)$ still yields negative values for an appropriate value of p_v* (in our numerical experiments we choose p_v^* in the range $0.01-0.1$), we apply another correction, namely

$$\mathrm{E}\left[V(t)|\, S(t)=K\right] = \mathcal{V}(K) + \epsilon - \min_K\left\{0, (1+\delta)\mathcal{V}(K)\right\},$$

with $0<\delta<1$. This correction is interpreted as follows: in the case that a part of $\mathcal{V}(K)$ is negative, we apply a vertical "shift" such that it becomes positive. If $\mathcal{V}(K)$ is completely non-negative, the vertical shift is zero[1]. This correction *guarantees* non-negativity of the approximation of the conditional expectation.

2.3. *Numerical experiments*

In this section we test the accuracy of the approximation of the conditional expectation in (26). We first test the method for a base case where an analytical reference value is available. Subsequently, we consider the SABR-LV and Heston-SLV models in a Monte Carlo simulation framework. In particular, given a pre-specified market, we add to an either poorly or satisfactorily calibrated

[1]Numerical experiments demonstrate that merely applying the third correction, i.e. applying a vertical shift, typically yields worse pricing results compared to *combining the second and third corrections* mentioned in Section 2.2.

'pure' SABR or Heston model the local volatility component, consisting of Dupire's local volatility and the conditional expectation approximation (26),

2.3.1. **The 2D-GBM model: a base case.** We start with testing the approximation of the conditional expectation (26) for a model which is given by two correlated Geometric Brownian Motions (GBMs):

$$\mathrm{d}Y_1(t) = \sigma_1 Y_1(t)\mathrm{d}W_1(t),\ \mathrm{d}Y_2(t) = \sigma_2 Y_2(t)\mathrm{d}W_2(t),\ Y_1(0) = y_{10},\ Y_2(0) = y_{20}, \tag{29}$$

with $\mathrm{d}W_1(t)\mathrm{d}W_2(t) = \rho \mathrm{d}t$. The expectation of $Y_2(t)$ conditional on the event $Y_1(t) = y_1$ is (An and Li 2015)

$$\mathrm{E}\left[Y_2(t)|Y_1(t) = y_1\right] = y_{20}\left(\frac{y_1}{y_{10}}\right)^{\rho\frac{\sigma_2}{\sigma_1}} \mathrm{e}^{t\left(\frac{1}{2}\rho\sigma_1\sigma_2 - \frac{1}{2}\sigma_2^2\rho^2\right)}. \tag{30}$$

Let[1] $y_{10} = 1$, $y_{20} = 0.05$, $\rho = -0.5$ and $t = 5$. As a first experiment, suppose we choose $N_{Y_2} = 6$ collocation points and $n = 7$ basis functions. In Figure 2.1 we compare the reference (30) and the approximation (26) obtained by the stochastic collocation – regression (SC–R) approach for a moderate case (left, $\sigma_1 = \sigma_2 = 0.3$) and a more extreme case (right, $\sigma_1 = \sigma_2 = 0.9$). An excellent fit is obtained.

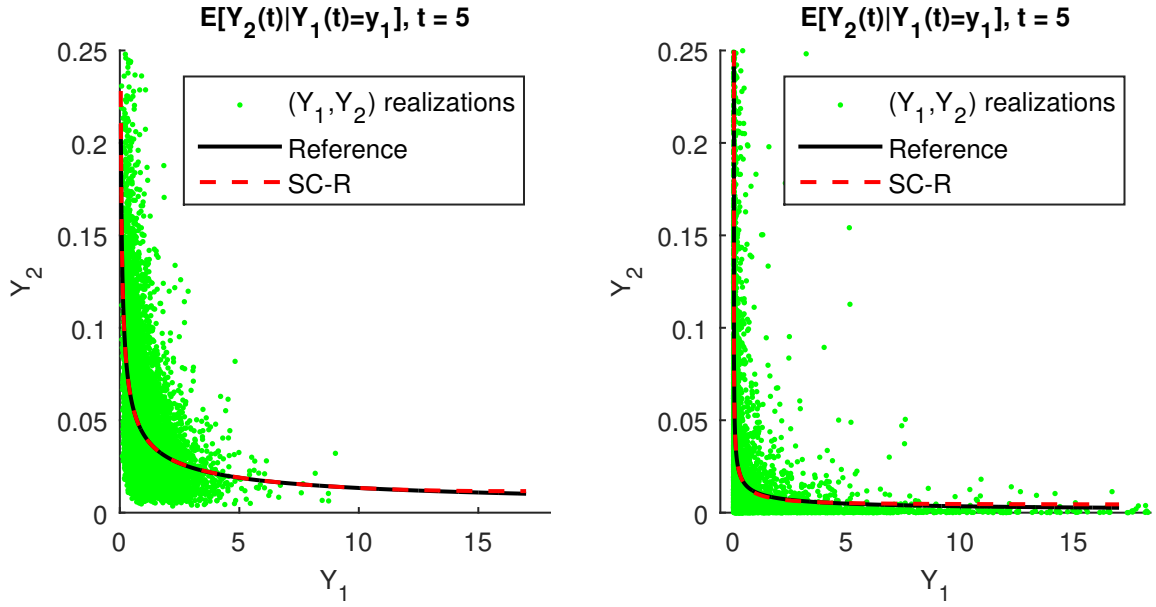

Figure 2.1. Comparison of the reference (30) and the SC–R approximation (26) for a moderate case (left) and a more extreme case (right).

The reference (30) allows for a numerical analysis of the errors $\bar{\epsilon}_1(\cdot)$ and $\bar{\epsilon}_2$, which are introduced by the stochastic collocation method and the regression step, respectively. We choose the parameter values just mentioned and $\sigma_1 = \sigma_2 = 0.3$. We make use of the result in the following lemma.

LEMMA 2.1 *Given the two-dimensional model (29). Let X and Z denote standard normal random variables and assume for an arbitrary t that the elements $Y_i(t) = y_i$, $i = 1, 2$, $X = x$, $Z = z$ are related by $y_1 = g(x)$, $y_2 = h(z)$, with $g(\cdot)$ and $h(\cdot)$ defined in (15) and (17), respectively. This implies that X and Z are jointly bivariate normally distributed.*

[1] In a stochastic volatility model these parameter values are representative choices for S_0, V_0 and $\rho_{s,v}$, respectively.

Proof. For a proof of Lemma 2.1, see Appendix B. □

Recall the error due to the stochastic collocation method: $\bar{\epsilon}_1(K) := \mathrm{E}\left[\epsilon_1(Z)|\, g(X) = K\right]$. In a Monte Carlo simulation framework, for a fixed t, let $\bar{\epsilon}_1(y_{1j})$ denote the error corresponding to the jth realization of $Y_1(t)$, y_{1j}. From (23) and (30) it follows that it is given by

$$\bar{\epsilon}_1(y_{1j}) = y_{20}\left(\frac{y_{1j}}{y_{10}}\right)^{\rho\frac{\sigma_2}{\sigma_1}} \mathrm{e}^{t\left(\frac{1}{2}\rho\sigma_1\sigma_2 - \frac{1}{2}\sigma_2^2\rho^2\right)} - f(y_{1j}),$$

with

$$f(y_{1j}) := a_0 + a_1\mathrm{E}\left[Z|\, X = g^{-1}\left(y_{1j}\right)\right] + \ldots + a_{N_{Y_2}-1}\mathrm{E}\left[Z^{N_{Y_2}-1}\middle|\, X = g^{-1}\left(y_{1j}\right)\right]. \tag{31}$$

As Z and X are jointly bivariate normally distributed, see Lemma 2.1, we are able to evaluate each conditional expectation in (31) analytically. For arbitrary $p \in \{1, 2, \ldots, N_{Y_2} - 1\}$, applying the Cholesky decomposition, straightforward calculus yields

$$\mathrm{E}\left[Z^p|\, X = g^{-1}(y_{1j})\right] = \sum_{k=0}^{p}\binom{p}{k}\rho^{p-k}\left(1-\rho^2\right)^{\frac{k}{2}}\left(g^{-1}(y_{1j})\right)^{p-k}\mu_k, \tag{32}$$

with $\mu_k = (k-1)!!$ if k is even and $\mu_k = 0$ if k is odd. The double exclamation marks stand for the "double factorial". For an even integer $n > 0$ it is defined as $n!! = n\cdot(n-2)\cdot(n-4)\ldots 6\cdot 4\cdot 2$ and for an odd integer $n > 0$ it is $n!! = n\cdot(n-2)\cdot(n-4)\ldots 5\cdot 3\cdot 1$ and, by an extension, $-1!! = 1$. Further, by definition, $0!! = 1$.

Given (32), for different N_{Y_2} values we compute $\mathcal{E}_1 := \log\left(\frac{1}{M}\sum_{j=1}^{M}|\bar{\epsilon}_1(y_{1j})|\right)$, where M denotes the total number of observations. In Figure 2.2 on the left-hand side $\mathcal{E}_1$ is displayed against the number of collocation points. An exponential convergence is observed, which is in line with Babuška *et al.* (2007), where in an elliptic PDE framework a rigorous proof for exponential convergence of the stochastic collocation method is provided. The error does not decrease further for $N_{Y_2} > 14$, as machine precision has been reached ($\exp(-36) \approx 2\cdot 10^{-16}$).

We proceed with analyzing $\bar{\epsilon}_2 := \sum_{p=1}^{N_{Y_2}-1}\widehat{\epsilon}_{2p}$, the error due to regression. Define $\bar{\epsilon}_{2j} := \sum_{p=1}^{N_{Y_2}-1}\widehat{\epsilon}_{2pj} = f(y_{1j}) - \widehat{f}(y_{1j}),\quad j = 1, 2, \ldots, M$, with $f(\cdot)$ given by (31), where the conditional expectations $\mathrm{E}\left[Z^p|\, X = g^{-1}(y_{1j})\right],\ p = 1, 2, \ldots, N_{Y_2} - 1$ are evaluated by the analytical formula (32), and $\widehat{f}(\cdot)$ denotes

$$\widehat{f}(y_{1j}) := a_0 + a_1\widehat{\mathrm{E}}\left[Z|\, X = g^{-1}\left(y_{1j}\right)\right] + \ldots + a_{N_{Y_2}-1}\widehat{\mathrm{E}}\left[Z^{N_{Y_2}-1}\middle|\, X = g^{-1}\left(y_{1j}\right)\right], \tag{33}$$

where $\widehat{\mathrm{E}}\left[Z^p|\, X = g^{-1}(y_{1j})\right],\ p = 1, 2, \ldots, N_{Y_2} - 1$ is obtained by OLS regression. We consider the logarithm of the mean squared error: $\mathcal{E}_2 := \log\left(\frac{1}{M}\sum_{j=1}^{M}\bar{\epsilon}_{2j}^2\right)$. We observe for $\mathcal{E}_2$ a convergence of order $\mathcal{O}(-\log(M))$, see the plot in the middle of Figure 2.2, where we consider $M = 1\cdot 10^3, 5\cdot 10^3, 1\cdot 10^4, 5\cdot 10^4, 1\cdot 10^5, 5\cdot 10^5, 1\cdot 10^6, 5\cdot 10^6, 1\cdot 10^7$.

Last, we study the dependence of $\mathcal{E}_2$ on the number of basis functions, see the right-hand plot of Figure 2.2, where we consider $n = 1, 2, \ldots, 12$ basis functions. We observe that for $n = 5$ the smallest error is achieved. For $n > 5$ the increase in $\mathcal{E}_2$ is due to overfitting, where oscillations in the approximation of the conditional expectation may occur.

In practice, for the Heston-SLV and SABR-LV models we typically choose $4 - 6$ collocation points; our numerical experiments confirm that with this number of collocation points sufficiently accurate results are obtained. In general, we choose the number of basis functions n in the range $5 - 9$, depending on how extreme the parameters of the calibrated 'pure' Heston or SABR model

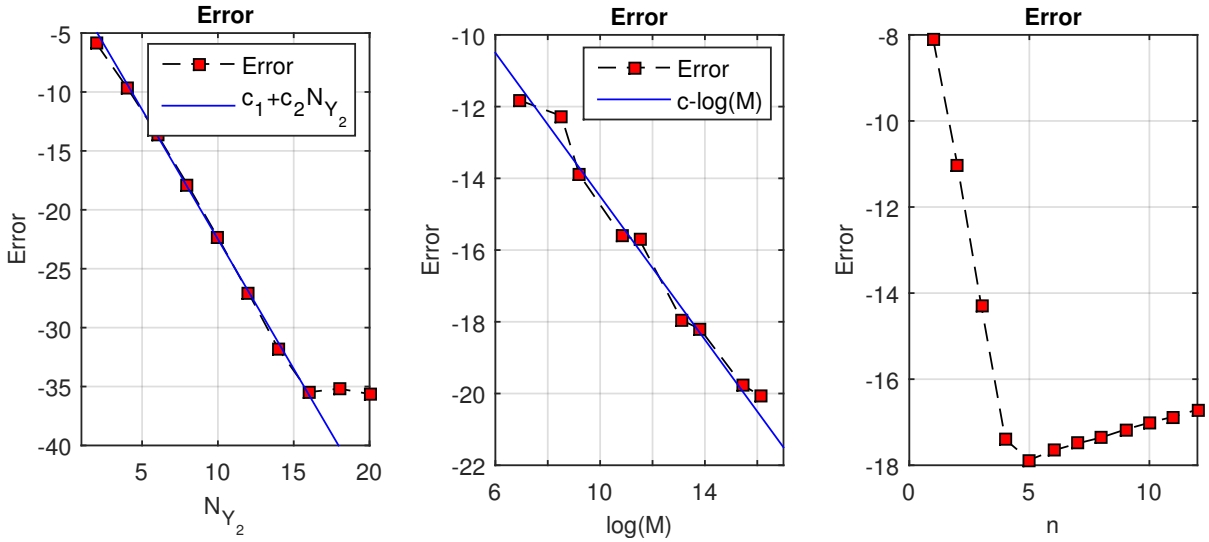

Figure 2.2. Left: the relation between the error $\mathcal{E}_1 = \log\left(\frac{1}{M}\sum_{j=1}^{M}|\overline{\epsilon}_1(y_{1j})|\right)$ and the number of collocation points N_{Y_2} ($M = 1 \cdot 10^6$). Middle: the relation between the error $\mathcal{E}_2 = \log\left(\frac{1}{M}\sum_{j=1}^{M}\overline{\epsilon}_{2j}^2\right)$ and the logarithm of the number of realizations M ($N_{Y_2} = 14$, $n = 5$). Right: the relation between the error $\mathcal{E}_2$ and the number of basis functions n ($N_{Y_2} = 14$, $M = 1 \cdot 10^6$).

are. In the follow-up section we consider the performance of the stochastic collocation – regression approach for the Heston-SLV and SABR-LV models in more detail.

2.3.2. **The SABR-LV and Heston-SLV models.** In this section we test the performance of the stochastic collocation – regression approach for the SABR-LV model (4)-(5) and the Heston-SLV model (7)-(8). Stochastic local volatility models, that are considered as the standard for pricing in an FX context, combine desirable features of a stochastic volatility model, e.g. preserving the shape of the forward volatility smile and reflecting more realistic smile dynamics, and the local volatility model, namely a perfect calibration to arbitrage-free European plain vanilla options.

We first consider the SABR-LV model. Let $S_0 = 1$, $V_0 = 0.05$, $\beta = 0.5$, $\gamma = 0.5$, $\rho = -0.5$ and $t = 2$. We generate for this parameter set (S, V)-realizations by simulating the 'pure' SABR model. Given the realizations at $t = 2$, for different numbers of basis functions we compare approximations obtained by the stochastic collocation – regression approach (26) and the non-parametric approach (Van der Stoep *et al.* 2014), using 10 bins, which serves as a reference. We consider $N_V = 4$ and $n = 3, 5, 7$. Results are displayed in Figure 2.3. For $n > 7$ no significant increase in accuracy was observed. For the Heston-SLV model we expect a similar increase in accuracy of the stochastic collocation – regression approach.

To assess whether for a given number of collocation points N_V and basis functions n the stochastic collocation – regression approach performs sufficiently accurate, we perform pricing experiments. In particular, given *a priori* specified market implied volatilities, we price European call options by a 'pure' Heston or SABR model[1] and the Heston-SLV ('H-SLV') and SABR-LV model, respectively. By definition of stochastic local volatility, the Heston-SLV and SABR-LV models should yield implied volatilities that *perfectly* match the ones corresponding to the market. At each time-step in the Monte Carlo simulation[2] we establish the conditional expectation according to (26). We generate synthetic market prices by the Heston model, which we assume to be calibrated perfectly to the market. For this we choose some parameter sets from Clark (2011) which may be encountered

[1]Implied volatilities for the 'pure' Heston model are obtained based on Fourier techniques.
[2]The Monte Carlo simulation consists of $2 \cdot 10^5$ paths (20 seeds, each seed constitutes 10^4 paths) and 200 time-steps per year, unless otherwise mentioned.

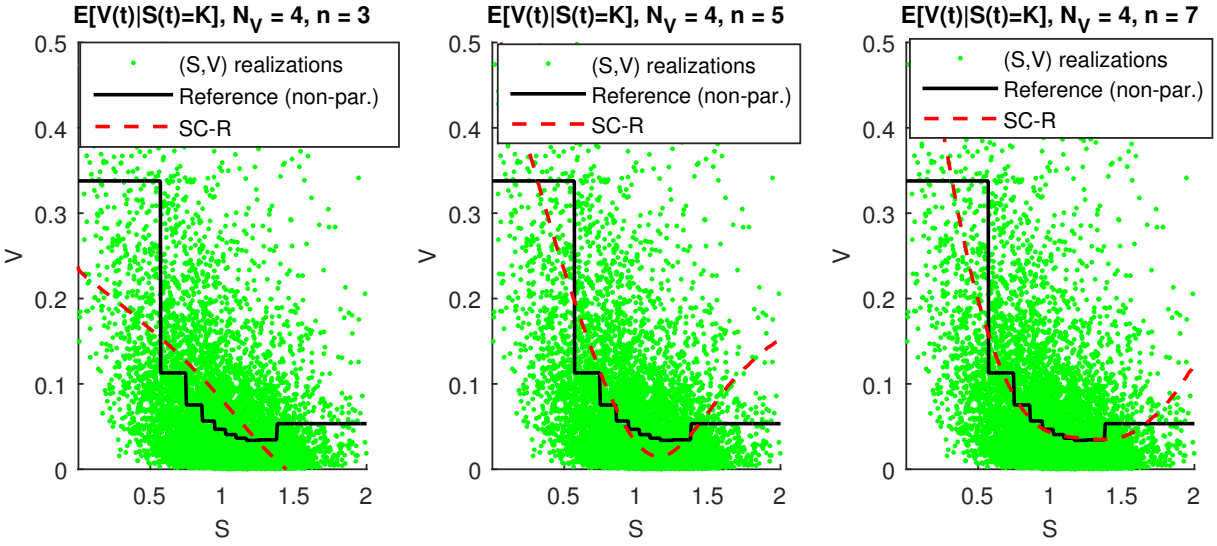

Figure 2.3. The conditional expectation approximation (26) obtained by the stochastic collocation – regression approach for $n = 3$ (left), $n = 5$ (middle) and $n = 7$ (right) compared to the non-parametric method (Van der Stoep *et al.* 2014). Number of collocation points is $N_V = 4$.

in typical FX markets, see Table 1 (market data as of 16 September 2008).

Given the market data, we assume a both satisfactorily and poorly calibrated Heston model (parameter values are 20% and 80% off, respectively). On top of this model we add the local volatility component to compensate for the calibration error. For each set in Table 1, we consider the implied volatilities corresponding to (1) the market (the perfectly calibrated Heston model with parameters given in Table 1), (2) the (satisfactorily or poorly) calibrated 'pure' Heston model, (3) the Heston-SLV model and (4) the traditional local volatility model. Given the expiry T, similar as in Piterbarg (2006) we consider the strikes $K_i = \exp(0.1\delta_i\sqrt{T})$ (we choose $S_0 = 1$ and zero interest rate), with $\delta_i = -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5$. We simulate the Heston-SLV model according to the *adapted QE scheme* given by equations (10)-(14). Implied volatilities for the 'pure' Heston model are obtained by a standard Fourier pricing technique.

For all cases we choose $N_V = 6$. For 'market' Sets 1 and 2 we choose $n = 5$ basis functions. For Set 3 $n = 7$ for the poorly calibrated case and $n = 15$ for the satisfactorily calibrated case. For Set 4 these numbers are $n = 5$ and $n = 9$, respectively. As mentioned earlier, in all pricing experiments we apply the first adaptation to the method described in Section 2.2. Further, for Sets 3 and 4, the satisfactorily calibrated case, we make use of the second and third adaptation specified in Section 2.2; we choose $p_v^* = 0.01$ and $\delta = 0.1$. For Sets 1 and 3 the results are provided by Tables 2 and 3, respectively. Given the standard deviations, we observe that for both the local volatility model and the Heston-SLV model for all strikes the reference is within the 95%-confidence interval[1]. A lower standard deviation can be obtained by increasing the number of Monte Carlo paths[2]. For Set 4 we report the implied volatilities in Figure 2.4. The results for Set 2 are essentially the same as these for Set 1 and therefore, to save some space, they are not presented.

We proceed with similar pricing experiments for the SABR-LV model; we consider the cases where the 'pure' SABR model is satisfactorily and poorly calibrated to the market data (generated by the Heston model with the parameters as specified in Table 1). Contrary to the Heston-SLV case, we report the results for Set 3 in Figure 2.5 and for Sets 2 and 4 in Tables 4 and 5, respectively.

[1]The boundaries of the 95%-confidence interval are $\mu(\overline{\sigma}_{1,\text{model}}, \overline{\sigma}_{2,\text{model}}, \overline{\sigma}_{3,\text{model}}, \ldots) \pm 1.96 \cdot \sigma(\overline{\sigma}_{1,\text{model}}, \overline{\sigma}_{2,\text{model}}, \overline{\sigma}_{3,\text{model}}, \ldots)$, with $\mu(\cdot)$ and $\sigma(\cdot)$ denoting the mean and standard deviation, respectively, and $\overline{\sigma}_{i,\text{model}}$ stands for the model implied volatility (obtained from Monte Carlo) corresponding to the ith seed.

[2]E.g., when repeating the experiment for the Heston-SLV model ($N_V = 6$, $n = 5$), given 'Heston market' Set 1, with 20 seeds, $5 \cdot 10^5$ paths per seed, we obtain the errors 0.02, 0.00, 0.01, 0.01, 0.00, 0.03, 0.01 and corresponding standard deviations 0.02, 0.01, 0.01, 0.01, 0.01, 0.01, 0.02.

Set	Ccypair	T	V_0	$\rho_{s,v}$	γ	κ	$\overline{V}$
1	EURGBP	1	0.01	0.23	0.21	1.50	0.01
2	EURUSD	2	0.02	−0.14	0.20	0.75	0.02
3	AUDJPY	3	0.07	−0.54	0.93	0.50	0.07
4	USDJPY	5	0.02	−0.71	0.39	0.30	0.02

Table 1. Heston parameters in typical FX markets (market data as of 16 September 2008, see Clark (2011)).

'Heston market' Set 1						
	Sat. calibrated			Poorly calibrated		
K	ϵ_{H}	$\epsilon_{\text{H-SLV}}$	ϵ_{LV}	ϵ_{H}	$\epsilon_{\text{H-SLV}}$	ϵ_{LV}
0.86	0.32	0.04 (0.09)	0.04 (0.11)	1.48	0.01 (0.11)	0.02 (0.09)
0.90	0.10	0.04 (0.08)	0.05 (0.09)	1.00	0.01 (0.09)	0.01 (0.08)
0.95	0.09	0.05 (0.08)	0.06 (0.09)	0.72	0.01 (0.09)	0.01 (0.09)
1.00	0.15	0.04 (0.09)	0.04 (0.09)	0.77	0.01 (0.11)	0.01 (0.10)
1.05	0.08	0.01 (0.10)	0.03 (0.10)	1.13	0.02 (0.12)	0.02 (0.11)
1.11	0.06	0.03 (0.13)	0.03 (0.12)	1.64	0.01 (0.12)	0.02 (0.15)
1.16	0.22	0.00 (0.17)	0.02 (0.14)	2.20	0.01 (0.15)	0.03 (0.19)

Table 2. Errors $\epsilon_{\text{model}} := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{model}}|$ in % corresponding to 'Heston market' Set 1. 'Sat.' stands for the satisfactorily calibrated Heston model and $\overline{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

'Heston market' Set 3						
	Sat. calibrated			Poorly calibrated		
K	ϵ_{H}	$\epsilon_{\text{H-SLV}}$	ϵ_{LV}	ϵ_{H}	$\epsilon_{\text{H-SLV}}$	ϵ_{LV}
0.77	1.14	0.01 (0.20)	0.04 (0.18)	1.09	0.13 (0.18)	0.05 (0.20)
0.84	1.26	0.05 (0.18)	0.04 (0.17)	2.17	0.09 (0.18)	0.05 (0.18)
0.92	1.35	0.07 (0.18)	0.04 (0.18)	3.19	0.06 (0.19)	0.04 (0.16)
1.00	1.33	0.11 (0.19)	0.04 (0.19)	4.01	0.03 (0.19)	0.03 (0.15)
1.09	1.03	0.14 (0.22)	0.01 (0.21)	4.30	0.01 (0.20)	0.02 (0.17)
1.19	0.34	0.07 (0.28)	0.01 (0.25)	3.66	0.05 (0.21)	0.02 (0.22)
1.30	0.44	0.10 (0.35)	0.01 (0.30)	2.18	0.06 (0.24)	0.04 (0.27)

Table 3. Errors $\epsilon_{\text{model}} := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{model}}|$ in % corresponding to 'Heston market' Set 3. 'Sat.' stands for the satisfactorily calibrated Heston model and $\overline{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

Given the standard deviations, we observe that for the SABR-LV model for all strikes – except for $K = 0.72$ and $K = 0.80$ for 'Heston market' Set 4 – the reference is within the 95%-confidence interval. Further, for the standard SABR model, only for 'Heston market' Set 2, the satisfactorily calibrated case, the reference is within the 95%-confidence interval (see footnote 1). Here we leave out the highly accurate results corresponding to Set 1, to save some space. All results are obtained with $N_V = 4$, $n = 5$ (Sets 1, 2) and $n = 7$ basis functions (Sets 3, 4) – these numbers correspond to the first experiment in this section. In all pricing experiments we apply the first adaptation to the method described in Section 2.2. Both the second and the third correction mentioned in Section 2.2 is *not* used. However, for Sets 3 and 4, in the calibration of the 'pure' SABR model we include a constraint on the vol-vol parameter, see Remark 2.2.

REMARK 2.2 (Limitations of the SC-R approach) *Considering expiries up to* 6 *years, for the Heston-SLV model the stochastic collocation – regression approach yields highly accurate results for the calibrated 'pure' Heston parameters that satisfy* $F := \frac{2\kappa\bar{V}}{\gamma^2} - 1 \geq -0.8$, *regardless of the 'Heston market' we assume. For extreme cases for which* $F \approx -0.8$ *we typically choose* $N_V = 6$, *the number of basis functions* n *in the range* $7 - 9$ *and make use of all enhancements described in Section 2.2, with* p_v^* *in the range* $0.01 - 0.1$ *and* $\delta = 0.1$. *Trivially, in the calibration one can control to which extent the Feller condition is violated by imposing constraints on the parameters, such that the stochastic collocation – regression approach works without the enhancements of Section 2.2 and for lower numbers of collocation points and basis functions. For our approach to work for*

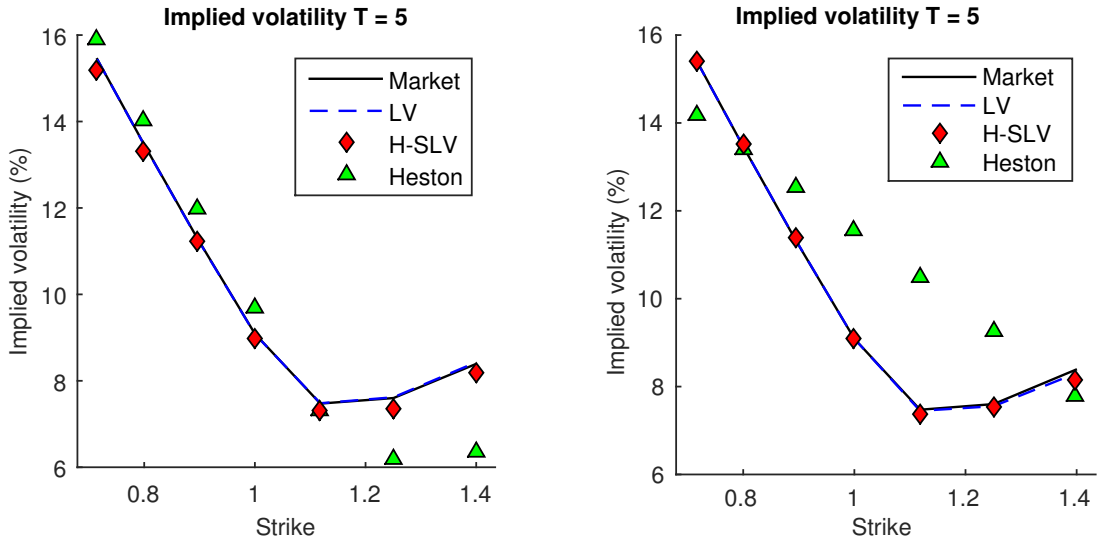

Figure 2.4. Black-Scholes implied volatilities corresponding to Set 4 with the satisfactorily (left) and poorly (right) calibrated Heston model. 'H-SLV' stands for the Heston-Stochastic Local Volatility model. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

'Heston market' Set 2				
	Sat. calibrated		Poorly calibrated	
K	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$
0.81	0.02 (0.23)	0.01 (0.25)	2.93 (0.21)	0.04 (0.25)
0.87	0.03 (0.18)	0.04 (0.19)	2.87 (0.14)	0.05 (0.19)
0.93	0.04 (0.14)	0.04 (0.16)	2.79 (0.11)	0.04 (0.16)
1.00	0.03 (0.13)	0.02 (0.14)	2.73 (0.10)	0.03 (0.15)
1.07	0.04 (0.14)	0.00 (0.15)	2.72 (0.11)	0.01 (0.16)
1.15	0.08 (0.14)	0.05 (0.16)	2.78 (0.12)	0.03 (0.17)
1.24	0.15 (0.17)	0.16 (0.20)	2.90 (0.14)	0.11 (0.21)

Table 4. Errors $\epsilon_{\text{model}} := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{model}}|$ in % corresponding to 'Heston market' Set 2. 'Sat.' stands for the satisfactorily calibrated Heston model and $\overline{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

'Heston market' Set 4				
	Sat. calibrated		Poorly calibrated	
K	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$	ϵ_{SABR}	$\epsilon_{\text{SABR-LV}}$
0.72	1.08 (0.21)	0.54 (0.22)	2.44 (0.22)	0.44 (0.21)
0.80	0.78 (0.14)	0.32 (0.13)	1.98 (0.14)	0.25 (0.13)
0.89	0.32 (0.09)	0.11 (0.08)	1.33 (0.08)	0.10 (0.08)
1.00	0.35 (0.06)	0.10 (0.06)	0.46 (0.05)	0.03 (0.06)
1.12	0.73 (0.06)	0.10 (0.08)	0.14 (0.06)	0.04 (0.07)
1.25	0.10 (0.09)	0.01 (0.09)	0.32 (0.08)	0.00 (0.09)
1.40	0.44 (0.13)	0.11 (0.14)	0.82 (0.13)	0.09 (0.14)

Table 5. Errors $\epsilon_{\text{model}} := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{model}}|$ in % corresponding to 'Heston market' Set 4. 'Sat.' stands for the satisfactorily calibrated Heston model and $\overline{\sigma}$ denotes the Black-Scholes implied volatility. Numbers in parentheses are standard deviations over the seeds.

the SABR-LV model, for 'Heston market' Sets 3 and 4 we need to impose in the calibration of the 'pure' SABR model the constraints $\gamma < 0.55$ and $\gamma < 0.4$, respectively, which seems very reasonable in practice.

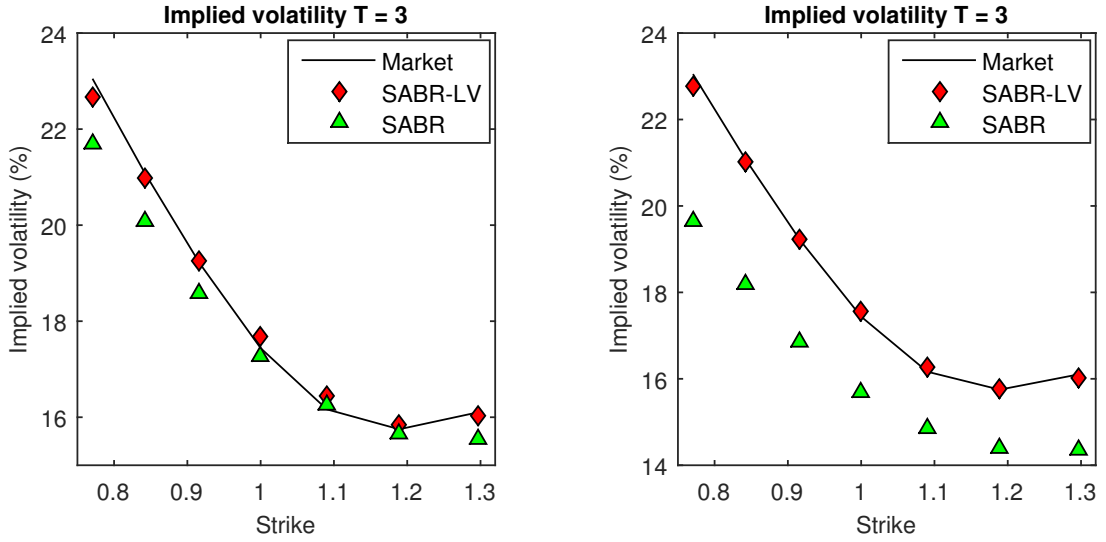

Figure 2.5. Black-Scholes implied volatilities corresponding to Set 3 with the satisfactorily (left) and poorly (right) calibrated SABR model. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

3. Local Volatility Model with Stochastic Rates

In this section we present an evaluation approach for the local volatility model extended with stochastic interest rates. *The method is similar as the one employed in a stochastic local volatility context*; based on the fact that cumulative distribution functions are equally distributed, we first project $S(\cdot)$ on a standard normal random variable. Subsequently we apply regression to approximate a conditional expectation.

As pointed out in Deelstra and Rayée (2012), in the long-dated FX options market the effect of interest rate volatility becomes increasingly relevant for a longer expiry and may become as important as that of the FX spot volatility. Further, also in an FX context, Piterbarg (Piterbarg 2006) considers for the pricing of *Power-Reverse Dual-Currency* (PRDC) swaps, the local volatility model incorporating stochastic interest rates, assuming that the domestic and foreign interest rates follow Hull-White dynamics. He states that FX options exhibit a significant volatility skew and, moreover, that PRDC swaps, due to their structure, are highly sensitive to it. Therefore, the assumption of lognormality of the FX rates in the standard three-factor pricing model is not appropriate to price and hedge long-dated FX products. Further, as pointed out in Benhamou *et al.* (2012), long-term callable path-dependent equity options require an appropriate modelling of the underlying asset process and, moreover, the early-exercise feature – in particular for a large time-span – suggests interest rates risk. Enhancing the local volatility model with stochastic interest rates is also the subject of research in Atlan (2006), Ren *et al.* (2007), Guyon and Henry-Labordère (2012), amongst others.

Similar as in e.g. Piterbarg (2006), Deelstra and Rayée (2012), Benhamou *et al.* (2012), let the interest rate be governed by Hull-White dynamics, which is also the case in the traditional three-factor pricing model. Under the risk-neutral Q-measure, the dynamics of the *Local Volatility-Hull White (LV-HW) model* are given by the following system of equations (see e.g. Atlan (2006)):

$$\mathrm{d}S(t)/S(t) = r(t)\mathrm{d}t + \sigma(t, S(t))\mathrm{d}W_s^{\mathrm{Q}}(t), \tag{34}$$

$$\mathrm{d}r(t) = \lambda\left(\theta(t) - r(t)\right)\mathrm{d}t + \eta\mathrm{d}W_r^{\mathrm{Q}}(t), \tag{35}$$

with $\mathrm{d}W_s^{\mathrm{Q}}(t)\mathrm{d}W_r^{\mathrm{Q}}(t) = \rho_{r,s}\mathrm{d}t$. In the interest rate process the speed of mean reversion λ and

the volatility coefficient η are related with the time-dependent term structure function $\theta(\cdot)$ via[1] $\theta(t) = f(0,t) + \frac{1}{\lambda}\frac{\partial}{\partial t}f(0,t) + \frac{\eta^2}{2\lambda^2}\left(1 - \mathrm{e}^{-2\lambda t}\right)$, which yields a model fit with the initial yield curve, where $f(0,t)$ denotes the initial instantaneous forward rate corresponding to expiry t, defined by $f(0,t) = -\partial \log\left(P(0,t)\right)/\partial t$ and $P(0,t)$ is the current value of the zero-coupon bond, which is given by

$$P(0,t) = \exp\left(-\int_0^t \psi(s)\mathrm{d}s + A(t)\right), \tag{36}$$

with $\psi(t) = r_0\mathrm{e}^{-\lambda t} + \lambda\int_0^t \theta(s)\mathrm{e}^{-\lambda(t-s)}\mathrm{d}s$ and $A(t) = \frac{\eta^2}{2\lambda^3}\left(\lambda t - 2(1-\mathrm{e}^{-\lambda t}) + \frac{1}{2}\left(1-\mathrm{e}^{-2\lambda t}\right)\right)$.

From the expression for the instantaneous forward rate, the initial interest rate r_0 is implied by the identity $r(t) = f(t,t)$. Further, the local volatility component reads

$$\sigma^2(t,K) = \frac{\frac{\partial C(t,K)}{\partial t} - K\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\middle|\mathcal{F}(t_0)\right]}{\frac{1}{2}K^2\frac{\partial^2 C(t,K)}{\partial K^2}}, \tag{37}$$

where $M(t)$ denotes the value of the moneyness account, determined by $\mathrm{d}M(t) = r(t)M(t)\mathrm{d}t$. As we always consider $t_0 = 0$, we leave out the filtration for notational purposes from now on.

In the local volatility component the expectation $\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right]$ is problematic in a calibration sense, as no direct link with the market quotes can be observed (Deelstra and Rayée 2012). Also, no analytical expressions for the joint distribution of $r(t)/M(t)$ and $S(t)$ are available. Further, the discretization scheme suggests that *for each time-step* in the simulation, the expectation, which in principle is a deterministic function of $s_{i,j}$, needs to be evaluated *for each path*. This is expensive and undesirable. In the following section we present a novel approach for the evaluation of the expectation, which is both efficient and accurate. *The method is similar to the one presented in the stochastic local volatility setting in Section 2.*

3.1. *Establishing* $\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right]$

In this section we determine an approximation for the *non-trivial* expectation in the local volatility component (37). Similar to the approach for evaluating stochastic local volatility models, the method essentially consists of two projection steps. We first apply a projection on a standard normal random variable, employing the equality in distribution of cumulative distribution functions, and subsequently we make use of ordinary least squares regression.

We start by applying a change of measure:

$$\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right] = P(0,t)\mathrm{E}^t\left[r(t)\mathbb{1}_{S(t)>K}\right]. \tag{38}$$

At the right-hand side, under the t-forward measure, $r(\cdot)$ is normally distributed with mean

$$\mu_r^t(t) = r(0)\mathrm{e}^{-\lambda t} + \int_0^t \widetilde{\theta}(u)\mathrm{e}^{-\lambda(t-u)}\mathrm{d}u, \quad \widetilde{\theta}(u) := \lambda\theta(u) + \frac{\eta^2}{\lambda}\left(\mathrm{e}^{-\lambda(t-u)} - 1\right)$$

[1]The expression for $\theta(\cdot)$ is obtained by decomposing the Hull-White model, see e.g. Pelsser (2000).

and standard deviation

$$\sigma_r^t(t) = \left(\frac{\eta^2}{2\lambda} \left(1 - \mathrm{e}^{-2\lambda t} \right) \right)^{1/2}.$$

Further, the CDF of $S(t)$ under the t-forward measure, denoted by $F^t_{S(t)}(\cdot)$, can be derived from the following well-known relation (see e.g. Gatheral (2006)):

$$\frac{\partial C(t,K)}{\partial K} = -P(0,t)\mathrm{Q}^t \left[S(t) > K \right],$$

where $C(t,K)$ is the price at $t = 0$ ('today') of a European call option with maturity t and strike K and $P(0,t)$ denotes the zero-coupon bond with expiry t. This relation directly implies

$$F^t_{S(t)}(K) = 1 - \mathrm{Q}^t \left[S(t) > K \right] = 1 + \frac{1}{P(0,t)} \frac{\partial C(t,K)}{\partial K}.$$

To evaluate the expectation at the right-hand side of equation (38), for a fixed t we project $S(t)$ onto a standard normal distribution $X \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$ via the function $g(\cdot)$, defined by[1] $g(\cdot) := F^{-1}_{S(t)}(F_X(\cdot))$, which ensures $S(t) \overset{\mathrm{d}}{=} g(X)$ and, moreover, for elements $S(t) = s$ and $X = x$: $s = g(x)$. This element-wise equality implies $x = g^{-1}(s) = F^{-1}_X(F^t_{S(t)}(s))$, which yields for the expectation in (38):

$$\mathrm{E}^{\mathrm{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t)>K} \right] = P(0,t)\mathrm{E}^t \left[r(t) \mathbb{1}_{g(X)>K} \right] = P(0,t)\mathrm{E}^t \left[r(t) \mathbb{1}_{X > g^{-1}(K)} \right]. \tag{39}$$

We proceed with the second projection step. Trivially, from (39) we write (as $F_X(g^{-1}(K)) = F^t_{S(t)}(K)$)

$$\begin{aligned} \mathrm{E}^{\mathrm{Q}} \left[\frac{r(t)}{M(t)} \mathbb{1}_{S(t)>K} \right] &= P(0,t)\mathrm{E}^t \left[r(t) \middle| X > g^{-1}(K) \right] \left(1 - F_X(g^{-1}(K)) \right) \\ &= P(0,t) \left(\mu_r^t(t) + \sigma_r^t(t)\mathrm{E}^t \left[Z \middle| X > g^{-1}(K) \right] \right) \left(1 - F^t_{S(t)}(K) \right). \end{aligned}$$

Similar to the approach presented in the stochastic local volatility setting, to evaluate the conditional expectation, we apply a projection on a set of basis functions $\psi_k(\cdot)$, $k = 1, 2, \ldots, n$ $(n < \infty)$ which depend on X. We again choose the simple set of orthogonal polynomials $\{1, x, x^2, \ldots, x^{n-1}\}$ and apply OLS regression to compute the corresponding coefficients, which yields

$$\begin{aligned} \mathrm{E} \left[Z \middle| X > g^{-1}(K) \right] = \widehat{\beta}_0 + \widehat{\beta}_1 \mathrm{E} \left[X \middle| X > g^{-1}(K) \right] + \widehat{\beta}_2 \mathrm{E} \left[X^2 \middle| X > g^{-1}(K) \right] \\ + \ldots + \widehat{\beta}_{n-1} \mathrm{E} \left[X^{n-1} \middle| X > g^{-1}(K) \right] + \epsilon. \end{aligned}$$

The truncated moments of X allow for an analytic evaluation, which we state in Result 3.1.

[1] For notation purposes we suppress the t-superscript in the inverse of the CDF of $S(t)$.

RESULT 3.1 (Moments of a truncated standard normal random variable) *Given* $X \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$ *and define* $m_i := \mathrm{E}\left[X^i \middle| X > a\right]$, *with* $a \in \mathrm{R}$. *The truncated moments are given by*

$$m_i = (i-1)m_{i-2} + \frac{a^{i-1} f_{\mathcal{N}(0,1)}(a)}{1 - F_{\mathcal{N}(0,1)}(a)},\ i = 1, 2, \ldots,$$

with $m_{-1} = 0$, $m_0 = 1$ *and* $f_{\mathcal{N}(0,1)}(\cdot)$ *and* $F_{\mathcal{N}(0,1)}$ *denoting the standard normal probability density and cumulative distribution functions, respectively.*

Combining results yields

$$\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right] = P(0,t)\Big(\mu_r^t(t) + \sigma_r^t(t)\left(\widehat{\beta}_0 + \widehat{\beta}_1 \mathrm{E}\left[X \middle| X > g^{-1}(K)\right] + \widehat{\beta}_2 \mathrm{E}\left[X^2 \middle| X > g^{-1}(K)\right]\right.$$
$$\left. + \ldots + \widehat{\beta}_{n-1}\mathrm{E}\left[X^{n-1} \middle| X > g^{-1}(K)\right]\right)\Big)\left(1 - F_{S(t)}^t(K)\right) + \epsilon. \tag{40}$$

The error ϵ is introduced in the regression step and is discussed in Appendix A. In Section 2.3.1 we demonstrate a decrease of this error with respect to the number of Monte Carlo realizations.

3.1.1. **Alternative approach.** Instead of applying regression, an alternative approach is based on the *assumption* that $r(\cdot)$ and X in (39) are governed by a joint bivariate normal distribution. For two jointly normally distributed random variables $X_1 \overset{\mathrm{d}}{=} \mathcal{N}(\mu_1, \sigma_1)$ and $X_2 \overset{\mathrm{d}}{=} \mathcal{N}(\mu_2, \sigma_2)$, correlated with correlation ρ, the following result holds:

$$\mathrm{E}\left[X_1 \mathbb{1}_{X_2>k}\right] = \left(\mu_1 + \rho\sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}\right)\left(1 - F_{X_2}(k)\right), \tag{41}$$

where $f_{\mathcal{N}(0,1)}(\cdot)$ and $F_{\mathcal{N}(0,1)}(\cdot)$ are the standard normal PDF and CDF, respectively, and $F_{X_2}(\cdot)$ is the CDF corresponding to the random variable X_2. A proof of this result is given in Appendix B.

By the result in (41) the expectation in (39) is approximated as follows:

$$\mathrm{E}^t\left[r(t)\mathbb{1}_{X>g^{-1}(K)}\right] = \left(\mu_r^t(t) + \rho_{r,X}^t(t)\sigma_r^t(t)\frac{f_{\mathcal{N}(0,1)}\left(g^{-1}(K)\right)}{1 - F_{\mathcal{N}(0,1)}\left(g^{-1}(K)\right)}\right)\left(1 - F_X(g^{-1}(K))\right) + \epsilon,$$

with $g^{-1}(K) = F_X^{-1}(F_{S(t)}^t(K))$, where the error term ϵ is introduced by assuming that $r(t)$ and X are jointly bivariate normally distributed under the t-forward measure. Further, as $F_X(g^{-1}(K)) = F_{S(t)}^t(K)$, we have

$$\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right] = P(0,t)\left(\mu_r^t(t) + \rho_{r,X}^t(t)\sigma_r^t(t)\frac{f_{\mathcal{N}(0,1)}\left(g^{-1}(K)\right)}{1 - F_{S(t)}^t(K)}\right)\left(1 - F_{S(t)}^t(K)\right) + \epsilon, \tag{42}$$

where the correlation parameter is numerically (i.e. based on the Monte Carlo paths) established by applying a change of measure:

$$\rho_{r,X}^t(t) \overset{\mathrm{def}}{=} \frac{\mathrm{E}^t\left[r(t)X\right] - \mathrm{E}^t\left[r(t)\right]\mathrm{E}^t\left[X\right]}{\sigma_r^t(t)\sigma_X^t} = \frac{\mathrm{E}^t\left[r(t)X\right]}{\sigma_r^t(t)} = \frac{1}{P(0,t)\sigma_r^t(t)}\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}X\right],$$

with $X = g^{-1}(K)$, $P(0,t)$ is defined in (36) and $\sigma_r^t(t) = \left(\frac{\eta^2}{2\lambda}(1-\mathrm{e}^{-2\lambda t})\right)^{1/2}$.

A comparison of (40) and (42) makes clear that the latter can be considered as a special case of the more generic expression in (40). We apply the SC-R approach in (40) or the alternative in (42) in a Monte Carlo simulation framework according to Algorithm 2. As for Algorithm 1, the indices $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, M$ denote the time-step and path, respectively.

for *each time-step* t_i, $i = 1, 2, \ldots, N$ **do**
1 Generate M pairs $(s_{i,j}, r_{i,j})$, $j = 1, 2, \ldots, M$ by going forward one time-step in the standard Euler discretization scheme of the Local Volatility-Hull White model.
2 Compute $\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t_i)}{M(t_i)}\mathbb{1}_{S(t_i)>s_{i,j}}\right]$ according to either (40) or the alternative in (42).
3 Establish the local volatility component $\sigma^2(t_i, s_{i,j})$ by equation (37) – use its value in step 1.
end
4 Price European call options based on the obtained values for $S(\cdot)$ at the time to maturity.

Algorithm 2: Pricing European call options by a Monte Carlo simulation of the LV-HW model, incorporating the SC-R approach (Section 3.1) or the alternative (Section 3.1.1).

Our numerical experiments in Section 3.2 indicate that the approximations (40) and (42) show a similar performance for the shorter expiries, whereas for the longer expiries the former outperforms the latter. The reason for this is the fact that the *error* due to the bivariate normality assumption becomes more pronounced for longer expiries, i.e. the joint distribution of $r(\cdot)$ and X in (39) then resembles less a bivariate normal distribution.

3.2. *Numerical experiments*

In this section we test the accuracy of the approximation in (40) and the alternative (42). We price European call options by means of a Monte Carlo simulation[1] of the Local Volatility-Hull White model. At each time-step in the simulation we either use the approximation (40) or approximate the expectation according to (42). We consider two sets of Hull-White parameters in the literature, see Table 6. In line with the literature we choose for both sets $r_0 = 0.02$.

We generate synthetic market data by applying Fourier techniques to the Heston model, which we assume to be calibrated perfectly to the market. We choose three sets of Heston parameters for which the Feller condition is strongly violated, namely the parameter sets presented in Andersen (2008), see Table 7. In the regression we choose $n = 5$ basis functions, so we consider the first four moments of the truncated standard normal distribution. Similar as in Piterbarg (2006), given the expiry T we consider the strikes $K_i = F_0^T \exp(0.1\delta_i\sqrt{T})$, with $F_0^T = S_0/P(0,T) = 1/P(0,T)$ (as $S_0 = 1$) denoting the initial forward and $\delta_i = -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5$.

For the two shortest expiries per Hull-White set the results are provided by Tables 8, 9 ('Heston market' Set 1), 10, 11 ('Heston market' Set 2), 12 and 13 ('Heston market' Set 3). We report the absolute error $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in %, with $\overline{\sigma}$ denoting the Black-Scholes implied volatility. The error $\epsilon_{\text{alternative}}$ corresponds to the alternative approach of Section 3.1.1. Given the standard

[1] The Monte Carlo simulation consists of $2 \cdot 10^5$ paths (20 seeds, each seed constitutes 10^4 paths) and 200 time-steps per year, unless otherwise mentioned.

Set	T	λ	η	$\rho_{r,s}$
A	1, 2, 5	0.01	0.01	0.6
B	5, 10, 15	0.01	0.007	−0.15

Table 6. Hull-White model parameters as in Grzelak and Oosterlee (2011) (Set A) and Piterbarg (2006), Grzelak and Oosterlee (2012) (Set B).

Set	V_0	$\rho_{s,v}$	γ	κ	$\overline{V}$
1	0.04	−0.9	1	0.5	0.04
2	0.04	−0.5	0.9	0.3	0.04
3	0.09	−0.3	1	1	0.09

Table 7. 'Heston market' parameters as in Andersen (2008).

'Heston market' Set 1, Hull-White Set A					
$T=1$			$T=2$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.00 (0.27)	0.01 (0.25)	0.84	0.01 (0.22)	0.03 (0.21)
0.92	0.01 (0.23)	0.01 (0.22)	0.90	0.01 (0.20)	0.02 (0.18)
0.97	0.02 (0.21)	0.02 (0.19)	0.97	0.01 (0.18)	0.01 (0.16)
1.02	0.04 (0.17)	0.04 (0.16)	1.04	0.03 (0.15)	0.01 (0.13)
1.07	0.05 (0.13)	0.05 (0.11)	1.12	0.04 (0.11)	0.06 (0.08)
1.13	0.05 (0.10)	0.05 (0.09)	1.20	0.01 (0.12)	0.02 (0.10)
1.19	0.04 (0.11)	0.04 (0.10)	1.29	0.02 (0.15)	0.03 (0.15)

Table 8. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 1, Hull-White Set A ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

'Heston market' Set 1, Hull-White Set B					
$T=5$			$T=10$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.04 (0.15)	0.13 (0.14)	0.76	0.01 (0.11)	0.38 (0.10)
0.88	0.03 (0.13)	0.11 (0.12)	0.89	0.00 (0.11)	0.31 (0.08)
0.90	0.02 (0.12)	0.07 (0.10)	1.04	0.01 (0.10)	0.21 (0.08)
1.11	0.00 (0.09)	0.02 (0.08)	1.22	0.03 (0.09)	0.05 (0.07)
1.24	0.01 (0.08)	0.09 (0.07)	1.43	0.06 (0.08)	0.26 (0.06)
1.38	0.01 (0.11)	0.10 (0.10)	1.68	0.01 (0.08)	0.50 (0.05)
1.55	0.02 (0.16)	0.03 (0.14)	1.96	0.01 (0.18)	0.28 (0.14)

Table 9. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 1, Hull-White Set B ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

'Heston market' Set 2, Hull-White Set A					
$T=1$			$T=2$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.09 (0.27)	0.09 (0.27)	0.84	0.05 (0.23)	0.07 (0.23)
0.92	0.08 (0.24)	0.07 (0.24)	0.90	0.05 (0.21)	0.06 (0.21)
0.97	0.06 (0.21)	0.06 (0.20)	0.97	0.05 (0.19)	0.05 (0.18)
1.02	0.05 (0.17)	0.04 (0.16)	1.04	0.04 (0.18)	0.03 (0.17)
1.07	0.05 (0.16)	0.04 (0.16)	1.12	0.05 (0.20)	0.04 (0.19)
1.13	0.05 (0.17)	0.05 (0.17)	1.20	0.09 (0.24)	0.08 (0.24)
1.19	0.08 (0.19)	0.07 (0.19)	1.29	0.14 (0.31)	0.13 (0.31)

Table 10. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 2, Hull-White Set A ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

deviations, we observe that for all LV-HW experiments both the SC-R approach and the alternative yield 95%-confidence intervals[2] that cover the reference implied volatility, except for one case: the alternative approach with 'Heston market' Set 1, Hull-White Set B, $T = 10$. For the expiries $T = 5$ and $T = 15$ corresponding to Hull-White Sets A and B, respectively, the results are displayed in Figures 3.1 and 3.2, respectively. With 'LV-HW alt.' we denote the alternative approach.

In general, for both the approximation (40) and its alternative (42) the results are highly satisfactory. For the shorter expiries the two methods show a comparable performance, however for $T = 10$ and $T = 15$ we observe that the regression-based approach outperforms the alternative – we clearly observe this in Figure 3.2. The reason for this is that the alternative approach relies on the bivariate normality assumption of $r(\cdot)$ and X in (39). The error introduced by the bivariate normality assumption becomes more pronounced for longer expiries, since the joint distribution of $r(\cdot)$ and X resembles less a bivariate normal distribution when going forward in time. In the left-hand plot of Figure 3.2 we observe a slight mismatch on the right-hand side of the strike range.

[2]The boundaries of the 95%-confidence interval are $\mu(\overline{\sigma}_{1,\text{model}}, \overline{\sigma}_{2,\text{model}}, \overline{\sigma}_{3,\text{model}}, \ldots) \pm 1.96 \cdot \sigma(\overline{\sigma}_{1,\text{model}}, \overline{\sigma}_{2,\text{model}}, \overline{\sigma}_{3,\text{model}}, \ldots)$, with $\mu(\cdot)$ and $\sigma(\cdot)$ denoting the mean and standard deviation, respectively, and $\overline{\sigma}_{i,\text{model}}$ stands for the model implied volatility (obtained from Monte Carlo) corresponding to the ith seed.

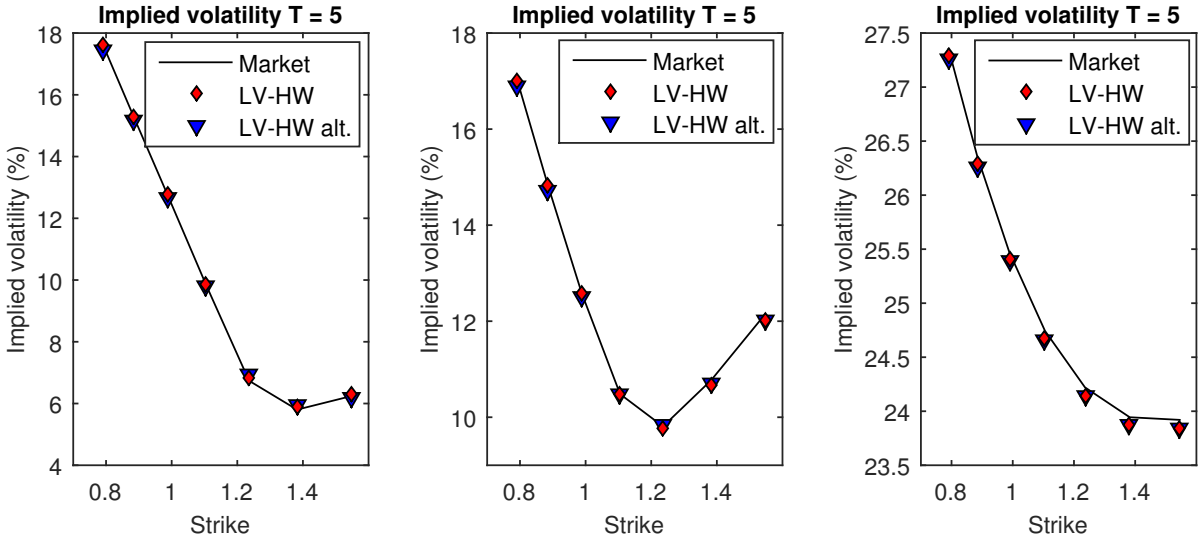

Figure 3.1. Black-Scholes implied volatilities corresponding to Hull-White Set A, $T = 5$ and the 'Heston market' Sets 1 (left), 2 (middle) and 3 (right), using 200 time-steps per year for Sets 1 and 2. For Set 3 500 time-steps per year are used, as 200 time-steps did not yield highly satisfactory results. 'LV-HW alt.' denotes the alternative approach presented in Section 3.1.1. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths).

'Heston market' Set 2, Hull-White Set B					
$T = 5$			$T = 10$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.08 (0.20)	0.11 (0.20)	0.76	0.05 (0.15)	0.20 (0.15)
0.88	0.07 (0.18)	0.09 (0.18)	0.89	0.04 (0.14)	0.14 (0.15)
0.90	0.05 (0.16)	0.05 (0.16)	1.04	0.01 (0.14)	0.04 (0.14)
1.11	0.03 (0.16)	0.00 (0.16)	1.22	0.02 (0.15)	0.09 (0.16)
1.24	0.04 (0.18)	0.00 (0.19)	1.43	0.00 (0.20)	0.13 (0.19)
1.38	0.06 (0.24)	0.04 (0.24)	1.68	0.05 (0.28)	0.05 (0.27)
1.55	0.10 (0.32)	0.09 (0.32)	1.96	0.06 (0.37)	0.05 (0.35)

Table 11. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 2, Hull-White Set B ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

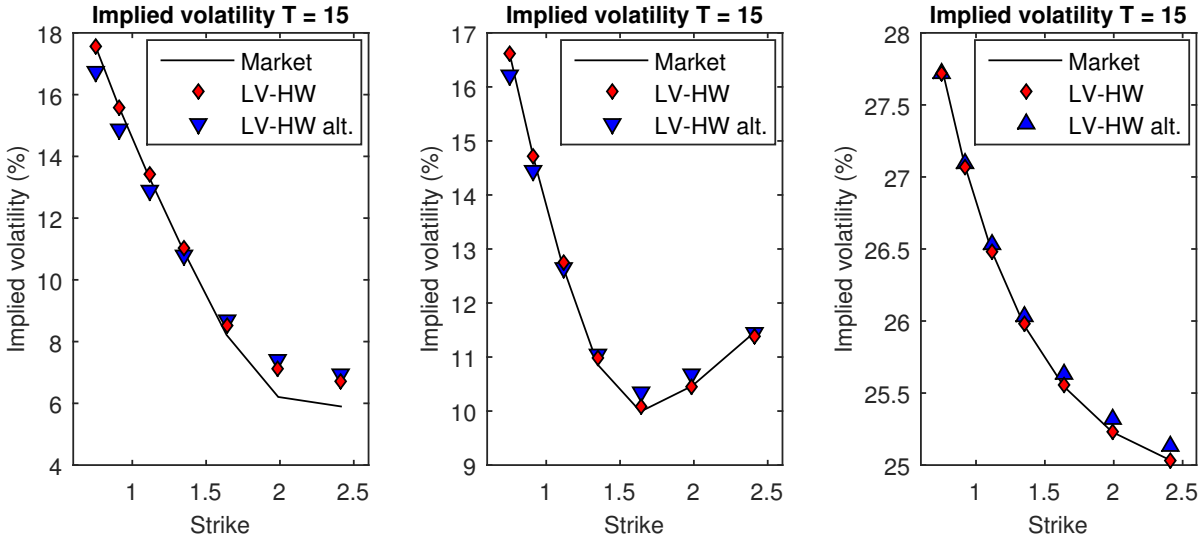

Figure 3.2. Black-Scholes implied volatilities corresponding to Hull-White Set B, $T = 15$ and the 'Heston market' Sets 1 (left), 2 (middle) and 3 (right). 'LV-HW alt.' denotes the alternative approach presented in Section 3.1.1. Results are obtained with $2 \cdot 10^5$ paths (2 seeds, each seed constitutes 10^5 paths) and 200 time-steps per year.

'Heston market' Set 3, Hull-White Set A					
$T=1$			$T=2$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.88	0.14 (0.26)	0.14 (0.26)	0.84	0.09 (0.29)	0.10 (0.29)
0.92	0.13 (0.24)	0.13 (0.24)	0.90	0.08 (0.28)	0.09 (0.28)
0.97	0.12 (0.23)	0.11 (0.23)	0.97	0.08 (0.28)	0.08 (0.27)
1.02	0.11 (0.24)	0.11 (0.24)	1.04	0.08 (0.29)	0.08 (0.28)
1.07	0.10 (0.25)	0.10 (0.25)	1.12	0.08 (0.31)	0.08 (0.31)
1.13	0.09 (0.27)	0.09 (0.27)	1.20	0.09 (0.34)	0.09 (0.34)
1.19	0.09 (0.29)	0.09 (0.29)	1.29	0.11 (0.37)	0.11 (0.37)

Table 12. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 3, Hull-White Set A ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

'Heston market' Set 3, Hull-White Set B					
$T=5$			$T=10$		
K	ϵ	$\epsilon_{\text{alternative}}$	K	ϵ	$\epsilon_{\text{alternative}}$
0.79	0.04 (0.27)	0.04 (0.28)	0.76	0.01 (0.35)	0.01 (0.38)
0.88	0.03 (0.28)	0.03 (0.29)	0.89	0.00 (0.36)	0.01 (0.38)
0.90	0.03 (0.30)	0.03 (0.30)	1.04	0.00 (0.38)	0.01 (0.40)
1.11	0.03 (0.31)	0.03 (0.32)	1.22	0.02 (0.40)	0.01 (0.43)
1.24	0.04 (0.32)	0.03 (0.33)	1.43	0.02 (0.43)	0.01 (0.45)
1.38	0.05 (0.34)	0.04 (0.35)	1.68	0.03 (0.46)	0.01 (0.48)
1.55	0.05 (0.38)	0.04 (0.39)	1.96	0.04 (0.49)	0.01 (0.52)

Table 13. Errors $\epsilon := |\overline{\sigma}_{\text{market}} - \overline{\sigma}_{\text{LV-HW}}|$ in % corresponding to the 'Heston market' Set 3, Hull-White Set B ($\overline{\sigma}$ denotes the Black-Scholes implied volatility). Numbers in parentheses are standard deviations over the seeds.

This is *not* due to the performance of the approximation methods, but due to general Monte Carlo bias[1].

4. Conclusion

In this article we considered in a Monte Carlo simulation framework two classes of hybrid local volatility models, namely stochastic local volatility models and the local volatility model extended with stochastic interest rates. For both model classes a *non-trivial* (conditional) expectation needs to be evaluated, which cannot be extracted from the market quotes and is expensive to compute. In this article we presented a novel, efficient approach to the evaluation of these expectations. The method essentially consists of two projection steps; the first projection employs the equality in distribution of cumulative distribution functions, which stands at the basis of the stochastic collocation method, the second projection step relies on standard regression techniques. By means of numerical experiments we confirm that our approach facilitates an efficient Monte Carlo evaluation and yields highly accurate pricing results for European-type options.

References

Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th edition*, 1972, Dover, New York.

An, Y. and Li, C., Approximating Local Volatility Functions of Stochastic Volatility Models: A Closed-Form Expansion Approach. *Probability in the Engineering and Informational Sciences*, 2015, **29**, 547–563.

Andersen, L., Simple and Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 2008, **11**, 1–42.

[1]We conclude this based on a simulation of the LV-HW model applying an 'exact', 'brute-force' approach to compute the expectation $\mathrm{E}^{\mathrm{Q}}\left[\frac{r(t)}{M(t)}\mathbb{1}_{S(t)>K}\right]$. The same bias was observed.

Atlan, M., Localizing volatilities. *ArXiV preprint math/0604316*, 2006.

Babuška, I., Nobile, F. and Tempone, R., A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 2007, **45**, 1005–1034.

Baker, G., Beneder, R. and Zilber, A., FX Barriers with Smile Dynamics. *Available at SSRN 964627*, 2004.

Beck, J., Tempone, R., Nobile, F. and Tamellini, L., On the Optimal Polynomial Approximation of Stochastic PDEs by Galerkin and Collocation Methods. *Mathematical Models and Methods in Applied Sciences*, 2012, **22**, 1–33.

Benhamou, E., Gobet, E. and Miri, M., Analytical Formulas for a Local Volatility Model with Stochastic Rates. *Quantitative Finance*, 2012, **12**, 185–198.

Benhamou, E., Rivoira, A. and Gruz, A., Stochastic Interest Rates for Local Volatility Hybrids Models. *Wilmott Magazine*, 2008.

Black, F. and Karasinski, P., Bond and Option Pricing when Short Rates are Lognormal. *Financial Analysts Journal*, 1991, **47**, 52–59.

Bloch, D.A. and Nakashima, Y., Multi-Currency Local Volatility Model. *Available at SSRN 1153337*, 2008.

Cameron, R.H. and Martin, W.T., The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals. *Annals of Mathematics*, 1947, pp. 385–392.

Clark, I.J., *Foreign Exchange Option Pricing: A Practitioners Guide*, 2011, John Wiley & Sons.

Deelstra, G. and Rayée, G., Local Volatility Pricing Models for Long-Dated FX Derivatives. *Applied Mathematical Finance*, 2012, pp. 1–23.

Derman, E., Ergener, D. and Kani, I., Static Options Replication. *The Journal of Derivatives*, 1995, **2**, 78–95.

Derman, E. and Kani, I., Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility. *International Journal of Theoretical and Applied Finance*, 1998, **1**, 61–110.

Dupire, B., Pricing with a Smile. *Risk*, 1994, **7**, 18–20.

Dupire, B., A Unified Theory of Volatility. *Derivatives Pricing: The Classic Collection*, 1996, pp. 185–196.

Fouque, J.P., Papanicolaou, G. and Sircar, K., Stochastic Volatility Correction to Black-Scholes. *Risk*, 2000, **13**, 89–92.

Ganapathysubramanian, B. and Zabaras, N., Sparse Grid Collocation Schemes for Stochastic Natural Convection Problems. *Journal of Computational Physics*, 2007, **225**, 652–685.

Gatheral, J., *The Volatility Surface: A Practitioner's Guide*, 2006, John Wiley & Sons, Ltd.

Greene, W.H., Econometric Analysis, 5th Edition. , 2002.

Grzelak, L.A. and Oosterlee, C.W., On the Heston Model with Stochastic Interest Rates. *SIAM Journal on Financial Mathematics*, 2011, **2**, 255–286.

Grzelak, L.A. and Oosterlee, C.W., On Cross-Currency Models with Stochastic Volatility and Correlated Interest Rates. *Applied Mathematical Finance*, 2012, **19**, 1–35.

Grzelak, L.A. and Oosterlee, C.W., From Arbitrage to Arbitrage-Free Implied Volatilities. *Journal of Computational Finance*, Forthcoming 2017.

Grzelak, L.A., Witteveen, J.A.S., Suarez-Taboada, M. and Oosterlee, C.W., The Stochastic Collocation Monte Carlo Sampler: Highly Efficient Sampling from 'Expensive' Distributions. *Available at SSRN 2529691*, 2014.

Guyon, J. and Henry-Labordère, P., Being Particular about Calibration. *Risk*, 2012, **25**, 88.

Gyöngy, I., Mimicking the One-Dimensional Marginal Distributions of Processes having an Itô Differential. *Probability Theory and Related Fields*, 1986, **71**, 501–516.

Hagan, P.S., Kumar, D., Lesniewski, A.S. and Woodward, D.E., Managing Smile Risk. *Wilmott Magazine*, 2002, pp. 84–108.

Henry-Labordère, P., Calibration of Local Stochastic Volatility Models to Market Smiles: a Monte-Carlo Approach. *Risk*, 2009, pp. 112–117.

Heston, S.L., A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 1993, **6**, 327–343.

Hull, J. and White, A., One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities. *Journal of Financial and Quantitative Analysis*, 1993, **28**, 235–254.

Jain, S. and Oosterlee, C.W., The Stochastic Grid Bundling Method: Efficient pricing of Bermudan options and their Greeks. *Applied Mathematics and Computation*, 2015, **269**, 412–431.

Jex, M., Henderson, R. and Wang, D., Pricing Exotics under the Smile. *Risk*, 1999, **12**, 72–75.

Johnson, S. and Lee, H., Capturing the Smile. *Risk*, 2003, **16**, 89–95.

Lipton, A. and McGhee, W., Universal Barriers. *Risk*, 2002, **15**, 81–85.

Lipton, A., The Vol Smile Problem. *Risk*, 2002, **15**, 61–66.

Lipton, A., Gal, A. and Lasis, A., Pricing of Vanilla and First-Generation Exotic Options in the Local Stochastic Volatility Framework: Survey and new Results. *Quantitative Finance*, 2014, **14**, 1899–1922.

Longstaff, F.A. and Schwartz, E.S., Valuing American Options by Simulation: a Simple Least-Squares Approach. *Review of Financial Studies*, 2001, **14**, 113–147.

Lorig, M., Pagliarani, S., Pascucci, A. *et al.*, Explicit Implied Volatilities for Multifactor Local-Stochastic Volatility Models. *Mathematical Finance, DOI: 10.1111/mafi.12105*, 2015.

Mardia, K.V., Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhyā: The Indian Journal of Statistics, Series B*, 1974, pp. 115–128.

Nobile, F., Tempone, R. and Webster, C.G., A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data. *SIAM Journal on Numerical Analysis*, 2008, **46**, 2309–2345.

Pascucci, A. and Mazzon, A., The Forward Smile in Local-Stochastic Volatility Models. *Available at SSRN*, 2015.

Pelsser, A., *Efficient Methods for Valuing Interest Rate Derivatives*, 2000, Springer Science & Business Media.

Piterbarg, V., Cross-Currency Exotics: Smiling Hybrids. *Risk*, 2006, **19**, 66–71.

Piterbarg, V., Markovian Projection Method for Volatility Calibration. *Risk*, 2007.

Rebonato, R., *Volatility and Correlation in the Pricing of Equity, FX, and Interest-rate Options*, 1999, John Wiley & Sons, Ltd.

Rebonato, R., McKay, K. and White, R., *The SABR/LIBOR Market Model: Pricing, Calibration and Hedging for Complex Interest-Rate Derivatives*, 2011, Wiley.com.

Ren, Y., Madan, D. and Qian, M.Q., Calibrating and Pricing with Embedded Local Volatility Models. *Risk*, 2007, **20**, 138–143.

Sankaran, S. and Marsden, A.L., A Stochastic Collocation Method for Uncertainty Quantification and Propagation in Cardiovascular Simulations. *Journal of Biomechanical Engineering*, 2011, **133**, 031001.

Schöbel, R. and Zhu, J., Stochastic Volatility with an Ornstein–Uhlenbeck Process: an Extension. *European Finance Review*, 1999, **3**, 23–46.

Steele, J.M., *Stochastic Calculus and Financial Applications*, Vol. 45, , 2001, Springer Science & Business Media.

Tataru, G. and Fisher, T., Stochastic Local Volatility. *Quantitative Development Group, Bloomberg Version 1*, 2010.

Tian, Y., Zhu, Z., Lee, G., Klebaner, F. and Hamza, K., Calibrating and Pricing with a Stochastic-Local Volatility Model. *The Journal of Derivatives*, 2015, **22**, 21–39.

Van der Stoep, A.W., Grzelak, L.A. and Oosterlee, C.W., The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation. *International Journal of Theoretical and Applied Finance*, 2014, **17**, 1450045.

Watanabe, S., Analysis of Wiener Functionals (Malliavin Calculus) and its Applications to Heat Kernels. *The Annals of Probability*, 1987, **15**, 1–39.

Witteveen, J.A.S. and Iaccarino, G., Simplex Stochastic Collocation with Random Sampling and Extrapolation for Nonhypercube Probability Spaces. *SIAM Journal on Scientific Computing*, 2012, **34**, A814–A838.

Xiu, D., Efficient Collocational Approach for Parametric Uncertainty Analysis. *Commun. Comput. Phys*, 2007, **2**, 293–309.

Xiu, D. and Hesthaven, J.S., High-Order Collocation Methods for Differential Equations with Random Inputs. *SIAM Journal on Scientific Computing*, 2005, **27**, 1118–1139.

Appendix A: Error analysis & discussion

In this section we briefly discuss the (asymptotics of the) errors in (26) which are due to the stochastic collocation method ($\bar{\epsilon}_1(\cdot)$) and the projection on an orthonormal basis and subsequently applying regression ($\bar{\epsilon}_2$).

A.1. *Stochastic collocation error*

The first error is $\bar{\epsilon}_1(K) := \mathrm{E}\left[\epsilon_1(Z)|\, X = g^{-1}(K)\right]$, with Z denoting a standard normal random variable. It is introduced by projecting, for a given t, the random variable $V(t)$ on Z via a Lagrange polynomial $h_{N_V}(\cdot)$, which interpolates through the collocation values $v_i = F^{-1}_{V(t)}(F_Z(z_i))$, where the *collocation points* z_i are chosen in an *optimal way*, namely based on the zeros of Hermite polynomials.

We start the analysis of $\bar{\epsilon}_1(\cdot)$ by the following (in)equalities:

$$\left|\mathrm{E}\left[h(Z) - h_{N_V}(Z)\right]\right| \leq \mathrm{E}\left[\left|h(Z) - h_{N_V}(Z)\right|\right] \text{ (by Jensen's inequality)} \tag{A1}$$

$$\leq \left(\mathrm{E}\left[\left(h(Z) - h_{N_V}(Z)\right)^2\right]\right)^{1/2}, \tag{A2}$$

where the latter inequality is a standard relation between L^p-norms, see e.g. Steele (2001).

As pointed out in the error analysis in Grzelak *et al.* (2014), the advantage of using optimal collocation points is that the stochastic collocation method can be connected to the computation of integrals by Gauss quadrature, which for the general function $\Psi(\cdot)$, weight function $f_Z(\cdot)$ and quadrature weights $\omega_i,\ i = 1, 2, \ldots, N_V$ reads:

$$\mathrm{E}\left[\Psi(Z)\right] = \int_{\mathrm{R}} \Psi(z) f_Z(z)\mathrm{d}z = \sum_{i=1}^{N_V} \Psi(z_i)\omega_i + \epsilon_{N_V}. \tag{A3}$$

In this article we choose the collocation variable $Z \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$, for which a simple relation between the 'stochastic collocation pairs' $\{z_i, \omega_i\}_{i=1}^{N_V}$ and the Gauss-Hermite quadrature pairs $\{z_i^H, \omega_i^H\}_{i=1}^{N_V}$ exists. Whereas in the stochastic collocation method the weight function $f_Z(z) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)$ is used, Gauss-Hermite quadrature is based on $f_Z(z) = \exp(-z^2)$. However, standard calculus yields $\int_{\mathrm{R}} \Psi(z)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}z^2}\mathrm{d}z = \int_{\mathrm{R}} \Psi(z\sqrt{2})\frac{1}{\sqrt{\pi}}\mathrm{e}^{-z^2}\mathrm{d}z$. From this one can show the relations $z_i^H = z_i/\sqrt{2}$ and $\omega_i^H = \omega_i\sqrt{\pi}$, which implies that for a standard normal collocation variable the error due to stochastic collocation is given by the error for the Gauss-Hermite quadrature, which is, see e.g. Abramowitz and Stegun (1972), given by:

$$\epsilon_{N_V} = \frac{N_V!\sqrt{\pi}}{2^{N_V}}\frac{\Psi^{(2N_V)}(\widehat{\xi})}{(2N_V)!}, \tag{A4}$$

with $\widehat{\xi} \in [\min(\boldsymbol{z})\max(\boldsymbol{z})]$, $\boldsymbol{z} = (z_1, z_2, \ldots, z_{N_V})$. By choosing $\Psi(z) = (h(z) - h_{N_V}(z))^2$, with $h(\cdot) = F^{-1}_{V(t)}(F_Z(\cdot))$ and $h_{N_V}(\cdot)$ the corresponding approximating Lagrange polynomial, equation (A3) yields the error:

$$\begin{aligned}
\mathrm{E}\left[\left(h(Z) - h_{N_V}(Z)\right)^2\right] &= \int_{\mathrm{R}} \left(h(z) - h_{N_V}(z)\right)^2 f_Z(z)\mathrm{d}z \\
&= \sum_{i=1}^{N_V} \left(h(z_i) - h_{N_V}(z_i)\right)^2\omega_i + \epsilon_{N_V} \\
&= \epsilon_{N_V},
\end{aligned}$$

as $h(z_i) = h_{N_V}(z_i)$. $\Psi(z) = (h(z) - h_{N_V}(z))^2$ can be written more explicitly as the square of the

standard Lagrange interpolation error, see e.g. Abramowitz and Stegun (1972):

$$\Psi(z) = \left(\frac{1}{N_V!} \left. \frac{\mathrm{d}^{N_V} h(z)}{\mathrm{d}z^{N_V}} \right|_{z=\widehat{\xi}} \prod_{i=1}^{N_V} (z - z_i) \right)^2, \tag{A5}$$

with $\widehat{\xi} \in [\min(\boldsymbol{z}) \max(\boldsymbol{z})]$, $\boldsymbol{z} = (z_1, z_2, \ldots, z_{N_V})$. This error may be bounded by choosing $z = \widehat{\xi}$ for which $\left| \frac{\mathrm{d}^{N_V} h(z)}{\mathrm{d}z^{N_V}} \right|$ attains its maximum.

Substituting (A5) in (A4) yields a complete specification for ϵ_{N_V}, and it can be shown that it converges to zero as $N_V \to \infty$ (under the condition that $\Psi(\cdot)$ is sufficiently smooth). Then, from the inequalities (A1) and (A2), combined with the identity

$$\mathrm{E}\left[h(Z) - h_{N_V}(Z)\right] = \int_{\mathrm{R}} \mathrm{E}\left[h(Z) - h_{N_V}(Z) \middle| X = x\right] f_X(x) \mathrm{d}x,$$

the error $\bar{\epsilon}_1(x) := \mathrm{E}\left[h(Z) - h_{N_V}(Z) \middle| X = x\right]$ converges to zero as $N_V \to \infty$, for arbitrary $x \in \mathrm{R}$.

A.2. *Regression error*

The second error term $\bar{\epsilon}_2 := \sum_{p=1}^{N_V - 1} \widehat{\epsilon}_{2p}$ is due to the projection of the unknown conditional expectations $\mathrm{E}\left[Z^p \middle| X = g^{-1}(K)\right]$ on a set of basis functions $\{1, x, x^2, \ldots, x^{n-1}\}$ and applying OLS regression. This is polynomial regression, which is a special case of multiple linear regression. In a Monte Carlo simulation framework, given M observations for the underlying $S(\cdot)$, s_j, we write, as we evaluate the local volatility component in $K = s_j$:

$$\mathrm{E}\left[Z^p \middle| X = x_j\right] = b_{0p} + b_{1p} x_j + \ldots + b_{n-1,p} x_j^{n-1} + \epsilon_{2pj}, \quad j = 1, 2, \ldots, M, \tag{A6}$$

with $x_j := g^{-1}(s_j)$ and ϵ_{2pj} is the *unobserved* error term corresponding to the jth realization. We apply OLS regression to compute $\widehat{\beta}_{kp}$, $k = 0, 1, \ldots, n-1$, $p = 1, 2, \ldots, N_V - 1$, which are estimates for b_{kp}. This yields $\widehat{\mathrm{E}}\left[Z^p \middle| X = x_j\right] = \widehat{\beta}_{0p} + \widehat{\beta}_{1p} x_j + \ldots + \widehat{\beta}_{n-1,p} x_j^{n-1} + \widehat{\epsilon}_{2pj}$, $j = 1, 2, \ldots, M$.

Let $\boldsymbol{x} = (1 \;\; x \;\; x^2 \;\; \ldots \;\; x^{n-1})$ be an $M \times n$ matrix and denote its jth row by $\boldsymbol{x}_j$. Under some standard assumptions for the regression model (A6), e.g. for $j = 1, 2, \ldots, M$ the errors ϵ_{2pj} should have conditional mean zero, i.e. $\mathrm{E}[\epsilon_{2pj} | \boldsymbol{x}] = 0$ ('strict exogeneity'), amongst others, the *Gauss-Markov theorem* – see e.g. Greene (2002) – states that $\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}} := \left(\widehat{\beta}_{0p}, \widehat{\beta}_{1p}, \ldots, \widehat{\beta}_{n-1,p}\right)$ is the Best Linear Unbiased Estimator ('BLUE') of $\boldsymbol{b}_{\boldsymbol{p}} := (b_{0p}, b_{1p}, \ldots, b_{n-1,p})$ amongst all $\boldsymbol{\beta}_{\boldsymbol{p}}$ candidates. $\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$ is 'best' in a least squares sense, i.e. it is the *unique* value for $\boldsymbol{\beta}_{\boldsymbol{p}}$ for which the sum of squared residuals $\sum_{j=1}^{M} \left(z_j^p - \boldsymbol{x}_j \boldsymbol{\beta}_{\boldsymbol{p}}^{\mathrm{T}} \right)^2$ is minimized, with $z_j = F_Z^{-1}(F_{V(t)}(v_j))$. If we additionally assume that the error terms ϵ_{2pj}, $j = 1, 2, \ldots, M$ in (A6) are independent and identically distributed with mean zero and finite variance σ^2, one can prove that $\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$ is a *consistent* estimator of $\boldsymbol{b}_{\boldsymbol{p}}$ (convergence in probability)

$$\lim_{M \to \infty} \mathrm{P}\left(\left| \widehat{\boldsymbol{\beta}}_{\boldsymbol{p}} - \boldsymbol{b}_{\boldsymbol{p}} \right| \geq \varepsilon \right) = 0 \quad \forall \varepsilon > 0,$$

and, moreover, applying the central limit theorem, that $\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$ is asymptotically normal (convergence in distribution):

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}} \xrightarrow{\mathrm{d}} \mathcal{N}\left(\boldsymbol{b}_{\boldsymbol{p}}, \frac{\sigma^2}{M} \mathbf{Q}^{-1} \right) \text{ if } M \to \infty, \text{ with } \mathbf{Q} \text{ defined by } \lim_{M \to \infty} \mathrm{P}\left(\left| \frac{\boldsymbol{x}'\boldsymbol{x}}{M} - \mathbf{Q} \right| \geq \varepsilon \right) \quad \forall \varepsilon > 0.$$

Q is a positive definite matrix.

REMARK A.1 (Number of basis functions) *Although $\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$ is the 'best' estimator of $\boldsymbol{b}_{\boldsymbol{p}}$ in the sense that it has the lowest variance compared to all other unbiased estimators $\boldsymbol{\beta}_{\boldsymbol{p}}$, the absolute error can still be significant though, e.g. due to an inappropriate choice of the basis functions or the number of basis functions. Tests for the significance of the polynomial terms, in particular the highest order term, can be conducted, where the null hypothesis states that $\widehat{\beta}_{kp} = 0$ for some $k = 0,\ 1, \ldots, n-1$. Related to this, one can employ either a forward selection procedure or a backward elimination procedure, in which the model is successively fit in increasing or decreasing order under statistical testing, respectively. Other criteria one can consider to test whether a multiple linear regression model is well-constructed are e.g. the well-known R^2-value (coefficient of determination) and the condition number of the matrix involved in the regression, which is a measure for the ill-posedness and multicollinearity of the problem.*

Appendix B: Proofs of Lemma 2.1 and the result in Section 3.1.1

In this section we provide proofs for Lemma 2.1 and the result in Section 3.1.1, which we state here as Lemmas B.1 and B.2, respectively.

LEMMA B.1 *Given the two-dimensional model (29). Let X and Z denote standard normal random variables and assume for an arbitrary t that the elements $Y_i(t) = y_i,\ i = 1, 2,\ X = x,\ Z = z$ are related by $y_1 = g(x),\ y_2 = h(z)$, with $g(\cdot)$ and $h(\cdot)$ defined in (15) and (17), respectively. This implies that X and Z are jointly bivariate normally distributed.*

Proof. We start with writing (29) in terms of independent Brownian motions $\widetilde{W}_1(\cdot)$ and $\widetilde{W}_2(\cdot)$:

$$\mathrm{d}Y_1(t) = \sigma_1 Y_1(t)\mathrm{d}\widetilde{W}_1(t),\ \mathrm{d}Y_2(t) = \sigma_2 Y_2(t)\left(\rho \mathrm{d}\widetilde{W}_1(t) + \sqrt{1-\rho^2}\mathrm{d}\widetilde{W}_2(t)\right), \tag{B1}$$

with $Y_1(0) = y_{10}$ and $Y_2(0) = y_{20}$. The solution to (B1) reads

$$Y_1(t) = y_{10}\exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1\widetilde{W}_1(t)\right),\ Y_2(t) = y_{20}\exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2\left(\rho\widetilde{W}_1(t) + \sqrt{1-\rho^2}\widetilde{W}_2(t)\right)\right), \tag{B2}$$

respectively. One can easily show $\log(Y_i(t)) \overset{\mathrm{d}}{=} \mathcal{N}(\overline{\mu}_i, \overline{\sigma}_i)$, with $\overline{\mu}_i := \log(y_{i0}) - \frac{1}{2}\sigma_i^2 t$ and $\overline{\sigma}_i := \sigma_i\sqrt{t}$, $i = 1, 2$. Further, the element-wise equality $y_1 = g(x)$, with $g(\cdot)$ specified in (15), implies, as $X \overset{\mathrm{d}}{=} \mathcal{N}(0,1)$,

$$x = \ g^{-1}(y_1) = \ F_X^{-1}(F_{Y_1(t)}(y_1)) = F_{\mathcal{N}(0,1)}^{-1}F_{\mathcal{N}(0,1)}\left(\frac{\log(y_1) - \overline{\mu}_1}{\overline{\sigma}_1}\right) = \frac{\log(y_1) - \overline{\mu}_1}{\overline{\sigma}_1},$$

which yields for y_1 and similarly for y_2, substituting the values of $\overline{\mu}_i$ and $\overline{\sigma}_i$, $i = 1, 2$:

$$y_1 = y_{10}\exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1\sqrt{t}x\right),\ y_2 = y_{20}\exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2\sqrt{t}z\right),$$

thus

$$Y_1(t) = y_{10}\exp\left(-\frac{1}{2}\sigma_1^2 t + \sigma_1\sqrt{t}X\right),\ Y_2(t) = y_{20}\exp\left(-\frac{1}{2}\sigma_2^2 t + \sigma_2\sqrt{t}Z\right). \tag{B3}$$

Equations (B2) and (B3) imply $X = \frac{1}{\sqrt{t}}\widetilde{W}_1(t)$ and $Z = \frac{1}{\sqrt{t}}\left(\rho\widetilde{W}_1(t) + \sqrt{1-\rho^2}\widetilde{W}_2(t)\right) = \rho X +$

$\sqrt{1-\rho^2}\widetilde{Z}$, with $\widetilde{Z} := \frac{1}{\sqrt{t}}\widetilde{W}_2(t)$. As we are able to express Z in terms of X via the Cholesky decomposition, Z and X are jointly bivariate normally distributed. □

LEMMA B.2 *For two jointly normally distributed random variables* $X_1 \overset{\mathrm{d}}{=} \mathcal{N}(\mu_1, \sigma_1)$ *and* $X_2 \overset{\mathrm{d}}{=} \mathcal{N}(\mu_2, \sigma_2)$ *correlated with correlation* ρ *the following result holds:*

$$\mathrm{E}\left[X_1 \mathbb{1}_{X_2>k}\right] = \left(\mu_1 + \rho\sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}\right)\left(1 - F_{X_2}(k)\right),$$

where $f_{\mathcal{N}(0,1)}(\cdot)$ *and* $F_{\mathcal{N}(0,1)}(\cdot)$ *are the standard normal PDF and CDF, respectively, and* $F_{X_2}(\cdot)$ *is the CDF corresponding to the random variable* X_2*.*

Proof. We start with writing

$$\mathrm{E}\left[X_1 \mathbb{1}_{X_2>k}\right] = \mathrm{E}\left[\mathrm{E}\left(X_1 \mathbb{1}_{X_2>k} \middle| X_2 > k\right)\right] = \mathrm{E}\left[\mathbb{1}_{X_2>k} \mathrm{E}\left(X_1 \middle| X_2 > k\right)\right]. \tag{B4}$$

For the inner expectation we set $X_2 = \sigma_2 Z_2 + \mu_2$ and $X_1 = \sigma_1\left(\rho Z_2 + \sqrt{1-\rho^2}Z_1\right) + \mu_1$, where Z_2 and Z_1 are independent standard normal random variables. The expression of X_1 in terms of Z_2 is established by assuming that X_1 and X_2 are bivariate normally distributed[1]. Straightforward calculus yields $\mathrm{E}\left(X_1 \middle| X_2 > k\right) = \rho\sigma_1 \mathrm{E}\left(Z_2 \middle| Z_2 > (k-\mu_2)/\sigma_2\right) + \mu_1$. The conditional expectation is given by $\mathrm{E}\left(Z_2 \middle| Z_2 > (k-\mu_2)/\sigma_2\right) = \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1-F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}$. So

$$\mathrm{E}\left(X_1 \middle| X_2 > k\right) = \rho\sigma_1 \frac{f_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)}{1 - F_{\mathcal{N}(0,1)}\left(\frac{k-\mu_2}{\sigma_2}\right)} + \mu_1.$$

Substituting this result in (B4) yields the result in the lemma. □

[1]The joint distribution of two normal random variables does not need to be bivariate normal. Only the reverse holds in general.