Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 21/09

Mike Felpel, Jörg Kienitz and Thomas A. McWalter

Effective Markovian Projection: Application to CMS Spread Options and Mid-Curve Swaptions

March 8, 2021

http://www.imacm.uni-wuppertal.de

Effective Markovian Projection: Application to CMS Spread Options and Mid-Curve Swaptions

Mike Felpel
Fachbereich Mathematik und Naturwissenschaften, Bergische Universtät Wuppertal
(mike.felpel@uni-wuppertal.de)

Jörg Kienitz
African Institute of Financial Markets and Risk Management, University of Cape Town
Fachbereich Mathematik und Naturwissenschaften, Bergische Universtät Wuppertal
(joerg.kienitz@math.uni-wuppertal.de)

Thomas A. McWalter
African Institute of Financial Markets and Risk Management, University of Cape Town
Department of Statistics, University of Johannesburg
(tom@analytical.co.za)

March 3, 2021

Abstract

Pricing of interest rate derivatives, such as CMS spread or mid-curve options, depends on the modelling of the underlying single rates. For flexibility and realism, these rates are often described in the framework of stochastic volatility models. In this paper we allow rates to be modelled within a class of general stochastic volatility models, which includes SABR, ZABR, free SABR and Heston models. We provide a versatile technique called Effective Markovian Projection, which allows a tractable model to be found that mimics the distribution of the more complex models used to price multi-rate derivatives. Three different numerical approaches are outlined and applied to relevant examples from practice. Finally, a new method that involves moment-matching of Johnson distributions is applied to facilitate closed form pricing formulas.

Keywords: Stochastic volatility, SABR, ZABR, Markovian Projection, Effective PDE, Approximation formula

Contents

1 Introduction

Markovian projection was introduced to the area of quantitative finance by [2, 27], where it was originally applied to produce closed-form approximations for European option prices on basket models for stocks, multi-factor interest rate models and hybrid interest rate/FX models. The method has since been extended and refined. The Heston stochastic volatility model has been considered by [6], while making the assumption of zero correlation between the Brownian motions driving the asset and the volatility processed. The displaced diffusion model was considered by [7], and the SABR-LIBOR model by [30]. Other works have investigated the standard SABR model [25, 26].

In the present paper we use results from our recent work on general stochastic volatility models [16] to extend the scope of application for Markovian projection. In particular, our novel approach allows pricing (and calibration) of multi-asset options using a large range of stochastic volatility models with asset dynamics specified by

$$\begin{cases} dF_t = v_t C(F_t)\, dW_t^{(1)}, & F_{t_0} = f, \\ dv_t = \mu(v_t)\, dt + \nu(v_t)\, dW_t^{(2)}, & v_{t_0} = \alpha, \\ \text{with} \quad d\langle W^{(1)}, W^{(2)} \rangle_t = \rho\, dt. \end{cases} \tag{1.1}$$

The model and parameters, as specified by the functions $C(\cdot)$, $\mu(\cdot)$ and $\nu(\cdot)$, are chosen to ensure the best fit to the current (discrete) market implied volatility surface and provide dynamics that are suitable for risk management and hedging of exotic contracts. The generality of the specification allows the necessary flexibility and control to ensure good calibration under varying market conditions. In our applications, we focus on multi-rate interest rate derivatives and consider CMS spread and mid-curve options as primary examples.

In summary, our research objectives are threefold: We provide a new Markovian projection technique, showcase applications on interest rate derivatives with multiple underlying rates and provide new results on moment-matching techniques outlining their application on general stochastic volatility models.

Our approach, to which we give the name *Effective Markovian Projection*, allows the projection of any general stochastic volatility model onto another model within the class specified by (1.1). This entails identifying the stochastic volatility model that, in some sense, best approximates the original dynamics. In [16] we derived an effective partial differential equation (PDE) for the dynamics specified by (1.1). This PDE may be interpreted in terms of a (parametric) local volatility model [12] using Markovian projection [18]. Owing to a common underlying structure of

the effective PDE, it is possible to find parametrizations for specific stochastic volatility models that provide good approximations to the general model.

When considering the application of Markovian projection to multi-factor baskets, we choose the reference model in such a way that it allows the effective modelling of general basket dynamics. To this end, we focus on the normal SABR model as our reference model since it can be shown that under a suitable numeraire the basket dynamics of normal SABR models is again a normal SABR model [21,22]. Having specified how to proceed in terms of the projection and the choice of the reference model, we provide numerical recipes for applying the method. Furthermore, to contextualise our developments, we first briefly introduce the classic application of Markovian projection.

The sequel consists of six sections. Section 2 provides the main theoretical results. After briefly summarizing the theoretical background of Markovian projection [18], we describe two novel applications. Initially, the classical approach is considered in the setting of general SABR dynamics, which slightly generalizes the results of [26]. We then introduce Effective Markovian Projection, which is based on effective probability theory [16, 20]. The numerical methods used for computing the projection are also introduced, in particular *ATM-Matching*, *Minimal Point Matching* and *N-Point Matching*. It should be emphasised that Effective Markovian Projection is directly applicable in the full general stochastic volatility setting. This is in contrast to classical Markovian projection, which must be tailored to the specific process, as seen for the General SABR model. It is known that this may produce approximations for volatility dynamics that are crude for long-dated instruments [26]. Section 3, focuses on examples from the interest rate derivatives world. Fixing notation and definitions, we consider CMS spread and mid-curve options as our main examples. With the underlying definitions at hand we consider a ZABR-type model as our base model and the normal SABR model as our reference model in Section 4, where we derive the corresponding projection and basket (spread) dynamics. Here we also specifically address numeraire issues. The resulting model amalgamates all the features of the base model and parametrizes the implied volatility surface in terms of a normal SABR model for each maturity, enabling the pricing of CMS spread and mid-curve options. Section 5 derives a method based on moment matching applied to the model class. Moment matching is applied to both NIG distributions and the family of Johnson distributions. The former approach has been considered previously by [9,14], and we show that there are some limitations in its application. The approach based on Johnson distributions provides a more stable and accurate result. The close connection of certain Johnson distributions to distributions arising from a SABR model has previously been recognized by [10]. Finally, Section 6 provides numerical illustrations of all

the methods considered in the previous sections. Here we focus exclusively on effective Markovian projection, since the classic approach has already been successfully applied to SABR models, see [26]. In particular, we demonstrate the additional flexibility in modelling and calibrating underlying rates in a multi-rate framework when considering base models of the ZABR type. We also illustrate the effects of changing model parameters. Section 7 concludes by summarizing the results and giving prospects for future research.

2 Effective Markovian Projection

In this section we present our main result and introduce a new approximation technique called Effective Markovian Projection (EMP). The idea underlying this technique is to combine the approximation results of singular perturbation techniques (see e.g. [16, 23]) with the general approximation technique of Markovian projection (see e.g. [3, 18]).

2.1 Markovian Projection

Before introducing EMP, let us first recall classical Markovian projection, which is based on the results in [3, 18]. Consider a stochastic process

$$dY_t = \alpha(t, Y_t)\, dt + \beta(t, Y_t)\, dW_t,$$

and assume that it has enough regularity to fit within the framework described in [3]. Here, the stochastic processes $\alpha(t, Y_t)$ and $\beta(t, Y_t)$ may depend on adapted Brownian motions other than, and possibly correlated with, W_t. It has been shown that there exists an SDE

$$dx_t = a(t, x_t)\, dt + b(t, x_t)\, dW_t,$$

with non-random coefficients and a solution x_t having the same one-dimensional probability distribution as Y_t, with the coefficients a and b satisfying

$$a(t, x) = \mathbb{E}[\alpha(t, Y_t)|Y_t = x] \qquad \text{and} \qquad b(t, x)^2 = \mathbb{E}[\beta(t, Y_t)^2|Y_t = x].$$

This existence result makes it possible to price European options by considering a one-dimensional local volatility model [11, 12]. Using the notation of the latter reference and considering an SDE of a forward price, i.e., $\alpha(t, Y_t) = 0$, the result states that

$$\sigma^2_{\text{local}}(t, x) = \mathbb{E}[\beta(t, Y_t)^2|Y_t = x]. \tag{2.1}$$

This projection onto a local volatility model, given in (2.1), is known as Markovian projection. The main challenge for practical applications then lies in the proper

evaluation of the conditional expectation. Possible approaches are based on Gaussian or least-square approximations, see [3]. We call the approach using these approximation techniques classical Markovian projection (CMP). Here we propose an alternative approach based on effective probability theory, see e.g. [16,20], which allows fast and accurate approximation. To better highlight the structural differences of our new approach, we first demonstrate an application of CMP to spread options under the SABR model.

2.2 Classic Markovian Projection for Generalized SABR

Using CMP, we derive a projection for the difference of a pair of general SABR stochastic volatility processes, which allows the modelling of spread options. The general SABR stochastic volatility processes, indexed by $i, j \in \{1, 2\}$, are governed by SDEs of the form:

$$\begin{cases} dF_{i,t} = v_{i,t} C_i(F_{i,t})\, dW_{i,t}^{(1)}, & F_{i,t_0} = f_i, \\ dv_{i,t} = \nu_i v_{i,t}\, dW_{i,t}^{(2)}, & v_{i,t_0} = \alpha_i, \\ \text{with} \quad d\langle W_i^{(1)}, W_j^{(2)} \rangle_t = \gamma_{ij}\, dt, & \\ \qquad\quad d\langle W_1^{(1)}, W_2^{(1)} \rangle_t = \rho\, dt, & \\ \text{and} \quad d\langle W_1^{(2)}, W_2^{(2)} \rangle_t = \xi\, dt, & \end{cases} \tag{2.2}$$

where C_i are suitably chosen functions. For example, the function may be chosen as the classical SABR backbone $C_i(x) = x^{\beta_i}$, for constants β_i.

While we do not explore it further here, a similar projection may be applied to d-dimensional basket options, where each of the individual processes have dynamics of the form (2.2). The d-dimensional basket would have $2d$ driving Brownian motions $W_{i,t}^{(j)}$, with $i = 1, \ldots, d$ and $j = 1, 2$, and it would have $(d^2 - d)/2$ correlations. As above, $W_{i,t}^{(1)}$ is the Brownian motion associated with the i-th forward and $W_{i,t}^{(2)}$ is associated with its stochastic volatility process.

In the case of a two-dimensional spread considered here, i.e.,

$$F_t = F_{1,t} - F_{2,t},$$

the spread is assumed to be governed by combined dynamics of the form:

$$\begin{cases} dF_t = u_t C(F_t)\, dB_t^{(1)}, & F_{t_0} = f, \\ du_t = \nu u_t\, dB_t^{(2)}, & u_{t_0} = 1, \\ \text{with} \quad d\langle B^{(1)}, B^{(2)} \rangle_t = \gamma dt, & \end{cases} \tag{2.3}$$

with the original stochastic volatility processes $v_{i,t}$ scaled by their initial values, $u_{i,t} = v_{i,t}/\alpha_i$, and a suitably specified function C. For example, we may choose $C(x) = p + q(x-a) + r(x-b)^2$, with constants p, q, r, a and b. The case of $q = r = 0$ leads to a Gaussian process, $r = 0$ to a displaced diffusion process, and the general case to a quadratic local volatility backbone with log-normal stochastic volatility. We apply CMP to find the parameters that provide the best match to the basket having combined dynamics given by (2.3). This results in the correlated Brownian motions $B_t^{(1)}$ and $B_t^{(2)}$ specified as

$$\begin{aligned} dB_t^{(1)} &= \sigma(t)^{-1} \left(u_{1,t}\alpha_1 C_1(F_{1,t})\, dW_{1,t}^{(1)} - u_{2,t}\alpha_2 C_2(F_{2,t})\, dW_{2,t}^{(1)} \right) \\ dB_t^{(2)} &= (\nu p)^{-1} \left(p_1\nu_1\rho_1\, dW_{1,t}^{(2)} - p_2\nu_2\rho_2\, dW_{2,t}^{(2)} \right). \end{aligned} \tag{2.4}$$

For the case of a displaced diffusion model we use $C(x) = p + q(x-f)$ with the parameters in (2.3) given by

$$\begin{gathered} p^2 = p_1^2 + p_2^2 - 2\rho p_1 p_2, \quad p_i = v_{i,t_0} C_i(f_i), \quad q_i = v_{i,t_0} C_i'(f_i), \\ \rho_1 = (p_1 - p_2\rho)/p, \rho_2 = (p_1\rho - p_2)/p, \\ q = (p_1^2 q_1\rho_1 + p_2^2 q_2\rho_2 - p_1 p_2\rho q_1\rho_1 - p_1 p_2\rho q_2\rho_2)/p^2 \\ \nu = \sqrt{(p_1\nu_1\rho_1)^2 + (p_2\nu_2\rho_2)^2 - 2\xi p_1 p_2\nu_1\nu_2\rho_1\rho_2}/p \qquad \text{and} \\ \gamma = \left(p_1^2\nu_1\rho_1\gamma_{11} + p_2^2\nu_2\rho_2\gamma_{22} - p_1 p_2\nu_1\rho_1\gamma_{12} - p_1 p_2\nu_2\rho_2\gamma_{21}\right)/(\nu p^2). \end{gathered}$$

The parameters are derived by approximating the variances of the spread as

$$\sigma^2(t) = u_{1,t}^2\alpha_1^2 C_1(F_{1,t})^2 + u_{2,t}^2\alpha_2^2 C_2(F_{2,t})^2 - 2\rho u_{1,t}u_{2,t}\alpha_1 C_1(F_{1,t})\alpha_2 C_2(F_{2,t}) \tag{2.5}$$

$$\text{and} \qquad u_t^2 = (p_1^2 u_{1,t}^2 + p_2^2 u_{2,t}^2 - 2\rho p_1 p_2 u_{1,t}u_{2,t})/p^2. \tag{2.6}$$

Some aspects of the derivation can be found in A.1 and the references cited therein. In particular for numerical examples we refer to [26]. Finally, we emphasise that the approximations presented are heavily dependent on the underlying models and thus only valid for the case of general SABR models.

2.3 Effective Markovian Projection for Generalized Stochastic Volatility

To motivate our new approximation technique, we now go a step further and relax the previous restriction of general SABR models and instead consider general

stochastic volatility models of the form

$$\begin{cases} dF_t = v_t C(F_t)\, dW_t^{(1)}, & F_{t_0} = f, \\ dv_t = \mu(v_t)\, dt + \nu(v_t)\, dW_t^{(2)}, & v_{t_0} = \alpha, \\ \text{with} \quad d\langle W^{(1)}, W^{(2)} \rangle_t = \rho\, dt, \end{cases} \tag{2.7}$$

which satisfy the assumptions of Theorem 1 of [16]. For convenience these assumptions can be found in Appendix A.2.

With this setup, we provide the central result of Effective Markovian Projection.

Proposition 2.1. *Given a general stochastic volatility model (2.7), satisfying the assumptions of Theorem 1 of [16], the local volatility function (2.1) is approximated as*

$$\sigma_{local}^2(t,x) \approx C(x)^2 a(t)^2 e^{G(t)} \big(1 + 2b(t)z(x) + c(t)z(x)^2\big), \tag{2.8}$$

where the coefficients a, b, c, z and G are specified in Appendix A.2.

We call this projection from a general stochastic volatility model onto a local volatility model the ***Effective Markovian Projection*** *(EMP).*

Proof. Let $p(t_0, f, \alpha, t, F, A)$ be the probability density function that $F_t = F$ and $v_t = A$ at time t, given that $F_{t_0} = f$ and $v_{t_0} = \alpha$ at time t_0. With this we can describe the conditional probability corresponding to (2.1) as

$$\begin{aligned} \sigma_{\text{local}}^2(t,x) = C(x)^2 \mathbb{E}\big[v_t^2 \big| F_t = x\big] &= C(x)^2 \frac{\mathbb{E}\big[v_t^2 \mathbb{I}_{\{F_t = x\}}\big]}{\mathbb{E}\big[\mathbb{I}_{\{F_t = x\}}\big]} \\ &= C(x)^2 \frac{\int_0^\infty A^2 p(t_0, f, \alpha, t, x, A)\, dA}{\int_0^\infty p(t_0, f, \alpha, t, x, A)\, dA} \\ &= C(x)^2 \frac{Q^{(2)}(t_0, f, \alpha, t, x)}{Q^{(0)}(t_0, f, \alpha, t, x)}, \end{aligned}$$

where the function $Q^{(k)}$ is defined by

$$Q^{(k)}(t_0, f, \alpha, t, F) := \int_0^\infty A^k p(t_0, f, \alpha, t, F, A)\, dA.$$

Applying Theorem 1 of [16] we have

$$Q^{(2)}(t_0, f, \alpha, t, x) \approx P(t,x) Q^{(0)}(t_0, f, \alpha, t, x),$$

with

$$P(t,x) = a(t)^2 e^{G(t)} \big(1 + 2b(t)z(x) + c(t)z(x)^2\big), \tag{2.9}$$

where the coefficients a, b, c, z and G are specified in [16] and provided in Appendix A.2. □

In the context of singular perturbation theory, this approximation is accurate to $\mathcal{O}(\varepsilon^2)$, which is the same order of accuracy provided by the SABR implied volatility formula (see [20]). To conclude this section, we provide further remarks on the applicability of the technique.

Remark 2.2. Although we have formulated EMP in terms of a stochastic volatility model, it may also be applied to a stochastic variance model (e.g. the Heston model). In this case the local volatility function is expressed as

$$\sigma^2_{\text{local}}(t,x) = C(x)^2 \frac{Q^{(1)}(t_0,f,\alpha,t,x)}{Q^{(0)}(t_0,f,\alpha,t,x)}.$$

As shown in [23], in this setting an analogous representation of

$$Q^{(1)} = P(t,x)Q^{(0)}(t_0,f,\alpha,t,x)$$

is applicable, yielding the representation of the local volatility function as in Proposition 2.1.

Remark 2.3. We stress that the backbone function $C(x)$ need not be restricted to a parametric function and can be chosen to be non-parametric. This allows application of EMP to stochastic local volatility models (see [1, 17, 28] for a non-exhaustive list), which we are currently investigating.

Remark 2.4. To guarantee that the resulting model is arbitrage-free, it is enough to ensure that the local volatility function remains positive. Assuming that the backbone function is chosen such that $C > 0$ then one requires the condition that $P > 0$.

2.4 Projection onto another Model

Using EMP we wish to project one model onto another. This is useful because some models are more tractable than others, for example possessing analytical solutions for various contingent claim prices. The idea is to consider matching the corresponding projected local volatility functions of both models. To this end, let us consider two local volatility functions σ^2_{local}, $\tilde{\sigma}^2_{\text{local}}$ of the form (2.8)

$$\begin{aligned}
\sigma^2_{\text{local}}(t,x) &= C(x)^2 a(t)^2 e^{G(t)}\big(1 + 2b(t)z(x) + c(t)z(x)^2\big)\\
\tilde{\sigma}^2_{\text{local}}(t,x) &= \tilde{C}(x)^2 \tilde{a}(t)^2 e^{\tilde{G}(t)}\big(1 + 2\tilde{b}(t)\tilde{z}(x) + \tilde{c}(t)\tilde{z}(x)^2\big).
\end{aligned}$$

Our goal is to specify the coefficients $\tilde{a}, \tilde{b}$ and $\tilde{c}$ in such a way that these two local volatility functions match each other as closely as possible. The other functions, e.g. $\tilde{C}$, are assumed to be model specific and are not modified. Note that a complete

matching would impose an x-dependence on these coefficients. This imposes a dependence of the forward dynamics on the volatility dynamics. Since in our later applications we aim to project a general SV model onto another SV model, we want to avoid this x-dependence at this point. Thus we focus on matching the local volatility function at specific points. To match three different coefficients, we need at least three different equations for the problem to be well posed. There is some freedom in choosing these equations and the choice can be individually tailored to the corresponding problem at hand. Some choices are presented in the following sections and are further examined for specific applications, which are presented in Section 6.

2.4.1 ATM Matching

The first approach is based exclusively on matching the at-the-money (ATM) initial point, $x = f$. Considering only one point we choose to match the value of the local volatility function in addition to the value of the first two derivatives at x. This corresponds to

$$\begin{aligned} \sigma^2_{\text{local}}(t,f) &= \tilde{\sigma}^2_{\text{local}}(t,f) \\ \partial_x \sigma^2_{\text{local}}(t,f) &= \partial_x \tilde{\sigma}^2_{\text{local}}(t,f) \\ \partial_{xx} \sigma^2_{\text{local}}(t,f) &= \partial_{xx} \tilde{\sigma}^2_{\text{local}}(t,f). \end{aligned} \tag{2.10}$$

Evaluating the first equation and noting that $z(f) = \tilde{z}(f) = 0$ we deduce that

$$\tilde{a}(t)^2 = \frac{C(f)^2}{\tilde{C}(f)^2} a(t)^2 e^{G(t)-\tilde{G}(t)}. \tag{2.11}$$

Evaluating the derivatives at $x = f$ then yields

$$\begin{aligned} \partial_x \sigma^2_{\text{local}}(t,f) &= 2\sigma^2_{\text{local}}(t,f) \left[\frac{C'(f)}{C(f)} + b(t) z'(f) \right] \\ \partial_{xx} \sigma^2_{\text{local}}(t,f) &= 2\sigma^2_{\text{local}}(t,f) \Bigg[\frac{C'(f)^2}{C(f)^2} + \frac{C''(f)}{C(f)} + 4\frac{C'(f)}{C(f)} b(t) z'(f) \\ &\qquad + b(t) z''(f) + c(t) z'(f)^2 \Bigg]. \end{aligned}$$

Taking into account that $z'(x) = \frac{1}{C(x)}$ we further simplify this to

$$\begin{aligned} \partial_x \sigma^2_{\text{local}}(t,f) &= 2\sigma^2_{\text{local}}(t,f) \left[\frac{1}{C(f)} \Big(C'(f) + b(t) \Big) \right] \\ \partial_{xx} \sigma^2_{\text{local}}(t,f) &= 2\sigma^2_{\text{local}}(t,f) \left[\frac{1}{C(f)^2} \Big(C'(f)^2 + C''(f) C(f) + 3C'(f) b(t) + c(t) \Big) \right]. \end{aligned} \tag{2.12}$$

Thus, matching (2.10) with the derivatives provided in (2.12), boils down to (2.11) and a matching of the bracketed terms. We shall refer to this algorithm as EMP-ATM matching.

2.4.2 Minimal Point Matching

The second matching approach we propose is the minimal point (EMP-MP) matching algorithm and is based on matching the local volatility functions of the two models at three distinct points. These three points are the minimal requirement for a well-posed specification for deducing the three different coefficients. When including in-the-money (ITM) and/or out-the-money (OTM) values, the method allows better fitting of the wings. As one of the matching points we take the ATM value, $x = f$, corresponding to the condition

$$\tilde{a}(t)^2 = \frac{C(f)^2}{\tilde{C}(f)^2} a(t)^2 e^{G(t)-\tilde{G}(t)}.$$

Inserting this into the local volatility functions, for each x we get the condition

$$2\tilde{b}(t)\tilde{z}(x) + \tilde{c}(t)\tilde{z}(x)^2 = \frac{C(x)^2\tilde{C}(f)^2}{\tilde{C}(x)^2 C(f)^2}\big(1 + 2b(t)z(x) + c(t)z(x)^2\big) - 1.$$

By selecting two additional points these equations allow an explicit solution where the new coefficients are given by

$$\begin{aligned}
\tilde{b}(t) &= \frac{1}{2}\frac{r(x_1)}{\tilde{z}(x_1)} - \frac{1}{2}\frac{r(x_2)\tilde{z}(x_1) - r(x_1)\tilde{z}(x_2)}{\tilde{z}(x_2)^2 - \tilde{z}(x_1)\tilde{z}(x_2)} \\
\tilde{c}(t) &= \frac{r(x_2)\tilde{z}(x_1) - r(x_1)\tilde{z}(x_2)}{\tilde{z}(x_1)\tilde{z}(x_2)^2 - \tilde{z}(x_1)^2\tilde{z}(x_2)},
\end{aligned}$$

with

$$r(x) = \frac{C(x)^2\tilde{C}(f)^2}{\tilde{C}(x)^2 C(f)^2}\big(1 + 2b(t)z(x) + c(t)z(x)^2\big) - 1.$$

2.4.3 N-Point Matching

The last matching algorithm we propose is based on numerical minimization. Specifying a set of more than three points, it is always possible to minimize an error function that quantifies the difference between the local volatility functions evaluated at these points. To mention a few possibilities, this includes error functions based on the absolute value, relative value or mean squared error. This procedure

may yield a better fit along the entire curve, but comes at the cost of requiring a minimization problem to be solved, since the coefficients are no longer explicit. Compared to a direct calibration of the underlying models, based, for example, on the implied volatility curve, this approach still yields computational advantages due to the simpler form of the local volatility function. We denote this approach the EMP-NP matching algorithm.

3 Interest Rate Derivatives

Having derived the general methodology we now demonstrate the applicability of the new approach by considering the pricing of semi-vanilla interest rate products, where the underlying swap rates are governed by more complex models of the ZABR type. We consider the displaced diffusion ZABR model (dZABR) as our primary example in this setting. The dZABR is specified by the following system of SDEs:

$$\begin{cases} dF_t = v_t(F_t + d)^\beta \, dW_t^{(1)}, & F_{t_0} = f, \\ dv_t = \nu v_t^\gamma \, dW_t^{(2)}, & v_{t_0} = \alpha, \\ \text{with} \quad d\langle W^{(1)}, W^{(2)} \rangle_t = \rho \, dt. \end{cases}$$

As far as the contingent claims is concerned, we focus our attention on options on CMS spreads and mid-curves. In the special case where the underlying rates are modelled using the normal SABR model (nSABR), options on CMS spreads have been intensively studied (see e.g. [21]). When the rates are governed by Black baskets the pricing of options on CMS spreads and mid-curves can be found in [5]. With our new approach we extend these pricing techniques to allow a general stochastic volatility model as the underlying model for the single rates. We demonstrate this in the case where each rate follows a dZABR under its own measure.

3.1 Interest Rate Notation

To introduce the previously mentioned pricing techniques, let us first specify the setting and corresponding notation. For the general tenor structure we consider some starting date T_1 and end date T_2. The payment dates in between are denoted by $T_1 < t_1, \ldots, t_n \leq T_2$. Furthermore, we denote the exercise date as T and assume the payment happens at time T_0. We use the notation $DF(T, T_0)$ to denote the discount factor from the payment date to the exercise date and define the annuity

factor $A(t, T_1, T_2)$ and the forward level $L(t, T_0, T_1, T_2)$ by

$$A(t, T_1, T_2) = \sum_{T_1 < t_n \leq T_2} \delta_n DF(t, t_n)$$
$$L(t, T_0, T_1, T_2) = \sum_{T_1 < t_n \leq T_2} \delta_n DF(t, T_0, t_n) = \frac{A(t, T_1, T_2)}{DF(t, T_0)},$$

respectively. Here $\delta_n = t_n - t_{n-1}$ denotes the day count fraction. Finally, the swap rates $R(t, T_1, T_2)$ are defined as

$$R(t, T_1, T_2) = \frac{DF(t, T_1) - DF(t, T_2)}{A(t, T_1, T_2)}. \tag{3.1}$$

Remark 3.1. Notice that the definition of the swap rates can be further generalized as in [21]. For our situation, however, this generalization is not necessary and we retain the notation used in [5].

3.2 Options on CMS Spreads

The first interest rate derivatives we consider are options on CMS spreads. To this end, we provide a short summary of the results of [21], where the goal is the evaluation of caplets, floorlets or swaplets on CMS spreads. We denote by $(R_i(t))_{i=1,2}$ the two swap rates $R(t, T_0, T_i)$. Moreover, we define the spread, $S(t)$, on the two rates as

$$S(t) = R_2(t) - R_1(t)$$

with initial value $s = r_2 - r_1$. Using a caplet as an example, the payoff function at the expiry date T is given by

$$V_{spread}(T; T, K) = DF(T, T_0)(S(T) - K)^+.$$

In turn, the value of the caplet at an earlier date is given as the conditional expectation under a suitable martingale measure. Under the forward measure, i.e., with the discount factor $DF(\cdot, T_0)$ as the numeraire, the value is expressed as

$$V_{spread}(t; T, K) = DF(t, T_0)\mathbb{E}^{T_0}\left[(S(T) - K)^+ \middle| t\right].$$

Without loss of generality we consider the value at the initial time, $t = 0$. To evaluate this expression, Hagan *et al.* introduced a suitable measure, H, under which

the spread process becomes a martingale[1]. By further analyzing the convexity correction the value of the caplet can be evaluated as (see e.g. [21, Equation 3.22])

$$V_{spread}(0;T,K) = DF(0,T_0)\big((1-(s-K)\lambda_s)C^S(0;T,K) + \lambda_s Q_C^S(0;T,K)\big). \quad (3.2)$$

Here the expressions C^S and Q_C^S denote the vanilla and quadratic calls of the spread under the martingale measure H, i.e.,

$$C^S(t;T,K) = \mathbb{E}^H\Big[(S(T)-K)^+\big|t\Big]$$
$$Q_C^S(t;T,K) = \mathbb{E}^H\Big[\big((S(T)-K)^+\big)^2\big|t\Big].$$

The coefficient λ_s denotes the convexity coefficient which, following the suggestion of Hagan *et al.* [21], is set in a manner consistent with the underlying swap rates

$$\lambda_s = \frac{\lambda_2 Q_S^2(0;T,r_2) - \lambda_1 Q_S^1(0;T,r_1)}{Q_S^S(0;T,s)}.$$

Here the functions Q_S^S and Q_S^i express the quadratic swaps

$$Q_S^S(t;T,K) = \mathbb{E}^H\Big[(S(T)-K)^2\big|t\Big]$$
$$Q_S^i(t;T,K) = \mathbb{E}^i\Big[(R_i(T)-K)^2\big|t\Big],$$

and the coefficients λ_i are the convexity coefficients of the underlying rates given by the approximation

$$\frac{L(0,T_0,T_0,T_i)}{L(T,T_0,T_0,T_i)} = \frac{L_i(0)}{L_i(T)} = 1 + \lambda_i(R_i(T)-R_i(0)) + \cdots. \quad (3.3)$$

Given the convexity coefficients and the explicit values of the vanilla and quadratic calls this provides a closed form pricing formula.

Remark 3.2. By construction, see [21], these formulae satisfy put-call parity.

[1] In the following applications the explicit form of the measure H is not of importance—only the dynamics of the spread measure expressed in this measure is relevant.

3.3 Options on Mid-curves

Another popular semi-vanilla product is the mid-curve call option. This option provides the holder with the right to enter into a swap with starting time T_1 and maturity T_2 at some exercise date T. As shown in [5, 15], the value of the mid-curve call option can be expressed as a call option on the mid-curve rate, $R_{mc}(T) = R(T, T_1, T_2)$, under the annuity measure, which uses numeraire $A_{mc}(T) = A(T, T_1, T_2)$. At time 0, this yields the value

$$V_{mc}(0, T, K) = A_{mc}(0)\mathbb{E}^{A_{mc}}\Big[(R_{mc}(T) - K)^+\Big]. \tag{3.4}$$

Note that by definition the mid-curve rate, $R_{mc}(T)$, is a martingale under the chosen annuity measure. To evaluate the expectation we do not assume an additional model for the mid-curve rate, but instead focus on the choice of the underlying swap rates. Therefore, in the rest of this section, we follow the approach presented in [5] and express the mid-curve rate as a weighted spread on swap rates. We start by considering the rates $R_i(T)$ and deduce that

$$R_{mc}(T) = R_2(T)\frac{A_2(T)}{A_{mc}(T)} - R_1(T)\frac{A_1(T)}{A_{mc}(T)}.$$

Defining new modified rates $\hat{R}_i(T)$ as

$$\hat{R}_i(T) = \frac{R_i(T)}{\hat{M}_i(T)} = \frac{R_i(T)}{\frac{A_{mc}(T)A_i(0)}{A_i(T)A_{mc}(0)}},$$

the mid-curve rate becomes

$$R_{mc}(T) = \hat{R}_2(T)\frac{A_2(0)}{A_{mc}(0)} - \hat{R}_1(T)\frac{A_1(0)}{A_{mc}(0)}.$$

At this point we highlight an important structural difference in comparison to the CMS spread options presented previously in Section 3.2, namely that under the chosen annuity measure the mid-curve rate, $R_{mc}(T)$, as well as the rates, $\hat{R}_i(T)$, are martingales at the same time. To evaluate the CMS spread options we worked under a measure where only the spread is a martingale, here we have a measure where the weighted spread and all the single rates are simultaneously martingales.

To characterize these new modified rates, $\hat{R}_i(T)$, we assume that the martingales $\hat{M}_i(T)$ are approximated by

$$\hat{M}_i(T) = \frac{A_{mc}(T)A_i(0)}{A_i(T)A_{mc}(0)} \approx 1 + \hat{\lambda}_i(R_i(T) - R_i(0)) + \cdots,$$

and, as demonstrated in [5], this allows us to recover the distributions of the modified rates from the original ones by considering approximate option prices using the relation

$$\mathbb{E}^{A_{mc}}\left[(\hat{R}_i(T)-K)^+\right] \approx \mathbb{E}^{A_i}\left[\left(R_i(T)(1-\hat{\lambda}_i K)-K(1-\hat{\lambda}_i r_i)\right)^+\right]. \tag{3.5}$$

4 Evaluation under ZABR-type Models

Having described the general framework for pricing interest rate derivatives, we now show how to apply EMP to a general stochastic volatility model for the underlying swap rates. As mentioned previously, we consider the dZABR model of [16] as our primary model.

4.1 Projection onto a Normal SABR Model

To provide analytical tractability, we consider a projection of the underlying dZABR model onto the nSABR model. To apply EMP, we first compute the corresponding coefficients for the function $P(t,x)$ specified in (2.9). For the dZABR model with $\beta \in (0,1)$ the coefficients, given in [16], are

$$\begin{aligned}
a(t) &= \alpha \\
b(t) &= \rho\nu\alpha^{\gamma-2} \\
c(t) &= \nu^2\alpha^{2(\gamma-2)}(1+(\gamma-1)\rho^2) \\
G(t) &= -\rho^2\nu^2\alpha^{2(\gamma-1)}(\gamma-1)t - \rho\nu\alpha^{\gamma}\beta(f+d)^{\beta-1}t,
\end{aligned}$$

and the transformed variable is

$$z(x) = \frac{1}{1-\beta}((x+d)^{1-\beta}-(f+d)^{1-\beta}).$$

Remark 4.1. Note that we have chosen the dZABR model as our primary example to ensure that notation remains relatively simple. More complicated examples such as the mean-reverting ZABR or Heston models can be considered, using the corresponding coefficients given, for example, in [16] and [23].

For the nSABR model these coefficients may be simplified further leading to

$$\tilde{\sigma}^2_{\text{local}}(t,x) = \tilde{\alpha}^2\left(1+2\frac{\tilde{\rho}\tilde{\nu}}{\tilde{\alpha}}\tilde{z}(x)+\frac{\tilde{\nu}^2}{\tilde{\alpha}^2}\tilde{z}(x)^2\right)$$

with $\tilde{z}(x) = x - f$.

Given these coefficients, we, in turn, can evaluate the local volatility function induced by EMP. We may now use any of the three projection procedures described in Section 2. For example, using the parameters $f = 0.005$, $\alpha = 0.3f^{1-\beta}$, $d = 0.002$, $\nu = 0.3$, $\rho = -0.3$, $T = 5$, $\gamma = 0.8$ and $\beta = 0.4$ we compute the local volatility function generated by each of the different projection procedures. The results are shown in Figure 1, and, as can be seen, yield good results in each case.

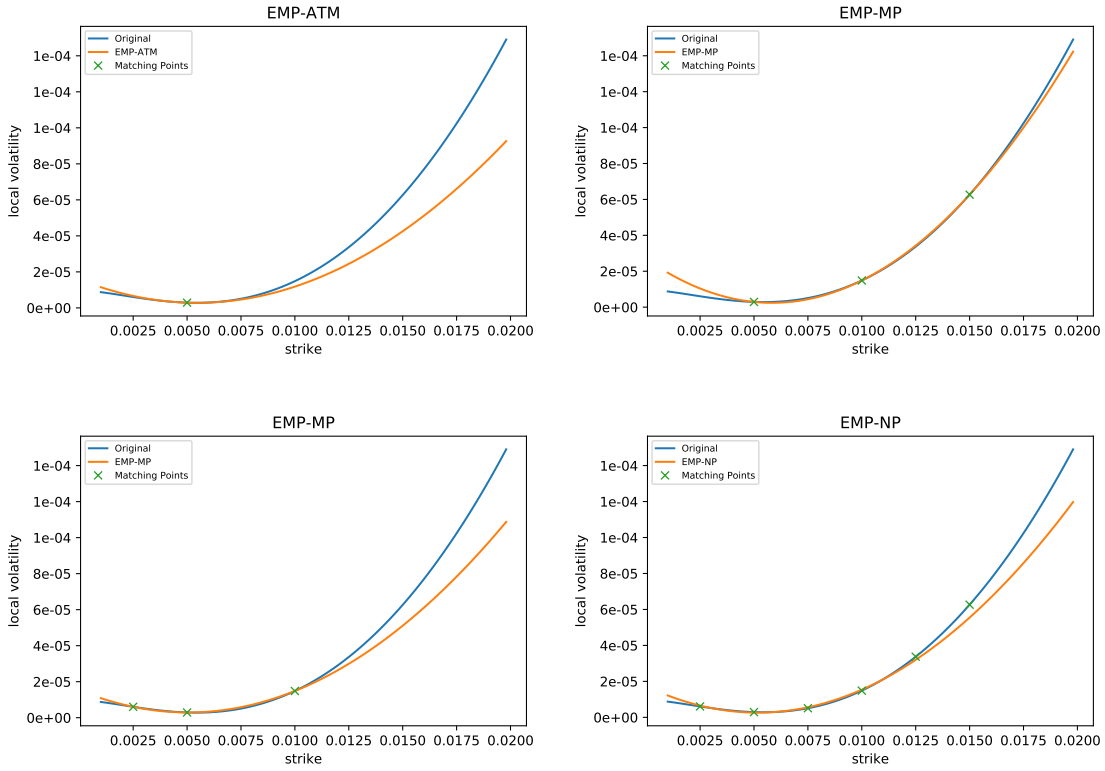

Figure 1: Local volatility function for the parameter set $f = 0.005, \alpha = 0.3f^{1-\beta}, d = 0.002, \nu = 0.3, \rho = -0.3, T = 5$, $\gamma = 0.8$ and $\beta = 0.4$. Green crosses indicate matching points.

4.2 Options on CMS Spreads

Given the pricing formula for options on CMS spreads in (3.2), we now consider valuation of vanilla and quadratic calls on the spread and swap rates. This depends heavily on the model chosen for the single rates. Assuming that each swap rate is modelled using dZABR, the call options can be valued using the methods described in [4, 16]. However, it is not clear how to value options on the spread process since the underlying model is not known. Therefore, in the explicit case, where each

rate follows an nSABR model under its respective measure, Hagan *et al.* [21, 22] showed that it is possible to approximate the dynamics of the spread using an nSABR model.[2] To achieve this, singular perturbation techniques were used and the resulting approximation has the same order of accuracy as the classical SABR implied volatility formulas. Applying these results and the following steps, we can evaluate CMS spread options using dZABR models for each of the rates:

(i) Given that each swap rate $R_i(t)$ is modeled using a dZABR model, we use EMP to deduce approximate rates $\tilde{R}_i(t)$ characterized in terms of nSABR models.

(ii) We approximate the spread, $S(t)$, on the swap rates as

$$S(t) \approx \tilde{S}(t) = \tilde{R}_1(t) - \tilde{R}_2(t).$$

Using the results in [21, 22] this allows for an approximation of the spread in terms of an nSABR model.

(iii) We evaluate the corresponding vanilla and quadratic calls using the explicit formulas presented in [21] and can, in turn, evaluate the value of the option on the CMS spread.

4.3 Options on Mid-Curves

To deduce the price of a call option on a mid-curve, we proceed in a fashion similar to that as above. In Section 3.3 it was shown that the underlying mid-curve rate can be expressed as a weighted spread of two modified rates, $\hat{R}_i(T)$. From an analytical perspective we cannot deduce a general dynamic of the rates $\hat{R}_i(T)$. We can, however, resort to numerical means in the spirit of [5]. Given that the underlying rates $R_i(T)$ follow a dZABR model under their own measure, we assume that the modified rates $\hat{R}_i(T)$ are governed by a dZABR model under the annuity measure A_{mc}. Using the parameters of the rates $R_i(T)$ we can then calibrate the model parameters of the rates $\hat{R}_i(T)$ using (3.5). Since we know that the rates $\hat{R}_i(T)$ are martingales we approximate them using EMP onto corresponding nSABR models. Given that the original weighted sum of the rates, $\hat{R}_i(T)$, is also a martingale under the A_{mc}-measure, we can use the more general basket result of Hagan *et al.* [21, 22] to deduce that the mid-curve rate, $R_{mc}(T)$, is approximated by an nSABR model. This reduces the evaluation of the option on mid-curves to a vanilla call option under an nSABR model.

[2] This actually also holds true when we replace the spread by a basket of nSABR models

5 Moment Matching

We now explore a second application of EMP in which we derive explicit approximating formulas for the density and characteristic functions of a general stochastic volatility model. Once again, the key step in this approach is the projection of the model onto an nSABR model. Thereafter we use approximation techniques based on a moment matching algorithm.

Concentrating on the nSABR for the moment, we fix a maturity and use explicit formulas for the first four moments available. As was demonstrated in [9] or [10], these moments can be characterized in their standardized forms, i.e., by the mean $\mathcal{M}$, variance $\mathcal{V}$, skew $\mathcal{S}$ and excess kurtosis $\mathcal{K}$ as

$$\begin{aligned}
\mathcal{M} &= \mathbb{E}\big[F_t\big] = f \\
\mathcal{V} &= \mu_2 = \frac{\tilde{\alpha}^2}{\tilde{\nu}^2}(x-1) \\
\mathcal{S} &= \frac{\mu_3}{\mu_2^{3/2}} = \tilde{\rho}(x+2)\sqrt{x-1} \\
\mathcal{K} &= \frac{\mu_4}{\mu_2^2} - 3 = (x-1)\left(\frac{1+4\tilde{\rho}^2}{5}(x^3+3x^2+6x+5)+1\right),
\end{aligned}$$

where

$$\mu_i = \mathbb{E}\big[(F_t - f)^i\big] \qquad \text{and} \qquad x = e^{\tilde{\nu}^2 t}.$$

Having these moments available, we proceed in the spirit of [9] and [29], and consider an approximation of this distribution using a parametric distribution of a specific form. In [9] this was demonstrated in the case where the approximating parametric distribution was the normal-inverse Gaussian (NIG) distribution. This distribution has four underlying parameters and explicit formulas for its moments. It is, therefore, possible to set up a well posed system of equations to determine the underlying parameters by matching the moments of the distributions. Corresponding algorithms to determine the parameters of the NIG distribution based on the nSABR moments can be found in [9] or [14].

Instead of considering the NIG distribution, we will use the Johnson's-S_U distribution (see [24]). Johnson's-S_U distribution defines a random variable, X, through a transformation of a standard normal random variable, Z, given by

$$X = \xi + \lambda \sinh\left(\frac{Z-\gamma}{\delta}\right).$$

The probability density function of this random variable, ϕ_S, is specified in terms

of the standard normal density function, ϕ_N, as

$$\phi_S(x) = \frac{\delta}{\lambda\sqrt{\left(\frac{x-\xi}{\lambda}\right)^2 + 1}} \phi_N\left(\gamma + \delta \sinh^{-1}\left(\frac{x-\xi}{\lambda}\right)\right).$$

As outlined in [10], both the distributions generated by the nSABR model, as well as Johnson's-S_U distribution can be seen as special cases of the hyperbolic normal stochastic volatility model. This underscores the similarity of the distributions and motivates our choice. As in the case of the NIG distribution, the Johnson's-S_U distribution has four underlying parameters and explicit formulas for the first four moments, see, e.g., [10] or [31]. This allows the use of the matching algorithm proposed by Tuenter [31] to determine the underlying parameters of the distribution. An outline of the algorithm is given in Algorithm 1, for all technical details we refer to the original paper [31].

Algorithm 1: Matching algorithm proposed by Tuenter [31].

Data: Moments $\mathcal{M}$, $\mathcal{V}$, $\mathcal{S}$, $\mathcal{K}$
Result: Johnson's-S_U distribution parameters δ, γ, λ, ξ
set $\beta_1 = \mathcal{S}^2$ and $\beta_2 = \mathcal{K} + 3$;
define $m(\omega) = -2 + \sqrt{4 + 2(\omega^2 - \frac{\beta_2+3}{\omega^2+2\omega+3})}$;
define $f(\omega) = (\omega - 1 - m(\omega))(\omega + 2 + 0.5m(\omega))^2$;
compute boundaries (ω_1, ω_2) using procedure in [31];
if $f(\omega_1) \leq \beta_1$ **then**
  break;
end
solve for ω^* s.t. $f(\omega^*) = \beta_1$;
set $m = m(\omega^*)$ and $\Omega = -\text{sgn}(\mathcal{S}) \sinh^{-1}(\sqrt{\frac{\omega^*+1}{2\omega^*}(\frac{\omega^*-1}{m} - 1)})$;
set $\delta = \frac{1}{\log(\omega^*)}$, $\gamma = \frac{\Omega}{\log(\omega^*)}$, $\lambda = \frac{\sqrt{\mathcal{V}}}{\omega^*-1}\sqrt{\frac{2m}{\omega^*+1}}$ and
$\xi = \mathcal{M} - \text{sgn}(\mathcal{S})\frac{\sqrt{\mathcal{V}}}{\omega^*-1}\sqrt{\omega^* - 1 - m}$;

Having determined the underlying parameters, we now have a completely specified probability distribution with explicit formulas for the density function. In the case of the NIG distribution, explicit formulas for the characteristic function are also available and allow the application of popular computation techniques using the fast Fourier transform (FFT) [8].

To extend these results to more complex models, such as dZABR, the following steps may be used:

(i) Using EMP, project the dZABR model onto the nSABR model.

(ii) Using the moment matching algorithm, fit a Johnson's-S_U or NIG distribution to the nSABR model.

(iii) Using the Johnson's-S_U or NIG distribution, explicit formulas for the density function and/or the characteristic function are available.

This procedure allows us to describe the distribution of the dZABR model at a fixed maturity by an explicit analytical formula.

6 Numerical Results

In order to evaluate the quality of our new approximation technique we perform a series of numerical experiments. We start by considering the various matching algorithms described in Section 2.4 and investigate their ability to fit the local volatility function. Thereafter, we explore the implications on the implied volatility surface and demonstrate the flexibility obtained when pricing various interest rate derivatives. We conclude by analysing the accuracy of the distribution matching approach proposed in Section 5.

For most of our examples we shall consider dZABR models specified using the parameter sets listed in Table 1. In the case where $\gamma = 1$ these models reduce to dSABR models. In this simpler setting, the main difference, when compared with the nSABR model, is the additional parameter β. We, therefore, highlight the dependence on this parameter and consider various values for it.

Parameter	Set 1	Set 2	Set 3	Set 4
f	0.005	0.005	0.005	0.005
α	$0.3f^{1-\beta}$	$0.3f^{1-\beta}$	$0.3f^{1-\beta}$	$0.3f^{1-\beta}$
d	0.002	0.002	0.002	0.002
ν	0.3	0.3	0.3	0.5
ρ	-0.3	-0.7	0	-0.3
T	5	5	10	1
γ	0.8	0.9	0.8	0.9

Table 1: Parameter values for dZABR models used in numerical experiments.

6.1 Local Volatility

To apply the different matching algorithms described in Section 2.4 we require additional parameters to be specified. In particular, other than the at-the-money value, two additional matching points are required to implement the EMP-MP algorithm. To demonstrate the influence of these points on the projection, we evaluate the local volatility function for the dZABR model using parameter set 1 with $\beta = 0.2$ and $\beta = 0.8$, and generate the curve using the EMP-MP algorithm for different values of the matching points. The errors observed between the original model and the approximating curves are presented in Tables 2 and 3.

x	dZABR	$x_1{=}0.5f$ $x_2{=}2f$	$x_1{=}0.5f$ $x_2{=}3f$	$x_1{=}2f$ $x_2{=}3f$
$0.5f$	$8.48e^{-6}$	0	0	$3.08e^{-6}$
$0.75f$	$4.57e^{-6}$	$-2.45e^{-7}$	$-3.48e^{-7}$	$1.04e^{-6}$
f	$2.58e^{-6}$	0	0	0
$1.5f$	$5.73e^{-6}$	$5.92e^{-7}$	$1.41e^{-6}$	$-4.36e^{-7}$
$2f$	$1.97e^{-5}$	0	$2.47e^{-6}$	0
$3f$	$8.37e^{-5}$	$-8.22e^{-6}$	0	0
$4f$	$1.99e^{-4}$	$-2.90e^{-5}$	$-1.17e^{-5}$	$-4.32e^{-6}$

Table 2: Values of the local volatility function for the dZABR model using parameter set 1 with $\beta = 0.2$, and the error in the EMP-MP local volatility function for different values of the matching points, x_1 and x_2.

x	dZABR	$x_1{=}0.5f$ $x_2{=}2f$	$x_1{=}0.5f$ $x_2{=}3f$	$x_1{=}2f$ $x_2{=}3f$
$0.5f$	$3.56e^{-6}$	0	0	$4.10e^{-6}$
$0.75f$	$3.55e^{-6}$	$-2.18e^{-7}$	$-3.55e^{-7}$	$1.49e^{-6}$
f	$3.75e^{-6}$	0	0	0
$1.5f$	$5.85e^{-6}$	$6.35e^{-7}$	$1.73e^{-6}$	$-7.33e^{-7}$
$2f$	$1.18e^{-5}$	0	$3.28e^{-6}$	0
$3f$	$4.09e^{-5}$	$-1.09e^{-5}$	0	0
$4f$	$1.01e^{-4}$	$-4.24e^{-5}$	$-1.94e^{-5}$	$-9.58e^{-6}$

Table 3: Values of the local volatility function for the dZABR model using parameter set 1 with $\beta = 0.8$, and the error in the EMP-MP local volatility function for different values of the matching points, x_1 and x_2.

For this example we observe a reasonably good fit in all cases. Here the match including the point $x_1 = 0.5f$ provides the best fit for the left wing, whereas the

match including the additional points $x_1 = 2f$ and $x_2 = 3f$ provides an excellent fit for the longer right wing. The latter also provides a reasonably good fit for the left wing. Thus, from now on, when we refer to the EMP-MP algorithm we shall use the additional points $x_1 = 2f$ and $x_2 = 3f$.

In Figure 2 we graph the local volatility functions generated by the original dZABR model in comparison with the local volatility functions generated by the EMP-MP, EMP-ATM and EMP-NP algorithms. When implementing the EMP-NP algorithm we minimized the relative mean squared error evaluated at equally spaced grid points from $0.5f$ to $3f$, using steps of length $0.5f$. It is observed that the EMP-MP and EMP-NP approaches seem to outperform the EMP-ATM matching procedure. With our choice of $x_1 = 2f$ and $x_2 = 3f$ as the additional matching points in the EMP-MP algorithm, we observe a better fit along the right wing, whereas the EMP-NP algorithm provides better results along the left wing.

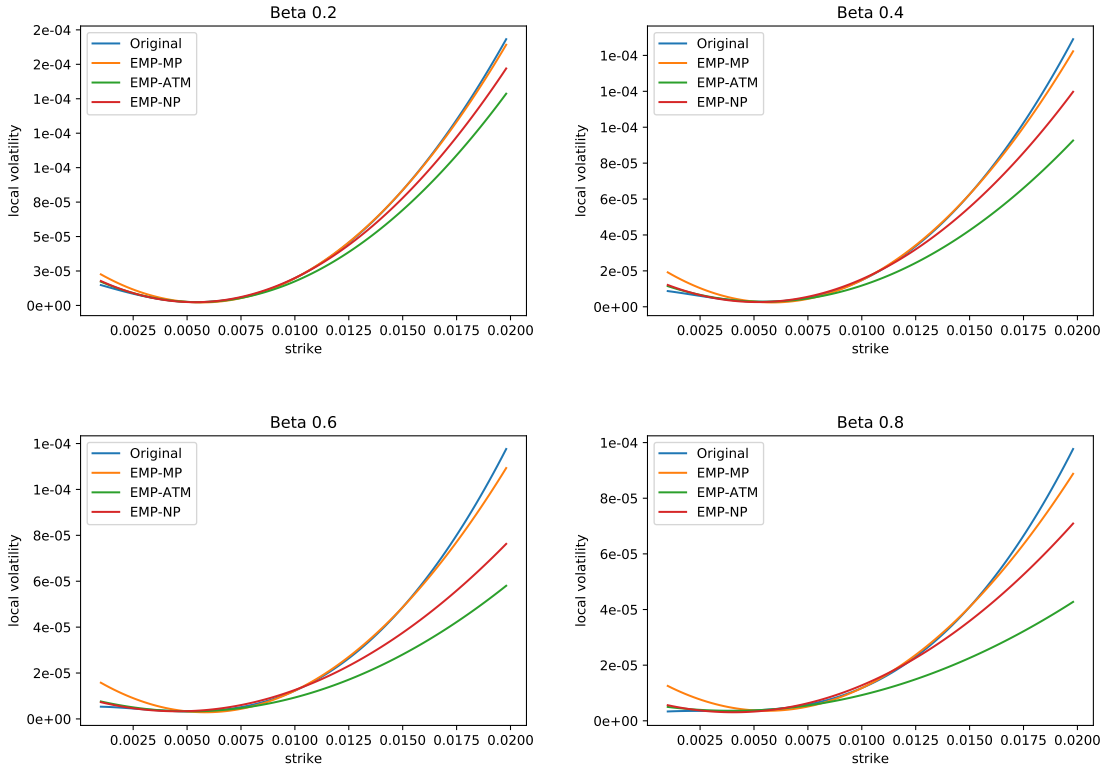

Figure 2: Local volatility functions for the different EMP algorithms using parameter set 1.

In conclusion we observe good fits for the projected dZABR model using EMP. Comparing the matching algorithms, the multiple point algorithms seem to out-

perform the ATM matching algorithm in terms of stability and quality of approximation.

6.2 Implied Volatility

The above results for the local volatility function seem promising. From a practical perspective it is, however, not clear how differences in these local volatility curves translate into pricing differences. We, therefore, generate implied volatility curves for call prices using each algorithm. In Figures 3 and 4 the normal implied volatility curves for the original model, using the framework described in [16] and [19], are compared with the curves generated by the nSABR model using the EMP-MP and EMP-NP algorithms, under the same framework and parameters. The curves are computed using parameter sets 1 and 2. We observe a good fit of the implied volatility curves for both approaches. In particular, the right wing is nearly perfectly matched. The greatest difference is observed in the left wing and for the largest value of $\beta = 0.8$.

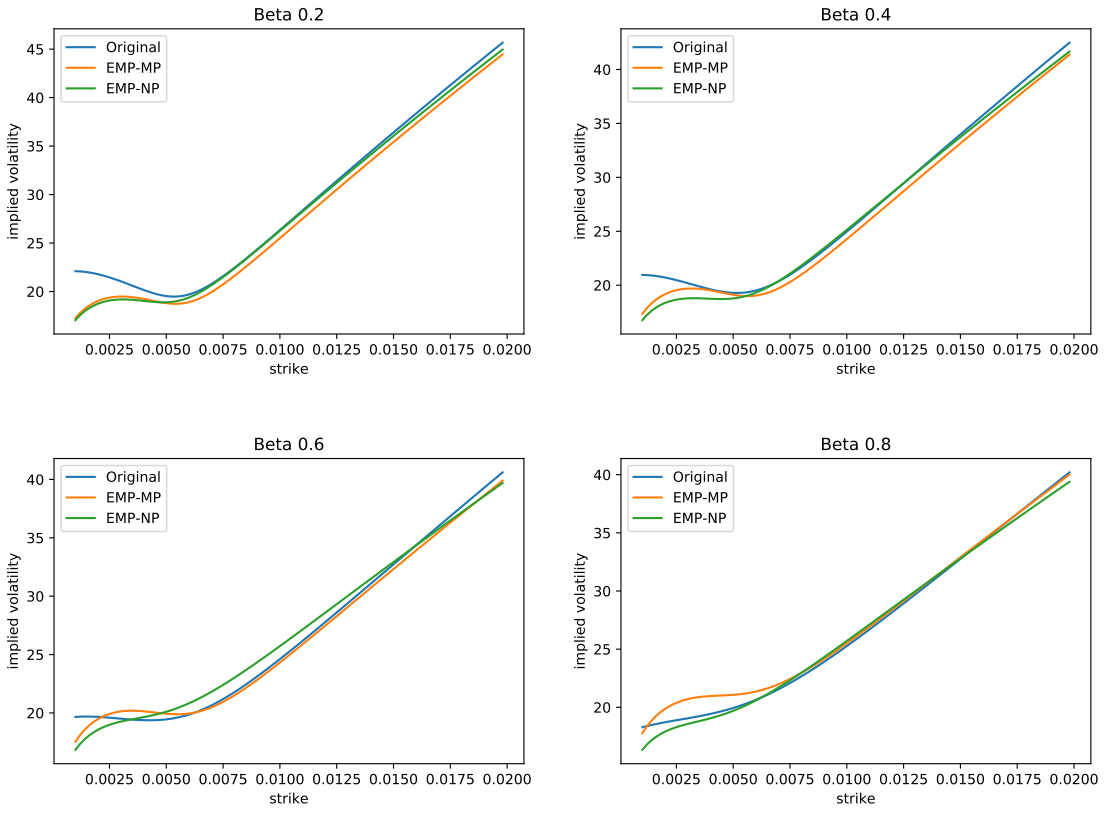

Figure 3: Normal implied volatility curves in bp computed using parameter set 1.

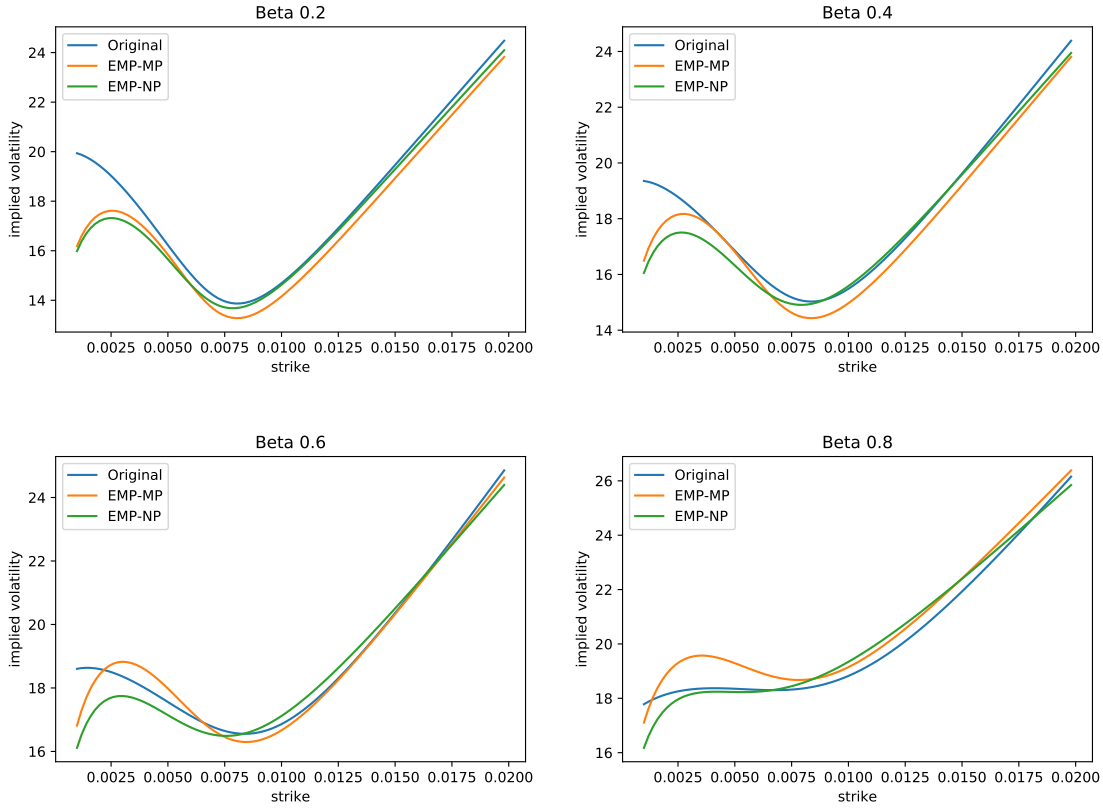

Figure 4: Normal implied volatility curves in bp computed using parameter set 2.

Similar results are observed in Tables 4 and 5, which show implied volatilities for parameter sets 3 and 4. Once again, the largest deviation in values is observed in the left wing, which is still reasonable given that we are using an arbitrage free approximation. In particular, for parameter set 3 we observe a better fit for the EMP-MP algorithm at all points considered. This demonstrates the influence of matching the entire local volatility curve on implied volatilities—even though the EMP-NP algorithm yields a better fit to the local volatility function in the area up to $3f$, larger strikes have an impact on the resulting implied volatility.

Remark 6.1. Note that the matching algorithms do not necessarily guarantee that the resulting value of ρ is bounded by -1 and 1. To ensure that the nSABR models are arbitrage free we imposed an additional bound of 0.999 on the absolute value of ρ.

	$\beta = 0.2$			$\beta = 0.8$		
x	dZABR	EMP-MP	EMP-NP	dZABR	EMP-MP	EMP-NP
$0.5f$	21.14	18.12	17.76	19.14	18.88	17.29
$0.75f$	21.49	19.67	19.27	20.37	20.78	19.07
f	22.19	20.96	20.73	21.85	22.35	20.89
$1.5f$	25.98	25.11	25.15	25.57	25.95	24.96
$2f$	30.96	30.22	30.22	29.83	30.15	29.18
$3f$	40.90	40.17	40.02	38.62	38.69	37.33
$4f$	50.30	49.50	49.19	47.23	46.86	45.00

Table 4: Normal implied volatility in bp for the dZABR model computed using parameter set 3 in comparison with the normal implied volatilities generated by EMP.

	$\beta = 0.2$			$\beta = 0.8$		
x	dZABR	EMP-MP	EMP-NP	dZABR	EMP-MP	EMP-NP
$0.5f$	21.20	22.58	21.18	20.28	25.18	20.13
$0.75f$	18.69	19.59	18.74	19.89	22.60	19.56
f	16.83	16.87	17.03	19.91	20.16	20.05
$1.5f$	18.37	17.55	18.99	22.10	19.95	23.25
$2f$	23.33	22.48	23.94	26.34	24.18	27.53
$3f$	33.80	32.85	34.15	36.28	33.92	36.29
$4f$	43.80	42.66	43.80	46.35	43.42	44.66

Table 5: Normal implied volatility in bp for the dZABR model computed using parameter set 4 in comparison with the normal implied volatilities generated by EMP.

6.3 Pricing of Interest Rate Derivatives

Having shown that the approximation techniques yield satisfactory results, we go a step further and demonstrate the possible benefits for interest rate derivative pricing. To demonstrate the advantages consider the following toy model:

6.3.1 Setup

For our concrete application we shall consider the $1y2y$- and the $1y5y$-swap rates as underlying rates and model them using dZABR models. We assume that the payment date corresponds to the exercise date $T = 1$. To show the additional

	$1y2y$	$1y5y$
f	0.003	0.005
α	0.0009	0.0015
ν	0.3	0.3
ρ	-0.5	-0.7

Table 6: Parameters of the nSABR models used to generate sample data for swap rates.

	V1	V2	V3
β	0.4	0.4	0.5
γ	0.8	0.9	0.8
d	0.002	0.002	0.002

Table 7: Additional parameters for three versions of the dZABR model.

flexibility of using dZABR models in comparison to, for example, nSABR models, we construct some sample data around the ATM values for each of the swap rates. These sample points were constructed using the nSABR model with the parameters presented in Table 6.

Next, we specify a few variations of the dZABR model for consideration. The values for the additional parameters in each of the versions are presented in Table 7. For each rate and each version the parameters α, ν and ρ are calibrated using the ATM samples. To be more precise we use sample points from 70%-ATM to 130%-ATM in steps of 5%. In Figure 5 and Figure 6 we graph the corresponding implied volatility curves for the rates. Call prices were computed using the explicit formulas presented in [21]. In these figures the additional flexibility available for controlling the wings is clearly visible.

To price the interest rate derivatives presented in Section 3 we must specify the correlation structure between swap rates. Here we assume a single driving Brownian motion for the stochastic volatility and impose a correlation structure between the rates of the form

$$\begin{bmatrix} 1 & \omega & \rho_{1y2y} \\ \omega & 1 & \rho_{1y5y} \\ \rho_{1y2y} & \rho_{1y5y} & 1 \end{bmatrix}.$$

Using the results described in Section 4, we evaluate vanilla and quadratic call options on a basket of nSABR models. To determine the dynamics of the spread we use the results of Hagan *et al.* [21] to express the spread, $S(T)$, of two nSABR

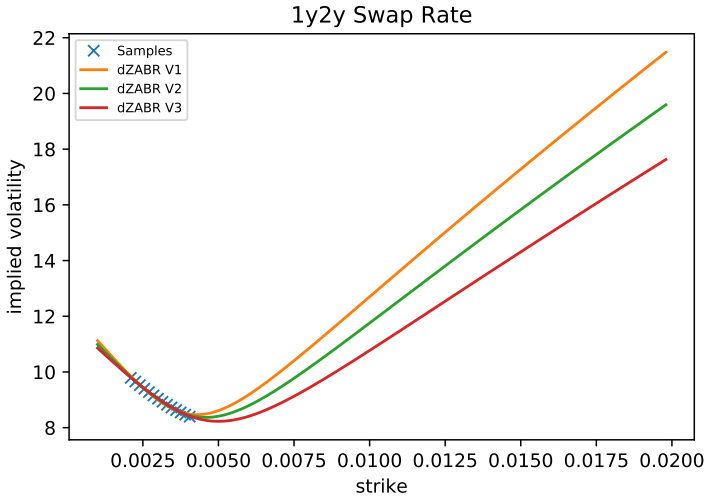

Figure 5: Normal implied volatility curves in bp for the dZABR models calibrated to the $1y2y$ swap rate.

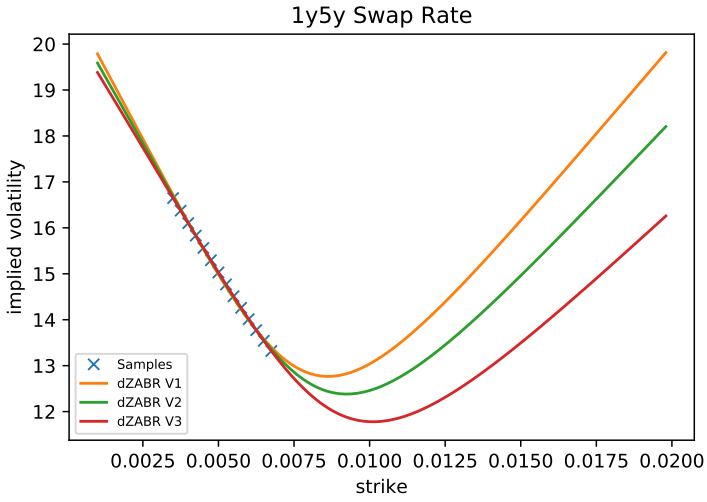

Figure 6: Normal implied volatility curves in bp for the dZABR models calibrated to the $1y5y$ swap rate.

models with parameters $(f_1, \alpha_1, \nu_1, \rho_1)$ and $(f_2, \alpha_2, \nu_2, \rho_2)$ as a nSABR model with parameters given by

$$\begin{aligned} f_s &= f_2 - f_1, \\ \alpha_s &= \sqrt{\alpha_2^2 - 2\omega\alpha_1\alpha_2 + \alpha_1^2}, \\ \nu_s &= \frac{1}{\alpha_s^2}(\alpha_2^2\nu_2 - \omega\alpha_1\alpha_2(\nu_1 + \nu_2) + \alpha_1^2\nu_1), \\ \rho_s &= \frac{1}{\alpha_s}(\alpha_2\rho_2 - \alpha_1\rho_1). \end{aligned}$$

These parameters are derived under the measure H at the exercise time T.

Remark 6.2. We note that these parameters are a rough approximation. For a more precise approximation the additional terms Γ and κ of [21] must be considered. Moreover, similar parameters may be computed for the weighted spread. For both generalizations we refer to [21].

6.3.2 CMS Spreads

Using the method described in Section 4, we can price call options on the CMS spread if we correctly specify the convexity coefficient. To determine the convexity coefficient we follow the approach of Antonov [5] and assume a flat yield curve. The convexity coefficients are then given by

$$\lambda_i = \frac{T_i - T}{2},$$

with a complete derivation of this expression provided in Appendix A.4.

Having defined all necessary terms we use the parametrization presented in Table 6 to compute the implied volatility of a call on a $5y - 2y$ spread with a maturity of 1 year. The correlation parameter ω was set to 0.5, and, in order to evaluate the expectations, we used the explicit formulas of [21]. In Figure 7 the implied volatility curves for all possible combinations of the dZABR rates are presented. The additional correlation parameter ω was calibrated to match the area from 90%-ATM to 110%-ATM with steps of size 2%. Here, we clearly see how the additional control of the wing of the swap rates allows good control of the wing of the CMS spread. In particular, this is useful when the quality of the data in the wings is questionable but the entire curve is required for risk management.

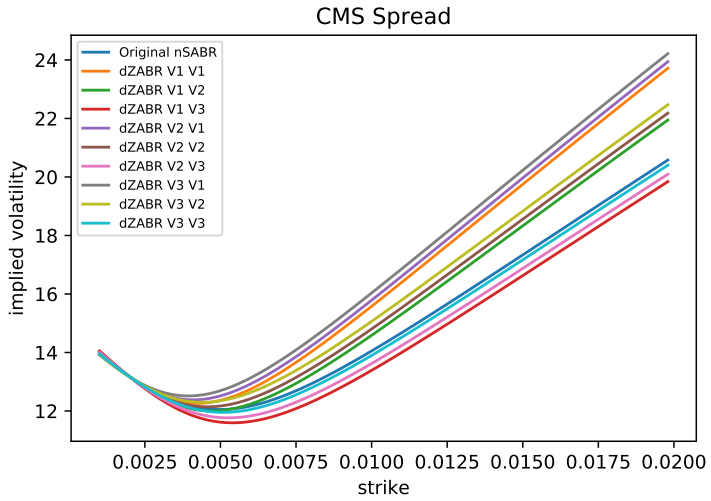

Figure 7: Normal implied volatility curves in bp for a caplet on the CMS spread. All possible combinations of the dZABR models specified in Table 7 are presented.

6.3.3 Mid-Curves

In a manner similar to that used above we approximate the convexity coefficients as

$$\hat{\lambda_i} = \frac{T + T_i - T_1 - T_2}{2}.$$

Using the approach described in Section 4.3 we evaluate the call prices on mid-curves. For the annuity factors we use further approximations and set the initial values to be

$$\begin{aligned} A_{mc}(0) &\approx T_2 - T_1 \\ A_1(0) &\approx T_1 - T \\ A_2(0) &\approx T_2 - T. \end{aligned}$$

In Table 8 implied volatilities are shown for selected combinations of the dZABR models. Again, better control of the right wing is achieved. To visualize the effects, we computed the implied volatility of the expectation of (3.4), i.e., without the additional factor coming from $A_{mc}(0)$. The results are shown in Figure 8.

Strikes	dZABR V1 V1	dZABR V2 V2	dZABR V3 V3
20	231.0	230.8	230.7
40	121.8	121.7	121.6
60	48.0	48.1	48.2
80	23.4	23.3	23.2
100	18.6	18.2	17.6
150	19.3	18.1	16.6
190	22.0	20.4	18.4

Table 8: Normal implied volatility of a call on the mid-curve for selected combinations of the dZABR models. All values are in bp.

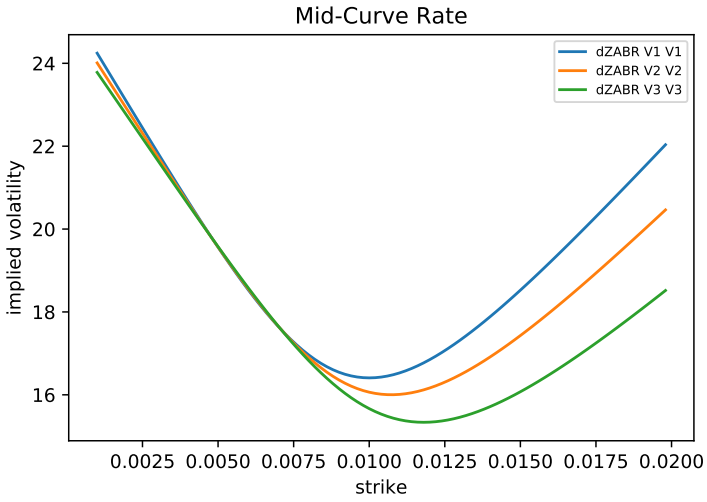

Figure 8: Normal implied volatility curves in bp for the caplet on the mid-curve without scaling.

6.4 Explicit Probability Density Function

For our last numerical demonstration we investigate the accuracy of various approximating distributions using the methodology described in Section 5. To be more precise, we compute the grid according to the framework described in [19] and [16] and deduce an average probability density over these intervals. In the notation of [19] this corresponds to the function $(\theta_j)_j$. For Johnson's-S_U distribution and the NIG distribution we evaluated the corresponding mass over the interval. The parameters of Johnson's-S_U distribution were computed using the algorithm described in Section 5, and the NIG parameters were determined using the method

provided in [14]. In Figure 9 the corresponding averaged probabilities are shown in comparison with the original dZABR distribution and the approximated nSABR distribution using the EMP-MP algorithm. The figures were generated using parameter set 2 of Table 6 with a slight modification of the maturity, which was set to $T = 1$.

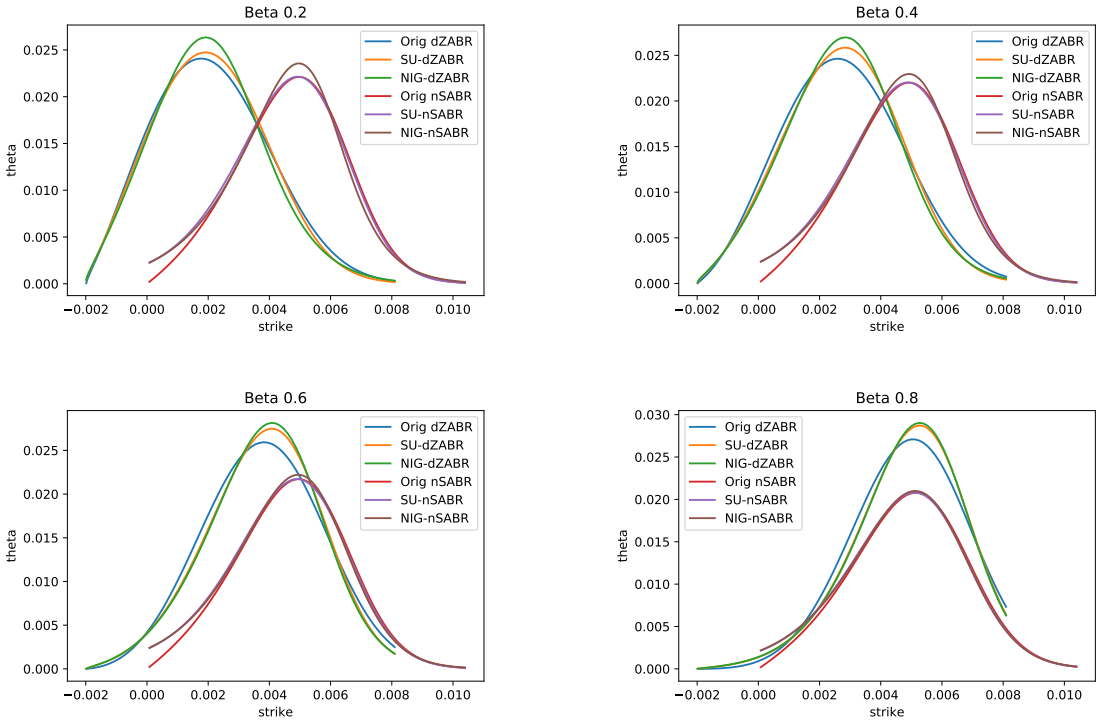

Figure 9: Probability over grid intervals for parameter set 2 and maturity $T = 1$.

As can be seen the approximation yields very good results and manages to capture the shape of the original distribution. In particular, both approximating distributions produce a nearly perfect fit to the nSABR model. The quality of dZABR approximation is of the same order as the EMP nSABR approximation of dZABR. From a practical perspective, however, this approach must be carefully applied when considering larger maturities. In Figures 10 and 11 the results of the approximations are shown for maturities $T = 5$ and $T = 7$, with all other parameters remaining the same.

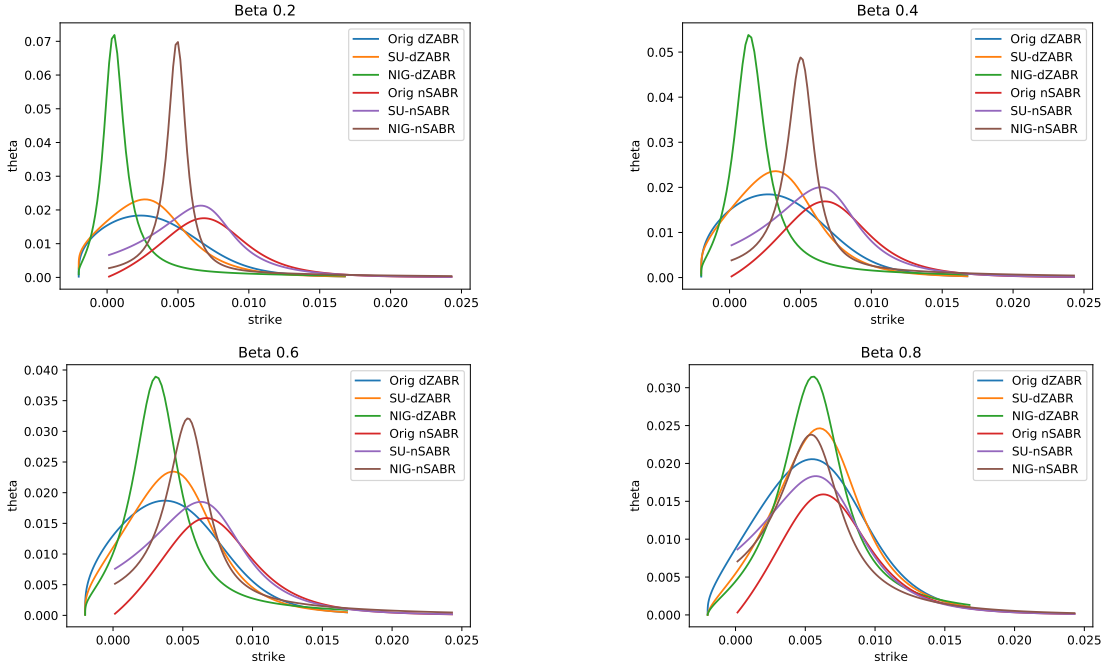

Figure 10: Probability over grid intervals for parameter set 2 and maturity $T = 5$.

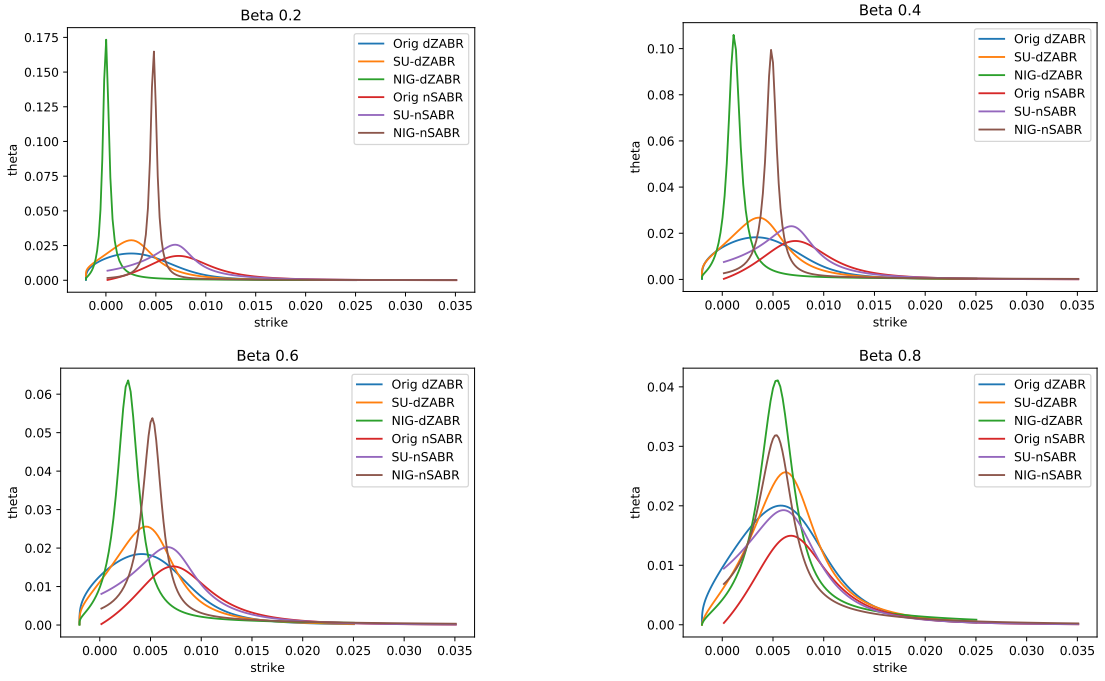

Figure 11: Probability over grid intervals for parameter set 2 and maturity $T = 7$.

Here, we observe that the Johnson distribution provides a better fit to the original distribution than the NIG distribution. For larger maturities this approximation is far from perfect—the corresponding moments are, however, perfectly matched. This behaviour can only be explained if the underling distributions cannot be characterized completely using only four moments. In particular, this implies one of two possibilities. The first possibility is that the higher moments of the distribution, which are not matched, become more important and should be taken into account when matching the parameters. The other possibility is that the underlying distribution may fall foul of the 'moment problem'. To provide some intuition, consider the first example provided in [13, Chapter 3.3.5] concerning the standard log-normal distribution, of a random variable X, which has probability density function

$$\phi_{logN}(x) = \frac{1}{\sqrt{2\pi}x} e^{-\frac{\log(x)^2}{2}} \text{ for } x \geq 0.$$

It was shown that the standard lognormal distribution, as well as its modified versions given by

$$\phi_a(x) = \phi_{logN}(x)(1 + a\sin(2\pi\log(x)))$$

with $a \in [-1, 1]$, all share the same moments, for all k, given by

$$\mathbb{E}[X^k] = e^{\frac{k^2}{2}}.$$

The problem is that the moments of the log-normal distribution grow too quickly, see [13, Theorem 3.3.25], where a sufficient condition for good behaviour is given by

$$\limsup_k \frac{\mathbb{E}[X^{2k}]^{\frac{1}{2k}}}{2k} = r < \infty. \tag{6.1}$$

Considering the leading order of the first four moments of the nSABR model, provided in Section 5, we assume the moments are of the form

$$\mathbb{E}[X^k] = \mathcal{O}(x^{\frac{(k-1)k}{2}}) = \mathcal{O}(e^{\tilde{\nu}^2 t \frac{(k-1)k}{2}}),$$

which does not obey the condition in (6.1). In particular, the growth of the moments increases with increasing maturity, influencing the speed of divergence. This may be translated into the quality of the matching algorithm, explaining the good fit for small maturities and the increasing discrepancies for larger maturities.

In conclusion, this suggests that our approach may not be universally applicable. In particular, when considering larger maturities the results must be analysed carefully, requiring a possible change of the matching algorithm to ensure the best fit. For small maturities, however, this approach can yield very good results, which allows for closed-form formulas describing the distribution and characteristic functions.

7 Conclusion

In this paper we have considered the setting of general stochastic volatility models as introduced by [16]. We have extended this work to allow pricing and calibration of multi-rate interest rate derivatives. Effective Markovian Projection was used with the normal SABR model acting as a reference model. This enabled us to apply the general setting to CMS spread and mid-curve options. Using a moment matching method based on Johnson distributions we were also able to make use of closed form solutions for pricing.

Future research will entail applying the Effective Markovian Projection method to non-parametric local stochastic volatility models with further application of these techniques for pricing derivatives.

A Appendix

A.1 Derivation of Classic Markovian Projection for general SABR

We adapt classical Markovian projection in the setting of the Heston model, see [6, 27], by following the same approach as used for the standard SABR model, see [26]. Since the derivation is very similar to that found in [6], we only present the main ideas and results.

We use the same notation as in Section 2.2. Specifically, we consider the scaled stochastic volatility $u_t := v_t/v_{t_0}$, with this scaling applied to the stochastic volatility processes u_i, $i \in \{1, 2\}$, as well. This means that we have $u_{t_0} = u_{i,t_0} = 1$. We start by considering the local volatility model

$$dF_t = \sigma(t) dB_t^{(1)}$$

with the Brownian motion defined in (2.4) and the local volatility function in (2.5). Imposing the desired form of (2.3) onto this model we can conclude

$$u_t^2 = \sigma^2(t)/C(F_t)^2 \approx (p_1^2 u_{1,t}^2 + p_2^2 u_{2,t}^2 - 2\rho p_1 p_2 u_{1,t} u_{2,t})/p^2$$

where we apply a freezing of the forward processes $F_{i,t}$, $i \in \{1, 2\}$, and use the abbreviations

$$p^2 = p_1^2 + p_2^2 - 2\rho p_1 p_2, \qquad p_i = v_{i,t_0} C_i(f_i), \qquad q_i = v_{i,t_0} C_i'(f_i).$$

For the case of a displaced diffusion model we use $C(x) = p + q(x - f)$ and to determine the parameters we consider the expression

$$\underbrace{\mathbb{E}[u_t^2|F_t = x]}_{=:I_1} \cdot C^2(x) = \underbrace{\mathbb{E}[\sigma^2(t)|F_t = x]}_{=:I_2}. \tag{A.1}$$

To compute the conditional expectations, Taylor series approximations are applied. The terms in the expression for $\sigma(\cdot)$ are of the form $u_{i,t}u_{j,t}C_{i,t}C_{j,t}$, and for u_t^2 they are of the form $u_{i,t}u_{j,t}$. Thus, by expanding around f_i, f_j, $u_{i,t_0} = 1$ and $u_{j,t_0} = 1$ we get

$$\begin{aligned} u_{i,t}u_{j,t}C_{i,t}(F_{i,t})C_{j,t}(F_{j,t}) \approx\ & C(f_i)C(f_j) \\ & +(F_{i,t} - f_i)C'(f_i)C(f_j) + (F_{j,t} - f_j)C_i(f_i)C_j'(f_j) \\ & +(u_{i,t} - 1)C_i(f_i)C_j(f_j) + (u_{j,t} - 1)C_i(f_i)C_j(f_j) \end{aligned}$$

and for $u_{i,t}u_{j,t}$ we have

$$u_{i,t}u_{j,t} \approx 1 + (u_{i,t} - 1) + (u_{j,t} - 1).$$

Next, we apply a Gaussian approximations by setting

$$\begin{aligned} d\bar{F}_t &= p\, d\bar{B}_t^{(1)} \\ d\bar{F}_{i,t} &= p_i\, dW_{i,t}^{(1)} \\ d\bar{u}_{i,t} &= \nu_i\, dW_{i,t}^{(2)} \\ d\bar{B}_t^{(1)} &= \Big(p_1\, dW_{1,t}^{(1)} - p_2\, dW_{2,t}^{(1)}\Big)/p \end{aligned}$$

with correlations given by

$$d\langle \bar{B}^{(1)}, W_1^{(1)}\rangle_t = \underbrace{(p_1 - p_2\rho)/p}_{=:\rho_1}\, dt, \qquad d\langle \bar{B}^{(1)}, W_2^{(1)}\rangle_t = \underbrace{(p_1\rho - p_2)/p}_{=:\rho_2}\, dt,$$

and

$$d\langle \bar{B}^{(1)}, dW_i^{(2)}\rangle_t = \underbrace{(p_1\gamma_{1i} - p_2\gamma_{2i})/p}_{=:\bar{\rho}_i}\, dt, \qquad i \in \{1, 2\}$$

Then, we derive approximations of the form:

$$\begin{aligned} \mathbb{E}[F_{i,t} - f_i|F_t = x] &\approx \mathbb{E}[\bar{F}_{i,t} - f_i|\bar{F}_t = x] \\ \mathbb{E}[u_{i,t} - 1|F_t = x] &\approx \mathbb{E}[\bar{u}_{i,t} - 1|\bar{F}_t = x]. \end{aligned}$$

This leads to the expressions

$$\mathbb{E}[F_{i,t} - f_i | F_t = x] \approx \frac{p_i \rho_i}{p}(x - f)$$
$$\mathbb{E}[u_{i,t} - 1 | F_t = x] \approx \frac{\nu_i \tilde{\rho}_i}{p}(x - f)$$

and finally,

$$\mathbb{E}[\sigma^2(t) | F_t = x] \approx p^2 + (x - f)K_1$$
$$\mathbb{E}[u_t^2 | F_t = x] \approx 1 + (x - f)K_2.$$

with

$$K_1 = 2(p_1^2(q_1\rho_1 + \nu_1\tilde{\rho}_1) + p_2^2(q_2\rho_2 + \nu_2\tilde{\rho}_2) - p_1p_2\rho(q_1\rho_1 + q_2\rho_2 + \nu_1\tilde{\rho}_1 + \nu_2\tilde{\rho}_2))/p$$
$$K_2 = 2(\nu_1 p_1(p_1 - p_2\rho)\tilde{\rho}_1 + \nu_2 p_2(p_2 - p_1\rho)\tilde{\rho}_2)/p^3.$$

From (A.1) we have

$$C(x) \approx \sqrt{\frac{p^2 + (x - f)K_1}{1 + (x - f)K_2}}.$$

Evaluating this function and its first derivative at the initial value yields

$$C(f) = p \qquad \text{and} \qquad C'(f) = q = (p_1^2 q_1\rho_1 + p_2^2 q_2\rho_2 - p_1p_2\rho q_1\rho_1 - p_1p_2\rho q_2\rho_2)/p^2.$$

From the variance process given by

$$u_t^2 = (p_1 u_{1,t}^2 + p_2^2 u_{2,t}^2 - 2\gamma p_1 p_2 u_{1,t} u_{2,t})/p^2$$

we deduce, via the Itō formula, that

$$\frac{du_t}{u_t} = \left(p_1\nu_1\rho_1\, dW_{1,t}^{(2)} - p_2\nu_2\rho_2\, dW_{2,t}^{(2)}\right)/p.$$

This leads to the result

$$du_t = \nu u_t dB_t^{(2)}$$

with

$$\nu = \sqrt{(p_1\nu_1\rho_1)^2 + (p_2\nu_2\rho_2)^2 - 2\xi p_1p_2\nu_1\nu_2\rho_1\rho_2}/p$$
$$\gamma = \left(p_1^2\nu_1\rho_1\gamma_{11} + p_2^2\nu_2\rho_2\gamma_{22} - p_1p_2\nu_1\rho_1\gamma_{12} - p_1p_2\nu_2\rho_2\gamma_{21}\right)/(\nu p^2)$$
$$dB_t^{(2)} = (\nu p)^{-1}\left(p_1\nu_1\rho_1\, dW_{1,t}^{(2)} - p_2\nu_2\rho_2\, dW_{2,t}^{(2)}\right).$$

A.2 Coefficients for General Stochastic Volatility Models

Here we recall the necessary assumptions that apply to Theorem 1 of [16].

Assumption I. The drift term, $\mu(\cdot)$, is differentiable, with derivative $\mu'(\cdot)$, and a solution $Y(t, t_0, \alpha)$ to the following PDE exists:

$$\begin{cases} \partial_t Y(t, t_0, \alpha) = \mu(Y(t, t_0, \alpha)) \\ Y(t, t, \alpha) = \alpha \\ Y(t_0, t_0, \alpha) = \alpha. \end{cases}$$

Assumption II. The function Y is differentiable and has an inverse function $y(t_0, t, a)$ such that

$$Y(t, t_0, \alpha) = a \qquad \Leftrightarrow \qquad \alpha = y(t_0, t, a).$$

Assumption III. The functions

$$X(t, t_0, \alpha) = \partial_\alpha Y(t, t_0, \alpha), \qquad Z(t, u) = Z(t, u, t_0, \alpha) = y(u, t, Y(t, t_0, \alpha)),$$
$$z(F) = \int_f^F \frac{1}{C(u)}\, du, \qquad s(t) = S(t_0, t, \alpha) = \int_{t_0}^t Z(t, u, t_0, \alpha)^2\, du$$

and

$$\psi(t, u, Z) = \nu\big(Z(t, u)\big) Z(t, u) X\big(t, u, Z(t, u)\big)$$

are well defined, $X\big(t, u, Z(t, u)\big)^{-1}$ exists, and the following integral functions are defined:

$$I_1(t) = \rho \int_{t_0}^t \psi(t, u, Z)\, du,$$
$$I_2(t) = 2 \int_{t_0}^t \nu\big(Z(t, u)\big)^2 X\big(t, u, Z(t, u)\big)^2 \int_u^t Z(t, v) X\big(t, v, Z(t, v)\big)^{-1}\, dv\, du,$$
$$I_3(t) = \rho \int_{t_0}^t \psi(t, u, Z) \int_u^t Z(t, v) X\big(t, v, Z(t, v)\big)^{-1}\, dv\, du,$$
$$I_4(t) = \rho^2 \int_{t_0}^t \psi(t, u, Z) \int_u^t \partial_Z\Big(\psi(t, v, Z)\Big) X\big(t, v, Z(t, v)\big)^{-1}\, dv\, du,$$
$$I_5(t) = \int_{t_0}^t \nu\big(Z(t, u)\big)^2 X\big(t, u, Z(t, u)\big)^2\, du.$$

Assumption IV. The function $C(\cdot)$ is differentiable at f, with derivative denoted by $C'(\cdot)$.

Given these assumptions the coefficients of Equation (2.9) are given by

$$a(t) = Y(t, t_0, \alpha), \qquad c(t) = b(t)^2 + \frac{1}{a(t)s(t)^2} I_2(t) - \frac{6b(t)}{s(t)^2} I_3(t) + \frac{2}{a(t)s(t)^2} I_4(t),$$
$$b(t) = \frac{1}{a(t)s(t)} I_1(t), \quad G(t) = -s(t)c(t) - s(t)b(t)\Gamma_0 + \frac{1}{a^2} I_5(t)$$

and

$$\Gamma_0 = -C'(f).$$

A.3 Basket Dynamics under suitable Numeraire

Let us consider the dynamics of the forward rate F_t given by a general SABR model

$$\begin{cases} dF_t = u_t C(F_t)\, dW_t, & F_{t_0} = f, \\ du_t = \nu u_t\, dZ_t, & u_{t_0} = 1, \\ \text{with} \quad d\langle W_t, Z_t \rangle_t = \gamma\, dt. \end{cases}$$

Let θ denote the corresponding model parameters, for instance after calibrating the model to given market data. When we consider forward swap rates we must determine each convexity-adjusted forward swap rate $\tilde{F}_t$. Given payment date T_p in the future, the convexity adjusted rate at T_p is determined by the conditional expectation of F_t with regard to the T_p-forward measure, thus,

$$\tilde{F}_t = \mathbb{E}^{T_p}[F_t].$$

The forward rate F_t is not a martingale with respect to the T_p-forward measure but the convexity adjusted rate $\tilde{F}_t$ is. Thus, given a parametrized model for F_t with model parameters θ we calibrate the respective model for $\tilde{F}_t$ and get model parameters $\tilde{\theta}$. This is similar to the method applied for the consideration in Section 3.3

Now, if we consider basket options we can consider each forward rate $F_{i,t}$ to be the convexity adjusted forward rate $\tilde{F}_{i,t}$ and safely assume that the corresponding dynamics are martingale, i.e., driftless with regard to the T_p-forward measure. Then, we proceed to price options on the basket by considering the basket for the driftless rates $\tilde{F}_{i,t}$.

A.4 Convexity Coefficients

To apply the pricing formulas of Section 3 we need to deduce the interdependence between the rates and the annuity expressed in terms of the convexity coefficients.

Due to the currently low interest regime, for our examples we do this by imposing the simplest form of a flat yield curve structure as in [5], i.e.,

$$DF(t,T) = e^{-(T-t)x(t)}.$$

Furthermore, we assume linear dependence of the functions on the factor x, which means that

$$DF(t,T) \approx 1 - (T-t)x(t) + \mathcal{O}(x^2).$$

Using the continuous representation of the annuity $A(t,T_1,T_2)$ we can approximate

$$\begin{aligned} A(t,T_1,T_2) &\approx \int_{T_1}^{T_2} DF(t,u)\,du \approx \int_{T_1}^{T_2} 1 - (u-t)x(t) + \mathcal{O}(x^2)\,du \\ &= (T_2 - T_1) - x(t)\int_{T_1}^{T_2} (u-t)\,du + \mathcal{O}(x^2), \end{aligned}$$

and can further deduce the linear dependence of the swap rates $R(t,T_1,T_2)$ on the factor x as

$$R(t,T_1,T_2) = \frac{x(t)(T_2-T_1)}{(T_2-T_1) - x(t)\int_{T_1}^{T_2}(u-t)\,du} + \mathcal{O}(x^2) = x(t) + \mathcal{O}(x^2). \tag{A.2}$$

With this we can express the dependence of the annuity on the rates as

$$A(t,T_1,T_2) \approx (T_2-T_1) - R(t,T_1,T_2)\int_{T_1}^{T_2}(u-t)\,du + \mathcal{O}(x^2).$$

To evaluate the convexity coefficients let us note that the desired terms are generally given by the form

$$\begin{aligned} M(x(t),x(s)) &= \frac{(a+b(t)x(t))(c+d(s)x(s))}{(c+d(t)x(t))(a+b(s)x(s))} \\ &\approx 1 + \frac{cb(t)-ad(t)}{ac}x(t) + \frac{ad(s)-cb(s)}{ac}x(s) + \mathcal{O}(x^2). \end{aligned} \tag{A.3}$$

For our concrete application this yields

$$\begin{aligned} M_i(x(T), x(t)) &= \frac{L_i(t)}{L_i(T)} = \frac{A_i(t)DF(T,T_0)}{A_i(T)DF(t,T_0)} \\ &\approx \frac{(1-(T_0-T)x(T))((T_i-T_0)-x(t)\int_{T_0}^{T_i}(u-t)\,du)}{((T_i-T_0)-x(T)\int_{T_0}^{T_i}(u-T)\,du)(1-(T_0-t)x(t))} \\ &\approx 1+\frac{-\tau_i(T_0-T)+\int_{T_0}^{T_i}(u-T)\,du}{\tau_i}x(T)+\frac{-\int_{T_0}^{T_i}(u-t)\,du+\tau_i(T_0-t)}{\tau_i}x(t) \\ &\approx 1+\frac{\int_{T_0}^{T_i}u\,du-\tau_i T_0}{\tau_i}(x(T)-x(t)). \end{aligned}$$

where we set $\tau_i = T_i - T_0$. This means we can approximate the convexity coefficients by

$$\lambda_i = \frac{1}{\tau_i}\int_{T_0}^{T_i} u\,du - T_0.$$

For the mid-curve options the convexity coefficient becomes

$$\begin{aligned} \hat{M_T^i}(x(T), x(0)) &= \frac{A^{mc}(T)A^i(0)}{A^i(T)A^{mc}(0)} \\ &\approx \frac{((T_2-T_1)-x(T)\int_{T_1}^{T_2}(u-T)\,du)((T_i-T_0)-x(0)\int_{T_0}^{T_i}u\,du)}{((T_i-T_0)-x(T)\int_{T_0}^{T_i}(u-T)\,du)((T_2-T_1)-x(0)\int_{T_1}^{T_2}u\,du)} \\ &\approx 1+\frac{\tau_{21}\int_{T_0}^{T_i}u\,du-\tau_i\int_{T_1}^{T_2}u\,du}{\tau_{21}\tau_i}(x(T)-x(0)). \end{aligned}$$

This means we have

$$\hat{\lambda_i} = \frac{\tau_{21}\int_{T_0}^{T_i}u\,du-\tau_i\int_{T_1}^{T_2}u\,du}{\tau_{21}\tau_i}.$$

References

[1] C. Alexander and L. M. Nogueira, *Stochastic local volatility*, in Proceedings of the Second IASTED International Conference: Financial Engineering and Applications, M. H. Hamza, ed., 2004, pp. 136–141.

[2] L. B. G. Andersen and V. Piterbarg, *Interest Rate Modeling – Volume I: Foundations and Vanilla Models*, Atlantic Financial Press, 2010.

[3] ——, *Interest Rate Modeling – Volume III: Products and Risk Management*, Atlantic Financial Press, 2010.

[4] J. Andreasen and B. Huge, *ZABR–Expansions for the masses*, Available at SSRN, (2011).

[5] A. Antonov, *Black basket analytics for mid-curves and spread-options*, Available at SSRN, (2020).

[6] A. Antonov, M. Arneguy, and N. Audet, *Markovian projection to a displaced volatility Heston model*, Available at SSRN, (2008).

[7] A. Antonov and T. Misirpashaev, *Markovian projection onto a displaced diffusion*, International Journal of Theoretical and Applied Finance, 12 (2009), pp. 507–522.

[8] P. Carr and D. B. Madan, *Option valuation using the fast Fourier transform*, Journal of Computational Finance, 2 (1999), pp. 61–73.

[9] X. Charvet and Y. Ticot, *Pricing with a smile: An approach using normal inverse Gaussian distributions with a SABR-like parameterisation*, Derivatives eJournal, (2011).

[10] J. Choi, C. Liu, and B. K. Seo, *Hyperbolic normal stochastic volatility model*, Journal of Futures Markets, 39 (2018), pp. 186–204.

[11] E. Derman and I. Kani, *Riding on a smile*, Risk, 7 (1994), pp. 32–39.

[12] B. Dupire, *Pricing with a smile*, Risk, 7 (1994), pp. 18–20.

[13] R. Durrett, *Probability: Theory and Examples*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, fourth ed., 2010.

[14] A. Eriksson, E. Ghysels, and F. Wang, *The normal inverse Gaussian distribution and the pricing of derivatives*, The Journal of Derivatives, 16 (2009), pp. 23–37.

[15] K. Feldman, *Change of measure in midcurve pricing*, Wilmott, 106 (2020), pp. 76–81.

[16] M. Felpel, J. Kienitz, and T. A. McWalter, *Effective stochastic volatility: Applications to ZABR-type models*, Quantitative Finance, on-line (2020), pp. 1–16.

[17] J. Göttker-Schnetmann and K. Spanderen, *Heston stochastic local volatility*. Quantlib Report, 2015.

[18] I. Gyoengy, *Mimicking the one-dimensional marginal distributions of processes having an Ito differential*, Probability Theory and Related Fields, 71 (1986), pp. 501–516.

[19] P. S. Hagan, *Conservative schemes for solving 1D PDEs*, Available at Researchgate, (2015).

[20] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward, *Arbitrage-free SABR*, Wilmott Magazine, 1 (2014), pp. 60–75.

[21] P. S. Hagan, A. S. Lesniewski, G. E. Skoufis, and D. E. Woodward, *CMS spread options*, Available at Researchgate, (2020).

[22] ——, *SABR for baskets*, Wilmott, (2021), pp. 2–13.

[23] P. S. Hagan, A. S. Lesniewski, and D. E. Woodward, *Implied volatility formulas for Heston models*, Wilmott Magazine, 98 (2018), pp. 44–57.

[24] N. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika, 36 (1949), pp. 149–176.

[25] M. Karlsmark, *Four essays in quantitative finance*, PhD Thesis, University of Copenhagen, (2013).

[26] J. Kienitz and D. Wetterau, *Financial Modeling - Theory, Implementation and Practice - (with Matlab source)*, Wiley, 2012.

[27] V. Piterbarg, *Markovian projection method for volatility calibration*, Available at SSRN, (2006).

[28] Y. F. Saporito, X. Yang, and J. P. Zubelli, *The calibration of stochastic local-volatility models: An inverse problem perspective*, Computers and Mathematics with Applications, 77 (2019), pp. 3054–3067.

[29] B. Tavin, *Implied distribution as a function of the volatility smile*, European Finance eJournal, (2011).

[30] O. Tsuchiya, *Markovian projection for the local stochastic volatility LIBOR market model*, Available at SSRN, (2015).

[31] H. J. H. Tuenter, *An algorithm to determine the parameters of S_U-curves in the Johnson system of probability distributions by moment matching*, Journal of Statistical Computation and Simulation, 70 (2001), pp. 325–347.