Journal of Computational Finance **20**(3), 31–49
DOI: 10.21314/JCF.2016.316

Risk journals

Research Paper

From arbitrage to arbitrage-free implied volatilities

Lech A. Grzelak[1,2] and Cornelis W. Oosterlee[1,3]

[1]Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands; email: L.A.Grzelak@tudelft.nl

[2]ING, Quantiative Analytics, Bijlmerplein 79, 1102 BH, Amsterdam, The Netherlands

[3]CWI: National Research Institute for Mathematics and Computer Science, Science Park 123, 1098 XG, Amsterdam, The Netherlands; email: C.W.Oosterlee@cwi.nl

(Received October 14, 2015; revised March 6, 2016; accepted March 7, 2016)

ABSTRACT

We propose a method for determining an arbitrage-free density implied by the Hagan formula. (We use the wording "Hagan formula" as an abbreviation of the Hagan–Kumar–Leśniewski–Woodward model.) Our method is based on the stochastic collocation method. The principle is to determine a few collocation points on the implied survival distribution function and project them onto the polynomial of an arbitrage-free variable for which we choose the Gaussian variable. In this way, we have equality in probability at the collocation points while the generated density is arbitrage free. Analytic European option prices are available, and the implied volatilities stay very close to those initially obtained by the Hagan formula. The proposed method is very fast and straightforward to implement, as it only involves one-dimensional Lagrange interpolation and the inversion of a linear system of equations. The method is generic and may be applied to other variants or other models that generate arbitrage.

Keywords: arbitrage-free density; collocation method; orthogonal projection; arbitrage-free volatility; SCMC sampler; implied volatility parameterization.

Corresponding author: L. A. Grzelak

Print ISSN 1460-1559 | Online ISSN 1755-2850

1 INTRODUCTION

When handling a large number of market volatility quotes, it is natural to express them in terms of some parametric form, so that a whole range of strikes can be explained by only a few parameters. Once the parametric equation is given, one can instantly obtain volatilities by evaluating the parametric function.

For several years, a market standard for volatility parameterization has been the well-known Hagan formula (Hagan *et al* 2002), which originates from a short-maturity heat kernel expansion. Although it is very easy to implement, the density implied by the approximation is not always free of arbitrage, especially not for very low strikes (in this case, it becomes negative or the density does not integrate to 1).

The pricing of specific financial derivatives, such as constant maturity swaps (CMSs), relies on integration of the payoff over the density that is implied by a volatility parameterization. For these CMS products, industrial practice is based on marginals, which should be properly defined and arbitrage free. In other words, these marginals cannot be based on the Hagan formula. In this paper, we propose an alternative.

Over the last decade, a number of model improvements have been introduced. For example, in Andreasen and Huge (2013), the stochastic alpha beta rho (SABR) model was solved using a one time-step finite difference approximation. In Hagan *et al* (2014), the density was arbitrage free, but the method required the numerical solution of a probability density function (PDF). Other improvements on the density have been introduced in Doust (2012) and Balland and Tran (2013).

Our approach differs from those available in the literature, as we do not seek a better analytic expression for the implied volatilities from the SABR model; instead, we project, by means of a coordinate transformation, the survival probability onto another stochastic variable, which leads to an arbitrage-free density. The concept can be expressed as follows: assuming that Y is a random variable corresponding to a model used for parameterizing volatilities, and X is a known random variable, eg, a Gaussian, we determine a coordinate transformation $y = g(x)$, for which European call prices (and their implied volatilities) agree, ie,

$$\int (y - K_i)^+ f_Y(y)\,\mathrm{d}y = \int (g(x) - K_i)^+ f_X(x)\,\mathrm{d}x, \quad i \in \{1, \ldots, N\}.$$

When the coordinate transformation is known, one can use it for pricing any plain vanilla product, while benefiting from the fact that the density $f_X(x)$ specified is free of arbitrage. In short, the proposed method can be used to approximate a random variable Y by a polynomial based on normal variables (or another variable), ie,

$$Y \sim a_0 + a_1 X + a_2 X^2 + a_3 X^3 + \cdots,$$

where the coefficients $a_0, a_1, a_2, \ldots$ are inferred from a mapping (which is explained in Section 2.2) and by solving a small system of equations.

Our preferred method for determining the mapping relies on the stochastic collocation method (Babuška *et al* 2007; Xiu and Hesthaven 2005). A mapping of a problematic variable Y to a variable X using only a few inversions of the cumulative distribution function (CDF) of Y has been presented in a financial context in Grzelak *et al* (2014), where a large number of Monte Carlo samples were generated based on only a few inversions of the original distribution. We start with the implied volatilities obtained using the Hagan formula and the density function governed by that.

This paper consists of a number of sections. In Section 2, the application of the stochastic collocation method, the projection of a "not properly defined variable" on a well-defined stochastic variable, is presented. In Section 3, we discuss different projection approaches based on the collocation method and show that the stochastic collocation method enables us to determine analytic European option prices. Section 4 contains numerical experiments with well-known parameter sets for the SABR model, and Section 5 concludes.

2 BASICS OF STOCHASTIC COLLOCATION AND IMPLIED DENSITY

2.1 Basics of stochastic collocation

Let us start with some of the intuition behind the collocation method. The method was developed to approximate an expensive-to-compute stochastic variable Y by means of a cheap variable X. An approximation is made based on the inversion of the CDF of Y at only a small set of collocation points, being the zeros of an orthogonal polynomial based on variable X.

The stochastic collocation method can be used to approximate a CDF. As any CDF is uniformly distributed, we have $F_Y(Y) \overset{\mathrm{d}}{=} F_X(X)$. This equality in distribution does not imply that X and Y are equal in distribution; rather, it implies that the CDFs follow the same uniform distribution. From the representation above, realizations of Y, y_n and X, x_n are connected via the following inversion relation:

$$y_n = F_Y^{-1}(F_X(x_n)). \tag{2.1}$$

The objective is now to determine an alternative relation that does not require many of the "expensive" inversions $F_Y^{-1}(\cdot)$ for all samples of X. Our task is to find a function $g(\cdot) = F_Y^{-1}(F_X(\cdot))$ such that

$$F_X(x) = F_Y(g(x)) \quad \text{and} \quad Y \overset{\mathrm{d}}{=} g(X),$$

where evaluations of function $g(\cdot)$ do not require the inversions $F_Y^{-1}(\cdot)$. With a mapping $g(\cdot)$ determined, the CDFs $F_X(x)$ and $F_Y(g(x))$ are equal not only in the distributional sense but also element-wise.

Sampling from the random variable Y can be decomposed into sampling from a cheap random variable X and a transformation to Y via $g(\cdot)$, ie, $y_n = g(x_n)$. It is important to choose $g(\cdot)$ to be a simple, basic function.

To find a proper mapping function, we need to extract some information from Y.

An efficient method for sampling from variable Y in terms of variable X is obtained by defining $g(\cdot)$ to be a polynomial expansion, ie,

$$y_n \approx g_N(x_n) = \sum_{i=1}^{N} y_i \ell_i(x_n), \qquad \ell_i(x_n) = \prod_{j=1, i\neq j}^{N} \frac{x_n - \bar{x}_j}{\bar{x}_i - \bar{x}_j}, \tag{2.2}$$

where x_n is a sample from X, $\bar{x}_i$ and $\bar{x}_j$ are so-called collocation points, and y_i is the exact evaluation at collocation point $\bar{x}_i$ in (2.1), ie, $y_i = F_Y^{-1}(F_X(\bar{x}_i))$ in (2.2). A particular choice for the collocation points $\bar{x}_i$ will be discussed in Section 3.

2.2 Implied density

The SABR model from Hagan *et al* (2002) is described by the following system of stochastic differential equations:

$$\begin{aligned} \mathrm{d}S(t) &= \sigma(t) S^{\beta}(t)\, \mathrm{d}W_1(t), \quad S(t_0) = S_0, \\ \mathrm{d}\sigma(t) &= \gamma\sigma(t)[\rho\, \mathrm{d}W_1(t) + \sqrt{1-\rho^2}\, \mathrm{d}W_2(t)], \quad \sigma(t_0) = \alpha, \end{aligned}$$

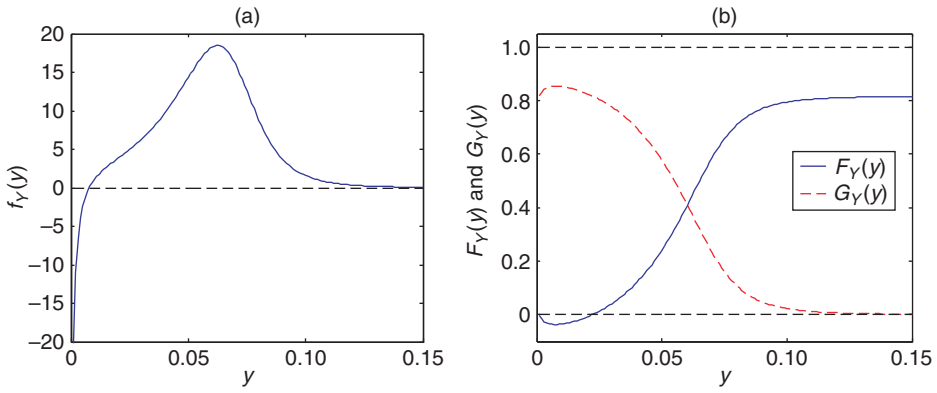
where $S(t)$ is the forward rate, S_0 is the initial forward rate, σ represents the stochastic volatility, and the parameters ρ, β, γ and α denote, respectively, the correlation, the skew, the volatility of volatility (vol-vol) and the overall level of the volatility parameters. The model is popular in the financial industry because it makes available an analytic, asymptotic implied volatility formula (derived with the help of perturbation theory), which is, for completeness, presented in Appendix A.

This implied volatility parameterization formula is often used in the financial industry for expressing market quotes, even for options with expiry times of twenty years or more. It is, however, a well-known fact that the accuracy of the so-called Hagan formula deteriorates with time, so the occurrence of implied densities giving rise to arbitrage opportunities increases as the option expiry times increase.

In this section, we discuss how to determine an arbitrage-free density based on $f_Y(\cdot)$ of Y. Typically, the implied density has problems around 0, where the absorption property is not properly handled in the formula.[1] The density deteriorates in a region near zero (see Figure 1(a)). We will map Y onto a random variable X, such that the mapping procedure takes place in those regions where the density of Y is properly defined.

[1] By using the so-called shifted variant of the model, ie, based on $S + \theta$ (see Hagan *et al* 2014), the instability at zero shifts toward $S = \theta$.

FIGURE 1 Hagan implied PDF, CDF and SDF.



(a) Probability density, with deterioration near zero. (b) Corresponding CDF (solid blue) and SDF (dashed red). Parameter values $\beta = 0.5$, $\alpha = 0.05$, $\rho = -0.7$, $\gamma = 0.4$, $F(t_0) = 0.05$ and $T = 7$.

The representation in (2.2), with $y_i = F_Y^{-1}(F_X(\bar{x}_i))$, is not yet well suited. The main problem comes from the fact that the implied CDF does not have the natural [0, 1] bounds, as shown in Figure 1(b). Since the density can become negative, CDF $F_Y(y)$ exhibits an upper bound that is less than 1.

Since $F_Y(\cdot)$ is not well defined, the inversion $F_Y^{-1}(F_X(\bar{x}_i))$ will give us incorrect mapping points. However, Figure 1 shows that, although $F_Y(y)$ does not have proper upper and lower bounds, the survival distribution function (SDF), defined by means of the European call options, $V_{\text{call}}(t_0, K)$, with strike K, as

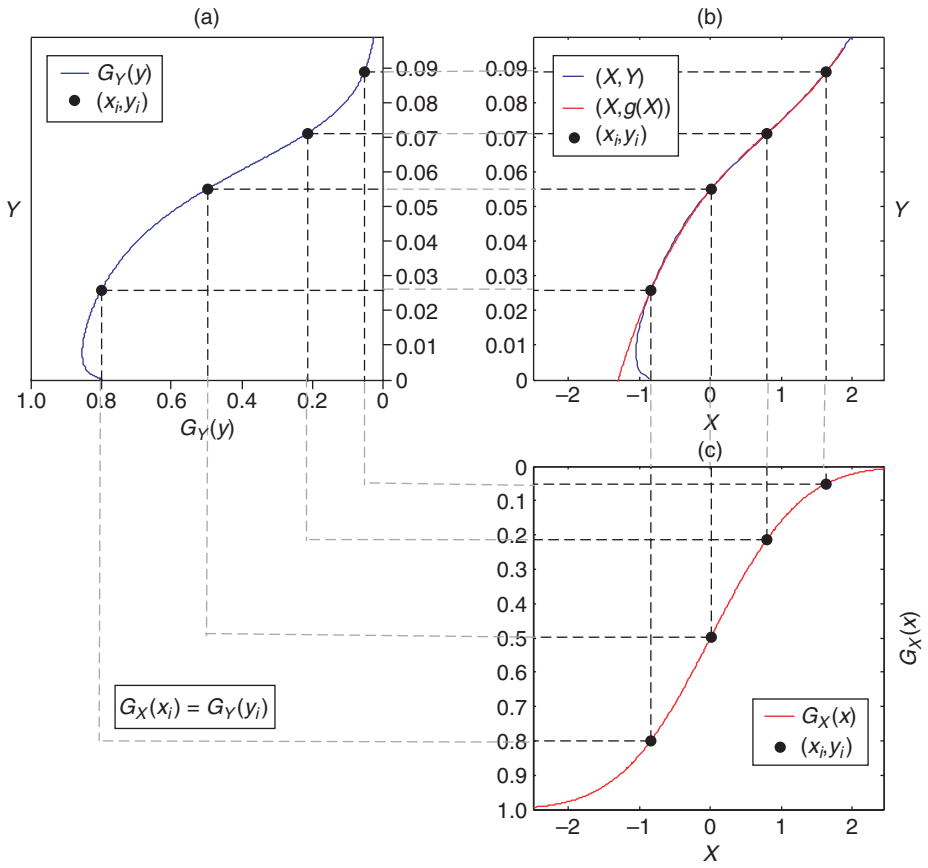
$$G_Y(y) = 1 - \int_{-\infty}^{y} f_Y(x)\,\mathrm{d}x = \int_{y}^{+\infty} f_Y(x)\,\mathrm{d}x = \boxed{-\frac{\partial V_{\text{call}}(t_0, K)}{\partial K}\bigg|_{K=y}}, \quad (2.3)$$

has as its natural limit a value of 0 for $y \to \infty$.[2] By focussing on the survival distribution $G_Y(\cdot)$, we can make use of the collocation mapping, which is given by

$$y \approx g_N(x) = \sum_{i=1}^{N} G_Y^{-1}(G_X(\bar{x}_i))\ell_i(x) =: \sum_{i=1}^{N} y_i \ell_i(x), \quad \ell_i(x) \text{ in (2.2)}. \quad (2.4)$$

[2] We could use that $\exists \hat{y} G_Y(\hat{y}) = 0.5$, so that $F_Y(\hat{y}) = 0.5$, to determine an adjustment factor for the CDF.

FIGURE 2 Illustration of the mappings of Y on $X \sim \mathcal{N}(0,1)$ with a polynomial $g_N(X)$.



(a) Survival probability implied from the Hagan model. (b) Relation between X and Y. (c) Survival probability for X. The parameters are taken as in Section 4.1.

REMARK 2.1 (Proper calculation of the survival distribution) To determine the survival probability, it is crucial not to use the representation $G_Y(y) = 1 - F_Y(y)$ but to integrate from y to infinity, as presented on the right-hand side of (2.3). By integration, we essentially find the zero region and ensure that the SDF has a proper limit value, under the assumption that the right-side tail converges to zero.

Because the inversion $G_Y^{-1}(G_X(\bar{x}_i))$ is only well defined in the part of $G_Y(\cdot)$ that is monotone, we set specific values $g_{\min}$ and $g_{\max}$ and choose the collocation points so that $G_X(\bar{x}_1) < g_{\max}$ and $G_X(\bar{x}_N) > g_{\min}$. In other words, the values of $g_{\min}$ and

$g_{\max}$ determine the range over which we can be confident about the quality of the mapping between the two variables.

When the limits $g_{\min}$ and $g_{\max}$ are prescribed, the collocation method maps the survival probability of Y onto a survival probability based on a polynomial of X. In Figure 2, an intuitive schematic picture of the projection routine is presented. Note that the Lagrange interpolation (part (b)) takes place at the variable level (X, Y), which is typically rather smooth and almost linear (see Grzelak *et al* (2014) for a more detailed discussion). In order to apply Lagrange interpolation between the nodes $(\bar{x}_i, y_i)$, it is important to make use of the optimal collocation points $\bar{x}_i$, ensuring that the polynomial has certain optimality properties and avoiding any oscillations. In Section 3, the details regarding the choice of these optimal points $\bar{x}_i$ will be given.

In the next subsection, we will discuss the relation between the densities of Y and X.

2.3 Recovery of the probability density function and pricing options

By the definition of function $g(x)$, we have

$$y = g(x) \overset{\text{def}}{=} G_Y^{-1}(G_X(x)), \quad \text{so } G_X(x) = G_Y(g(x)) =: G_Y(y). \tag{2.5}$$

Differentiating (2.5) with respect to x results in

$$\frac{\mathrm{d}G_X(x)}{\mathrm{d}x} = -f_X(x), \qquad \frac{\mathrm{d}G_Y(g(x))}{\mathrm{d}x} = -f_Y(g(x))\frac{\mathrm{d}g(x)}{\mathrm{d}x}.$$

The relation between the densities is therefore given by

$$f_Y(g(x)) = f_X(x)\left(\frac{\mathrm{d}g(x)}{\mathrm{d}x}\right)^{-1} \approx f_X(x)\left(\frac{\mathrm{d}g_N(x)}{\mathrm{d}x}\right)^{-1}, \tag{2.6}$$

with $g_N(x)$ as in (2.4); the derivative $\mathrm{d}g_N(x)/\mathrm{d}x$ is known analytically (see below in (2.7)).

By means of the mapping $y \to x$, we can make use of the cheap PDF of X, $f_X(\cdot)$. In order to evaluate $f_X(x)$, however, the mapping $x \approx g_N^{-1}(y)$ needs to be calculated.

REMARK 2.2 (Efficient evaluation of $x = g^{-1}(y)$) Since mapping $y = g(x)$ is bijective and $g(x)$ is strictly increasing, $g^{-1}(y)$ is too. This implies that the arguments x can be obtained by the inverse interpolation (Michalup 1950) of $g(x)$ against y, which can be done at essentially no cost.

With $g_N(x)$ the Lagrange polynomial, its derivative reads as follows:

$$\frac{\mathrm{d}g_N(x)}{\mathrm{d}x} = \sum_{i=1}^{N} y_i \frac{\mathrm{d}\ell_i(x)}{\mathrm{d}x} = \sum_{i=1}^{N} y_i \ell_i(x) \sum_{j=1, j\neq i}^{N} \frac{1}{x - \bar{x}_j}. \tag{2.7}$$

The density can be further simplified to

$$f_Y(g(x)) \approx f_X(x)\left(\sum_{i=1}^{N} y_i \ell_i(x) \sum_{j=1, j\neq i}^{N} \frac{1}{x-\bar{x}_j}\right)^{-1}. \tag{2.8}$$

LEMMA 2.3 (Arbitrage-free collocation variable) *The PDF as given in (2.6) is free of arbitrage and integrates to* 1.

PROOF *The proof is straightforward: by integrating the density $f_Y(y)$, which is not well defined, and changing variables $y = g(x)$, we find*

$$\int_{-\infty}^{+\infty} f_X(x)\left(\frac{\mathrm{d}g_N(x)}{\mathrm{d}x}\right)^{-1} \mathrm{d}g_N(x) = 1. \tag{2.9}$$

The proper limits of the corresponding functions $G_X(x)$ and $F_X(x)$ are guaranteed by the choice of X. □

Using the results above, we can price European-style payoffs highly efficiently, as

$$\begin{aligned} V(t_0, y_0) &= \mathbb{E}[V(T, Y(T)) \mid \mathcal{F}(t_0)] = \int_0^{\infty} V(T, y) f_Y(y)\, \mathrm{d}y \\ &= \int_{G_X^{-1}(1)}^{G_X^{-1}(0)} V(T, g(x)) f_Y(g(x)) \frac{\mathrm{d}g(x)}{\mathrm{d}x}\, \mathrm{d}x, \end{aligned} \tag{2.10}$$

which, by (2.6) and the approximation $g(x) \approx g_N(x)$, gives

$$V(t_0, y_0) = \int_{G_X^{-1}(1)}^{G_X^{-1}(0)} V(T, g(x)) f_X(x)\, \mathrm{d}x \approx \int_{G_X^{-1}(1)}^{G_X^{-1}(0)} V(T, g_N(x)) f_X(x)\, \mathrm{d}x. \tag{2.11}$$

Although the pricing of options is generally done numerically by integrating the expression in (2.11), European put and call option prices are known analytically when X is a Gaussian variable (see Section 3.4).

3 DETAILS OF THE STOCHASTIC COLLOCATION

We start with a discussion on how to choose a proper set of collocation points. With these collocation points determined, we can derive analytic European option prices. As a last item of this section, we discuss a recalibration procedure for the exact fitting of implied volatilities.

3.1 Collocation points determined by the strikes

We first consider the case in which, for a given set of strikes $K = (y_1, y_2, \ldots, y_N)^{\mathrm{T}}$, a calibration with the Hagan formula needs to be performed. At this point, we assume

that the corresponding SDF is monotone in the interval $[y_1, y_N]$, or, in other words, that the SDF is well defined in that region. Typically, the strikes at which the calibration then takes place are not too close to the zero region, where the implied density may not be arbitrage free. At the set of collocation points $(\bar{x}_1, \ldots, \bar{x}_N)^{\mathrm{T}}$, the survival probabilities implied by the model and by the collocation variable X will be identical.

If we take the collocation variable to be a standard normally distributed variable, $X \sim \mathcal{N}(0,1)$, the coordinate transformation is then given by (2.4), with $y_i = G_Y^{-1}(G_{\mathcal{N}(0,1)}(\bar{x}_i))$ and G_Y as in (2.3). We determine the collocation points $\bar{x}_i$ at which we have

$$G_Y^{-1}(G_{\mathcal{N}(0,1)}(\bar{x}_i)) = y_i \;\Rightarrow\; \bar{x}_i = G_{\mathcal{N}(0,1)}^{-1}(G_Y(y_i)). \tag{3.1}$$

With a set of strike values (ie, points y_i), we only need to determine the corresponding survival probabilities $G_Y(y_i)$ and use them in an inversion procedure based on a standard normal variable, which is trivial.[3]

When the collocation points $\bar{x}_i$ are given by means of the specified strikes $K_i = y_i$, the constructed Lagrange polynomial may not be stable for large values of N ($N \geq 5$). For large values of N, it is recommended to use the quadrature points as presented in Section 3.2, where the use of collocation points preserves interpolation stability.

3.2 Collocation points obtained by quadrature

An alternative to choosing collocation points according to the strikes available in the market is to determine the collocation points $\bar{x}_i$ in an optimal manner based on the zeros of an orthogonal polynomial. As explained in Grzelak *et al* (2014), since we choose X to be normally distributed, the $\bar{x}_i$ points will be the zeros of the Hermite polynomials (Abramowitz and Stegun 1972).

However, we need to make sure that the collocation points do not require inversions $G_Y^{-1}(G_X(\bar{x}_i))$ at points where $G_Y(\cdot)$ is not well defined. Therefore, we will apply a so-called grid-stretching technique, as introduced in Grzelak *et al* (2014), and define a stretched variable $Z := a + bX$, with $X \sim \mathcal{N}(0,1)$. We specify two parameters $g_{\min}$ and $g_{\max}$ to define the range in which the inversion can be safely performed. We subsequently determine values a and b so that $G_Z(\bar{x}_1) = g_{\max}$, $G_Z(\bar{x}_N) = g_{\min}$, which implies that $1 - F_{\mathcal{N}(a,b)}(\bar{x}_1) = g_{\max}$ and $1 - F_{\mathcal{N}(a,b)}(\bar{x}_N) = g_{\min}$. After some basic algebraic operations, we arrive at

$$\begin{aligned} b &= \frac{\bar{x}_1 - \bar{x}_N}{F_{\mathcal{N}(0,1)}^{-1}(1 - g_{\max}) - F_{\mathcal{N}(0,1)}^{-1}(1 - g_{\min})}, \\ a &= \bar{x}_1 - bF_{\mathcal{N}(0,1)}^{-1}(1 - g_{\max}). \end{aligned} \tag{3.2}$$

[3] Note that $G_{\mathcal{N}(0,1)}^{-1}(\cdot) = -F_{\mathcal{N}(0,1)}^{-1}(\cdot)$, with $F_{\mathcal{N}(0,1)}(\cdot)$, is the CDF of a standard normal variable.

So, $G_Z(\bar{x}_i) \in [g_{\min}, g_{\max}]$, and the following y_i values will be obtained:

$$y_i = G_Y^{-1}\left(G_{\mathcal{N}(0,1)}\left(\frac{\bar{x}_i - a}{b}\right)\right), \tag{3.3}$$

where $\bar{x}_i$ are the optimal collocation points of $X \sim \mathcal{N}(0,1)$. These points are available based on the moments of a standard normal variable (see Grzelak *et al* (2014) for details).

Note that Gauss quadrature has the optimal polynomial degree for integration. By the grid-stretching approach described above, one may lose the theoretical optimality properties. However, grid stretching maintains the stability of interpolation, as it only leads to a linear shift of quadrature nodes.

3.3 Absorption at zero

A consequence of using a polynomial based on a normally distributed random variable X for the approximation of variable Y is that an arbitrage-free density may give rise to a positive probability of negative values, ie, $\mathbb{P}[g(X) < 0] > 0$.

An absorption at zero boundary condition can be easily incorporated into the methodology by the following constraint on the function $g_N(X)$:

$$\hat{g}_N(X) = \begin{cases} g_N(X) & \text{for } g_N(X) > 0 \Leftrightarrow X > g_N^{-1}(0), \\ 0 & \text{for } g_N(X) \leqslant 0 \Leftrightarrow X \leqslant g_N^{-1}(0), \end{cases} \tag{3.4}$$

with $X > g_N^{-1}(0)$ corresponding to the condition $Y > 0$, and $Y \approx g_N(X)$.

The polynomial $\hat{g}_N(X)$ has an atom at $X = g_N^{-1}(0)$, which corresponds to an atom at $Y = 0$. The corresponding probability mass is given by

$$\mathbb{P}[Y = 0] \approx \mathbb{P}[\hat{g}_N(X) = 0] = \mathbb{P}[g_N(X) \leqslant 0].$$

The absorption boundary condition in (3.4) assigns the probability of $g_N(X)$ becoming negative to one point at 0.

For $K > 0$, both polynomials $g_N(X)$ and $\hat{g}_N(X)$ will yield the same European option prices. This can be seen by prescribing $V(T, Y(T)) = \max(Y(T) - K, 0)$ in (2.11), which then yields

$$\max(Y(T) - K, 0) \approx \max(\hat{g}_N(X) - K, 0) = \max(g_N(X) - K, 0),$$

as $\hat{g}_N(X) = g_N(X)$ for $g_N(X) > 0$. The difference between $g_N(X)$ and $\hat{g}_N(X)$ will only be present when dealing with negative values.

In Figure 3 (Section 4.1), an illustrative example of this will be given.

3.4 Analytic European option prices for normal collocation variable

Before we give the analytic expression for European option prices, we recall the formulas for the moments of a truncated normal distribution.

RESULT 3.1 (The moments for a truncated univariate normal distribution) Let $X \sim \mathcal{N}(0,1)$ and $a \in (-\infty, +\infty)$. Then, the expression for the moments $m_i := \mathbb{E}[X^i \mid X > a]$ reads

$$m_i = (i-1)m_{i-2} + \frac{a^{i-1} f_{\mathcal{N}(0,1)}(a)}{1 - F_{\mathcal{N}(0,1)}(a)}, \quad i = 1, \ldots, \tag{3.5}$$

with $m_{-1} = 0$, $m_0 = 1$ and $f_{\mathcal{N}(0,1)}(x)$ and $F_{\mathcal{N}(0,1)}(x)$ the standard normal probability and CDFs, respectively.

In the following lemma, we show that European option prices under $g_N(X)$ with $X \sim \mathcal{N}(0,1)$ are known analytically.

LEMMA 3.1 (European call option prices) *With the collocation random variable $X \sim \mathcal{N}(0,1)$ for $g_N(X)$, European call prices are analytically available. They are given by*

$$\begin{aligned} V_{\text{call}}(t_0, K) &\approx \int_{G_X^{-1}(1)}^{G_X^{-1}(0)} (g_N(x) - K)^+ f_{\mathcal{N}(0,1)}(x)\,\mathrm{d}x \\ &= G_{\mathcal{N}(0,1)}(c_K)\left[\sum_{i=0}^{N-1} a_i \mathbb{E}[X^i \mid X > c_K] - K\right], \end{aligned}$$

with $c_K = g_N^{-1}(K)$, $G_{\mathcal{N}(0,1)}(c_K) = 1 - F_{\mathcal{N}(0,1)}(c_K)$ and $\mathbb{E}[X^i \mid X > c_k]$ the moments of the truncated normal variable, given in Result 3.1, and where a_i, $i \in \{0, \ldots, N-1\}$, are (constant) coefficients obtained by inverting the Vandermonde matrix, V, in the matrix equation $V\boldsymbol{a} = \boldsymbol{y}$. The kth row of matrix V is given by $(1, \bar{x}_k^1, \bar{x}_k^2, \ldots, \bar{x}_k^{N-1})$, with $\bar{x}_i$ the predetermined collocation points.

PROOF *The integration domain $(G_X^{-1}(1), G_X^{-1}(0))$ for a normally distributed random variable X is given by $(-\infty, +\infty)$. A Lagrange function $g_N(x)$ can be expressed in terms of monomials, $\boldsymbol{m}(x) = (1, x, x^2, \ldots, x^{N-1})^{\mathrm{T}}$, as*

$$g_N(x) = a_0 + a_1 x + \cdots + a_{N-1} x^{N-1}, \quad \text{with } g_N(\bar{x}_i) = y_i, \tag{3.6}$$

with coefficients a_i, $i \in \{0, \ldots, N-1\}$. The coefficients $\boldsymbol{a} = (a_0, \ldots, a_{N-1})^{\mathrm{T}}$ are determined by the interpolation conditions, $g_N(\bar{x}_i) = y_i$, for $i = 1, \ldots, N$. These

coefficients can be found as solutions of linear system $V\boldsymbol{a} = \boldsymbol{y}$*, ie,*

$$\begin{pmatrix} 1 & \bar{x}_1^1 & \bar{x}_1^2 & \dots & \bar{x}_1^{N-1} \\ 1 & \bar{x}_2^1 & \bar{x}_2^2 & \dots & \bar{x}_2^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{x}_{N-1}^1 & \bar{x}_{N-1}^2 & \dots & \bar{x}_{N-1}^{N-1} \\ 1 & \bar{x}_N^1 & \bar{x}_N^2 & \dots & \bar{x}_N^{N-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-2} \\ a_{N-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \\ y_N \end{pmatrix}, \tag{3.7}$$

where matrix V *is the Vandermonde matrix.*

Given the monomial representation of function $g_N(x)$*, we set* $c_K = g_N^{-1}(K)$*, for which the option pricing equation becomes*

$$\int_{-\infty}^{+\infty} \max(g_N(x) - K, 0)\, f_{\mathcal{N}(0,1)}(x)\, \mathrm{d}x = \int_{c_K}^{+\infty} g_N(x)\, f_{\mathcal{N}(0,1)}(x)\, \mathrm{d}x - K\mathbb{P}[X > c_K].$$

By definition of $g_N(x)$ *in (3.6), we have*

$$\begin{aligned} &\int_{c_K}^{+\infty} g_N(x)\, f_{\mathcal{N}(0,1)}(x)\, \mathrm{d}x - K\mathbb{P}[X > c_K] \\ &\qquad = \sum_{i=0}^{N-1} a_i \int_{c_K}^{+\infty} x^i\, f_{\mathcal{N}(0,1)}(x)\, \mathrm{d}x - K\mathbb{P}[X > c_K] \\ &\qquad = \sum_{i=0}^{N-1} a_i \mathbb{E}[X^i 1_{X > c_K}] - K\mathbb{P}[X > c_K], \end{aligned} \tag{3.8}$$

where $X \sim \mathcal{N}(0,1)$*.*

The last integral defines $\mathbb{E}[X^i 1_{X>c_K}]$*, which is equivalent to*

$$\begin{aligned} \mathbb{E}[X^i 1_{X>c_K}] &= \mathbb{E}[X^i \mid X > c_K]\mathbb{P}[X > c_K] \\ &= \mathbb{E}[X^i \mid X > c_K](1 - F_{\mathcal{N}(0,1)}(c_K)). \end{aligned} \tag{3.9}$$

By combining (3.9) and (3.8), the proof is complete. □

By the put–call parity and the lemma above, put option prices are also available in closed form. Moreover, with analytic European option prices, the calculation of the corresponding implied volatility is a trivial exercise.

REMARK 3.2 (Inversion of the Vandermonde matrix) In order to calculate these European option prices, one needs to invert the Vandermonde matrix in (3.7). For a large number of collocation points ($N > 8$), this can be problematic, and therefore it is important to check the condition number of this matrix.

3.5 Method enhancement and recalibration

We cannot guarantee that the arbitrage-free density obtained by stochastic collocation yields the same implied volatilities as those given by the Hagan formula. Here, we address this issue.

Let $\Omega_Y = [\beta, \alpha, \rho, \gamma]$ be the set of parameters that is initially obtained by calibrating with the Hagan formula to find the implied volatilities, $\sigma_Y^{\Omega}(\bar{K})$, with $\bar{K}$ the set of market strike values. This set of parameters will thus give us density $f_Y^{\Omega}(y)$, which may not be free of arbitrage. By the collocation method, we subsequently obtain a density $\hat{f}_X^{\Omega}(x)$ that is arbitrage free but does not necessarily return the implied volatilities for which $\sigma_X^{\Omega}(\bar{K}) = \sigma_Y^{\Omega}(\bar{K})$.

In essence, function $g(\cdot)$ not only depends on X but is also a function of the parameters, ie, $g(X;\Omega)$; so, by changing Ω, we may obtain a different set of option prices and corresponding volatilities. Therefore, we define here an optimization problem in which we search for $\hat{\Omega}$, so that $g(X;\hat{\Omega})$ gives the same implied volatilities as those from the market.

In other words, we describe an optimization procedure in which we determine the set of parameters, $\hat{\Omega} = [\hat{\beta}, \hat{\alpha}, \hat{\rho}, \hat{\gamma}]$, so that the volatilities from the market and the collocation method agree, ie,

$$\min_{\hat{\Omega}} \sum_{\bar{K}} (\sigma_{\text{Mrkt}}(\bar{K}) - \sigma_{g(X,\hat{\Omega})}(\bar{K}))^2. \tag{3.10}$$

Typically, $\hat{\Omega}$ will be close to the initial set Ω, and a local-search algorithm, such as the Nelder–Mead method (Nelder and Mead 1965), will therefore be sufficient to find the optimal solution.

The optimization problem in (3.10) can also be expressed in terms of European option prices. The gradients required during optimization in that case may be derived analytically.

Additionally, one may include constraints on the coefficients of polynomial $g_N(X)$, ensuring, for example, the martingale property of the model, ie, $\mathbb{E}[g_N(X)] = S_0$, with S_0 the forward price. Since $g_N(X)$ is given as a polynomial, we have

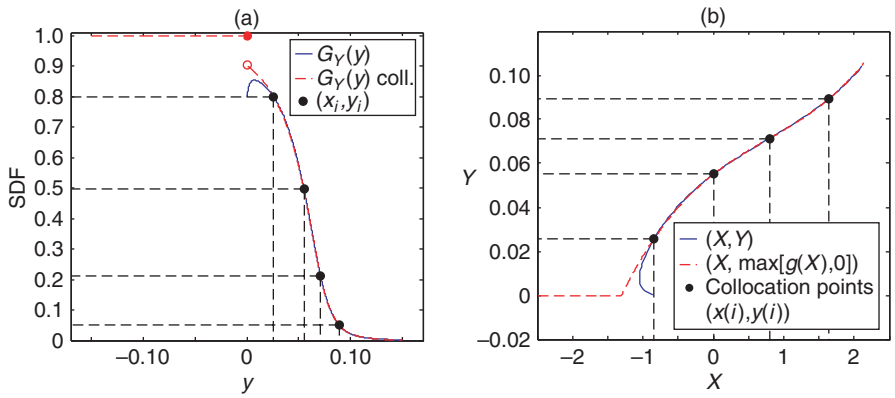
$$\mathbb{E}[g_N(X)] = \sum_{i=0}^{N-1} a_i \mathbb{E}[X^i] = S_0, \tag{3.11}$$

which, for example, with $N = 4$ and $X \sim \mathcal{N}(0,1)$, gives us $S_0 = a_0 + a_2$.

4 NUMERICAL EXPERIMENTS

In this section, we discuss several numerical experiments, ranging from a simple illustrative example to testing the introduced method with model parameters that are often used in the literature and considered to be extreme and realistic.

FIGURE 3 Hagan versus the collocation implied survival probability and the relation between Y, X and $g(X)$.



(a) SDFs obtained from the implied density and the collocation method. (b) Relation between (X, Y) and $(X, \hat{g}_N(X))$, as given in (3.4). The parameters are set as $\beta = 0.5$, $\alpha = 0.05$, $\rho = -0.7$, $\gamma = 0.4$, $F(t_0) = 0.05$ and $T = 7$.

4.1 Illustrative numerical experiment

We continue with the experiment in Section 2.2, with the following set of parameters: $\beta = 0.5, \alpha = 0.05, \rho = -0.7, \gamma = 0.4, F(t_0) = 0.05$ and $T = 7$.

For the implied survival probabilities, we take $g_{\max} = 0.8$ and $g_{\min} = 0.05$, implying that the density from the parameterization is well defined in the range $[G_Y^{-1}(g_{\max}), G_Y^{-1}(g_{\min})]$. We choose $N = 4$ collocation points. Since we project the model on a polynomial of normal variables, we first determine the optimal collocation points, $\bar{x}_i$, for $X \sim \mathcal{N}(0,1)$. These points are well known, as they are the zeros of the Hermite polynomials (scaled by $\sqrt{2}$):

$$(\bar{x}_1^0, \bar{x}_2^0, \bar{x}_3^0, \bar{x}_4^0) = (-2.3344, -0.7420, 0.7420, 2.3344).$$

For details on how to obtain these points, we refer to Grzelak *et al* (2014). Using (3.2), $a = -0.7541$ and $b = 1.8777$, and by $\bar{x}_i = (\bar{x}_0 - a)/b$, the collocation points are given by

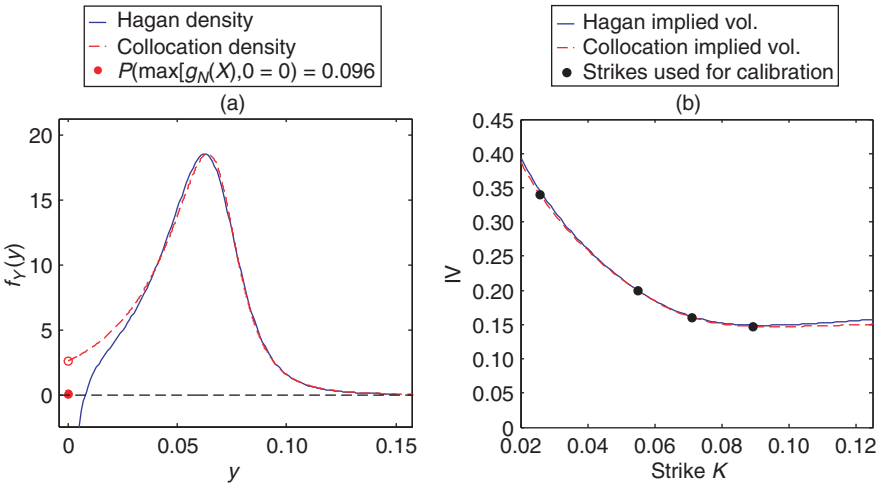
$$(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (-0.8416, 0.0065, 0.7968, 1.6448).$$

For each collocation point we calculate, see (3.3), $y_i = G_Y^{-1}(G_X(\bar{x}_i))$:

$$(y_1, y_2, y_3, y_4) = (0.0258, 0.0551, 0.0713, 0.0894).$$

Points y_i should be considered as optimal points at which the survival probabilities are mapped. It is easy to check that $G_X(\bar{x}_1) = G_Y(y_1) = g_{\max} = 0.8$ and $G_X(\bar{x}_4) =$

FIGURE 4 Hagan versus the collocation implied density, and the Hagan and collocation implied volatilities.



(a) The implied density and the density by the collocation method. (b) The corresponding implied volatilities. The parameters are as in Figure 3.

$G_Y(y_4) = g_{\min} = 0.05$. In Figure 2, the mapping procedure is illustrated, and in Figure 3, the absorption technique is shown. The corresponding PDF and implied volatilities are depicted in Figure 4.

With collocation points $\bar{x}_i$ and nodes y_i, the mapping is $y \approx g_4(x) = \sum_{i=1}^{4} y_i \ell_i(x)$, and the price of a European call option with strike K is given by Lemma 3.1.

In Figure 3, the recovered SDFs and the relation between the variables X and Y are presented. The stochastic collocation method guarantees an exact match at the collocation points, which are indicated by the black dots. Between the collocation points, the variables are interpolated by Lagrange interpolation. Outside the range of collocation points, an extrapolation, determined by X, takes place.

From a computational perspective, the collocation method is highly efficient. The complete simulation takes about 0.01 seconds, which is highly satisfactory.[4]

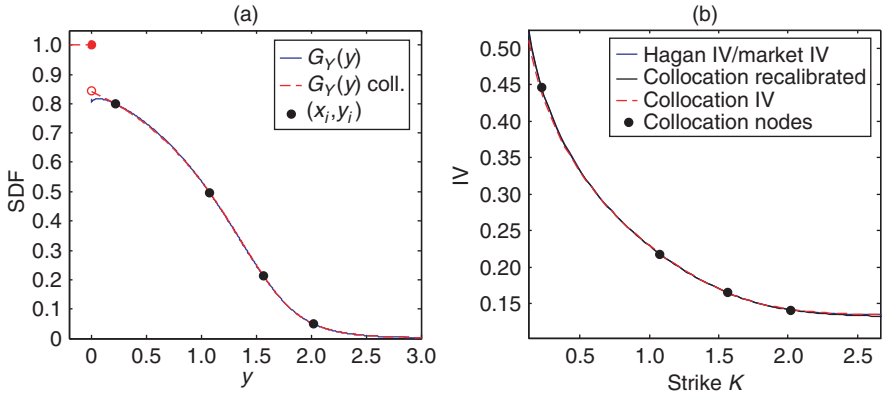
In the considered example, the application of the collocation method already gave highly satisfactory results, without the recalibration step. It will be shown in the next section that the results can be further improved.

[4] i5-4670 CPU @ 3.40GHz with 8.00GB ram simulated with MATLAB.

TABLE 1 Model parameters chosen in the experiments.

Parameters	β	α (ATM)	ρ (Corr)	γ (vol-vol)	$F(t_0)$	T
Set I as in Antonov and Spector (2012)	0.6	0.25	−0.8	0.3	1	10
Set II as in Hagan *et al* (2014)	0.25	0.35	−0.1	1	1	1
Set III as in Balland and Tran (2013)	0.2	0.26	−0.5	0.35	1	15

FIGURE 5 Survival probabilities and implied volatilities for Set I (as given in Antonov and Spector (2012)), with and without recalibration.



(a) Hagan versus collocation survival probability. (b) Implied volatilities. The experiment was performed with $g_{\min} = 0.05$ and $g_{\max} = 0.8$. $G_Y(y)$ stands for the implied survival probability, while $G_Y(y)$ coll. indicates the survival probability obtained from the collocation method.

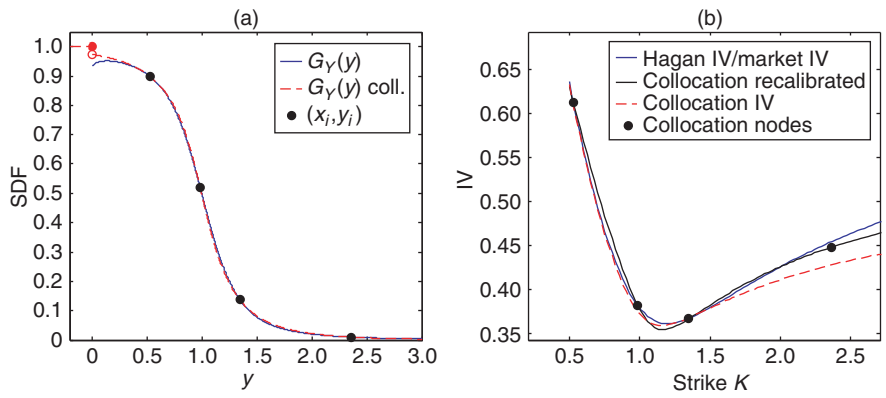
4.2 Market examples

Here, we test our method with parameters that are well known from the literature. Different parameter combinations are presented in Table 1, where the option expiry varies from one year to fifteen years. In the experiments, we show the generated densities and the corresponding implied volatilities.

In all experiments, we consider four collocation points that are determined based on the strikes $y_i = K_i$, as presented in Section 4.1. The method thus employs only four collocation points to reproduce the implied volatilities from the market.

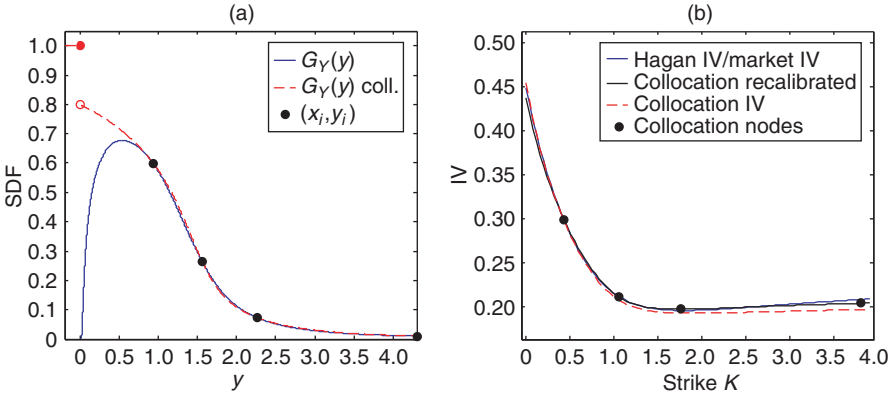
In Figures 5, 6 and 7, the obtained survival probabilities (SDFs) and the implied volatilities are presented. In all three cases, the SDFs from the collocation method

FIGURE 6 Survival probabilities and implied volatilities for Set II (as given in Hagan *et al* (2014)), with and without recalibration.



(a) Hagan versus collocation survival probability. (b) Implied volatilities. The experiment was performed with $g_{\min} = 0.01$ and $g_{\max} = 0.9$.

FIGURE 7 Survival probabilities and implied volatilities for Set III (as given in Balland and Tran (2013)), with and without recalibration.



(a) Hagan versus collocation survival probability. (b) Implied volatilities. The experiment was performed with $g_{\min} = 0.01$ and $g_{\max} = 0.6$.

are as desired, ie, they are monotone and their limits are 0 and 1. The resulting implied volatilities are not all very close to the market values. This can be improved by performing the recalibration step. In all three cases, the recalibration results in an

almost perfect implied volatility match at the collocation points. We also note that the tail asymptotics and the level of curvature and skewness were preserved by the stochastic collocation method. With as few as four collocation points, in all examples, a wide range of implied volatility shapes were generated.

Because the recalibration step requires only local optimization iterations, it is very fast. The full projection and calibration procedure takes less than 0.1 seconds for all three cases.

5 CONCLUSIONS

In this paper, we have presented an application of the stochastic collocation method for obtaining an arbitrage-free density based on the Hagan formula. Our method relies on the availability of a survival distribution function, which is not necessarily well defined on the whole domain; this is projected onto a Gaussian variable. The method presented gives implied volatilities in accordance with those obtained by the model; however, in some cases, a recalibration step is required to guarantee a perfect fit. The method is easy to implement, as it only relies on Lagrange interpolation and the solution of a linear system of equations.

APPENDIX A. IMPLIED VOLATILITY FOR THE STOCHASTIC ALPHA BETA RHO MODEL

The approximating implied volatility derived in Hagan *et al* (2002) reads as follows:

$$\hat{\sigma}^H(T,K) = A(K)\frac{z(K)}{\chi(z(K))} + B(T,K),$$
$$\hat{\sigma}^H(T,S_0) = \frac{\alpha}{S_0^{1-\beta}}B(T,S_0),$$

where

$$z(K) = \frac{\gamma}{\alpha}(S_0K)^{(1-\beta)/2}\log(S_0/K),$$
$$\chi(z(K)) = \log\left(\frac{\sqrt{1-2\rho z(K)+z^2(K)}+z(K)-\rho}{1-\rho}\right),$$
$$A(K) = \alpha\left(S_0K^{(1-\beta)/2}\left(1+\frac{(1-\beta)^2}{24}\log^2(S_0/K) + \frac{(1-\beta)^4}{1920}\log^4(S_0/K)\right)\right)^{-1},$$
$$B(T,K) = \left\{1+\left(\frac{(1-\beta)^2}{24}\frac{\alpha^2}{(S_0K)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\gamma\alpha}{(S_0K)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\gamma^2\right)\right\}T.$$

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Abramowitz, M., and Stegun, I. A. (eds) (1972). Modified Bessel functions I and K. In *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th edn. Dover, New York.

Andreasen, J., and Huge, B. (2013). Expanded forward volatility. *Risk* **26**(1), 101–107.

Antonov, A., and Spector, M. (2012). Advanced analytics for the SABR model. Working Paper, Social Science Research Network (http://doi.org/bkqb).

Babuška, I., Nobile, F., and Tempone, R. (2007). A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis* **45**, 1005–1034 (http://doi.org/c6m9gv).

Balland, P., and Tran, Q. (2013). SABR goes normal. *Risk* **26**(5), 76–81.

Doust, P. (2012). No-arbitrage SABR. *The Journal of Computational Finance* **15**(3), 3–31 (http://doi.org/bj65).

Grzelak, L. A., Witteveen, J. A. S., Suarez, M., and Oosterlee, C. W. (2014). The stochastic collocation Monte Carlo sampler: highly efficient sampling from "expensive" distributions. Working Paper, Social Science Research Network. URL: http://ssrn.com/abstract=2529691 (http://doi.org/bkqc).

Hagan, P. S., Kumar, D., Leśniewski, A. S., and Woodward, D. E. (2002). Managing smile risk. *Wilmott Magazine*, July, 84–108 (http://doi.org/bkqd).

Hagan, P. S., Kumar, D., Leśniewski, A. S., and Woodward, D. E. (2014). Arbitrage-free SABR. *Wilmott Magazine*, January, 60–75.

Michalup, E. (1950). On inverse linear interpolation. *Scandinavian Actuarial Journal* **1950**(1–2), 98–100 (http://doi.org/fxrp7s).

Nelder, J., and Mead, R. (1965). A simplex method for function minimization. *Computer Journal* **7**, 308–313 (http://doi.org/bkqf).

Xiu, D., and Hesthaven, J. S. (2005). High-order collocation methods for differential equations with random inputs. *SIAM Journal on Scientific Computing* **27**, 1118–1139 (http://doi.org/dvwftm).