Article

Computing Black Scholes with Uncertain Volatility—A Machine Learning Approach

Kathrin Hellmuth ![ORCID] and Christian Klingenberg *![ORCID]

Department of Mathematics, University of Würzburg, 97074 Würzburg, Germany;
kathrin.hellmuth@mathematik.uni-wuerzburg.de
* Correspondence: klingen@mathematik.uni-wuerzburg.de

Abstract: In financial mathematics, it is a typical approach to approximate financial markets operating in discrete time by continuous-time models such as the Black Scholes model. Fitting this model gives rise to difficulties due to the discrete nature of market data. We thus model the pricing process of financial derivatives by the Black Scholes equation, where the volatility is a function of a finite number of random variables. This reflects an influence of uncertain factors when determining volatility. The aim is to quantify the effect of this uncertainty when computing the price of derivatives. Our underlying method is the generalized Polynomial Chaos (gPC) method in order to numerically compute the uncertainty of the solution by the stochastic Galerkin approach and a finite difference method. We present an efficient numerical variation of this method, which is based on a machine learning technique, the so-called Bi-Fidelity approach. This is illustrated with numerical examples.

Keywords: numerical finance; Black Scholes equation; uncertainty quantification; uncertain volatility; polynomial chaos; Bi-Fidelity method

1. Introduction

In modern financial markets, traders can choose from a large variety of financial derivatives. This term denotes financial instruments that have a value determined by so-called underlying variables or assets such as stocks, the oil price or the weather. Originally, derivatives were invented to reduce the risk of uncertain prices, especially in agricultural markets where one could have long periods between sowing and harvest, see Chapter 1 and Chapter I in [1,2], respectively.

As the derivative market was growing, the need for a pricing formula for derivatives also increased in the 20th century. A breakthrough was made by Black, Scholes [3] and Merton [4] when they contemporaneously formulated a model allowing the evaluation of derivatives. They were later awarded the Nobel prize in economics for their work. Derived from this model, the Black Scholes equation

$$\frac{\partial V(S,t)}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V(S,t)}{\partial S^2} + rS\frac{\partial V(S,t)}{\partial S} - rV(S,t) = 0, \quad S \in (0,\infty), \, t \in [0,T], \quad (1)$$

explains the behaviour of the price V of the derivative by means of a partial differential equation (PDE). This derivative is allowed to depend on the time t up to maturity T and only one underlying stochastic asset (e.g., a stock, an index or some commodity price). The price of the asset is denoted by S and is assumed to follow a geometric Brownian motion. The constant r denotes the risk free rate of interest in the market and $\sigma \in \mathbb{R}$ is the so-called volatility of the stochastic asset. Later, this model was extended to multiple underlying assets and adjusted for certain kinds of underlying variables, for example, interest rates; see [5].

Comparison to real data soon showed that the volatility σ of one and the same stochastic asset can take values that differ more than explainable by rounding errors, etc.; see [6–8].

The most popular approaches to deal with this are to model the volatility either as local volatility, i.e., a function $\sigma(S, t)$ as in [9–12] or as a stochastic process; compare, e.g., the famous Heston model [13] or [6,7,14].

These models as well as others are formulated in continuous-time. Because prices in financial markets are only visible in trades, the markets always operate in discrete time; compare, e.g., [15]. The models above, therefore, represent approximations of reality. In order to use them, they have to be fitted to the market using discrete data. This fitting procedure, however, introduces uncertainty in the recovered values, e.g., by rounding errors and by the interplay of the random nature of stochastic assets and their discrete observations. Because the interest rate r is easy to determine even from discrete data, we focus on the volatility σ for considerations of uncertainty.

Also in [16–18], the authors investigated uncertainty in the volatility: They modelled it as a one-dimensional random variable $\Sigma(\omega) = \Theta(\omega)$ or a function of a one-dimensional random variable $\Sigma(\Theta(\omega))$ for ω in the underlying probability space. Then the price process $V(S, t, \Theta)$ also depends on Θ and follows the stochastic version of the Black Scholes equation

$$\frac{\partial V(S, t, \Theta)}{\partial t} + \frac{1}{2}\Sigma(\Theta)^2 S^2 \frac{\partial^2 V(S, t, \Theta)}{\partial S^2} + rS\frac{\partial V(S, t, \Theta)}{\partial S} - rV(S, t, \Theta) = 0. \tag{2}$$

This equation can be derived by inserting the stochastic volatility into the Black Scholes model, where the Brownian motion is independent of Θ. It can be studied by means of uncertainty quantification: The solution V is developed in a generalized Polynomial Chaos (gPC) expansion

$$V(S, t, \Theta(\omega)) = \sum_{n=0}^{\infty} v_n(S, t) p_n(\Theta(\omega)) \tag{3}$$

for orthonormal polynomials p_n w.r.t. the distribution of Θ and coefficients given by the expected value $v_n(S, t) = E(V(S, t, \Theta)p_n(\Theta))$. If Θ has a density $\mu : \mathcal{D} \to \mathbb{R}$, one can alternatively calculate the coefficients by

$$v_n(S, t) = \int_{\mathcal{D}} V(S, t, x) p_n(x) \mu(x)\, dx.$$

In [16], these integrals are directly computed by a quadrature rule. The required solutions $V(\cdot, \cdot, x_j)$ in the quadrature points x_j are calculated as the solutions of the deterministic Black Scholes Equation (1) with $\sigma = x_j$. This classifies the method as a Stochastic Collocation method.

In the articles [17,18], however, the stochastic Galerkin method is applied for computing the coefficients $v_n(S, t)$. By inserting the gPC expansion (3) into the stochastic Black Scholes Equation (2), multiplying the equation by an orthogonal polynomial $p_k(\Theta)$ and applying the expected value on both sides, deterministic PDEs for the coefficients $v_n(S, t)$ are derived

$$0 = \frac{\partial v_k(S, t)}{\partial t} + \frac{1}{2}S^2 \sum_{n=0}^{\infty} \frac{\partial^2 v_n(S, t)}{\partial S^2} E\left((\Sigma(\Theta))^2 p_k(\Theta)p_n(\Theta)\right) + rS\frac{\partial v_k(S, t)}{\partial S} - rv_k(S, t). \tag{4}$$

After truncating of the system and the coupling term to a finite number of indices, the system is solved numerically by the method of lines in [17] and the finite element method in [18].

In our work, we extend the model used above to the volatility $\Sigma(\Theta_1, \ldots, \Theta_L)$ depending on many finitely random variables $\Theta_1, \ldots, \Theta_L$. This leads to the stochastic Black Scholes equation

$$
\begin{aligned}
0 = \quad & \frac{\partial V(S, t, \Theta_1, \ldots, \Theta_L)}{\partial t} + \frac{1}{2}\Sigma^2(\Theta_1, \ldots, \Theta_L)S^2\frac{\partial^2 V(S, t, \Theta_1, \ldots, \Theta_L)}{\partial S^2} \\
& + rS\frac{\partial V(S, t, \Theta_1, \ldots, \Theta_L)}{\partial S} - rV(S, t, \Theta_1, \ldots, \Theta_L).
\end{aligned}
\tag{5}
$$

A model like this might, for instance, occur when the volatility is modelled as a random variable that also depends on certain stochastic factors as in [19–22]. We propose a momentum constrained maximum likelihood technique to fit the volatility distribution to real data and compare our results to market data.

The solution is then derived in the same way as in (4) and calculated numerically by a finite difference method. The computational cost for multiple similar calculations is reduced by a Bi-Fidelity technique, which can be considered as a machine learning approach.

This paper thus extends the existing literature [17,18] to a finite number of random variables affecting the volatility. To deal with the increasing computational effort, a Bi-Fidelity approach is introduced in numerical computations. The effectiveness is illustrated in numerical examples with an explicit finite difference scheme. Another novelty is the comparison of our model to real market data. In order to fit the stochastic volatility of our model to the observed market volatility, a momentum constrained maximum likelihood approach is proposed.

The outline of the article is as follows: After introducing gPC to finitely many random variables, a method of fitting the stochastic volatility to real data is described in Section 2. The stochastic Galerkin method is used to solve Equation (5) in Section 3. However, computational costs can be high. Thus, we introduce a Bi-Fidelity numerical technique to compute this more efficiently in Section 4. The paper is rounded out with numerical results illustrating the effectiveness of this technique and the fit to real market data in Section 5.

2. Fitting the Random Volatility to Real Market Data

For the convenience of the reader and in order to introduce notation, we briefly recall the fundamentals of generalized polynomial chaos (gPC). We then propose a method to fit the volatility distribution to real data.

Denote by $\Theta_1, \ldots, \Theta_L$ random variables with joint distribution function $\bar{F} : \bar{\mathcal{D}} \to \mathbb{R}$ for a multivariate domain $\bar{\mathcal{D}} \subset \mathbb{R}^L$. For a function $\bar{f} : \bar{\mathcal{D}} \to \mathbb{R}$, the following notation is used for integration with respect to (w.r.t.) $\bar{F}$:

$$
\langle \bar{f} \rangle := \int_{\bar{\mathcal{D}}} \bar{f}(x_1, \ldots, x_L)\, d\bar{F}(x_1, \ldots, x_L) = E(\bar{f}(\Theta_1, \ldots, \Theta_L)).
$$

Consider a system of polynomials $\{\bar{p}_\alpha : \bar{\mathcal{D}} \to \mathbb{R} \mid \alpha = (\alpha_1, .., \alpha_L) \in \mathbb{N}_0^L\}$, where the polynomial $\bar{p}_\alpha(x_1, \ldots, x_L)$ has degree in the i-th variable $deg_{x_i}(\bar{p}_\alpha) = \alpha_i$. In adaption to Definition 8.24 in [23], we call this an *infinite system of orthogonal polynomials w.r.t.* $\bar{F}$, if for all multi indices $\alpha, \beta \in \mathbb{N}_0^L$ one has

$$
\langle \bar{p}_\alpha \bar{p}_\beta \rangle = 0 \qquad \text{for } \alpha \neq \beta,
$$
$$
\langle \bar{p}_\alpha^2 \rangle =: \bar{\gamma}_\alpha > 0.
$$

The existence of orthogonal polynomial systems follows from the Gram Schmidt algorithm, if for all $\alpha = (\alpha_1, \ldots, \alpha_L) \in \mathbb{N}_0^L$ the moments $\langle x_1^{\alpha_1} \cdot \ldots \cdot x_L^{\alpha_L} \rangle$ are finite. Uniqueness of the orthogonal polynomials is then given up to multiplication by constants. If the Θ_i are independent, they are in particular given by the product of the orthogonal polynomials w.r.t. the distribution of each Θ_i.

In the following, $L^p_{dF}(D, H)$ denotes the space of all functions $D \to H$ that are p-times integrable w.r.t. the measure dF for some $D \subset \mathbb{R}^n$ and codomain H. If dF is not explicitly defined, the Lebesgue measure is chosen. If D and H are not defined, then D equals the domain of F and H equals $\mathbb{R}$.

It is well known that under certain circumstances, orthogonal polynomials span the space $L^2_{d\bar{F}}$. They are thus called a complete orthogonal basis of $L^2_{d\bar{F}}$.

This is, for example, the case, if $\bar{F}$ is absolutely continuous, has finite moments and either $(\Theta_1, \ldots, \Theta_L)$ realizes in a compact domain almost surely or the density of $\bar{F}$ is exponentially integrable. For details, see [24]. If the Θ_i are independent, the orthogonal polynomials w.r.t. $\bar{F}$ span $L^2_{d\bar{F}}$, if all orthogonal polynomial systems w.r.t. the density of Θ_i span the corresponding L^2 spaces. This is due to the tensor product representation of $L^2_{d\bar{F}}$ in case of independency of the Θ_i, see Example E.10 in [25].

Assuming such circumstances to be given, the gPC expansion can be defined.

Theorem 1 (adaption of section 11.3 in [23]). *Let $\Theta_1, \ldots, \Theta_L : \Omega \to \mathbb{R}$ be random variables with joint distribution $\bar{F}$ such that the orthogonal polynomials $(\bar{p}_\alpha)_{\alpha \in \mathbb{N}_0^L}$ w.r.t. $\bar{F}$ form a complete basis of $L^2_{d\bar{F}}$. Denote by $\mathcal{H}$ an arbitrary Hilbert space, e.g., the real numbers $\mathbb{R}$ or a space of the form $L^p(D, \mathbb{R})$, $p = 0, 1, 2$, for some domain $D \subset \mathbb{R}^n$. Then every random variable $X : \Omega \to \mathcal{H}$ with*

$$X =^d \tilde{X}(\Theta_1, \ldots, \Theta_L) \tag{6}$$

in distribution for a function $\tilde{X} \in L^2_{d\bar{F}}(\bar{\mathcal{D}}, \mathcal{H})$ can be represented in the generalized Polynomial Chaos (gPC) form

$$X =^d \sum_{\alpha \in \mathbb{N}_0^L} x_\alpha \bar{p}_\alpha(\Theta_1, \ldots, \Theta_L) \qquad with \qquad x_\alpha = \frac{\langle X \bar{p}_\alpha \rangle}{\langle \bar{p}_\alpha^2 \rangle} \in \mathcal{H}. \tag{7}$$

The proof follows in analogy to the proof for independent continuous random variables in Section 11.3 in [23] from the tensor product decomposition $L^2_{d\bar{F}} \otimes \mathcal{H} \cong L^2_{d\bar{F}}(\bar{\mathcal{D}}, \mathcal{H})$.

Assuming $\Sigma \in L^2_{d\bar{F}}$, Theorem 1 gives the decomposition of the volatility

$$\Sigma(\Theta_1, \ldots, \Theta_L) := \sum_{\alpha \in \mathbb{N}_0^L} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \ldots, \Theta_L). \tag{8}$$

To fit the model to the data, we truncate the series by bounding the maximum polynomial degree $|\alpha| := \alpha_1 + \ldots + \alpha_L$ by $K \in \mathbb{N}_0$

$$\Sigma^K(\Theta_1, \ldots, \Theta_L) := \sum_{\alpha \in \mathbb{N}_0^L, |\alpha| \leq K} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \ldots, \Theta_L). \tag{9}$$

We propose a momentum constrained maximum likelihood approach to fit the gPC coefficients σ_α to discrete real-world data. This facilitates the computation.

The values of the volatility of an asset can be obtained, e.g., by calculating the implied volatilities from corresponding European options. This generates a dataset of implied volatilities representing observations of the random variable $\Sigma^K(\Theta_1, \ldots, \Theta_L)$. When fitting the volatility to the data, we constrain ourselves to those tuples of coefficients corresponding to a volatility $\Sigma^K(\Theta_1, \ldots, \Theta_L)$ whose first moments coincide with the empirical moments of the dataset. The choice of these constraints reduces the dimension of the parameter space for the likelihood function while important characteristics of the distribution—the statistical moments—are maintained.

We illustrate the technique on the simple case of two independent random variables and truncation at $K = 1$:

Example 1. *Let* Θ_1, Θ_2 *be two independent random variables of known distribution and w.l.o.g. expected value* 0 *and variance* 1. *Assume the densities of their distributions exist and denote them by* f_1, f_2, *respectively. Now consider the random volatility truncated to maximum polynomial degree* 1

$$\Sigma^1(\Theta_1, \Theta_2) = \sigma_{00}\bar{p}_{00}(\Theta_1, \Theta_2) + \sigma_{01}\bar{p}_{01}(\Theta_1, \Theta_2) + \sigma_{10}\bar{p}_{10}(\Theta_1, \Theta_2) = \sigma_{00} + \sigma_{01}\Theta_2 + \sigma_{10}\Theta_1$$

when choosing orthonormal polynomials. Assume that values of the volatility $y_1, \ldots, y_M$ *are given. The plain maximum likelihood method would maximize the joint density h of the realizations* $y_1, \ldots, y_M$ *of* $\Sigma^1(\Theta_1, \Theta_2)$ *w.r.t. the three coefficients* $(\sigma_{00}, \sigma_{01}, \sigma_{10}) \in U \subset \mathbb{R}^3 \setminus \{0\}$:

$$(\sigma_{00}, \sigma_{01}, \sigma_{10}) = \underset{(\sigma_{00},\sigma_{01},\sigma_{10})\in U}{\operatorname{argmax}} h(y_1, \ldots, y_M \mid (\sigma_{00}, \sigma_{01}, \sigma_{10}))$$

$$= \underset{(\sigma_{00},\sigma_{01},\sigma_{10})\in U}{\operatorname{argmax}} \prod_{i=1}^{M} \frac{1}{|\sigma_{10}\sigma_{01}|} \int_{t\in\mathbb{R}} f_1\left(\frac{x - \sigma_{00} - t}{\sigma_{10}}\right) f_2\left(\frac{t}{\sigma_{01}}\right) dt.$$

Constraining the maximum likelihood estimator to be exact in the first moment, i.e., the expected value, gives

$$\mathrm{E}(\Sigma^1(\Theta_1, \Theta_2)) = \sigma_{00} \overset{!}{=} \bar{y}_M := \mathrm{mean}(y_1, \ldots, y_M).$$

This reduces the optimization to two variables

$$(\sigma_{01}, \sigma_{10}) = \underset{(\sigma_{01},\sigma_{10})\in\hat{U}}{\operatorname{argmax}} \prod_{i=1}^{M} \frac{1}{|\sigma_{10}\sigma_{01}|} \int_{t\in\mathbb{R}} f_1\left(\frac{x - \bar{y}_M - t}{\sigma_{10}}\right) f_2\left(\frac{t}{\sigma_{01}}\right) dt,$$

where $\hat{U} = U \cap \{\bar{y}_M\} \times \mathbb{R}^2$.

Further constraints can be used to reduce the complexity even more: Claiming, e.g., the variance to coincide with the empirical variance $\mathrm{Var}(\Sigma^1(\Theta_1, \Theta_2)) \overset{!}{=} S_M^2$ *gives the additional relation*

$$\sigma_{10} = \pm\sqrt{S_M^2 - \sigma_{10}^2}$$

reducing the optimization to one variable and the sign of σ_{10}. *This ends our example.*

3. Deriving the System of PDEs for the gPC Coefficients

The stochastic Galerkin method is applied to the Black Scholes Equation (5) with uncertain volatility in order to transform the stochastic PDE into a system of deterministic PDEs for the gPC coefficients of the solution $V(S, t, \Theta_1, \ldots, \Theta_L)$.

To do so, one has to assume $\Sigma \in L^2_{d\bar{F}}$ and $V \in L^2_{d\bar{F}}(\bar{\mathcal{D}}, L^2((0, \infty) \times [0, T], \mathbb{R}))$, such that Theorem 1 can be applied. In analogy to the one-dimensional case in [17,18], the thus derived gPC expansions are inserted in the Black Scholes Equation (5). Multiplication of the equation by $\bar{p}_\delta(\Theta_1, \ldots, \Theta_L)$ and application of the expected value, for each $\delta \in \mathbb{N}_0^L$ at a time, yields the equations

$$0 = \frac{\partial v_\delta(S,t)}{\partial t} + \frac{1}{2}S^2 \sum_{\alpha,\beta,\gamma\in\mathbb{N}_0^L} \sigma_\alpha\sigma_\beta \frac{\partial^2 v_\gamma(S,t)}{\partial S^2} M_{\alpha\beta\gamma\delta} + rS\frac{\partial v_\delta(S,t)}{\partial S} - rv_\delta(S,t)$$

due to orthogonality of the p_α. Note that the Galerkin multiplication tensor $M_{\alpha\beta\gamma\delta} := \frac{\langle \bar{p}_\alpha\bar{p}_\beta\bar{p}_\gamma\bar{p}_\delta \rangle}{\langle \bar{p}_\delta^2 \rangle}$ exists since the integrated functions are all polynomials in L variables.

In order to solve the system, the boundary conditions and the final condition corresponding to the considered financial derivative are transformed to conditions on the gPC coefficients v_i. Usually, they are deterministic and thus appear in the coefficient $v_{(0,\ldots,0)}$, whereas all other coefficients vanish.

We truncated the gPC expansions up to maximum polynomial degrees $K, N \in \mathbb{N}_0$ and obtained representation (9) for Σ^K and

$$V^N(S, t, \Theta_1, \ldots, \Theta_L) := \sum_{\delta \in \mathbb{N}_0^L, |\delta| \leq N} v_\delta^N(S, t) \bar{p}_\delta(\Theta_1, \ldots, \Theta_L) \tag{10}$$

for coefficients $v_\delta^N \in L^2((0, \infty) \times [0, T], \mathbb{R})$.

The system of equations for the truncated gPC coefficients v_δ^N, $\delta \in \mathbb{N}_0^L$ with $|\delta| \leq N$, is then given by

$$0 = \frac{\partial v_\delta^N(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{\substack{\alpha, \beta, \gamma \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K, \\ |\gamma| \leq N}} \sigma_\alpha \sigma_\beta \frac{\partial^2 v_\gamma^N(S, t)}{\partial S^2} M_{\alpha\beta\gamma\delta} + rS \frac{\partial v_\delta^N(S, t)}{\partial S} - r v_\delta^N(S, t), \tag{11}$$

which can be evaluated numerically. For demonstrative purposes, we use a finite difference scheme, see Appendix A.

Note, however, that convergence of the truncated stochastic Galerkin solution V^N in (10) to the true solution V as $N \to \infty$ is not obvious and could not be proven to date. It is a topic open to further research. From now on, we assume convergence.

4. A Bi-Fidelity Approach for Calculating the Stochastic Galerkin Solution to the Black Scholes Equation with Random Volatility

If the volatility depends on just $L = 2$ random variables, the stochastic Galerkin (SG) solution truncated at maximum degree N already has $(N+1)(N+2)/2$ gPC coefficients. Thus, $(N+1)(N+2)/2$ coupled equations have to be solved to obtain the approximate SG solution. The number of equations and with it the computational cost rapidly increase as N or L increases.

This becomes a problem especially if the SG solutions for many options shall be computed at a time, e.g., for risk management evaluations of derivatives with different underlying assets. The Bi-Fidelity approach provides a solution to this problem, if the same type of option (e.g., European Call options) with the same maturity T and interest rate r, but different distributions of the volatility model $\Sigma(\Theta_1, \ldots, \Theta_L)$ are considered. A situation like this arises, for instance, when comparing financial derivatives of the same type and maturity date but with different underlying stochastic assets.

In general, given a PDE depending on a random variable Ξ, the Bi-Fidelity method aims to approximate the desired high fidelity solution at a certain realization z of Ξ by pre-stored high and low fidelity solutions in some other realizations of Ξ and the computationally cheaper low fidelity solution in z. This can be considered a machine learning approach because the properties of the equation are learned offline by picking suitable realizations for the stored approximation data.

The application of Bi-Fidelity techniques to various problems is an active area of research [26,27]. In the setting of uncertainty quantification for PDE models, it is frequently described in the context of uncertainty quantification via Stochastic Collocation methods; see, e.g., [28,29] for the general procedure or [30–33] for applications. The combination with the stochastic Galerkin method works similarly; however, it is not very common in literature.

At first, the random variable Ξ has to be assigned. In our case, it is not given by $(\Theta_1, \ldots, \Theta_L)$, since we still want our solution to be a random variable depending on the Θ_i in order to explore its stochastic behaviour. Instead, we suppose the distribution of $\Sigma(\Theta_1, \ldots, \Theta_L)$ to change from calculation to calculation, as it would be the case for different underlying assets, without changes in the distributions of the Θ_i. This could reflect different sensitivities of different stochastic assets to the factors Θ_i. By representation (9) of the truncated gPC expansion of Σ, a variation in the distribution of the volatility means a

variation in at least one of the gPC coefficients σ_α, $|\alpha| \leq K$. Hence, the random variable Ξ describes volatility models of the form (9) by their gPC coefficients σ_α, $|\alpha| \leq K$.

Then, high and low fidelity models have to be defined. The high fidelity model is the one we are actually interested in. We choose a high-resolution numerical solution to (11) derived by the explicit finite difference scheme (A3) with a large number of grid points in the S-space M_ζ^H and in time N_τ^H. The low fidelity model, i.e., the cheaper model that is less trusted but used for the approximation rule, is chosen to be the same numerical solution on a coarse grid with small M_ζ^L and N_τ^L.

Note, however, that N_τ^L must not be chosen too small; otherwise, the stability of the scheme could be violated for a large number of volatility models. The reason for this requirement will become clear at step 1 of the offline data generating steps.

It is also possible to consider different numerical schemes or even different but similar underlying equations for the high and low fidelity models. However, important characteristics of the solution must be shared between the models.

Now one can proceed with the typical Bi-Fidelity algorithm as described in [28–30]. Below, the application of this algorithm is explained, where the volatility is assumed to depend on $L = 2$ random variables Θ_1, Θ_2 for a better readability. An extension to more random variables can easily be achieved. The truncation number $K = 1$ is chosen such that the random variable Ξ represents the gPC coefficients σ_{00}, σ_{10} and σ_{01} of the volatility as in Example 1.

Since the actual computational effort lies in the calculation of the transformed system of Equation (A2), the Bi-Fidelity approach is applied directly on the transformed $\bar{\mathbf{v}}$. Thus, a transformation back to the original variables $\mathbf{v}^N$, S and t is performed only once for the Bi-Fidelity solution, reducing the computational cost. For the calculation of the scheme, initial conditions and the Galerkin multiplication tensors are pre-stored and reused.

The following three steps describe the *offline learning phase* of the algorithm in which the stored approximation data are generated [28–30]. These steps have to be executed only once.

Step 1:   At first, the codomain of Ξ is described by finite intervals $\sigma_{00} \in [a_{00}, b_{00}]$, $\sigma_{10} \in [a_{10}, b_{10}]$, $\sigma_{01} \in [a_{01}, b_{01}]$.

The intervals can, for instance, be constructed by calculation of $\sigma_{00}, \sigma_{10}, \sigma_{01}$ for some of the stochastic assets of interest. Alternatively, one can think of possible values of σ_{00} inspired by similar experiments and choose bounds of σ_{10} and σ_{01} such that the variance of $\Sigma(\Theta_1, \Theta_2)$ is bounded by some predefined value. We used this approach in the calculations of Section 5.

After that, a large set Y of possible realizations of Ξ has to be chosen such that it is a good 'cover' of the possible values of Ξ. One can use Monte Carlo sampling or a structured grid on the codomain of Ξ.

For every volatility model described by a $y \in Y$, the low fidelity solution $\bar{\mathbf{v}}^L(y)$ is computed, if the corresponding system of equations is parabolic and the low fidelity scheme is stable.

Step 2:   Since one can usually not afford to calculate the high fidelity solution in every $y \in Y$, one has to determine the most important points. Let $A \in \mathbb{N}$ denote the number of high fidelity computations one can afford, then this can be achieved by choosing $z_0 := \mathrm{argmax}_{y \in Y} d^L(\bar{\mathbf{v}}^L(y), 0))$ and

$$z_{i+1} := \underset{y \in Y}{\mathrm{argmax}}\, d^L(\bar{\mathbf{v}}^L(y), \bar{\mathbf{V}}^L(z_1, \ldots, z_i)), \quad i = 0, \ldots, A-1. \tag{12}$$

The notation $\bar{\mathbf{V}}^L(\hat{Y}) := \mathrm{span}(\bar{\mathbf{v}}^L(\hat{y}) \mid \hat{y} \in \hat{Y})$ for $\hat{Y} \subset Y$ is used for the span of the solutions to previously picked $z_i \in \hat{Y}$. Then $d^L(u, V) := \inf_{v \in V} \|u - v\|^L$ is the distance of a point $v \in \bar{\mathbf{V}}^L(Y)$ to the set $V \subset \bar{\mathbf{V}}^L(Y)$. We used a greedy procedure for the point selection; for further details on the computation, compare Algorithm 1 in [29].

This step selects the points $z_1, \ldots, z_A$ whose solutions span the 'largest' subspace $\bar{\mathbf{V}}^L(z_1, \ldots, z_A)$ of $\bar{\mathbf{V}}^L(Y)$.

Step 3: The high fidelity solution is calculated in the thus derived points $z_1, \ldots, z_A$. Note that N_τ^H has to be chosen large enough such that the numerical scheme is stable for all volatility models z_i. The parabolicity of the system of PDEs does not need to be checked again, as it has been checked in step 1 already. The pairs of high and low fidelity solutions $\bar{\mathbf{v}}^H(z_i), \bar{\mathbf{v}}^L(z_i)$ are stored.

Assume now that a certain volatility model z is given and the corresponding Bi-Fidelity solution of the Black Scholes equation with uncertain volatility shall be computed. This is performed in the *online phase* as follows [28–30]:

Step 1: The low fidelity solution $\bar{\mathbf{v}}^L(z)$ is calculated by scheme (A3). Note that the system of equations needs to be parabolic and the scheme has to be stable for a reasonable calculation.

Step 2: The low fidelity solution $\bar{\mathbf{v}}^L(z)$ is projected onto $\bar{\mathbf{V}}^L(z_1, \ldots, z_A)$ leading to the projection formula

$$\bar{\mathbf{v}}^L(z) \approx P_{\bar{\mathbf{V}}^L(z_1, \ldots, z_A)} \bar{\mathbf{v}}^L(z) = \sum_{k=1}^{A} c_k \bar{\mathbf{v}}^L(z_k)$$

with projection coefficients $c_k \in \mathbb{R}$. Here $P_{\mathbf{V}}\mathbf{y}$ denotes the orthogonal projection of $\mathbf{y}$ onto $\mathbf{V}$. Details of the computation of the c_k can be found in [29], for instance.

Step 3: Finally, the Bi-Fidelity solution is constructed by applying the same projection law to the stored high fidelity solutions

$$\bar{\mathbf{v}}^{BF}(z) := \sum_{k=1}^{A} c_k \bar{\mathbf{v}}^H(z_k).$$

After deriving $\bar{\mathbf{v}}^{BF}$, it has to be transformed back to the original variables $\mathbf{v}$, S and t.

5. Numerical Results

This section presents numerical solutions to the Black Scholes equation with uncertain volatility. For the sake of simplicity, the volatility is assumed to depend on the two independent random variables Θ with standard normal distribution and Δ uniformly distributed on $[-0.5, 0.5]$. The properties of such models are investigated and the model is tested on real data. Furthermore, the error of the Bi-Fidelity approximation is investigated and its computation time is compared to the high fidelity model.

For more convenient reading, times t and the maturity T are given in days, whereas for the computations, these values were multiplied by $1/251$ to go over to years.

5.1. Results for the Extended Model

The numerical solution to the truncated system of Equation (11) for a European Call option with a strike price *strike* $= 100$ and maturity $T = 20$ in a market with risk-free rate of interest $r = 0$ is visualized in Figure 1a,b by plotting its mean and variance.

The volatility of the underlying stochastic asset is modelled by

$$\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta \tag{13}$$

for Θ standard normally distributed and Δ uniformly distributed on $[0.5, 0.5]$. The random variables are modelled to be independent. For the gPC expansion of the solution, the truncation number $N = 5$ was chosen, for which system (11) is parabolic. The numbers of grid points $M_\zeta = 200$ in ζ and $N_\tau = 319$ in τ were chosen such that the applied explicit finite difference scheme (A3) is stable.

Contour lines were drawn at a height of quarters of the maximum absolute value and the borders of the smoothing area, i.e., the area where the solution differs from its final

condition $V(S, T) = (S - strike)^+$, were drawn in red. These lines will be present in each of the following surface plots. Note that the expected value surface resembles the solution of the deterministic Black Scholes equation for $\sigma = 0.5$ in Figure 1c, but the smoothing area is larger.
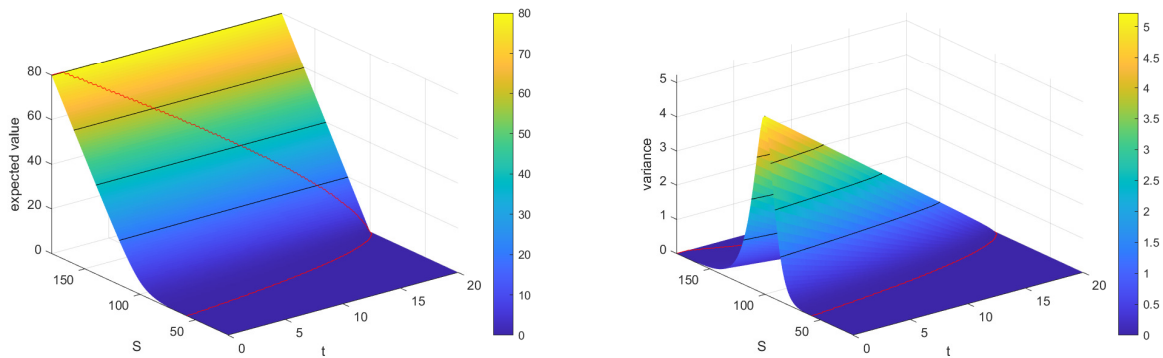
(**a**) Expected value surface for the stochastic solution.

(**b**) Variance surface for the stochastic solution.
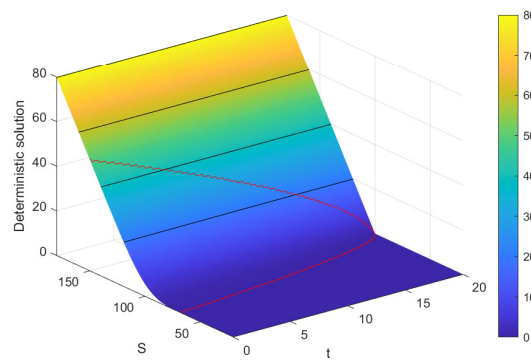
(**c**) Deterministic solution.

Figure 1. Solutions to the Black Scholes equation for a European Call option with $T = 20, strike = 100$ and $r = 0$ for the volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normally and Δ uniformly distributed, in (**a**,**b**) and the deterministic model $\sigma = 0.5$ in (**c**) calculated with $K = 1, N = 5$, $M_\zeta = 200, N_\tau = 319$.

Experiments showed that the qualitative shape of the expected value and variance is characteristic for solutions to the Black Scholes equation with random volatility (5) of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$. These models lead to solutions that 'lie between' the solutions for volatility that depends on Θ or Δ only and has the same mean and variance.

The higher σ_{10} is in comparison to σ_{01}, the closer the solution is to the solution for volatility depending on Θ only and the further away it is from the solution for the model depending on Δ only, and vice versa. An increase in the mean σ_{00} of the volatility while keeping its variance constant was observed to enlarge the smoothing area and thus the spread of the variance, which in turn flattens it.

A rise in the variance $\sigma_{10}^2 + \sigma_{01}^2/12$ of the volatility with constant mean σ_{00}, however, seemed to scale up the variance of the SG solution by the same factor. Meanwhile, the expected value of the SG solution was marginally increased within the smoothing area.

5.2. Comparison to Real Market Data

The model is compared to market prices of a European Call option, whose end of day values are considered from 7 January 2019 to 20 September 2019. (The values were obtained from https://www.finanzen.net/, accessed on 21 September 2019). Its underlying

asset is the DAX index, and the strike price and maturity are given by *strike* = 10,275 and
$T = 180$ days, respectively.

A volatility model of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$ was fitted to the daily
implied volatilities by using the moment constrained maximum likelihood approach from
Example 1 with the two moments mean and variance. This lead to the volatility model

$$\Sigma(\Theta, \Delta) = 0.2292 + 0.1126\Theta + 0.0115\Delta, \tag{14}$$

whose fitted density is shown in Figure 2a together with a histogram density estimator.
The SG solution was computed using the truncation number $N = 5$ and the numbers of
grid points $M_\zeta = 200$ and $N_\tau = 678$. With these values, the numerical scheme is stable and
the system of Equation (11) is parabolic.

Figure 2b shows the market prices and the expected value of the SG solution as
well as the range expected value plus/minus standard deviation and the solution to the
deterministic Black Scholes equation with volatility $\sigma = E(\Sigma(\Theta, \Delta))$. A more detailed plot
of those graphs for the last 55 days of the option is given in Figure 2c. One observes that
the expected value of the SG solution is very close to the data on these days but slightly
above the data at earlier times. However, the data are always in the range expected value
plus/minus standard deviation, as one would expect from stochastic theory.
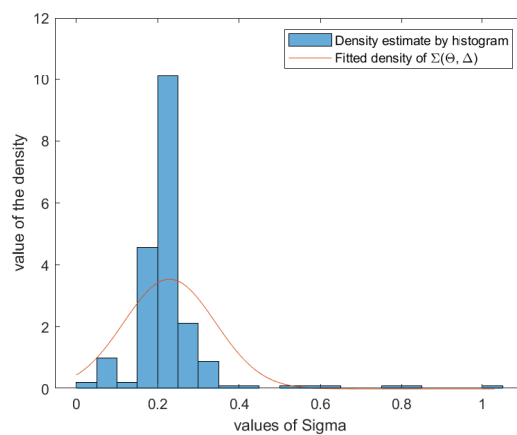
(**a**)
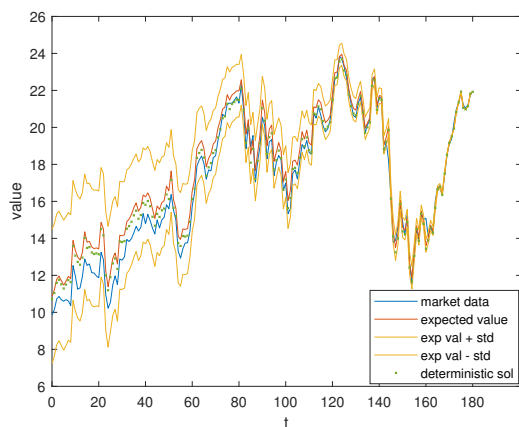
(**b**)
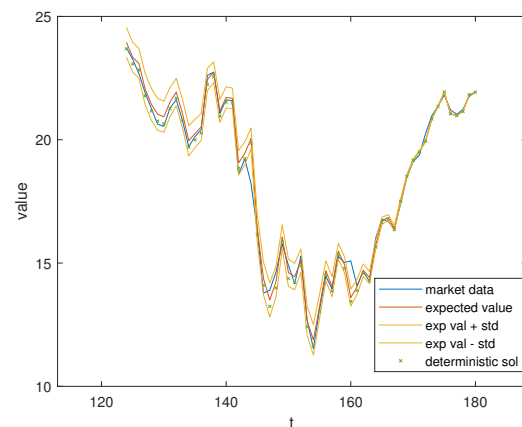
(**c**)

Figure 2. Comparing the stochastic model to real market data. (**a**) Histogram density estimator and
density of $\Sigma(\Theta, \Delta)$ fitted to the implied volatilities by constrained maximum likelihood. (**b**) Market
values of the option together with the expected value of the SG solution and the range expected value
plus minus standard deviation. (**c**) Detailed look on the last 55 days.

A comparison to the deterministic solution shows that it also lies above the market data for early times. Recall that unlike the deterministic solution, the SG solution allows realizations to differ from the expected value within a certain range.

5.3. Comparing Bi-Fidelity Solution and High Fidelity Solution

The Bi-Fidelity solution of the Black Scholes equation with uncertain volatility (5) following volatility model (13) for a European Call option is compared to its high fidelity solution. After that, a simulation is performed to find the error in expected value and in variance between the Bi-Fidelity solution and the high fidelity solution. The error is characterized in size and shape by a Monte Carlo simulation for different volatility models. Finally, the computation times for high fidelity and Bi-Fidelity model are compared.

We go back to the toy model of a market with interest rate $r = 0$ and maturity $T = 23$ of the option. The strike price was set to $strike = 100$ and the gPC expansion of the solution was truncated after a total polynomial degree of $N = 5$ as before.

A rather coarse grid with $M_\zeta^L = 50$ and $N_\tau^L = 150$ was chosen for the low fidelity model. This N_τ^L is high enough such that the vast majority of all low fidelity computations performed in the examples explained below was stable. In the case of instability, the corresponding sample point was removed from the set of low fidelity sample points. The high fidelity solution was computed on a fine grid with $M_\zeta^H + 1 = 350 + 1$ grid points in S direction. The number of grid points $N_\tau^H + 1 = 5853 + 1$ in the time direction was chosen such that all high-resolution computations for important volatility models were stable.

The low fidelity sample points represented volatility models $\Sigma_i(\Theta, \Delta) = \sigma_{00}^{(i)} + \sigma_{10}^{(i)}\Theta + \sigma_{01}^{(i)}\Delta$ with

$$
\begin{aligned}
\sigma_{00}^{(i)} &\in \{0 < 0.05\lambda \le 0.8 \,|\, \lambda \in \mathbb{N} \setminus \{0\}\}, \\
\sigma_{10}^{(i)} &\in \left\{0.05\lambda \le \sqrt{\sigma_{00}/2} \,|\, \lambda \in \mathbb{N}_0\right\} \quad \text{and} \\
\sigma_{01}^{(i)} &\in \left\{0.05\lambda \le \sqrt{12(\sigma_{00}/2 - \sigma_{10}^2)} \,|\, \lambda \in \mathbb{N}_0\right\}.
\end{aligned}
\tag{15}
$$

The coefficients were chosen such that $Var(\Sigma(\Theta, \Delta)) \le \sigma_{00}^{(i)}/2$.

Figure 3a,b shows the expected value surfaces of the high fidelity and the Bi-Fidelity solution for the volatility model $\Sigma(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$. At first glance they seem to approximately coincide.
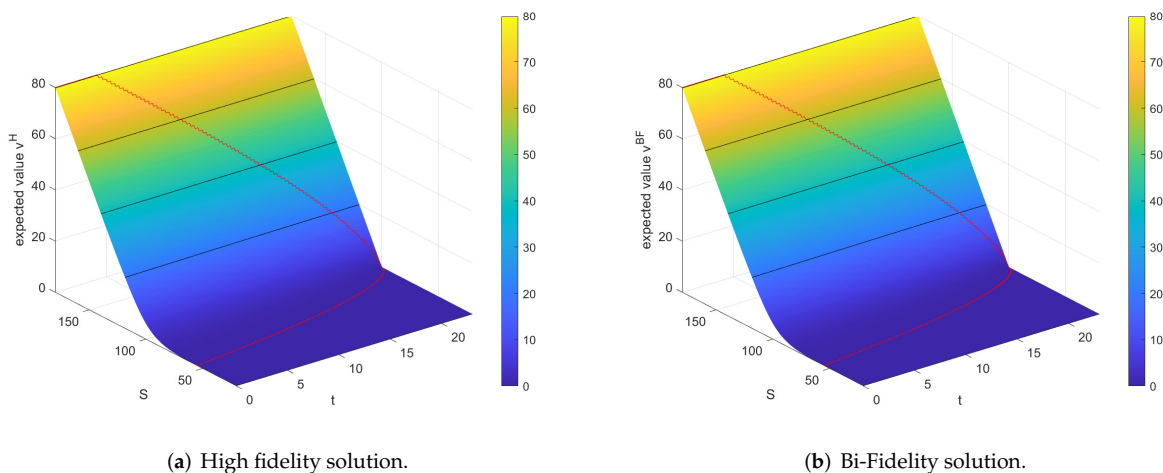

(**a**) High fidelity solution.

(**b**) Bi-Fidelity solution.

Figure 3. Expected value surfaces for high fidelity and Bi-Fidelity solution.

To study the deviations, the absolute difference in expected values is displayed in Figure 4a close to the strike price and Figure 4b for a wider range of S values. One observes a difference of size 10^{-3} within the smoothing area that seems to increase in absolute value as $S \to \infty$. Figure 4c shows the difference for all values of S and t. The maximum absolute value of the absolute difference is less than 0.3 and occurs close to $S = \infty$, where the option values tend to infinity. Therefore, a difference of 0.3 in these regions means a small deviation. The difference in the smoothing area of size $3 \cdot 10^{-3}$ is larger compared to the values attained in this region that are close to zero. Recall, however, that the solution is multiplied by *strike* when transforming back the variables. Hence, an error of size 10^{-3} at strike 100 means an error of size $10^{-5} \cdot strike$.

(**a**) close to the strike price

(**b**) for a wider range of S values
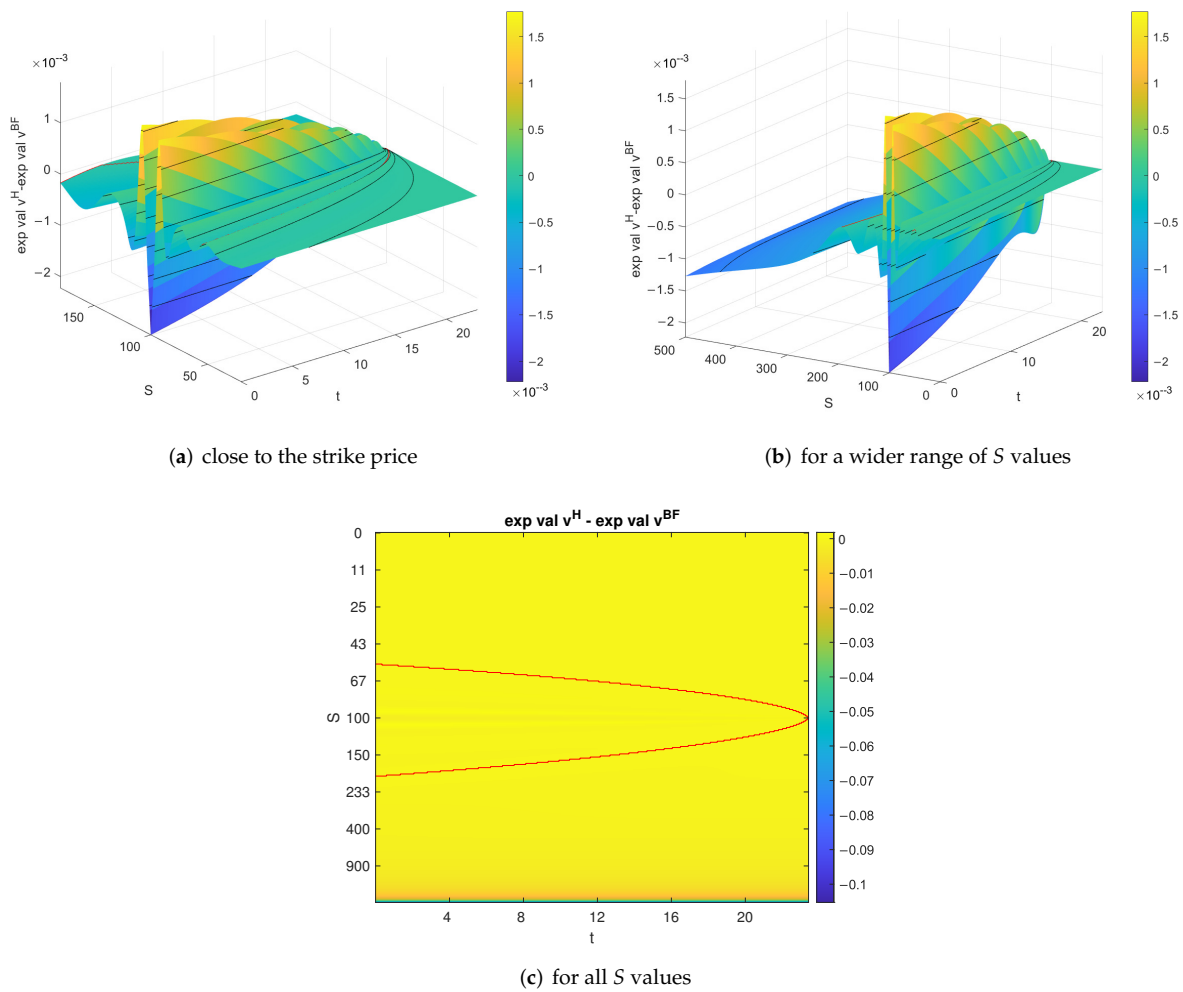

(**c**) for all S values

Figure 4. Absolute difference in expected value of high fidelity and Bi-Fidelity solution.

The variances of high and Bi-Fidelity solution are considered in Figure 5a,b, respectively. The high fidelity variance seems to be a little bit steeper than the Bi-Fidelity variance.
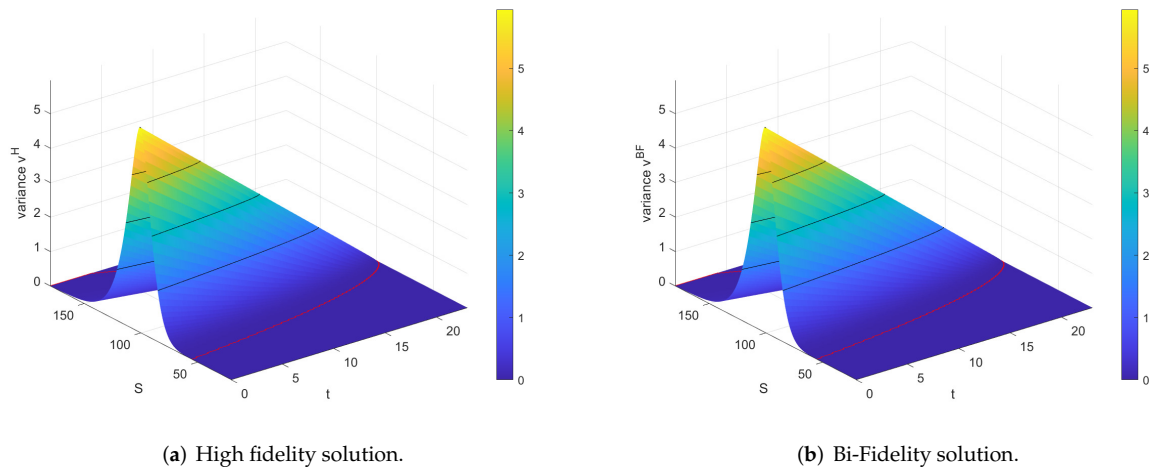
(**a**) High fidelity solution.

(**b**) Bi-Fidelity solution.

Figure 5. Variance surface for high fidelity and Bi-Fidelity solution.

We examine the absolute difference in variance as represented in Figure 6a to lie in the smoothing area. Figure 6b, showing the difference for all S and t values, supports this conclusion. The error is again of size $10^{-3} = 10^{-7} \cdot strike^2$.
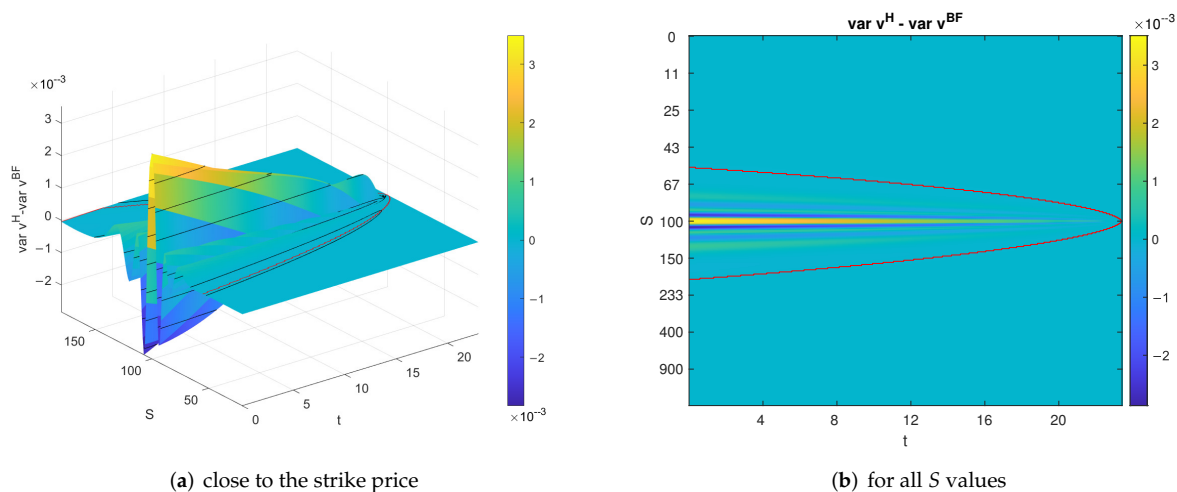

(**a**) close to the strike price

(**b**) for all S values

Figure 6. Absolute difference in variance of high fidelity and Bi-Fidelity solution.

Finally, a simulation was performed to obtain the characteristic size and shape of the Bi-Fidelity error. For this purpose, 300 volatility models of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$ were generated randomly from uniformly distributed random variables $\sigma_{00} \in [0, 0.8], \sigma_{10} \in [0, \sqrt{\sigma_{00}/2}], \sigma_{01} \in [0, \sqrt{12(\sigma_{00}/2 - \sigma_{10}^2)}]$ 'covered' by the grid in (15). Both high and Bi-Fidelity solutions were calculated for each of these volatility models.

The mean absolute difference of the expected values is represented in Figure 7a close to the strike price and Figure 7b for a larger range of S values. Figure 7c is a plot of the error for all S and t values. The smoothing area is not plotted since it differs for every volatility model. The shape of the error is characterized by an oscillation of size $10^{-3} = 10^{-5} \cdot strike$ close to the strike price and a steady increase in absolute value for $S \to \infty$. The maximum absolute difference lies close to $S = \infty$ and has a size of $10^{-2} = 10^{-4} \cdot strike$, which is small in relative terms. This coincides with the error shape in Figure 4a–c and thus seems to be characteristic for the considered Bi-Fidelity model.
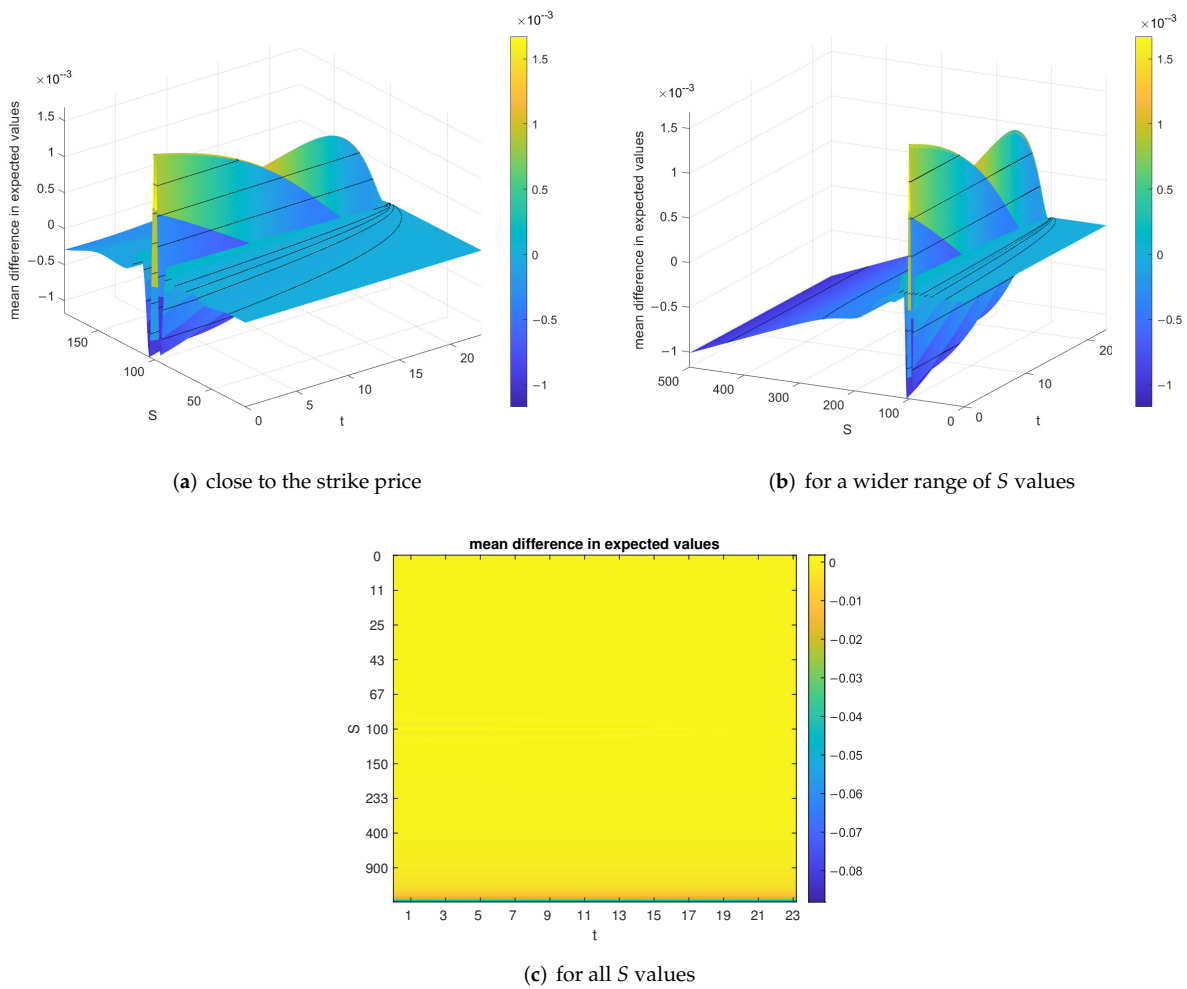
(**a**) close to the strike price

(**b**) for a wider range of *S* values

(**c**) for all *S* values

Figure 7. Mean absolute difference in expected value of high fidelity and Bi-Fidelity solution.

The characteristic error in the variances derived by the same 300 volatility models is displayed in Figure 8a. It shows some oscillation close to the strike price of size $10^{-2} = 10^{-6} \cdot strike^2$ but vanishes elsewhere, as one can observe in Figure 8b.
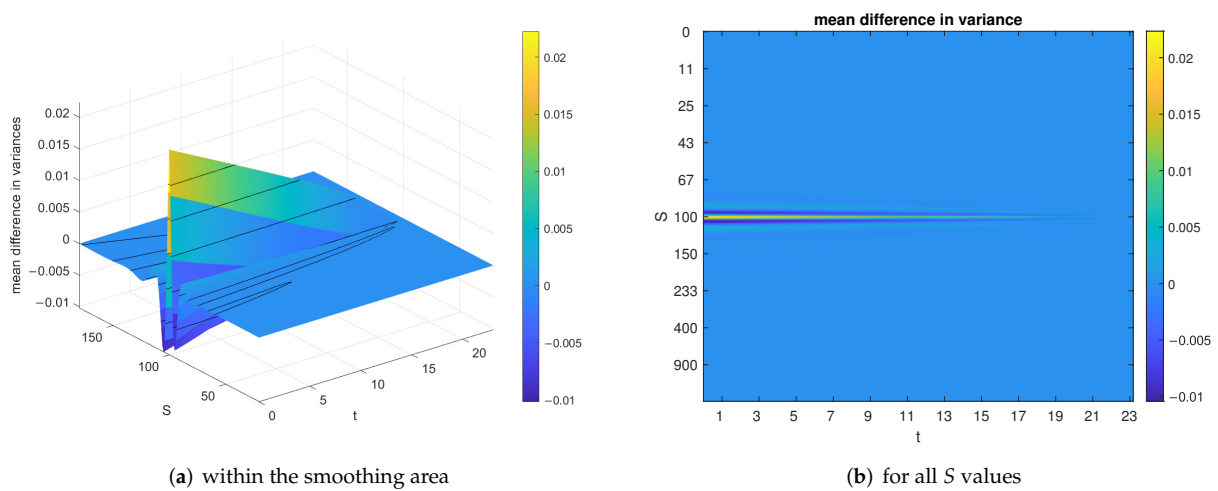

(**a**) within the smoothing area

(**b**) for all *S* values

Figure 8. Mean absolute difference in variance of high fidelity and Bi-Fidelity solution.

This error can possibly be reduced by adding more approximation pairs for the Bi-Fidelity computation or by choosing a low fidelity model closer to the high fidelity model.

Remark 1. *Because the focus of this paper is the illustration of the general technique, a European Call option is used as an example in the above error analysis. If one were to consider other financial options, this might lead to different results. In practice, the method has to be fitted to the specific question and the considered option type and market settings (r, T, the range of the volatility gPC coefficients). In particular, the choice of the numerical method and the desired accuracy can vary for different questions. A modeller could, e.g., be interested in obtaining a very accurate model. They would thus prescribe an error threshold, choose a higher-order scheme for numerics and then search for the grid sizes such that the error threshold is not exceeded (up to some percentile). A trader instead might, for instance, only be interested in a rough prognosis of end-of-day values. They can, therefore, work with coarser grids and less accurate schemes. Because the method behaves differently for different option types, numerical schemes and market settings, a general recommendation on the choice of hyper-parameters cannot be given. Instead, the method would have to be investigated for each specific application.*

In general, however, one can say that the Bi-Fidelity error increases with the difference in the low and high fidelity model. In our case this means choosing an even coarser grid as a low fidelity model and/or an even finer grid for the high fidelity models might introduce an additional model approximation error. However, choosing the low fidelity model very close to the high fidelity model cancels the computational advantage of the technique. A trade-off has to be made considering the specific situation.

Furthermore, increasing the number A of approximation data pairs should increase the accuracy up to some point. If A is not prescribed by computational resources, one could, e.g., add new approximation points in (12) until the considered distance is below some threshold for all remaining realizations.

5.4. Comparison of Computation Times

For demonstration, the above Bi-Fidelity model and the high fidelity model with the same number of grid points $M_\zeta^H = 350$ and $N_\tau^H = 5853$ were calculated in the same 300 randomly generated volatility models. Every model $\Sigma^{(i)}(\Theta, \Delta) = \sigma_{00}^{(i)} + \sigma_{10}^{(i)}\Theta + \sigma_{01}^{(i)}\Delta$ belonging to iteration $i \in \{1, \ldots, 300\}$ was generated such that it satisfies the same bounds on the coefficients $\sigma_{00}^{(i)} \in (0, 0.8], \sigma_{10}^{(i)} \in [0, \sqrt{\sigma_{00}/2}]$ and $\sigma_{01}^{(i)} \in [0, \sqrt{12(\sigma_{00}/2 - \sigma_{10}^2)}]$ as for the low fidelity sample points in (15). The $\Sigma^{(i)}$ should thus be 'covered' by the low fidelity sample points, which enables a Bi-Fidelity computation. In every calculation, the stability of the scheme w.r.t. the chosen time step is checked. The computation times for both models are plotted in Figure 9.

The mean computation time for the high fidelity model is 173.99 s, whereas the Bi-Fidelity model achieved a mean computation time of 10.68 s per volatility model. Hence, the application of the Bi-Fidelity method accelerated our computations by a factor of 16.3 in mean. For finer high fidelity grids, this difference should further increase. However, choosing a finer grid means introducing a larger difference in the high and low fidelity model, which could lead to larger errors.
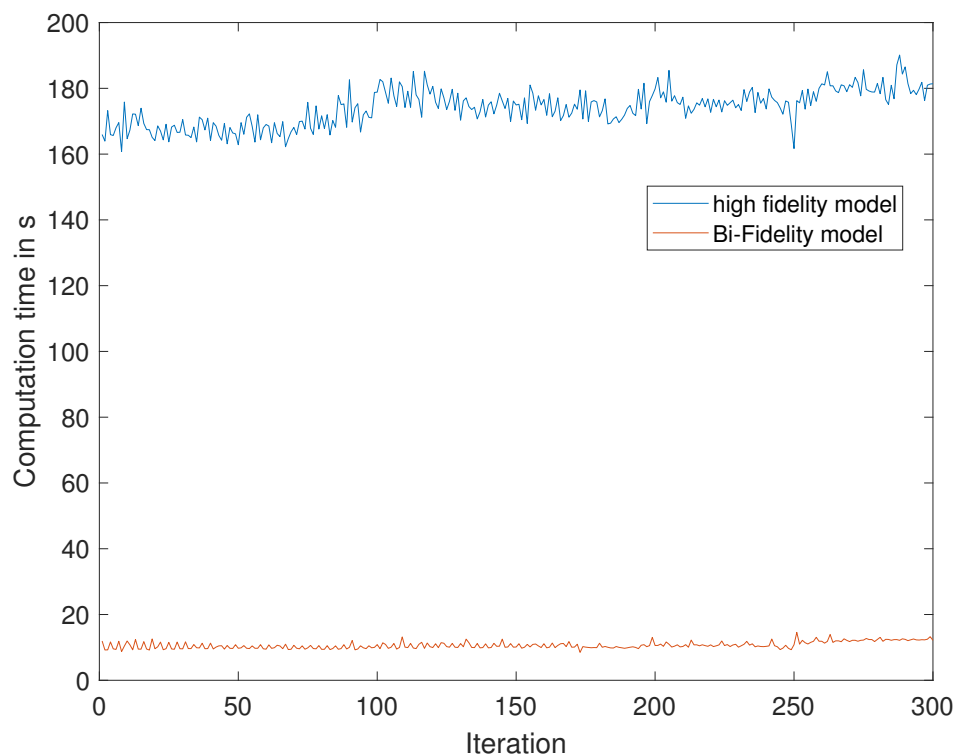
Figure 9. Computation times for the high fidelity model and the Bi-Fidelity model evaluated in the same volatility model.

6. Conclusions

When the volatility in the Black Scholes equation is determined by discrete market data, uncertainty is introduced due to the estimation procedure. We modelled this uncertainty by a dependence on a finite number of random variables representing random factors of influence. A possibility to fit this uncertain volatility to market data was demonstrated. Afterward, the Black Scholes equation with uncertain volatility was used to model the price process of a derivative. Under certain assumptions, the random volatility and the stochastic solution can be represented by their generalized Polynomial Chaos (gPC) expansions allowing the application of the stochastic Galerkin method. The resulting deterministic system of PDEs for the gPC coefficients was truncated and solved numerically by a finite difference scheme.

Numerical examples showed that the expected value of this stochastic model fitted real market data in a similar way as a deterministic model. However, the stochastic solution allows deviations from its expected value within a certain range and it can be used for calculations of further stochastic quantities such as the variance of the solution or in risk management applications.

However, computation can become costly for a large number of random variables or a late truncation. This is due to the fast increase in the number of gPC coefficients. Therefore, a machine learning technique was presented to reduce the computation cost for computing the solutions for different volatility models within the same setting (option type, maturity, interest rate, maximum polynomial degree). The so-called Bi-Fidelity approach approximates a costly solution on the basis of a computationally cheaper solution and some pre-stored costly solutions for wisely selected volatility models.

For a European Call option, the maximum absolute difference in the expected value of the Bi-Fidelity solution to the desired solution was experimentally observed to be of size $10^{-5} \cdot strike$ in mean close to the strike price and increase to size $10^{-4} \cdot strike$ in mean for $S \to \infty$, where the expected value also tends to ∞. The maximum difference in variance attained a value of size $10^{-6} \cdot strike^2$ in mean. Meanwhile, the mean computation time was decreased by a factor of 16.3.

A topic that is still open to further research is the convergence of the truncated gPC expansion of the stochastic solution to the true solution as the truncation number goes to infinity. If convergence is assumed to hold, it is also possible to solve the deterministic system of PDEs for the gPC coefficients with a different numerical technique and apply the Bi-Fidelity approach to this solution. Furthermore, one could think of applying the technique used in this paper to the Black Scholes equation with uncertain volatility and interest rates, when there are doubts concerning its true value, or to familiar equations such as the Black Scholes equation for multiple assets or the bond equation.

Appendix A. Finite Difference Scheme

For demonstrative purposes, European Call options with strike price *strike* and maturity T will be considered to present the finite difference scheme used for solving the system of Equation (11).

For an easier implementation, system (11) is rewritten in vector form. This is performed via a bijection ϕ from the set $\{0, \ldots, |I| - 1\}$ of positions in the vector to the set of multi indices $I := \{\delta \in \mathbb{N}_0^L \,|\, |\delta| \leq N\}$, as described in Section 5.2 in [34]. Define $\mathbf{v} := (v_{\phi(0)}^N, \ldots, v_{\phi(|I|-1)}^N)^T$, then one can represent system (11) by

$$\mathbf{0}_{|I|} = \frac{\partial \mathbf{v}(S,t)}{\partial t} + \frac{1}{2}S^2 \mathbf{A} \frac{\partial^2 \mathbf{v}(S,t)}{\partial S^2} + rS \frac{\partial \mathbf{v}(S,t)}{\partial S} - r\mathbf{v}(S,t),$$

where the coupling matrix $\mathbf{A}$ is given by

$$\mathbf{A}[n,l] = \sum_{\substack{\alpha,\beta \in \mathbb{N}_0^L, \\ |\alpha|,|\beta| \leq K}} \sigma_\alpha \sigma_\beta M_{\alpha\beta(\phi(l))(\phi(n))}, \qquad \text{for } n, l = 0, \ldots, |I| - 1. \tag{A1}$$

The boundary conditions and final values have to be transformed to vectors as well. If the deterministic part with multiindex $(0, \ldots, 0)$ is in the first vector position and a European Call option is considered, they are given by

$$\mathbf{v}(S, T) = \begin{pmatrix} (S - strike)^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; S \in (0, \infty),$$

$$\mathbf{v}(S, t) \xrightarrow{S \to 0} \mathbf{0}_{|I|}, \qquad t \in [0, T], \quad \text{and}$$

$$\frac{1}{S}\mathbf{v}(S, t) \xrightarrow{S \to \infty} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad t \in [0, T].$$

The system has to be transformed into a finite domain. For the European Call option, this can be achieved by the following transformation of variables

$$\zeta := \frac{S}{S + strike},$$
$$\tau := T - t,$$
$$\bar{\mathbf{v}}(\zeta, \tau) := \frac{\mathbf{v}(S, t)}{S + strike} = \frac{(1 - \zeta)\mathbf{v}(strike \cdot \zeta / (1 - \zeta), T - \tau)}{strike},$$

which can be found, e.g., in Chapter 2.2.5 in [35] for the deterministic Black Scholes equation. This leads to a PDE for $\bar{\mathbf{v}}$ given by:

$$\frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} = \frac{1}{2}\zeta^2(1 - \zeta)^2 \mathbf{A}\frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2} + r\zeta(1 - \zeta)\frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} - r(1 - \zeta)\bar{\mathbf{v}}(\zeta, \tau), \qquad \text{(A2)}$$
$$\zeta \in (0, 1), \tau \in [0, T],$$

with boundary and initial conditions

$$\bar{\mathbf{v}}(\zeta, 0) = \begin{pmatrix} (2\zeta - 1)^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \zeta \in (0, 1),$$

$$\bar{\mathbf{v}}(\zeta, \tau) \xrightarrow{\zeta \to 0} \mathbf{0}_{|I|}, \qquad \tau \in [0, T], \quad \text{and}$$

$$\bar{\mathbf{v}}(\zeta, \tau) \xrightarrow{\zeta \to 1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \tau \in [0, T].$$

In order to solve the system, we choose a finite difference scheme because it is easy to implement for practitioners. An equidistant grid

$$\zeta_m := \frac{m}{M_\zeta} = m\Delta\zeta, \qquad m = 0, \ldots, M_\zeta,$$
$$\tau^n := T\frac{n}{N_\tau} = n\Delta\tau, \qquad n = 0, \ldots, N_\tau,$$

with $\Delta\zeta := 1/M_\zeta, \Delta\tau := T/N_\tau$ was selected. The numbers $M_\zeta, N_\tau \in \mathbb{N}$ were chosen large enough to represent the solution in a proper way and in the right proportion to

obtain a stable scheme. The partial derivatives are approximated component wise by finite differences, as it was done for the deterministic solution in Chapter 8.1.1 in [35], with

$$\text{forward differences for}\quad \frac{\partial \bar{\mathbf{v}}}{\partial \tau}(\zeta_m, \tau^n) \approx \frac{\bar{\mathbf{v}}(\zeta_m, \tau^{n+1}) - \bar{\mathbf{v}}(\zeta_m, \tau^n)}{\Delta \tau}\quad \text{and}$$

$$\text{central differences for}\quad \frac{\partial \bar{\mathbf{v}}}{\partial \zeta}(\zeta_m, \tau^n) \approx \frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{2\Delta \zeta}$$

$$\text{and for}\quad \frac{\partial^2 \bar{\mathbf{v}}}{\partial \zeta^2}(\zeta_m, \tau^n) \approx \frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - 2\bar{\mathbf{v}}(\zeta_m, \tau^n) + \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{(\Delta \zeta)^2},$$

for $m = 1, \ldots, M_\zeta - 1, n = 0, \ldots, N_\tau - 1$. This yields the explicit finite difference scheme

$$
\begin{aligned}
\bar{\mathbf{v}}(\zeta_m, \tau^{n+1}) = \quad \Delta\tau\Bigg(\quad & \frac{1}{2}\zeta_m^2(1-\zeta_m)^2 \mathbf{A}\frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - 2\bar{\mathbf{v}}(\zeta_m, \tau^n) + \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{(\Delta\zeta)^2} \\
& + r\zeta_m(1-\zeta_m)\frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{2\Delta\zeta} - r(1-\zeta_m)\bar{\mathbf{v}}(\zeta_m, \tau^n) \Bigg) \\
& + \quad \bar{\mathbf{v}}(\zeta_m, \tau^n),
\end{aligned}
\tag{A3}
$$

for $m = 1, \ldots, M_\zeta - 1, n = 0, \ldots, N_\tau - 1$ with initial value

$$\bar{\mathbf{v}}(\zeta_m, 0) = \begin{pmatrix} (2\zeta_m - 1)^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad m = 1, \ldots, M_\zeta - 1.$$

The remaining values for $m \in \{0, M_\zeta\}$, i.e., $\zeta_m \in \{0, 1\}$, are given by the boundary conditions $\bar{\mathbf{v}}(0, \tau^n) = \mathbf{0}_{N+1}$ and $\bar{\mathbf{v}}(1, \tau^n) = (1, 0, \ldots, 0)^T$ for all n.

Consistency of the scheme can easily be verified. By the Lax–Richtmyer Equivalence theorem, see for instance [36] Theorem 1.5.1, convergence of the numerical solution is given, if M_ζ and N_τ are chosen to obtain a stable scheme (A3) and if the system of Equation (A2) is well-posed. Well-posedness is in particular given for a parabolic system, i.e., when all real parts of the eigenvalues of $\mathbf{A}$ are positive.

The Galerkin multiplication tensor and thus the entries of the coupling matrix $\mathbf{A}$ can be computed by a suitable quadrature method. Gaussian quadrature was used to obtain the numerical results in Section 5.

References

1. Whaley, R. *Derivatives: Markets, Valuation, and Risk Management*; Wiley Finance; Wiley: Hoboken, NJ, USA, 2007.
2. Crawford, G.; Sen, B. *Derivatives for Decision Makers: Strategic Management Issues*; Wiley Series in Financial Engineering; Wiley: New York, NY, USA, 1996.
3. Black, F.; Scholes, M. The Pricing of Options and Corporate Liabilities. *J. Political Econ.* **1973**, *81*, 638–654. [CrossRef]
4. Merton, R.C. The Theory of Rational Option Pricing. *Bell J. Econ. Manag. Sci.* **1973**, *4*, 141–183. [CrossRef]
5. Cox, J.C.; Ingersoll, J.E., Jr.; Ross, S.A. A theory of the term structure of interest rates. *Econometrica* **1985**, *53*, 385–407. [CrossRef]
6. Rubinstein, M. Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978. *J. Financ.* **1985**, *40*, 455–480. [CrossRef]
7. Scott, L.O. Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. *J. Financ. Quant. Anal.* **1987**, *22*, 419–438. [CrossRef]
8. Günther, M.; Jüngel, A. Chapter 4 Die Black-Scholes-Gleichung and 8 Einige weiterführende Themen. In *Finanzderivate mit MATLAB*, 2nd ed.; Vieweg + Teubner: Wiesbaden, Germany, 2010.
9. Dupire, B. Pricing with a smile. *Risk* **1994**, *7*, 18–20.
10. Coleman, T.; Li, Y.; Verma, A. Reconstructing the unknown local volatility function. *J. Comput. Financ.* **1999**, *2*, 77–102. [CrossRef]
11. Crepey, S. Calibration of the local volatility in a trinomial tree using Tikhonov regularization. *Inverse Probl.* **2002**, *19*, 91–127. [CrossRef]
12. Hanke, M.; Rösler, E. Computation of Local Volatilities from Regularized Dupire Equations. *Int. J. Theor. Appl. Financ.* **2005**, *8*, 207–221. [CrossRef]

13. Heston, S.L. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Rev. Financ. Stud.* **1993**, *6*, 327–343. [CrossRef]

14. Hull, J.C.; White, A. The Price of Options on Assets with Stochastic Volatilities. *J. Financ.* **1987**, *42*, 281–300. [CrossRef]

15. Mishura, Y.; Ralchenko, K. *Discrete-Time Approximations and Limit Theorems: In Applications to Financial Markets*; De Gruyter: Berlin, Germany, 2021. [CrossRef]

16. Namihira, M.; Kopriva, D.A. Computation of the effects of uncertainty in volatility on option pricing and hedging. *Int. J. Comput. Math.* **2012**, *89*, 1281–1302. [CrossRef]

17. Pulch, R.; van Emmerich, C. Polynomial chaos for simulating random volatilities. *Math. Comput. Simul.* **2009**, *80*, 245–255. [CrossRef]

18. Drakos, S. Uncertain Volatility Derivative Model Based on the Polynomial Chaos. *J. Math. Financ.* **2016**, *6*, 55–63. [CrossRef]

19. Zhang, Y.; Ding, S.; Scheffel, E. Policy impact on volatility dynamics in commodity futures markets: Evidence from China. *J. Futures Mark.* **2018**, *38*, 1227–1245. [CrossRef]

20. Bazzana, F.; Collini, A. How does HFT activity impact market volatility and the bid-ask spread after an exogenous shock? An empirical analysis on S&P 500 ETF. *N. Am. J. Econ. Financ.* **2020**, *54*, 101240. [CrossRef]

21. Xie, D.; Cui, Y.; Liu, Y. How does investor sentiment impact stock volatility? New evidence from Shanghai A-shares market. *China Financ. Rev. Int.* **2021**. [CrossRef]

22. Miloș, M.C. Impact of MiFID II on the Market Volatility—Analysis on Some Developed and Emerging European Stock Markets. *Laws* **2021**, *10*, 55. [CrossRef]

23. Sullivan, T.J. Introduction to Uncertainty Quantification. In *Texts in Applied Mathematics*; Springer: Cham, Switzerland 2015; Volume 63.

24. Rahman, S. A polynomial chaos expansion in dependent random variables. *J. Math. Anal. Appl.* **2018**, *464*, 749–775. [CrossRef]

25. Janson, S. *Gaussian Hilbert Spaces*; Cambridge University Press: Cambridge, UK, 1997.

26. De, S.; Maute, K.; Doostan, A. Bi-fidelity stochastic gradient descent for structural optimization under uncertainty. *Comput. Mech.* **2020**, *66*, 745–771. [CrossRef]

27. Fairbanks, H.R.; Jofre, L.; Geraci, G.; Iaccarino, G.; Doostan, A. Bi-fidelity approximation for uncertainty quantification and sensitivity analysis of irradiated particle-laden turbulence. *J. Comput. Phys.* **2020**, *402*, 108996. [CrossRef]

28. Zhu, X.; Narayan, A.; Xiu, D. Computational Aspects of Stochastic Collocation with Multifidelity Models. *SIAM/ASA J. Uncertain. Quantif.* **2014**, *2*, 444–463. [CrossRef]

29. Narayan, A.; Gittelson, C.; Xiu, D. A Stochastic Collocation Algorithm with Multifidelity Models. *SIAM J. Sci. Comput.* **2014**, *36*, 495–521. [CrossRef]

30. Liu, L.; Zhu, X. A bi-fidelity method for the multiscale Boltzmann equation with random parameters. *J. Comput. Phys.* **2020**, *402*, 108914. [CrossRef]

31. Gamba, I.M.; Jin, S.; Liu, L. Error estimates of a Bifidelity method for kinetic equations with random parameters and multiple scales. *Int. J. Uncertain. Quantif.* **2021**, *11*, 57–75. [CrossRef]

32. Gao, H.; Zhu, X.; Wang, J.X. A bi-fidelity surrogate modeling approach for uncertainty propagation in three-dimensional hemodynamic simulations. *Comput. Methods Appl. Mech. Eng.* **2020**, *366*, 113047. [CrossRef]

33. Liu, L.; Pareschi, L.; Zhu, X. A bi-fidelity stochastic collocation method for transport equations with diffusive scaling and multi-dimensional random inputs. *arXiv* **2021**, arXiv:2107.09250.

34. Xiu, D. *Numerical Methods for Stochastic Computations*; Princeton University Press: Princeton, NJ, USA, 2010.

35. Zhu, Y.; Wu, X.; Chern, I.L.; Sun, Z. *Derivative Securities and Difference Methods*, 2nd ed.; Springer: New York, NY, USA, 2013.

36. Strikwerda, J.C. *Finite Difference Schemes and Partial Differential Equations*, 2nd ed.; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2004.