

MASTERTHESIS

for the acquisition of the academic degree Master of Science
in Mathematics

**Computing the Black Scholes
equation with uncertain volatility
using the stochastic Galerkin method and a
Bi-Fidelity approach**



Julius-Maximilians-Universität Würzburg
Faculty of Mathematics and Computer Science
Chair of Mathematics VI, Mathematical Fluid Dynamics

submitted by

Kathrin Hellmuth

Theodor-Heuss-Ring 11, 97215 Uffenheim

supervisor

Professor Dr. Christian Klingenberg

Würzburg, 30th April 2020

Acknowledgements

I give thanks to the Max Weber-Programm Bayern, that supported me since the beginning of my Bachelor studies in a conceptual and in the first two semesters also in a financial manner.

My thanks also goes to the Hanns-Seidel-Stiftung. After acceptance to the scholarship program in my second semester of Bachelor studies, I received conceptual and from the third semester on also financial support from resources of the Federal Ministry of Education and Research.

Furthermore, I acknowledge my gratitude to Professor Dr. Christian Klingenberg for the selection of my topic I enjoyed to work on and for his dedicated supervision. There was a great cooperative atmosphere in his work group and I learned a lot from the seminar talks. I also thank him for his advices regarding my professional future.

Professor Klingenberg also enabled the contact to Dr. Liu Liu from the University of Texas at Austin via Email and Skype. She gave me ideas and recommendations of literature on the convergence of the truncated Stochastic Galerkin solution and on the application of the Bi-Fidelity technique and sent me pieces of code for the implementation of Gaussian quadrature rules. I am very grateful for her help.

I would also like to thank my former teacher, Mr. Armin Forster, who advised me to study maths from the 8th grade on. I am still happy with my choice.

Another big thank you goes to my parents Renate and Kilian, who always supported me and enabled my studies.

I also owe my thanks to my brother Johannes who helped me installing Matlab on my laptop and lent me his mouse, when mine was broken.

Besides that I give thanks to my friends, who helped me balancing university and social life.

A special thanks goes to Anna, who proofread this thesis.

Moreover, a big thank you goes to my boyfriend Alex. I draw a lot of strength from our relation.

Finally, I thank my dog Kessy who forced me to go for a walk with her every day and thus gave me a healthy daily routine for my afternoon break.

Contents

Acknowledgements	i
Table of Contents	ii
List of Figures	iv
List of Algorithms	xi
1 Introduction	1
2 The Black Scholes Equation	5
2.1 The deterministic case	5
2.2 Techniques for deriving the volatility	7
2.3 Volatility as a random variable	8
3 Towards the Stochastic Galerkin method	9
3.1 Orthogonal polynomials	10
3.2 Generalized Polynomial Chaos expansion	14
4 Application of the Stochastic Galerkin method	20
5 Numerical implementation	26
5.1 Computation of the coupling matrix	28
5.2 Deriving a finite difference scheme	33
6 Numerical results	41
7 Extension of the model to volatility depending on finitely many independent random variables	58
7.1 The Black Scholes equation with uncertain volatility depending on multiple random variables	58
7.2 Generalized Polynomial Chaos expansions in finitely many independent random variables	59
7.3 Application of the Stochastic Galerkin method in multiple random variables	61
7.4 Numerics	64
7.5 Numerical results for the extended model	71

8	Integration of a Bi-Fidelity approach	85
8.1	General Procedure	86
8.2	Application to the Black Scholes equation with uncertain volatility and numerical implementation	91
8.3	Comparing Bi-Fidelity solution and high fidelity solution	104
9	Summary and Conclusion	127
	List of Abbreviations	129
	Bibliography	130

List of Figures

- 3.1 Graphs of the first Hermite polynomials He_0 , He_1 , He_2 , He_3 , He_3 and He_3 together with the density μ 11
- 3.2 Graphs of the first Legendre polynomials $Le_0 = \mu$, Le_1 , Le_2 , Le_3 , Le_4 and Le_5 12
- 6.1 Expected value surface for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 42
- 6.2 Expected value at some times t for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ 42
- 6.3 Solution of the deterministic Black Scholes equation for a European Call option with $\sigma = 0.3$, $T = 20$, $strike = 100$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 43
- 6.4 Absolute difference of the expected values of the SG solution and the deterministic solution for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed and $\sigma = 0.3$ respectively, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area of the stochastic solution circled in red. . . . 44
- 6.5 Variance surface for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area circled in red. 45
- 6.6 Variance at some times t for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ 45
- 6.7 Expected value surface for a European Call option with $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 46
- 6.8 Expected values at $t = 0$ for European Call option in the models $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, and $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ 47

6.9	Absolute difference in expected values of the models $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, and $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum absolute value and the smoothing area for normal $\Sigma(\Theta)$ circled in red.	47
6.10	Variance surface for a European Call option with $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area circled in red.	48
6.11	Variances at $t = 0$ for European Call option in the models $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, and $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$	49
6.12	Absolute difference of variances of the models for $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, and $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-0.5, 0.5]$ distributed, $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area for normal $\Sigma(\Theta)$ circled in red.	49
6.13	Expected values at $t = 0$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$	50
6.14	Absolute difference in the expected values for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_2 is circled in red.	51
6.15	Expected values at $S = \textit{strike} = 100$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$	51
6.16	Absolute difference in expected values for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_3 is circled in red.	52
6.17	Variances at $t = 0$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$	53
6.18	Absolute difference in variances for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$, Θ normal distributed, with $T = 20$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_2 is circled in red.	53
6.19	Implied volatilities of the European Call option data with fitted normal distribution.	54
6.20	Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation.	55

6.21	Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation for the last 55 days of the option.	56
6.22	Market values of the option together with the deterministic solution and the expected value of the SG solution.	57
7.1	Expected value surface for a European Call option with $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of its maximum value and its smoothing area circled in red.	72
7.2	Variance surface for a European Call option with $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of its maximum value and smoothing area circled in red.	73
7.3	Expected values in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 , $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$	74
7.4	Expected values in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 zoomed in, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$	74
7.5	Variances in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 , $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$	75
7.6	Absolute difference in the expected values for a European Call option for the models Σ_0 and Σ_1 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_0 circled in red.	76
7.7	Absolute difference in the expected values for a European Call option for the models Σ_1 and Σ_3 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_1 circled in red.	77
7.8	Absolute difference in variance for a European Call option for the models Σ_0 and Σ_1 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_0 circled in red.	78
7.9	Absolute difference in variance for a European Call option for the models Σ_1 and Σ_3 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_1 circled in red.	79
7.10	Histogram density estimator of $\Sigma(\Theta, \Delta)$ derived by the implied volatilities together with the density of $\Sigma(\Theta, \Delta)$ fitted to the implied volatilities by maximum likelihood.	81
7.11	Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation.	82
7.12	Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation for the last 55 days.	83

7.13	Market values of the option together with the expected value of the SG solution and deterministic solution.	84
8.1	Expected value surface of the high fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.	105
8.2	Expected value of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.	106
8.3	Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red.	107
8.4	Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for a wider range of S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red.	108
8.5	Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red.	109
8.6	Variance of the high fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.	110
8.7	Variance of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, <i>strike</i> = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.	111

8.8 Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and the high fidelity smoothing area circled in red. 112

8.9 Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red. 113

8.10 Expected value of the high fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 114

8.11 Expected value of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 115

8.12 Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red. 116

8.13 Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red. 117

8.14 Variance of the high fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 118

8.15 Variance of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red. 119

8.16 Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and the high fidelity smoothing area circled in red. 120

8.17 Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red. 121

8.18 Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value. 122

8.19 Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for a wider range of S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value. 123

8.20 Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$. 124

8.21 Mean absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value. 125

8.22 Mean absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$ 126

List of Algorithms

1	Deriving the Gauss Quadrature nodes	31
2	Calculate the Galerkin Multiplication Tensor \mathbf{M}	32
3	Compute the coupling matrix A	33
4	Test for Stability	38
5	Computing the truncated gPC coefficients \mathbf{v} of V	39
6	Compute the cardinality $ I $	65
7	Compute the Galerkin multiplication tensor matrix \mathbf{M} for L random variables	67
8	Calculation of the coupling matrix A for L random variables	68
9	Alternative calculation of the coupling matrix A for $L = 2$ random variables	69
10	Computing the truncated gPC coefficients \mathbf{v} of V for L random variables .	70
11	Bi-Fidelity Offline step 1: Compute the low fidelity solutions in many Ξ values	94
12	Bi-Fidelity Offline Step 2: Identify the most important Ξ values $z_1, \dots, z_A \in Y$	96
13	Bi-Fidelity Offline step 3: Compute the high fidelity solution in the important points	100
14	Bi-Fidelity Online step 2: Calculate of the projection coefficients c_k	102
15	Bi-Fidelity Online step 3: Construct the Bi-Fidelity solution	103

Chapter 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 or 'derived' by other, so called underlying variables (see [Hul12] Chapter 1), which explains the name 'derivative'. Underlying variables can have many different forms, for example one can find¹ derivatives on commodity prices like the oil price or the price of coffee or sugar, on exchange rates for currencies, on interest rates or even on the weather. But most often the underlying variables are given as shares of companies or stock market indices. They are also called underlying assets if the variable is an asset occurring in the market .

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. However, they also open up opportunities for speculation. The following historical examples show derivatives used for both purposes.

The first record of a derivative-like trade goes back to a Babylonian law from late 18th century BC protecting debtors in case of crop failure by remitting their interest payments for that year. The annual interest payment of the debtor can be seen as a derivative paying the interest in case of a normal harvest and nothing in case of crop failure, as explained in [Wha07] Chapter 1.

Another reference on derivative trading can be found in Aristotle's Politics. He tells the story of the philosopher Thales who predicted a good olive harvest long before by his astronomical knowledge. He therefore paid a deposit for the right to rent all olive oil presses for the next harvesting season so he can sublet them to the olive farmers. A good harvest would lead to a high demand for oil presses such that he could charge a high fee and make a lot of money. However, if the harvest was bad and the demand for oil presses was low, he would be forced to rent the presses for a small fee. His return on investing the deposit hence depended on the harvest and therefore represented an agricultural derivative used for speculative purposes. A more detailed description can be found in [CS96] Chapter I. Fortunately for him, Thales' plan worked out and it was a good olive harvest.

¹Derivatives and corresponding underlying variables can be found on <https://www.finanzen.net/> for instance.

An example explaining the minimization of risk by derivatives was given by Dutch tulip merchants and growers in the 17th century. Tulips were a popular trading asset and luxury item back then leading to a rapid increase in prices. Merchants of the bulbs therefore managed their risk in rising purchase prices by buying Call options on the tulips, i.e. rights to buy the tulip bulbs in the future for a predefined price. The growers conversely bought Put options, rights to sell the bulbs at a predefined price in the future, to manage the risk of falling prices, see e.g. [Wha07] Chapter 1.

By means of such options as well as forward contracts, i.e. contracts for delivering a certain good in the future for a predefined price, farmers and merchants of different agricultural assets as grain, cotton, butter, eggs, coffee etc. managed their risks at the Chicago Board of Trade in 19th century America for instance, compare [Wha07] Table 1.1. In the middle of the 20th century, the first derivatives on metals like silver and lead were introduced and not long after that, in the 1970s, also foreign currencies and stocks served as underlying assets, see [Wha07] Table 1.1.

With the rapidly evolving market also the need for a pricing formula for derivatives grew in the 20th century, since incorrectly priced derivatives are themselves potential risks. A breakthrough was made by Black, Scholes [BS73] and Merton [Mer73] in 1973 when they contemporaneously formulated a model allowing the evaluation of derivatives, for which they were later awarded the Nobel prize in economics. Derived from this model, the Black Scholes equation

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

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, whose price is denoted by S and follows a geometric Brownian motion (e.g. a stock, an index or some commodity price). 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. For European Call and Put options, they could even derive a closed form solution to this PDE and thus an explicit pricing formula. Later, this model was extended to multiple underlying assets and adjusted for certain kinds of underlying variables like interest rates, see e.g. [CIR85].

With this pricing model at hand, the derivative market increased up to its nowadays size with 66 million transactions and a corresponding gross volume of €735 trillion just in the European Union (EU) in 2018, according to EU supervising agencies [SA19].

However, the comparison to real data soon showed that the volatility σ of one and the same stochastic asset can take values that differ more than one can explain by rounding errors etc., see e.g. [Rub85], [Sco87] and [GJ10]. The most popular approaches to deal with this are to model the volatility either as local volatility, i.e. a function of the asset price S and the time t , (see [Dup94], [CLV99], [Cre02] and [HR05]) or as a stochastic process, compare e.g. the famous Heston model [Hes93] or [Rub85], [Sco87] and [HW87].

Another approach is used in [NK12], [PvE09] and [Dra16]: The volatility is modelled as a random variable $\Sigma(\omega) = \Theta(\omega)$ (in [NK12]) or a function of a random variable $\Sigma(\Theta(\omega))$ (in [PvE09] and [Dra16]) for ω from a probability space. The resulting stochastic version of the Black Scholes equation is then studied by means of uncertainty quantification. Since a similar model is used in this thesis, the methods of these three papers are briefly explained below.

In all these papers, the solution V , now also depending on Θ , is developed in a generalized Polynomial Chaos (gPC) expansion

$$V(S, t, \Theta(\omega)) = \sum_{n=0}^{\infty} v_n(S, t) p_n(\Theta(\omega)) \quad (1.1)$$

for orthogonal 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 $\mu : \mathcal{D} \rightarrow \mathbb{R}$, one can alternatively calculate the coefficients by

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

For this purpose, a quadrature rule with nodes x_j and weights w_j is used in [NK12] for the above integral $v_n(S, t) \approx \sum_{j=1}^J V(S, t, x_j) p_n(x_j) \mu(x_j) w_j$. The solutions $V(S, t, x_j)$ needed in the quadrature formula are the solutions of the *deterministic* Black Scholes equation with $\sigma = x_j$ (recall that the stochastic volatility model $\Sigma = \Theta$ is used). With the thus derived coefficients $v_n(S, t)$, the approximate solution is calculated by $V(S, t, \Theta(\omega)) \approx \sum_{n=0}^N v_n(S, t) p_n(\Theta(\omega))$ for some $N \in \mathbb{N}$. This method is a so called Stochastic Collocation method, since the equation is exactly satisfied by the solution in some points, the quadrature nodes x_i , and the overall stochastic solution is constructed from these exact solutions, see e.g. [Sul15] Section 13.2.

In [PvE09] and [Dra16] however, the Stochastic Galerkin method is used for computing the coefficients $v_n(S, t)$. By inserting the gPC expansion 1.1 into the stochastic Black Scholes equation, 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

$$0 = \frac{\partial v_k(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{n=0}^{\infty} \frac{\partial^2 v_n(S, t)}{\partial S^2} E((\Sigma(\Theta))^2 p_k(\Theta) p_n(\Theta)) + rS \frac{\partial v_k(S, t)}{\partial S} - r v_k(S, t). \quad (1.2)$$

After the system and the coupling term were truncated to attain a maximum index N , they are solved numerically by the method of lines in [PvE09] and the finite element method in [Dra16]. The expected value and the variance of the solution $V(S, t, \Theta)$ are then calculated.

Now this thesis provides a more detailed explanation on the theoretical background of

the Stochastic Galerkin method and its application to the Black Scholes equation with random volatility. To obtain numerical results, an explicit finite difference scheme will be used. The model will be extended to volatility depending on finitely many independent random variables and a possibility to reduce the computational cost when calculating the gPC coefficients of the solution V will be shown.

At the beginning, the original deterministic Black Scholes equation and the reason for modelling the volatility not as a constant but a random variable are given as well as the stochastic Black Scholes equation used in the following chapters.

Then, the theoretical background of the gPC expansion is described in detail including facts on orthogonal polynomials that will be useful for computations and conditions on when they span the space of twice integrable functions w.r.t. the density of the random variable.

In chapter 4, the Stochastic Galerkin method is applied to stochastic Black Scholes equation explaining also the general procedure.

The numerics for solving the system of PDEs for the gPC coefficients $v_n(S, t)$ of the stochastic Black Scholes solution is shown in the next chapter followed by numerical results obtained by the programs from the numerics chapter.

Subsequently, the model is extended to volatility depending on finitely many random variables in chapter 7. Again, the theoretical background of the gPC expansion in this case is explained, the Stochastic Galerkin method is applied and the numerics for solving the system of PDEs for the gPC coefficients of the solution is illustrated, before numerical results are presented.

Finally, we introduce a Bi-Fidelity technique in order to save computational cost, since the computational cost rises fast if a finer grid is used or more gPC coefficients should be computed, especially when volatility depends one more than one random variable. This technique is applied to the model derived in the previous chapter and numerical experiments are done.

Chapter 2

The Black Scholes Equation

In financial mathematics, the Black Scholes equation is a very famous partial differential equation (PDE) that describes the development of the price of a financial derivative underlying just one stochastic asset. Together with the corresponding Black Merton Scholes model, it marked a breakthrough in the theory of option pricing and is - in modified versions - still used in practice, see for example [Eli19] footnote 22. The Black Scholes equation is introduced in this chapter and common techniques for deriving the value of volatility are explained. Since these techniques reveal significant deviations in the derived volatility values, models extending the Black Scholes equation to non constant volatility are mentioned and the model used in this work is derived.

2.1 The deterministic case

The Black Scholes equation has its origin in the Black Merton Scholes model first mentioned by Black and Scholes [BS73] and Merton [Mer73] in the year 1973 who were awarded the Nobel prize in economics for their work in 1997 [Nob97].

The model describes a market with two assets, the deterministic asset B and the stochastic asset S . The market is assumed to be

- free of arbitrage, i.e. money cannot come from nowhere, which is a necessary assumption for reasonable pricing,
- frictionless, meaning there are no transaction costs, taxes, fees etc., and
- liquid, i.e. at any time there exists someone you can buy an asset from or sell it to.

It is assumed to allow

- buying and selling of assets at any time and at any fraction.
- short selling, which means selling an asset without having it at hand. Usually the price of the asset is then paid instead of transferring the asset.
- borrowing money from the deterministic asset at the same rate of interest that one would get for investing money in the deterministic asset.

The evolution of the stochastic asset S , which is supposed not to pay any dividends, and the deterministic asset B are specified by means of stochastic differential equations

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt, \end{aligned} \tag{2.1}$$

where the model parameter μ is the *drift rate* and σ the *volatility* of S and r is called the *risk free rate of interest* in the market. The process W_t is a standard Brownian motion. For valuation purposes, the drift rate is insignificant and can therefore be neglected in the following. The risk free rate of interest can easily be obtained by bank data, whereas the volatility has to be reconstructed from historical asset or option prices.

By using Itô-calculus one can now derive a PDE that describes the development of a general financial derivative's price $V(S, t)$ depending on the time t and the price of the underlying asset S at that moment, see e.g. [Mer73], [Hul12] Chapter 14 or [GJ10]. The so called *Black-Scholes equation* states, that

$$\frac{\partial V(S, t)}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V(S, t)}{\partial S^2} + r S \frac{\partial V(S, t)}{\partial S} - r V(S, t) = 0, \quad S \in (0, \infty), t \in [0, T]. \tag{2.2}$$

Solving this PDE according to the derivative's condition at maturity T and its boundary conditions in S , one obtains the price process of the derivative. Since the Black-Scholes equation is a parabolic PDE, the existence and uniqueness of weak solutions in $L^2([0, T] \times (0, \infty), \mathbb{R})$ is guaranteed, see for instance [Eva98]. However, in general it can only be obtained numerically.

Example 1 ([GJ10] Chapter 4.2.2. and Satz 4.12). A European Call option grants its holder the right to buy an asset with price S at the time T in future for a predefined strike price *strike*.

At time T , the output of the option will therefore be $(S - \textit{strike})^+ := \max(S - \textit{strike}, 0)$. This should coincide with the fair price $V(S, T)$ at time T to avoid arbitrage, since a higher (lower) price \hat{V} would guarantee the vendor (buyer) an instant positive output $\hat{V} - (S - \textit{strike})^+ > 0$ ($(S - \textit{strike})^+ - \hat{V} > 0$) at no cost. This is the final condition. For the boundary conditions in S , consider at first $S_t = 0$ at one time $t \in [0, T]$. Then the asset price dynamic shows $S_u = 0$ for all $u \geq t$, so that the final output is 0. But in an arbitrage free market, a zero output is worth nothing, hence $V(0, t) = 0$ for all $t \in [0, T]$. For very high S however, it is very likely that the Call will be exercised at T with output $S_T - \textit{strike} > 0$. This output is approximately worth $S_t - \textit{strike} e^{r(T-t)}$ at time t and therefore $V(S, t) \rightarrow \infty$ for $S \rightarrow \infty$.

Thus, if one postulates the Black-Scholes model for the considered market and the asset price S , the option price can be calculated by solving the Black-Scholes equation 2.2 with the final and boundary conditions

$$\begin{aligned} V(S, T) &= (S - \textit{strike})^+, \quad S \in (0, \infty), \\ V(S, t) &\xrightarrow{S \rightarrow 0} 0, \quad t \in [0, T] \quad \text{and} \\ \frac{V(S, t)}{S} &\xrightarrow{S \rightarrow \infty} 1, \quad t \in [0, T]. \end{aligned}$$

For this special case, the PDE can be solved analytically to obtain the Black-Scholes formula for European Call options

$$V(S, t) = S\Phi(d_1) - \text{strike} e^{-r(T-t)}\Phi(d_2), \quad S > 0, \quad 0 \leq t < T,$$

with the distribution function Φ of the standard normal distribution

$$\Phi(x) = \frac{1}{2\pi} \int_{-\infty}^x e^{-s^2/2} ds, \quad x \in \mathbb{R},$$

and

$$d_{1/2} = \frac{\ln(S/\text{strike}) + (r \pm \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

2.2 Techniques for deriving the volatility

Within the years, different techniques were developed to derive a value of the volatility from historical market data as described in [GJ10]. The two most popular will be portrayed in this section.

The *historic volatility* makes use of the fact, that σ^2 is just the variance of the logarithmic yearly difference of the asset prices:

$$d \ln(S_t) = \frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

Because the Brownian motion has independent increments $W_{t_1} - W_{t_2}$ with variance $t_1 - t_2$ for $t_1 > t_2$ and the drift term grows deterministically, the variance of

$$d \ln(S_{t_1}) - d \ln(S_{t_2}) = \mu(t_1 - t_2) + \sigma(W_{t_1} - W_{t_2})$$

is given by $\sigma^2(t_1 - t_2)$. Hence, it makes sense to determine the volatility from n sequentially collected daily asset prices S_i corresponding to time t_i for $i = 1, \dots, n$. The squared historic volatility is computed as the sample variance of daily logarithmic asset price differences $y_i := \ln(S_{i+1}) - \ln(S_i)$ multiplied by the mean number of trading days in one year td , i.e.

$$\sigma_{hist} = \sqrt{td} \left(\frac{1}{n-2} \sum_{i=0}^{n-1} (y_i - \bar{y})^2 \right),$$

where \bar{y} is the arithmetic mean of the y_i . This is a consistent and convergent estimate for σ due to the properties of the sample variance.

Another approach makes use of option pricing, as the Black-Scholes equation can be solved analytically in the case of European options, see example 1. They resemble optional future contracts, i.e. the right to buy (Call option) or sell (Put option) an asset at a predefined strike price *strike* at the maturity T . As the price of these options depends

monotonously on σ before maturity, it is possible to fit their calculated prices to the market prices by varying σ . This gives a well-defined value (see [GJ10] section 4.4.2) which is called the *implied volatility* σ_{impl} . If one refers to the market price of a European option at a specific time \tilde{t} and value of the stochastic asset \tilde{S} as \tilde{V} , then σ_{impl} satisfies

$$V_{\sigma_{impl}}(\tilde{S}, \tilde{t}) = \tilde{V}. \quad (2.3)$$

Here, the notation V_x should illustrate, that the value x was used for σ in the formula for the option price V .

Either way, the computation of volatility depends on random observations of S or on the values of European options, which could be influenced by rounding errors or supply and demand, leading to slightly different prices. Indeed, as shown in [Rub85], [Sco87] and [GJ10] Tabelle 4.1 e.g., the implied volatility takes significantly different values.

Trying to explain this, [Dup94], [CLV99], [Cre02] and [HR05] for instance modelled the volatility as a function of S and t . The so called *local volatility* hence exchanges every σ in the Black-Scholes model by $\sigma(S, t)$, leading to some extra terms in the Black-Scholes equation in order to deal with the volatility risk.

Another common way to cope with the variability in σ is to model it as a stochastic process just as S . This approach was e.g. used in the Heston model [Hes93] as well as in [Rub85], [Sco87] and [HW87].

2.3 Volatility as a random variable

In this work, however, the volatility will be described in a different way: It will be assumed to be a random variable Σ that is independent of the asset price S and does not change in time. This could e.g. reflect the uncertainty in deriving the right value for σ . If one uses the implied volatility to obtain values for σ , as we will do in this work, the independence is a quite rational assumption, because the asset price S is already considered during the computation of option prices. Therefore, one could assume the asset price to be already 'priced in' and as a result, the implied volatility could be assumed to be independent on it.

In this model, deriving the Black-Scholes equation works the same way as it did in the case of a constant volatility. The Black-Scholes equation with random volatility is then given by

$$\frac{\partial V(S, t)}{\partial t} + \frac{1}{2} \Sigma^2 S^2 \frac{\partial^2 V(S, t)}{\partial S^2} + rS \frac{\partial V(S, t)}{\partial S} - rV(S, t) = 0 \quad (2.4)$$

for $S \in (0, \infty)$, $t \in [0, T]$, where Σ is a random variable reflecting the random volatility. It should be pointed out that the solution of this equation $V(S, t)$ is a random variable, which takes values according to the value of Σ . If Ω is the sample space of Σ , i.e. $\Sigma : \Omega \rightarrow \mathbb{R}$, then $V(S, t)$ is for all $(S, t) \in (0, \infty) \times [0, T]$ a function $\Omega \rightarrow \mathbb{R}$, too. Furthermore, a realization of the solution $V(S, t)(\omega)$, $\omega \in \Omega$, equals the solution to the deterministic Black-Scholes equation with $\sigma = \Sigma(\omega)$, i.e. in the notation as in equation 2.3 one has

$$V(S, t)(\omega) = V_{\Sigma(\omega)}(S, t), \quad \omega \in \Omega. \quad (2.5)$$

Chapter 3

Towards the Stochastic Galerkin method

This chapter explains the ideas behind the Stochastic Galerkin method, that is demonstrated by applying it to the Black Scholes equation with random volatility 2.4 in the next chapter.

Since the Stochastic Galerkin method is based on the general Polynomial Chaos (gPC) expansion of random variables, that makes use of orthogonal polynomials, these topics have to be discussed first.

The aim of this chapter is therefore to derive another representation of random variables $X : \Omega \rightarrow \mathcal{F}$, where Ω is a probability space and \mathcal{F} is a Hilbert space, in our case \mathbb{R} or a space $L^p(D, \mathbb{R})$ of Lebesgue L^p functions from some domain D to \mathbb{R} for $p = 0, 1, 2$. This representation should only depend on another continuous random variable $\Theta : \Omega \rightarrow \mathbb{R}$ with density $\mu : \mathcal{D} \rightarrow \mathbb{R}$ for $\mathcal{D} \subset \mathbb{R}$ and be of the type $\sum_{k=0}^{\infty} x_k p_k(\Theta)$ for $x_k \in \mathcal{F}$ and p_k orthogonal polynomials w.r.t. μ . It should equal X in distribution, since this work always considers equality of random variables in distribution.

To make sure X can be represented by Θ in a proper way, we assume $X = \tilde{X}(\Theta)$ in distribution with $\tilde{X} \in L^2_{\mu}(\mathcal{D}, \mathcal{F})$.

Notation:

As usual, $L^2_{\mu} := L^2_{\mu}(\mathcal{D}, \mathbb{R})$ describes the space of all functions $\mathcal{D} \rightarrow \mathbb{R}$ with finite L^2_{μ} norm $\|\cdot\|_{\mu}$ modulo μ null sets. The norm is defined by

$$\|f\|_{\mu}^2 := \int_{\mathcal{D}} (f(x))^2 \mu(x) dx \quad \text{for } f : \mathcal{D} \rightarrow \mathcal{F},$$

which is actually the natural norm defined by the scalar product

$$\langle f, g \rangle_{\mu} := \int_{\mathcal{D}} f(x)g(x)\mu(x) dx \quad \text{for } f, g \in L^2_{\mu}.$$

For sake of easier readability, the integral of $f \in L^2_{\mu}$ will also be written in the following way

$$\langle f \rangle_{\mu} := \int_{\mathcal{D}} f(x)\mu(x) dx.$$

Note that with this notation, the subsequent relations hold for $f, g \in L^2_\mu$:

$$\begin{aligned}\langle f, g \rangle_\mu &= \langle fg \rangle_\mu, \\ \|f\|_\mu^2 &= \langle f, f \rangle_\mu = \langle f^2 \rangle_\mu.\end{aligned}$$

For $\mathcal{F} = L^p(D, \mathbb{R})$, the space $L^2_\mu(\mathcal{D}, L^p(D, \mathbb{R}))$ and the above notation is defined accordingly, with the difference that the integrals lie in $L^p(D, \mathbb{R})$ and finite norm is meant to hold pointwise for all points in D . Note that the random variable X is then actually a stochastic process $D \times \Omega \rightarrow \mathbb{R}$. Therefore, values of X will also be written as $X(t, \omega)$ for $t \in D$ and $\omega \in \Omega$.

Furthermore, the difference between f and $f(x)$ will be dealt with in a loose fashion such that sometimes $f(x)$ will denote the function instead of its value in x . It should be clear from the context if the function or the value is meant.

3.1 Orthogonal polynomials

Orthogonal Polynomials were extensively studied in literature, e.g. by [Sze59] and [Chi78] and [Fre71], because they are of great advantage in approximation theory and numerics for PDEs, where they serve as basis functions for the Galerkin approach for instance. We will use them to approximate a random variable by another one via a series of orthogonal polynomials evaluated in the latter random variable.

Definition 1 ([Chi78]). Let $\mathcal{N} := \mathbb{N}_0$ or $\mathcal{N} := \{0, 1, \dots, N\}$, $N \in \mathbb{N}$, and μ be a probability density on $\mathcal{D} \subset \mathbb{R}$. Then, the set $\{p_n \mid n \in \mathcal{N}\}$ of polynomials in x defined on \mathcal{D} is called an *orthogonal system of polynomials* with respect to the probability density μ , if for all $n \in \mathcal{N}$, the polynomial p_n has degree $\deg(p_n) = n$ and it satisfies

$$\begin{aligned}\langle p_n p_m \rangle_\mu &= \int_{\mathcal{D}} p_n(x) p_m(x) \mu(x) dx = 0 && \text{for } m \neq n \text{ and} \\ \langle p_n^2 \rangle_\mu &= \int_{\mathcal{D}} p_n^2(x) \mu(x) dx = \|p_n\|_{L^2_\mu}^2 > 0.\end{aligned}\tag{3.1}$$

The p_n are called the *orthogonal polynomials* w.r.t. μ .

Using the Kronecker delta function $\delta_{xy} := 1_{\{x=y\}}$, $x, y \in \mathbb{R}$, condition 3.1 can be summarized: $\langle p_n p_m \rangle_\mu = \gamma_n \delta_{nm}$ with $\gamma_n := \|p_n\|_{L^2_\mu}^2 > 0$.

Orthogonal polynomials can be defined for a greater class of distributions, for instance distributions that do not necessarily have densities or are not necessarily probability distributions, see [Chi78] or [Fre71]. For our purposes however, the above definition is sufficient and more illustrative, because only probability distributions with densities will be considered in this work.

The existence of orthogonal systems w.r.t. an arbitrary probability density function μ is guaranteed by the Gram Schmidt algorithm applied to the monomials $x^n, n \in \mathcal{N}$, as long as all corresponding moments $\int_{\mathcal{D}} x^n \mu(x) dx$ are finite, see e.g. [Fre71].

It is obvious, that orthogonal systems are not unique, take for example $\{2p_n \mid n \in \mathcal{N}\}$ for a set of orthogonal polynomials $\{p_n \mid n \in \mathcal{N}\}$. However, the Gram Schmidt algorithm shows that the polynomials are unique up to multiplication by constants, see also [Chi78].

Some important orthogonal polynomials with corresponding density functions are given in the next example.

Example 2 ([Chi78], [Sul15]). We will consider orthogonal polynomials for the probability density of a standard normal and a uniformly on $[-1/2, 1/2]$ distributed random variable.

- (a) The monic orthogonal polynomials of the standard normal distribution with density function $\mu : \mathbb{R} \rightarrow \mathbb{R}$,

$$\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

are the *Hermite polynomials*

$$\text{He}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}, \quad \text{for } n \in \mathbb{N}.$$

It holds, that $\langle \text{He}_n \text{He}_m \rangle_\mu = n! \delta_{nm}$. The first six polynomials are given by

$$\begin{aligned} \text{He}_0(x) &= 1, \\ \text{He}_1(x) &= x, \\ \text{He}_2(x) &= x^2 - 1, \\ \text{He}_3(x) &= x^3 - 3x, \\ \text{He}_4(x) &= x^4 - 6x^2 + 3, \\ \text{He}_5(x) &= x^5 - 10x^3 + 15x. \end{aligned}$$

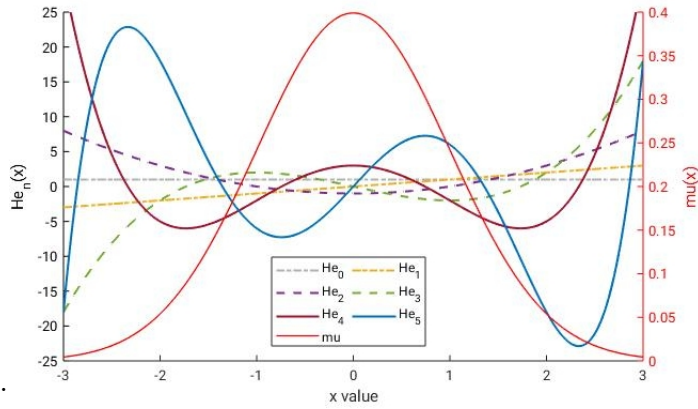


Figure 3.1: Graphs of the first Hermite polynomials He_0 , He_1 , He_2 , He_3 , He_4 and He_5 together with the density μ .

- (b) The orthogonal polynomials of the uniform distribution on $[-1/2, 1/2]$ with density function $\mu(x) = 1$ for all $x \in [-1/2, 1/2]$ are called the *Legendre polynomials* Le_n . They are in particular the solutions y to the differential equations

$$(1 - x^2) \frac{d^2}{dx^2} y + 2x \frac{d}{dx} y + n(n + 1)y = 0$$

for $n \in \mathbb{N}$ and satisfy $\langle \text{Le}_n \text{Le}_m \rangle_\mu = \frac{1}{2n+1} \delta_{nm}$. The Legendre polynomials of degree smaller than six are

$$\begin{aligned}
\mathbf{Le}_0(x) &= 1, \\
\mathbf{Le}_1(x) &= 2x, \\
\mathbf{Le}_2(x) &= 6x^2 - \frac{1}{2}, \\
\mathbf{Le}_3(x) &= 20x^3 - 3x, \\
\mathbf{Le}_4(x) &= 70x^4 - 15x^2 + \frac{3}{8}, \\
\mathbf{Le}_5(x) &= 252x^5 - 70x^3 + \frac{15}{4}x.
\end{aligned}$$

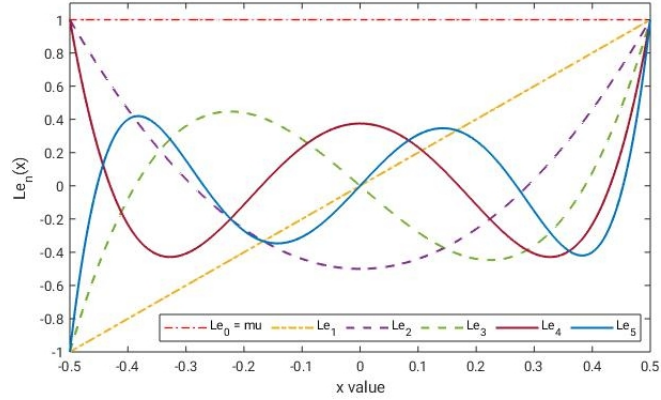


Figure 3.2: Graphs of the first Legendre polynomials $\mathbf{Le}_0 = \mu$, \mathbf{Le}_1 , \mathbf{Le}_2 , \mathbf{Le}_3 , \mathbf{Le}_4 and \mathbf{Le}_5 .

Orthogonality of the p_n in particular enables a fast computation of integrals of p_n with another polynomial.

Theorem 1 ([Chi78], Theorem I.2.1). *For a probability density function μ and a set \mathcal{N} like in definition 1 and polynomials $\{p_n \mid n \in \mathcal{N}\}$ on \mathcal{D} with $\deg(p_n) = n$, it is equivalent that is*

- (i) $\{p_n \mid n \in \mathcal{N}\}$ is an orthogonal system of polynomials w.r.t. μ ,
- (ii) $\langle p_n q \rangle_\mu = \int_{\mathcal{D}} p_n(x)q(x)\mu(x) dx = 0$ for every polynomial q on \mathcal{D} of degree smaller than n and unequal to zero for $n = \deg(q)$.

Proof. We prove this in two steps.

(i) \Rightarrow (ii) Let $\{p_n \mid n \in \mathcal{N}\}$ be an orthogonal system of polynomials w.r.t. μ . Due to $\deg(p_k) = k$, the set $\{p_k \mid k = 0, \dots, m\}$ is a basis for the vector subspace of polynomials of degree smaller or equal to $m \in \mathcal{N}$. Then every polynomial q on \mathcal{D} of degree m can be written as a linear combination of the p_k , $k = 0, \dots, m$, i.e.

$$q(x) = \sum_{k=0}^m c_k p_k(x), \quad c_k \in \mathbb{R} \text{ for } k = 0, \dots, m \text{ and } c_m \neq 0.$$

Now, for $m < n$ linearity of the integral gives

$$\langle p_n q \rangle_\mu = \sum_{k=0}^m c_k \langle p_k p_n \rangle_\mu = 0$$

by orthogonality of the $\{p_n, n \in \mathcal{N}\}$ and for $m = n$

$$\langle p_n q \rangle_\mu = \sum_{k=0}^m c_k \langle p_k p_n \rangle_\mu = c_m \langle p_m^2 \rangle_\mu = c_m \gamma_m \neq 0.$$

(ii) \Rightarrow (i) Assuming (ii) holds and choosing $q = p_m$, one obtains that $\{p_n \mid n \in \mathcal{N}\}$ form an orthogonal system of polynomials w.r.t. μ .

□

To calculate the orthogonal polynomials w.r.t. a probability distribution μ , one could use the Gram Schmidt algorithm. However, there is an easier and computationally faster way given by the recursion formulas for orthogonal polynomials.

Theorem 2 ([Chi78], Theorem I.4.1). *Let μ be a probability density on $\mathcal{D} \subset \mathbb{R}$ and $\{p_n \mid n \in \mathcal{N}\}$ a corresponding orthogonal system of monic polynomials, that means the coefficient of the highest potency of x in $p_n(x)$ is 1. Then for $n \in \mathcal{N}$ the p_n satisfy a recursion relation*

$$p_{n+1}(x) = (x - a_n)p_n(x) - b_n p_{n-1}(x)$$

$$\text{for } a_n = \frac{\langle x p_n^2(x) \rangle_\mu}{\langle p_n^2(x) \rangle_\mu} \in \mathbb{R}, \quad b_n = \frac{\langle p_n^2(x) \rangle_\mu}{\langle p_{n-1}^2(x) \rangle_\mu} > 0.$$

Here, one defines $p_{-1}(x) := 0$.

Proof. Because $x p_n(x)$ is a monic polynomial of degree $n + 1$, it can be written as a linear combination of the $p_k, k = 0, \dots, n + 1$:

$$x p_n(x) = \sum_{k=0}^{n+1} c_{n,k} p_k(x) \quad \text{with } c_{n,k} = \frac{\langle x p_n(x) p_k(x) \rangle_\mu}{\langle p_k^2(x) \rangle_\mu}$$

and $c_{n,n+1} = 1$ due to the monic property. For $k < n - 1$, the polynomial $x p_k(x)$ is of degree $k + 1 < n$ and hence $c_{n,k} = 0$ by theorem 1 and we are left with

$$x p_n(x) = p_{n+1}(x) + c_{n,n} p_n(x) + c_{n,n-1} p_{n-1}(x)$$

for $n \geq 1$. By replacing $x p_{n-1}$ by $\sum_{k=0}^n c_{n-1,k} p_k(x)$ with $c_{n-1,n} = 1$ and using the orthogonality of the p_k , one has

$$c_{n,n-1} = \frac{\langle x p_n(x) p_{n-1}(x) \rangle_\mu}{\langle p_{n-1}^2(x) \rangle_\mu} = \frac{\sum_{k=0}^n c_{n-1,k} \langle p_n(x) p_k(x) \rangle_\mu}{\langle p_{n-1}^2(x) \rangle_\mu} = c_{n-1,n} \frac{\langle p_n^2(x) \rangle_\mu}{\langle p_{n-1}^2(x) \rangle_\mu} = b_n.$$

Since $c_{n,n} = a_n$, the result follows. □

Note, that also non monic orthogonal polynomials p_n satisfy a recursion relation of the form

$$p_{n+1}(x) = (c_n x - a_n) p_n(x) - b_n p_{n-1}(x) \tag{3.2}$$

for $a_n, b_n, c_n \in \mathbb{R}$. This can easily be shown using the formula for the monic polynomials.

Example 3 ([Chi78], [Sul15]). The Hermite and Legendre polynomials satisfy the recursion formulas

$$\begin{aligned}\text{He}_{n+1}(x) &= x\text{He}_n(x) - n\text{He}_{n-1}(x), \\ \text{Le}_{n+1}(x) &= \frac{2n+1}{n+1}2x\text{Le}_n(x) - \frac{n}{n+1}\text{Le}_{n-1}(x).\end{aligned}$$

The monic Legendre polynomials $\tilde{\text{Le}}_n := \text{Le}_n / (2^n \prod_{i=1}^{n-1} \frac{2i+1}{i+1})$ satisfy the recurrence relation

$$\tilde{\text{Le}}_{n+1}(x) = x\tilde{\text{Le}}_n(x) - \frac{n^2}{4(4n^2-1)}\tilde{\text{Le}}_{n-1}(x).$$

These results will be of great advantage for the implementation of the final model obtained by the Stochastic Galerkin method, see program 2.

3.2 Generalized Polynomial Chaos expansion

The reason we considered orthogonal polynomials last section is, that they are used for general Polynomial Chaos (gPC) expansions. A special case that considers the normal distribution, the so called Wiener-Hermite Polynomial Chaos, was first published by Wiener in 1938 [Wie38]. Further derivations and more general definitions can be found in many books about uncertainty quantification, e.g. [Sul15], [Xiu10] or [GHO19].

General polynomial Chaos expansions try to represent a random variable $X : \Omega \rightarrow \mathcal{F}$ for some space \mathcal{F} by means of a series of orthogonal polynomials p_k in another continuous random variable $\Theta : \Omega \rightarrow \mathbb{R}$ with coefficients $x_k \in \mathcal{F}$

$$X = \sum_{k=0}^{\infty} x_k p_k(\Theta).$$

The reason such expansions are studied is, that usually the density $\mu : \mathcal{D} \rightarrow \mathbb{R}$ of Θ will be chosen to belong to an easier distribution than the one of X , making it less complicated to deal with the (potentially truncated) series than with X itself.

However the distributions of X and Θ must be of a 'similar type'. Imagine e.g. a constant random variable Θ . It would be impossible to represent a Gaussian random variable X , that takes values in every non trivial interval of \mathbb{R} with probability greater than zero, but takes one fix value with probability zero, by means of a constant series of orthogonal polynomials evaluated in Θ .

Thus, put in mathematical terms, one assumes that

$$X = \tilde{X}(\Theta) \quad \text{in distribution, with } \tilde{X} \in L^2_{\mu}(\mathcal{D}, \mathcal{F}). \quad (3.3)$$

Furthermore, it is assumed that the distribution of Θ and hence its density μ is chosen such that infinitely many orthogonal polynomials exist. This is true for the standard normal distribution or the uniform distribution on $[-1/2, 1/2]$ for instance.

If there exist only finitely many orthogonal polynomials, it will not be possible to write every random variable $\tilde{X}(\Theta)$ in the desired form, take for instance a polynomial in Θ of degree higher than the highest degree of orthogonal polynomials.

In order to write every random variable of the form $\tilde{X}(\Theta)$ in such a way, the orthogonal polynomials w.r.t. μ need to span L_μ^2 .

Definition 2 ([Isk18] Definition 6.22). A *complete basis* $(\phi_n)_{n \in \mathbb{N}_0}$ of L_μ^2 is a sequence of linearly independent functions in L_μ^2 , such that $\overline{\text{span}(\phi_n \mid n \in \mathbb{N}_0)} = L_\mu^2$.

In the following, conditions on when orthogonal polynomials form such a complete basis are given. Note, that linear independence of orthogonal polynomials on non trivial intervals is always given by the condition on their degrees and the fact that non zero polynomials can only have as many zeros as their degree. For the proofs of completeness, the following lemma will be useful.

Lemma 1 ([Isk18] Theorem 6.26). *Let μ be a density function on $\mathcal{D} \subset \mathbb{R}$. For linearly independent functions $(\phi_n)_{n \in \mathbb{N}_0}$ the following is equivalent:*

- (i) $(\phi_n)_{n \in \mathbb{N}_0}$ is a complete basis of L_μ^2 ,
- (ii) if $f \in L_\mu^2$ is orthogonal to every ϕ_n , i.e. $\langle f \phi_n \rangle_\mu = 0$ for all n , then f has to be $0 \in L_\mu^2$.

Proof. We can assume w.l.o.g. that the ϕ_n are orthonormal. If they are not, apply the Gram Schmidt algorithm and normalise the polynomials to go over to the orthonormalised basis for which the assertions hold, if and only if they hold for the ϕ_n .

- (i) \Rightarrow (ii) Assume $(\phi_n)_{n \in \mathbb{N}_0}$ is a complete orthonormal basis and $f \in L_\mu^2$ is orthogonal to every ϕ_n . By completeness, for each $\varepsilon > 0$ there exist $N \in \mathbb{N}_0$ and $s_N \in \text{span}(\phi_0, \dots, \phi_N)$ with $\|s_N - f\| < \varepsilon$. As known from linear algebra, the orthogonal projection $P_n(f) := \sum_{k=0}^n \langle f \phi_k \rangle_\mu \phi_k$ of f onto the orthonormal basis ϕ_0, \dots, ϕ_n is the best approximation of f in $\text{span}(\phi_0, \dots, \phi_n)$ and hence satisfies for $n \geq N$

$$\|P_n(f) - f\|_\mu = \inf_{s \in \text{span}(\phi_0, \dots, \phi_n)} \|s - f\|_\mu \leq \inf_{s \in \text{span}(\phi_0, \dots, \phi_N)} \|s - f\|_\mu \leq \|s_N - f\| < \varepsilon.$$

Therefore, $P_n(f) \rightarrow f$ for $n \rightarrow \infty$. The continuity of the L_μ^2 -norm together with the orthonormality of the ϕ_k then gives

$$\|f\|_\mu^2 = \lim_{n \rightarrow \infty} \|P_n(f)\|_\mu^2 = \lim_{n \rightarrow \infty} \sum_{k=0}^n \langle f \phi_k \rangle_\mu^2 = 0.$$

Hence, $f = 0$ in L_μ^2 .

- (ii) \Rightarrow (i) Define $\mathcal{S} := \text{span}(\phi_n \mid n \in \mathbb{N}_0) \subset L_\mu^2$ and \mathcal{S}^\perp as the orthogonal complement of \mathcal{S} in L_μ^2 . Per definition, $f \in L_\mu^2$ is an element of \mathcal{S}^\perp if and only if $\langle f \phi_n \rangle_\mu = 0$ for all n . By (ii), only $f = 0$ satisfies this condition and therefore $\mathcal{S}^\perp = \{0\}$ and $L_\mu^2 = \mathcal{S} \oplus \mathcal{S}^\perp = \mathcal{S}$, which proves the assertion.

□

If μ is defined on a bounded interval $[a, b] \subset \mathbb{R}$, an infinite system of orthogonal polynomials is always a complete basis of L^2_μ , which will be shown below as in [KS51].

The famous Theorem of Weierstraß already suggests, that polynomials might be a good choice for approximation, since they lie dense in the continuous functions on compact intervals, which are again dense in L^2_μ on compact intervals:

Theorem 3 (Theorem of Weierstraß, [Sul15] Theorem 8.20 or [Isk18] Korollar 6.12). *For any continuous function $f : [a, b] \rightarrow \mathbb{R}$ on a compact interval $[a, b] \subset \mathbb{R}$ and any $\varepsilon > 0$, there exists a polynomial p such that*

$$\sup_{a \leq x \leq b} |f(x) - p(x)| < \varepsilon.$$

The proof will be omitted here and can be found in various literature as in [Isk18].

Theorem 4 (Completeness on bounded intervals, adapted from [KS51] Kapitel VIII. §7). *Let $[a, b] \subset \mathbb{R}$ be a bounded interval and μ a probability density on it. Then every infinite system of orthogonal polynomials w.r.t. μ is a complete basis of L^2_μ .*

Proof. Assume $(p_n)_{n \in \mathbb{N}_0}$ is a system of orthogonal polynomials w.r.t. μ . By Lemma 1, it remains to show that for a function $f \in L^2_\mu$, the condition $\langle f p_n \rangle_\mu = 0$ for all $n \in \mathbb{N}_0$ implies $f = 0$.

By the density of continuous functions in L^2_μ (see e.g. [AE08] Theorem X.4.18) and the theorem 3 of Weierstraß, for every $\varepsilon > 0$ there exists a continuous function $g : [a, b] \rightarrow \mathbb{R}$ with $\|f - g\|_\mu < \varepsilon/2$ and a polynomial $q(x)$ with $\sup_{a \leq x \leq b} |g(x) - q(x)| < \varepsilon/2$. With the triangle inequality, it follows

$$\begin{aligned} \|f - q\|_\mu &\leq \|f - g\|_\mu + \|g - q\|_\mu = \|f - g\|_\mu + \left(\int_a^b |g(x) - q(x)|^2 \mu(x) dx \right)^{1/2} \\ &< \frac{\varepsilon}{2} + \left(\left(\frac{\varepsilon}{2} \right)^2 \int_a^b \mu(x) dx \right)^{1/2} = \varepsilon. \end{aligned}$$

Furthermore, $\langle f p_n \rangle_\mu = 0$ implies $\langle f q \rangle_\mu = 0$, as one can develop $q(x) = \sum_{k=0}^{\deg(q)} c_k p_k(x)$ in the orthogonal polynomials. This together with Hölders inequality (see e.g. [AE08] X.4.2) yields

$$\begin{aligned} \|f\|_\mu^2 &= \langle f^2 \rangle_\mu = \langle f^2 \rangle_\mu - \langle f q \rangle_\mu \\ &= \langle f(f - q) \rangle_\mu = \int_a^b f(x)(f(x) - q(x)) \mu(x) dx \\ &\leq \left(\int_a^b |f(x)|^2 \mu(x) dx \right)^{1/2} \left(\int_a^b |f(x) - q(x)|^2 \mu(x) dx \right)^{1/2} = \|f\|_\mu \|f - q\|_\mu \\ &< \varepsilon \|f\|_\mu. \end{aligned}$$

Dividing by $\|f\|_\mu < \infty$ shows that $\|f\|_\mu = 0$ and therefore $f = 0$ in L^2_μ . □

For unbounded intervals, [KS51] described a sufficient condition for the completeness.

Theorem 5 (Completeness on \mathbb{R} , [KS51], Kapitel VIII. §7). *Let $\mu : \mathbb{R} \rightarrow \mathbb{R}$ be a probability density and $v(t) = \int_{-\infty}^{\infty} |x|^t \mu(x) dx$. If*

$$\lim_{n \rightarrow \infty} \frac{1}{n} v(2n)^{1/2n} = 0, \quad (3.4)$$

then all infinite systems of orthogonal polynomials form a complete basis of L_{μ}^2 .

Note that the function v represents the absolute moments of a distribution given by the density function μ , which for even natural numbers coincide with the (non absolute) moments. Thus, expressed in a sloppy way, the condition states that the even moments grow slowly enough.

Proof. Let f be a function in L_{μ}^2 and $(p_n)_{n \in \mathbb{N}_0}$ a set of orthogonal polynomials. Assume that μ satisfies condition 3.4 and that $\langle f p_n \rangle_{\mu} = 0$ for all n . In the following, we try to show, that this implies $f = 0$ in order to apply theorem 1.

Define $\alpha^2 := \int_{-\infty}^{\infty} |f(x)|^2 \mu(x) dx \geq 0$, then the inequality of Hölder, see e.g. [AE08] Theorem X.4.2, yields

$$\begin{aligned} \left| \int_{-\infty}^{\infty} f(x) x^n \mu(x) dx \right| &\leq \left(\int_{-\infty}^{\infty} |f(x)|^2 \mu(x) dx \right)^{1/2} \left(\int_{-\infty}^{\infty} |x^n|^2 \mu(x) dx \right)^{1/2} \\ &= \alpha v^{1/2}(2n). \end{aligned} \quad (3.5)$$

Now consider the function $F : \mathbb{R} \rightarrow \mathbb{C}$,

$$F(t) = \int_{-\infty}^{\infty} f(x) e^{ixt} \mu(x) dx$$

for $i = \sqrt{-1}$. This function exists since the Hölder inequality gives

$$\begin{aligned} |F(t)| &\leq \int_{-\infty}^{\infty} |f(x)| |e^{ixt}| \mu(x) dx = \int_{-\infty}^{\infty} |f(x)| \mu(x) dx \\ &\leq \left(\int_{-\infty}^{\infty} |f(x)|^2 \mu(x) dx \right)^{1/2} \left(\int_{-\infty}^{\infty} |1|^2 \mu(x) dx \right)^{1/2} = \alpha \cdot 1. \end{aligned}$$

Inserting the series representation of the e function, we would like to show that F can also be written as a series

$$F(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \int_{-\infty}^{\infty} f(x) x^n \mu(x) dx.$$

This is true as long as this series converges, hence we calculate Hadamard's radius of convergence $(\lim_{n \rightarrow \infty} \sqrt[n]{|c_n|})^{-1}$, where c_n are the coefficient of the potencies of t in the series, see e.g. [AE06a] Theorem II.9.2. By equation 3.5 they satisfy

$$|c_n| = \frac{|i|^n}{n!} \left| \int_{-\infty}^{\infty} f(x) x^n \mu(x) dx \right| \leq \frac{1}{n!} \alpha v^{1/2}(2n).$$

Hence, we have

$$0 \leq \lim_{n \rightarrow \infty} \sqrt[n]{|c_n|} \leq \lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n!} \alpha v(2n)^{1/2}} = \lim_{n \rightarrow \infty} \left(\sqrt[n]{\frac{n^n}{n!} \alpha \frac{1}{n} v(2n)^{1/2n}} \right) = 0,$$

where the last equality arises from assumption 3.4 together with the boundedness of

$$\sqrt[n]{\frac{n^n}{n!} \alpha} \leq \sqrt[n]{e^n \max(\alpha, 1)^n} = e \max(\alpha, 1)$$

due to the series representation of e^n . The radius of convergence ∞ now tells us, that the series converges for every $t \in \mathbb{R}$ and thus it is a representation of $F(t)$ on \mathbb{R} .

Furthermore, our assumption $\langle f p_n \rangle_\mu = 0$ for all n is equivalent to $\langle f x^n \rangle_\mu = 0$ for all n , what can be seen by developing x^n in the orthogonal polynomials p_0, \dots, p_n and using linearity of the integral $\langle \cdot \rangle_\mu$. Therefore, all coefficients c_n in the series representation vanish and one obtains $F(t) = 0$ for all $t \in \mathbb{R}$.

Since F is - up to some constant - equal to the so called Fourier transform of $f(x)\mu(x)$, it is zero if and only if $f(x)\mu(x)$ is zero everywhere except for Lebesgue null sets. This result can e.g. be found in [KS51] Kapitel VIII. §7 or [AE08] Theorem X.9.12. But μ is non negative, implying f to be zero on $\{x \mid \mu(x) > 0\}$ except for some Lebesgue null sets, which are μ null sets, too, and hence $f = 0$ in L^2_μ . \square

For intervals that are bounded on one side, the proof can be modified and leads to similar results, see [KS51] Kapitel VIII §7.

There exists further research on the completeness of orthogonal polynomial systems, for instance a generalization concerning the support of the densities [Cra71]. However for our purposes the above results are sufficient.

Example 4. 1. The Legendre polynomials Le_n form a complete basis of L^2_μ for the density function $\mu : [-1/2, 1/2] \rightarrow \mathbb{R}$, $\mu(x) = 1$.

This is a direct consequence of Theorem 4.

2. The Hermite polynomials He_n provide a complete basis of L^2_μ for the density μ of the standard normal distribution on \mathbb{R} .

To show this, we want to apply Theorem 5. The moments $v(2n)$ can be found in [PR96]:

$$v(2n) = \frac{(2n)!}{2^n n!}.$$

They fulfil the requirements of Theorem 5, since

$$\begin{aligned} 0 \leq \frac{1}{n} v(2n)^{1/2n} &= \frac{1}{n} \left(\frac{(2n)!}{2^n n!} \right)^{1/2n} = \frac{1}{\sqrt{2n}} \left(\frac{(2n)!}{n!} \right)^{1/2n} = \frac{(2n(2n-1)\dots(n+1))^{1/2n}}{\sqrt{2n}} \\ &\leq \frac{((2n)^n)^{1/2n}}{\sqrt{2n}} = \frac{\sqrt{2n}}{\sqrt{2n}} = \frac{1}{\sqrt{n}} \rightarrow 0 \end{aligned}$$

for $n \rightarrow \infty$.

With the above, we can define gPC expansions in a reasonable way.

Theorem 6 (gPC expansion). *Let Θ be a random variable with density $\mu : \mathcal{D} \rightarrow \mathbb{R}$, such that there exists a complete basis of corresponding orthogonal polynomials p_n for L_μ^2 . Denote an arbitrary Hilbert space by \mathcal{F} (take for instance \mathbb{R} or $L^p(D, \mathbb{R})$, $p = 0, 1, 2$). Then for every random variable $X : \Omega \rightarrow \mathcal{F}$ that can be written in the form $X = \tilde{X}(\Theta)$ with $\tilde{X} \in L_\mu^2(\mathcal{D}, \mathcal{F})$, it holds that*

$$X = \tilde{X}(\Theta) = \sum_{k=0}^{\infty} x_k p_k(\Theta), \quad (3.6)$$

where the $x_k \in \mathcal{F}$ are given by

$$x_k = \frac{\langle X p_k \rangle_\mu}{\langle p_k^2 \rangle_\mu}. \quad (3.7)$$

Definition 3. Representation 3.6 is called the *general Polynomial Chaos expansion* of X w.r.t. Θ .

Note, that for a real valued random variable $X : \Omega \rightarrow \mathbb{R}$, the gPC expansion is the orthogonal projection of \tilde{X} onto $\text{span}(p_n \mid n \in \mathbb{N})$ evaluated in Θ .

Proof. Example E.12 in [Jan97] shows the isometry $L_\mu^2 \otimes \mathcal{F} \cong L_\mu^2(\mathcal{D}, \mathcal{F})$ by the map $g \otimes f \mapsto gf$ for $g \in L_\mu^2$, $f \in \mathcal{F}$. But every element $x \in L_\mu^2 \otimes \mathcal{F}$ can be written as a linear combination of simple tensors

$$x = \sum_{k=1}^m c_k (u_k \otimes v_k), \quad m \in \mathbb{N}, c_k \in \mathbb{R}, u_k \in L_\mu^2, v_k \in \mathcal{F}.$$

By the isometry, then every $\tilde{X} \in L_\mu^2(\mathcal{D}, \mathcal{F})$ can be represented in the form

$$\tilde{X} = \sum_{k=1}^m c_k u_k v_k, \quad m \in \mathbb{N}, c_k \in \mathbb{R}, u_k \in L_\mu^2, v_k \in \mathcal{F}.$$

Since the p_n span L_μ^2 , one can find $u_k^{(n)} \in \mathbb{R}$ such that $u_k = \sum_{n=0}^{\infty} u_k^{(n)} p_n$. Inserting this in the above equation yields

$$\tilde{X} = \sum_{k=1}^m c_k \left(\sum_{n \in \mathbb{N}_0} u_k^{(n)} p_n \right) v_k = \sum_{k=1}^m \sum_{n \in \mathbb{N}_0} c_k u_k^{(n)} p_n v_k = \sum_{n \in \mathbb{N}_0} p_n \left(\sum_{k=1}^m c_k u_k^{(n)} v_k \right) = \sum_{n \in \mathbb{N}_0} p_n x_n,$$

where the linear combination $x_n := \sum_{k=1}^m c_k u_k^{(n)} v_k$ of vectors in \mathcal{F} lies in \mathcal{F} itself.

Due to orthogonality of the p_n , the x_n are the coefficients of the orthogonal projection of \tilde{X} onto the span of all p_n which verifies representation 3.7. \square

The above derived gPC expansion allows us to study the Black Scholes equation with uncertain volatility 2.4 by the Stochastic Galerkin method.

Chapter 4

Application of the Stochastic Galerkin method

The Stochastic Galerkin method serves to transform a differential equation with stochastic variables into several deterministic differential equations that can be solved numerically. The solution for the equation in stochastic variables can then be reconstructed from the solutions of the deterministic ones. In this chapter, the method will be explained in detail while applying it to the Black Scholes equation with random volatility.

The Stochastic Galerkin method became popular in uncertainty quantification within the last years. Ghanem and Spanos are mentioned by [Sul15] to be the first to have enhanced the method's popularity with the book [GS91] and several articles like [GS97]. Later, many authors in uncertainty quantification treated the method in their books like [Xiu10], [Sul15] and [GHO19] just to mention some. The theory applied in this chapter was taken from the first two of them.

In the Black Scholes equation with random volatility 2.4

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

the volatility $\Sigma : \Omega \rightarrow \mathbb{R}$ and the solution $V : (0, \infty) \times [0, T] \times \Omega \rightarrow \mathbb{R}$ are modelled to be a random variable and a stochastic process respectively on a common probability space. We require the distribution of Σ to be given and try to find the solution V on that basis.

In the following the Stochastic Galerkin method will be explained and applied in 6 steps.

Step 1: Writing the stochastic variables as gPC expansions

First, we rewrite all variables modelled as random variables or stochastic processes as gPC expansions w.r.t. a continuous random variable $\Theta : \Omega \rightarrow \mathbb{R}$ with density $\mu : \mathcal{D} \rightarrow \mathbb{R}$. Assume, all conditions of theorem 6 are satisfied for Σ and V , i.e. there exists a complete basis of orthogonal polynomials p_n of L^2_μ as well as an L^2_μ functions $\tilde{\Sigma} : \mathcal{D} \rightarrow \mathbb{R}$ and a

$L^2_\mu(\mathcal{D}, L^2((0, \infty) \times [0, T], \mathbb{R}))$ function $\tilde{V} : (0, \infty) \times [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ with

$$\Sigma = \tilde{\Sigma}(\Theta) \quad \text{and} \quad V(S, t, \cdot) = \tilde{V}(S, t, \Theta) \quad (4.1)$$

for all S and t in distribution. Then, theorem 6 shows that we can write Σ and V in the gPC form

$$\Sigma = \sum_{k=0}^{\infty} \sigma_k p_k(\Theta) \quad \text{and} \quad V(S, t, \cdot) = \sum_{k=0}^{\infty} v_k(S, t) p_k(\Theta),$$

with $\sigma_k \in \mathbb{R}$ and $v_k \in L^2((0, \infty) \times [0, T], \mathbb{R})$.

It should be mentioned that for Σ , condition 4.1 can be checked, as Σ 's distribution is known, whereas the distribution of V should be derived by this method and is therefore unknown. Thus, one can not verify condition 4.1 for V at this step. Instead one postulates it without proof, applies the Stochastic Galerkin method and verifies the assumption afterwards. However, this is a typical approach when applying the Stochastic Galerkin method.

Step 2: Inserting the gPC expansions in the differential equation

In the next step, these expansions are inserted in the PDE. Here one takes into account that the partial derivatives of V in S and t pass over to the coefficients v_k

$$\begin{aligned} \frac{\partial V(S, t, \cdot)}{\partial t} &= \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial t} p_k(\Theta), \\ \frac{\partial V(S, t, \cdot)}{\partial S} &= \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial S} p_k(\Theta) \quad \text{and} \\ \frac{\partial^2 V(S, t, \cdot)}{\partial S^2} &= \sum_{k=0}^{\infty} \frac{\partial^2 v_k(S, t)}{\partial S^2} p_k(\Theta), \end{aligned}$$

because theorem 6 is still applicable for these.

That is true since partial differentiation in S or t does not affect the integrability of \tilde{V} with $V = \tilde{V}(\Theta)$ on \mathcal{D} . If $\tilde{V} \in L^2_\mu(\mathcal{D}, L^2\{(0, \infty) \times [0, T], \mathbb{R}\}) \cong L^2_\mu(\mathcal{D}, L^2\{(0, \infty), L^2([0, T], \mathbb{R})\})$ holds, see [AE08] Theorem X.6.22 for the isomorphism, it follows that $\partial \tilde{V} / \partial t$ lies in $L^2_\mu(\mathcal{D}, L^2\{(0, \infty), L^1([0, T], \mathbb{R})\})$. The descending integrability in the differentiation variable can be found in [Eva98] Chapter 5.2.3 Theorem 1, whereas the constant integrability in the other variables is a consequence of Bemerkung VII.2.7 in [AE06b]. Thus, integrability on \mathcal{D} stays the same and $\partial V / \partial t = \partial \tilde{V}(\Theta) / \partial t$ can be written as function of Θ . This function lies in $L^2_\mu(\mathcal{D}, L^2\{(0, \infty), L^1([0, T], \mathbb{R})\})$. Because $L^2\{(0, \infty), L^1([0, T], \mathbb{R})\}$ is a Hilbert space, theorem 6 can be applied to $\partial \tilde{V} / \partial t$ and its k th gPC coefficient is given by

$$\begin{aligned} \frac{1}{\langle p_k^2 \rangle_\mu} \left\langle \frac{\partial \tilde{V}(\Theta)}{\partial t} p_k \right\rangle_\mu &= \frac{1}{\langle p_k^2 \rangle_\mu} \int_{\mathcal{D}} \frac{\partial \tilde{V}(x)}{\partial t} p_k(x) \mu(x) dx = \frac{1}{\langle p_k^2 \rangle_\mu} \frac{\partial}{\partial t} \int_{\mathcal{D}} \tilde{V}(x) p_k(x) \mu(x) dx \\ &= \frac{\partial}{\partial t} \frac{1}{\langle p_k^2 \rangle_\mu} \int_{\mathcal{D}} \tilde{V}(x) p_k(x) \mu(x) dx = \frac{\partial v_k}{\partial t} \end{aligned}$$

due to independence of p_k and μ of t . The same arguments show the existence of gPC expansion for the considered partial derivatives in S .

So, after inserting the above gPC expansions, the Black Scholes equation becomes

$$0 = \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial t} p_k(\Theta) + \frac{1}{2} \left(\sum_{k=0}^{\infty} \sigma_k p_k(\Theta) \right)^2 S^2 \sum_{k=0}^{\infty} \frac{\partial^2 v_k(S, t)}{\partial S^2} p_k(\Theta) \quad (4.2)$$

$$+ rS \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial S} p_k(\Theta) - r \sum_{k=0}^{\infty} v_k(S, t) p_k(\Theta).$$

The reason of this becomes clear in the next step.

Step 3: Multiplication of the equation with an orthogonal polynomial evaluated in Θ and application of the expectation, for all orthogonal polynomials

The equation is multiplied by $p_n(\Theta)$, and the expected value is applied on both sides of the PDE 4.2. This is done for each $n \in \mathbb{N}_0$. Keeping in mind that for function f integrable w.r.t. μ , the expected value of $f(\Theta)$ is given by $E(f(\Theta)) = \langle f \rangle_\mu$, one obtains deterministic equations for all coefficients $v_n, n \in \mathbb{N}_0$

$$0 = \langle 0 \cdot p_n \rangle_\mu = \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial t} \langle p_k p_n \rangle_\mu + \frac{1}{2} S^2 \sum_{i,k,l=0}^{\infty} \sigma_i \sigma_k \frac{\partial^2 v_l(S, t)}{\partial S^2} \langle p_i p_k p_l p_n \rangle_\mu$$

$$+ rS \sum_{k=0}^{\infty} \frac{\partial v_k(S, t)}{\partial S} \langle p_k p_n \rangle_\mu - r \sum_{k=0}^{\infty} v_k(S, t) \langle p_k p_n \rangle_\mu.$$

Due to orthogonality, all $\langle p_k p_n \rangle_\mu$ with $k \neq n$ vanish and one obtains a countably infinite system of coupled PDEs similar to the Black Scholes equation

$$0 = \frac{\partial v_n(S, t)}{\partial t} \langle p_n^2 \rangle_\mu + \frac{1}{2} S^2 \sum_{i,k,l=0}^{\infty} \sigma_i \sigma_k \frac{\partial^2 v_l(S, t)}{\partial S^2} \langle p_i p_k p_l p_n \rangle_\mu$$

$$+ rS \frac{\partial v_n(S, t)}{\partial S} \langle p_n^2 \rangle_\mu - r v_n(S, t) \langle p_n^2 \rangle_\mu.$$

This can be reformulated by dividing by $\langle p_n^2 \rangle_\mu > 0$ and renaming the so called Galerkin multiplication tensor

$$M_{ikln} := \frac{\langle p_i p_k p_l p_n \rangle_\mu}{\langle p_n^2 \rangle_\mu}. \quad (4.3)$$

These values exist for all $i, k, l, n \in \mathbb{N}_0$ even if the p_k are just in L_μ^2 , since their products $p_i p_k p_l p_n$ are polynomials again and can therefore be written as linear combinations of the orthogonal polynomials $p_0, \dots, p_{i+k+l+n}$. The linearity of the integral shows that $\langle p_i p_k p_l p_n \rangle_\mu$ is finite.

With this, we get the equations we will be working with for numerical evaluation

$$0 = \frac{\partial v_n(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{i,k,l=0}^{\infty} \sigma_i \sigma_k \frac{\partial^2 v_l(S, t)}{\partial S^2} M_{ikln} + rS \frac{\partial v_n(S, t)}{\partial S} - r v_n(S, t), \quad (4.4)$$

$$n \in \mathbb{N}_0, S \in (0, \infty), t \in [0, T].$$

The term containing the multiplication tensors will be called coupling term, as the equations are coupled by it.

Step 4: Transforming the boundary values to the solution's gPC coefficients

To solve these equations, boundary values need to be stated. They are given by the particular financial derivative to be priced and are usually deterministic.

In this work, European Call options, as in example 1, will be considered for demonstrative purposes. In the deterministic case, the final value $V(S, T) = (S - \text{strike})^+$, $S \in (0, \infty)$ and the boundary values $V(0, t) = 0$ and $V(S, t)/S \rightarrow 1$ for $S \rightarrow \infty$, $t \in [0, T]$ are independent of the value of σ . Equation 2.5 shows, that the boundary values of the solution to the Black Scholes equation with random volatility 2.4 are deterministic, too. It follows that they appear only in the coefficient v_0 , whose polynomial $p_0 \equiv 1$ is deterministic, since all other orthogonal polynomials lead to stochastic behaviour of the solution. Therefore, the other coefficients vanish at the boundaries. The final value and the boundary conditions for the system of equations 4.4 can be summarized as

$$\begin{aligned} v_0(S, T) &= (S - \text{strike})^+, & S \in (0, \infty), \\ v_n(S, T) &= 0, & S \in (0, \infty), \\ v_0(S, t), v_n(S, t) &\xrightarrow{S \rightarrow 0} 0, & t \in [0, T], \\ v_0(S, t)/S &\xrightarrow{S \rightarrow \infty} 1, & t \in [0, T], \\ v_n(S, t)/S &\xrightarrow{S \rightarrow \infty} 0, & t \in [0, T] \end{aligned} \quad (4.5)$$

for all $n \in \mathbb{N} \setminus \{0\}$.

Step 5: Solving the system of differential equations

Now, one can solve the system 4.4 (approximately), what is typically done numerically. Therefore the gPC expansions and hence the system of equations have to be truncated to finite length.

For the Black Scholes equation, one first approximates Σ by its truncated gPC expansion

$$\Sigma^K := \sum_{k=0}^K \sigma_k p_k(\Theta)$$

for one $K \in \mathbb{N}_0$, which is possible because of the convergence of gPC expansions given by Theorem 6. Then, we search for the solution of system 4.4 of the form

$$V^N(S, t, \omega) := \sum_{k=0}^N v_k^N(S, t) p_k(\Theta(\omega)), \quad S \in (0, \infty), t \in [0, T], \omega \in \Omega \quad (4.6)$$

for some $N \in \mathbb{N}_0$ and $v_k^N \in L^2((0, \infty) \times [0, T], \mathbb{R})$ for that we hope to get a good approximation V^N of V . Since V is not known, one cannot choose N to bound the error $V^N - V$ to some prescribed size as one can for K .

The system of equations for the v_k^N is then given by

$$0 = \frac{\partial v_n^N(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{i,k=0}^K \sum_{l=0}^N \sigma_i \sigma_k \frac{\partial^2 v_l^N(S, t)}{\partial S^2} M_{ikln} + rS \frac{\partial v_n^N(S, t)}{\partial S} - r v_n^N(S, t), \quad (4.7)$$

$$n = 0, \dots, N, S \in (0, \infty), t \in [0, T],$$

so to say a truncated version of the original infinite system 4.4.

Remark 1. Note that the space of all functions $f : (0, \infty) \times [0, T] \times \mathcal{D}$ of the type $f(S, t, x) = \sum_{k=0}^N f_k(S, t) p_k(y)$ with $f_k \in L^2((0, \infty) \times [0, T], \mathbb{R})$ is a subspace of the space $L^2_\mu(\mathcal{D}, L^2((0, \infty) \times [0, T], \mathbb{R}))$ that is spanned in x direction by the first $N \in \mathbb{N}$ orthogonal polynomials. In this notation, solving the truncated Stochastic Galerkin system 4.7 is equivalent to searching the solution to the original infinite system 4.4 in the subspace explained above. The truncated solution is therefore optimal in a certain sense.

Remark 2. However, convergence of this solution V^N , with coefficients derived by 4.7, to V as $N \rightarrow \infty$ is not trivial, as in general $v_n^N \neq v_k$ for $k = 0, \dots, N$. This can be seen by comparing the PDEs 4.7 and 4.4 for v_n^N and v_k especially in the coupling term. Thus, the error between V^N and V consists of an error made by solving a slightly different PDE plus the gPC truncation error whose size cannot be prescribed.

We tried to estimate this total error and proof convergence by adapting the convergence proof in [GX08]. But our attempt failed, since unlike the problem in the paper, second derivatives appear in the coupling term and its coefficient depends on S , ruining some important arguments. After consulting further literature on the convergence of truncated Stochastic Galerkin systems for different PDEs like [DJI19], [SJ18], [LJ18] and [AN16] for ordinary differential equations, it turned out to be a very problem related task for PDEs. The proof for the Black Scholes equation with random volatility is therefore still open to further research.

Assuming convergence is given in order to consider the truncated Stochastic Galerkin solution as meaningful, the numerics used in this work will be explained in the next chapter.

For the following computations, one usually summarizes the system of PDEs to one vector valued PDE defining

$$\mathbf{v}(S, t) := \begin{pmatrix} v_0^N(S, t) \\ v_1^N(S, t) \\ \vdots \\ v_N^N(S, t) \end{pmatrix}.$$

System 4.7 can be reformulated to

$$\mathbf{0}_{N+1} = \frac{\partial \mathbf{v}(S, t)}{\partial t} + \frac{1}{2} S^2 \mathbf{A} \frac{\partial^2 \mathbf{v}(S, t)}{\partial S^2} + rS \frac{\partial \mathbf{v}(S, t)}{\partial S} - r \mathbf{v}(S, t), \quad (4.8)$$

where bold variables indicate vectors in \mathbb{R}^{N+1} like $\mathbf{0}_{N+1} = (0, \dots, 0)^T$ or matrices in $\mathbb{R}^{(N+1) \times (N+1)}$ like the so called coupling matrix \mathbf{A} with

$$\mathbf{A}[n, l] := \sum_{i, k=0}^K \sigma_i \sigma_k M_{ikln}, \quad n, l = 0, \dots, N. \quad (4.9)$$

This system is parabolic, if the real values of all eigenvalues of \mathbf{A} are positive, see [Str04] Chapter 6.2. Then, by [Str04] Theorem 6.2.1, it is *well posed*, i.e. for every final value $\mathbf{v}(S, T)$, there exists a unique solution depending continuously on the initial data. Because of this and the Lax-Richtmeyer theorem showing convergence of a finite difference solution to the continuous solution, the later derived numerics will only consider the case of parabolicity, see next section.

The boundary conditions and the final condition in time corresponding to the European Call option can be written as

$$\begin{aligned} \mathbf{v}(S, T) &= \begin{pmatrix} (S - \text{strike})^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad S \in (0, \infty), \\ \mathbf{v}(S, t) &\xrightarrow{S \rightarrow 0} \mathbf{0}_{N+1}, \quad t \in [0, T], \quad \text{and} \\ \frac{1}{S} \mathbf{v}(S, t) &\xrightarrow{S \rightarrow \infty} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad t \in [0, T]. \end{aligned} \quad (4.10)$$

Step 6: Reconstructing the stochastic solution

After one obtained the v_k^N numerically, one can reconstruct V^N by formula 4.6. One can also consider more illustrative quantities directly from the gPC coefficients like

- the expected value

$$E(V^N) = E \left(\sum_{k=0}^N v_k^N p_k(\Theta) \right) = \left\langle \sum_{k=0}^N v_k^N p_k \right\rangle_{\mu} = \sum_{k=0}^N v_k^N \langle p_k \rangle_{\mu} = v_0^N, \quad (4.11)$$

since $\langle p_k \rangle_{\mu} = \langle 1 p_k \rangle_{\mu} = \langle p_0 p_k \rangle_{\mu} = \gamma_0 \delta_{0k}$ with $\gamma_0 = \int_{\mathcal{D}} \mu(x) dx = 1$,

- or the variance

$$\begin{aligned} \text{Var}(V^N(S, t)) &= E((V^N - E(V^N))^2) = \left\langle \left(\sum_{k=0}^N v_k^N p_k - v_0^N \right)^2 \right\rangle_{\mu} \\ &= \sum_{k, j=1}^N v_k^N v_j^N \langle p_k p_j \rangle_{\mu} = \sum_{k=1}^N (v_k^N)^2 \gamma_k. \end{aligned}$$

Chapter 5

Numerical implementation

This chapter will explain numerical methods used to obtain an approximate solution to the truncated system of PDEs 4.8 for the gPC coefficients $\mathbf{v} := (v_0^N, \dots, v_N^N)$

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

$S \in (0, \infty), t \in [0, T]$, that was obtained by the Stochastic Galerkin method in the last chapter. The explicit finite difference scheme that will be derived by discretization of equation 4.8 will be investigated for properties like consistency, stability and convergence needed to ensure meaningful computations, see section 5.2.

All derivations consider a European Call option with maturity T and strike price $strike > 0$, leading to final and boundary conditions as in 4.10

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

$$\mathbf{v}(S, t) \xrightarrow{S \rightarrow 0} \mathbf{0}_{N+1}, \quad t \in [0, T], \quad \text{and}$$

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

For the computation, the domain of the $\mathbf{v} : (0, \infty) \times [0, T] \rightarrow \mathbb{R}^{N+1}$ has to be adapted to a finite domain. This is done via a transformation of variables analogue to the one in [ZWCS13] Chapter 2.2.5

$$\begin{aligned} \zeta &= \frac{S}{S + strike}, \\ \tau &= T - t, \end{aligned} \tag{5.1}$$

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

Then $\bar{\mathbf{v}}$ is defined on the finite domain $(0, 1) \times [0, T]$. Later, the results will be transformed back by

$$\begin{aligned} S &= \text{strike} \frac{\zeta}{1 - \zeta}, \\ t &= T - \tau, \\ \mathbf{v}(S, t) &= (S + \text{strike}) \bar{\mathbf{v}}(\zeta, \tau) = (S + \text{strike}) \bar{\mathbf{v}}\left(\frac{S}{S + \text{strike}}, T - t\right). \end{aligned} \tag{5.2}$$

Computing the partial derivatives needed in 4.7 in terms of $\bar{\mathbf{v}} := (\bar{v}_0^N, \dots, \bar{v}_N^N)^T$, ζ and τ yields as in [ZWCS13]

$$\begin{aligned} \frac{\partial \mathbf{v}(S, t)}{\partial t} &= -\frac{\text{strike}}{1 - \zeta} \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} \\ \frac{\partial \mathbf{v}(S, t)}{\partial S} &= \bar{\mathbf{v}}(\zeta, \tau) + (1 - \zeta) \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} \\ \frac{\partial^2 \mathbf{v}(S, t)}{\partial S^2} &= \frac{(1 - \zeta)^3}{\text{strike}} \frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2}. \end{aligned}$$

Inserting this in PDE 4.8, it follows that $\bar{\mathbf{v}}$ satisfies

$$\begin{aligned} \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} &= \frac{1}{2} \zeta^2 (1 - \zeta)^2 \mathbf{A} \frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2} + r \zeta (1 - \zeta) \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} - r (1 - \zeta) \bar{\mathbf{v}}(\zeta, \tau), \\ &\zeta \in (0, 1), \tau \in [0, T], \end{aligned} \tag{5.3}$$

with \mathbf{A} as defined before in 4.9.

As discussed for equation 4.8, this PDE is also parabolic according to the definition in [Str04] chapter 6.2., if the real parts of all eigenvalues of \mathbf{A} are positive. Then the equation is *well posed*, meaning for all initial values there exists a unique solution depending continuously on the initial value, see [Str04] Theorem 6.2.1 and chapter 1.2. Since we wish to have this property, in the following we assume that the real parts of all eigenvalues of \mathbf{A} are positive. In real computations, this has to be checked.

The boundary conditions and the final condition in time 4.10 transform to

$$\begin{aligned} \bar{\mathbf{v}}(\zeta, 0) &= \begin{pmatrix} (2\zeta - 1)^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, & \zeta \in (0, 1), \\ \bar{\mathbf{v}}(\zeta, \tau) &\xrightarrow{\zeta \rightarrow 0} \mathbf{0}_{N+1}, & \tau \in [0, T], \quad \text{and} \\ \bar{\mathbf{v}}(\zeta, \tau) &\xrightarrow{\zeta \rightarrow 1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, & \tau \in [0, T]. \end{aligned}$$

With this, $\bar{\mathbf{v}}$ can be continuously extended to $[0, 1] \times [0, T]$ by $\bar{\mathbf{v}}(0, \tau) := \mathbf{0}_{N+1}$ and $\bar{\mathbf{v}}(1, \tau) := (1, 0, \dots, 0)^T$. This simplifies the boundary conditions and the scheme derived below. Now, we can start deriving a scheme to compute $\bar{\mathbf{v}}$ and finally \mathbf{v} . However, at first the coupling matrix \mathbf{A} has to be computed, as it will be a coefficient in the scheme.

5.1 Computation of the coupling matrix

Whichever scheme might be used, usually one has to compute the coupling matrix \mathbf{A} from 4.9

$$\mathbf{A}[n, l] := \sum_{i,k=0}^K \sigma_i \sigma_k M_{ikln}, \quad n, l = 0, \dots, N,$$

where the M_{ikln} were defined to be

$$M_{ikln} := \frac{\langle p_i p_k p_l p_n \rangle_\mu}{\langle p_n^2 \rangle_\mu},$$

as it will occur in the scheme. One way to do this is explained in this section.

The difficulty lies in the computation of the integrals occurring here, but suitable quadrature methods were e.g. explained in [Sul15], [Wil62] or [GST07]. In this work we will use Gauss quadrature formulas since they enable exact computation of $\langle p_i p_k p_l p_n \rangle_\mu$.

Definition 4 ([Sul15] Definition 9.8). A quadrature formula

$$\int_{\mathcal{D}} f(x) \mu(x) dx \approx \sum_{i=1}^G w_i f(x_i)$$

with G nodes $x_1, \dots, x_G \in \mathcal{D}$ and weights $w_1, \dots, w_G \in \mathbb{R}$ for any real valued function f on \mathcal{D} is called a G -point Gaussian quadrature rule, if the nodes x_i are the zeros of the μ -orthogonal polynomial p_G and the weights are the integrals over the Lagrange basis polynomials, i.e.

$$w_i := \int_{\mathcal{D}} \prod_{j=1, j \neq i}^G \frac{x - x_j}{x_i - x_j} \mu(x) dx. \quad (5.4)$$

Note that all zeros of orthogonal polynomials are distinct, see e.g. [Sul15] Theorem 8.16 (a).

The Gaussian quadrature rule is the most powerful in the sense of order of accuracy.

Theorem 7 ([Sul15] Lemma 9.2, Theorem 9.9). (a) *There is no quadrature rule with G nodes of order greater or equal to $2G$, i.e. there do not exist nodes $x_1, \dots, x_G \in \mathcal{D}$ and weights $w_1, \dots, w_G \in \mathbb{R}$, such that the quadrature rule is exact*

$$\int_{\mathcal{D}} q(x) \mu(x) dx = \sum_{i=1}^G w_i q(x_i)$$

for all polynomials q with degree smaller or equal to $2G$.

- (b) The G -point Gauss quadrature rule has order of accuracy $2G - 1$, i.e. for all polynomials of degree less than or equal to $2G - 1$, the quadrature gives the correct value for the integral.

Proof. (a) Let $x_1, \dots, x_G \in \mathcal{D}$ and $w_1, \dots, w_G \in \mathbb{R}$ be arbitrary and consider the polynomial $q(x) := ((x - x_1)\dots(x - x_G))^2$ of order $2G$. Since the x_i are zeros of q , the quadrature rule gives

$$\sum_{i=1}^G w_i q(x_i) = 0 < \int_{\mathcal{D}} q(x) \mu(x) dx,$$

where the integral is greater than zero by monotonicity of the integral, because q is a non-negative function with zeros only in the x_i , i.e. q is positive on \mathcal{D} apart from some μ null sets.

- (b) Let q be a polynomial of degree $\deg(q) \leq 2G - 1$ and p_G denote the orthogonal polynomial w.r.t. μ of degree G . Dividing q by p_G with remainder we obtain

$$q(x) = g(x)p_G(x) + r(x),$$

where g and r are polynomials of degree less than or equal to $G - 1$. Integration w.r.t. $\mu(x) dx$ and application of the quadrature rule gives

$$\begin{aligned} \int_{\mathcal{D}} q(x) \mu(x) dx &= \int_{\mathcal{D}} g(x) p_G \mu(x) dx + \int_{\mathcal{D}} r(x) \mu(x) dx \\ \sum_{i=1}^G w_i q(x_i) &= \sum_{i=1}^G w_i g(x_i) p_G(x_i) + \sum_{i=1}^G w_i r(x_i), \end{aligned}$$

where the x_i are the zeros of p_G and the w_i are defined as in 5.4. Therefore, the first term on the right hand side of the equation for the Gauss quadrature rules vanishes. But the first term on the right hand side of the integral equation also equals zero, since p_G is orthogonal to every polynomial of degree smaller than G , see Theorem 1, in particular to g .

It remains to show $\int_{\mathcal{D}} r(x) \mu(x) dx = \sum_{i=1}^G w_i r(x_i)$. This is true, as r coincides with its Lagrange interpolation on the $G > \deg(r)$ points x_1, \dots, x_G , see [Sul15] Theorem 8.18 together with the fact that polynomials p are uniquely determined by their values in $\deg(p)$ distinct points. Therefore

$$\begin{aligned} \int_{\mathcal{D}} q(x) \mu(x) dx &= \int_{\mathcal{D}} r(x) \mu(x) dx = \int_{\mathcal{D}} \sum_{i=1}^G \left(r(x_i) \prod_{j=1, j \neq i}^G \frac{x - x_j}{x_i - x_j} \right) dx \\ &= \sum_{i=1}^G r(x_i) \int_{\mathcal{D}} \frac{x - x_j}{x_i - x_j} dx = \sum_{i=1}^G r(x_i) w_i = \sum_{i=1}^G w_i q(x_i). \end{aligned}$$

□

The Gaussian quadrature with $K + N + 1$ nodes enables exact computation of the M_{ikln} , since only polynomials of degree up to $K + K + N + N = 2(K + N)$ are integrated.

Computation of the Gauss quadrature nodes and weights:

The computation of the nodes is done via a companion matrix as described in [GW69] Chapter 2 or [GST07] Chapter 5.3. Since the zeros of the considered orthogonal polynomials coincide with the zeros of the monic or normalized orthogonal polynomials, we will work with the latter two below. Assume the monic orthogonal polynomials \hat{p} w.r.t. μ satisfy the recurrence relation from Theorem 2

$$\hat{p}_{n+1}(x) = (x - a_n)\hat{p}_n(x) - b_n\hat{p}_{n-1}(x), \quad \text{where } b_n = \frac{\hat{\gamma}_n}{\hat{\gamma}_{n-1}}, n \geq 0.$$

Then the normalized orthogonal polynomials $\tilde{p}_n = \hat{p}_n/\hat{\gamma}_n$, where $\hat{\gamma}_n := \langle \hat{p}_n^2 \rangle_\mu$, satisfy

$$\begin{aligned} \tilde{p}_{n+1}(x) &= \frac{\hat{p}_{n+1}(x)}{\sqrt{\hat{\gamma}_{n+1}}} = (x - a_n) \frac{\sqrt{\hat{\gamma}_n}}{\sqrt{\hat{\gamma}_{n+1}}} \tilde{p}_n(x) - b_n \frac{\sqrt{\hat{\gamma}_{n-1}}}{\sqrt{\hat{\gamma}_{n+1}}} \tilde{p}_{n-1}(x) \\ &= (x - a_n) \frac{\sqrt{\hat{\gamma}_n}}{\sqrt{\hat{\gamma}_{n+1}}} \tilde{p}_n(x) - \frac{\hat{\gamma}_n}{\sqrt{\hat{\gamma}_{n-1}}\sqrt{\hat{\gamma}_{n+1}}} \tilde{p}_{n-1}(x) \\ \Leftrightarrow x\tilde{p}_n(x) &= \sqrt{\frac{\hat{\gamma}_{n+1}}{\hat{\gamma}_n}} \tilde{p}_{n+1}(x) + a_n \tilde{p}_n(x) + \sqrt{\frac{\hat{\gamma}_n}{\hat{\gamma}_{n-1}}} \tilde{p}_{n-1}(x) \\ &=: \beta_n \tilde{p}_{n+1}(x) + \alpha_n \tilde{p}_n(x) + \beta_{n-1} \tilde{p}_{n-1}(x) \end{aligned} \quad (5.5)$$

with $\beta_n := \sqrt{\hat{\gamma}_{n+1}/\hat{\gamma}_n}$ and $\alpha_n := a_n$. Written in matrix form, this gives

$$x \begin{pmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \tilde{p}_2(x) \\ \vdots \\ \tilde{p}_{G-2}(x) \\ \tilde{p}_{G-1}(x) \end{pmatrix} = \begin{pmatrix} \alpha_0 & \beta_0 & & & & \\ \beta_0 & \alpha_1 & \beta_1 & & & \\ & \beta_1 & \alpha_2 & \beta_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{G-3} & \alpha_{G-2} & \beta_{G-2} \\ & & & & \beta_{G-2} & \alpha_{G-1} \end{pmatrix} \begin{pmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \tilde{p}_2(x) \\ \vdots \\ \tilde{p}_{G-2}(x) \\ \tilde{p}_{G-1}(x) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \beta_{G-1} \tilde{p}_G(x) \end{pmatrix},$$

where empty entries are zeros.

Denoting the matrix in the above equation by \mathbf{R} and the vector of polynomials by $\tilde{\mathbf{p}}(x)$, a value x_i is a zero of \tilde{p}_G and therefore also of p_G if and only if $x_i \tilde{\mathbf{p}}(x_i) = \mathbf{R} \tilde{\mathbf{p}}(x_i)$. Expressed differently, the zeros of p_G are exactly the eigenvalues of \mathbf{R} . The nodes are therefore determined by calculating the eigenvalues of \mathbf{R} .

For the weights, we consider eigenvectors of \mathbf{R} and proceed as in [GST07]. Because the \tilde{p}_n are orthonormal and $\tilde{p}_n \tilde{p}_m$ is of maximum degree $2(G - 1) < 2G - 1$, the Gauss quadrature can be applied on the orthogonality condition 3.1

$$\delta_{nm} = \int_{\mathcal{D}} \tilde{p}_n(x) \tilde{p}_m(x) \mu(x) dx = \sum_{i=1}^G w_i \tilde{p}_n(x_i) \tilde{p}_m(x_i).$$

In matrix notation, this is equivalent to

$$\mathbf{I}_G = \tilde{\mathbf{P}}\mathbf{W}\tilde{\mathbf{P}}^T$$

for the matrix $\tilde{\mathbf{P}}[i, j] = \tilde{p}_i(x_{j+1})$ and the diagonal matrix $\mathbf{W}[i, i] = w_{i+1}$ for $i, j = 0, \dots, G-1$. Since \mathbf{I}_G is invertible, \mathbf{P} must have this property, too, and we can write

$$\mathbf{W} = (\tilde{\mathbf{P}}^T \tilde{\mathbf{P}})^{-1},$$

$$\text{i.e. } w_j = \left(\sum_{k=0}^{G-1} (\tilde{p}_k(x_j))^2 \right)^{-1}.$$

Assume now, $\mathbf{v}^{(j)} = (v_0^{(j)}, \dots, v_{G-1}^{(j)})^T$ is an eigenvector of \mathbf{R} corresponding to the eigenvalue x_j with Euclidean norm $\|\mathbf{v}^{(j)}\|_E = 1$ and note that $\tilde{\mathbf{p}}(x_j)$ is one as well. Because all eigenvalues have multiplicity 1, as the zeros of orthogonal polynomials are all different see e.g. [Sul15] Theorem 8.16 (a), eigenvectors are unique up to multiplication by a scalar, i.e. we obtain

$$\mathbf{v}^{(j)} = \frac{v_0^{(j)}}{p_0(x_j)} \tilde{\mathbf{p}}(x_j) = v_0^{(j)} \tilde{\mathbf{p}}(x_j),$$

from comparing the first entries and inserting $\tilde{p}_0 \equiv 1$. Now, one can compute the w_j by

$$w_j = \left(\sum_{k=0}^{G-1} (\tilde{p}_k(x_j))^2 \right)^{-1} = \|\tilde{\mathbf{p}}_k(x_j)\|_E^{-2} = \left\| \frac{1}{v_0^{(j)}} \mathbf{v}^{(j)} \right\|_E^{-2} = \frac{(v_0^{(j)})^2}{\|\mathbf{v}^{(j)}\|_E^2} = (v_0^{(j)})^2.$$

This tells us that the weight w_j corresponding to the node x_j is given by the squared first entry of the normalized eigenvector of \mathbf{R} corresponding to the eigenvalue x_j .

In summary, using Pseudo code in Matlab notation, where blue text denotes comments and descriptions, one computes the G Gauss quadrature weights $\mathbf{w} := (w_1, \dots, w_G)^T$ and nodes $\mathbf{x} := (x_1, \dots, x_G)^T$ by

Algorithm 1 Deriving the Gauss Quadrature nodes

Input: Number G of Gauss quadrature points.

Output: Gauss quadrature nodes $\mathbf{x} = (x_1, \dots, x_G)^T$ and weights $\mathbf{w} = (w_1, \dots, w_G)^T$.

Calculating the companion matrix \mathbf{R} :

Calculate $\alpha := (\alpha_0, \dots, \alpha_{G-1})^T$ and $\beta := (\beta_0, \dots, \beta_{G-2})^T$ by formula 5.5;

$\mathbf{R} := \text{diag}(\beta, 1) + \text{diag}(\alpha) + \text{diag}(\beta, -1)$;

Computing its eigenvectors (columns of \mathbf{V}) and eigenvalues (diagonal entries of \mathbf{L}):

$[\mathbf{V} \ \mathbf{L}] = \text{eig}(\mathbf{R})$;

Saving the nodes x_i and weights w_i in sorted column vectors \mathbf{x} and \mathbf{w} :

$[\mathbf{x} \ \text{SortIndex}] = \text{sort}(\text{diag}(\mathbf{L}))$;

$\mathbf{V} = \mathbf{V}(:, \text{SortIndex})'$;

$\mathbf{w} = \mathbf{V}(:, 1) \cdot \hat{\ }^2$;

After we derived the $G = K + N + 1$ Gauss nodes x_1, \dots, x_{K+N+1} and weights w_1, \dots, w_{K+N+1} , we can calculate the M_{ikln} by Gaussian quadrature. Denote by \mathbf{M} the 4 dimensional tensor with

$$\mathbf{M}[i, k, l, n] = M_{ikln} = \frac{\langle p_i p_k p_l p_n \rangle_\mu}{\langle p_n^2 \rangle_\mu} = \frac{\langle p_i p_k p_l p_n \rangle_\mu}{\gamma_n} = \frac{1}{\gamma_n} \sum_{r=1}^{K+N+1} w_r p_i(x_r) p_k(x_r) p_l(x_r) p_n(x_r).$$

Assume the γ_n are known, see for instance example 2 for Hermite and Legendre polynomials. If they are not they can be calculated by Gauss quadrature, too. Hence, we need to calculate all values $p_n(x_r)$ where x_r is a Gauss quadrature node. This is done by using the recursion formula for the orthogonal polynomials 3.2. One sets

$$\begin{aligned} \mathbf{p}_0 &:= (p_0(x_1), \dots, p_0(x_{K+N+1})) := (1, \dots, 1), \\ \mathbf{p}_1 &:= (p_1(x_1), \dots, p_1(x_{K+N+1})) := (x_1, \dots, x_{K+N+1}) = \mathbf{x}^T \end{aligned}$$

and calculates the values of the polynomials p_2, \dots, p_{K+N} by the recursion formula

$$\begin{aligned} \mathbf{p}_{n+1} &= (c_n \mathbf{x}^T \cdot - a_n) \cdot (p_n(x_1), \dots, p_n(x_{K+N+1})) - b_n (p_{n-1}(x_1), \dots, p_{n-1}(x_{K+N+1})) \\ &= (c_n \mathbf{x}^T \cdot - a_n) \cdot \mathbf{p}_n - b_n \mathbf{p}_{n-1}. \end{aligned}$$

Here \cdot means element wise execution of the operator \circ . In Matlab Pseudo code, we have

Algorithm 2 Calculate the Galerkin Multiplication Tensor \mathbf{M}

Input: Truncation numbers K, N of the truncated gPC expansions of V and Σ .

Output: Galerkin multiplication tensor \mathbf{M} .

Compute the $K + N + 1$ -point Gauss quadrature nodes \mathbf{x} and weights \mathbf{w} by program 1.

Calculate the matrix \mathbf{P} with $\mathbf{P}[i, j] = p_i(x_{j+1})$ by the recursion formula of the p_n :

$\mathbf{P}(0, :) = \text{ones}(1, K + N)$;

$\mathbf{P}(1, :) = \mathbf{x}'$;

for $i = 1 : \max(K - 1, N - 1)$ **do**

$\mathbf{P}(i + 1, :) = (c_n * \mathbf{x}' \cdot - a_n) \cdot \mathbf{P}[i, :] - b_n \mathbf{P}[i - 1, :]$;

end for

Calculate \mathbf{M} by Gauss quadrature:

$\mathbf{M} = \text{zeros}(K + 1, K + 1, N + 1, N + 1)$;

for $n = 0 : N$ **do**

Compute γ_n by a given formula or by Gauss quadrature $\mathbf{P}[n, :] \cdot^2 * \mathbf{w}$;

for $i, k = 0 : K, l = 0 : N$ **do**

$\mathbf{M}(i, k, l, n) = (\mathbf{P}(i, :) \cdot \mathbf{P}(k, :) \cdot \mathbf{P}(l, :) \cdot \mathbf{P}(n, :)) * \mathbf{w} / \gamma_n$

end for

end for

Note that if the model is computed for multiple options with the same density μ and truncation numbers N, K , the Galerkin multiplication tensor \mathbf{M} has to be calculated only once, since it only depends on μ, N and K . This would reduce computational cost. The multiplication tensor can also be stored for later usage.

With this, we calculate \mathbf{A} with

$$\mathbf{A}[n, l] := \sum_{i,k=0}^K \sigma_i \sigma_k M_{ikln} = \sigma \mathbf{M}[:, :, l, n] \sigma^T, \quad n, l = 0, \dots, N,$$

where $\sigma := (\sigma_0, \dots, \sigma_K)$ is the row vector of the first $K + 1$ gPC coefficients of Σ and $\mathbf{M}[:, :, l, n]$ stands for the matrix of all elements on \mathbf{M} , whose last two indices are l, n .

Algorithm 3 Compute the coupling matrix A

Input: Truncation numbers K, N of the truncated gPC expansions of V and Σ , vector of gPC coefficients $\sigma := (\sigma_0, \dots, \sigma_K)$ of Σ .

Output: Coupling matrix \mathbf{A} .

Compute \mathbf{M} by program 2.

Calculate \mathbf{A} :

$\mathbf{A} = \text{zeros}(N + 1, N + 1)$;

for $l, n = 0 : N$ **do**

$\mathbf{A}(n, l) = \sigma \mathbf{M}[:, :, l, n] \sigma'$;

end for

5.2 Deriving a finite difference scheme

The PDE 5.3 for the vector of transformed gPC coefficients $\bar{\mathbf{v}} : [0, 1] \times [0, T] \rightarrow \mathbb{R}^{N+1}$

$$\frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} = \frac{1}{2} \zeta^2 (1 - \zeta)^2 \mathbf{A} \frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2} + r \zeta (1 - \zeta) \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} - r (1 - \zeta) \bar{\mathbf{v}}(\zeta, \tau)$$

is now discretised on $M_\zeta + 1$ equidistant values in ζ and $N_\tau + 1$ equidistant values in τ , where $M_\zeta, N_\tau \in \mathbb{N}$ are chosen large enough to represent the solution in a proper way and in the right proportion to obtain a stable scheme, see later. Define

$$\begin{aligned} \zeta_m &:= \frac{m}{M_\zeta} = m \Delta \zeta, \quad m = 0, \dots, M_\zeta \\ \tau^n &:= T \frac{n}{N_\tau} = n \Delta \tau, \quad n = 0, \dots, N_\tau, \end{aligned}$$

and $\Delta\zeta := 1/M_\zeta$, $\Delta\tau := T/N_\tau$. The partial derivatives are approximated component wise by finite differences as in [ZWCS13] Chapter 8.1.1. We use

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

for $m = 1, \dots, M_\zeta - 1$, $n = 0, \dots, N_\tau - 1$. Approximating equation 5.3 in $(\zeta, \tau) = (\zeta_m, \tau^n)$ by these finite differences, we obtain an explicit finite difference scheme

$$\begin{aligned} \bar{v}(\zeta_m, \tau^{n+1}) = & \Delta\tau \left(\frac{1}{2} \zeta_m^2 (1 - \zeta_m)^2 \mathbf{A} \frac{\bar{v}(\zeta_{m+1}, \tau^n) - 2\bar{v}(\zeta_m, \tau^n) + \bar{v}(\zeta_{m-1}, \tau^n)}{(\Delta\zeta)^2} \right. \\ & \left. + r \zeta_m (1 - \zeta_m) \frac{\bar{v}(\zeta_{m+1}, \tau^n) - \bar{v}(\zeta_{m-1}, \tau^n)}{2\Delta\zeta} - r(1 - \zeta_m) \bar{v}(\zeta_m, \tau^n) \right) \\ & + \bar{v}(\zeta_m, \tau^n), \end{aligned} \quad (5.6)$$

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

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

The remaining values for $m \in \{0, M_\zeta\}$, i.e. $\zeta_m \in \{0, 1\}$, are given by the boundary conditions $\bar{v}(0, \tau^n) = \mathbf{0}_{N+1}$ and $\bar{v}(1, \tau^n) = (1, 0, \dots, 0)^T$ for all n . Note that this is an explicit scheme, as the value at time τ^{n+1} is derived only from values at time τ^n .

In the following, we will show some properties of this scheme needed to provide a reasonable approximation to the true solution.

Consistency:

Consistency expresses the fact that the considered scheme represents the original PDE, if the discretization step widths go to zero, i.e. the scheme is a reasonable discretization of the PDE.

Definition 5 ([Str04] Definition 1.4.2). Consider a PDE in the variables x, t of the form $\mathcal{L}(u) = f$ for some differential operator \mathcal{L} and a source term f together with a finite difference scheme $L_{\Delta x, \Delta t}(v) = f$ with discretization step width Δx in x and Δt in t . This scheme is called *consistent*, if for any smooth function $\phi(x, t)$ we have pointwise

$$\mathcal{L}(\phi) - L_{\Delta x, \Delta t}(\phi) \rightarrow 0 \quad \text{as } \Delta x, \Delta t \rightarrow 0.$$

For our scheme, consider a C^2 function $\phi : [0, 1] \times [0, T] \rightarrow \mathbb{R}^{N+1}$. The Taylor formula, see [AE06a] Korollar IV.3.3), gives

$$\begin{aligned}\frac{\partial \phi}{\partial \tau}(\zeta_m, \tau^n) &= \frac{\phi(\zeta_m, \tau^{n+1}) - \phi(\zeta_m, \tau^n)}{\Delta \tau} + \mathbf{o}_\tau(1), \text{ for } \Delta \tau \rightarrow 0, \\ \frac{\partial \phi}{\partial \zeta}(\zeta_m, \tau^n) &= \frac{\phi(\zeta_{m+1}, \tau^n) - \phi(\zeta_{m-1}, \tau^n)}{2\Delta \zeta} + \mathbf{o}_\zeta(\Delta \zeta), \text{ for } \Delta \zeta \rightarrow 0 \quad \text{and} \\ \frac{\partial^2 \phi}{\partial \zeta^2}(\zeta_m, \tau^n) &= \frac{\phi(\zeta_{m+1}, \tau^n) - 2\phi(\zeta_m, \tau^n) + \phi(\zeta_{m-1}, \tau^n)}{(\Delta \zeta)^2} + \mathbf{o}_\zeta(1), \text{ for } \Delta \zeta \rightarrow 0.\end{aligned}$$

Now, one brings all terms in the PDE 5.3 and the finite difference scheme 5.6 for ϕ to the left hand side and obtains the form $L(\phi) = \mathbf{0}_{N+1}$ for both equations. Subtracting the left hand sides evaluated in the points (ζ_m, τ^n) yields

$$\begin{aligned}& \frac{\partial \phi(\zeta_m, \tau^n)}{\partial \tau} - \frac{1}{2}\zeta_m^2(1 - \zeta_m)^2 \mathbf{A} \frac{\partial^2 \phi(\zeta_m, \tau^n)}{\partial \zeta^2} - r\zeta_m(1 - \zeta_m) \frac{\partial \phi(\zeta_m, \tau^n)}{\partial \zeta} + r(1 - \zeta_m)\phi(\zeta_m, \tau^n) \\ & - \left(\frac{\phi(\zeta_m, \tau^{n+1}) - \phi(\zeta_m, \tau^n)}{\Delta \tau} - \frac{1}{2}\zeta_m^2(1 - \zeta_m)^2 \mathbf{A} \frac{\phi(\zeta_{m+1}, \tau^n) - 2\phi(\zeta_m, \tau^n) + \phi(\zeta_{m-1}, \tau^n)}{(\Delta \zeta)^2} \right. \\ & \left. - r\zeta_m(1 - \zeta_m) \frac{\phi(\zeta_{m+1}, \tau^n) - \phi(\zeta_{m-1}, \tau^n)}{2\Delta \zeta} + r(1 - \zeta_m)\phi(\zeta_m, \tau^n) \right) \\ & = \left(\frac{\partial \phi(\zeta_m, \tau^n)}{\partial \tau} - \frac{\phi(\zeta_m, \tau^{n+1}) - \phi(\zeta_m, \tau^n)}{\Delta \tau} \right) \\ & - \frac{1}{2}\zeta_m^2(1 - \zeta_m)^2 \mathbf{A} \left(\frac{\partial^2 \phi(\zeta_m, \tau^n)}{\partial \zeta^2} - \frac{\phi(\zeta_{m+1}, \tau^n) - 2\phi(\zeta_m, \tau^n) + \phi(\zeta_{m-1}, \tau^n)}{(\Delta \zeta)^2} \right) \\ & - r\zeta_m(1 - \zeta_m) \left(\frac{\partial \phi(\zeta_m, \tau^n)}{\partial \zeta} - \frac{\phi(\zeta_{m+1}, \tau^n) - \phi(\zeta_{m-1}, \tau^n)}{2\Delta \zeta} \right) \\ & = \mathbf{o}_\tau(1) - \frac{1}{2}\zeta_m^2(1 - \zeta_m)^2 \mathbf{A} \mathbf{o}_\zeta(1) - r\zeta_m(1 - \zeta_m) \mathbf{o}_\zeta(\Delta \zeta).\end{aligned}$$

Because of $\zeta_m \in [0, 1]$, the coefficients $\zeta_m^2(1 - \zeta_m)^2$ and $\zeta_m(1 - \zeta_m)$ are bounded and therefore the difference goes to zero as $\Delta \zeta$ and $\Delta \tau$ do. This shows consistency.

Stability:

Stability tells us, that small perturbation in the initial data cause small errors in the solution, see [TW98] chapter 1.2.1. We require this for our computation to be useful, since e.g. rounding errors can introduce small errors at any time step that could destroy the computation if stability is not given.

Definition 6 (adapted from [Str04] Definition 1.5.1). Consider a one step finite difference scheme $\mathbf{u}(x_m, t^{n+1}) = f_{\Delta x, \Delta t}(\mathbf{u}(x_0, t^n), \dots, \mathbf{u}(x_M, t^n))$ in N dimensions with discretization step width Δx in x and Δt in t . Define $\tilde{\mathbf{u}}^n := (\mathbf{u}^T(x_0, t^n), \dots, \mathbf{u}^T(x_M, t^n))^T \in \mathbb{R}^{(M+1)N}$ for all n . The finite difference scheme is said to be *stable*, if there exists an integer J such that for some norm $\|\cdot\|$ for any $T > 0$, a constant C_T exists with

$$\|\tilde{\mathbf{u}}^n\|^2 \leq C_T \sum_{j=0}^J \|\tilde{\mathbf{u}}^j\|^2 \quad \text{for every } n \text{ with } 0 \leq n\Delta t \leq T.$$

In the following, we find a sufficient condition for our scheme 5.6 to be stable. At first we rewrite our scheme for $\tilde{\mathbf{v}}^n$ in the notation of definition 6 and obtain

$$\begin{aligned}
\tilde{\mathbf{v}}^{n+1} &= \begin{pmatrix} \mathbf{v}(\zeta_0, \tau^{n+1}) \\ \mathbf{v}(\zeta_1, \tau^{n+1}) \\ \vdots \\ \mathbf{v}(\zeta_{M_\zeta-1}, \tau^{n+1}) \\ \mathbf{v}(\zeta_{M_\zeta}, \tau^{n+1}) \end{pmatrix} \tag{5.7} \\
&= \left[\mathbf{I}_{(M+1)(N+1)} + \Delta\tau \left\{ \frac{1}{2(\Delta\zeta)^2} \begin{pmatrix} \mathbf{A} & & & & \\ & \mathbf{A} & & & \\ & & \ddots & & \\ & & & \mathbf{A} & \\ & & & & \mathbf{A} \end{pmatrix} \right. \right. \\
&\quad \cdot \begin{pmatrix} \eta_0^2 \mathbf{I} & & & & \\ & \eta_1^2 \mathbf{I} & & & \\ & & \ddots & & \\ & & & \eta_{M_\zeta-1}^2 \mathbf{I} & \\ & & & & \eta_{M_\zeta}^2 \mathbf{I} \end{pmatrix} \begin{pmatrix} -2\mathbf{I} & \mathbf{I} & & & \\ \mathbf{I} & -2\mathbf{I} & \mathbf{I} & & \\ & & \ddots & \ddots & \ddots \\ & & & \mathbf{I} & -2\mathbf{I} & \mathbf{I} \\ & & & & \mathbf{I} & -2\mathbf{I} \end{pmatrix} \\
&\quad + \frac{r}{2\Delta\zeta} \begin{pmatrix} \eta_0 \mathbf{I} & & & & \\ & \eta_1 \mathbf{I} & & & \\ & & \ddots & & \\ & & & \eta_{M_\zeta-1} \mathbf{I} & \\ & & & & \eta_{M_\zeta} \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{0} & -\mathbf{I} & & & \\ \mathbf{I} & \mathbf{0} & -\mathbf{I} & & \\ & & \ddots & \ddots & \ddots \\ & & & \mathbf{I} & \mathbf{0} & -\mathbf{I} \\ & & & & \mathbf{I} & \mathbf{0} \end{pmatrix} \\
&\quad \left. - r \begin{pmatrix} (1-\zeta_0)\mathbf{I} & & & & \\ & (1-\zeta_1)\mathbf{I} & & & \\ & & \ddots & & \\ & & & (1-\zeta_{M_\zeta-1})\mathbf{I} & \\ & & & & (1-\zeta_{M_\zeta})\mathbf{I} \end{pmatrix} \right\} \begin{pmatrix} \mathbf{v}(\zeta_0, \tau^n) \\ \mathbf{v}(\zeta_1, \tau^n) \\ \vdots \\ \mathbf{v}(\zeta_{M_\zeta-1}, \tau^n) \\ \mathbf{v}(\zeta_{M_\zeta}, \tau^n) \end{pmatrix} \\
&=: \mathbf{C}\tilde{\mathbf{v}}^n
\end{aligned}$$

where $\eta_m := \zeta_m(1-\zeta_m)$ and $\mathbf{I} := \mathbf{I}_{(N+1)}$ is the unit matrix and $\mathbf{0} := \mathbf{0}_{(N+1)\times(N+1)}$ the zero matrix in $\mathbb{R}^{(N+1)\times(N+1)}$ and $\mathbf{C} \in \mathbb{R}^{(N+1)(M_\zeta+1)\times(N+1)(M_\zeta+1)}$ denotes the matrix in square brackets.

With the above recursion formula we have $\tilde{\mathbf{v}}^n = \mathbf{C}^n \tilde{\mathbf{v}}^0$ and it follows for any vector norm and its corresponding operator norm

$$\|\tilde{\mathbf{v}}^n\|^2 = \|\mathbf{C}^n \tilde{\mathbf{v}}^0\|^2 \leq \|\mathbf{C}^n\|^2 \|\tilde{\mathbf{v}}^0\|^2.$$

Therefore, stability is given if $\|\mathbf{C}^n\|^2$ is bounded for $n \rightarrow \infty$.

The Kreiss matrix theorem, to be found e.g. in [Str04] Theorem 9.4.1, tells us that this is the case if and only if \mathbf{C} is unitary trigonalisable, i.e. a unitary matrix \mathbf{U} and a

upper tridiagonal matrix \mathbf{S} exist with $\mathbf{S} = \mathbf{UCU}^{-1}$, and the entries of \mathbf{S} satisfy

$$|\mathbf{S}[i, i]| \leq 1 \quad \text{and} \quad (5.8)$$

$$|\mathbf{S}[i, j]| \leq C_b \min(1 - |\mathbf{S}[i, i]|, 1 - |\mathbf{S}[j, j]|) \quad (5.9)$$

for some boundary constant C_b for all i, j . Schur's theorem, see [HJ13] Theorem 2.3.1, shows that every matrix is unitary trigonalisable over \mathbb{C} . Furthermore the diagonal entries of \mathbf{S} are the eigenvalues of \mathbf{C} , therefore \mathbf{C} satisfies the first condition 5.8 if and only if all eigenvalues λ_k of \mathbf{C} satisfy $|\lambda_k| \leq 1$.

Since we want to determine the maximal $\Delta\tau$ for which the scheme is stable for a given M_ζ , we reformulate the condition $|\lambda_k| \leq 1$ such that it is a condition for $\Delta\tau$. To do this, we write $\mathbf{C} = \mathbf{I}_{(M+1)(N+1)} + \Delta\tau\mathbf{B}$, where \mathbf{B} is the matrix in curly brackets in equation 5.7. Note that \mathbf{B} is independent of $\Delta\tau$ and that the eigenvalues λ_k of \mathbf{C} satisfy $\lambda_k = 1 + \Delta\tau\nu_k$, where ν_k denotes an eigenvalue of \mathbf{B} . I.e. $|\lambda_k| \leq 1$ is equivalent to

$$0 \leq \nu_k \text{ and } 0 \leq \Delta\tau \leq \frac{-2}{\nu_k} \quad \text{for all } \nu_k \text{ eigenvalue of } \mathbf{B}.$$

Where division by zero in the second condition is understood as $-2/0 = -2/0^- = \infty$.

If condition 5.8 is given, condition 5.9 is equivalent to

$$|\mathbf{S}[\tilde{i}, j]| = |\mathbf{S}[j, \tilde{i}]| = 0 \quad (5.10)$$

for all j , if $|\mathbf{S}[\tilde{i}, \tilde{i}]| = 1$. The direction 5.9 \Rightarrow 5.10 is clear, whereas if condition 5.10 holds, one can choose C_b large enough to fulfil condition 5.9:

$$C_b \geq \max_{i, j \leq i \text{ with } |\mathbf{S}[i, i]|, |\mathbf{S}[j, j]| < 1} \frac{|\mathbf{S}[i, j]|}{\min(1 - |\mathbf{S}[i, i]|, 1 - |\mathbf{S}[j, j]|)}.$$

In summary, the scheme is stable, if $\|\mathbf{C}^n\|^2$ is bounded which is equivalent to

$$\begin{aligned} 0 \leq \nu_k \text{ and } 0 \leq \Delta\tau \leq \frac{-2}{\nu_k}, \quad & \text{for all } \nu_k \text{ eigenvalue of } \mathbf{B} \quad (5.11) \\ |\mathbf{S}[i, i]| = 1 \Rightarrow |\mathbf{S}[i, j]| = |\mathbf{S}[j, i]| = 0, \quad & \mathbf{S} = \mathbf{UCU}^{-1} \text{ upper trigonal with } \mathbf{U} \text{ unitary.} \end{aligned}$$

This has to be checked before the computation of the finite difference scheme, for instance by a program like the following given in Matlab notation.

Algorithm 4 Test for Stability

Input: Interest rate r , coupling matrix \mathbf{A} , number of discretization points M_ζ in ζ .

Output: Boolean variable *stable* to determine if stability is possible, *dtaumax* maximum $\Delta\tau$ allowed for stability to hold.

Calculate the matrix \mathbf{B} from equation 5.7.

```
stable = 1;
S = schur(B);
```

Condition 2: Save the positions of diagonal entries equal to 1 in **diag1**, set them to 0 in a copy of **S**:

```
diag1 = (diag(S) == 1);
```

```
S2 = S;
```

```
S2(diag1, diag1) = 0;
```

Select all entries in the same column (**Scol**) or row (**Srow**) as a diagonal entry 1:

```
Scol = S2(:, diag1);
```

```
Srow = S2(diag1, :);
```

```
S3 = [Scol Srow'];
```

Test them to be numerically equal to zero, i.e. in absolute value smaller than a maximum value *maxval* > 0 reflecting the computer precision:

```
if max(max(abs(S3))) > maxval then
```

```
    stable = 0;
```

```
    dtaumax = 0;
```

```
end if
```

Condition 1: Test the eigenvalues **EV** of **B** to be non-positive for stability, calculate *dtaumax* in that case:

```
EV = diag(S);
```

```
if sum((EV > 0)) > 0 then
```

```
    stable = 0;
```

```
    dtaumax = 0;
```

```
else if sum((EV ≠ 0)) == 0 then
```

```
    dtaumax = 1;
```

```
else
```

```
    dtaumax = 2/max(abs(EV(EV = 0)));
```

```
end if
```

Convergence:

Convergence of a finite difference scheme expresses the fact, that the finite difference solution converges to the solution of the approximated PDE as the discretization step size goes to zero, i.e. the scheme provides an approximate solution to the PDE.

Definition 7. A one step finite difference scheme for a PDE is called *convergent* if for any solution $u(x, t)$ of the PDE it holds for all solutions $w(x_m, t^n)$ of the finite difference

scheme for which the initial condition $w_0(x_m) \rightarrow u_0(x)$ as $x_m \rightarrow x$, one has

$$w(x_m, t^n) \rightarrow u(x, t) \quad \text{for } (x_m, t^n) \xrightarrow{\Delta x, \Delta t \rightarrow 0} (x, t).$$

The Lax-Richtmyer Equivalence Theorem provides a useful tool to investigate the convergence of a finite difference scheme.

Theorem 8 (Lax-Richtmyer Equivalence Theorem, see [Str04] Theorem 1.5.1). *A consistent finite difference scheme for a well posed PDE is convergent if and only if it is stable.*

A proof can be found in [Str04] Chapter 10.5 for instance.

With this, our derived finite difference scheme 5.6 is convergent if it is stable and equation 5.3 is parabolic. Hence, to ensure a reasonable computation, the stability condition 5.11 and the positivity of the real parts of all eigenvalues of \mathbf{A} have to be checked.

After a successful computation of $\bar{\mathbf{v}}$, one transforms $\bar{\mathbf{v}}$ back to \mathbf{v} by

$$\begin{aligned} S &= \text{strike} \frac{\zeta}{1 - \zeta}, \\ t &= T - \tau, \\ \mathbf{v}(S, t) &= (S + \text{strike}) \bar{\mathbf{v}}(\zeta, \tau) = \text{strike} \frac{1}{1 - \zeta} \bar{\mathbf{v}}(\zeta, \tau). \end{aligned} \tag{5.2}$$

In total, the pseudo code for the computation of \mathbf{v} given in Matlab notation is

Algorithm 5 Computing the truncated gPC coefficients \mathbf{v} of V

Input: Truncation numbers K, N of the truncated gPC expansions of V and Σ , vector of gPC coefficients $\sigma := (\sigma_0, \dots, \sigma_K)$ of Σ , interest rate r , maturity T of the option, its strike price *strike* and number of discretization points M_ζ in ζ .

Output: Tensor \mathbf{v} with $\mathbf{v}(k, m, n) = v_k(S_m, t^n)$, $k = 0, \dots, N$, $m = 0, \dots, M_\zeta - 1$, $n = 0, \dots, N_\tau$, vectors $\mathbf{S} := (S_0, \dots, S_{M_\zeta-1})$ and $\mathbf{t} = (t^0, \dots, t^{N_\tau})$ where S_m, t^n are the transformed back ζ_m, τ^n .

Calculate the coupling matrix \mathbf{A} by program 3.

Check for parabolicity:

```
if min(real(eig(A))) ≤ 0 then
    error('Error: Equation not parabolic');
end if
```

Code continues on the next page.

Check for stability by program 4 and obtain the output variables *stable* and *dtaumax*
if *stable* == 0 **then**

 error('Error: Scheme not stable');

end if

Calculate the minimum N_τ to have stability:

$N_\tau = \text{ceil}(T/dtaumax)$;

$\bar{\mathbf{v}} = \text{zeros}(N + 1, M_\zeta + 1, N_\tau + 1)$;

Setting the initial condition for $\bar{\mathbf{v}}$:

$\bar{\mathbf{v}}(0, :, 0) = \max(2/M_\zeta * (0 : M_\zeta) - 1, 0)$;

Calculating $\bar{\mathbf{v}}$ for the remaining times by scheme 5.6:

for $n = 0 : N_\tau - 1$ **do**

 Boundary conditions:

$m = 0$;

$\bar{\mathbf{v}}(:, m, n + 1) = \text{zeros}(N + 1, 1)$;

$m = M_\zeta$;

$\bar{\mathbf{v}}(:, m, n + 1) = \text{zeros}(N + 1, 1)$;

$\bar{\mathbf{v}}(0, m, n + 1) = 1$;

 Scheme:

for $m = 1 : M_\zeta - 1$ **do**

$dvdzeta = M_\zeta/2 * (\bar{\mathbf{v}}(:, m + 1, n) - \bar{\mathbf{v}}(:, m - 1, n))$;

$d2vdzeta2 = M_\zeta^2 * (\bar{\mathbf{v}}(:, m + 1, n) - 2 * \bar{\mathbf{v}}(:, m, n) + \bar{\mathbf{v}}(:, m - 1, n))$;

$\bar{\mathbf{v}}(:, m, n + 1) = T/N_\tau * (1/2 * (m/M_\zeta)^2 * (M_\zeta - m)^2 / M_\zeta^2 * A * d2vdzeta2 + r * m / M_\zeta * (M_\zeta - m) / M_\zeta * dvdzeta - r * (M_\zeta - m) / M_\zeta * \bar{\mathbf{v}}(:, m, n)) + \bar{\mathbf{v}}(:, m, n)$;

end for

end for

Transforming back $\bar{\mathbf{v}}$ to \mathbf{v} , calculating \mathbf{S} and \mathbf{t} :

$\mathbf{v} = \text{zeros}(N + 1, M_\zeta, N_\tau + 1)$;

for $n = 0 : N_\tau$ **do**

$\mathbf{v}(:, :, n) = \text{strike} * (1 : -1/M_\zeta : 1/M_\zeta).^(-1) * \bar{\mathbf{v}}(:, 1 : M_\zeta, N_\tau - n)$;

end for

$\mathbf{S} = \text{strike} * (0 : M_\zeta - 1) ./ (M_\zeta : -1 : 1)$;

$\mathbf{t} = 0 : T/N_\tau : T$;

Chapter 6

Numerical results

This chapter shows numerical results obtained by the programs of last chapter for the truncated system of equations 4.7. For our numerical experiments, we will always consider European Call options. For more convenient reading, the maturity T and times t will be given in days instead of years as needed for the computation. In the computation, these times were multiplied by $1/251$ since there were 251 trade days in 2019.

In the following, a European call option with strike price 100 and maturity $T = 20$ days is considered. Note that this corresponds to a period of slightly less than one month in real time, since financial markets are usually closed on weekends and holidays. The interest rate in the market is assumed to be $r = 0$ and the volatility of the stochastic asset is supposed to follow a normal distribution with mean $\sigma_0 = 0.3$ and standard deviation $\sigma_1 = 0.1$, i.e. we have the gPC expansion

$$\Sigma(\Theta) = 0.3 + 0.1\Theta,$$

where Θ is standard normal distributed with density function $\mu(x) = 1/\sqrt{2\pi}e^{-x^2/2}$. Therefore, Hermite polynomials are used as orthogonal polynomials, see example 2. The Stochastic Galerkin (SG) solution was truncated after degree $N = 5$ for which the system of equations for the gPC coefficients 5.3 is parabolic. For the calculation a grid with $M_\zeta + 1 = 200 + 1$ values in ζ and $N_\tau + 1 = 84 + 1$ in τ was used. This was chosen such that the finite difference scheme is stable.

The expected value surface of the Stochastic Galerkin solution is displayed in figure 6.1, where contour lines were drawn in black at each quarter of the maximum value. The 'smoothing area' denotes the area around $S = 100$, where the expected value of the SG solution in $t < T$ differs from the final condition of the option, since the final condition is smoothed by the parabolicity of PDE 4.7. It is visualized by drawing its borders in red. Figure 6.2 shows the expected value in the times $t = 0, t = 5, t = 10, t = 15$ and $t = T = 20$. One recognizes that the value in $t = T = 20$ coincides with the output of the option that is given by $(S - 100)^+$ which is the final condition. The non differentiable point at $S = \textit{strike} = 100, t = T = 20$ gets smoothed as t moves away. This is what one would expect from a parabolic PDE. For small $S < 100$ the solution is approximately

zero and for big $S > 100$ it approximately attains the value $S - strike$. These properties are consistent with the boundary conditions.

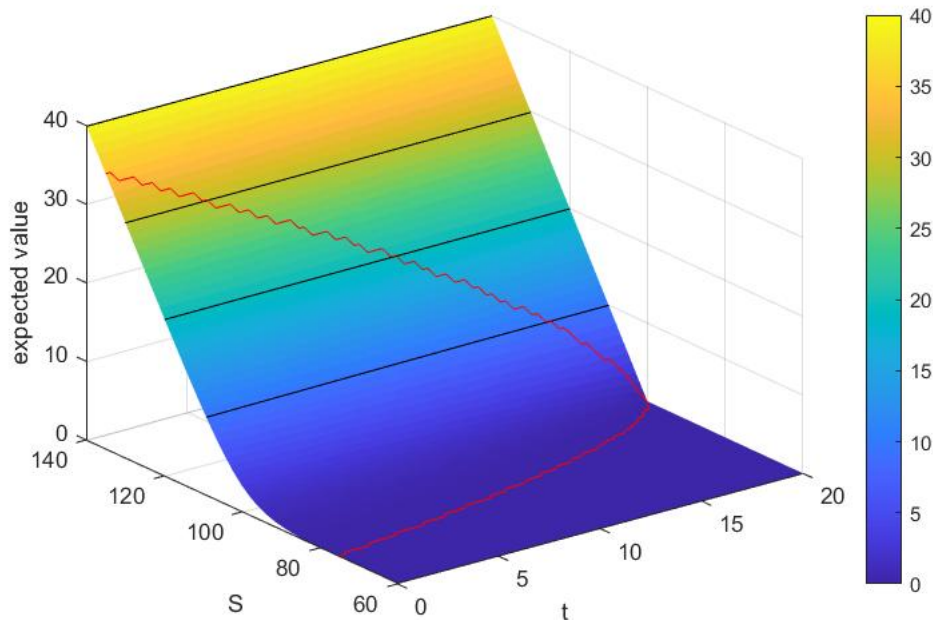


Figure 6.1: Expected value surface for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

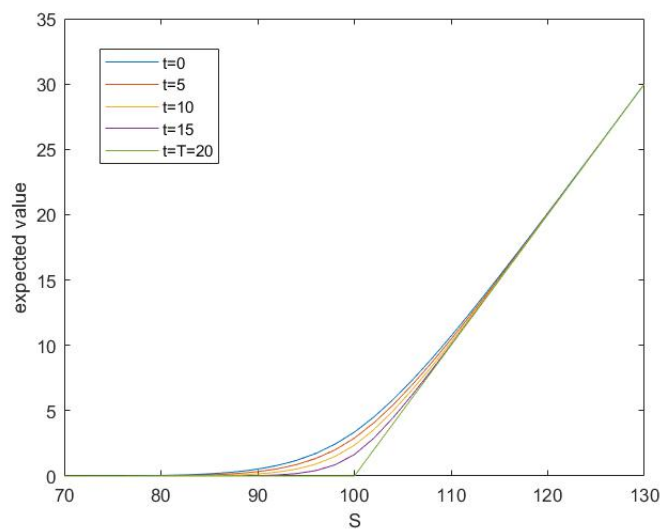


Figure 6.2: Expected value at some times t for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$.

Note that the expected value surface is similar to the solution of the deterministic Black Scholes equation for $\sigma = 0.3 = E(\Sigma(\Theta))$ in figure 6.3, however a difference between the expected value of the SG solution and the deterministic solution exists as shown in figure 6.4. This difference between expected value of the SG solution and the deterministic solution is due to the coupling term occurring in the PDE 4.4 for the 0th gPC coefficient which equals the expected value of the SG solution, see equation 4.11.

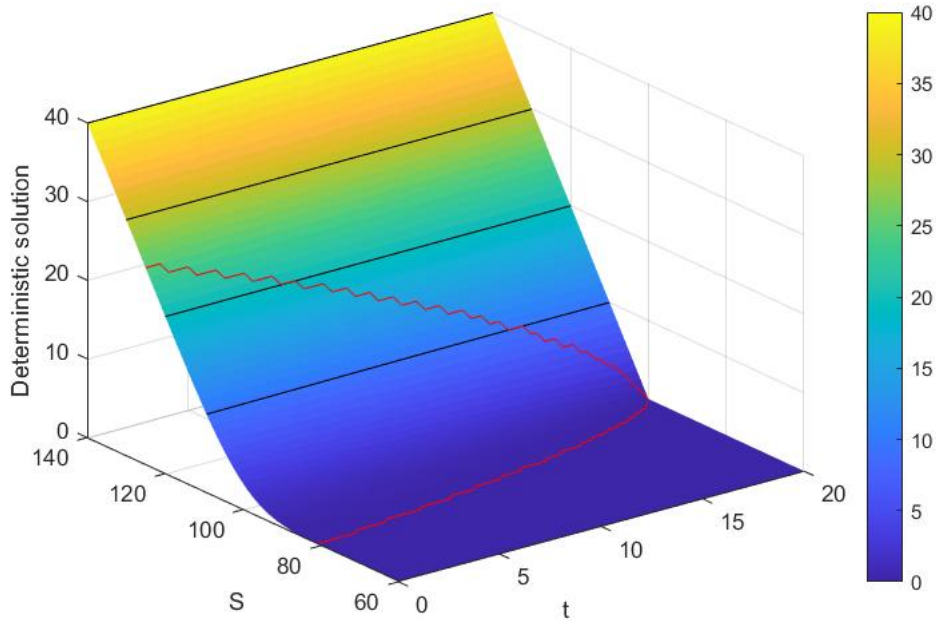


Figure 6.3: Solution of the deterministic Black Scholes equation for a European Call option with $\sigma = 0.3, T = 20, strike = 100$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

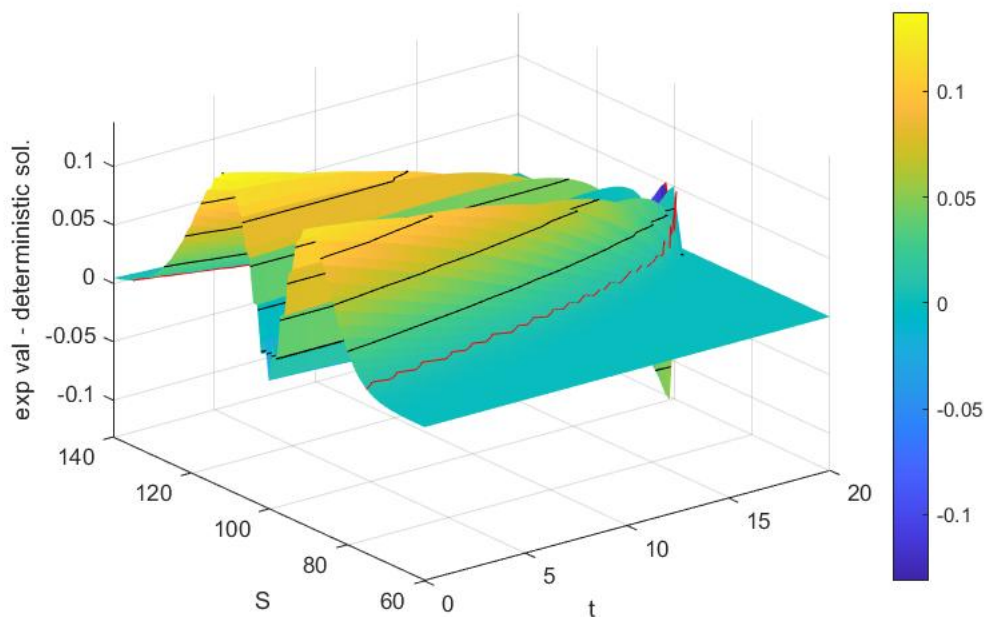


Figure 6.4: Absolute difference of the expected values of the SG solution and the deterministic solution for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed and $\sigma = 0.3$ respectively, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area of the stochastic solution circled in red.

The variance of the SG solution is displayed in figure 6.5 as a surface and figure 6.6 in the times $t = 0, t = 5, t = 10, t = 15$ and $t = T = 20$. Observe that it decreases as $t \rightarrow T$ and vanishes in $t = T$. This is a desired behaviour since the value at maturity is the final condition and therefore deterministic.

The angular shape and the slight asymmetry w.r.t. $S = 100$ is due to the interpolation procedure used to adapt the values in S direction for reasonable plotting, since the original coordinates in S are not linear but determined by the transformation of variables 5.2 applied to a equidistant grid in ζ .

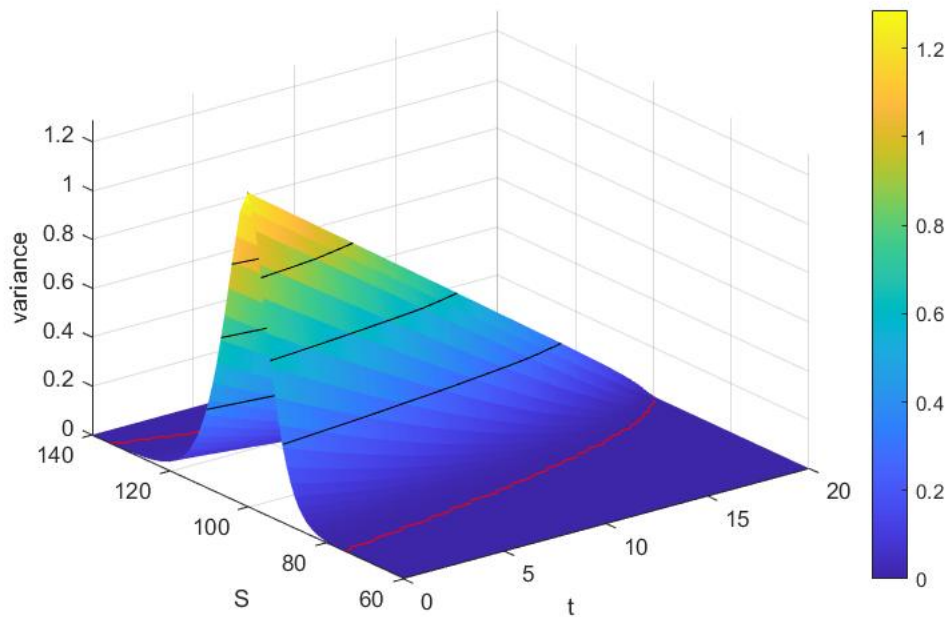


Figure 6.5: Variance surface for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area circled in red.

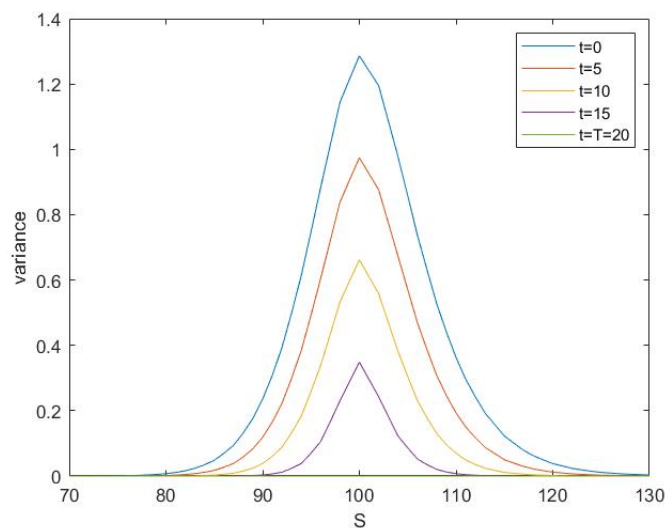


Figure 6.6: Variance at some times t for a European Call option with $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ normal distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$.

For comparison, the same model will be considered for a uniformly on $[-0.5, 0.5]$ distributed random variable Δ . The corresponding orthogonal polynomials are Legendre polynomials from example 2. One can easily check that the variance of Δ is $1/12$. Thus,

to obtain similar conditions, we exchange the model for the volatility by $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$ to obtain a distribution of the volatility that has the same mean and variance as for the standard normal distributed Θ . All the other parameters stay the same.

The expected value of the SG solution in this model is shown in figure 6.7. It looks very similar to the one for the normal distributed Θ in figure 6.1. The comparison of the expected values for both models in $t = 0$ in figure 6.8 and the plot of the absolute difference of the expected values in figure 6.9 confirm that there is practically no difference in the expected values. This leads to the experimental conclusion that for the expected value of the solution, the expected value of the volatility is most important and its actual distribution is rather irrelevant.

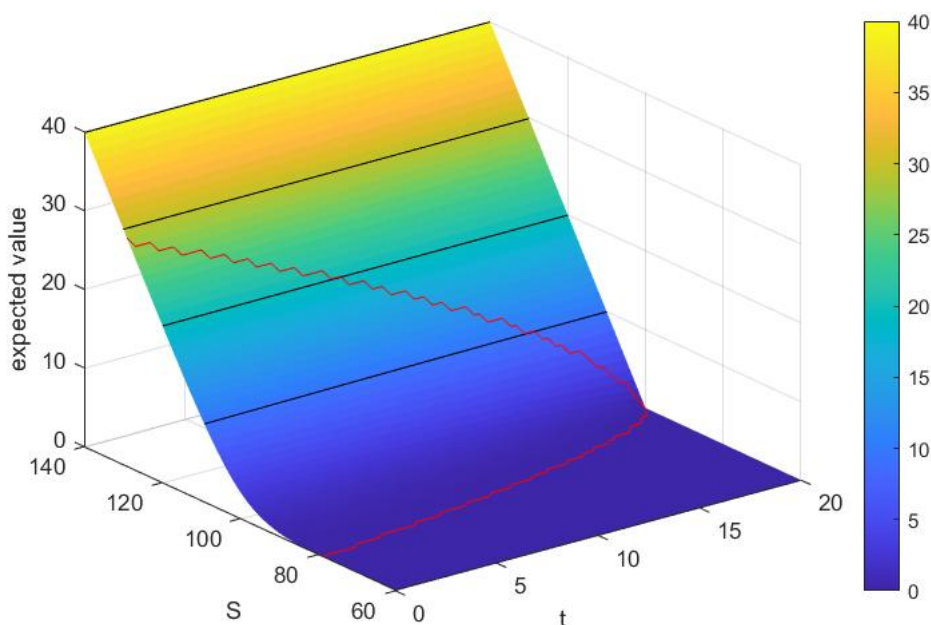


Figure 6.7: Expected value surface for a European Call option with $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

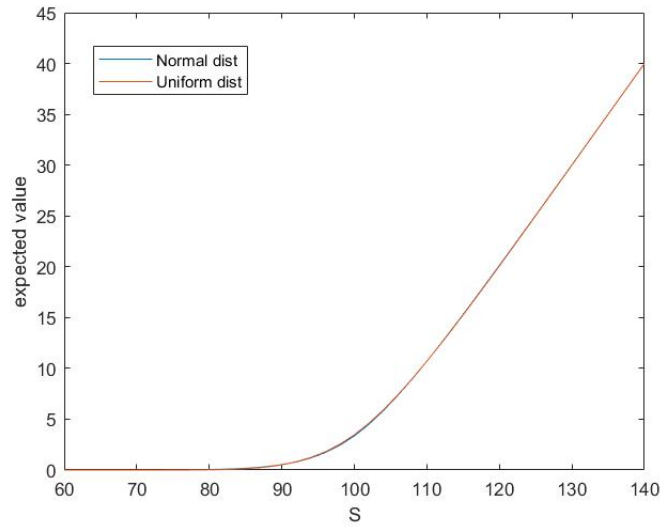


Figure 6.8: Expected values at $t = 0$ for European Call option in the models $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, and $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$.

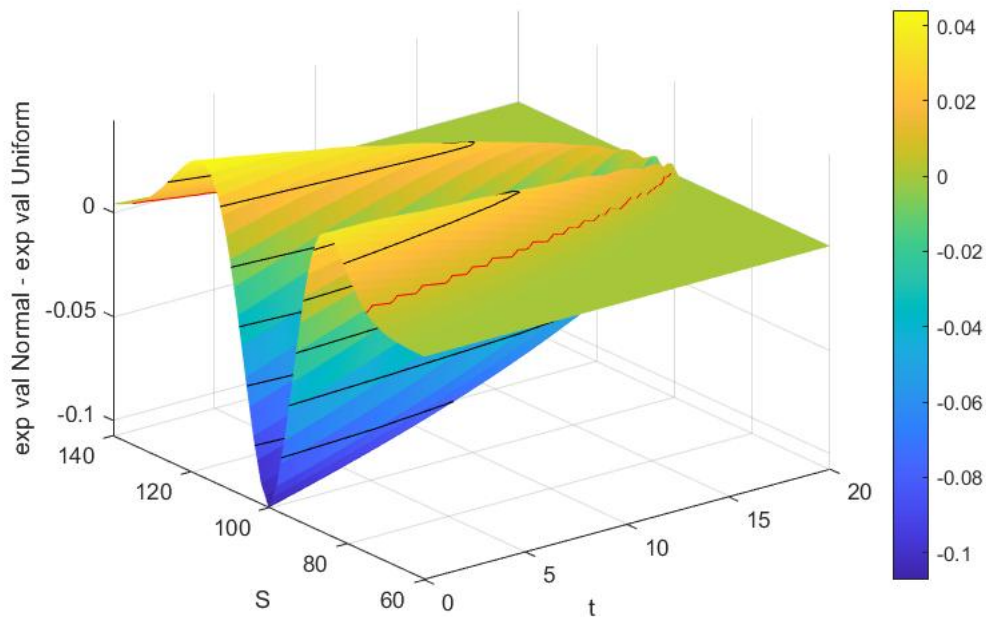


Figure 6.9: Absolute difference in expected values of the models $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, and $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum absolute value and the smoothing area for normal $\Sigma(\Theta)$ circled in red.

Figure 6.10 displays the variance of the SG solution where the volatility is modelled by a uniform distribution. Its shape is similar to the variance for the normal distributed volatility in figure 6.5, but the maximum attained value of the variance is approximately 0.6 whereas solution for the normal distributed $\Sigma(\Theta)$ attains a variance of roughly 1.3. This difference is evident when looking at the plot of variances of both models at time $t = 0$ in figure 6.11 or plotting the absolute difference of both variances for all $S \in [60, 140]$ and $t \in [0, T]$ in figure 6.12.

The reason for the lower variance in case of a uniformly distributed volatility could be that the uniform distribution generates bounded realizations, i.e. every realization of Δ lies in $[-0.5, 0.5]$ whereas the realizations of Θ spread over the whole real numbers. To be precise, the probability that Θ attains a value outside of $[-0.5, 0.5]$ is given by

$$1 - \int_{-0.5}^{0.5} \frac{1}{2\pi} e^{-\frac{x^2}{2}} dx \approx 1 - 0.3829 = 0.6170.$$

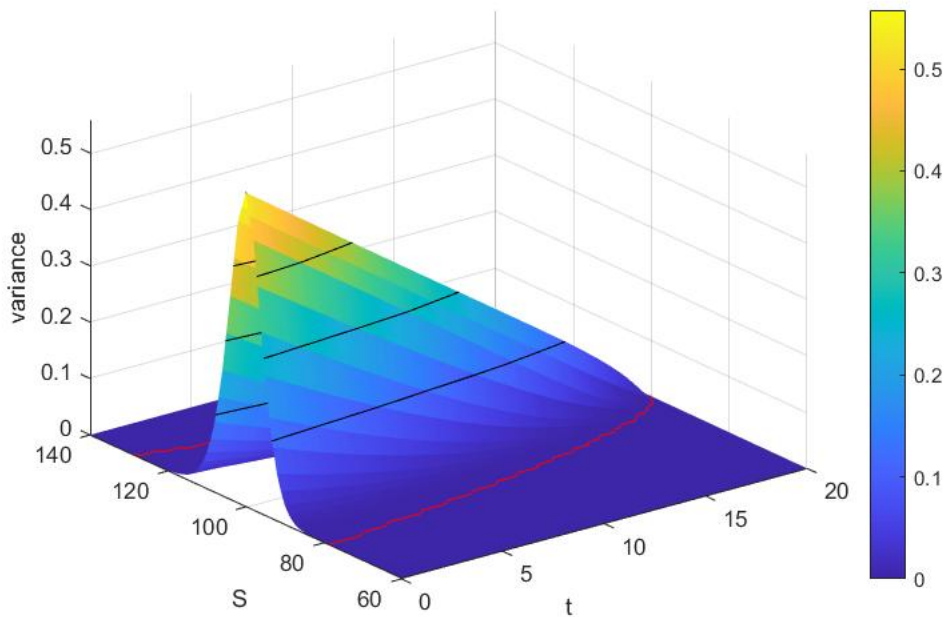


Figure 6.10: Variance surface for a European Call option with $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, $T = 20$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area circled in red.

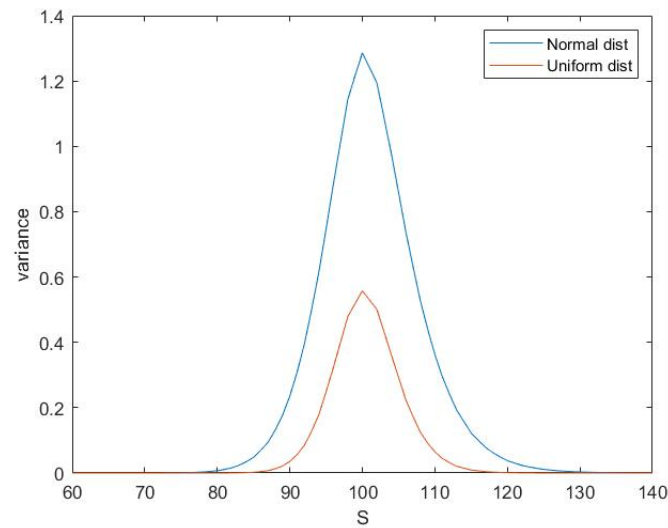


Figure 6.11: Variances at $t = 0$ for European Call option in the models $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-1/2, 1/2]$ distributed, and $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$.

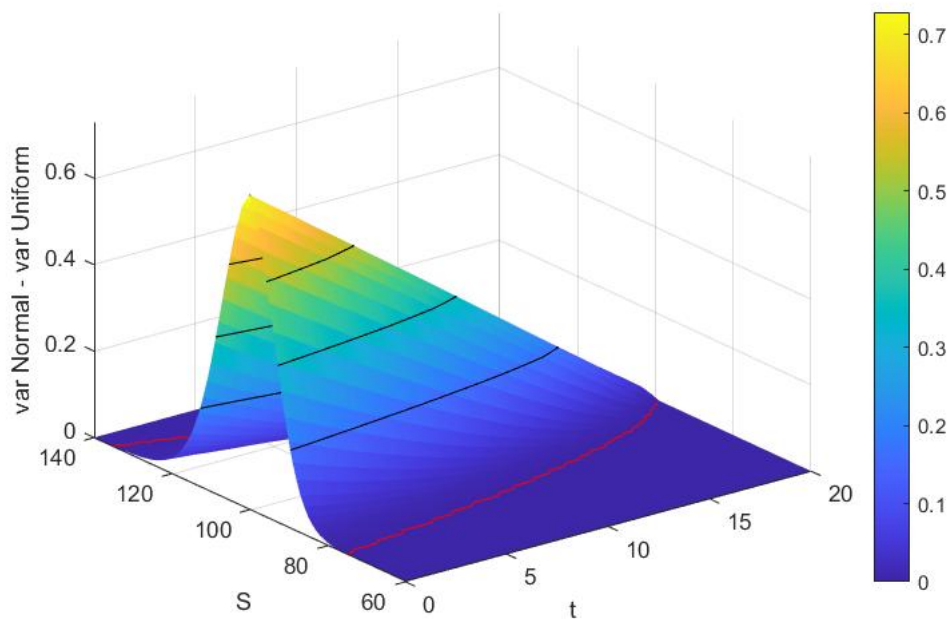


Figure 6.12: Absolute difference of variances of the models for $\Sigma(\Theta) = 0.3 + 0.1\Theta$, Θ standard normal distributed, and $\Sigma(\Delta) = 0.3 + 0.1\sqrt{12}\Delta$, Δ uniformly on $[-0.5, 0.5]$ distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ with contour lines at quarters of its maximum value and the smoothing area for normal $\Sigma(\Theta)$ circled in red.

Comparing the expected value and variance for different volatility models of the form $\Sigma(\Theta) = \sigma_0 + \sigma_1\Theta$:

From now on, let Θ be standard normal distributed again. In the following, the influence of σ_0 and σ_1 will be determined by experiments.

Therefore, the same European Call option with strike price $strike = 100$ and maturity $T = 20$ in the same market with interest rate $r = 0$ will be considered for three different volatility models

$$\begin{aligned}\Sigma_1(\Theta) &= 0.3 + 0.1\Theta, \\ \Sigma_2(\Theta) &= 0.5 + 0.1\Theta \quad \text{and} \\ \Sigma_3(\Theta) &= 0.3 + 0.2\Theta.\end{aligned}$$

The truncation numbers $K = 1, N = 5$ were chosen and a grid with $M_\zeta + 1 = 200 + 1$ points in ζ direction and $N_\tau + 1 = 400 + 1$ in τ direction was used for all models. This framework guarantees stability of the scheme and parabolicity of the system of equations 4.8 for every volatility model Σ_1, Σ_2 and Σ_3 , such that the system can be solved numerically in a proper way.

Figure 6.13 shows the expected values of the corresponding SG solutions at $t = 0$. As one can see, the expected values for Σ_1 and Σ_3 nearly coincide, whereas the one corresponding to Σ_2 significantly differs. The smoothing area is larger for Σ_2 as one would expect in the deterministic model for a larger value of σ . This leads to the conclusion that for the expected value, almost exclusively the 0th gPC coefficient σ_0 matters. This coincides with the observation from comparing the models for normal and uniformly distributed volatilities, as σ_0 is the mean of the volatility. The absolute difference between the expected values in the models Σ_1 and Σ_2 is shown in figure 6.14, where one can see that mainly lies in the smoothing area.

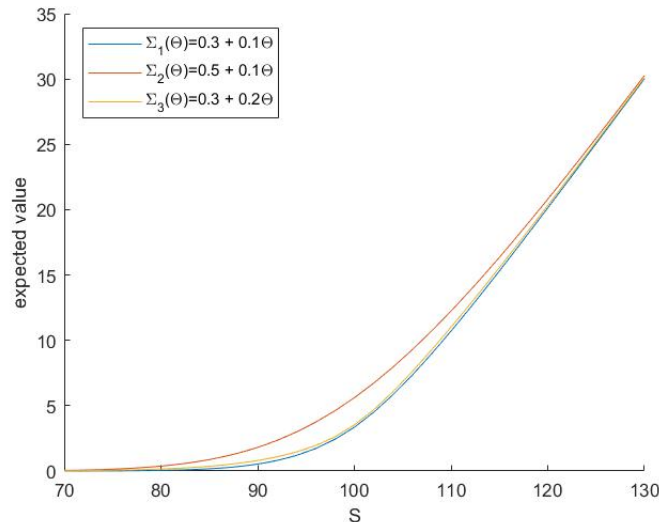


Figure 6.13: Expected values at $t = 0$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$.

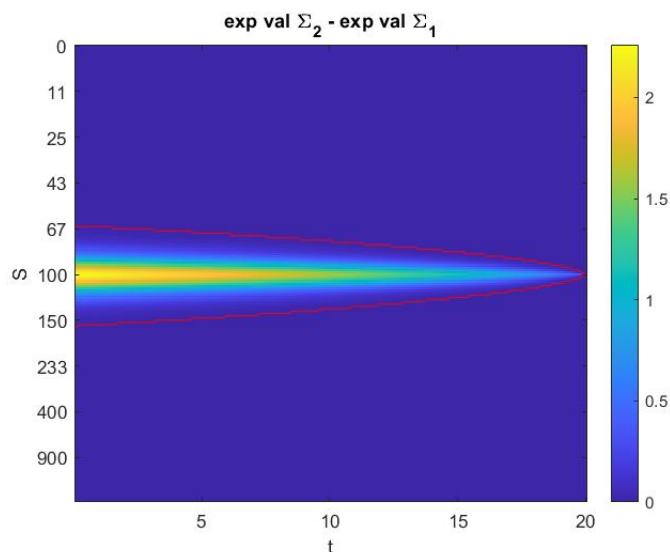


Figure 6.14: Absolute difference in the expected values for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_2 is circled in red.

However, figure 6.15, that displays the expected values at $S = strike = 100$, illustrates that there also is a small difference between the expected value for Σ_1 and Σ_3 , which is caused by the different values in σ_1 . The absolute difference of the solution's expected values for Σ_1 and Σ_3 is displayed in figure 6.16 for all $S \in [60, 140]$ and $t \in [0, T]$ showing that it mainly lies in the neighbourhood of $S = strike = 100$.

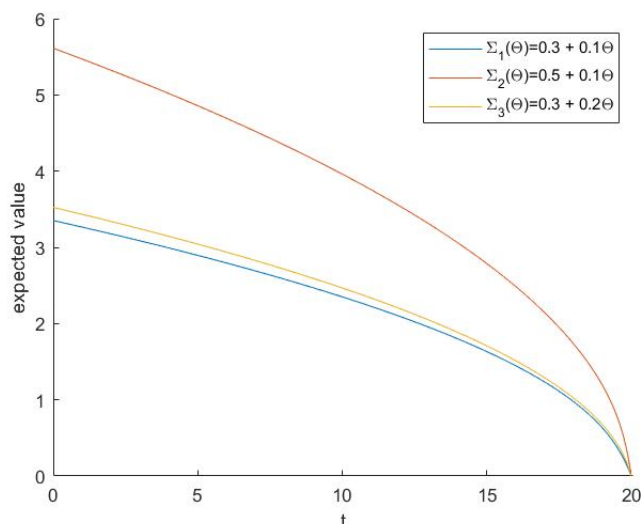


Figure 6.15: Expected values at $S = strike = 100$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$.

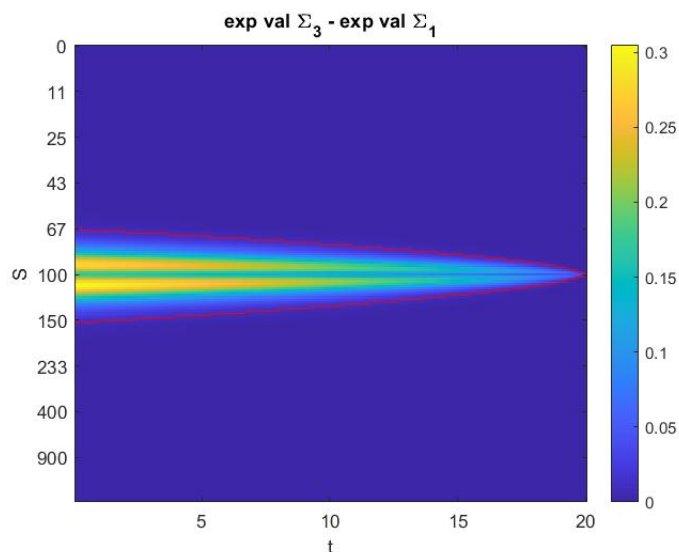


Figure 6.16: Absolute difference in expected values for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_3 is circled in red.

The variance of the solution for Σ_3 is approximately four times as high as for Σ_1 , as the plot of the variances in $t = 0$ in figure 6.17 shows. This coincides with the relation of variances $Var(\Sigma(\Theta)) = \sigma_1^2$ of the different models for $\Sigma(\Theta)$. Hence, σ_1 determines the maximum value of the variance of the corresponding SG solution.

The variance for Σ_2 attains the same maximum as Σ_1 , however the curve is flatter. Figure 6.18 displays the absolute differences of variances of the SG solutions for Σ_1 and Σ_2 . One can identify the difference to lie in the neighbourhood of the strike price $S = 100$. This coincides with the larger smoothing area in the case of Σ_2 that one could observe in figure 6.13. The shape of the variance is therefore determined by the value of σ_0 : the larger it is, the flatter the curve.

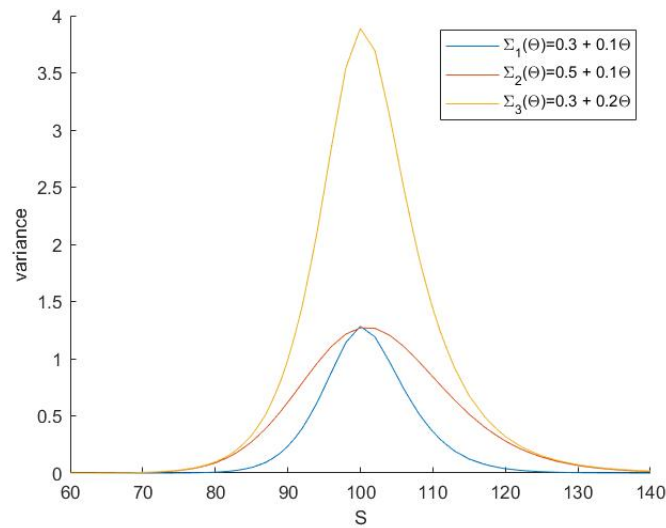


Figure 6.17: Variances at $t = 0$ for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$, $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$ and $\Sigma_3(\Theta) = 0.3 + 0.2\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 400$.

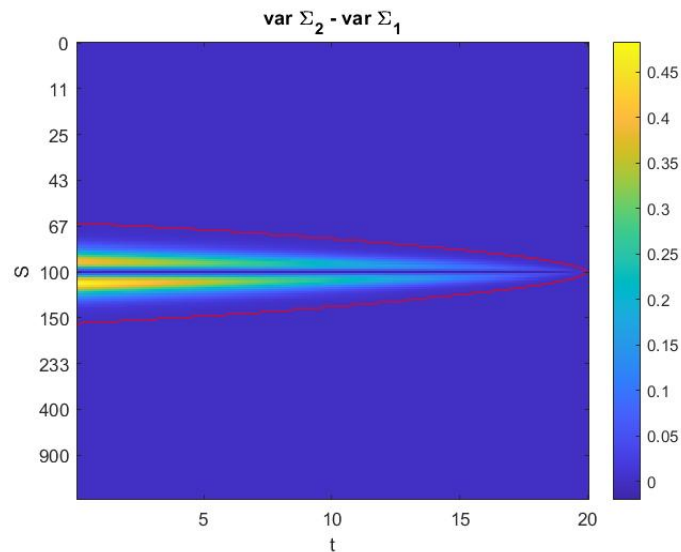


Figure 6.18: Absolute difference in variances for a European Call option for the models $\Sigma_1(\Theta) = 0.3 + 0.1\Theta$ and $\Sigma_2(\Theta) = 0.5 + 0.1\Theta$, Θ normal distributed, with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 84$ where the smoothing area for Σ_2 is circled in red.

Comparison to real market data:

As a last step, the model will be tested on real data. For this, the end of day values of a European Call option on the DAX index issued on the January 7th 2019 with strike price $strike = 10275$ and maturity on September 20th 2019, i.e. $T = 180$ days, were

obtained¹. The implied volatilities were calculated by formula 2.3. They are visualized as a histogram in figure 6.19. A normal distribution was fitted to these data by maximum likelihood, i.e. one searched for σ_0 and σ_1 such that $\Sigma(\Theta) = \sigma_0 + \sigma_1\Theta$ with a standard normal distributed random variable Θ is most likely to attain the values of the implied volatilities as realizations. From statistics, it is known that these optimal values are given by the arithmetic mean of the implied volatilities $\sigma_0 = 0.2292$ and their standard deviation $\sigma_1 = 0.1126$, see e.g. [Kre05] Beispiel 13.1.

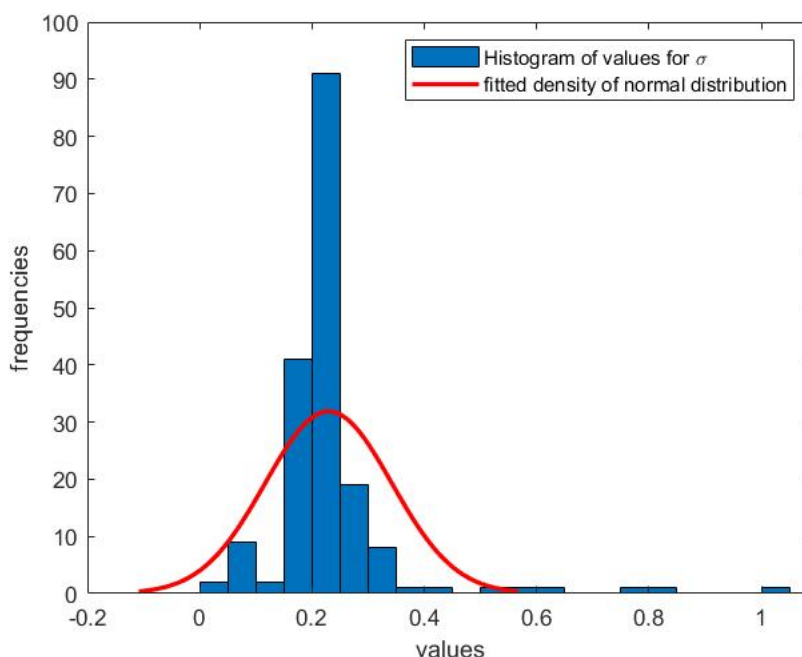


Figure 6.19: Implied volatilities of the European Call option data with fitted normal distribution.

With this model for $\Sigma(\Theta)$, the SG solution can be computed. The truncation numbers $K = 1$ and $N = 5$, for which the system of equations 4.8 is parabolic, are chosen and the number of grid points in ζ is set to $M_\zeta = 200$. The number $N_\tau = 678$ of grid points in τ is chosen to be large enough such that the scheme is stable.

Figure 6.20 shows the real market data together with the expected value of the SG solution and the range expected value plus/minus the standard deviation. A closer plot of the last 55 days of the option can be found in figure 6.21. One observes that the expected value is very close to the real data in this time span. For earlier times, the expected value lies almost everywhere above the data, but the data always lies in the range expected value plus/minus the standard deviation. This is consistent with stochastic theory that tells us by the inequality of Tschebyscheff, see [Kre05] Satz 3.15, that the probability that a random variable lies in the interval of its expected value plus/minus twice its standard deviation is greater than 75%.

¹The values were taken from the website <https://www.finanzen.net/>.

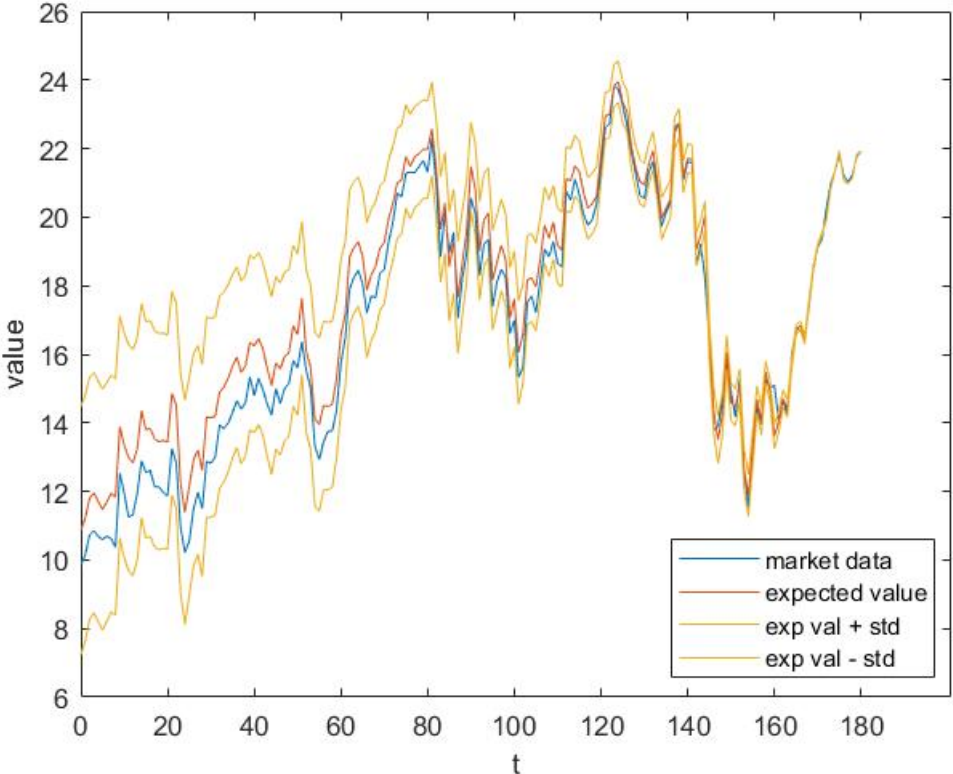


Figure 6.20: Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation.

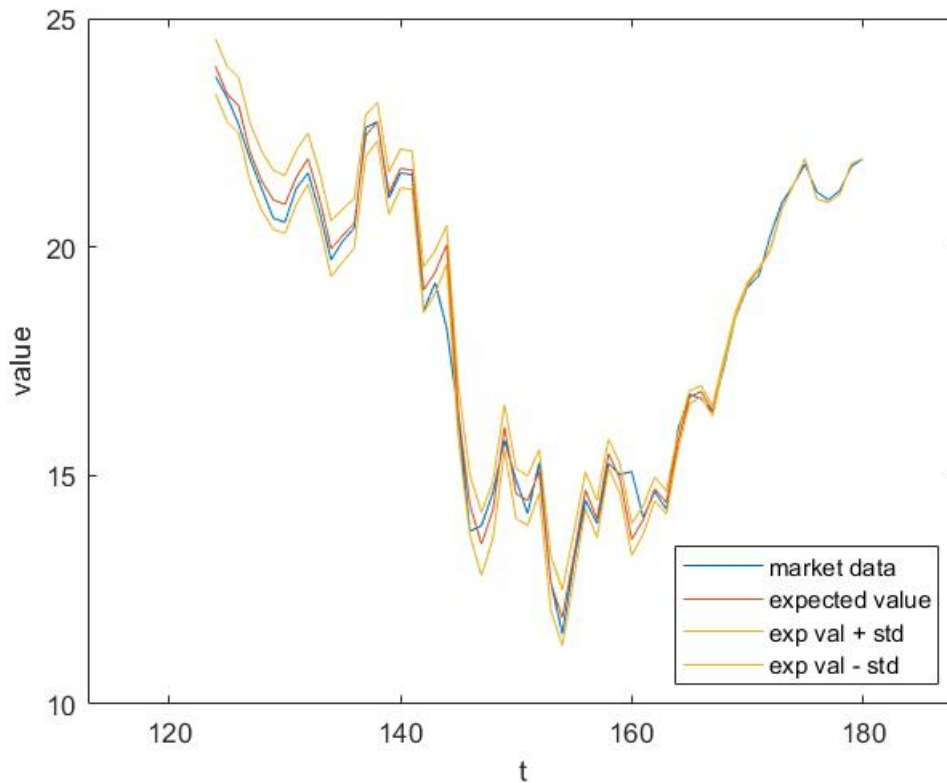


Figure 6.21: Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation for the last 55 days of the option.

Figure 6.22 compares the market prices, the expected value of the SG solution and the prices according to a deterministic model, where the average of all implied volatilities was taken as a value for the constant volatility. Note that also the deterministic prices lie above the market prices. However, our stochastic SG solution allows realizations that differ from the expected value within a certain range as the market prices do. Clearly, this is not possible for a deterministic model.

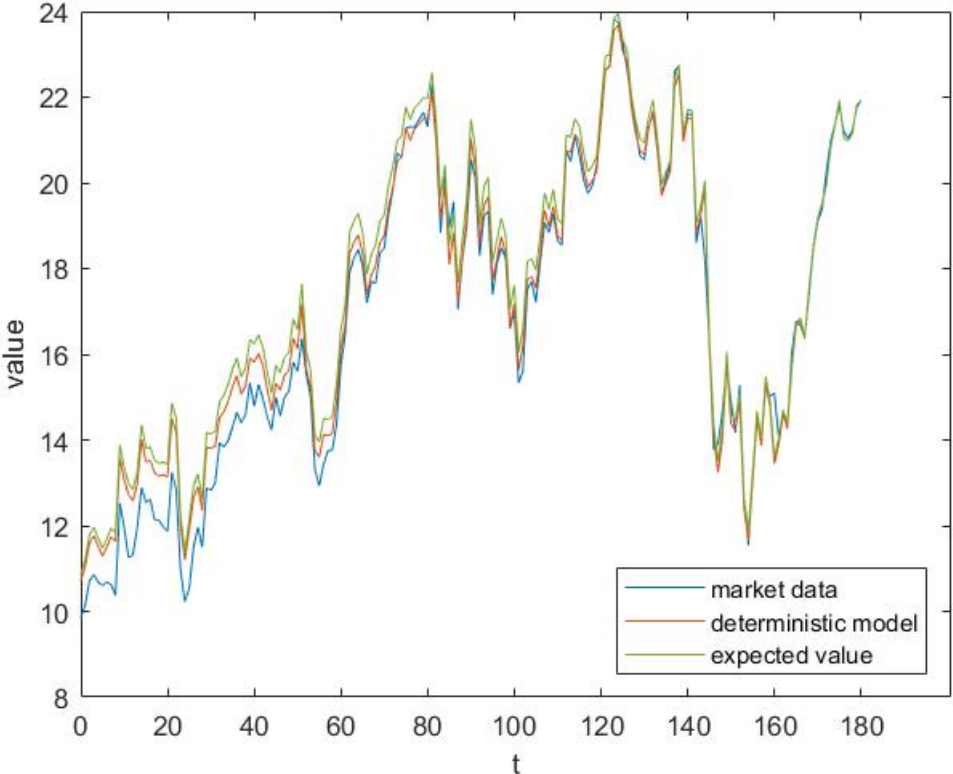


Figure 6.22: Market values of the option together with the deterministic solution and the expected value of the SG solution.

Chapter 7

Extension of the model to volatility depending on finitely many independent random variables

The Stochastic Galerkin method allows the stochastic parameter to depend on more than one random variable, as described in [Xiu10] Chapter 6.1 or [Sul15] Chapter 8.7. Such a model will be applied to the Black Scholes equation with uncertain volatility in this chapter by formulating the Black Scholes equation with uncertain volatility depending on finitely many independent random variables. Afterwards, the theoretical background of gPC expansions in multiple random variables will be explained and the Stochastic Galerkin method will be applied. Then, a numerical implementation of this model will be derived followed by numerical examples.

7.1 The Black Scholes equation with uncertain volatility depending on multiple random variables

In economics, there is plenty of research on impact factors on the volatility for different assets. For example, the impact of ratings from rating agencies on the volatility of a country's stock market and its bonds was studied in [RM98]. Using time series models, [RTA16] investigated the influence of central bank intervention on the volatility of exchange rates and [RKB12] examined the effects on the volatility of asset prices when futures on those assets, i.e. contracts to buy or sell the asset for a predefined price at a fixed point in the future, were introduced.

Returning to our model, assume information on influence factors on the volatility of the stochastic asset is given. The actual state of such an influence factor may be difficult to obtain from the market. Thus, we will model the influence factor as a random variable. By introducing dependency of the volatility on this random variable, the volatility model can be extended to use information from the influence factor. Nonetheless, the volatility might still behave stochastic even if a state of the influence factor is fixed. In this case,

one might model the volatility to also depend on another random variable. This mimicks the stochastic nature of the volatility.

In total, one obtains models for the volatility of the form

$$\Sigma(\Theta_1, \dots, \Theta_L),$$

for a natural number $L \in \mathbb{N}$, where the random variables $\Theta_1, \dots, \Theta_L$ are assumed to be continuous.

If the volatility is modelled as above, the Black Scholes equation with uncertain volatility is given by

$$0 = \frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial t} + \frac{1}{2} \Sigma^2(\Theta_1, \dots, \Theta_L) S^2 \frac{\partial^2 V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S^2} + rS \frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S} - rV(S, t, \Theta_1, \dots, \Theta_L). \quad (7.1)$$

Below, we will study this equation by applying the Stochastic Galerkin method.

7.2 Generalized Polynomial Chaos expansions in finitely many independent random variables

In this section, the generalized Polynomial Chaos (gPC) expansion of functions of more than one random variable will be explained, because we will need them for the Stochastic Galerkin method.

Denote by $\Theta_1, \dots, \Theta_L$ continuous random variables with joint density, i.e. density of $(\Theta_1, \dots, \Theta_L)$, given by $\bar{\mu} : \bar{\mathcal{D}} \rightarrow \mathbb{R}$ for a multivariate domain $\bar{\mathcal{D}} \subset \mathbb{R}^L$. As in the case of one random variable, for a function $f : \bar{\mathcal{D}} \rightarrow \mathbb{R}$ we will use the following notation for integration with respect to the density $\bar{\mu}$:

$$\langle \bar{f} \rangle_{\bar{\mu}} := \int_{\bar{\mathcal{D}}} \bar{f}(x_1, \dots, x_L) \bar{\mu}(x_1, \dots, x_L) d(x_1, \dots, x_L).$$

Orthogonal polynomials w.r.t. $\bar{\mu}$ can be defined as follows:

Definition 8 (adapted from [Sul15] Definition 8.24). Let $\bar{\mu} : \bar{\mathcal{D}} \rightarrow \mathbb{R}$ be a multivariate probability density defined on the domain $\bar{\mathcal{D}} \subset \mathbb{R}^L$. Then a system of polynomials $\{\bar{p}_\alpha : \bar{\mathcal{D}} \rightarrow \mathbb{R} \mid \alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L\}$, where the polynomial $\bar{p}_\alpha(x_1, \dots, x_L)$ has degree in the i -th variable $\deg_{x_i}(\bar{p}_\alpha) = \alpha_i$, is called an *infinite system of orthogonal polynomials*, if for all multi indices $\alpha, \beta \in \mathbb{N}_0^L$ one has

$$\begin{aligned} \langle \bar{p}_\alpha \bar{p}_\beta \rangle_{\bar{\mu}} &= 0 & \text{for } \alpha \neq \beta, \\ \langle \bar{p}_\alpha^2 \rangle_{\bar{\mu}} &=: \bar{\gamma}_\alpha > 0. \end{aligned}$$

Existence of orthogonal polynomial systems again follows from the Gram Schmidt algorithm, if for all $\alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L$ the moments $\langle x_1^{\alpha_1} \cdot \dots \cdot x_L^{\alpha_L} \rangle_{\bar{\mu}}$ are finite. Hence,

uniqueness of the orthogonal polynomials is given up to multiplication by constants.

In the following, assume for simplicity that $\Theta_1, \dots, \Theta_L$ are independent continuous random variables with corresponding densities $\mu_i : \mathcal{D}_i \rightarrow \mathbb{R}$, $\mathcal{D}_i \subset \mathbb{R}$ for $i = 1, \dots, L$. Then their joint density is given by $\bar{\mu}(x_1, \dots, x_L) = \mu_1(x_1) \cdot \dots \cdot \mu_L(x_L)$. Therefore, infinite orthogonal polynomial systems w.r.t. $\bar{\mu}$ exist if and only if they exist w.r.t. every μ_i and the orthogonal polynomials w.r.t. $\bar{\mu}$ are of the form

$$\bar{p}_\alpha(x_1, \dots, x_L) = p_{\alpha_1}^{(1)}(x_1) \cdot \dots \cdot p_{\alpha_L}^{(L)}(x_L), \quad \alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L, \quad (7.2)$$

where the $p_{\alpha_i}^{(i)}$ are the orthogonal polynomials w.r.t. μ_i of degree α_i , $i = 1, \dots, L$, see e.g. [Xiu10] Chapter 5.2. The $\bar{\gamma}_\alpha$ are products of the $\gamma_{\alpha_i}^{(i)}$ corresponding to the polynomials $p_{\alpha_i}^{(i)}$, i.e. $\bar{\gamma}_\alpha = \gamma_{\alpha_1}^{(1)} \cdot \dots \cdot \gamma_{\alpha_L}^{(L)}$.

The space $L_{\bar{\mu}}^2$ of all square integrable functions w.r.t. the density $\bar{\mu}$ from $\bar{\mathcal{D}} = \mathcal{D}_1 \times \dots \times \mathcal{D}_L$ to \mathbb{R} is by iterative application of Example E.10 from [Jan97] isometric to the tensor product

$$L_{\bar{\mu}}^2 := L_{\mu_1 \times \dots \times \mu_L}^2(\mathcal{D}_1 \times \dots \times \mathcal{D}_L, \mathbb{R}) \cong L_{\mu_1}^2 \otimes \dots \otimes L_{\mu_L}^2.$$

Example E.5 in the same book shows, that if the $p_n^{(i)}$, $n \in \mathbb{N}_0$, form a complete basis of $L_{\mu_i}^2$ for every $i = 1, \dots, L$, then the \bar{p}_α , $\alpha \in \mathbb{N}_0^L$, as defined in equation 7.2 form a complete basis of $L_{\bar{\mu}}^2$. In this case, the gPC expansion can be defined as in the case of dependence on one random variable:

Theorem 9. *For $i = 1, \dots, L$ let $\Theta_i : \Omega_i \rightarrow \mathbb{R}$ be a continuous random variable with corresponding density $\mu_i : \mathcal{D}_i \rightarrow \mathbb{R}$ such that the Θ_i are independent and the orthogonal polynomials $(p_n^{(i)})_{n \in \mathbb{N}}$ w.r.t. μ_i form a complete basis of $L_{\mu_i}^2$. Denote by \mathcal{F} 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 $X : \Omega_1 \times \dots \times \Omega_L \rightarrow \mathcal{F}$ that equals a function $\tilde{X} \in L_{\mu_1 \times \dots \times \mu_L}^2(\mathcal{D}_1 \times \dots \times \mathcal{D}_L, \mathcal{F})$ evaluated in the Θ_i in distribution*

$$X = \tilde{X}(\Theta_1, \dots, \Theta_L)$$

can be represented in the form

$$X = \sum_{\alpha \in \mathbb{N}_0^L} x_\alpha \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) = \sum_{i_1, \dots, i_L=0}^{\infty} x_{(i_1, \dots, i_L)} p_{i_1}^{(1)}(\Theta_1) \dots p_{i_L}^{(L)}(\Theta_L) \quad (7.3)$$

where the $\bar{p}_\alpha = p_{\alpha_1}^{(1)} \dots p_{\alpha_L}^{(L)}$, $\alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L$, are the orthogonal polynomials w.r.t. the joint density of the Θ_i given by $\bar{\mu} = \mu_1 \cdot \dots \cdot \mu_L$. The x_α are given by

$$x_\alpha = \frac{\langle X \bar{p}_\alpha \rangle_{\bar{\mu}}}{\langle \bar{p}_\alpha^2 \rangle_{\bar{\mu}}} \in \mathcal{F}. \quad (7.4)$$

Again equality of random variables is meant in distribution. The proof works analogously to the proof of the gPC expansion theorem 6 for dependency on one random variable.

Definition 9. A representation as in equation 7.3 is again called *generalized Polynomial Chaos expansion* of X .

If one does not assume independence of the Θ_i , one can still define gPC expansions. For further details, see [Rah18]. However, for our purposes the assumption of independence is sufficient.

7.3 Application of the Stochastic Galerkin method in multiple random variables

In this section, the Stochastic Galerkin method will be applied to the Black Scholes equation with uncertain volatility depending on $L \in \mathbb{N}$ random variables 7.1

$$0 = \frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial t} + \frac{1}{2} \Sigma^2(\Theta_1, \dots, \Theta_L) S^2 \frac{\partial^2 V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S^2} + rS \frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S} - rV(S, t, \Theta_1, \dots, \Theta_L).$$

The aim of this is to transform the stochastic PDE into a system of deterministic PDEs for the gPC coefficients of the solution $V(S, t, \Theta_1, \dots, \Theta_L)$. The application of the Stochastic Galerkin method will be explained in six steps, as it was done in chapter 4 for dependency on one random variable. In literature, the general procedure of application is explained in [Xiu10] chapter 6.1 for instance.

In the following, let $\Theta_1, \dots, \Theta_L$ be independent continuous random variables with density functions μ_1, \dots, μ_L such that the orthogonal polynomials $(p_n^{(i)})_{n \in \mathbb{N}}$ w.r.t. every μ_i form a complete basis of $L^2_{\mu_i}$. Denote by $\bar{p}_\alpha(x_1, \dots, x_L) = p_{\alpha_1}^{(1)}(x_1) \dots p_{\alpha_L}^{(L)}(x_L)$ the orthogonal polynomials of degree $\alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L$ w.r.t. the joint density $\bar{\mu}(x_1, \dots, x_L) = \mu_1(x_1) \dots \mu_L(x_L)$ of the Θ_i .

Additionally, assume that $\Sigma \in L^2_{\bar{\mu}}$ and $V \in L^2_{\bar{\mu}}(\bar{\mathcal{D}}, L^2((0, \infty) \times [0, T], \mathbb{R}))$ for the domain $\bar{\mathcal{D}} = \mathcal{D}_1 \times \dots \times \mathcal{D}_L$ of $\bar{\mu}$, where \mathcal{D}_i denotes the domain of μ_i .

Step 1: Writing the stochastic variables as gPC expansions

At first, represent $\Sigma(\Theta_1, \dots, \Theta_L)$, $V(S, t, \Theta_1, \dots, \Theta_L)$ and the partial derivatives of V in S and t by their gPC expansions. This is possible because of theorem 9. The partial derivatives pass over to the gPC coefficients by the same reason as in the case of dependency on one random variable, see chapter 4 Step 2.

$$\begin{aligned} \Sigma(\Theta_1, \dots, \Theta_L) &= \sum_{\alpha \in \mathbb{N}_0^L} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ V(S, t, \Theta_1, \dots, \Theta_L) &= \sum_{\alpha \in \mathbb{N}_0^L} v_\alpha(S, t) \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ \frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial t} &= \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial v_\alpha(S, t)}{\partial t} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \end{aligned} \tag{7.5}$$

$$\begin{aligned}\frac{\partial V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S} &= \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial v_\alpha(S, t)}{\partial S} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ \frac{\partial^2 V(S, t, \Theta_1, \dots, \Theta_L)}{\partial S^2} &= \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial^2 v_\alpha(S, t)}{\partial S^2} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L)\end{aligned}$$

Step 2: Inserting the gPC expansions in the differential equation

Now, one can insert the gPC expansions in the Black Scholes equation 7.1 to obtain

$$\begin{aligned}0 &= \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial v_\alpha(S, t)}{\partial t} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ &+ \frac{1}{2} \left(\sum_{\alpha \in \mathbb{N}_0^L} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \right)^2 S^2 \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial^2 v_\alpha(S, t)}{\partial S^2} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ &+ rS \sum_{\alpha \in \mathbb{N}_0^L} \frac{\partial v_\alpha(S, t)}{\partial S} \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) - r \sum_{\alpha \in \mathbb{N}_0^L} v_\alpha(S, t) \bar{p}_\alpha(\Theta_1, \dots, \Theta_L).\end{aligned}\tag{7.6}$$

Step 3: Multiplication of the equation with an orthogonal polynomial evaluated in $\Theta_1, \dots, \Theta_L$ and applying the expectation, for all orthogonal polynomials

The equation is multiplied by $\bar{p}_\delta(\Theta_1, \dots, \Theta_L)$ and the expected value is applied on both sides of PDE 7.6, for each $\delta \in \mathbb{N}_0^L$ at a time. Because of $\langle \bar{p}_\alpha \bar{p}_\delta \rangle_{\bar{\mu}} = 0$ for $\alpha \neq \delta$ by orthogonality, these terms drop out. Dividing by $\langle \bar{p}_\delta^2 \rangle_{\bar{\mu}} = \bar{\gamma}_\alpha > 0$, one obtains a coupled system of deterministic equations for all coefficients $v_\delta, \delta \in \mathbb{N}_0^L$,

$$0 = \frac{\partial v_\delta(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{\alpha, \beta, \gamma \in \mathbb{N}_0^L} \sigma_\alpha \sigma_\beta \frac{\partial^2 v_\gamma(S, t)}{\partial S^2} \frac{\langle \bar{p}_\alpha \bar{p}_\beta \bar{p}_\gamma \bar{p}_\delta \rangle_{\bar{\mu}}}{\langle \bar{p}_\delta^2 \rangle_{\bar{\mu}}} + rS \frac{\partial v_\delta(S, t)}{\partial S} - r v_\delta(S, t).$$

Introducing the Galerkin multiplication tensor

$$M_{\alpha\beta\gamma\delta} := \frac{\langle \bar{p}_\alpha \bar{p}_\beta \bar{p}_\gamma \bar{p}_\delta \rangle_{\bar{\mu}}}{\langle \bar{p}_\delta^2 \rangle_{\bar{\mu}}},$$

that exists since the integrated functions are all polynomials in L variables, the coupled system can be reformulated

$$0 = \frac{\partial v_\delta(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{\alpha, \beta, \gamma \in \mathbb{N}_0^L} \sigma_\alpha \sigma_\beta \frac{\partial^2 v_\gamma(S, t)}{\partial S^2} M_{\alpha\beta\gamma\delta} + rS \frac{\partial v_\delta(S, t)}{\partial S} - r v_\delta(S, t).\tag{7.7}$$

Step 4: Transforming the boundary conditions to the solution's gPC coefficients

The boundary conditions have to be converted to boundary conditions for the system of PDEs 7.7 for the gPC coefficients. They are given by the particular financial derivative one wants to price and usually they are deterministic.

Recall, that the price of a European Call option as in example 1 satisfies the final condition $V(S, T) = (S - \text{strike})^+$, $S \in (0, \infty)$, and the boundary conditions $V(0, t) = 0$ and $V(S, t)/S \rightarrow 1$ for $S \rightarrow \infty$, $t \in [0, T]$. The final and boundary conditions for the Stochastic Galerkin solution are deterministic too by the same argumentation as in chapter 4. So, they only appear in the coefficient $v_{(0, \dots, 0)}$, whose corresponding orthogonal polynomial $\bar{p}_{(0, \dots, 0)} \equiv 1$ is the only polynomial independent of all Θ_i and therefore deterministic when evaluated in $\Theta_1, \dots, \Theta_L$. Thus, the boundary condition for the system of equations 7.7 are

$$\begin{aligned} v_{(0, \dots, 0)}(S, T) &= (S - \text{strike})^+, & S \in (0, \infty), \\ v_\alpha(S, T) &= 0, & S \in (0, \infty), \\ v_{(0, \dots, 0)}(S, t), v_\alpha(S, t) &\xrightarrow{S \rightarrow 0} 0, & t \in [0, T], \\ v_{(0, \dots, 0)}(S, t)/S &\xrightarrow{S \rightarrow \infty} 1, & t \in [0, T], \\ v_\alpha(S, t)/S &\xrightarrow{S \rightarrow \infty} 0, & t \in [0, T] \end{aligned}$$

for all $(0, \dots, 0) \neq \alpha \in \mathbb{N}_0^L$.

Step 5: Solving the system of differential equations

In order to solve system 7.7 numerically, we truncate the gPC expansions of Σ and V to a finite number of terms. The truncation is done such that the polynomials attain a maximum total degree $\text{deg}_{x_1} + \dots + \text{deg}_{x_L}$ as it was done in [Xiu10] Chapter 6.1. Define $|\alpha| := \alpha_1 + \dots + \alpha_L$ for $\alpha = (\alpha_1, \dots, \alpha_L) \in \mathbb{N}_0^L$, then the total degree of \bar{p}_α equals $|\alpha|$. This leads to the truncated gPC expansion

$$\begin{aligned} \Sigma^K(\Theta_1, \dots, \Theta_L) &:= \sum_{\alpha \in \mathbb{N}_0^L, |\alpha| \leq K} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \dots, \Theta_L) \\ V^N(S, t, \Theta_1, \dots, \Theta_L) &:= \sum_{\delta \in \mathbb{N}_0^L, |\delta| \leq N} v_\delta^N(S, t) \bar{p}_\delta(\Theta_1, \dots, \Theta_L) \end{aligned} \quad (7.8)$$

for fixed $K, N \in \mathbb{N}_0$ and $v_\delta^N \in L^2((0, \infty) \times [0, T], \mathbb{R})$.

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

$$0 = \frac{\partial v_\delta^N(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{\substack{\alpha, \beta, \gamma \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K, \\ |\gamma| \leq N}} \sigma_\alpha \sigma_\beta \frac{\partial^2 v_\gamma^N(S, t)}{\partial S^2} M_{\alpha\beta\gamma\delta} + rS \frac{\partial v_\delta^N(S, t)}{\partial S} - r v_\delta^N(S, t). \quad (7.9)$$

The numerical methods used to solve this are explained in the next section.

As in the case of dependency on one random variable, convergence of the truncated Stochastic Galerkin solution 7.8 to the true solution 7.5 as $N \rightarrow \infty$ is open to further research. However, one assumes convergence to be given in order to trust the numerical solution to represent the true solution.

Step 6: Reconstructing the stochastic solution

The truncated solution V^N can be approximately reconstructed from the numerically obtained v_δ^N by formula 7.8. However, mean and variance of the truncated solution can also be computed directly from the gPC coefficients:

- The expected value is given by $E(V^N) = v_{(0,\dots,0)}^N$ and
- the variance by $Var(V^N(S, t)) = \sum_{\delta \in \mathbb{N}_0^L, |\delta| \leq N, \delta \neq (0,\dots,0)} (v_\delta^N)^2 \bar{\gamma}_\delta$,

as one can show by the same arguments as in the case of dependency on one random variable.

7.4 Numerics

In this section, the numerics used for solving the system of equations 7.9 is presented. Again European Call options with strike price *strike* and maturity T will be considered for demonstrative purposes.

In order to reuse as much of the code developed for dependence on one random variable as possible, we will at first rewrite system 7.9 in vector form as in equation 4.8. To put the gPC coefficients of V in a vector, one needs to define an order in which they will appear in the vector. This is done via a bijection ϕ from the set $\{0, \dots, |I| - 1\}$ of positions in the vector to the set of multi indices $I := \{\delta \in \mathbb{N}_0^L \mid |\delta| \leq N\}$.

In this work, a bijective mapping was used that is equivalent to sorting the multi indices as shown in table 7.1 and assigning them their sorting index, as suggested by [Sul15] Chapter 8.7.

vector position	multi index
0	(0, 0, ..., 0, 0)
1	(0, 0, ..., 0, 1)
⋮	⋮
N	(0, 0, ..., 0, N)
$N + 1$	(0, 0, ..., 1, 0)
$N + 2$	(0, 0, ..., 1, 1)
⋮	⋮
$2N$	(0, 0, ..., 1, $N - 1$)
⋮	⋮
$ I - 1$	($N, 0, \dots, 0$)

Table 7.1: Mapping ϕ from the set of vector positions to the multi indices

Note that $|I|$ can easily be computed by a program like the following.

Algorithm 6 Compute the cardinality $|I|$

Input: Truncation number N .

Output: cardinality $cardI = |I|$ with $I = \{\delta \in \mathbb{N}_0^L \mid |\delta| \leq N\}$.

```

cardI = 0;
for i1 = 0 : N do
    for i2 = 0 : N - i1 do
        ...
        for iL = 0 : N - i1 - ... - iL-1 do
            cardI = cardI + 1;
        end for
    end for
end for
end for
    
```

Define $\mathbf{v} := (v_{\phi(0)}^N, \dots, v_{\phi(|I|-1)}^N)^T$, then one can represent system 7.9 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

$$\mathbf{A}[n, l] = \sum_{\substack{\alpha, \beta \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K}} \sigma_\alpha \sigma_\beta M_{\alpha\beta(\phi(l))(\phi(n))}, \quad \text{for } n, l = 0, \dots, |I| - 1. \quad (7.10)$$

This system has the same shape as the vector form PDE 4.8 for dependency on one random variable. Hence, it is parabolic and therefore well posed, if the real values of all eigenvalues of \mathbf{A} are positive, which has to be checked for convergence of the numerical scheme.

The boundary conditions and the final value in vector form are

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

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

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

After applying the same transformation of variables 5.1 as in chapter 5

$$\begin{aligned}\zeta &= \frac{S}{S + \text{strike}}, \\ \tau &= T - t, \\ \bar{\mathbf{v}}(\zeta, \tau) &= \frac{\mathbf{v}(S, t)}{S + \text{strike}} = \frac{(1 - \zeta)\mathbf{v}(\text{strike} \zeta / (1 - \zeta), T - \tau)}{\text{strike}},\end{aligned}$$

the PDE for $\bar{\mathbf{v}}$ can be formulated:

$$\frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} = \frac{1}{2} \zeta^2 (1 - \zeta)^2 \mathbf{A} \frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2} + r \zeta (1 - \zeta) \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} - r (1 - \zeta) \bar{\mathbf{v}}(\zeta, \tau), \quad (7.11)$$

$$\zeta \in (0, 1), \tau \in [0, T],$$

with boundary and initial conditions

$$\begin{aligned}\bar{\mathbf{v}}(\zeta, 0) &= \begin{pmatrix} (2\zeta - 1)^+ \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \zeta \in (0, 1), \\ \bar{\mathbf{v}}(\zeta, \tau) &\xrightarrow{\zeta \rightarrow 0} \mathbf{0}_{|I|}, \quad \tau \in [0, T], \quad \text{and} \\ \bar{\mathbf{v}}(\zeta, \tau) &\xrightarrow{\zeta \rightarrow 1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \tau \in [0, T].\end{aligned}$$

This has the same shape as equation 5.3 and therefore the same numerical methods can be used to solve it, as explained later. The only differences lie the dimension of the vector $\bar{\mathbf{v}} \in \mathbb{R}^{|I|}$ and definition 7.10 of the coupling matrix \mathbf{A} .

Before one can apply a numerical scheme, the coupling matrix given in equation 7.10

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

has to be computed, since the scheme will depend on \mathbf{A} . To calculate the Galerkin multiplication tensor with entries $M_{\alpha\beta\gamma\delta}$, one can take advantage of the methods in chapter 5, since by independence of the Θ_i one has

$$\begin{aligned}M_{\alpha\beta\gamma\delta} &= \frac{\langle \bar{p}_\alpha \bar{p}_\beta \bar{p}_\gamma \bar{p}_\delta \rangle_{\bar{\mu}}}{\bar{\gamma}_\delta} = \frac{\langle p_{\alpha_1}^{(1)}(x_1) \dots p_{\alpha_L}^{(L)}(x_L) \cdot \dots \cdot p_{\delta_1}^{(1)}(x_1) \dots p_{\delta_L}^{(L)}(x_L) \rangle_{\mu_1(x_1) \dots \mu_L(x_L)}}{\gamma_{\delta_1}^{(1)} \dots \gamma_{\delta_L}^{(L)}} \\ &= \frac{\langle p_{\alpha_1}^{(1)} p_{\beta_1}^{(1)} p_{\gamma_1}^{(1)} p_{\delta_1}^{(1)} \rangle_{\mu_1}}{\gamma_{\delta_1}^{(1)}} \dots \frac{\langle p_{\alpha_L}^{(L)} p_{\beta_L}^{(L)} p_{\gamma_L}^{(L)} p_{\delta_L}^{(L)} \rangle_{\mu_L}}{\gamma_{\delta_L}^{(L)}} = M_{\alpha_1 \beta_1 \gamma_1 \delta_1}^{(1)} \dots M_{\alpha_L \beta_L \gamma_L \delta_L}^{(L)},\end{aligned}$$

where $M_{ikln}^{(j)}$ denotes the entry of the Galerkin multiplication tensor corresponding to density μ_j , $j = 1, \dots, L$. With this, one can compute the Galerkin multiplication tensor \mathbf{M} with $\mathbf{M}[i, k, l, n] = M_{\phi(i)\phi(k)\phi(l)\phi(n)}$, $l, n = 0, \dots, |I| - 1$ and $i, k = 0, \dots, |J| - 1$ with $J := \{\alpha \in \mathbb{N}_0^L \mid |\alpha| \leq K\}$ by the following program given in Pseudo Matlab Code.

Algorithm 7 Compute the Galerkin multiplication tensor matrix \mathbf{M} for L random variables

Input: Truncation number N, K of the truncated gPC expansions of V and Σ .

Output: Galerkin multiplication tensor \mathbf{M} .

Compute $\mathbf{M}^{(j)}$, $j = 1, \dots, L$, with $\mathbf{M}^{(j)}[i, k, l, n] = M_{ikln}^{(j)}$ by program 2 from chapter 5.

Compute \mathbf{M} : The value of $\phi^{-1}(\alpha)$ gets stored in phi_α , respectively for β, γ and δ
 $phi_\alpha = -1$; $phi_\beta = -1$; $phi_\gamma = -1$; $phi_\delta = -1$;

Compute $cardI := |I|$, $cardJ := |J|$ by program 6.

$\mathbf{M} = \text{zeros}(cardI, cardI, cardJ, cardJ)$;

for $\alpha_1, \beta_1 = 0 : K, \gamma_1, \delta_1 = 0 : N$ **do**

for $\alpha_2 = 0 : (K - \alpha_1), \beta_2 = 0 : (K - \beta_1), \gamma_2 = 0 : (N - \gamma_1), \delta_2 = 0 : (N - \delta_1)$ **do**

...

for $\alpha_L = 0 : (K - \alpha_1 - \dots - \alpha_{L-1})$ **do**

$phi_\alpha = phi_\alpha + 1$;

for $\beta_L = 0 : (K - \beta_1 - \dots - \beta_{L-1})$ **do**

$phi_\beta = phi_\beta + 1$;

for $\gamma_L = 0 : (N - \gamma_1 - \dots - \gamma_{L-1})$ **do**

$phi_\gamma = phi_\gamma + 1$;

for $\delta_L = 0 : (N - \delta_1 - \dots - \delta_{L-1})$ **do**

$phi_\delta = phi_\delta + 1$;

$\mathbf{M}[phi_\alpha, phi_\beta, phi_\gamma, phi_\delta] = \mathbf{M}^{(1)}[\alpha_1, \beta_1, \gamma_1, \delta_1] * \dots * \mathbf{M}^{(L)}[\alpha_L, \beta_L, \gamma_L, \delta_L]$;

end for

end for

end for

end for

end for

end for

Now one can compute \mathbf{A} by

$$\mathbf{A}[\phi^{-1}(\delta), \phi^{-1}(\gamma)] = \sum_{\substack{\alpha, \beta \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K}} \sigma_\alpha \sigma_\beta M_{\alpha\beta\gamma\delta}$$

for $\gamma, \delta \in I$ as follows.

Algorithm 8 Calculation of the coupling matrix A for L random variables

Input: Truncation number N, K of the truncated gPC expansions of V and Σ , vector $\sigma = [\sigma_{\phi_0}, \dots, \sigma_{\phi(|J|-1)}]$ of the gPC coefficients of Σ ordered by ϕ .

Output: Coupling matrix \mathbf{A} .

Compute the Galerkin multiplication tensor \mathbf{M} by program 7.

Compute the cardinality $\text{card}I = |I|$ by program 6.

$\mathbf{A} = \text{zeros}(\text{card}I, \text{card}I)$;

Again, phi_γ denotes the value $\phi^{-1}(\gamma)$, respectively for δ .

$\text{phi}_\gamma = -1$; $\text{phi}_\delta = -1$;

for $\gamma_1, \delta_1 = 0 : N$ **do**

for $\gamma_2 = 0 : (N - \gamma_1), \delta_2 = 0 : (N - \delta_1)$ **do**

 ...

for $\gamma_L = 0 : (N - \gamma_1 - \dots - \gamma_L), \delta_L = 0 : (N - \delta_1 - \dots - \delta_L)$ **do**

$\text{phi}_\gamma = \text{phi}_\gamma + 1$; $\text{phi}_\delta = \text{phi}_\delta + 1$;

$\mathbf{A}[\text{phi}_\delta, \text{phi}_\gamma] = \sigma \mathbf{M}[:, :, \text{phi}_\gamma, \text{phi}_\delta] \sigma'$;

end for

end for

end for

For $L = 2$ and σ given in matrix form

$$\sigma = \begin{pmatrix} \sigma_{00} & \sigma_{01} & \dots & \sigma_{0(K-1)} & \sigma_{0K} \\ \sigma_{10} & \sigma_{11} & \dots & \sigma_{1(K-1)} & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma_{(K-1)0} & \sigma_{(K-1)1} & \dots & 0 & 0 \\ \sigma_{K0} & 0 & \dots & 0 & 0 \end{pmatrix} \quad (7.12)$$

one can alternatively calculate \mathbf{A} by

$$\begin{aligned} \mathbf{A}[\phi^{-1}(\delta), \phi^{-1}(\gamma)] &= \sum_{\substack{\alpha, \beta \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K}} \sigma_\alpha \sigma_\beta M_{\alpha\beta\gamma\delta} = \sum_{\alpha_1, \beta_1=0}^K \sum_{\alpha_2=0}^{K-\alpha_1} \sum_{\beta_2=0}^{K-\beta_1} \sigma_{\alpha_1\alpha_2} \sigma_{\beta_1\beta_2} M_{\alpha_1\beta_1\gamma_1\delta_1}^{(1)} M_{\alpha_2\beta_2\gamma_2\delta_2}^{(2)} \\ &= \sum_{\alpha_1, \beta_1=0}^K \sum_{\alpha_2=0}^{K-\alpha_1} \sum_{\beta_2=0}^{K-\beta_1} \sigma[\alpha_1, \alpha_2] \mathbf{M}^{(2)}[\alpha_2, \beta_2, \gamma_2, \delta_2] \sigma[\beta_1, \beta_2] \mathbf{M}^{(1)}[\alpha_1, \beta_1, \gamma_1, \delta_1] \\ &= \sum_{\alpha_1=0}^K (\sigma \mathbf{M}^{(2)}[:, :, \gamma_2, \delta_2] \sigma^T (\mathbf{M}^{(1)}[:, :, \gamma_1, \delta_1])^T) [\alpha_1, \alpha_1] \\ &= \text{tr}(\sigma \mathbf{M}^{(2)}[:, :, \gamma_2, \delta_2] \sigma^T (\mathbf{M}^{(1)}[:, :, \gamma_1, \delta_1])^T), \end{aligned}$$

where tr denotes the trace of a matrix and $\mathbf{M}^{(2)}[:, :, \gamma_2, \delta_2]$ stands for the matrix of all elements on $\mathbf{M}^{(2)}$, whose last two indices are γ_2, δ_2 . This does not require the computation

of \mathbf{M} but directly uses $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$. Note that ϕ^{-1} is given by $(\alpha_1, \alpha_2) \mapsto \alpha_1(N+1) - \alpha_1(\alpha_1 - 1)/2 + \alpha_2$ for $L = 2$. In Pseudo code, \mathbf{A} is then calculated as follows:

Algorithm 9 Alternative calculation of the coupling matrix A for $L = 2$ random variables

Input: Truncation number N, K of the truncated gPC expansions of V and Σ , matrix σ as in equation 7.12 of the gPC coefficients of Σ ordered by ϕ .

Output: Coupling matrix \mathbf{A} .

Compute $\mathbf{M}^{(j)}$, $j = 1, \dots, L$, with $\mathbf{M}^{(j)}[i, k, l, n] = M_{ikln}^{(j)}$ by program 2 from chapter 5.

Compute the cardinality $cardI = |I|$ by program 6.

$\mathbf{A} = \text{zeros}(cardI, cardI);$

for $\gamma_1, \delta_1 = 0 : N$ **do**

for $\gamma_2 = 0 : (N - \gamma_1), \delta_2 = 0 : (N - \delta_1)$ **do**

$\mathbf{A}[(\delta_1(N+1) - \delta_1(\delta_1 - 1)/2 + \delta_2), (\gamma_1(N+1) - \gamma_1(\gamma_1 - 1)/2 + \gamma_2)]$
 $= \text{tr}(\sigma \mathbf{M}^{(2)}[:, :, \gamma_2, \delta_2] \sigma' (\mathbf{M}^{(1)}[:, :, \gamma_1, \delta_1])')$;

end for

end for

After \mathbf{A} was computed, one can apply the same explicit finite difference scheme as in chapter 5 to solve equation 7.11:

$$\begin{aligned} \bar{\mathbf{v}}(\zeta_m, \tau^{n+1}) = & \Delta\tau \left(\frac{1}{2} \zeta_m^2 (1 - \zeta_m)^2 \mathbf{A} \frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - 2\bar{\mathbf{v}}(\zeta_m, \tau^n) + \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{(\Delta\zeta)^2} \right. \\ & \left. + r\zeta_m(1 - \zeta_m) \frac{\bar{\mathbf{v}}(\zeta_{m+1}, \tau^n) - \bar{\mathbf{v}}(\zeta_{m-1}, \tau^n)}{2\Delta\zeta} - r(1 - \zeta_m)\bar{\mathbf{v}}(\zeta_m, \tau^n) \right) \\ & + \bar{\mathbf{v}}(\zeta_m, \tau^n). \end{aligned} \quad (5.6)$$

By the Lax-Richtmyer Equivalence Theorem 8, convergence of the numerical solution to the solution of the truncated system of PDEs 7.9 is given in the sense of Definition 7, if the scheme is consistent and stable and the system of equations is well posed. To show well posedness, one checks the real parts of the eigenvalues of \mathbf{A} to be positive, since then the equation is parabolic and therefore well posed. Consistency of the scheme 5.6 is shown in the same way as in chapter 5 and stability can be investigated by the same program 4 as in chapter 5. In total, a program to calculate the gPC coefficients of the truncated Stochastic Galerkin solution 7.8 is given below in Matlab Pseudo Code.

Algorithm 10 Computing the truncated gPC coefficients \mathbf{v} of V for L random variables

Input: Truncation numbers K, N of the truncated gPC expansions of V and Σ , vector of gPC coefficients $\sigma := (\sigma_{\phi(0)}, \dots, \sigma_{\phi(|J|-1)})$ of Σ (or matrix of gPC coefficients of Σ in the same form as in equation 7.12 for $L = 2$ and usage of program 9 for \mathbf{A}), interest rate r , maturity of the option T , and strike price *strike*, number of discretization points M_ζ in ζ .

Output: Tensor \mathbf{v} with $\mathbf{v}(k, m, n) = v_k(S_m, t^n)$, $k = 0, \dots, N$, $m = 0, \dots, M_\zeta - 1$, $n = 0, \dots, N_\tau$, vectors $\mathbf{S} := (S_0, \dots, S_{M_\zeta-1})$ and $\mathbf{t} = (t^0, \dots, t^{N_\tau})$ where S_m, t^n are the transformed back ζ_m, τ^n .

Calculate the coupling matrix \mathbf{A} by program 8 (or program 9 for $L = 2$ and σ as in 7.12).

Check for parabolicity:

```
if min(real(eig(A))) ≤ 0 then
    error('Error: Equation not parabolic');
end if
```

Check for stability by program 4, this yields the output variables *stable* and *dtaumax*.

```
if stable == 0 then
    error('Error: Scheme not stable');
end if
```

Calculate the minimum N_τ to have stability:

```
 $N_\tau = \text{ceil}(T/dtaumax);$ 
```

Compute $\bar{\mathbf{v}}$:

```
[cardI, ~] = size(A);
 $\bar{\mathbf{v}} = \text{zeros}(cardI, M_\zeta + 1, N_\tau + 1);$ 
```

Setting the initial condition for $\bar{\mathbf{v}}$:

```
 $\bar{\mathbf{v}}(0, :, 0) = \max(2/M_\zeta * (0 : M_\zeta) - 1, 0);$ 
```

Code continues on the next page.

Calculating \bar{v} for the remaining times by scheme 5.6:

for $n = 0 : N_\tau - 1$ **do**

Boundary conditions:

$m = 0;$

$\bar{v}(:, m, n + 1) = \text{zeros}(\text{card}I, 1);$

$m = M_\zeta;$

$\bar{v}(:, m, n + 1) = \text{zeros}(\text{card}I, 1);$

$\bar{v}(0, m, n + 1) = 1;$

Scheme:

for $m = 1 : M_\zeta - 1$ **do**

$dvdzeta = M_\zeta/2 * (\bar{v}(:, m + 1, n) - \bar{v}(:, m - 1, n));$

$d2vdzeta2 = M_\zeta^2 * (\bar{v}(:, m + 1, n) - 2 * \bar{v}(:, m, n) + \bar{v}(:, m - 1, n));$

$\bar{v}(:, m, n + 1) = T/N_\tau * (1/2 * (m/M_\zeta)^2 * (M_\zeta - m)^2 / M_\zeta^2 * A * d2vdzeta2 + r * m / M_\zeta * (M_\zeta - m) / M_\zeta * dvdzeta - r * (M_\zeta - m) / M_\zeta * \bar{v}(:, m, n)) + \bar{v}(:, m, n);$

end for

end for

Transforming back \bar{v} , calculating **S** and **t**:

$\mathbf{v} = \text{zeros}(\text{card}I, M_\zeta, N_\tau + 1);$

for $n = 0 : N_\tau$ **do**

$\mathbf{v}(:, :, n) = \text{strike} * (1 : -1/M_\zeta : 1/M_\zeta).^(-1) * \bar{v}(:, 1 : M_\zeta, N_\tau - n);$

end for

$\mathbf{S} = \text{strike} * (0 : M_\zeta - 1) ./ (M_\zeta : -1 : 1);$

$\mathbf{t} = 0 : T/N_\tau : T;$

7.5 Numerical results for the extended model

In this section, numerical solutions to the truncated system of equations 7.9 are visualized by plotting their means and variances. These solutions were obtained by the programs described in the last section. European Call options will be considered for demonstrative purposes. As in chapter 6, times t and the maturity T will be given in days for more convenient reading, whereas for the computations, these values were multiplied by 1/251 to go over to years.

At first we consider a European Call option with strike price $\text{strike} = 100$ and maturity $T = 20$ in a market with risk free rate of interest $r = 0$. The volatility of the underlying stochastic asset is modelled to be a linear combination of a standard normal distributed random variable Θ and a uniformly on $[-0.5, 0.5]$ distributed random variable Δ independent of Θ given by

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

For this model, we search for an approximate solution of the truncated system of equations 7.9 of the form 7.8 for $N = 5$. For this N , system 7.9 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 5.6 is stable.

The expected value surface of the Stochastic Galerkin (SG) solution is displayed in figure 7.1. Contour lines were drawn at height of quarters of the maximum absolute value and the borders of the smoothing area were drawn in red. These lines will be present in each of the following surface plots. Note that the surface is very similar to the expected value surfaces considered in chapter 6.

Figure 7.2 displays the variance of the Stochastic Galerkin solution whose shape resembles the characteristic shape in the case of dependence on one random variable in figure 6.5 for instance. The difference of this model compared to the models for dependence on one random variable will be studied below.

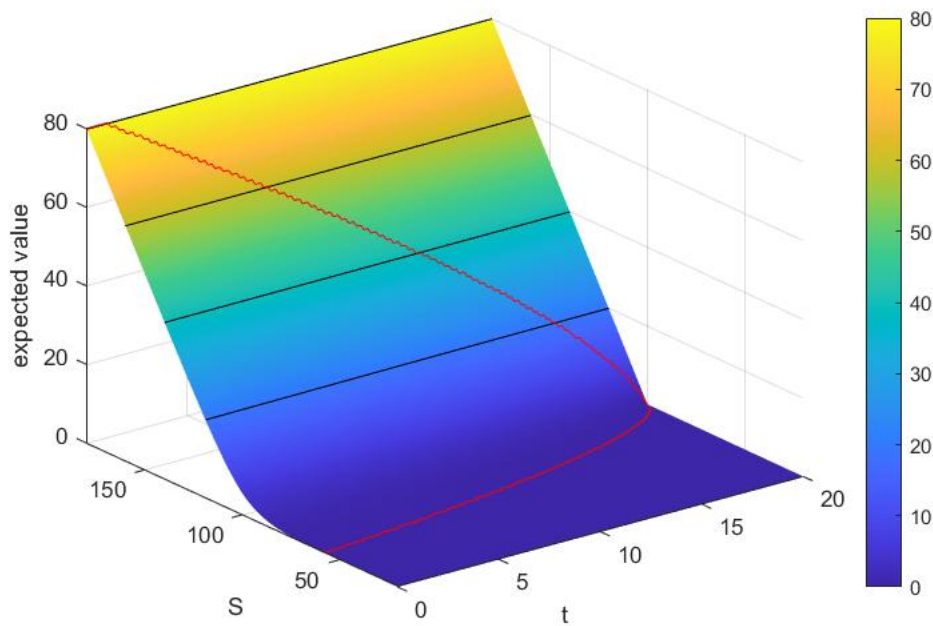


Figure 7.1: Expected value surface for a European Call option with $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

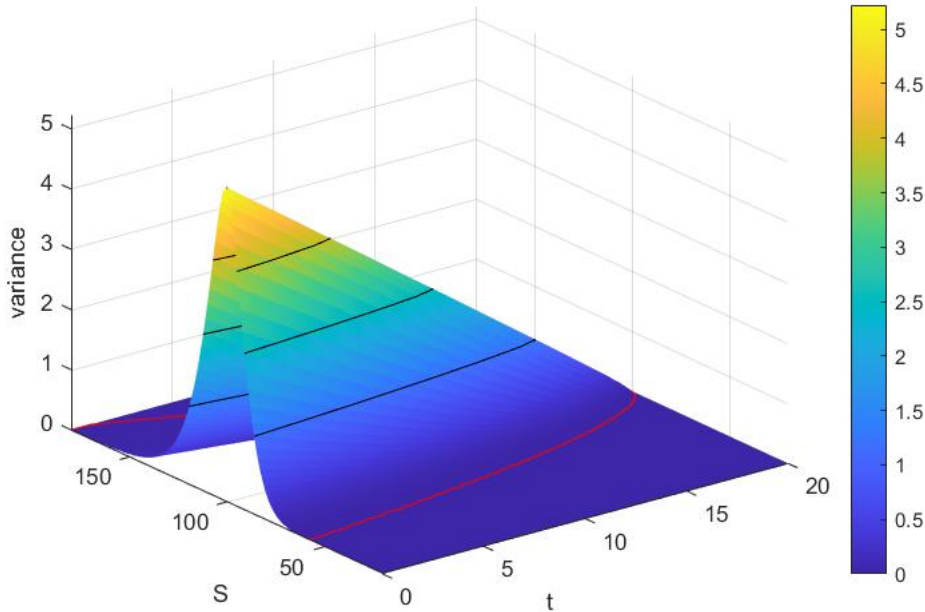


Figure 7.2: Variance surface for a European Call option with $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of its maximum value and smoothing area circled in red.

Comparing the expected value and variance for different volatility models of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$:

One can compare the SG solution for the volatility models

$$\begin{aligned} \Sigma_0(\Theta, \Delta) &= \Sigma_0(\Theta) = 0.5 + \sqrt{0.05}\Theta, \\ \Sigma_1(\Theta, \Delta) &= 0.5 + 0.2\Theta + 0.2\sqrt{12}\Delta \\ \Sigma_2(\Theta, \Delta) &= 0.5 + 0.1\Theta + 0.2\sqrt{12}\Delta \quad \text{and} \\ \Sigma_3(\Theta, \Delta) &= \Sigma_3(\Delta) = 0.5 + \sqrt{0.05 \cdot 12}\Delta. \end{aligned}$$

Note that all $\Sigma_i(\Theta, \Delta)$ have the same variance $\sigma_{01}^2/12 + \sigma_{10}^2 = 0.05$ and Σ_1 is the model considered before.

For these volatility models, the systems of equations 7.9 are parabolic for $N = 5$ and the finite difference scheme remains stable if one chooses $M_\zeta = 200$, $N_\tau = 319$. The expected values at $t = 0$ of the SG solutions for all four models are illustrated in figure 7.3 for all $S \in [80, 120]$. Since the differences are not easily observable, figure 7.4 shows an enlarged area of this plot around $S = 100$. Here one can see that the expected values for the volatility models Σ_1, Σ_2 properly depending on two random variables lie between the expectations of the one random variable models Σ_1, Σ_3 . A larger value of σ_{10} in comparison to σ_{01} leads to a solution more similar to the solution in case of dependency on Θ only given by model Σ_0 . Vice versa, for larger values of σ_{01} the solution lies closer to the solution for Σ_3 that depends on Δ only.

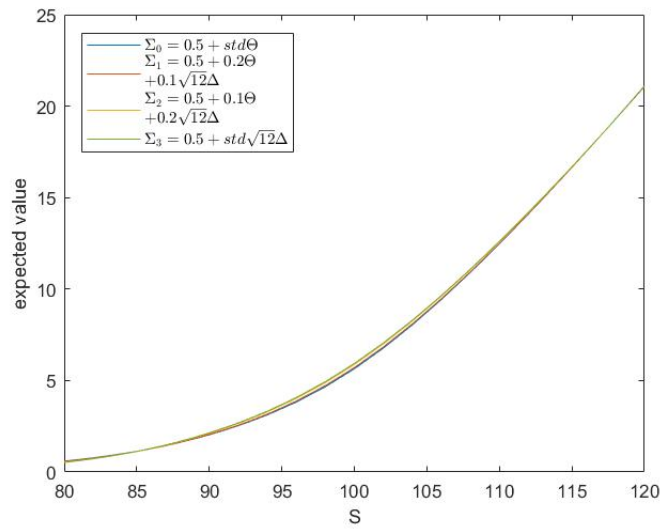


Figure 7.3: Expected values in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 , $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$.

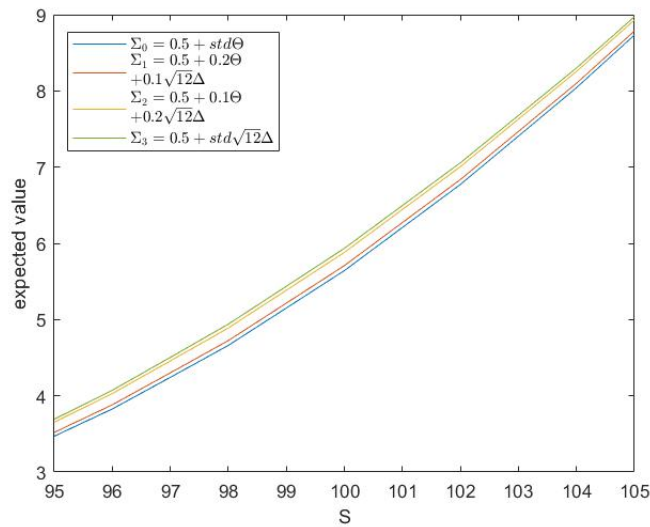


Figure 7.4: Expected values in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 zoomed in, $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$.

The same behaviour can be observed for the variances in figure 7.5.

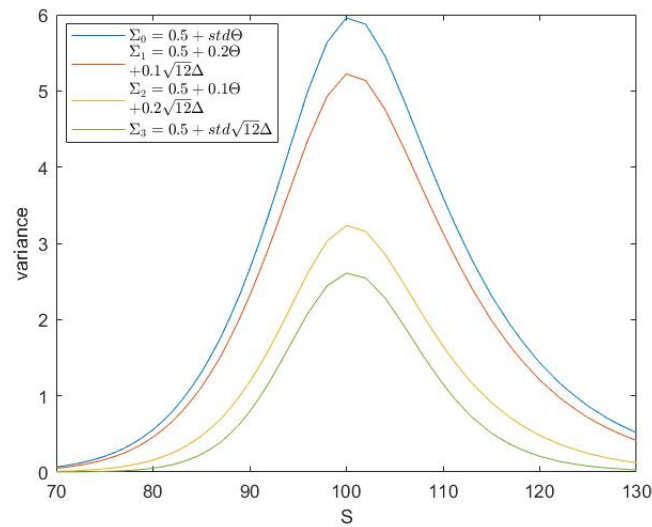


Figure 7.5: Variances in $t = 0$ for a European Call option for the models $\Sigma_0, \Sigma_1, \Sigma_2$ and Σ_3 , $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$.

To get a notion of the differences between model Σ_1 and the models Σ_0 and Σ_3 depending on one random variable only, the absolute differences in the expected values of the solutions for Σ_0 and Σ_1 and for Σ_1 and Σ_3 are displayed in figure 7.6 and figure 7.7 respectively. As one would expect from figures 7.3 and 7.4, the differences mostly lie in the smoothing area and the difference between the expected values for Σ_1 and Σ_0 is not as large as the one between the solutions for Σ_1 and Σ_3 .

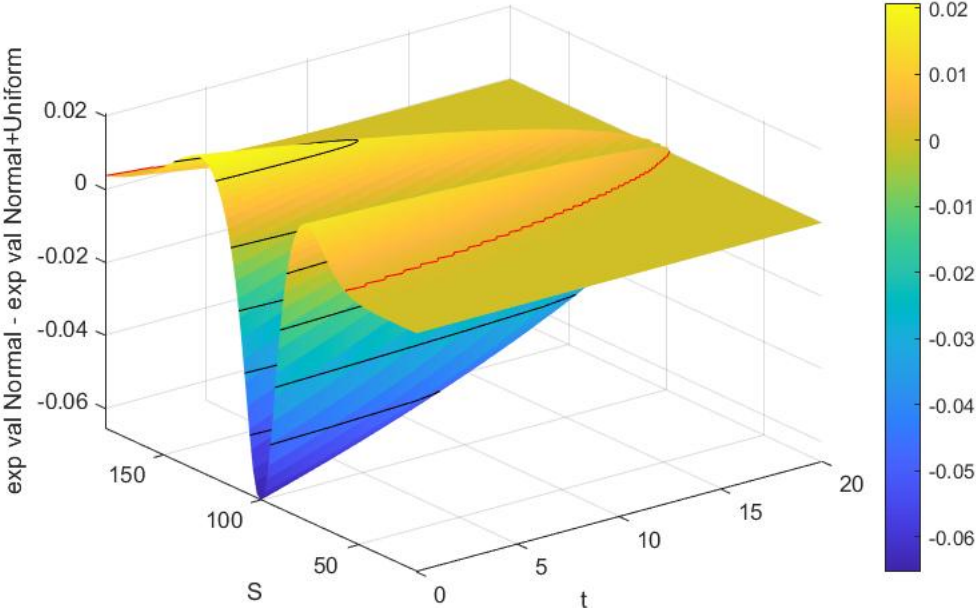


Figure 7.6: Absolute difference in the expected values for a European Call option for the models Σ_0 and Σ_1 with $T = 20, strike = 100, K = 1, N = 5, M_\zeta = 200, N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_0 circled in red.

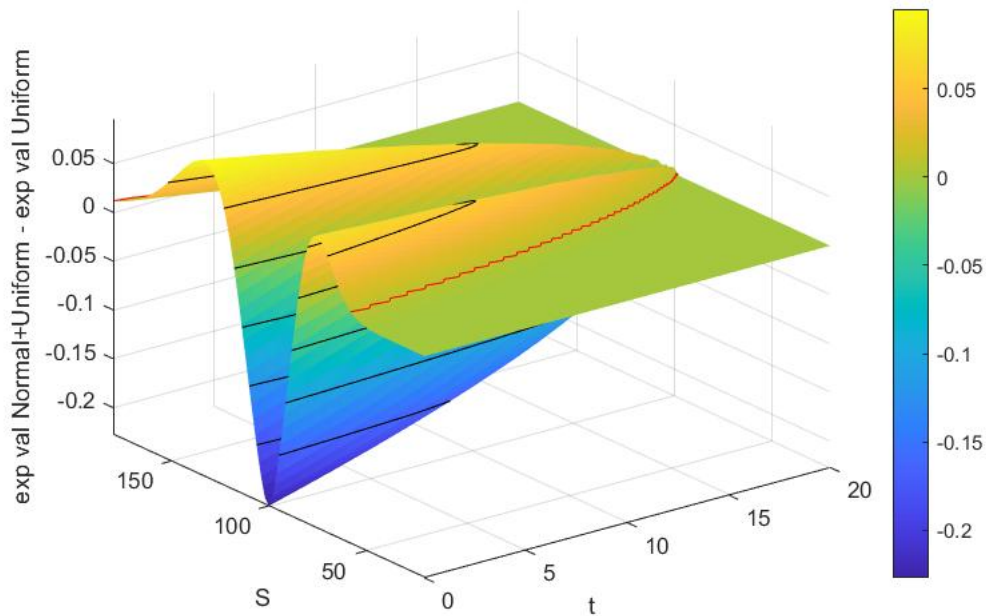


Figure 7.7: Absolute difference in the expected values for a European Call option for the models Σ_1 and Σ_3 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_1 circled in red.

Also the differences in variance between these three models, as illustrated in figures 7.8 and 7.9, are located in the smoothing area and the variance of model Σ_1 is closer to the variance for model Σ_0 than to the one for model Σ_3 .

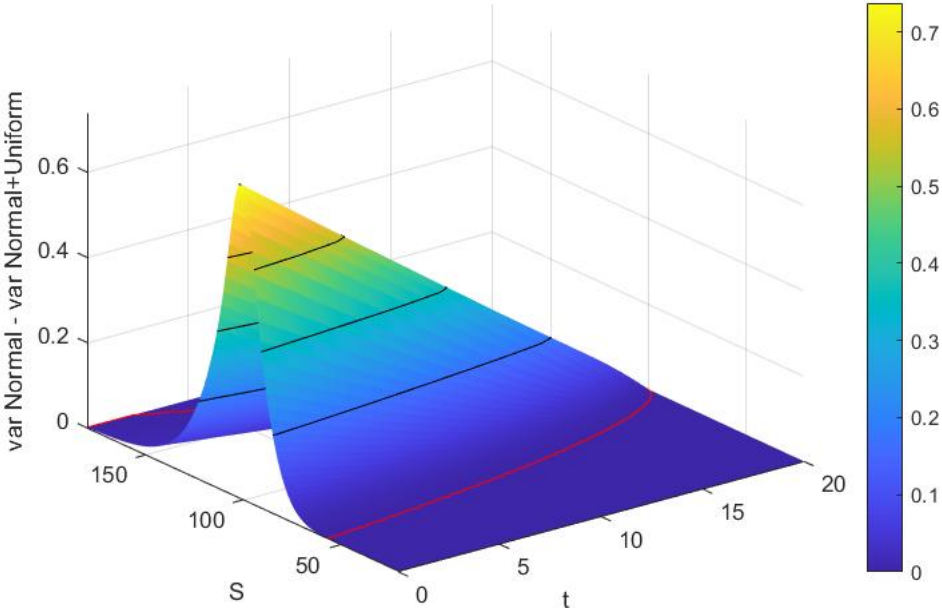


Figure 7.8: Absolute difference in variance for a European Call option for the models Σ_0 and Σ_1 with $T = 20, strike = 100, K = 1, N = 5, M_\zeta = 200, N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_0 circled in red.

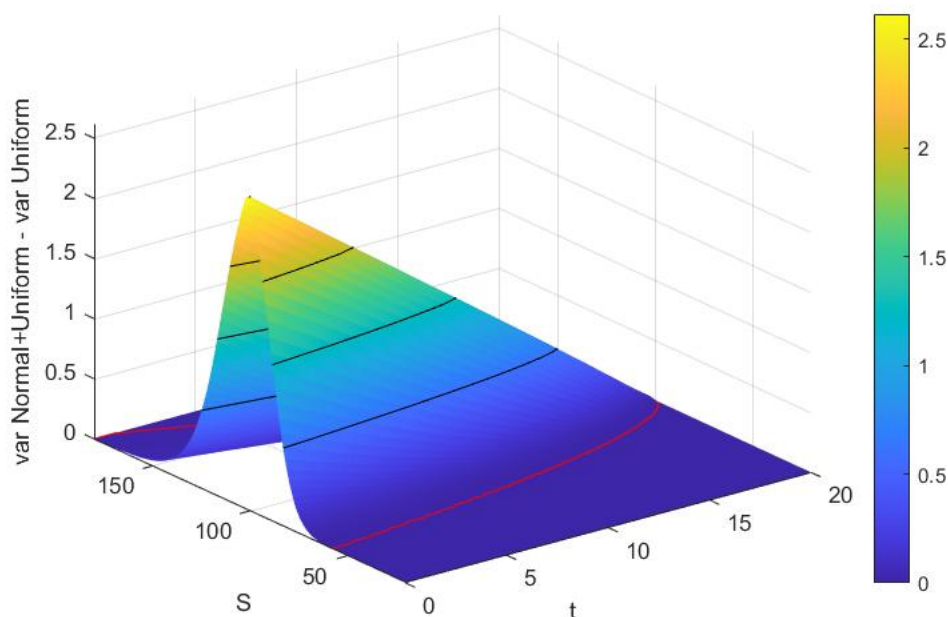


Figure 7.9: Absolute difference in variance for a European Call option for the models Σ_1 and Σ_3 with $T = 20$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta = 200$, $N_\tau = 319$ with contour lines at quarters of the maximum absolute difference and smoothing area for Σ_1 circled in red.

In summary, the experiment shows that volatility models of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$ lead to SG solutions that 'lie between' the solutions for volatility depending on Θ or Δ only that have the same mean and variance of the volatility. 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, vice versa.

Comparison to real market data:

The model will now be compared to market prices of a European Call option. Consider again the end of day values of a European Call option on the DAX index from January 7th 2019 to September 20th 2019 with strike price $strike = 100$ and maturity $T = 180$ days ¹ as in chapter 6.

In the following, a volatility model of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$ for independent random variables Θ and Δ , where Θ is standard normal distributed and Δ follows a uniform distribution on $[-0.5, 0.5]$, is fitted to the data. To do this, the implied volatilities $\sigma_{impl}^{(i)}$, $i = 1, \dots, 180$, were calculated by formula 2.3.

Consider at first σ_{00} , which coincides with the expected value of $\Sigma(\Theta, \Delta)$. By the law of large numbers, see e.g. [Kre05] Satz 12.4, the mean of the implied volatilities converges

¹The values were obtained from <https://www.finanzen.net/>.

to σ_{00} as the number of implied volatilities converges to infinity. This justifies the choice $\sigma_{00} = \text{mean}(\sigma_{impl}^{(i)}, i = 1, \dots, 180)$.

The values of σ_{10} and σ_{01} are derived by a constrained maximum likelihood approach as follows: As constraint, the variance of $\Sigma(\Theta, \Delta)$ given by

$$\text{Var}(\Sigma(\Delta, \Theta)) = \text{Var}(\sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta) = \sigma_{10}^2 \text{Var}(\Theta) + \sigma_{01}^2 \text{Var}(\Delta) = \sigma_{10}^2 \cdot 1 + \sigma_{01}^2 \cdot \frac{1}{12}$$

should coincide with the empirical variance S_{180}^2 of the implied volatilities, which is again motivated by the law of large numbers. Hence, one can represent σ_{01} by σ_{10} in the following form

$$\sigma_{01} = \sqrt{12(S_{180}^2 - \sigma_{10}^2)}.$$

Thus, we search for $\sigma_{10} \in (0, \sqrt{S_{180}^2})$ that maximizes the probability that 180 independent realizations $\Sigma(\Theta_1, \Delta_1), \dots, \Sigma(\Theta_{180}, \Delta_{180})$ of $\Sigma(\Theta, \Delta)$ attain the values $\sigma_{impl}^{(1)}, \dots, \sigma_{impl}^{(180)}$. Since $\Sigma(\Theta, \Delta)$ is continuous, this is done by maximizing the joint density $f_{\sigma_{10}}$ of $\Sigma(\Theta_1, \Delta_1), \dots, \Sigma(\Theta_{180}, \Delta_{180})$ evaluated in the point $\sigma_{impl}^{(1)}, \dots, \sigma_{impl}^{(180)}$ in σ_{01} . Denoting the density of $a\Theta + \sqrt{12(S_{180}^2 - a^2)}\Delta$ by g_a , this joint density can be rewritten by independence of the realizations as

$$f_{\sigma_{10}}(\sigma_{impl}^{(1)}, \dots, \sigma_{impl}^{(180)}) = \prod_{i=1}^{180} g_{\sigma_{10}}(\sigma_{impl}^{(i)} - \sigma_{00}).$$

The density g_a can be calculated as the convolution of the density $h_1(x) = 1/(\sqrt{2\pi}a)e^{-x^2/(2a^2)}$, $x \in \mathbb{R}$, of the random variable $a\Theta$ and the density $h_2(x) = 1/\sqrt{12(S_{180}^2 - a^2)}$, $x \in [-0.5\sqrt{12(S_{180}^2 - a^2)}, 0.5\sqrt{12(S_{180}^2 - a^2)}]$, of the random variable $\sqrt{12(S_{180}^2 - a^2)}\Delta$, see e.g. [Kre05] Satz 11.8, by

$$\begin{aligned} g_a(x) &= \int_{t \in \mathbb{R}} h_1(t)h_2(x-t) dt = \int_{x-\frac{1}{2}\sqrt{12(S_{180}^2 - a^2)}}^{x+\frac{1}{2}\sqrt{12(S_{180}^2 - a^2)}} h_1(t) \frac{1}{\sqrt{12(S_{180}^2 - a^2)}} dt \\ &= \frac{1}{\sqrt{12(S_{180}^2 - a^2)}} \left(H_1 \left(x + \frac{1}{2}\sqrt{12(S_{180}^2 - a^2)} \right) - H_1 \left(x - \frac{1}{2}\sqrt{12(S_{180}^2 - a^2)} \right) \right). \end{aligned}$$

Here, $H_1(x) := \int_{-\infty}^x h_1(t) dt$ denotes the cumulative distribution function of the normal distribution with mean 0 and standard deviation a which is available in most calculation software.

In summary, we choose

$$\begin{aligned}
 \sigma_{00} &= \text{mean}(\sigma_{impl}^{(i)}, i = 1, \dots, 180), \\
 \sigma_{10} &= \underset{a \in (0, \sqrt{S_{180}^2})}{\operatorname{argmax}} f_a(\sigma_{impl}^{(1)}, \dots, \sigma_{impl}^{(180)}) = \underset{a \in (0, \sqrt{S_{180}^2})}{\operatorname{argmax}} \prod_{i=1}^{180} g_a(\sigma_{impl}^{(i)} - \sigma_{00}) \\
 &= \underset{a \in (0, \sqrt{S_{180}^2})}{\operatorname{argmax}} \prod_{i=1}^{180} \frac{1}{\sqrt{12(S_{180}^2 - a^2)}} \left(H_1(\sigma_{impl}^{(i)} - \sigma_{00} + \frac{1}{2}\sqrt{12(S_{180}^2 - a^2)}) \right. \\
 &\quad \left. - H_1(\sigma_{impl}^{(i)} - \sigma_{00} - \frac{1}{2}\sqrt{12(S_{180}^2 - a^2)}) \right), \\
 \sigma_{01} &= \sqrt{12(S_{180}^2 - \sigma_{10}^2)}.
 \end{aligned} \tag{7.13}$$

Applying these formulas to our data leads to the volatility model

$$\Sigma(\Theta, \Delta) = 0.2292 + 0.1126\Theta + 0.0115\Delta. \tag{7.14}$$

The fitted density of $\Sigma(\Theta, \Delta)$ is shown in figure 7.10 together with a histogram density estimator.

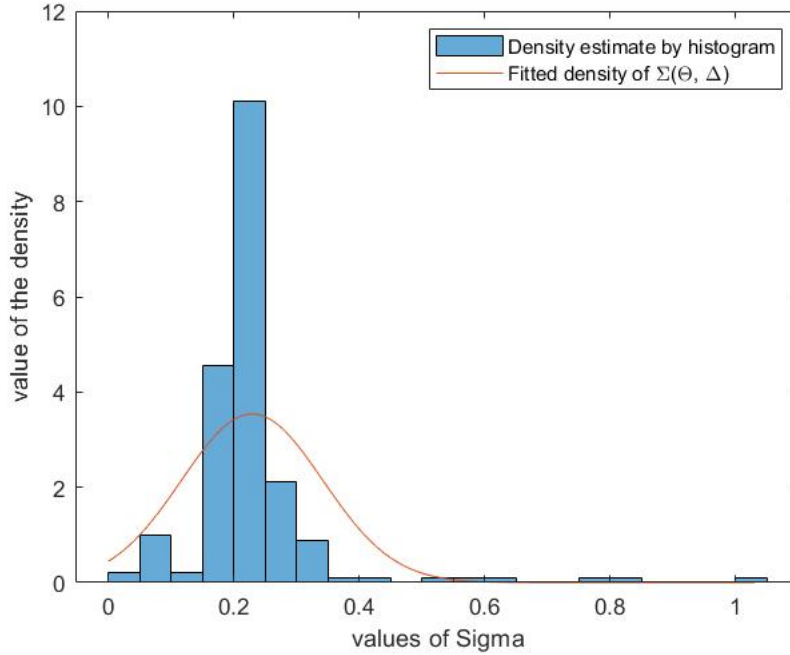


Figure 7.10: Histogram density estimator of $\Sigma(\Theta, \Delta)$ derived by the implied volatilities together with the density of $\Sigma(\Theta, \Delta)$ fitted to the implied volatilities by maximum likelihood.

With volatility model 7.14, one can compute the SG solution. The truncation number $N = 5$ and the numbers of grid points $M_\zeta = 200$ and $N_\tau = 678$ were chosen to obtain a

stable scheme for the parabolic system of equations 7.9.

Figure 7.11 shows the market prices and the expected value of the SG solution as well as the range expected value plus/minus standard deviation. A more detailed plot of those graphs for the last 55 days of the option is given in figure 7.12. Again one observes that the expected value of the SG solution is very close to the data in these days but slightly above the data at earlier times. However, the data is always in the range expected value plus/minus standard deviation, as one would expect from stochastic theory.

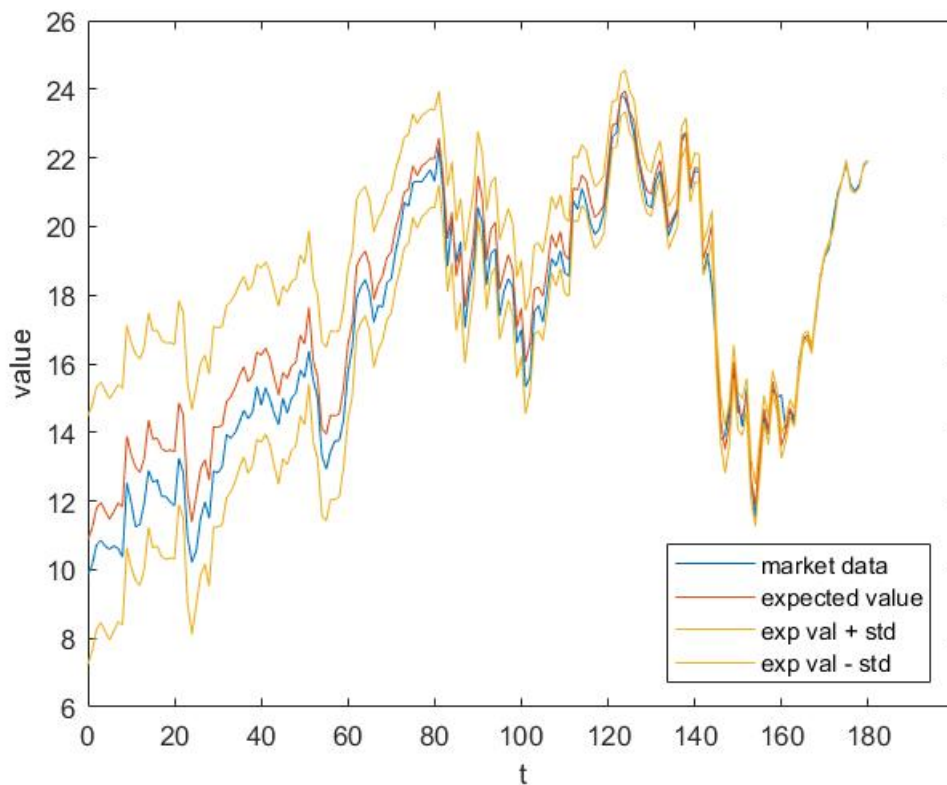


Figure 7.11: Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation.

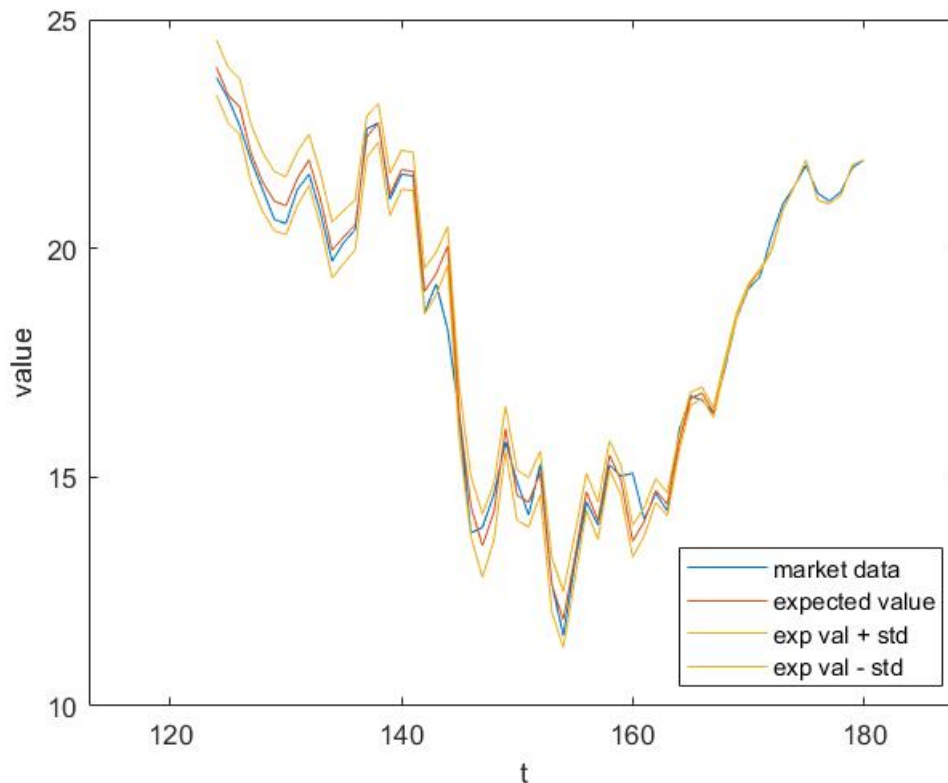


Figure 7.12: Market values of the option together with the expected value of the SG solution and the range expected value plus minus standard deviation for the last 55 days.

A comparison of the SG solution to the market data and the solution to the deterministic Black Scholes equation 2.2, where the mean of the implied volatilities was taken as a value for the volatility, is given in figure 7.13. This figure shows, that also the deterministic solution 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.

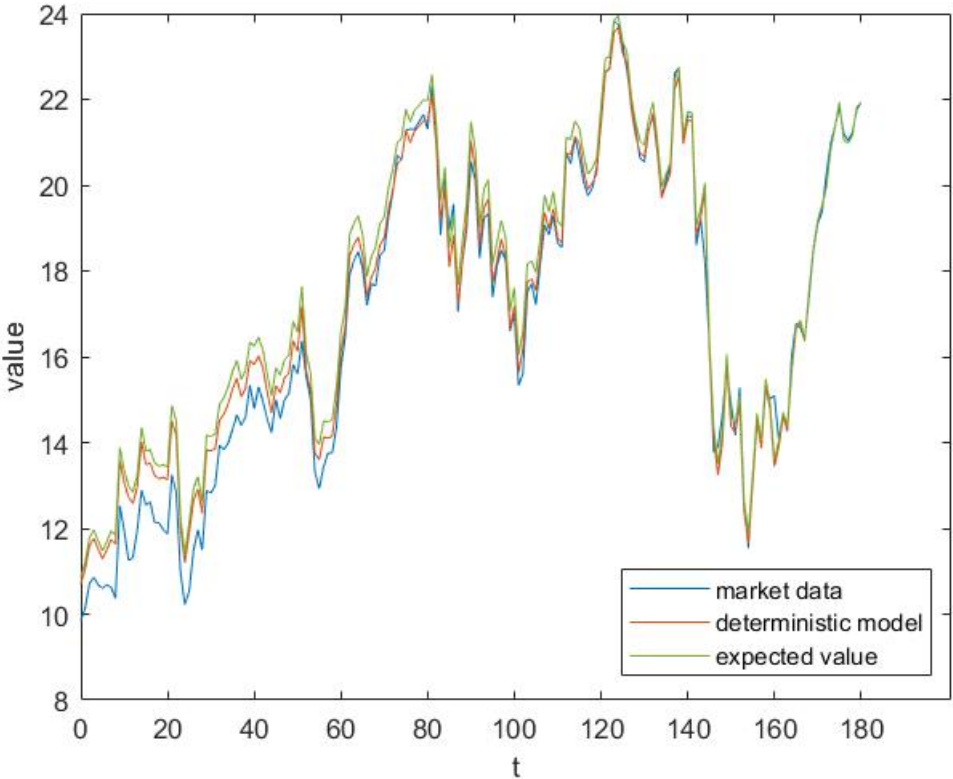


Figure 7.13: Market values of the option together with the expected value of the SG solution and deterministic solution.

Chapter 8

Integration of a Bi-Fidelity approach

In the previous chapter, an explicit finite difference scheme for the computation of system

$$0 = \frac{\partial v_\delta^N(S, t)}{\partial t} + \frac{1}{2} S^2 \sum_{\substack{\alpha, \beta, \gamma \in \mathbb{N}_0^L, \\ |\alpha|, |\beta| \leq K, \\ |\gamma| \leq N}} \sigma_\alpha \sigma_\beta \frac{\partial^2 v_\gamma^N(S, t)}{\partial S^2} M_{\alpha\beta\gamma\delta} + rS \frac{\partial v_\delta^N(S, t)}{\partial S} - r v_\delta^N(S, t) \quad (7.9)$$

for $\delta \in I = \{(\delta_1, \dots, \delta_L) \in \mathbb{N}_0^L \mid \delta_1 + \dots + \delta_L \leq N\}$ was presented. Here, the number of random variables, the volatility depends on, was denoted by $L \in \mathbb{N}$ and the truncation number of the gPC expansion of V by $N \in \mathbb{N}$. Recall that for $L = 2$, one already had $|I| = (N + 1)(N + 2)/2$ gPC coefficients of V and hence $(N + 1)(N + 2)/2$ coupled equations. This number of equations $|I|$ and with it the computational cost rapidly increase as N or L increases.

This becomes a problem especially if one wants to compute the SG solutions for many options at a time. A solution to this problem is given by applying a Bi-Fidelity approach, if system 7.9 has to be solved many times for the same type of option (e.g. European Call options) with the same maturity T and interest rate r , but varying coefficients σ_α of the volatility model $\Sigma(\Theta_1, \dots, \Theta_L) = \sum_{\alpha \in \mathbb{N}_0^L} \sigma_\alpha \bar{p}_\alpha(\Theta_1, \dots, \Theta_L)$. A situation like this arises for instance when comparing financial derivatives of the same type and maturity date, but with different underlying stochastic assets, compare e.g. the mechanism of deriving the gPC coefficients of $\Sigma(\Theta, \Delta)$ in 7.13.

In this section, at first the general procedure of applying a Bi-Fidelity approach will be explained. Afterwards, it will be applied to our problem 7.9 and the numerical implementation be explained. Finally, some numerical examples will be given comparing the Bi-Fidelity solution to the high fidelity solution.

In literature, the method is frequently described for uncertainty quantification via Stochastic Collocation methods, see e.g. [Zhu14] and [Nar14] for the general procedure and [LZ20] for an application to the multi-scale Boltzmann equation. However, the same procedure can be applied to equations derived by a Stochastic Galerkin method, if one takes care of the classification of the random variable, as it will become clear in section 8.2.

8.1 General Procedure

The general procedure of applying a Bi-Fidelity approach for solving a PDE will be explained in this section, loosely following [Nar14] and [Zhu14].

Consider the general initial value problem depending on a possibly multidimensional random variable $\Xi : \Omega \rightarrow \mathbb{D}$

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t, \Xi) &= \mathcal{L}(u)(x, t, \Xi), \quad x \in D, t \in (0, T], \\ \mathcal{B}(u)(x, t, \Xi) &= 0, \quad x \in \partial D, t \in (0, T], \\ u(x, 0, \Xi) &= u_0(t, \Xi), \quad t \in (0, T], \end{aligned}$$

for some spacial domain D , a differential operator \mathcal{L} , a boundary operator \mathcal{B} and some initial conditions given by u_0 .

In order to apply a Bi-Fidelity approach, one needs to define two different models describing the solution of this initial value problem. The solution one is actually interested in, that is expensive to calculate, will be called high fidelity solution and denoted by u^H . Then, a second model that is much cheaper to compute, but one does not trust the solution as much as the accurate high fidelity solution, is needed. However, to obtain information about the high fidelity solution, the low fidelity model has to be similar enough to the high fidelity model as explained later. The solution corresponding to the low fidelity model will be called the low fidelity solution and be denoted by u^L .

One could for instance choose the solution of an easier but similar PDE, as it was done in [LZ20], or a less accurate numerical scheme by using lower order methods or a coarser grid as low fidelity solution.

Below, $u^L(\xi)$ with $\xi \in \mathbb{D}$ stands for the function $u^L(\cdot, \cdot, \xi)$. Note that each $u^L(\xi)$ is the (approximate) solution to a deterministic PDE and therefore lies in a Hilbert space $U^L := \text{span}(u^L(\xi) \mid \xi \in \mathbb{D})$ with corresponding inner product $\langle \cdot, \cdot \rangle^L$ and induced norm $\| \cdot \| ^L$. The same notation is used for the high fidelity solution.

The aim of applying the Bi-Fidelity procedure is to approximate the high fidelity solution in an arbitrary but fixed value $z \in \mathbb{D}$ of the random variable Ξ by information from the low fidelity solution in z together with some stored high fidelity solutions to certain Ξ values z_1, \dots, z_A , where $A \in \mathbb{N}$ is the number of high fidelity computations one can afford. To obtain a good approximation for as many $z \in \mathbb{D}$ as possible, the choice of z_1, \dots, z_A is important. The following procedure describes the selection of the values z_1, \dots, z_A and the corresponding low and high fidelity solutions $u^L(z_i), u^H(z_i), i = 1, \dots, A$ in three steps:

Step 1: Compute the low fidelity solutions in many Ξ values

Let B denote the number of affordable low fidelity simulations with $B \gg A$. A set $Y = \{y_i \in \mathbb{D} \mid i = 1, \dots, B\}$ of $B \in \mathbb{N}$ possible values of Ξ is chosen such that the y_i 'cover' \mathbb{D} well enough. Usually, one either takes Monte Carlo sample points, i.e. independent realizations of Ξ , or the points of a deterministic grid on \mathbb{D} . The low fidelity solution

$u^L(y_i)$ is computed in every point y_i .

Step 2: Select important values of Ξ

Since it would be too costly to compute the high fidelity solution in every point y_i , $i = 1, \dots, B$, one has to detect the A most important points. For a subset $\hat{Y} \subset Y$, define the spanned solution space $U^L(\hat{Y}) := \text{span}(u^L(\hat{y}) \mid \hat{y} \in \hat{Y})$. With this notation, one searches the points whose solutions differ the most in the sense that they span the largest subspace of $U^L(Y)$. This is guaranteed by the choice

$$z_{i+1} := \underset{y \in Y}{\operatorname{argmax}} d^L(u^L(y), U^L(Z_i)), \quad i = 0, \dots, A - 1, \quad (8.1)$$

see e.g. [Nar14], where $d^L(v, U) := \inf_{u \in U} \|v - u\|^L$ is the distance of a point $v \in U^L(Y)$ to the set $U \subset U^L(Y)$ induced by the norm $\|\cdot\|^L$ and $Z_i := \{z_1, \dots, z_i\}$ is the set of the first i important values z_1, \dots, z_i with $Z_0 := \emptyset$.

Step 3: Compute the high fidelity solution in the important points

The high fidelity solution is computed in the important values z_1, \dots, z_A of Ξ . By the assumed similarity of the high and low fidelity solutions, the $u^H(z_i)$ are expected to also approximately span the largest subspace of $U^H(Y)$ in the same sense as in Step 2.

These steps have to be performed only once, since the results, the high fidelity solutions $u^H(z_i)$ and low fidelity solutions $u^L(z_i)$ in the z_i , $i = 1, \dots, A$, are stored. Thus, they are called 'Offline' steps in [LZ20]. They build the data basis for later computations of an approximate high fidelity solution in a value $z \in \mathbb{D}$ of Ξ in the so called 'Online' steps.

Assume therefore, a value $z \in \mathbb{D}$ of Ξ is given and the corresponding high fidelity solution $u^H(z)$ shall be derived. To save computational cost, one computes an approximation $u^{BF}(z)$, the Bi-Fidelity solution, instead of $u^H(z)$. Using the data stored from the Offline steps, one can calculate $u^{BF}(z)$ as follows:

Step 1: Compute the low fidelity solution in the point z

At first, the low fidelity solution $u^L(z)$ for the value z of Ξ is computed. Note that this was assumed to be much cheaper than computing the high fidelity solution.

Step 2: Project the low fidelity solution onto $U^L(Z_A)$

The low fidelity solution $u^L(z)$ is then projected onto $U^L(Z_A)$ leading to the projection law

$$u^L(z) \approx P_{U^L(Z_A)} u^L(z) = \sum_{i=1}^A c_i u^L(z_i)$$

for some constants $c_i := c_i(z) \in \mathbb{R}$ depending on z . Here $P_X x := P_X(x)$ denotes the orthogonal projection of an element x of some space onto the subspace X .

Step 3: Calculate the Bi-Fidelity solution

By applying the same projection law to the high fidelity solutions, one obtains the Bi-Fidelity solution

$$u^{BF}(z) := \sum_{i=1}^A c_i u^H(z_i), \quad (8.2)$$

where the c_i are the coefficients derived in the last step. For sufficient similarity between the high and low fidelity models, a good choice of y_1, \dots, y_B and A large enough, this should lead to a good approximation of $u^H(z)$.

A bound on the approximation error $\|u^H(z) - u^{BF}(z)\|^H$ will be given in Theorem 10 by similar arguments as in [Nar14]. To investigate the approximation error, at first the projection error $\|u^H(z) - P_{U^H(Z_A)} u^H(z)\|^H$ and the interpolation error $\|P_{U^H(Z_A)} u^H(z) - u^{BF}(z)\|^H$ will be considered in the following lemmas. Define for a finite subset $\tilde{Y} \subset \mathbb{D}$

$$\begin{aligned} z_{i+1}^L &:= \operatorname{argmax}_{y \in \tilde{Y}} d^L(u^L(y), U^L(Z_i^L)), \\ z_{i+1}^H &:= \operatorname{argmax}_{y \in \tilde{Y}} d^H(u^H(y), U^H(Z_i^H)), \end{aligned} \quad (8.3)$$

$i = 0, \dots, A - 1$, with $Z_i^L = \{z_1, \dots, z_i\}$ and $Z_0^L := \emptyset$, respectively for the high fidelity model.

Lemma 2. *Let A be even and assume that there exist constants $0 < K_L, K_H \leq 1$ such that the low and high fidelity model satisfy*

$$\begin{aligned} d^H(u^H(z), U^H(\mathcal{E})) &\geq K^L d^L(u^L(z), U^L(\mathcal{E})) \quad \text{and} \\ d^L(u^L(z), U^L(\mathcal{E})) &\geq K^H d^H(u^H(z), U^H(\mathcal{E})) \end{aligned} \quad (8.4)$$

for all $z \in \mathbb{D}$ and every finite subset \mathcal{E} of \mathbb{D} . Then one has for any finite $\tilde{Y} \subset \mathbb{D}$

$$\max_{z \in \tilde{Y}} \|u^H(z) - P_{U^H(Z_A^L)} u(z)\| \leq \frac{\sqrt{2}}{K^L K^H} \sqrt{d_{A/2}(u(\tilde{Y}))}, \quad (8.5)$$

where

$$d_n(X) := \inf_{\substack{U \text{ subspace of } U^H \\ \dim(U)=n}} \sup_{x \in X} d^H(x, U)$$

denotes the Kolmogorov n -width.

Condition 8.4 tells us that the distances of the solutions are comparable for high and low fidelity model.

Proof. (adapted from [Nar14] Theorem 4.2). By the assumptions on the distances and the definition of z_{i+1}^L , the following inequality is true for all $i = 0, \dots, A - 1$

$$\begin{aligned} d^H(u^H(z_{i+1}^L), U^H(Z_i^L)) &\geq K^L d^L(u^L(z_{i+1}^L), U^L(Z_i^L)) = K^L \max_{z \in \tilde{Y}} d^L(u^L(z), U^L(Z_i^L)) \\ &\geq K^L \max_{z \in \tilde{Y}} K^H d^H(u^H(z), U^H(Z_i^L)) \\ &= K^L K^H \max_{z \in \tilde{Y}} d^H(u^H(z), U^H(Z_i^L)). \end{aligned}$$

Since $0 < K^L K^H \leq 1$ holds, this shows that the point selection procedure by equation 8.1 is a weak greedy procedure as defined in [DPW13] equation 1.8. Therefore, applying Corollary 3.3 (i) in the same paper yields equation 8.5. \square

To bound the interpolation error $\|P_{U^H(Z_A)} u^H(z) - u^{BF}(z)\|^H$ in $z \in \mathbb{D}$, rather technical conditions are needed. For their description, define the following matrices and vectors:

- The Gramian matrix $\mathbf{G}^L \in \mathbb{R}^{A \times A}$ corresponding to the inner product $\langle \cdot, \cdot \rangle^L$ of to the Hilbert space U^L is given by $\mathbf{G}^L[i-1, j-1] = \langle u^L(z_i), u^L(z_j) \rangle^L$, $i, j = 1, \dots, A$. Note that \mathbf{G}^L is invertible by choice of z_i .
- The vector $\mathbf{f}^L \in \mathbb{R}^A$ is defined by $\mathbf{f}^L[i-1] = \langle u^L(z_i), u^L(z) \rangle$, $i = 1, \dots, A$.
- The matrix $\mathbf{G}^H \in \mathbb{R}^{A \times A}$ and the vector $\mathbf{f}^H \in \mathbb{R}^A$ are defined analogously for the high fidelity solutions. Since \mathbf{G}^H is semi positive definite, its semi positive root $\sqrt{\mathbf{G}^H} = \mathbf{O}\sqrt{\mathbf{\Lambda}}\mathbf{O}$ can be explained by the spectral decomposition $\mathbf{G}^H = \mathbf{O}\mathbf{\Lambda}\mathbf{O}^T$ for an orthogonal matrix \mathbf{O} and a diagonal matrix $\mathbf{\Lambda}$. Furthermore, denote the Moore Penrose pseudo inverse of \mathbf{G}^H by $(\mathbf{G}^H)^\dagger$ (see e.g. [BIG03] Chapter 1.1), since \mathbf{G}^H is not necessarily invertible.
- The orthogonal projection matrix onto the kernel of \mathbf{G}^H is denoted by \mathbf{Q} .

The Euclidean norm $\|\cdot\|$ is chosen as a vector norm.

Lemma 3 ([Nar14] Lemma 4.3). *Suppose that there exist small $\varepsilon_1, \varepsilon_2 > 0$ such that*

$$\left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger (\mathbf{f}^H - \mathbf{f}^L) \right\| \leq \varepsilon_1 \left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|$$

holds for all $z \in \mathbb{D}$ and the Gramian matrices satisfy

$$\left\| \sqrt{\mathbf{G}^H} (\mathbf{G}^L)^{-1} \sqrt{\mathbf{G}^H} - \mathbf{I}_A \right\| < \varepsilon_2. \quad (8.6)$$

Then the interpolation error in $z \in \mathbb{D}$ with $\|P_{U^H(Z_A)} u^H(z)\|^H > 0$ is bounded by

$$\|P_{U^H(Z_A)} u^H(z) - u^{BF}(z)\|^H \leq (\varepsilon_1 + \varepsilon_2 + \varepsilon_1 \varepsilon_2) \|P_{U^H(Z_A)} u^H(z)\|^H + \left\| \sqrt{\mathbf{G}^H} (\mathbf{G}^L)^{-1} \mathbf{Q} \mathbf{f}^L \right\|. \quad (8.7)$$

Formulated in an illustrative manner, condition 8.6 means that the projection coefficients of $P_{U^L(Z_A)} u^L(z)$ and of $P_{U^H(Z_A)} u^H(z)$ behave similarly, whereas condition 8.7 states that the Gramian matrices are similar enough. Details can be found in [Nar14] chapter 4.2.

Remark 3. Note that the last term on the right hand side of inequality 8.7 vanishes if \mathbf{G}^H is invertible, because then the orthogonal projection onto the kernel of \mathbf{G}^H is $\mathbf{Q} = \mathbf{0}_{A \times A}$. By the density of invertible matrices in the set of all quadratic matrices of a certain dimension, it is a very rare event that \mathbf{G}^H is not invertible, which neither occurred in our numerical experiments nor in those in [Nar14].

Proof. Denoting the projection coefficients of $P_{U^H(Z_A)}u^H(z)$ by c_i^H , i.e. $P_{U^H(Z_A)}u^H(z) = \sum_{i=1}^A c_i^H u^H(z_i)$, one obtains

$$\begin{aligned} (\|P_{U^H(Z)}u^H(z) - u^{BF}(z)\|^H)^2 &= \left(\left\| \sum_{i=1}^A c_i^H u^H(z_i^L) - c_i u^H(z_i^L) \right\|^H \right)^2 \\ &= \sum_{i,j=1}^A (c_i^H - c_i)(c_j^H - c_j) \langle u^H(z_i^L), u^H(z_j^L) \rangle^H \\ &= (\mathbf{c}^H - \mathbf{c})^T \mathbf{G}^H (\mathbf{c}^H - \mathbf{c}) = \|\sqrt{\mathbf{G}^H}(\mathbf{c}^H - \mathbf{c})\|^2. \end{aligned}$$

for $\mathbf{c}, \mathbf{c}^H \in \mathbb{R}^A$ with $\mathbf{c}^H[i-1] = c_i^H$ and $\mathbf{c}[i-1] = c_i$. Because of $\mathbf{f}[i-1] = \langle u^L(z_i), u^L(z) \rangle = \langle u^L(z_i), P_{U^L(Z_A)}u^L(z) \rangle$, one has

$$\mathbf{G}^L \mathbf{c} = \mathbf{f}^L. \quad (8.8)$$

In analogy, $\mathbf{G}^H \mathbf{c}^H = \mathbf{f}^H$ holds. Now consider $\sqrt{\mathbf{G}^H}(\mathbf{c}^H - \mathbf{c})$. Because the kernels of \mathbf{G}^H and $\sqrt{\mathbf{G}^H}$ coincide, one can rewrite $\mathbf{Q} = \mathbf{I}_A - \sqrt{\mathbf{G}^H} \left(\sqrt{\mathbf{G}^H} \right)^\dagger$, see e.g. [BIG03] Ex.42. Using this representation, a little calculation shows

$$\begin{aligned} \sqrt{\mathbf{G}^H}(\mathbf{c}^H - \mathbf{c}) &= \left(\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\sqrt{\mathbf{G}^H} - \mathbf{I}_A \right) \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \\ &\quad + \left(\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\sqrt{\mathbf{G}^H} - \mathbf{I}_A \right) \left(\sqrt{\mathbf{G}^H} \right)^\dagger (\mathbf{f}^L - \mathbf{f}^H) \\ &\quad + \left(\sqrt{\mathbf{G}^H} \right)^\dagger (\mathbf{f}^L - \mathbf{f}^H) + \sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\mathbf{f}^L. \end{aligned}$$

Inserting assumptions 8.6 and 8.7 yields

$$\begin{aligned} \|P_{U^H(Z)}u^H(z) - u^{BF}(z)\|^H &= \|\sqrt{\mathbf{G}^H}(\mathbf{c}^H - \mathbf{c})\|^H \leq \varepsilon_2 \left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|^H + \varepsilon_2 \varepsilon_1 \left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|^H \\ &\quad + \varepsilon_1 \left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|^H + \|\sqrt{\mathbf{G}^H}(\mathbf{G}^L)^{-1}\mathbf{Q}\mathbf{f}^L\|^H. \end{aligned}$$

To obtain inequality 8.7, it remains to show $\left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|^H = \|P_{U^H(Z_A)}u^H(z)\|^H$, which is by true because

$$\begin{aligned} (\|P_{U^H(Z_A)}u^H(z)\|^H)^2 &= \sum_{i,j=1}^A c_i^H c_j^H \langle u^H(z_i), u^H(z_j) \rangle^H = \mathbf{c}^H \mathbf{G}^H \mathbf{c}^H = \|\sqrt{\mathbf{G}^H} \mathbf{c}^H\|^2 \\ &= \left\| \left(\sqrt{\mathbf{G}^H} \right)^\dagger \mathbf{f}^H \right\|^2, \end{aligned}$$

where the last equality is obtained by inserting $\mathbf{c}^H = (\mathbf{G}^H)^\dagger \mathbf{f}^H$. \square

In summary, one can now formulate the following theorem on bounds of the Bi-Fidelity approximation error.

Theorem 10 (adapted from [Nar14], Theorem 4.4). *Assume that A is an even natural number, that there exist constants $0 < K_L, K_H \leq 1$ such that*

$$\begin{aligned} d^H(u^H(z), U^H(\mathcal{E})) &\geq K^L d^L(u^L(z), U^L(\mathcal{E})) \quad \text{and} \\ d^L(u^L(z), U^L(\mathcal{E})) &\geq K^H d^H(u^H(z), U^H(\mathcal{E})) \end{aligned}$$

for all $z \in \mathbb{D}$ and every finite subset \mathcal{E} of \mathbb{D} and that the assumptions of lemma 3 are satisfied for $\varepsilon_1, \varepsilon_2$. Then the Bi-Fidelity approximation is bounded in the following way

$$\begin{aligned} \|u^H(z) - u^{BF}(z)\|^H &\leq \frac{\sqrt{2}}{K^L K^H} \sqrt{d_{A/2}(u(Y))} + (\varepsilon_1 + \varepsilon_2 + \varepsilon_1 \varepsilon_2) \|P_{U^H(Z)} u^H(z)\|^H \\ &\quad + \left\| \sqrt{\mathbf{G}^H} (\mathbf{G}^L)^{-1} \mathbf{Q} \mathbf{f}^L \right\|. \end{aligned} \quad (8.9)$$

Proof. The triangle inequality shows that the approximation error can be decomposed into the projection error and the interpolation error considered above

$$\|u^H(z) - u^{BF}(z)\|^H \leq \|u^H(z) - P_{U^H(Z_A)} u^H(z)\|^H + \|P_{U^H(Z_A)} u^H(z) - u^{BF}(z)\|^H.$$

Inserting the derived estimates yields the desired upper bound for the approximation error. \square

To show convergence of the Bi-Fidelity solution to the high fidelity solution, the particular equation to be solved and the chosen models have to be investigated.

Nevertheless, let us consider the terms on the right hand side of equation 8.9 for the general model. For the first term, suppose that the Kolmogorov width $d_{A/2}(u(Y))$ becomes small if $A, B \rightarrow \infty$ and Y is chosen such that it is a good 'cover' of \mathbb{D} . If this is not the case, one cannot expect to find a good approximation of $u^H(z)$ for all z by considering finitely many data points $u^H(z_1), \dots, u^H(z_A)$. The third term vanishes, if \mathbf{G}^H is invertible by the same arguments as in remark 3. Invertibility of \mathbf{G}^H is for instance automatically given if one chooses the low fidelity model equal to the high fidelity model. Also the coefficient $(\varepsilon_1 + \varepsilon_2 + \varepsilon_1 \varepsilon_2)$ of the second term can be chosen equal to zero, if high and low fidelity model coincide, as conditions 8.6 and 8.7 show.

This way one can expect that for a good choice of the low fidelity model, the last two terms on the right hand side of equation 8.9 become small. If one can in addition show that they converge to zero as the low fidelity model converges to the high fidelity model and that the Kolmogorov n -width goes to zero as n becomes large, then inequality 8.9 establishes convergence of the Bi-Fidelity solution to the high fidelity solution.

8.2 Application to the Black Scholes equation with uncertain volatility and numerical implementation

As explained at the beginning of this chapter, computation of the truncated SG solution of the Black Scholes equation with uncertain volatility depending on $L \in \mathbb{N}$ random

variables 7.9

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

$\delta \in I = \{(\delta_1, \dots, \delta_L) \in \mathbb{N}_0^L \mid \delta_1 + \dots + \delta_L \leq N\}$, becomes costly if large truncation numbers N are chosen or if L is large. To reduce this cost in case one wants to compute the SG solution multiple times for the same option type with the same maturity T and interest rate r , a Bi-Fidelity approach is applied as explained in the last section. Therefore, high and low fidelity models have to be defined and the random variable Ξ has to be assigned. This will be done in the following and the numerical implementation of the procedure will be explained.

Define the high fidelity model as a high resolution numerical solution to 7.9 derived by the explicit finite difference scheme 5.6

$$\begin{aligned} \bar{v}(\zeta_m, \tau^{n+1}) = & \Delta\tau \left(\frac{1}{2} \zeta_m^2 (1 - \zeta_m)^2 \mathbf{A} \frac{\bar{v}(\zeta_{m+1}, \tau^n) - 2\bar{v}(\zeta_m, \tau^n) + \bar{v}(\zeta_{m-1}, \tau^n)}{(\Delta\zeta)^2} \right. \\ & \left. + r \zeta_m (1 - \zeta_m) \frac{\bar{v}(\zeta_{m+1}, \tau^n) - \bar{v}(\zeta_{m-1}, \tau^n)}{2\Delta\zeta} - r(1 - \zeta_m) \bar{v}(\zeta_m, \tau^n) \right) \\ & + \bar{v}(\zeta_m, \tau^n), \quad \text{for } m = 1, \dots, M_\zeta - 1, n = 0, \dots, N_\tau - 1, \end{aligned}$$

i.e. a high M_ζ^H and corresponding to that, for stability, a high N_τ^H are chosen. The low fidelity model is taken to be the same numerical solution on a coarse grid, i.e. with small M_ζ^L and N_τ^L . This is computationally cheaper, since on the one hand, by a smaller N_τ^L less iteration steps of the scheme are needed to reach the solution at maturity T , and on the other hand a smaller M_ζ^L yields less applications of scheme 5.6 at every single time step.

However, N_τ^L must not be chosen too small. Otherwise, due to stability problems there might not exist a large number of low fidelity solutions covering \mathbb{D} well enough. In addition, the calculation of the Bi-Fidelity solution might not work for many Ξ values, since the computation of the low fidelity solution in Online step 1 would be unstable.

Now, one has to assign the random variable Ξ . It does not coincide with $(\Theta_1, \dots, \Theta_L)$, since we want our solution and therefore also the Bi-Fidelity solution to be a random variable depending the Θ_i such that we can explore its behaviour. Instead, we suppose the distribution of $\Sigma(\Theta_1, \dots, \Theta_L)$ to change from calculation to calculation, as it would be the case when different underlying assets are considered. By truncation after total polynomial degree K , the gPC expansion of $\Sigma^K(\Theta_1, \dots, \Theta_L)$ is of the form

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

A variation in the distribution of the volatility therefore means a variation in at least one of the gPC coefficients σ_α , $|\alpha| \leq K$, since the distribution of the Θ_i is supposed to stay

the same. Hence, the random variable Ξ in this case describes the values of all σ_α with $|\alpha| \leq K$.

Below, the numerical implementation of the Bi-Fidelity method for calculating the SG solution of the Black Scholes equation is explained, where the volatility is assumed to depend on $L = 2$ random variables Θ_1, Θ_2 for a better readability. In case of dependence on more than two random variables, the programs can easily be adapted. The truncation number $K = 1$ is chosen such that the random variable Ξ represents the gPC coefficients σ_{00}, σ_{10} and σ_{01} . The values of Ξ are hence assumed to be of the form

$$z = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & 0 \end{pmatrix} \quad (8.10)$$

as in 7.12. This definition enables us to directly insert z as a volatility model in the computation of $\bar{\mathbf{v}}$, see programs 9 and 10.

Note that the actual computational effort lies in the calculation of the transformed system of equations 7.11 for $\bar{\mathbf{v}} = (\bar{v}_{\phi(0)}^N, \dots, \bar{v}_{\phi(|I|-1)}^N)^T$

$$\frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \tau} = \frac{1}{2} \zeta^2 (1 - \zeta)^2 \mathbf{A} \frac{\partial^2 \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta^2} + r \zeta (1 - \zeta) \frac{\partial \bar{\mathbf{v}}(\zeta, \tau)}{\partial \zeta} - r (1 - \zeta) \bar{\mathbf{v}}(\zeta, \tau), \quad (8.11)$$

$$\zeta \in (0, 1), \tau \in [0, T].$$

This is why the Bi-Fidelity approach will be applied on the calculation of $\bar{\mathbf{v}}$. In this way, a transformation back to the original variables \mathbf{v}^N, S and t has to be done only once for the Bi-Fidelity solution. If one applied the Bi-Fidelity procedure on the original system, the transformation back would have to be performed for every low fidelity solution from Offline Step 1, increasing the computational cost of the Offline part without providing advantages like higher accuracy.

The numerical implementation, as suggested by [Nar14], will be explained in detail for each Step from section 8.1.

Step 1: Compute the low fidelity solutions in many Ξ values

The set of low fidelity sample points $Y = \{y_i \in \mathbb{D} \mid i = 1, \dots, B\}$ is chosen as a structured grid on the set of possible values of $\sigma_{00}, \sigma_{10}, \sigma_{01}$. To do this, assumptions on the range of the gPC coefficients have to be made, i.e. one has to choose finite intervals such that $\sigma_{00} \in [a_{00}, b_{00}], \sigma_{10} \in [a_{10}, b_{10}], \sigma_{01} \in [a_{01}, b_{01}]$.

This can for example be done by experiments, i.e. calculating $\sigma_{00}, \sigma_{10}, \sigma_{01}$ for some of the later interesting stochastic assets by formula 7.13 and taking the minimum (maximum) of these calculated values for σ_{00} minus (plus) some safety margin as lower (upper) bound for σ_{00} , accordingly for σ_{10} and σ_{01} .

As an alternative, one can think of possible values of σ_{00} inspired by experiments e.g., and choose bounds of σ_{10} and σ_{01} such that the variance of $\Sigma(\Theta_1, \Theta_2)$ is bounded by $\sigma_{00}/2$ or another predefined value.

The grid points y_k were stored in a tensor $\hat{\Sigma}$, where

$$\hat{\Sigma}[:, :, k - 1] = y_k = \begin{pmatrix} \sigma_{00}^{(k)} & \sigma_{01}^{(k)} \\ \sigma_{10}^{(k)} & 0 \end{pmatrix}$$

describes the volatility model $\Sigma(\Theta_1, \Theta_2) = \sigma_{00}^{(k)} + \sigma_{10}^{(k)}\Theta_1 + \sigma_{01}^{(k)}\Theta_2$. For every of these volatility models y_k , the low fidelity solution $\bar{\mathbf{v}}^{L(k)}$, i.e. the numerical solution on a coarse grid $(\zeta_m^L, \tau^{L,n})_{m=0, \dots, M_\zeta^L, n=0, \dots, N_\tau^L}$, was computed. The solution is stored in a tensor $\hat{\mathbf{v}}^L$ by $\hat{\mathbf{v}}^L[i, m, n, k - 1] = \bar{v}_{\phi(i)}^{L(k)}(\zeta_m^L, \tau^{L,n}) = \bar{\mathbf{v}}^{L(k)}[i, m, n]$.

To save computational cost in this step, the programs 9 and 10 for calculating the numerical solution \mathbf{v}^L and the coupling matrix \mathbf{A} were slightly modified:

Program 9 was changed such that Galerkin multiplication tensors are not calculated at every simulation. Instead they are given as input variables.

In program 10, the transformation back of the variables from $\bar{\mathbf{v}}$ to \mathbf{v} was left out, because the Bi-Fidelity method is applied to $\bar{\mathbf{v}}$. In addition, the Galerkin multiplication tensors were added as input variables to pass them over to the calculation of \mathbf{A} by the modified program 9 and the initial condition for $\bar{\mathbf{v}}$ is not calculated within the program but given as an input variable, too. This saves computational cost, since the initial conditions for options of the same type coincide and the Galerkin multiplication for fixed K and N are always the same. Furthermore, the check for parabolicity and stability was altered. Instead of displaying an error, a return variable *stable* equals 0, if the equation is not parabolic or the scheme is not stable, and 1 else. The variable *stable* is necessary to determine whether a volatility model is a suitable low fidelity sample point or not.

Algorithm 11 Bi-Fidelity Offline step 1: Compute the low fidelity solutions in many Ξ values

Input: Truncation numbers K, N of the truncated gPC expansions of V and Σ , interest rate r , maturity of the option T and numbers of discretization points in the low fidelity model M_ζ^L for ζ and N_τ^L for τ .

Output: Tensor $\hat{\Sigma} \in \mathbb{R}^{2 \times 2 \times B}$ of low fidelity sample points with $\hat{\Sigma}[:, :, k - 1] = y_k$ is the k th volatility model, tensor of corresponding low fidelity solutions $\hat{\mathbf{v}}^L \in \mathbb{R}^{|I| \times (M_\zeta^L + 1) \times (N_\tau^L + 1) \times B}$.

Compute the Galerkin multiplication tensors as well as the initial condition for the option only once to save computational cost.

Derive the Galerkin multiplication tensors $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$ corresponding to the distribution of Θ_1 and Θ_2 by program 2.

initcondlow = zeros($(N + 1) * (N + 2) / 2, M_\zeta^L + 1$);

initcondlow(1, :) = max($2 / M_\zeta^L * (0 : M_\zeta^L) - 1, 0$);

Code continues on the next page.

Compute $\hat{\Sigma}$ and $\hat{\mathbf{v}}^L$ assuming $\sigma_{00} \in [a_{00}, b_{00}]$, $a_{00}, b_{00} \in \mathbb{R}^+$ and a variance of $\Sigma(\Theta_1, \Theta_2)$ less than $\sigma_{00}/2$. The linear grid has step width *step* in every direction.

$\hat{\Sigma} = []$;

$\hat{\mathbf{v}}^L = []$;

for $\sigma_{00} = a_{00} : \textit{step} : b_{00}$ **do**

for $\sigma_{10} = 0 : \textit{step} : \sqrt{\sigma_{00}/2}$ **do**

for $\sigma_{01} = 0 : \textit{step} : \sqrt{12 * (\sigma_{00}/2 - \sigma_{10}^2)}$ **do**

sigma = $[\sigma_{00} \ \sigma_{01}; \sigma_{10} \ 0]$;

 Calculate $\bar{\mathbf{v}}^L$ by modified program 10 as described above with volatility model explained by **sigma** and $M_\zeta^L + 1$ grid points in ζ and $N_\tau^L + 1$ in τ .

 If the computation was stable and the equation was parabolic, the values **sigma** and $\bar{\mathbf{v}}$ are added to the tensors $\hat{\Sigma}$ and $\hat{\mathbf{v}}^L$.

if *stable* == 1 **then**

$\hat{\Sigma} = \text{cat}(3, \hat{\Sigma}, \mathbf{sigma})$;

$\hat{\mathbf{v}}^L = \text{cat}(4, \hat{\mathbf{v}}^L, \bar{\mathbf{v}}^L)$;

end if

end for

end for

end for

Step 2: Select important grid points

The algorithm for the selection of important low fidelity sample points as proposed in [Nar14] algorithm 1 was adapted to our problem.

At the k th selection step, the vector \mathbf{d} saves the squared distances $(d^L(\bar{v}^L(y), \bar{V}^L(Z_k)))^2$, where $Z_k := \{z_1, \dots, z_k\}$ is the set of the first k important sample points z_1, \dots, z_k and $\bar{V}^L(\hat{Y}) := \text{span}\{\bar{v}^L(y) \mid y \in \hat{Y}\}$ denotes the solution space spanned by the $\hat{Y} \subset Y$. Note that selection rule 8.1 defines z_{k+1} as the y with maximum corresponding $d^L(\bar{v}^L(y), \bar{V}^L(Z_k))$. However, by taking the maximum over the values $(d^L(\bar{v}^L(y), \bar{V}^L(Z_k)))^2$ of \mathbf{d} in each step $k = 1, \dots, A$ one obtains the same values due to the strict monotonicity of $x \mapsto x^2$.

To calculate \mathbf{d} , at first the distance d^L has to be chosen. In our case, the distance d^L induced by the Euclidean inner product $\langle a, b \rangle^L := \sum_{i,m,n} a(i, m, n)b(i, m, n)$ for $a, b \in \mathbb{R}^{|I| \times (M_\zeta^L + 1) \times (N_\tau^L + 1)}$ was taken. Then the values of \mathbf{d} are derived iteratively: as an initial value, the entries of \mathbf{d} are set to

$$\mathbf{d}[i - 1] = (d^L(\bar{v}^L(y_i), \bar{V}^L(Z_0)))^2 = (d^L(\bar{v}^L(y_i), 0))^2 = (\|\hat{\mathbf{v}}(y_i)\|^L)^2, \quad i = 1, \dots, B,$$

before the first step $k = 1$ for the norm $\|\cdot\|^L$ induced by the inner product $\langle \cdot, \cdot \rangle^L$. Then, in a loop over $k = 1, \dots, A$ the z_k are detected as the y_i with maximum corresponding $\mathbf{d}[i - 1]$, $i \geq k$, and their position is stored. In the tensors $\hat{\Sigma}$ and $\hat{\mathbf{v}}^L$ and the vector \mathbf{d} , the k th position w.r.t. the sample points and the position of the maximum are exchanged such that the z_i are in the first positions of $\hat{\Sigma}$. Finally, $\mathbf{d}[i - 1]$ are updated to the distances $(d^L(\bar{v}^L(\tilde{y}_i), \bar{V}^L(Z_k)))^2$, $i > k$. Here the \tilde{y}_i are the new sorted sample points after exchanging z_k with y_k .

At the end of the k th step, this leads to vectors and tensors of the form

$$\begin{aligned}
\hat{\Sigma}[:, :, j-1] &= z_j, & j &= 1, \dots, k, \\
\hat{\Sigma}[:, :, i-1] &= \tilde{y}_i, & i &= k+1, \dots, B, \\
\hat{\mathbf{v}}^L[:, :, :, j-1] &= \bar{\mathbf{v}}^L(z_j), & j &= 1, \dots, k, \\
\hat{\mathbf{v}}^L[:, :, :, i-1] &= \bar{\mathbf{v}}^L(\tilde{y}_i), & i &= k+1, \dots, B, \\
\mathbf{d}[j-1] &= (d^L(\bar{\mathbf{v}}^L(z_j), \bar{\mathbf{V}}^L(Z_{j-1})))^2, & j &= 1, \dots, k, \quad \text{and} \\
\mathbf{d}[i-1] &= (d^L(\bar{\mathbf{v}}^L(\tilde{y}_i), \bar{\mathbf{V}}^L(Z_k)))^2, & i &= k+1, \dots, B.
\end{aligned}$$

Algorithm 12 Bi-Fidelity Offline Step 2: Identify the most important Ξ values $z_1, \dots, z_A \in Y$

Input: Tensor $\hat{\Sigma} \in \mathbb{R}^{2 \times 2 \times B}$ of low fidelity sample points (set Y_i) and tensor of corresponding low fidelity solutions $\hat{\mathbf{v}}^L \in \mathbb{R}^{|I| \times (M_\zeta^L + 1) \times (N_\tau^L + 1) \times B}$ from Offline Step 1, number A of affordable high fidelity simulations.

Output: Tensor $\hat{\Sigma}_A \in \mathbb{R}^{2 \times 2 \times A}$ of important sample points with $\hat{\Sigma}_A[:, :, k-1] = z_k$ is the k th important Σ models, tensor of corresponding low fidelity solutions $\hat{\mathbf{v}}_A^L \in \mathbb{R}^{|I| \times (M_\zeta^L + 1) \times (N_\tau^L + 1) \times A}$ and Cholesky decomposition factor \mathbf{L} with $\mathbf{G}^L = \mathbf{L}\mathbf{L}^T$ for the Gramian matrix \mathbf{G}^L from last section.

Initialization.

```

[~, ~, B] = size( $\hat{\Sigma}$ );
 $\mathbf{L}$  = zeros(B, A);
 $\mathbf{d}$  = zeros(B, 1);
for  $k = 1 : B$  do
     $\mathbf{d}(k-1) = \sum_{i,m,n} (\hat{\mathbf{v}}^L(i, m, n, k))^2$ ;
end for

```

Code continues on the next page.

Selection procedure: m denotes the maximum $d^L(u^L(y_i), U^L(Z_k))$ value and p its position in \mathbf{d} .

for $k=1:A$ **do**

$[m \ p] = \max(\mathbf{d}(k-1 : B-1));$

$p = p + k - 1;$

 To avoid ill conditioning, check if $m \neq 0$, i.e. if m is bigger than machine precision eps .

if $m < eps$ **then**

$k = k - 1;$

break;

end if

Exchange the entries of \mathbf{d} , \mathbf{L} , $\hat{\Sigma}$ and $\hat{\mathbf{v}}^L$ corresponding to the k th and p th sample point by each other.

$\mathbf{d}([k-1 \ p]) = \mathbf{d}([p \ k-1]);$

$\mathbf{L}([k-1 \ p], :) = \mathbf{L}([p \ k-1], :);$

$\hat{\mathbf{v}}^L(:, :, :, [k-1 \ p]) = \hat{\mathbf{v}}^L(:, :, :, [p \ k-1]);$

$\hat{\Sigma}(:, :, [k-1 \ p]) = \hat{\Sigma}(:, :, [p \ k-1]);$

Calculate the new distances $\mathbf{d}(i-1) = (d^L(\bar{v}^L(\tilde{y}_i), \bar{V}^L(Z_k)))^2$ for $i > k$. It will be verified after this program that the following calculations lead to the right value.

$\mathbf{LLT} = \mathbf{L}(:, 0 : (A-2)) * \mathbf{L}(:, 0 : (A-2))';$

for $t = (k+1) : B$ **do**

$rt = \sum_{i,m,n} \hat{\mathbf{v}}^L(i, m, n, k) \hat{\mathbf{v}}^L(i, m, n, t) - \mathbf{LLT}(t, k); = \langle \hat{\mathbf{v}}^L(:, :, :, k), \hat{\mathbf{v}}^L(:, :, :, t) \rangle^L - \mathbf{LLT}(t, k).$

$\mathbf{L}(k-1, k-1) = \sqrt{\mathbf{d}(k-1)};$

$\mathbf{L}(t-1, k-1) = rt / \mathbf{L}(k-1, k-1);$

$\mathbf{d}(t-1) = \mathbf{d}(t-1) - \mathbf{L}(t-1, k-1)^2;$

end for

end for

To avoid ill conditioning, if one m was smaller than eps .

$A = k;$

Truncate $\hat{\Sigma}$, $\hat{\mathbf{v}}^L$ and \mathbf{L} such that they contain only the A important points z_k , corresponding low fidelity solutions or the corresponding rows of \mathbf{L} .

$\hat{\mathbf{v}}_A^L = \hat{\mathbf{v}}^L(:, :, :, 0 : A-1);$

$\hat{\Sigma}_A = \hat{\Sigma}(:, :, 0 : A-1);$

$\mathbf{L} = \mathbf{L}(0 : A-1, :);$

In the following, the update step of \mathbf{d} will be explained in order to show that \mathbf{d} actually

saves the right values at the end of the k th step

$$\mathbf{d}_{new} = [(d^L(\bar{v}^L(z_1), \bar{V}^L(Z_0)))^2, \dots, (d^L(\bar{v}^L(z_k), \bar{V}^L(Z_{k-1})))^2, \quad (8.12)$$

$$(d^L(\bar{v}^L(\tilde{y}_{k+1}), \bar{V}^L(Z_k)))^2, \dots, (d^L(\bar{v}^L(\tilde{y}_B), \bar{V}^L(Z_k)))^2].$$

Denote a matrix or a vector \mathbf{X} at the beginning of step k by \mathbf{X}_{old} and the same matrix or vector after the k th step by \mathbf{X}_{new} . Then, \mathbf{X}_{new} from step k will become \mathbf{X}_{new} at the beginning of the $(k+1)$ st step.

The following induction shows that the update step from program 12 produces the right \mathbf{d}_{new} .

For step $k=1$, compute at first the updated matrix \mathbf{L}_{new} . Recall that it was initialized by $\mathbf{L}_{old} = \mathbf{0}_{B \times A}$ and therefore $\mathbf{LLT} = \mathbf{0}_B$. Denote by \tilde{u}_k^L the Gram Schmidt orthonormalized solution $\bar{v}^L(z_k)$, $k=1, \dots, A$. Then one has

$$\begin{aligned} \mathbf{L}_{new}(k-1, k-1) &= \sqrt{\mathbf{d}_{old}[0, 0]} = d^L(\bar{v}^L(z_1), \bar{V}^L(Z_0)) = d^L(\bar{v}^L(z_k), \bar{V}^L(Z_{k-1})) \\ \mathbf{L}_{new}(t-1, k-1) &= \frac{rt}{\mathbf{L}(k-1, k-1)} = \frac{\langle \bar{v}^L(z_k), \bar{v}^L(\tilde{y}_t) \rangle^L - \mathbf{LLT}[t-1, k-1]}{d^L(\bar{v}^L(z_k), \bar{V}^L(Z_{k-1}))} \\ &= \frac{\langle \bar{v}^L(z_1), \bar{v}^L(\tilde{y}_t) \rangle^L}{d^L(\bar{v}^L(z_1), \bar{V}^L(Z_0))} = \left\langle \frac{\bar{v}^L(z_1)}{\|\bar{v}^L(z_1)\|^L}, \bar{v}^L(\tilde{y}_t) \right\rangle^L = \langle \tilde{u}_1^L, \bar{v}^L(\tilde{y}_t) \rangle^L \\ &= \langle \tilde{u}_k^L, \bar{v}^L(\tilde{y}_t) \rangle^L = \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_k^L \rangle^L \end{aligned}$$

for $t > k=1$. With this, the updated \mathbf{d}_{new} is given by

$$\begin{aligned} \mathbf{d}_{new}[t-1] &= \mathbf{d}_{old}[t-1] - (\mathbf{L}[t-1, k-1])^2 = (\|\bar{v}^L(\tilde{y}_t)\|^L)^2 - (\langle \tilde{u}_1^L, \bar{v}^L(\tilde{y}_t) \rangle^L)^2 \\ &= \langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(\tilde{y}_t) \rangle^L - \langle \langle \tilde{u}_1^L, \bar{v}^L(\tilde{y}_t) \rangle^L \tilde{u}_1^L, \bar{v}^L(\tilde{y}_t) \rangle^L \\ &= \langle \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_1)}(\bar{v}^L(\tilde{y}_t)), \bar{v}^L(\tilde{y}_t) \rangle^L \\ &= \langle \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_1)}(\bar{v}^L(\tilde{y}_t)), \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_1)}(\bar{v}^L(\tilde{y}_t)) \rangle^L \\ &= (\|\bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_1)}(\bar{v}^L(\tilde{y}_t))\|^L)^2 = (d^L(\bar{v}^L(\tilde{y}_t), \bar{V}^L(Z_1)))^2, \end{aligned}$$

where the equality in the second to last line follows from the fact that $P_{\bar{V}^L(Z_1)}$ is an orthogonal projection and therefore $\langle P_{\bar{V}^L(Z_1)}(x), x - P_{\bar{V}^L(Z_1)}(x) \rangle^L = 0$ for all $x \in \mathbb{R}^{|I| \times (M_\xi^L + 1) \times (N_T^L + 1)}$, since $Id - P_{\bar{V}^L(Z_1)}$ is the orthogonal projection on the orthogonal complement of $\bar{V}^L(Z_1)$, if Id denotes the identity.

Assume now, that for a fixed but arbitrary $k \in 1, \dots, A-1$, at the end of the k th it-

eration the matrix \mathbf{L} and the vector \mathbf{d} are given by

$$\begin{aligned}
\mathbf{L}_{new}[j-1, j-1] &= d^L(\bar{v}^L(z_j), \bar{V}^L(Z_{j-1})) && \text{for } j = 1, \dots, k, \\
\mathbf{L}_{new}[i-1, j-1] &= \langle \bar{v}^L(\tilde{y}_i), \tilde{u}_j^L \rangle^L && \text{for } j = 1, \dots, k, i > j, \\
\mathbf{L}_{new}[i-1, j-1] &= 0, && \text{for } j = 1, \dots, k, i < j, \\
\mathbf{L}_{new}[i-1, j-1] &= 0, && \text{for } j = k+1, \dots, A, i = 1, \dots, B, \quad \text{and} \\
\mathbf{d}_{new}[j-1] &= (d^L(\bar{v}^L(z_j), \bar{V}^L(Z_{j-1})))^2 && \text{for } j = 1, \dots, k, \\
\mathbf{d}_{new}[i-1] &= (d^L(\bar{v}^L(\tilde{y}_i), \bar{V}^L(Z_k)))^2 && \text{for } i = k+1, \dots, B.
\end{aligned}$$

Consider the $(k+1)$ st iteration. Here, \mathbf{L}_{new} and \mathbf{d}_{new} from the k th step as given above become \mathbf{L}_{old} and \mathbf{d}_{old} of the $(k+1)$ st step. At first, compute rt for $t > k+1$:

$$\begin{aligned}
rt &= \langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) \rangle^L - \mathbf{L}\mathbf{L}\mathbf{T}[t-1, k+1-1] \\
&= \langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) \rangle^L - \sum_{l=1}^{A-1} \mathbf{L}_{old}[t-1, l-1] \mathbf{L}_{old}^T[l-1, k+1-1] \\
&\stackrel{(1)}{=} \langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) \rangle^L - \sum_{l=1}^k \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_l^L \rangle^L \langle \bar{v}^L(z_{k+1}), \tilde{u}_l^L \rangle^L \\
&= \left\langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) - \sum_{l=1}^k \langle \bar{v}^L(z_{k+1}), \tilde{u}_l^L \rangle^L \tilde{u}_l^L \right\rangle^L \\
&= \langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) - P_{\bar{V}^L(Z_k)} \bar{v}^L(z_{k+1}) \rangle^L,
\end{aligned}$$

where equality (1) follows since $\mathbf{L}_{old}^T[l-1, k+1-1] = 0$ for $l > k$. With

$$\begin{aligned}
\mathbf{L}_{new}[k+1-1, k+1-1] &= \sqrt{\mathbf{d}_{old}[k+1-1]} = d^L(\bar{v}^L(z_{k+1}), \bar{V}^L(Z_k)) \\
&= \|\bar{v}^L(z_{k+1}) - P_{\bar{V}^L(Z_k)} \bar{v}^L(z_{k+1})\|^L,
\end{aligned}$$

the matrix \mathbf{L}_{new} is of the same form as \mathbf{L}_{old} for $k+1$ instead of k , since for $t > k+1$ one has

$$\begin{aligned}
\mathbf{L}_{new}[t-1, k+1-1] &= \frac{rt}{\mathbf{L}_{new}[k+1-1, k+1-1]} \\
&= \frac{\langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(z_{k+1}) - P_{\bar{V}^L(Z_k)} \bar{v}^L(z_{k+1}) \rangle^L}{\|\bar{v}^L(z_{k+1}) - P_{\bar{V}^L(Z_k)} \bar{v}^L(z_{k+1})\|^L} = \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_{k+1}^L \rangle^L.
\end{aligned}$$

Using this, the update of vector \mathbf{d} is given by

$$\begin{aligned}
\mathbf{d}_{new}[t-1] &= \mathbf{d}_{old}[t-1] - \mathbf{L}[t-1, k+1-1]^2 \\
&= \left(d^L(\bar{v}^L(\tilde{y}_t), \bar{V}^L(Z_k)) \right)^2 - \left(\langle \bar{v}^L(\tilde{y}_t), \tilde{u}_{k+1}^L \rangle^L \right)^2 \\
&=^{(1)} \left\langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_k)} \bar{v}^L(\tilde{y}_t) \right\rangle^L - \left\langle \bar{v}^L(\tilde{y}_t), \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_{k+1}^L \rangle^L \tilde{u}_{k+1}^L \right\rangle^L \\
&= \left\langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(\tilde{y}_t) - \sum_{j=1}^k \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_j^L \rangle^L \tilde{u}_j^L - \langle \bar{v}^L(\tilde{y}_t), \tilde{u}_{k+1}^L \rangle^L \tilde{u}_{k+1}^L \right\rangle^L \\
&= \left\langle \bar{v}^L(\tilde{y}_t), \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_{k+1})} \bar{v}^L(\tilde{y}_t) \right\rangle^L \\
&=^{(2)} \left\langle \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_{k+1})} \bar{v}^L(\tilde{y}_t), \bar{v}^L(\tilde{y}_t) - P_{\bar{V}^L(Z_{k+1})} \bar{v}^L(\tilde{y}_t) \right\rangle^L \\
&= \left(d^L(\bar{v}^L(\tilde{y}_t), \bar{V}^L(Z_{k+1})) \right)^2,
\end{aligned}$$

where the equalities (1) and (2) follow by the same argument as in the calculation of \mathbf{d}_{new} in step $k = 1$.

In summary, it was shown by induction that \mathbf{d}_{new} is of the desired form 8.12 and therefore, taking z_{k+1} as the \tilde{y}_t , $t > k$, with maximum value corresponding to it in \mathbf{d}_{old} implements the point selection rule explained in equation 8.1.

Step 3: Compute the high fidelity solution in the important points

As a next step, the high fidelity solution, i.e. the high resolution numerical solution with $M_\zeta^H + 1$ grid points in ζ , is computed for every important volatility model $z_k = \hat{\Sigma}_A(:, :, k-1)$, $k = 1, \dots, A$, by scheme 5.6. At first, the number of grid points $N_\tau^H + 1$ in τ is chosen such that the numerical scheme is stable for all volatility models z_k . Then the numerical solution $\bar{\mathbf{v}}^{H,k}$ with $\bar{\mathbf{v}}^{H,k}[i, m, n] = \bar{v}_{\phi(i)}^H(\zeta_m^H, \tau^{H,n})$ is calculated for the volatility model z_k and stored in the tensor $\hat{\mathbf{v}}^H \in \mathbb{R}^{|I| \times (M_\zeta^H + 1) \times (N_\tau^H + 1) \times A}$ with $\hat{\mathbf{v}}^H[:, :, :, k-1] = \bar{\mathbf{v}}^{H,k}$ for $k = 1, \dots, A$.

Algorithm 13 Bi-Fidelity Offline step 3: Compute the high fidelity solution in the important points

Input: Truncation number N of the truncated gPC expansions of V , interest rate r , maturity of the option T , number of discretization points in the high fidelity model M_ζ^H for ζ , Galerkin multiplication tensors $\mathbf{M}^{(1)}$ and $\mathbf{M}^{(2)}$ as well as the important volatility models collected in $\hat{\Sigma}_A$.

Output: Tensor of high fidelity solutions $\hat{\mathbf{v}}^H \in \mathbb{R}^{|I| \times (M_\zeta^H + 1) \times (N_\tau^H + 1) \times A}$ corresponding to the important volatility models in $\hat{\Sigma}_A$.

[Code continues on the next page.](#)

Compute the minimum N_τ^L such that the finite difference scheme 5.6 is stable for all important volatility models z_k . The variable *stable* saves whether the scheme can be stabilized by choosing a high N_τ^H , the variable *dtaumax* saves the maximum attainable value of $\Delta\tau^H = T/N_\tau^H$ when stability holds for all z_k .

dtaumax = *Inf*;

for $k = 1 : A$ **do**

 Compute the coupling matrix \mathbf{A} for volatility model $\mathbf{sigma} = \hat{\Sigma}(:, :, k-1)$ by modified program 9 from Offline step 1.

 Compute *stable* and *dt* by the stability determining program 4. Then *dt* denotes the maximum $\Delta\tau^H$ value allowed for stability when considering the equation with volatility model z_k .

if *stable* == 0 **then**

 error('Unstable computation of high fidelity solutions.');

end if

if *dt* < *dtaumax* **then**

dtaumax = *dt*;

end if

end for

$N_\tau^H = \text{ceil}(T/dtaumax)$;

Generate the high fidelity solutions.

initcondhigh = zeros(($N + 1$) * ($N + 2$)/2, $M_\zeta^H + 1$);

initcondhigh(1, :) = max(2/ M_ζ^H * (0 : M_ζ^H) - 1, 0);

$\hat{\mathbf{v}}^H = \text{zeros}((N + 1) * (N + 2)/2, M_\zeta^H + 1, N_\tau^H + 1, A)$;

for $k = 1 : A$ **do**

 Calculate $\bar{\mathbf{v}}^H$ by modified program 10 from Offline step 1 for a volatility model explained by $\hat{\Sigma}(:, :, k-1)$ with $M_\zeta^H + 1$ grid points in ζ and $N_\tau^H + 1$ in τ .

 By choice of N_τ^H , the scheme is stable. The parabolicity of the equation follows from the parabolicity check when calculating the low fidelity solutions. Therefore no further check of these conditions is necessary.

$\hat{\mathbf{v}}^H = \text{cat}(4, \hat{\mathbf{v}}^H, \bar{\mathbf{v}}^H)$;

end for

Store the data, i.e. save $r, T, N, M_\zeta^L, N_\tau^L, M_\zeta^H, N_\tau^H, A, \hat{\Sigma}_A, \hat{\mathbf{v}}^H, \hat{\mathbf{v}}_A^L$ and \mathbf{L} for later usage.

For the online steps, assume a volatility model $z \in \mathbb{D}$ is given in the form 8.10 and one wants to compute the Bi-Fidelity solution of the Black Scholes equation with uncertain volatility of the form $\Sigma(\Theta_1, \Theta_2) = \sigma_{00} + \sigma_{10}\Theta_1 + \sigma_{01}\Theta_2$. This is done as follows:

Step 1: Compute the low fidelity solution in the point z

The low resolution numerical solution $\bar{\mathbf{v}}^L(z)$ is calculated by scheme 5.6. It is important that the same number of grid points $M_\zeta^L + 1$ and $N_\tau^L + 1$ in ζ and τ direction are used as in the computation of the low fidelity solutions in Offline step 1. The calculation is done via

modified program 10 as explained in Offline step 1 and stability is checked by the output variable *stable*. An error is displayed, if the scheme is unstable and the calculation is aborted, as it is done in program 13.

Step 2: Project the low fidelity solution onto $\bar{V}^L(Z_A)$

In this step, the projection coefficients c_k are derived with $P_{\bar{V}^L(Z_A)}\bar{v}^L(z) = \sum_{k=1}^A c_k \bar{v}^L(z_k)$. To do this, recall the relation

$$\mathbf{G}^L \mathbf{c} = \mathbf{f}^L \quad (8.8)$$

for $\mathbf{c} = (c_1, \dots, c_A)^T$ and \mathbf{G}^L and \mathbf{f}^L defined as in the last section. Moreover, the matrix \mathbf{L} from Offline step 2 was computed such that $\mathbf{G}^L = \mathbf{L}\mathbf{L}^T$. Since it was stored, one can use it to easily compute the projection coefficients by

$$\mathbf{c} = (\mathbf{L}^T)^{-1} \mathbf{L}^{-1} \mathbf{f}^L.$$

Note that \mathbf{L} is invertible because \mathbf{G}^L is by choice of the z_k . This is the theoretical background of the following program that resembles an implementation of algorithm 2 from [Nar14] in Matlab notation.

Algorithm 14 Bi-Fidelity Online step 2: Calculate of the projection coefficients c_k

Input: Calculated low fidelity solution $\bar{\mathbf{v}}^L(z)$ from Online Step 1, number A of projection coefficients, low fidelity solutions $\hat{\mathbf{v}}_A^L$ for the important volatility models z_k as collected in $\hat{\Sigma}_A$, matrix \mathbf{L} from Offline step 2.

Output: Vector $\mathbf{c} = (c_1, \dots, c_A)^T$ of projection coefficients of $P_{\bar{V}^L(Z_A)}\bar{v}^L(z) = \sum_{k=1}^A c_k \bar{v}^L(z_k)$.

Compute \mathbf{f}^L

$\mathbf{f}^L = \text{zeros}(A, 1)$;

for $k = 1 : A$ **do**

$\mathbf{f}^L(k) = \sum_{i,m,n} \bar{\mathbf{v}}^L(i, m, n) * \hat{\mathbf{v}}^L(i, m, n, k); = \langle \bar{\mathbf{v}}^L(z), \hat{\mathbf{v}}^L[:, :, :, k] \rangle^L$

end for

$\mathbf{c} = \text{inv}(\mathbf{L})' * \text{inv}(\mathbf{L}) * \mathbf{f}^L$;

Step 3: Calculate the Bi-Fidelity solution

Finally, the Bi-Fidelity solution is constructed by formula

$$\bar{v}^{BF}(z) := \sum_{k=1}^A c_k \bar{v}^H(z_k). \quad (8.2)$$

Algorithm 15 Bi-Fidelity Online step 3: Construct the Bi-Fidelity solution

Input: Truncation number N , computed vector $\mathbf{c} = (c_1, \dots, c_A)^T$ of projection coefficients and high fidelity solutions $\hat{\mathbf{v}}^H$ in important volatility models.

Output: Bi-Fidelity solution $\bar{\mathbf{v}}^{BF}$ for the volatility model declared by z with $\bar{\mathbf{v}}^{BF}[i, m, n] = \bar{v}_{\phi(i)}^{BF}(\zeta_m^H, \tau^{H,n})$.

$\bar{\mathbf{v}}^{BF} = \text{zeros}((N + 1 * (N + 2)/2), M_{\zeta}^H + 1, N_{\tau}^H + 1);$

for $k = 1 : A$ **do**

$\bar{\mathbf{v}}^{BF} = \bar{\mathbf{v}}^{BF} + \mathbf{c}[k - 1] * \hat{\mathbf{v}}^H[:, :, :, k - 1];$

end for

After deriving $\bar{\mathbf{v}}^{BF}$, it has to be transformed back to the original variables \mathbf{v} , S and t by equation 5.2. This can be done by the same code used for the transformation back at the end of program 10.

8.3 Comparing Bi-Fidelity solution and high fidelity solution

This section will present Bi-Fidelity solutions of the Black Scholes equation with uncertain volatility 7.1 for a European Call option in two volatility models. They will be compared to the corresponding high fidelity solutions derived by direct application of program 10. After that, a simulation will be done to find the mean size and shape of the error in expected value and in variance between the Bi-Fidelity solution and the high fidelity solution.

The volatility was assumed to depend on two random variables Θ and Δ with standard normal distribution and uniform distribution on $[-0.5, 0.5]$ respectively. Its gPC expansion was truncated after total polynomial degree $K = 1$.

The interest rate in the market was supposed to be $r = 0$ and a maturity of $T = 23$ days was chosen. This describes a period of slightly more than a month, since stock markets are usually closed on weekends and holidays. Again, for computations the maturity and times t given in days were multiplied with $1/251$ to obtain yearly values. 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 it was done in previous chapters on numerical results.

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 computations of low fidelity solutions performed in the examples explained below was stable. In 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 grid with $M_\zeta^H + 1 = 200 + 1$ grid points in ζ direction. The number of grid points $N_\tau^H + 1 = 1908 + 1$ in τ direction was derived by program 13 such that all high resolution computations for important volatility models were stable.

The low fidelity sample points were of the form

$$y_i = \begin{pmatrix} \sigma_{00}^{(i)} & \sigma_{01}^{(i)} \\ \sigma_{10}^{(i)} & 0 \end{pmatrix}$$

as in equation 8.10 with

$$\begin{aligned} \sigma_{00}^{(i)} &\in \{0 < 0.05\lambda \leq 0.8 \mid \lambda \in \mathbb{N} \setminus \{0\}\}, \\ \sigma_{10}^{(i)} &\in \left\{0.05\lambda \leq \sqrt{\sigma_{00}/2} \mid \lambda \in \mathbb{N}_0\right\} \quad \text{and} \\ \sigma_{01}^{(i)} &\in \left\{0.05\lambda \leq \sqrt{12(\sigma_{00}/2 - \sigma_{10}^2)} \mid \lambda \in \mathbb{N}_0\right\}. \end{aligned} \tag{8.13}$$

The Bi-Fidelity solutions can then be derived by the numerics explained in the last chapter.

At first, the volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$ known from section 7.5 will be considered. Figures 8.1 and 8.2 show the expected value surfaces of the high fidelity and the Bi-Fidelity solution with contour lines at each quarter of the maximum absolute value and the borders of the smoothing area plotted in red. The expected values seem to approximately coincide.

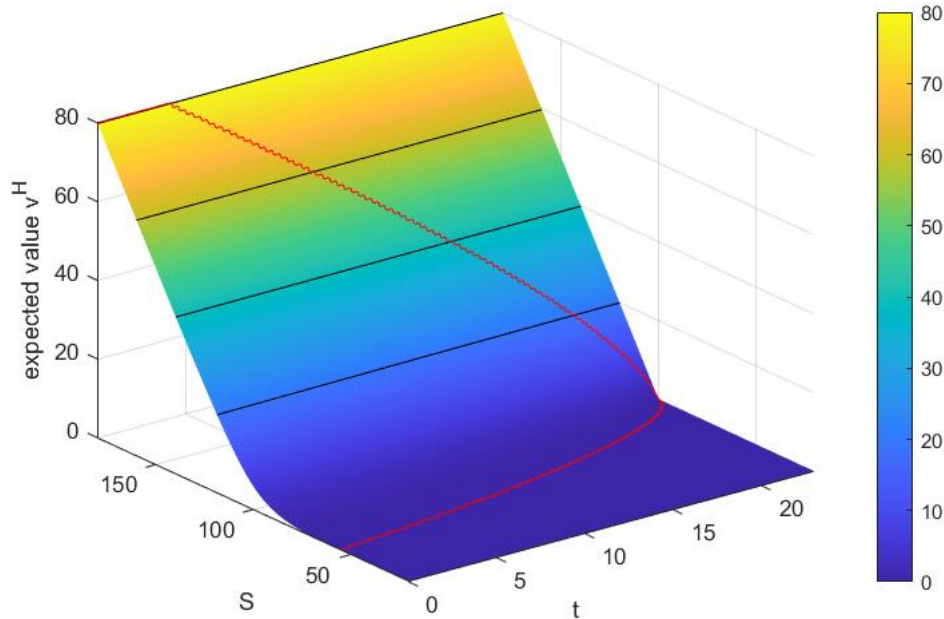


Figure 8.1: Expected value surface of the high fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

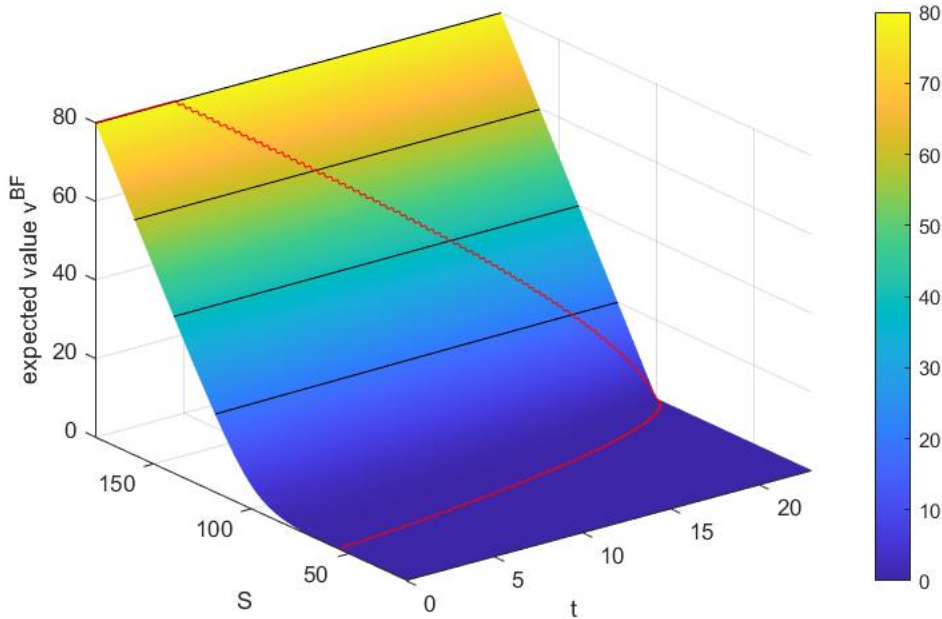


Figure 8.2: Expected value of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

To study their deviations, the absolute difference in expected values is displayed in figures 8.3 close to the strike price and figure 8.4 for a wider range of S values. One can observe that there is some difference of size 10^{-3} within the smoothing area, but for $S \rightarrow \infty$ the difference of the two solutions seems to increase in absolute value. Figure 8.5 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 tends to infinity. Therefore, a difference of 0.3 in these regions means 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 multiplies by $strike$ when transforming back the variables. Hence, an error of size 10^{-3} at strike 100 means an error of size $10^{-5} \cdot strike$.

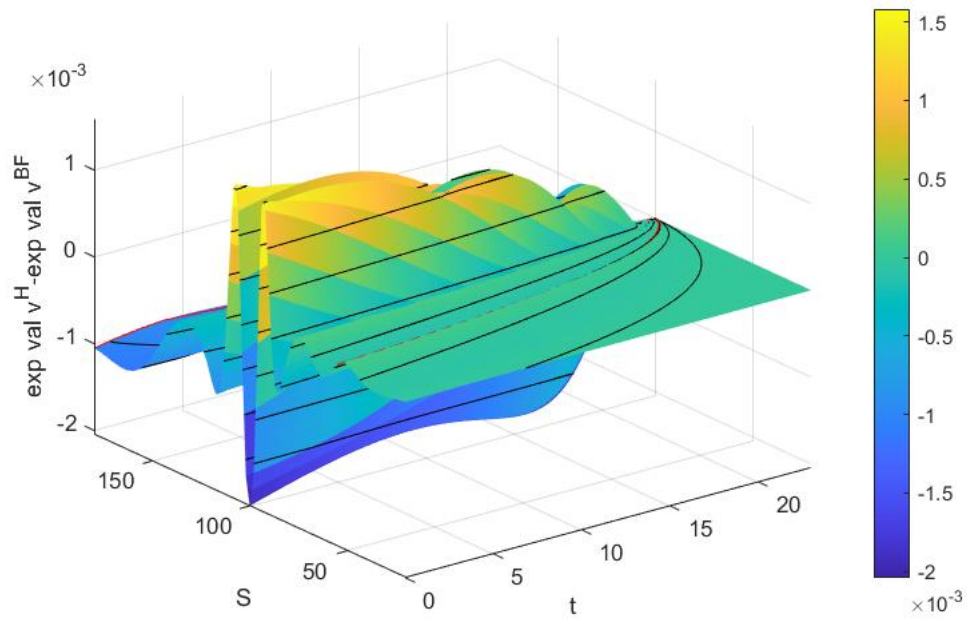


Figure 8.3: Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red.

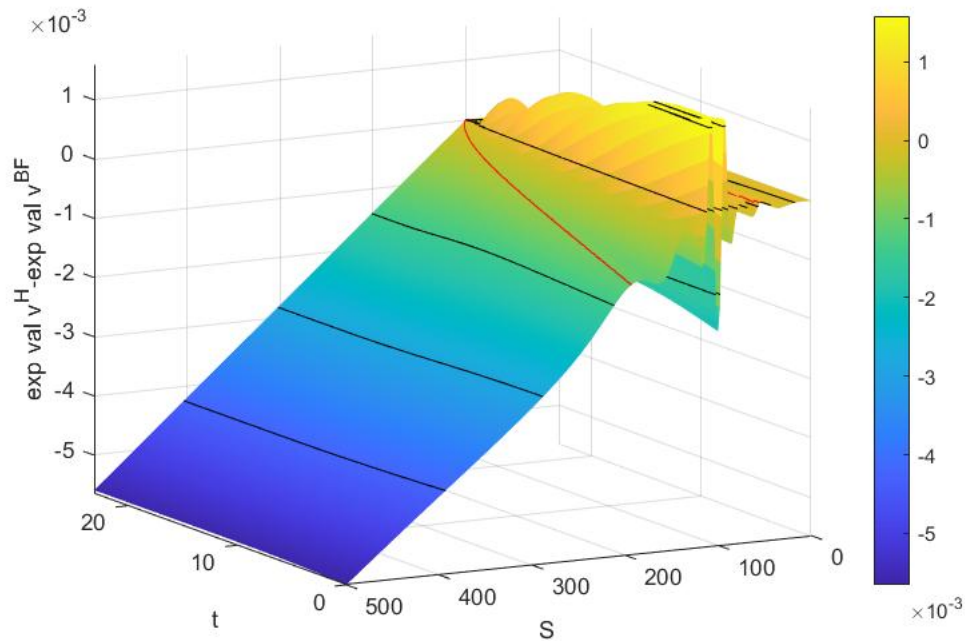


Figure 8.4: Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for a wider range of S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red.

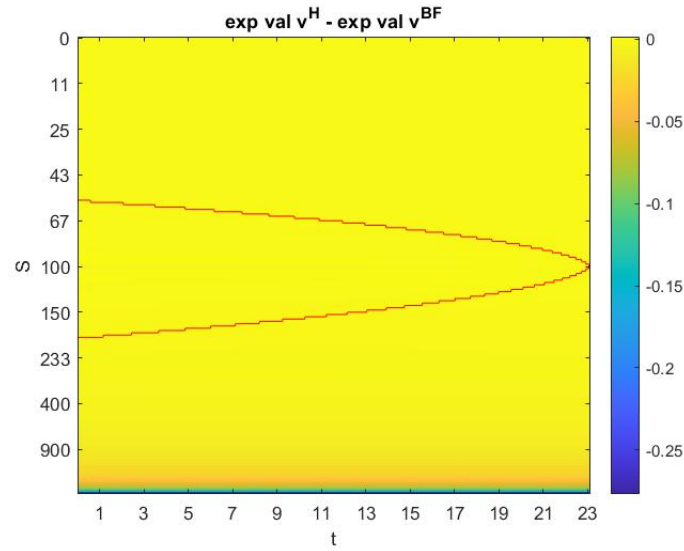


Figure 8.5: Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red.

Now look at the variances. They are shown in figures 8.6 and 8.7. Again they look similar.

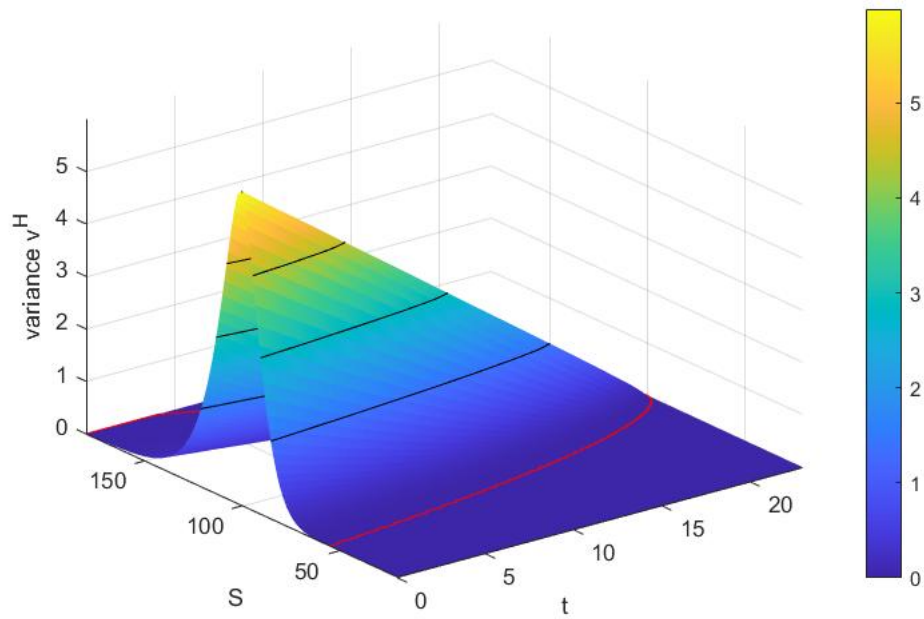


Figure 8.6: Variance of the high fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

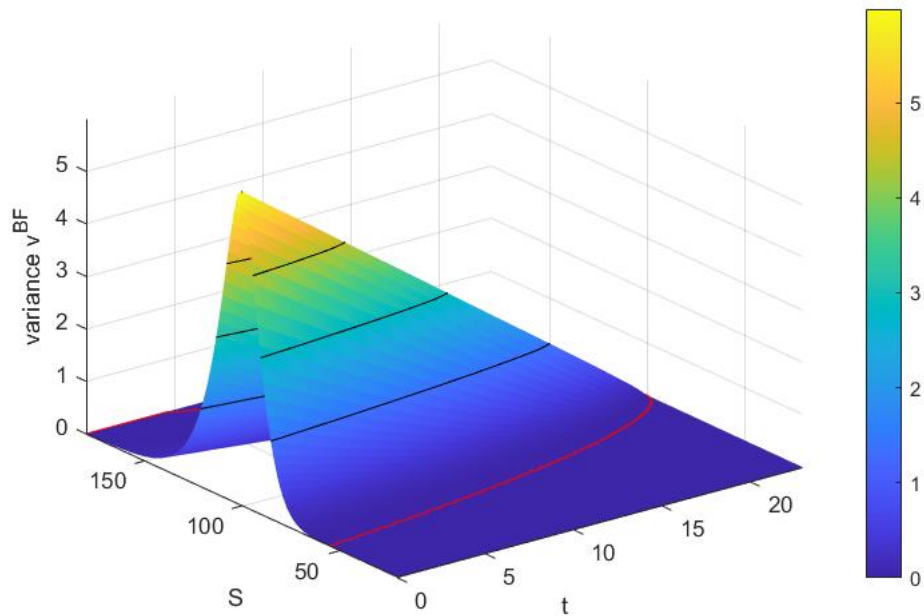


Figure 8.7: Variance of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

We examine the absolute difference in variance as represented in figure 8.8 to lie in the smoothing area. Figure 8.9 showing the difference for all S and t values supports this conclusion. The error is again of size $10^{-3} = 10^{-7}strike^2$.

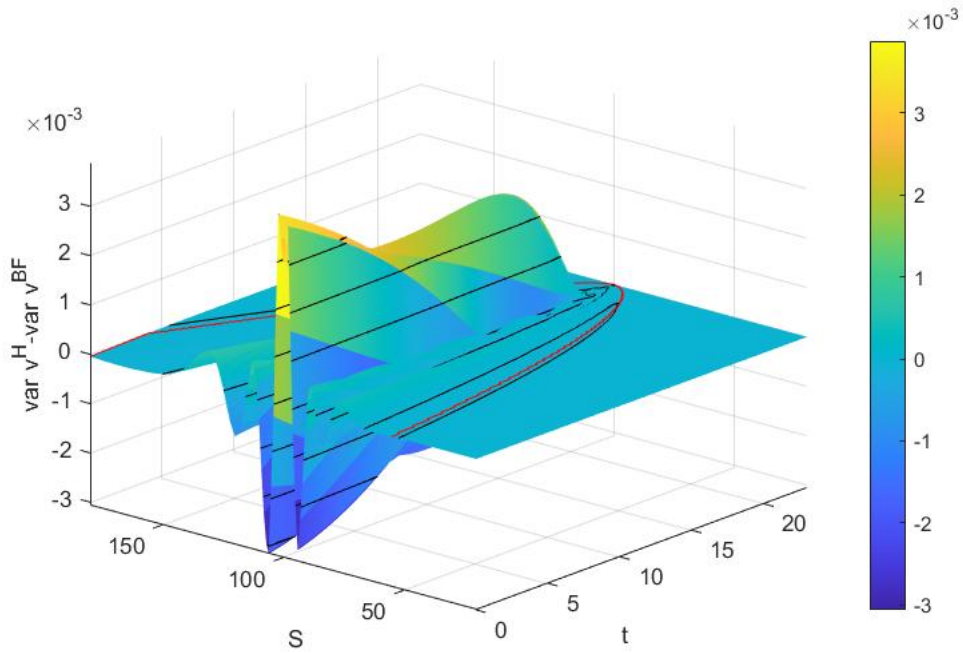


Figure 8.8: Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and the high fidelity smoothing area circled in red.

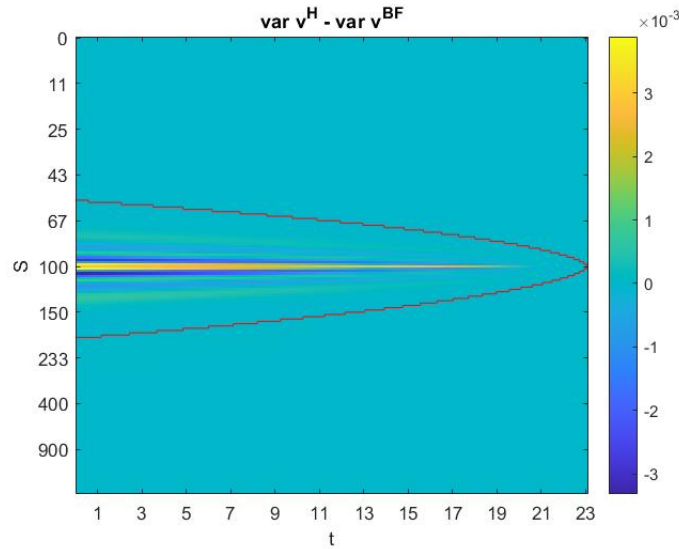


Figure 8.9: Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_1(\Theta, \Delta) = 0.5 + 0.2\Theta + 0.1\sqrt{12}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red.

Now a second volatility model will be considered. The model is given by $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, where $g = (1 + \sqrt{5})/2$ denotes the golden ratio. The expected value of the high fidelity and the Bi-Fidelity solution can be found in figures 8.10 and 8.11 respectively. Again the difference does not seem to be large.

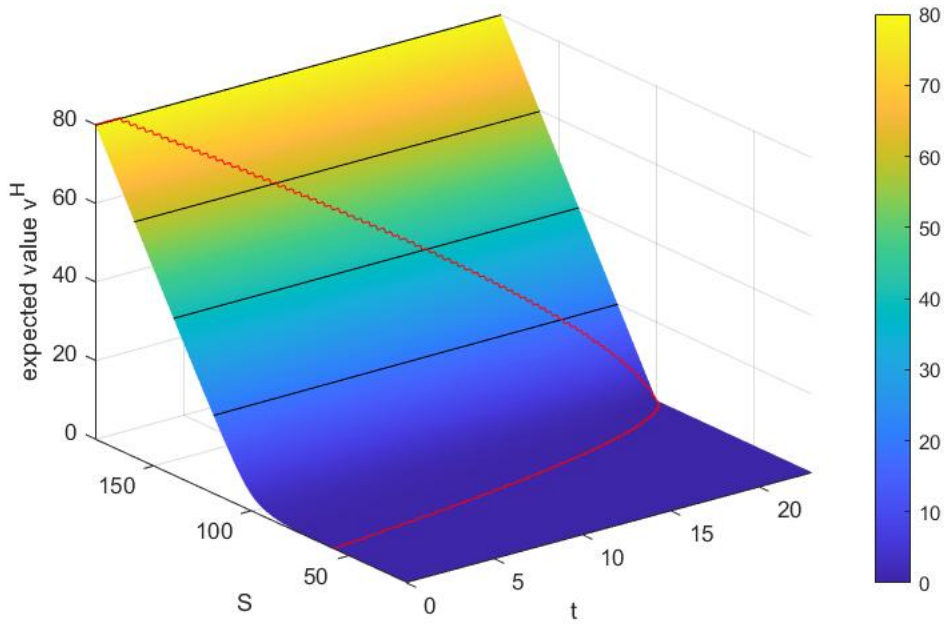


Figure 8.10: Expected value of the high fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

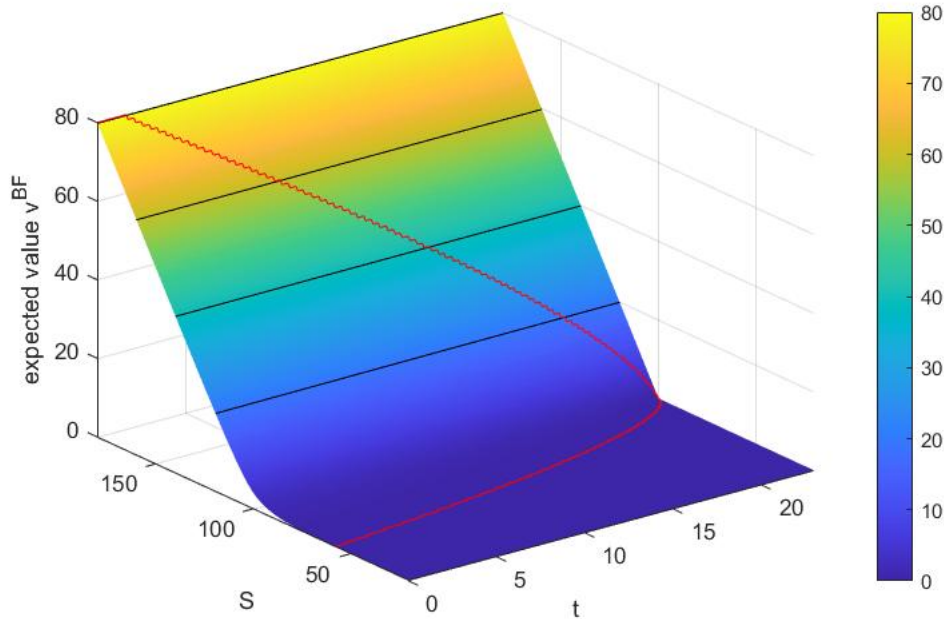


Figure 8.11: Expected value of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

The absolute difference is investigated in figure 8.12 close to the strike price and figure 8.13 for all times and S values. The figures show that the difference is also of the size $10^{-3} = 10^{-5}strike$ and lies in the smoothing area and close to $S = \infty$.

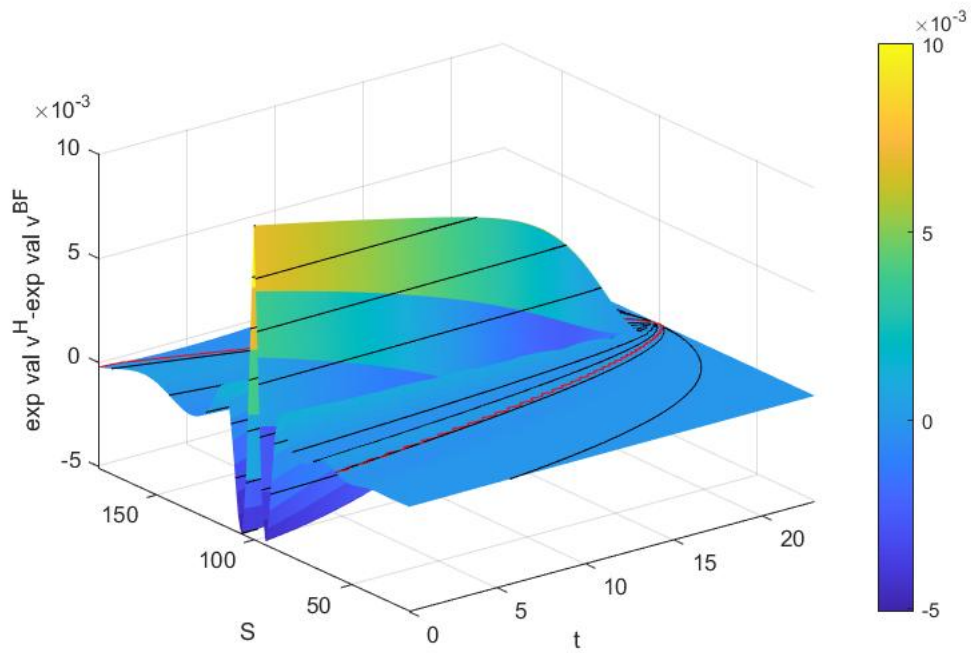


Figure 8.12: Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum absolute value and the high fidelity smoothing area circled in red.

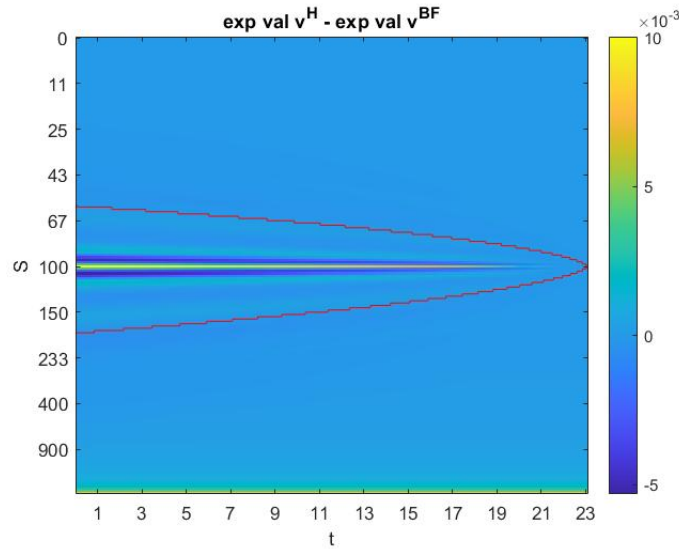


Figure 8.13: Absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red.

The variances are displayed in figures 8.14 and 8.15. Their absolute difference is represented in figure 8.16 close to the strike price and figure 8.17 for all S and t . This time the maximum absolute value of the difference in variance is of size $10^{-2} = 10^{-6}strike^2$. This indicates that the Bi-Fidelity solution for model Σ_2 is not as close to the high fidelity solution as it was for model Σ_1 . This hypothesis is supported by the Bi-Fidelity errors as explained in section 8.1. The error for model Σ_1 attains the value $3.4739 \cdot 10^{-5}$, whereas the error for model Σ_2 equals $6.9868 \cdot 10^{-5}$.

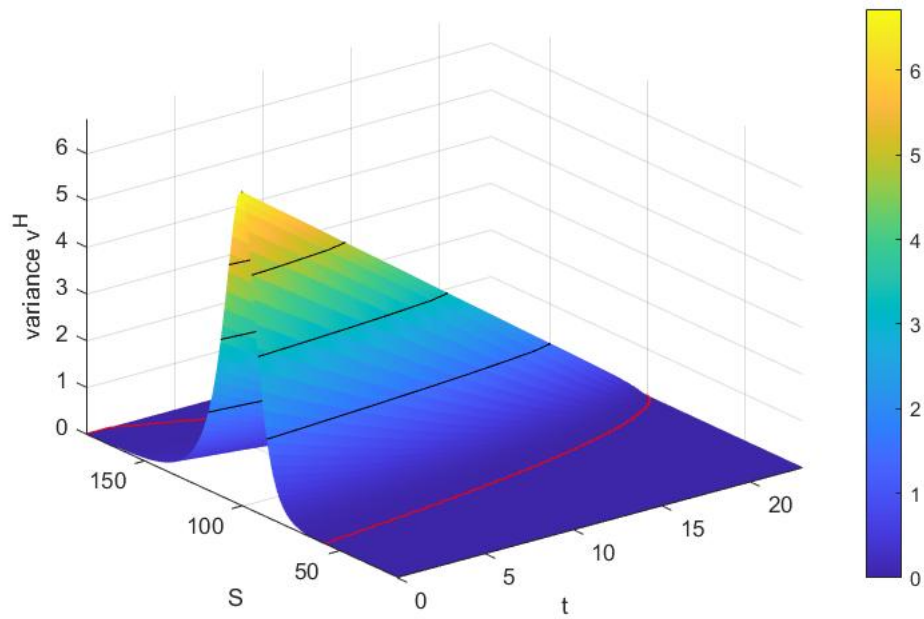


Figure 8.14: Variance of the high fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

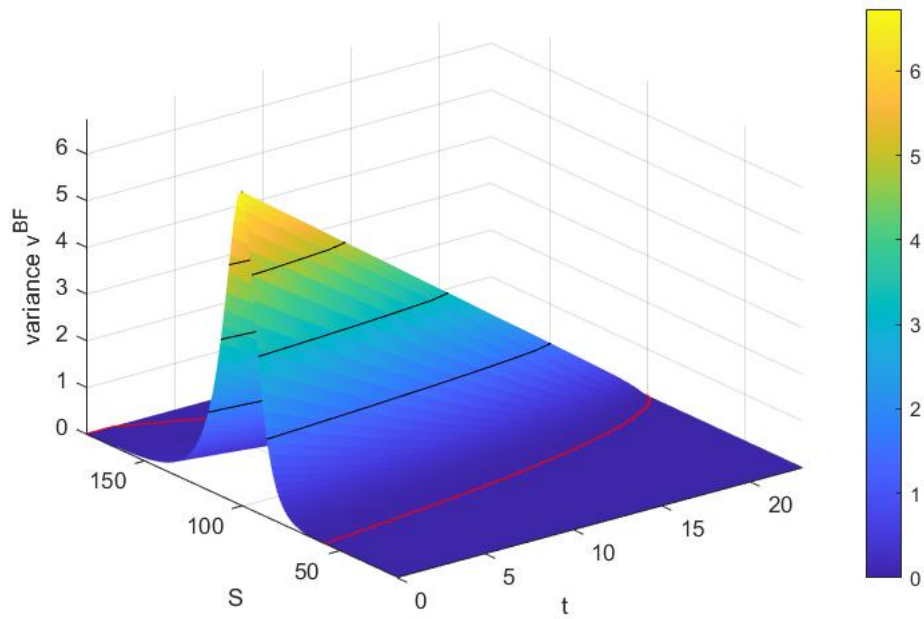


Figure 8.15: Variance of the Bi-Fidelity solution for a European Call option for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and its smoothing area circled in red.

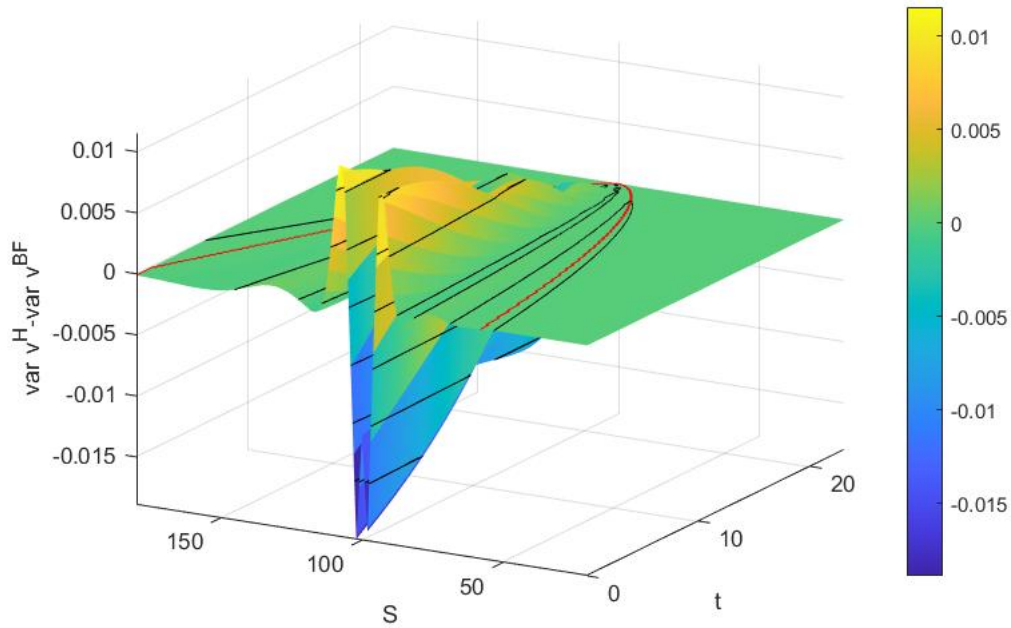


Figure 8.16: Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with contour lines at quarters of its maximum value and the high fidelity smoothing area circled in red.

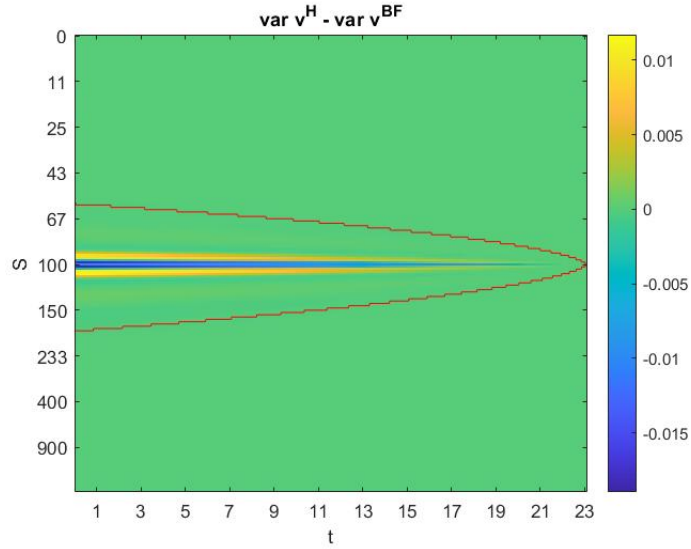


Figure 8.17: Absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for volatility model $\Sigma_2(\Theta, \Delta) = 0.1\pi + 0.1e\Theta + 0.1g\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_\zeta^L = 50$, $N_\tau^L = 150$, low fidelity sample points as in 8.13, $M_\zeta^H = 200$ and $N_\tau^H = 1908$ with the high fidelity smoothing area circled in red.

Finally, a simulation of the errors was done to obtain the mean size and shape of the Bi-Fidelity error. For this purpose, 100 volatility models of the form $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$ were generated randomly by obtaining the coefficients σ_{ij} as realizations of uniform random variables such that $\sigma_{00} \in [0, 0.8]$, $\sigma_{10} \in [0, \sqrt{\sigma_{00}/2}]$, $\sigma_{01} \in [0, \sqrt{12(\sigma_{00}/2 - \sigma_{10}^2)}]$. The mean absolute difference of the expected value of the Bi-Fidelity solution and the expected value of the high fidelity solution is represented in figure 8.18 close to the strike price and figure 8.19 for a larger range of S values. Figure 8.20 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 figures shows a shape of the error in expectation similar to the error shape in the previous volatility models. There is an oscillation of size $10^{-4} = 10^{-6}strike$ close to the strike price and for $S \rightarrow \infty$ the error steadily increases in absolute value. The maximum absolute difference lies close to $S = \infty$ and has a size of $10^{-1} = 10^{-3}strike$, which is small in relative terms. This shape seems to be characteristic for the considered Bi-Fidelity model.

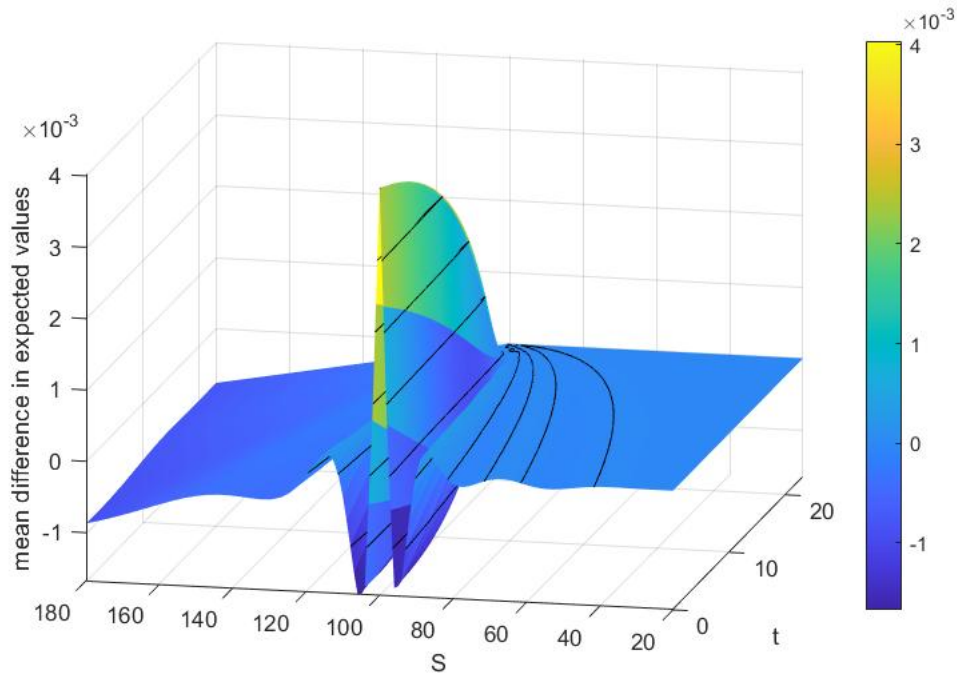


Figure 8.18: Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option close to the strike price for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$ with contour lines at quarters of its maximum absolute value.

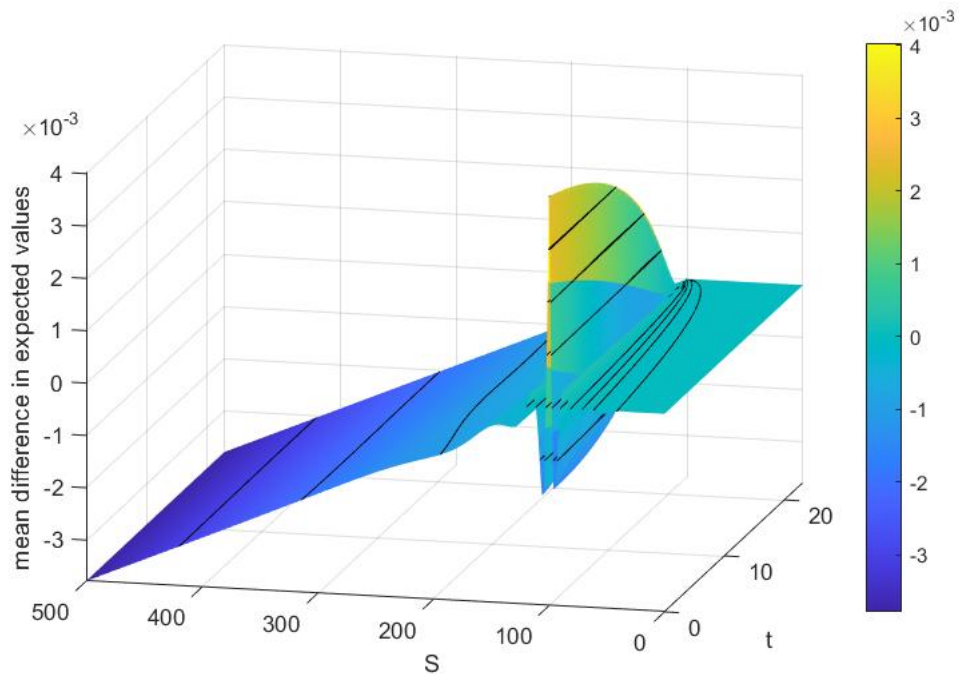


Figure 8.19: Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for a wider range of S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$ with contour lines at quarters of its maximum absolute value.

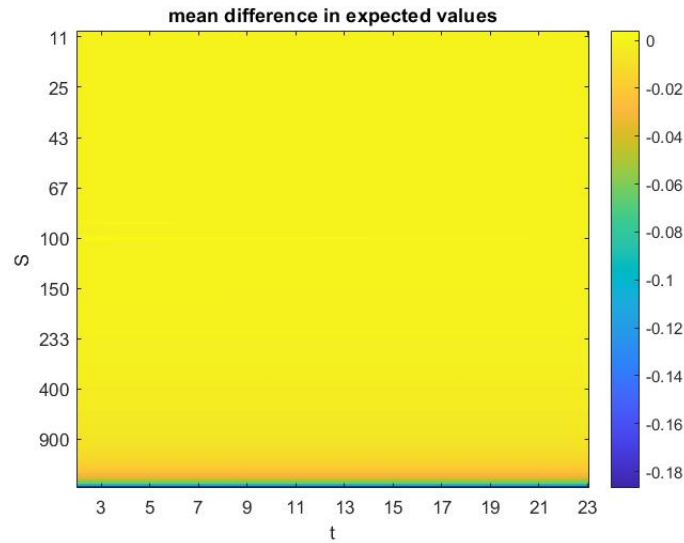


Figure 8.20: Mean absolute difference in expected value of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$.

The characteristic error in variances derived by the same 100 volatility models is displayed in figure 8.21. It shows some oscillation close to the strike price of size $10^{-2} = 10^{-6}strike^2$, but vanishes elsewhere, as one can observe in figure 8.22.

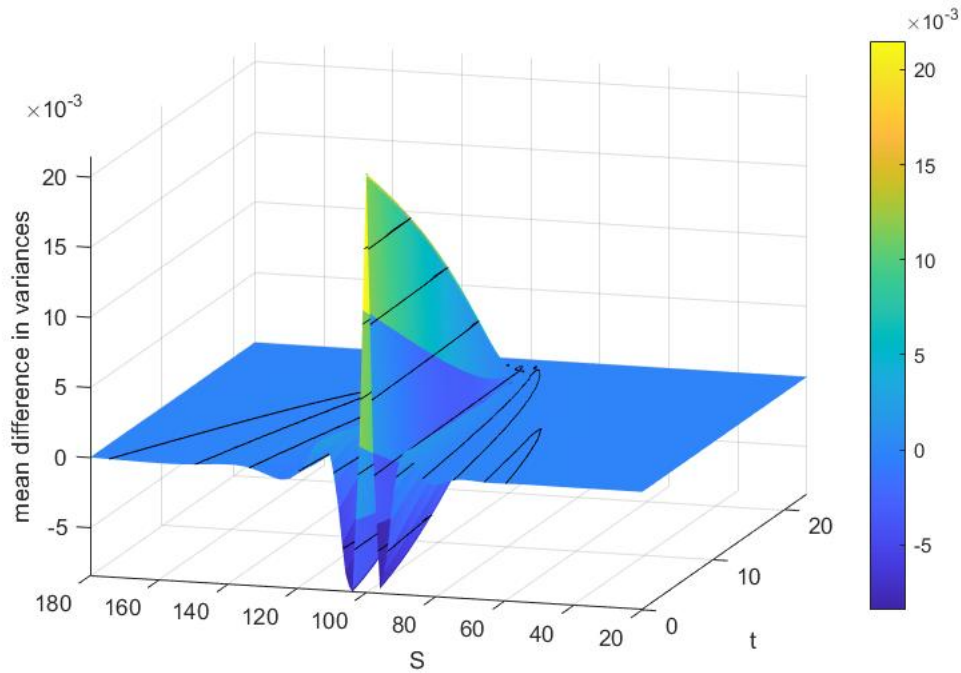


Figure 8.21: Mean absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, $strike = 100$, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$ with contour lines at quarters of its maximum value.

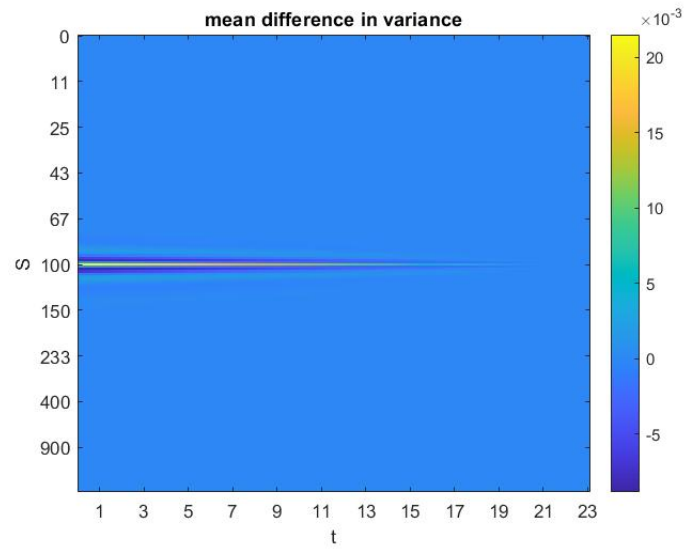


Figure 8.22: Mean absolute difference in variance of the high fidelity and the Bi-Fidelity solution for a European Call option for all S values for 100 volatility models $\Sigma(\Theta, \Delta) = \sigma_{00} + \sigma_{10}\Theta + \sigma_{01}\Delta$, Θ normal distributed, Δ uniform distributed, $T = 23$, *strike* = 100, $K = 1$, $N = 5$, $M_{\zeta}^L = 50$, $N_{\tau}^L = 150$, low fidelity sample points as in 8.13, $M_{\zeta}^H = 200$ and $N_{\tau}^H = 1908$.

Chapter 9

Summary and Conclusion

In this thesis, the Black Scholes equation was considered, where the volatility was assumed to be a random variable, that can be written as a L^2_μ function of another random variable whose density μ is easy to compute. The theoretical background of the generalized Polynomial Chaos expansion was illustrated, since it provides a basis of the Stochastic Galerkin method. This method was then applied to the Black Scholes equation with uncertain volatility in order to transform it to a system of deterministic PDEs. Techniques for solving the system of PDEs were shown and their implementation was explained, before numerical examples were presented. These examples experimentally investigated the influence of the distribution of the volatility on the distribution of the solution and compared the calculated solution to real market data.

Afterwards, the model was extended to volatility depending on finitely many independent random variables. Again, the theoretical background of the Stochastic Galerkin method was explained and the method was applied to the new model. The numerics was adapted and numerical examples were given and compared to those for the model with one random variable.

Since computational cost rises fast as the number of random variables rises, a Bi-Fidelity technique was introduced in order to save computational cost. This method was adapted to the model with volatility depending on two random variables, the numerical implementation was explained and experiments were done comparing the Bi-Fidelity solution to the true solution.

In literature, the Stochastic Galerkin approach is applied to the Black Scholes equation in [PvE09] and [Dra16], where the method of lines or the finite elements method are used for solving the system of PDEs for the gPC coefficients. In this thesis, an explicit finite difference scheme was used. The papers consider dependence on one random variable. This model was extended to volatility depending on finitely many independent random variables in this thesis. Furthermore, a Bi-Fidelity technique for reducing computational cost was presented.

A topic that is still open to further research is the convergence of the truncated Stochastic Galerkin solution to the true solution as mentioned in remark 2. Since the system of

PDEs for the gPC coefficients of the solution is coupled, a truncation (as necessary for computations) changes the calculation of the coefficients and therefore does not lead to the right coefficients. Hence, theorem 6 showing the convergence of the gPC expansion can not be applied.

An extension of the model to random variables that are not necessarily independent could be done. There already exists theory on gPC expansions with dependent random variables, see [Rah18].

Furthermore, the computational cost and the fit of the Stochastic Galerkin model to real data could be compared to the model derived by stochastic Collocation in [NK12] or to other volatility models like local volatility or the Heston model for instance.

An application of this model could be given in derivative pricing risk management. When banks or other institutions emit derivatives, deviations in the volatility as observable in the markets (see e.g. [Rub85], [Sco87] and [GJ10] Tabelle 4.1) mean a risk to them. If they e.g. worked with the original Black Scholes equation and underestimated the volatility, their calculated price would be too low. A calculation with the model derived in this work, however, would give them the possibility to choose a price such that the fair emission price is lower or equal to this price with a certain probability p . To do this, they would have to calculate the (approximate) distribution of $V(0, S_0)$, where S_0 is the value of the underlying asset at time 0, by the truncated gPC expansion 4.6 and take the p quantile as their emission price.

List of Abbreviations

BC	Before Christ
e.g.	for example
etc.	etcetera
EU	European Union
gPC	generalized Polynomial Chaos
i.e.	id est
PDE	partial differential equation
SG	Stochastic Galerkin
w.l.o.g.	without loss of generality
w.r.t.	with respect to

Bibliography

- [AE06a] Herbert Amann and Joachim Escher. *Analysis I*, volume 1 of *Grundstudium Mathematik*. Birkhäuser, 3rd edition, 2006.
- [AE06b] Herbert Amann and Joachim Escher. *Analysis II*, volume 2 of *Grundstudium Mathematik*. Birkhäuser, 2nd edition, 2006.
- [AE08] Herbert Amann and Joachim Escher. *Analysis III*, volume 3 of *Grundstudium Mathematik*. Birkhäuser, 2nd edition, 2008.
- [AN16] Christophe Audouze and Prasanth Nair. A priori error analysis of stochastic Galerkin projection schemes for randomly parametrized ordinary differential equations. *International Journal for Uncertainty Quantification*, 6(4):pp. 287–312, January 2016.
- [BIG03] Adi Ben-Israel and Thomas N.E. Greville. *Generalized inverses*, volume 15 of *CMS books in mathematics*. Springer, 2nd edition, 2003.
- [BS73] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):pp. 638–654, May-June 1973.
- [Chi78] Theodore S. Chihara. *An introduction to orthogonal polynomials*, volume 13 of *Mathematics and its applications*. Gordon and Breach, 1978.
- [CIR85] John C. Cox, Jr. Ingersoll, Jonathan E., and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):pp. 385–407, March 1985.
- [CLV99] Thomas Coleman, Yuying Li, and Arun Verma. Reconstructing the unknown local volatility function. *The Journal of Computational Finance*, 2(3):pp. 77–102, September 1999.
- [Cra71] B. D Craven. Stone’s theorem and completeness of orthogonal systems. *Journal of the Australian Mathematical Society*, 12(2):pp. 211–223, 1971.
- [Cre02] S. Crepey. Calibration of the local volatility in a trinomial tree using Tikhonov regularization. *Inverse Problems*, 19(1):pp. 91–127, December 2002.
- [CS96] G. Crawford and B. Sen. *Derivatives for Decision Makers: Strategic Management Issues*. Wiley Series in Financial Engineering. Wiley, 1996.

- [DJI19] Esther S. Daus, Shi Jin, and Liu Liu. Spectral Convergence of the Stochastic Galerkin Approximation to the Boltzmann Equation with Multiple Scales and Large Random Perturbation in the Collision Kernel. *Kinetic and Related Models*, 12(4):pp. 909–922, 2019.
- [DPW13] Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk. Greedy Algorithms for Reduced Bases in Banach Spaces. *Constructive Approximation*, 37(3):pp. 455–466, 2013.
- [Dra16] Stefanos Drakos. Uncertain volatility derivative model based on the polynomial chaos. *Journal of Mathematical Finance*, 6(1):pp. 55–63, 2016.
- [Dup94] Bruno Dupire. Pricing with a smile. *Risk magazine*, (7):pp. 18–20, 1994.
- [Eli19] Elisabeth Roegele, Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin). Allgemeinverfügung gemäß Artikel 42 MiFIR bezüglich binärer Optionen. https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Aufsichtsrecht/Verfuegung/vf_190701_allgvfg_Binaere_Optionen.html, July 2019.
- [Eva98] Lawrence C Evans. *Partial differential equations*, volume 19 of *Graduate studies in mathematics*. American Mathematical Society, 1998.
- [Fre71] Geza Freud. *Orthogonal polynomials*. Pergamon Pr., 1971.
- [GHO19] Roger Ghanem, David Higdon, and Houman Owhadi. *Handbook of uncertainty quantification*. Springer, 2019.
- [GJ10] Michael Günther and Ansgar Jüngel. *Finanzderivate mit MATLAB*. Vieweg + Teubner, 2nd edition, 2010.
- [GS91] Roger G Ghanem and Pol D Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer New York, 1991.
- [GS97] Roger G Ghanem and Pol D Spanos. Spectral techniques for stochastic finite elements. *Archives of Computational Methods in Engineering*, 4(1):pp. 63–100, 1997.
- [GST07] Amparo Gil, Javier Segura, and N. M. Temme. *Numerical methods for special functions*. Society for Industrial and Applied Mathematics (SIAM), 2007.
- [GW69] Gene H. Golub and John H. Welsch. Calculation of Gauss quadrature rules. *Mathematics of Computation*, 23:pp. 221–230, 1969.
- [GX08] David Gottlieb and Dongbin Xiu. Galerkin method for wave equations with uncertain coefficients. *COMMUNICATIONS IN COMPUTATIONAL PHYSICS*, 3(2):pp. 505–518, February 2008.

- [Hes93] Steven L. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2):pp. 327–343, January 1993.
- [HJ13] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge Univ. Press, 2nd edition, 2013.
- [HR05] Martin Hanke and Elisabeth Rösler. Computation of Local Volatilities from Regularized Dupire Equations. *International Journal of Theoretical and Applied Finance*, 8(2):pp. 207–221, March 2005.
- [Hul12] John C. Hull. *Optionen, Futures und andere Derivate*. wi, Wirtschaft: Higher Education. Pearson, 8th edition, 2012.
- [HW87] John C. Hull and Alan White. The Price of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, 42(2):pp. 281–300, June 1987.
- [Isk18] Armin Iske. *Approximation*. Springer-Lehrbuch Masterclass. 2018.
- [Jan97] Svante Janson. *Gaussian Hilbert Spaces*. Cambridge University Press, June 1997.
- [Kre05] Ulrich Krengel. *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. Vieweg Studium : Aufbaukurs Mathematik. Vieweg, 8th, expanded edition, 2005.
- [KS51] Stefan Kaczmarz and Hugo Steinhaus. *Theorie der Orthogonalreihen*, volume 6 of *Monografie matematyczne*. Chelsea Publ., 2. edition, 1951.
- [LJ18] Liu Liu and Shi Jin. Hypocoercivity based Sensitivity Analysis and Spectral Convergence of the Stochastic Galerkin Approximation to Collisional Kinetic Equations with Multiple Scales and Random Inputs. *SIAM Multiscale Modeling and Simulation*, 16(3):pp. 1085–1114, 2018.
- [LZ20] Liu Liu and Xueyu Zhu. A bi-fidelity method for the multiscale Boltzmann equation with random parameters. *Journal of Computational Physics*, 402, February 2020.
- [Mer73] Richard C. Merton. The Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science*, 4(1):pp. 141–183, Spring 1973.
- [Nar14] Narayan, Akil and Gittelsohn, Claude and Xiu, Dongbin. A Stochastic Collocation Algorithm with Multifidelity Models. *SIAM Journal on Scientific Computing*, 36(2):pp. 495–521, 2014.
- [NK12] Motoi Namihira and David A. Kopriva. Computation of the effects of uncertainty in volatility on option pricing and hedging. *International Journal of Computer Mathematics*, 89(9):pp. 1281–1302, 2012.

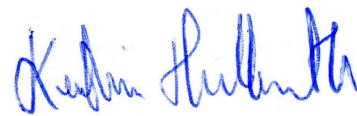
- [Nob97] NobelPrize.org. Awarding the Nobel Prize in Economics to Merton and Scholes. Press Release, October 1997.
- [PR96] Jagdish K. Patel and Campbell B. Read. *Handbook of the normal distribution*, volume 150 of *Statistics*. Dekker, second, revised and expanded edition, 1996.
- [PvE09] Roland Pulch and Cathrin van Emmerich. Polynomial chaos for simulating random volatilities. *Mathematics and Computers in Simulation*, 80(2):pp. 245–255, 2009.
- [Rah18] Sharif Rahman. A polynomial chaos expansion in dependent random variables. *Journal of Mathematical Analysis and Applications*, 464(1):pp. 749–775, August 2018.
- [RKB12] Namita Rajput, Ruhi Kakkar, and Geetanjali Batra. Futures Trading and Its Impact on Volatility of Indian Stock Market. *Asian Journal of Finance and Accounting*, 5(1), December 2012.
- [RM98] Helmut Reisen and Julia Maltzan. Sovereign credit ratings, emerging market risk and financial market volatility. *Intereconomics*, 33(2):pp. 73–82, 1998.
- [RTA16] Smita Roy Trivedi and P. Apte. Central Bank Intervention in USD/INR Market: Estimating Its Reaction Function and Impact on Volatility. *Asia-Pacific Financial Markets*, 23(3):pp. 263–279, 2016.
- [Rub85] Mark Rubinstein. 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. *The Journal of Finance*, 40(2):pp. 455–480, June 1985.
- [SA19] European Securities and Markets Authority. *ESMA Annual Statistical Report: EU Derivatives Markets 2019*. Publications Office, 2019.
- [Sco87] Louis O. Scott. Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. *The Journal of Financial and Quantitative Analysis*, 22(4):pp. 419–438, December 1987.
- [SJ18] Ruiwen Shu and Shi Jin. Uniform regularity in the random space and spectral accuracy of the stochastic Galerkin method for a kinetic-fluid two-phase flow model with random initial inputs in the light particle regime. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(5):pp. 1651–1678, September-October 2018.
- [Str04] John C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematics, second edition, January 2004.
- [Sul15] T. J. Sullivan. *Introduction to Uncertainty Quantification*, volume 63 of *Texts in applied mathematics*. Springer, 2015.

- [Sze59] Gabor Szegő. *Orthogonal polynomials*, volume 23 of *Colloquium publications*. American Math. Soc., revised edition, 1959.
- [TW98] Aslak Tveito and Ragnar Winther. *Introduction to partial differential equations*, volume 29 of *Texts in applied mathematics*. Springer, 1998.
- [Wha07] R.E. Whaley. *Derivatives: Markets, Valuation, and Risk Management*. Wiley Finance. Wiley, 2007.
- [Wie38] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):pp. 897–936, October 1938.
- [Wil62] Herbert S. Wilf. *Mathematics for the physical sciences*. Wiley, 1962.
- [Xiu10] Dongbin Xiu. *Numerical Methods for Stochastic Computations*. Princeton University Press, 2010.
- [Zhu14] Zhu, Xueyu and Narayan, Akil and Xiu, Dongbin. Computational Aspects of Stochastic Collocation with Multifidelity Models. *SIAM/ASA Journal on Uncertainty Quantification*, 2(11):pp. 444–463, 2014.
- [ZWCS13] You-lan Zhu, Xiaonan Wu, I-Liang Chern, and Zhi-zhong Sun. *Derivative Securities and Difference Methods*. Springer, second edition, 2013.

Declaration of Originality

I confirm that this assignment is my own work and that I have not sought or used inadmissible help of third parties to produce this work and that I have only used the referenced sources. This work has not yet been submitted to another examination institution for obtaining an academic degree and has not yet been published.

Würzburg, 30th April 2020



Kathrin Hellmuth