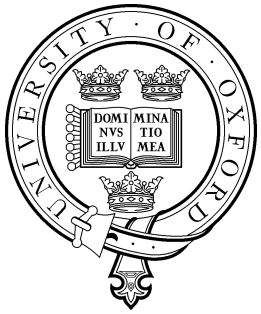
Expansion Methods for High-Dimensional PDEs in Finance



Rasmus Wissmann

St Catherine's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2014

To my parents.

Still round the corner there may wait

A new road or a secret gate

- J. R. R. Tolkien

Acknowledgements

I am indebted to everyone supporting me in the making of this thesis. It has been an exceptional intellectual, professional and personal journey.

First an foremost I am grateful to my supervisor Dr. Christoph Reisinger. He has always been incredibly supportive and provided me with the initial suggestion to look into expansion-type methods. Throughout the past years we have worked closely together on the topics documented in this thesis and he has helped me shape and develop my understanding of quantitative finance, numerical mathematics and computational methods.

The Mathematical and Computational Finance group in Oxford has been a very pleasant and supportive environment. I would like to thank all its members for the numerous seminars, conferences, discussions and insightful informal coffee breaks. I am grateful for the chance to work as a tutor and teaching assistant. Professor Ben Hambly, Professor Mike Giles and Dr. Lajos Gergely Gyurkó provided invaluable feedback and inspiration during my transfer and confirmation vivas. Among fellow DPhil students I would like to particularly mention and thank Jan Hendrik Witte, Alan White, and Pedro Vitoria.

This work would not have been possible without the generous financial support of the University of Oxford, the Mathematical Institute, St. Catherine's College, the Engineering and Physical Sciences Research Council, Nomura, the German National Academic Foundation (Studienstiftung des deutschen Volkes), the Oxford Man-Institute for Quantitative Finance, and G-Research.

Lastly, I would like to thank the great people of University College and St. Catherine's College, my close friends Andreas Bachmeier, Emanuel Ferm, Jonas Hirsch, Joanna Palermo and Pia-Ramona Wojtinnek, as well as my family. They have enabled this journey and made it more enjoyable than I could have hoped.

Abstract

We develop expansion methods as a new computational approach towards high-dimensional partial differential equations (PDEs), particularly of such type as arising in the valuation of financial derivatives. The proposed methods are extended from [41] and use principal component analysis (PCA) of the underlying process in combination with a Taylor expansion of the value function into solutions to low-dimensional PDEs. They enable calculation of highly accurate approximate solutions with computational complexity polynomial in the number of dimensions for PDEs with a low number of dominant principal components.

For the case of PDEs with constant coefficients, we show existence of expansion solutions and prove theoretical error bounds. We give a precise characterisation of when our methods can be applied and construct specific examples of a first and second order version. We provide numerical results showing that the empirically observed convergence speeds are in agreement with the theoretical predictions. For the case of PDEs with varying coefficients, we give a heuristic motivation using the Parametrix approach and empirically test the methods' accuracy for a range of variable parameter stock models.

We demonstrate the applicability of our expansion methods to real-world securities pricing problems by considering path-dependent and early-exercise options in the LIBOR market model. Using the example of Bermudan swaptions and Ratchet floors, which are considered difficult benchmark problems, we give a careful analysis of the numerical accuracy and computational complexity. We are able to demonstrate that for problems with medium to high dimensionality, around $60 - 100$, and moderate time horizons, the presented PDE methods deliver results comparable in accuracy to benchmark state-of-the-art Monte Carlo methods in similar or (significantly) faster run time.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation and Problem Definition

This thesis is concerned with the numerical solution of high-dimensional partial differential equations (PDEs) in mathematical finance applications via expansion methods. Specifically, we deal with the problem of calculating the value of a function v at a given point (x_0, T), where v is the solution of the Cauchy problem for the parabolic PDE

$$\frac{\partial}{\partial t} v(x,t) = \mathcal{L}(x,t)v(x,t)\,,\ (x,t) \in \mathbb{R}^N \times [0,T] \tag{1.1}$$

$$\mathcal{L}(x,t) = \sum_{i,j=1}^{N} A_{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} B_i(x,t)\frac{\partial}{\partial x_i} \tag{1.2}$$

$$v(x,0) = g(x)\,\forall x \in \mathbb{R}^N. \tag{1.3}$$

The spatial dimensionality of the problem is typically on the order of $N \sim 5 - 100$. While we focus on this type of PDE, our basic ideas are applicable to other types as well.

Our primary motivation for solving such high-dimensional parabolic PDEs derives from pricing multi-asset derivatives. These are financial instruments whose payout depends on the value of a number of underlying financial assets at one or multiple future dates. The value of the financial derivative itself thus depends on the dynamics of the underlyings. Using the mathematical theory of financial markets and arbitrage-

free pricing of traded assets, the dynamics of these assets can be described as stochastic processes. In most common models, the values of financial derivatives are then equivalently characterised as the expected value of a payoff functional under these stochastic processes or as the solution of an associated partial (integro-)differential equation.

Consider for example a set of financial assets with value processes X_i, $1 \leq i \leq N$, satisfying

$$\mathrm{d}X_i(t)/X_i(t) = \mu_i(X,t)\,\mathrm{d}t + \sigma_i(X,t)\,\mathrm{d}W_i^{\mathcal{Q}}(t)$$

on a probability space $\{\Omega, \mathcal{F}, \mathcal{P}\}$ with filtration $\{\mathcal{F}_t\}, t \in [0,T]$, $T \in \mathbb{R}^+ \cup \infty$ and $X = (X_1, \ldots, X_N)$. Here $\sigma : \mathbb{R}^N \times [0,T] \to \mathbb{R}_0^{N,+}$ is the volatility, $\mu : \mathbb{R}^N \times [0,T] \to \mathbb{R}^N$ is the drift, $W^{\mathcal{Q}}$ is a standard Brownian motion under the risk-neutral measure $\mathcal{Q}$ and $\rho : \mathbb{R}^N \times [0,T] \to \mathbb{R}^{N\times N}$ is the correlation matrix, i.e.,

$$\langle\, \mathrm{d}W_i^{\mathcal{Q}},\, \mathrm{d}W_j^{\mathcal{Q}} \rangle = \rho_{ij}\,\mathrm{d}t \quad \forall i,j \in 1,\ldots,N.$$

A European option is characterised by its payout function $G : \mathbb{R}^N \to \mathbb{R}$, which determines the amount $G(X(T))$ its holder receives at time $t = T$. The arbitrage-free value of the option relative to a numéraire $\mathcal{N}$ is then given by

$$v(X(t),t) = \mathbb{E}^{\mathcal{Q}}\left(\frac{G(X(T))}{\mathcal{N}(T)} \middle| \mathcal{F}_t \right), \tag{1.4}$$

assuming that standard technical conditions hold[1]. Here $G(\cdot)$ is the absolute payoff at time T. By the Feynman-Kac theorem, v satisfies the parabolic PDE

$$\frac{\partial}{\partial t}v(x,t) = -\sum_{i=1}^{N} \mu_i x_i \frac{\partial}{\partial x_i} v(x,t) - \sum_{i,j=1}^{N} \tfrac{1}{2}\sigma_i\sigma_j\rho_{ij}x_i x_j \frac{\partial^2}{\partial x_i \partial x_j} v(x,t) \tag{1.5}$$

$$v(x,T) = g(x)\,\forall x \in \mathbb{R}^N \tag{1.6}$$

on $\mathbb{R}^N \times [0,T]$. For simplicity of notation, we have used the relative payoff $g(\cdot) = G(\cdot)/\mathcal{N}(T)$. For path-dependent or early-exercise derivatives additional conditions

[1] See, e.g., [14].

are introduced. We can bring (1.5)-(1.6) to overlap with (1.1)-(1.3) by reversing time.

1.2 State of the Art

The solution of (1.1)-(1.3) at a given coordinate point (x_0, T), $v(x_0, T)$, can be calculated analytically only for simple and low-dimensional models and financial derivatives. In general, exact solution techniques are not available. Analytic approximations and simplifications exist, but are only practical for a limited class of problems and are difficult to develop. An illustrative example of such advanced analytical techniques is the model introduced in [35] for the PDEs associated with the specific problem of pricing a path-dependent interest rate derivative called a Ratchet cap. This inspired our investigation into this type of financial derivative, as described in Chapter 6.

In spite of isolated successes in such special cases, accurate closed-form approximations for complex multi-asset models are generally not possible and one has to resort to computational approaches. The two dominant classes of such methods in derivative pricing are Monte Carlo methods (see, e.g., [16]), for estimating the expectation (1.4) via simulation, and discretisation methods (see, e.g., [1, 44]) for approximating the solution to the respective PDE (1.5)–(1.6). The latter group broadly includes lattice, spectral and Fourier methods.

Simulation methods scale favourably with the dimension of the process and are well suited to track path-dependent quantities which determine the payoff of some exotic derivatives depending not only one the final value of its asset, but also on values at intermediate times. Disadvantages of simulation methods are that the convergence in the number of samples is typically slow and that they require additional approximations to deal with early-exercise derivatives. This type of options can be exercised before maturity and thus raise the question of what optimal exercise strategies look like.

Conversely, PDE discretisation methods incorporate early-exercise features easily and allow fast convergence in the number of grid points used in each direction, which makes them very efficient for low-dimensional problems. However, they quickly become intractable as the dimensionality increases: The effort to solve N-dimensional PDEs numerically with standard grid-based methods grows exponentially with N, a behaviour known as the curse of dimensionality. Even sophisticated PDE methods tailored to high-dimensional approximation, such as those based on sparse grids, are typically not able to deal with practical problems where N exceeds about five to eight, see [21, 23, 32, 41]. Given especially the advantages in dealing with early-exercise, it would be not only of academic interest but also practically very relevant, to be able to solve generic derivative pricing problems with PDE methods for dimensionality on the order of $5 - 100$.

1.3 Expansion Methods

In this thesis we propose and investigate a hybrid analytical-numerical expansion method. We first transform the N-dimensional PDE problem for v into a standard form for a new function u and approximate its solution as a sum of solutions u^ν of lower-dimensional PDEs. The latter are then evaluated numerically as u^ν_{num}. This results in a chain of approximations

$$v(x_0, T) \approx u(z_0, T) \approx \sum_{(w,\nu)\in\xi} w \cdot u^\nu(z_0, T) \approx \sum_{(w,\nu)\in\xi} w \cdot u^\nu_{num}(z_0, T),$$

where ξ is a index set describing the specific expansion method.

Computing what is typically a lower order polynomial number $|\xi| = O(N^p)$ of low-dimensional PDEs of maximum dimension $d \ll N$ is vastly (often exponentially) more efficient than calculating the original solution. In fact, we will see that $p = 1$ and $d = 2$ is often sufficient for practically adequate accuracy. We will analyse the individual steps in this approach and discuss how to construct accurate expansion

schemes for a range of applications. Overall, this expansion method was inspired by, and is a generalisation of, a Principal Component Analysis(PCA)-based approach first introduced in [41].

Broadly speaking, the underlying principle of this and related approaches is an anchored ANOVA-type decomposition (see [37]) of a solution $u(z)$, $z \in \mathbb{R}^N$, into

$$\begin{aligned} u(z) &= u_0(a) + \sum_{i=1}^{N} u_i(a; z_i) + \sum_{\substack{i,j=1 \\ i<j}}^{N} u_{i,j}(a; z_i, z_j) + \ldots + u_{1,\ldots,N}(a; z_1, \ldots, z_N) \\ &= \sum_{v \subseteq \{1,\ldots,N\}} u_v(a; z^v), \end{aligned} \tag{1.7}$$

where we associate $u_\emptyset$ with u_0, $u_{\{i\}}$ with u_i etc. The terms on the right-hand side each only depend on a subset of the coordinates, $z^v = (z_{i_1}, \ldots, z_{i_{|v|}})$, and a chosen 'anchor' $a = (a_1, \ldots, a_N)$. This has been successfully applied to quadrature problems from finance in [18], and its relation to the PDE expansions in [41], which form the basis for the present work, is highlighted in [39] and [42].

Much of the theoretical work in this thesis can be understood as using Taylor expansions to derive and analyse efficient anchored ANOVA-type decompositions for parabolic PDEs with coefficients and initial conditions typical for derivative pricing applications. The central source of complexity reduction will always be the approximation of the solution of a high-dimensional PDE through a linear combination of solutions of derived lower-dimensional PDEs. It is thus helpful to keep equation (1.7) in mind, in particular if one is already familiar with this kind of techniques.

Key to the efficiency of this approximation as a numerical method is that the relative importance of u_v decays rapidly with increasing $|v|$, as $|v|$ is the dimension of the coordinate space of u_v. This can be achieved by a coordinate transformation of the underlying stochastic process and of the corresponding forward or backward PDE. Optimal linear transformations taking into account the payoff function are analysed in [24], while here we consider the principal components of the covariance matrix

Σ of the Brownian driver of the process. The accuracy of the approximate solution obtained by truncating (1.7) after a small number of terms with small $|v|$ then depends largely on the (relative) sizes of the eigenvalues λ_i of Σ, $1 \leq i \leq N$. This will be motivated in Chapter 3 by expanding the value function in λ_i. In this approach we follow [41], which first introduced this idea for vanilla basket options.

1.4 Outline and Contributions

The goal underlying our work was to turn the basic idea of a PCA-based approximation for vanilla basket options into a widely applicable method for numerically solving high-dimensional PDEs, in particular in the context of pricing financial derivatives. To achieve this we had to generalise the methods to encompass a broad class of relevant PDEs and boundary conditions and to provide a theoretical justification of their applicability and error analysis. We then had to empirically study the accuracy, convergence and run times achieved in numerical experiments and to compare them to our theoretical predictions.

In the context of pricing financial derivatives, we wanted to understand and characterise the applicability of expansion methods, and to numerically verify their performance for practically relevant securities within realistic market models. For this we had to relate, contrast and benchmark expansion methods with and against existing methods, such as Monte Carlo, ANOVA and factor models. Lastly, we wanted to provide a usable implementation of PDE expansion methods as proof of concept and reference.

This describes the work plan guiding our research. The rest of this thesis is structured as follows:

- *Chapter 2* introduces expansion methods for high-dimensional PDEs and pricing of financial derivatives. In it we explain the motivation from Principal Component Analysis (PCA) of the underlying covariance matrix, demonstrate

the connection to the multi-dimensional heat equation and describe related approaches.

- *Chapter 3* explores the specific case of constant-coefficient PDEs. We construct expansion methods via Taylor series expansion and derive theoretical error bounds.

- *Chapter 4* applies the approach to European options in the Black-Scholes model. We carefully analyse the empirical convergence and error behaviour for a first and a second order expansion method, and in doing so demonstrate alignment between theoretical predictions and numerical results.

- *Chapter 5* generalises the application of expansion methods to non-constant coefficient PDEs. We motivate the approach via theoretical considerations and provide a proof of concept by conducting numerical experiments for a range of variable coefficient models.

- *Chapter 6* applies the expansion method to two kinds of financial derivatives, Bermudan swaptions and Ratchet floors, in the LIBOR market model. These practically relevant examples are challenging test cases and among other things demonstrate that our PDE-based approach can deliver superior performance in cases where Monte Carlo-based methods struggle.

- *Chapter 7* introduces our publicly available pricing framework, EMDP: Expansion Method Derivative Pricer. Written in MATLAB, this implementation makes available many of the concepts introduced in this thesis.

- *Chapter 8* summarises our results and describes possible directions for future research.

In line with the above stated goal of developing a general framework for applying the PDE expansion method, the main contributions of our work are

- extending the PDE expansion method for derivative pricing from simple, log-normal equity basket models to complex, practically relevant applications with high-dimensional underlying processes, such as path-dependent and early-exercise options on the LIBOR curve;
- proving applicability and error bounds for the PCA-based PDE expansion method for constant coefficients based on properties of the initial condition and the PDE coefficients themselves;
- investigating the error terms for non-constant PDE coefficients and providing an approximate solution via localisation;
- conducting extensive numerical studies and demonstrating agreement between theoretical bounds and experimental results;
- benchmarking the PDE expansion method against widely used Monte Carlo methods for options on the yield curve and thereby demonstrating for the first time that the PDE expansion method can outperform state-of-the-art Monte Carlo methods for such complex and high-dimensional applications;
- presenting a systematic and generic approach to construct higher order approximations and giving numerical results demonstrating clearly the accuracy improvement achieved;
- providing a MATLAB-based reference implementation as proof-of-concept and basis for future research.

The following publications are based on the work presented in this thesis:

- Numerical Valuation of Derivatives in High-Dimensional Settings via PDE Expansions, C. Reisinger and R. Wissmann, 2013, to appear in J. of Comp. Fin., http://arxiv.org/abs/1209.1909

- Error bounds for a localised PDE expansion method, C. Reisinger and R. Wissmann, 2014, in preparation for submission

- EMPD: Expansion Method Derivative Pricer, a MATLAB-based framework for financial derivative pricing based on the expansion methods introduced in this thesis, https://github.com/rasmuswissmann/EMDP .

Chapter 2

Expansion Methods

2.1 Problem Definition

Consider a N-dimensional parabolic PDE

$$\frac{\partial v}{\partial t} = \mathcal{D}v \tag{2.1}$$

$$\mathcal{D} = \sum_{i,j=1}^{N} A_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} B_i \frac{\partial}{\partial x_i} \tag{2.2}$$

$$v(x,0) = h(x) \quad \forall x \in \mathbb{R}^N, \tag{2.3}$$

for a function $v : \mathbb{R}^N \times [0,T] \to \mathbb{R}$, $T > 0$. The coefficients A_{ij} and B_i in the differential operator $\mathcal{D}$ are functions from $\mathbb{R}^N \times [0,T]$ to $\mathbb{R}$.[1] $h : \mathbb{R}^N \to \mathbb{R}$ specifies the initial condition at $t = 0$.

We are interested in approximating the value of $v(x_0, T)$ for a given, fixed $x_0 \in \mathbb{R}^N$. For this purpose, we split the full problem into a number of lower-dimensional problems and combine their solutions into an approximate solution $\tilde{v}^{\xi}(x_0, T)$ of the form

$$\tilde{v}^{\xi}(x_0, T) = \sum_{(w,\nu)\in\xi} w\ \tilde{v}^{\nu}(x_0, T), \tag{2.4}$$

where $\xi = \{(w_1, \nu_1), (w_2, \nu_2), \ldots, (w_n, \nu_n)\}$, $w_i \in \mathbb{R}$, $\nu_i \subseteq \{1, \ldots, N\}$ for all $1 \leq i \leq n$ and each $\tilde{v}^{\nu}(x,t)$ is the solution of a PDE with low dimensionality of order $O(|\nu|)$.

[1] For simplicity we have omitted a zero order term, but this could easily be included. In fact, a term of the form $c(t)v(x,t)$ can directly be accounted for by the substitution $\exp(\int_0^t c(\tau)d\tau)v(x,t)$.

Our investigation of expansion methods was initially motivated by the problem of pricing complex financial derivatives, which requires the calculation of the expected payout under a probabilistic measure determined by the market model. This can equivalently be expressed as a PDE problem of the form (2.1)-(2.3). The A_{ij} are then related to the covariance matrix of the stochastic processes describing the underlying financial assets. Due to the nature of financial markets, there are often only one or few dominant factors driving the dynamics, essentially forming a hidden low-dimensional structure approximating the full problem. Our expansion methods are designed to exploit such structures and do so automatically, i.e., without having to manually identify them explicitly for every given market structure.

Sticking with this theme, we will use examples motivated from financial derivative pricing to illustrate core ideas throughout this thesis as well as in our numerical experiments. It is worth stressing that expansion methods are by no means restricted to PDEs in mathematical finance.

2.2 Transformation

As a first step, we transform the PDE into a standard form by using an orthogonal transformation and subsequent translation of the spatial coordinates. For constant coefficients, this allows us to diagonalize the PDE and reduce it to the multi-dimensional heat equation.

Definition 1. *For a given orthogonal matrix $Q \in \mathbb{R}^{N\times N}$ and differential equation (2.2), define a shift function $\beta : \mathbb{R}^N \times [0,T] \to \mathbb{R}^N$ componentwise by*

$$\beta_i(x,t) := \sum_j^N Q_{ji} \int_0^t B_j(x,t')dt' \tag{2.5}$$

for $1 \leq i \leq N$.

We then introduce new spatial coordinates z via

$$z = Q^T x + \beta(x_0, t) \tag{2.6}$$

and set $z_0 = Q^T x_0 + \beta(x_0, T)$. We write $x(z,t) = Q(z - \beta(x_0,t))$ for the inverse transform.

Proposition 2. *The PDE problem (2.1)–(2.3) transforms into*

$$\frac{\partial u}{\partial t} = \mathcal{L}u \tag{2.7}$$

$$\mathcal{L} = \sum_{k,l=1}^{N} \lambda_{kl} \frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k=1}^{N} \kappa_k \frac{\partial}{\partial z_k} \tag{2.8}$$

$$u(z,0) = g(z) := h(x(z,0)) \tag{2.9}$$

for a function $u : \mathbb{R}^N \times [0,T] \to \mathbb{R}$, $T > 0$ *where* $u(z,t) = v(x,t)$. *Here we have introduced the PDE coefficient functions*

$$\lambda_{kl}(z,t) = \sum_{i,j=1}^{N} Q_{ik} Q_{jl} A_{ij}(x(z,t),t) \tag{2.10}$$

$$\kappa_k(z,t) = \sum_{i=1}^{N} Q_{ik} \left[B_i(x(z,t),t) - B_i(x_0,t) \right]. \tag{2.11}$$

Proof. Directly via coordinate transformation. See Appendix A.1. □

Remark 3. *In the important special case that the* $B_j(x,t)$ *do not depend on the spatial coordinates* x, *the difference in equation (*2.11*) vanishes identically and thus* $\kappa(z,t) \equiv 0$. $\mathcal{L}(z,t)$ *then contains second order terms only.*

What is the point of this transformation, given that both PDE problems are structurally similar? The answer is that dependent on the properties of the original PDE coefficients, it is often possible to concentrate a significant part of the dynamics of the PDE in just a few dimensions with an appropriate choice of Q. This is particularly true if the PDE models a process that naturally has a hidden lower-dimensional structure. Furthermore — as we shall see — the transformation often allows us to simplify the structure of the transformed PDE.

2.3 The Constant Coefficient Case: N-dimensional Heat Equation

To illustrate the advantages of using the transformed problem, consider the case of constant PDE coefficients A_{ij} and B_i. We can then choose Q to be the matrix of eigenvectors of the constant, positive semidefinite[2] and symmetric coefficient matrix $A = (A_{ij})_{1\leq i,j\leq N}$ sorted by eigenvalue size[3], i.e.,

$$Q = (q_1, \ldots, q_N), \;\; Aq_i = \lambda_i q_i, \;\; \lambda_1 \geq \ldots \geq \lambda_N \geq 0,$$

and get

$$\lambda(z,t) = const. = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}.$$

$\mathcal{L}$ simplifies to the N-dimensional heat operator

$$\mathcal{L} = \sum_{k=1}^{N} \lambda_k \frac{\partial^2}{\partial z_k^2} \tag{2.12}$$

and we get

$$\frac{\partial u}{\partial t} = \sum_{k=1}^{N} \lambda_k \frac{\partial^2 u}{\partial z_k^2} \tag{2.13}$$

$$u(z,0) = g(z) \tag{2.14}$$

for $z \in \mathbb{R}^N$, $t \in (0,T)$. The diffusion coefficients are the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$ of $(A_{ij}(x_0,t_0))_{1\leq i,j\leq N}$.

The solution of (2.13)–(2.14) is well known and can explicitly be written as

$$u(z,t) = (\Phi^N \star g)(z,t) \;\; = \;\; \int_{\mathbb{R}^N} \Phi^N(y,t)\, g(z-y)\, dy,$$

with the Green's function

$$\Phi^N(y,t) = \prod_{k=1}^{N} \frac{\exp(-y_k^2/4\lambda_k t)}{\sqrt{4\pi\lambda_k t}}.$$

[2]Positive definiteness is guaranteed by the PDE being parabolic.

[3]If Σ has eigenvalues with multiplicity larger than 1, then this decomposition is not unique. However, this does not pose a problem for our approach: We can simply choose any such matrix Q.

Similarly, for a function u^ν that only contains the diffusion terms relating to a index set $\nu \subseteq \{1, \dots, N\}$,

$$u^\nu(z,t) = \int_{\mathbb{R}^N} \Phi^\nu(y,t)\, g(z-y)\, dy,$$
$$\Phi^\nu(y,t) = \prod_{k\in\nu} \frac{\exp(-y_k^2/4\lambda_k t)}{\sqrt{4\pi\lambda_k t}} \prod_{k\in\{1,\dots,N\}\setminus\nu} \delta(y_k), \tag{2.15}$$

where $\delta(\cdot)$ is the Dirac delta function.

For non-constant coefficients, an exact transformation from (2.7) to (2.13) may not always be possible, but locally – around $(z_0, 0)$ – (2.13) will still be a good approximation to (2.7). We will elaborate on this in Section 2.5.

2.4 Dimensionality Reduction via Expansion

In the following, we introduce the basic ideas behind the expansion in equation (2.4). We will first give a general framework as well as a description of the resulting error, before moving to a specific natural choice for the lower-dimensional PDE problems. Overall, we are looking to calculate approximate PDE solutions via expansions of PDE problems into related, lower-dimensional ones. Solutions of PDEs restricted to a given subset of dimensions are particularly well-suited for this.

Definition 4. *Given a PDE problem of the form (2.7)–(2.9), an index set $\nu \subseteq \{1, \dots, N\}$, and a differential operator $\mathcal{L}^\nu(z,t)$ of the form (2.7), such that it only contains derivatives in directions z_j, $j \in \nu$, define u^ν as the solution of the $|\nu|$-restricted PDE problem*

$$\frac{\partial}{\partial t} u^\nu(z,t) = \mathcal{L}^\nu(z,t) u^\nu(z,t) \tag{2.16}$$
$$u^\nu(z,0) = g(z). \tag{2.17}$$

Note that since $\mathcal{L}^\nu$ only operates on the dimensions in the index set ν, the problem of calculating $u^\nu(z_0, T)$ *for a fixed* $z_0 \in \mathbb{R}^N$ is of spatial dimension $|\nu|$. Calculating the full solution $u^\nu(z,T)$ for all values of $z \in \mathbb{R}^N$ is still an N-dimensional problem.

Definition 5. *For a given* $\xi = \{(w_1, \nu_1), (w_2, \nu_2), \ldots, (w_n, \nu_n)\}$*, where* $w_i \in \mathbb{R}$ *and* $\nu_i \subseteq \{1, \ldots, N\}$ *for all* $1 \leq i \leq n$*, we define*

$$u^{\xi}(z,t) = \sum_{(w,\nu)\in\xi} w\, u^{\nu}(z,t). \tag{2.18}$$

We call ξ *an* expansion encoding.

The idea is that ξ encodes an approximate solution for u via expansion into solutions u^{ν} of lower-dimensional PDEs. Setting $\tilde{v}^{\nu}(x_0, T) = u^{\nu}(z(x_0, T), T) = u^{\nu}(z_0, T)$ now gives an approximate solution for $v(x_0, T)$ of the form specified in equation (2.4).

Example 6. *A principal component approximation using the first* r *terms can be encoded by*

$$\xi = \{(1, \{1, \ldots, r\})\}.$$

The full solution $u^{\xi} = u$ *without expansion error is given by* $r = N$ *and* $\xi = \{(1, \{1, \ldots, N\})\}$.

Example 7. *For the expansion*

$$\begin{aligned}\xi = \{&(1 + r - N, \{1, \ldots, r\}), (1, \{1, \ldots, r, r+1\}), (1, \{1, , \ldots, r, r+2\}),\\ &(1, \{1, , \ldots, r, r+3\}), \ldots, (1, \{1, , \ldots, r, N\})\},\end{aligned} \tag{2.19}$$

the approximation is

$$\begin{aligned}u^{\xi} &= (1 + r - N)u^{\{1,\ldots,r\}} + \sum_{k=r+1}^{N} u^{\{1,\ldots,r,k\}}\\ &= u^{\{1,\ldots,r\}} + \sum_{k=r+1}^{N} \left(u^{\{1,\ldots,r,k\}} - u^{\{1,\ldots,r\}}\right).\end{aligned}$$

In Sections 3.1 and 3.3 we will show how to systematically derive it from a first order Taylor expansion around the first r *eigenvalues.*

We will use the notation

$$\hat{u}^\xi := u^\xi - u \tag{2.20}$$

for the expansion error and observe that it is – as the difference of two PDE solutions – itself the solution of a PDE:

Theorem 8. *The expansion error $\hat{u}^\xi(z,t) = u^\xi(z,t) - u(z,t)$ solves the non-homogeneous PDE*

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}^\xi(z,t) &= \mathcal{L}(z,t)\hat{u}^\xi(z,t) + f^\xi(z,t,u^\nu) \quad \forall (z,t) \in \mathbb{R}^N \times (0,T] \\ \hat{u}^\xi(z,0) &= \left(\sum_{(w,\nu)\in\xi} w - 1\right) g(z) \quad \forall z \in \mathbb{R}^N\end{aligned}$$

with source term

$$\begin{aligned}f^\xi(z,t,u^\nu) &= \sum_{(w,\nu)\in\xi} w f^\nu(z,t,u^\nu) \\ f^\nu(z,t,u^\nu) &= \left[\mathcal{L}^\nu(z,t) - \mathcal{L}(z,t)\right] u^\nu(z,t).\end{aligned}$$

Proof. By definition

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}^\xi(z,t) &= \frac{\partial}{\partial t}(u^\xi(z,t) - u(z,t)) \\ &= \sum_{(w,\nu)\in\xi} w\mathcal{L}^\nu(z,t)u^\nu(z,t) - \mathcal{L}(z,t)u(z,t) \\ &= \mathcal{L}(z,t)\hat{u}^\xi(z,t) + \sum_{(w,\nu)\in\xi} w\left[\mathcal{L}^\nu(z,t) - \mathcal{L}(z,t)\right] u^\nu(z,t)\end{aligned}$$

for all $(z,t) \in \mathbb{R}^N \times (0,T]$ and

$$\hat{u}^\xi(z,0) = \sum_{(w,\nu)\in\xi} w u^\nu(z,0) - u(z,0) = \left(\sum_{(w,\nu)\in\xi} w - 1\right) g(z).$$

□

Remark 9. *An expansion with expansion encoding ξ should satisfy $\sum_{(w,\nu)\in\xi} w = 1$. The expansion error $\hat{u}^\xi(z,t)$ then vanishes at $t = 0$. This is true for the encoding in Examples 6 and 7.*

We now move on to a particular choice for the lower-dimensional operator $\mathcal{L}^\nu$. We call it the *natural restriction* of the original operator $\mathcal{L}$ to the index set ν:

Definition 10. *Given an index set $\nu \subseteq \{1, \dots, N\}$ and a differential operator*

$$\mathcal{L}(z,t) = \sum_{k,l=1}^{N} \lambda_{kl}(z,t) \frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k=1}^{N} \kappa_k(z,t) \frac{\partial}{\partial z_k}$$

define the natural restriction $\mathcal{L}^\nu_{nat}$ of $\mathcal{L}$ to ν as

$$\mathcal{L}^\nu_{nat}(z,t) = \sum_{(k,l)\in\nu\times\nu} \lambda_{kl}(z,t) \frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k\in\nu} \kappa_k(z,t) \frac{\partial}{\partial z_k}.$$

Corollary 11. *If $\mathcal{L}^\nu$ is the natural restriction $\mathcal{L}^\nu_{nat}$ of $\mathcal{L}$ to ν, then the function $f^\nu(z,t,u^\nu)$ in Theorem 8 is given by*

$$\begin{aligned} f^\nu(z,t,u^\nu) = &- \sum_{(k,l)\in\{1,\dots,N\}^2\setminus\nu\times\nu} \lambda_{kl}(z,t) \frac{\partial^2}{\partial z_k \partial z_l} u^\nu(z,t) \\ &- \sum_{k\in\{1,\dots,N\}\setminus\nu} \kappa_k(z,t) \frac{\partial}{\partial z_k} u^\nu(z,t). \end{aligned}$$

Theorem 12. *If $\mathcal{L}^\nu$ is the natural restriction $\mathcal{L}^\nu_{nat}$ of $\mathcal{L}$ to ν, then the expansion error $\hat{u}^\xi$ is the solution of the non-homogeneous PDE*

$$\begin{aligned} \frac{\partial}{\partial t}\hat{u}^\xi(z,t) = &\mathcal{L}^{\nu_\xi}(z,t)\hat{u}^\xi(z,t) + f^{\nu_\xi}(z,t,u^\nu) \\ &+ \tilde{f}(z,t,u) \quad \forall (z,t) \in \mathbb{R}^N \times (0,T] \\ \hat{u}^\xi(z,0) = &\left(\sum_{(w,\nu)\in\xi} w - 1 \right) g(z) \quad \forall z \in \mathbb{R}^N \end{aligned} \tag{2.21}$$

with source terms

$$\begin{aligned} f^{\nu_\xi}(z,t,u^\nu) &= \sum_{(w,\nu)\in\xi} w \left[\mathcal{L}^\nu(z,t) - \mathcal{L}^{\nu_\xi}(z,t) \right] u^\nu \\ \tilde{f}(z,t,u) &= \left[\mathcal{L}^{\nu_\xi}(z,t) - \mathcal{L}(z,t) \right] u(z,t). \end{aligned}$$

Here we have set $\nu_\xi = \bigcap_{(w,\nu)\in\xi} \nu$ and

$$\mathcal{L}^{\nu_\xi}(z,t) = \sum_{(k,l)\in\nu_\xi\times\nu_\xi}^{N} \lambda_{kl}(z,t) \frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k\in\nu_\xi} \kappa_k(z,t) \frac{\partial}{\partial z_k}.$$

Proof. By definition

$$\begin{aligned}
\frac{\partial}{\partial t}\hat{u}^{\xi}(z,t) &= \frac{\partial}{\partial t}\left[u^{\xi}(z,t) - u(z,t)\right] \\
&= \sum_{(w,\nu)\in\xi} w\mathcal{L}^{\nu}(z,t)u^{\nu}(z,t) - \mathcal{L}(z,t)u(z,t) \\
&= \left[\mathcal{L}^{\nu_\xi}(z,t) - \mathcal{L}^{\nu_\xi}(z,t)\right] \sum_{(w,\nu)\in\xi} wu^{\nu}(z,t) \\
&\quad + \sum_{(w,\nu)\in\xi} w\mathcal{L}^{\nu}(z,t)u^{\nu}(z,t) \\
&\quad + \left[\mathcal{L}^{\nu_\xi}(z,t) - \mathcal{L}^{\nu_\xi}(z,t)\right] u(z,t) - \mathcal{L}(z,t)u(z,t) \\
&= \mathcal{L}^{\nu_\xi}(z,t)\left[u^{\xi}(z,t) - u(z,t)\right] \\
&\quad + \sum_{(w,\nu)\in\xi} w\left[\mathcal{L}^{\nu}(z,t) - \mathcal{L}^{\nu_\xi}(z,t)\right] u^{\nu}(z,t) \\
&\quad + \left[\mathcal{L}^{\nu_\xi}(z,t) - \mathcal{L}(z,t)\right] u(z,t) \\
&= \mathcal{L}^{\nu_\xi}(z,t)\hat{u}^{\xi}(z,t) + f^{\nu_\xi}(z,t,u^{\nu}) + \tilde{f}(z,t,u)
\end{aligned}$$

for all $(z,t) \in \mathbb{R}^N \times (0,T]$ and as before

$$\hat{u}^{\xi}(z,0) = \sum_{(w,\nu)\in\xi} wu^{\nu}(z,0) - u(z,0) = \left(\sum_{(w,\nu)\in\xi} w - 1\right) g(z).$$

□

2.5 Localization

For a constant coefficient PDE the transformation (2.6) in Chapter 2.2 leads to the fully diagonal N-dimensional heat operator (2.12) as a particularly simple form of differential operator. If faced with a non-constant coefficient PDE problem, we can localize the differential operator, for example by "freezing" its coefficients to a given point (z_0, t_0). This will lead to a localization error in addition to the expansion error.

Definition 13. *For a PDE problem of the form (2.7)–(2.9) with time- and space-dependent coefficients and given a point $(z_0, t_0) \in \mathbb{R}^N \times [0,T]$, we can derive the*

localized *constant-coefficient PDE by replacing* $\mathcal{L}(z,t)$ *with* $\mathcal{L}^{loc}_{z_0,t_0} = \mathcal{L}(z_0,t_0)$ *given by*

$$\mathcal{L}^{loc}_{z_0,t_0} = \sum_{k,l=1}^{N} \lambda_{kl}(z_0,t_0)\frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k=1}^{N} \kappa_k(z_0,t_0)\frac{\partial}{\partial z_k}.$$

$u_{loc}(z,t)$ *is then defined as the solution of*

$$\begin{aligned}\frac{\partial}{\partial t}u^{loc}(z,t) &= \mathcal{L}^{loc}_{z_0,t_0}u(z,t)\\ u^{loc}(z,0) &= g(z).\end{aligned}$$

Theorem 14. *The localization error* $\hat{u}(z,t) = u_{loc}(z,t) - u(z,t)$ *satisfies the non-homogeneous PDE*

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}(z,t) &= \mathcal{L}^{loc}_{z_0,t_0}\hat{u}(z,t) + f(z,t,z_0,u) \quad \forall (z,t) \in \mathbb{R}^N \times (0,T]\\ \hat{u}(z,0) &= 0 \quad \forall z \in \mathbb{R}^N\end{aligned}$$

with source term

$$f(z,t,z_0,u) = \left[\mathcal{L}^{loc}_{z_0,t_0} - \mathcal{L}(z,t)\right] u(z,t).$$

Proof. By definition

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}(z,t) &= \frac{\partial}{\partial t}(u^{loc}(z,t) - u(z,t))\\ &= \mathcal{L}^{loc}_{z_0,t_0}u^{loc}(z,t) - \mathcal{L}(z,t)u(z,t))\\ &= \mathcal{L}^{loc}_{z_0,t_0}\hat{u}(z,t) + \left[\mathcal{L}^{loc}_{z_0,t_0} - \mathcal{L}(z,t)\right] u(z,t)\end{aligned}$$

for all $(z,t) \in \mathbb{R}^N \times (0,T]$ and

$$\hat{u}(z,0) = u^{loc}(z,0) - u(z,0) = g(z) - g(z) = 0.$$

□

Given the zero boundary condition, the size of $\hat{u}(z,t)$ is directly related to the size of its source term. With Φ denoting the Green's function of $\mathcal{L}^{loc}_{z_0,t_0}$ it follows that

$$\hat{u}(z,t) = \int_0^t \int_{\mathbb{R}^N} \Phi(y,\tau) f(z-y, t-\tau, z_0, u) dy d\tau.$$

Writing out f explicitly we get

$$\begin{aligned} f(z,t,z_0,u) &= \left[\mathcal{L}^{loc}_{z_0,t_0} - \mathcal{L}(z,t)\right] u(z,t) \\ &= \sum_{k,l=1}^{N} \left[\lambda_{kl}(z_0,t_0) - \lambda_{kl}(z,t)\right] \frac{\partial^2}{\partial z_k \partial z_l} u(z,t) \\ &\quad + \sum_{k=1}^{N} \left[\kappa_k(z_0,t_0) - \kappa_k(z,t)\right] \frac{\partial}{\partial z_k} u(z,t). \end{aligned}$$

The size of the localization error depends on the spatial and temporal variation of the coefficient functions.

Remark 15. *This complete localization approach seems crude but works well for important classes of real world examples, such as the PDE associated with pricing financial products in LIBOR markets, see Chapter 6. It is also possible to only localize κ but allow λ to change over time.*

With this choice of localization method and the expansion approach from the previous section, localization and expansion are commuting operations, i.e.,

$$\mathcal{L}(z,t) \to \sum_{(w,\nu)\in\xi} \mathcal{L}^{\nu}(z,t) \to \sum_{(w,\nu)\in\xi} \mathcal{L}^{\nu,loc}_{z_0,t_0}$$

and

$$\mathcal{L}(z,t) \to \mathcal{L}^{loc}_{z_0,t_0} \to \sum_{(w,\nu)\in\xi} \mathcal{L}^{loc,\nu}_{z_0,t_0}$$

lead to the same differential operator.

2.6 Relation to Derivatives Pricing

Consider again the derivative pricing example introduced in Section 1.1. The corresponding Feynman-Kac PDE was given by (1.5)-(1.6)

$$\begin{aligned} \frac{\partial}{\partial t} v(x,t) &= -\sum_{i=1}^{N} \mu_i x_i \frac{\partial}{\partial x_i} v(x,t) - \sum_{i,j=1}^{N} \tfrac{1}{2} \sigma_i \sigma_j \rho_{ij} x_i x_j \frac{\partial^2}{\partial x_i \partial x_j} v(x,t) \\ v(x,T) &= h(x) \, \forall x \in \mathbb{R}^N \end{aligned}$$

on $\mathbb{R}^N \times [0,T]$. This captures the case of non path-dependent European options, which can be exercised at maturity only. For financial derivatives with path-dependency or more complex early-exercise features, additional constraints and model equations become necessary. In many cases, this will result in a sequence of problems of this type. For more details see Chapter 6.5.

Assume now that ρ and σ are constant and μ is a function of t alone. Let Σ be the covariance matrix, $\Sigma_{ij} = \sigma_i \rho_{ij} \sigma_j$ for all $1 \leq i, j \leq N$, and let $Q \in \mathbb{R}^{N \times N}$ be the orthogonal matrix of eigenvectors of Σ. W.l.o.g., let the eigenvalues λ_i be sorted in descending order, i.e., $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$. Then the coordinate transformation

$$\tau = T - t \ , \quad z = Q^T \ln(x) + \beta(\tau), \tag{2.22}$$

where

$$\beta_i(\tau) = -\sum_{j=1}^{N} Q_{ji} \left(\frac{\tau \sigma_j^2}{2} + \int_0^\tau \mu_j(s)\, \mathrm{d}s \right), \quad 1 \leq i \leq N, \tag{2.23}$$

leads to

$$\frac{\partial u}{\partial \tau} - \frac{1}{2} \sum_{i=1}^{N} \lambda_i \frac{\partial^2 u}{\partial z_i^2} = 0 \quad \forall (z, \tau) \in \mathbb{R}^N \times [0,T],$$

where $v(X,t) = u(z,\tau)$ and

$$u(z,0) = g(z) := h\left(\exp[Qz]\right).$$

Here the rotation with Q^T eliminates mixed derivatives and the translation by $\beta(\tau)$ eliminates the first order terms. This can be seen by a straightforward calculation of the partial derivatives in the new coordinates – see Appendix A.1 or [38].

2.7 Anchored ANOVA decompositions

Representing a bottleneck for computational approaches, high-dimensional PDEs and partially equivalent problems such as high-dimensional integrals have been – and

remain – the focus of intense research efforts. This lead to the development of various approximation techniques, some of which are related to our expansion methods. For example, the expansion methods developed in this thesis automatically make use – via PCA-based transformation and Taylor expansion - of the low-dimensional structure implicit in many practical problems, rather than detecting and incorporating this low-dimensionality explicitly as in factor method.

An approach closely related to our expansion methods are Anchored ANOVA decompositions. They are used in [18] to obtain dimension-adaptive approximations to option values expressed as integrals over high-dimensional spaces. In [42], a relation to the expansions from [41] is pointed out by utilising the integral representation of the solution to the heat equation, and [39] discusses these ideas jointly in the PDE context. We devote some room here to contrast these two closely related approaches, because it illustrates some of the complexity-reducing ideas (and limitation) underlying them.

We formulate the problem in slightly more general terms here as befits the applications later on. Consider the situation where $Z = (Z_1(t), \ldots, Z_N(t))_{t\geq 0}$ is an N-dimensional Markov process and where the time t value u of a contingent claim is fully determined by $Z(t)$ and the time to maturity $T-t$. We therefore write this value as $u(Z(t), T-t)$. Define, for a set $v = \{i_1, \ldots, i_m\} \subseteq \{1, \ldots, N\}$, auxiliary processes $Z^v = (Z_1^v, \ldots, Z_N^v)$ which are "frozen" in the coordinates with indices not in v, that is, $Z_i^v(t) = Z_i(0)$ for all $t \geq 0$, $i \notin v$, and we impose that the joint law of $\{Z_i^v, i \in v\}$, is identical to the joint law of $\{Z_i, i \in v\}$, *given* that $Z_j(t) = Z_j(0)$ for all $j \notin v$. To be more specific, in the common case where Z is defined through a stochastic differential equation (SDE) of the type

$$\mathrm{d}Z_i(t) = \mu_i(Z(t), t)\,\mathrm{d}t + \sigma_i(Z(t), t)\,\mathrm{d}W_i(t), \tag{2.24}$$

we define Z^v by

$$\mathrm{d}Z_i^v(t) = \begin{cases} \mu_i(Z^v(t),t)\,\mathrm{d}t + \sigma_i(Z^v(t),t)\,\mathrm{d}W_i(t) & i \in v, \\ 0 & \text{else}, \end{cases} \tag{2.25}$$

which are constant in directions $i \in \{1,\ldots,N\}\backslash v$, i.e., $Z_i^v(t) = Z_i^v(0)$ for those i.

A particular example studied in [18, 42], which is related to the valuation of European-style derivatives, is

$$u(z,\tau) = u(z,T-t) = \mathbb{E}[g(Z(T))|Z(t) = z], \tag{2.26}$$

where $\tau = T - t$ is the time to maturity, g is the payoff function and the expectation is taken with respect to an explicitly known probability measure. In that case, we define approximations based on the process Z^v in (2.25) as

$$\widehat{u}_v(a; z^v, \tau) = \mathbb{E}[g(Z_T^v)|Z_i^v(t) = z_i \;\forall\, i \in v; Z_i^v(t) = a_i \;\forall\, i \notin v], \tag{2.27}$$

where $a \in \mathbb{R}^N$ with $a_i = Z_i(0)$, $z^v = (z_{i_1},\ldots,z_{i_{|v|}})$, i.e., we anchor the solution at the initial value of this stochastic process.

The forward and backward PDEs for processes of the type (2.24) are second order linear parabolic. To get from the PDE for (2.26) to the one for (2.27), the coefficients of all derivatives in directions x_i are set to zero for $i \notin v$, as per (2.25).

The coordinates z in (2.22) were chosen specifically as the principal components of the covariance matrix of a diffusion process X, for constant μ and σ, in which case $\mu_{Z,i} = 0$ and $\sigma_{Z,i}^2 = \lambda_i$. A similar construction is used and analysed in [23]. The PDE satisfied by $\widehat{u}_v$ in (2.27) is the heat equation with diffusion coeffiicents

$$\lambda^v = \sum_{i \in v} \lambda_i e_i. \tag{2.28}$$

As the solution $\widehat{u}_v$, $v = \{i_1,\ldots,i_{|v|}\}$, only depends on the sub-vector of coordinates $z^v = (z_{i_1},\ldots,z_{i_{|v|}})$ non-trivially,

$$\begin{aligned} u_v(a; z^v, \tau) &= \widehat{u}_v - \sum_{w \subset v} u_w \\ &= \sum_{w \subseteq v} (-1)^{|w|-|v|} \widehat{u}_w \end{aligned}$$

is well-defined and gives a suitable anchored ANOVA decomposition of u (see [18] and the references therein) as given in (1.7), where τ is an additional argument of all terms.

For optimal stopping problems, such as the Bermudan swaptions studied later, the analogue to (2.26) and (2.27) is

$$\begin{aligned} u(z,\tau) &= \sup_{\mathcal{T}\in\mathbb{T}} \mathbb{E}[g(Z(\mathcal{T}))|Z(t)=z], \\ \widehat{u}_v(a;z^v,\tau) &= \sup_{\mathcal{T}\in\mathbb{T}^v} \mathbb{E}[g(Z^v(\mathcal{T}))|Z_i^v(t)=z_i\ \forall\, i\in v; Z_i^v(t)=a_i\ \forall\, i\notin v], \end{aligned}$$

where $\mathbb{T}, \mathbb{T}^v$ are suitable sets of stopping times.

For path-dependent options, the process Z has to be set up to include the path-dependent quantity in order to bring it back into the assumed Markovian framework. In Chapter 6.4, we demonstrate this on the example of a Ratchet floor. As the path-dependent quantity is reset at discrete time points, the corresponding component of Z is a jump-process instead of following an SDE of the form (2.24).

The general principle is that we define $\widehat{u}_v$ as the "solution of the problem with Z replaced by Z^v". To re-iterate, the key is that Z^v changes only in $|v|$ dimensions and is constant in the remaining $N-|v|$ dimensions. This means that u_v can be found by solving $|v|$-dimensional (e.g., PDE) problems instead of the N-dimensional one.

We will later use Taylor expansion to construct specific expansion methods. The link between (1.7) and this approach in equation (3.7) can be established if we pick $\lambda^0 = 0$, set $v = \{i_1, \ldots, i_k\}$, λ^v as in (2.28), and, inductively (skipping z and τ as argument

of Δ for brevity),

$$
\begin{aligned}
&\Delta^{(i_1)}(u;\lambda,\lambda^0) \\
&= \delta\lambda_i \frac{u(z,\tau;\lambda^0+\delta\lambda_{i_1}e_{i_1}) - u(z,\tau;\lambda^0)}{\delta\lambda_{i_1}}\,, \\
&\Delta^{(i_1,\ldots,i_k,i_{k+1})}(u;\lambda,\lambda^0) \\
&= \Delta^{(i_{k+1})}(\Delta^{(i_1,\ldots,i_k)};\lambda,\lambda^v) \\
&= \delta\lambda_{i_{k+1}} \frac{\Delta^{(i_1,\ldots,i_k)}(u;\lambda,\lambda^v+\delta\lambda_{i_{k+1}}e_{i_{k+1}}) - \Delta^{(i_1,\ldots,i_k)}(u;\lambda,\lambda^v)}{\delta\lambda_{i_{k+1}}} \\
&= \sum_{w\subseteq v\cup\{k+1\}} (-1)^{|w|-|v|-1} u(z,\tau;\lambda^w),
\end{aligned}
\tag{2.29}
$$

for $i_l \neq i_m$, $1 \leq l < m \leq k+1$, and 0 otherwise. Only terms of mixed first order are present and absorb all higher order terms – see next paragraph. Then, the precise relation between ANOVA terms and the finite difference approximation to the Taylor expansion is

$$
\begin{aligned}
\widehat{u}_v(z;z^v,0) &= u(z,0;\lambda^v), \\
u_v(z;z^v,0) &= \Delta^{(i_1,\ldots,i_k)}(u;z,0;\lambda,\lambda^0).
\end{aligned}
$$

The relation between ANOVA and multi-variate Taylor expansions in the coordinates is discussed in [17]. The twist here is to apply the expansion in $\delta\lambda_i$ instead of z_i.
A further point to note is that if the above expansion is truncated to include terms up to $|v| = n < N$, it is only of first order accurate in $\delta\lambda_i$. For relatively large $\delta\lambda_i$ and smooth solutions, the inclusion of higher order Taylor terms in individual and mixed directions and higher order finite difference formulae may be preferable as we will see in Chapter 6.4. The extra cost is small as typically the dimensionality of PDEs involved will not increase. What distinguishes the above expansion from other finite difference approximations is that it is an exact decomposition, i.e., if we include all terms up to degree N, we recover the exact solution irrespective of its smoothness.

In a variation to (1.7), we can consider a decomposition, where in addition to the anchor a, all contributions may also depend on the first coordinate,

$$\begin{aligned} u(z) = & u_0^{(1)}(a; z_1) + \sum_{i=2}^{N} u_i^{(1)}(a; z_1; z_i) + \sum_{\substack{i,\, j\, =\, 2 \\ i\, <\, j}}^{N} u_{i,j}^{(1)}(a; z_1; z_i, z_j) \\ & + \ldots + u_{2,\ldots,N}^{(1)}(a; z_1; z_2, \ldots, z_N), \end{aligned}$$

and, generalising this from one to $r \geq 1$ coordinates,

$$\begin{aligned} u(z) = & u_0^{(r)}(a; z_1, \ldots, z_r) \\ & + \sum_{p=1}^{N-r} \sum_{\substack{\{i_1, \ldots, i_p\} \\ \subseteq \{r+1, \ldots, N\}}} u_{i_1,\ldots,i_p}^{(r)}(a; z_1, \ldots, z_r; z_{i_1}, \ldots, z_{i_p}). \end{aligned}$$

Clearly, in relation to Chapter 3.1, this corresponds to using $\lambda^0 = (\lambda_1, 0, \ldots, 0)$ and $\lambda^0 = (\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$, resp., and an adaptation of the finite difference formulae.

The goal is to find a decomposition where the contributions $u_{i_1,\ldots,i_p}^{(r)}$ decay fast with increasing r and increasing p, in order for the approximations

$$\begin{aligned} u^{(r,s)}(z) = & u_0^{(r)}(a; z_1, \ldots, z_r) \\ & + \sum_{p=1}^{s} \sum_{\substack{\{i_1, \ldots, i_p\} \\ \subseteq \{r+1, \ldots, N\}}} u_{i_1,\ldots,i_p}^{(r)}(a; z_1, \ldots, z_r; z_{i_1}, \ldots, z_{i_p}) \end{aligned} \tag{2.30}$$

for $s \leq N - r$ to be accurate for small $r + s$. The first order, first eigenvalue method used in Chapter 6 corresponds to $r = s = 1$. In [11], a natural link between ANOVA decompositions and dimension adaptive sparse grids is exploited to construct *a priori* as well as *a postiori* optimal approximations to high-dimensional functions.

The effect of higher-dimensional terms in the cases $r = 2$, $s = 1$ and $r = 1$, $s = 2$ is illustrated in [42] for equity basket options, extending the case $r = 1$, $s = 1$ in [41]. The data there have in common with our set-up in Chapter 6 the presence of a dominant eigenvalue, such that the case $(r, s) = (1, 2)$ gives a notable improvement

over $(1,1)$ for arithmetic average basket options by capturing the second order terms in the small eigenvalues, while $(r,s) = (2,1)$ does not give a big accuracy gain.
We will give numerical results up to $r = 5$ and $s = 3$ in Chapter 6.4, in the context of the LIBOR Market Model.

Chapter 3

Constant Coefficient PDEs

3.1 Method Construction via Taylor Series

It is possible to derive coherent expansions of high accuracy by Taylor expanding the solution $u(z,t) = u(z,t,\lambda)$ in the eigenvalues λ [40].

Lemma 16. *Let $u(z,t,\cdot) \in C^m$ and fix $\lambda_0 \in \mathbb{R}^N$. Using multi-index notation we write $D^\alpha_\lambda u = \frac{\partial^{|\alpha|}u}{\partial\lambda^\alpha} = \frac{\partial^{|\alpha|}u}{\partial\lambda_1^{\alpha_1}\ldots\partial\lambda_N^{\alpha_N}}$ and $(\lambda-\lambda_0)^\alpha = (\lambda_1-\lambda_{0,1})^{\alpha_1}\cdot\ldots\cdot(\lambda_N-\lambda_{0,N})^{\alpha_N}$. Then*

$$u(z,t,\lambda) = \sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{(\lambda-\lambda_0)^\alpha}{\alpha!}D^\alpha_\lambda u(z,t,\lambda_0) + R_m(z,t,\lambda,\lambda_0,u) \tag{3.1}$$

In general, the remainder term $R_m(z,t,\lambda,\lambda_0,u)$ can be written as

$$\begin{aligned} R_m(z,t,\lambda,\lambda_0,u) &= \sum_{|\alpha|=m}(\lambda-\lambda_0)^\alpha R^\alpha(z,t,\lambda,\lambda_0,u) \\ \lim_{\lambda\to\lambda_0} R^\alpha(z,t,\lambda,\lambda_0,u) &= 0. \end{aligned}$$

If $u(z,t,\cdot) \in C^{m+1}$, then an explicit form can be given as

$$\begin{aligned} R_m(z,t,\lambda,\lambda_0,u) &= \sum_{|\alpha|=m+1}(\lambda-\lambda_0)^\alpha R^\alpha(z,t,\lambda,\lambda_0,u) \\ R^\alpha(z,t,\lambda,\lambda_0,u) &= \frac{|\alpha|}{\alpha!}\int_0^1 (1-s)^{|\alpha|-1}(D^\alpha_\lambda u)(z,t,\lambda_0+s(\lambda-\lambda_0))ds. \end{aligned}$$

Proof. This is a standard result of multivariate calculus. See for example [30]. □

We will analyse the existence of the λ-derivatives in detail in Chapter 3.2. For the time being, we note that they then satisfy PDEs themselves, which can be derived by differentiating the heat equation with respect to λ. For instance,

$$v_j = \left.\frac{\partial u}{\partial \lambda_j}\right|_{\lambda=\lambda_0}$$

is the solution of

$$\frac{\partial v_j}{\partial t} = \sum_{k=1}^{N} \lambda_{0,k} \frac{\partial^2 v_j}{\partial z_k^2} + w_j, \tag{3.2}$$

$$v_j(z,0) = 0, \tag{3.3}$$

where

$$w_j = \left.\frac{\partial^2 u}{\partial z_j^2}\right|_{\lambda=\lambda_0}$$

is itself the solution to

$$\frac{\partial w_j}{\partial t} = \sum_{k=1}^{N} \lambda_{0,k} \frac{\partial^2 w_j}{\partial z_k^2}, \tag{3.4}$$

$$w_j(z,0) = \frac{\partial^2 g}{\partial z_j^2}. \tag{3.5}$$

Now if λ_0 is chosen such that $\lambda_{0,k} = 0$ for all k except for a very few, (3.2) and (3.4) are low-dimensional PDEs which can be solved with standard methods. For instance, if $\lambda_0 = (\lambda_1, 0, \ldots, 0)$, (3.2) is a one-dimensional PDE in z_1, with a forcing term given by the solution to the other one-dimensional PDE (3.4).

This argument can in principle be repeated for higher derivatives in λ, including mixed derivatives, of arbitrary order. We use a different approach for the remainder of this thesis based on a difference approximation with finite step-size of the λ-derivatives.

Difference Stencils in λ

To make practical use of the Taylor expansion, we need to evaluate the partial derivatives in equation (3.1). One method is to choose suitable finite difference approximations Δ_λ^α. For example, Hilber et al. [23] propose to use high order compact finite difference stencils introduced in [33].

Denoting the step size in each direction i by $\delta\lambda_i$ and the approximation orders of Δ_λ^α by the multi-index set d_α, we have

$$D_\lambda^\alpha u(z,t,\lambda_0) = \Delta_\lambda^\alpha(u,z,t,\delta\lambda,\lambda_0) + \sum_{\beta\in d_\alpha} O(|\delta\lambda^\beta|) \tag{3.6}$$

Clearly, $D_\lambda^0 u(z,t,\lambda_0) = \Delta^0(u,z,t,\delta\lambda,\lambda_0) = u(z,t,\lambda_0)$. Combining equations (3.6) and (3.1) gives

$$\begin{aligned} u(z,t,\lambda) = &\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{(\lambda-\lambda_0)^\alpha}{\alpha!}\Delta^\alpha(u,z,t,\delta\lambda,\lambda^0) \\ &+\sum_{k=1}^{m}\sum_{|\alpha|=k}\sum_{\beta\in d_\alpha} O\left(|(\lambda-\lambda_0)^\alpha\delta\lambda^\beta|\right) + R_m(z,t,\lambda,\lambda_0,u) \end{aligned} \tag{3.7}$$

If we choose $\delta\lambda = \lambda - \lambda_0$ and assume $u \in C^{m+1}$ in λ, then the leading order errors are of size

$$\sum_{|\alpha|=m+1} O\left(|(\lambda-\lambda_0)^\alpha|\right) + \sum_{k=1}^{m}\sum_{|\alpha|=k}\sum_{\beta\in d_\alpha} O\left(|(\lambda-\lambda_0)^{\alpha+\beta}|\right)$$

The order of the second sum is determined by the expansion order (assuming sufficient smoothness) and will be the limiting factor to the overall order in practice. Choosing higher order stencils, the order of the terms in the second sum can be increased, provided enough smoothness in λ, i.e., derivatives up to order $\alpha+\beta$ exist. This comes at the expense of additional computational complexity, as the PDE solution has to be computed for different values of λ to evaluate the finite difference formula. As the dimensionality of the PDEs is the same, this is a constant multiplicative factor and therefore not crucial.

More specifically, computing $\Delta_\lambda^\alpha u(z,t,\delta\lambda,\lambda_0)$ requires the values $u(z,t,\lambda')$ for different points λ', depending on the finite difference scheme. This can be translated into an expansion encoding using the ν/ξ-notation introduced in Chapter 2.4.

Example 17. *For $\alpha = (0,\dots,0,1,0,\dots,0)$, where the k-th entry is 1, a firstorder finite difference approximation is given by*

$$\Delta_\lambda^\alpha u(z,t,\delta\lambda,\lambda_0) = \frac{u(z,t,\lambda_0+\delta\lambda_k e_k) - u(z,t,\lambda_0)}{\delta\lambda_k},$$

where e_k is the k-th canonical unit vector. Now choose $1 \leq r \leq N$ and set $\lambda_0 = (\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$, $\delta\lambda = \lambda - \lambda_0$, and $m = 1$. Furthermore, assume $u(z,t,\lambda)$ is twice continuously differentiable in λ. Then

$$\begin{aligned}
u(z,t,\lambda) &= D_\lambda^0 u(z,t,\lambda_0) + \sum_{|\alpha|=1} \frac{(\lambda-\lambda_0)^\alpha}{\alpha!} D_\lambda^\alpha u(z,t,\lambda_0) + R_1(z,t,\lambda,\lambda_0,u) \\
&= u(z,t,\lambda_0) \\
&\quad + \sum_{k=r+1}^{N} \lambda_k \frac{u(z,t,\lambda_0+\lambda_k e_k) - u(z,t,\lambda_0) - \lambda_k^2 R^{2\alpha_k}(z,t,\lambda,\lambda_0,u)}{\lambda_k} \\
&\quad + \sum_{k=r+1}^{N} \lambda_k^2 R^{2\alpha_k}(z,t,\lambda,\lambda_0,u) + \sum_{k,l=r+1,k\neq l}^{N} \lambda_k\lambda_l R^{\alpha_k+\alpha_l}(z,t,\lambda,\lambda_0,u) \\
&= (1+r-N)u(z,t,\lambda_0) + \sum_{k=r+1}^{N} u(z,t,\lambda_0+\lambda_k e_k) \\
&\quad + \sum_{k,l=r+1,k\neq l}^{N} \lambda_k\lambda_l R^{\alpha_k+\alpha_l}(z,t,\lambda,\lambda_0,u) \\
&= u^\xi(z,t,\lambda) + O(\lambda_{r+1}^2 + \ldots + \lambda_N^2)
\end{aligned}$$

for

$$\begin{aligned}
\xi = \{&(1+r-N,\{1,\ldots,r\}),(1,\{1,\ldots,r,r+1\}),(1,\{1,\ldots,r,r+2\}), \\
&\ldots,(1,\{1,\ldots,r,N-1\}),(1,\{1,\ldots,r,N\})\}.
\end{aligned} \tag{3.8}$$

Error Bounds

Definition 18. *Let*

$$\begin{aligned}
C^b &= \{g : \mathbb{R}^n \to \mathbb{R} : \ |g(z)| \leq c \ \exp(k|z|^\alpha) \text{ for some } c,k>0,\ 0 \leq \alpha < 2\}, \\
C^{j,k,mix} &= \left\{g \in C^b : \ \partial_{i_1}^j \ldots \partial_{i_k}^j g \in C^b,\ \forall 1 \leq i_1 < \ldots < i_k \leq n\right\}.
\end{aligned}$$

We now state the main result for the expansion error in Example 7 and 17.

Theorem 19. *For the heat equation (2.13), (2.14), assume $g \in C^{2,2,mixed}$. Then the expansion error $\hat{u}^\xi$ from (2.20), for u^ξ from (2.18) with ξ from (2.19), satisfies*

$$\left|\hat{u}^\xi(z,t)\right| \leq t^2 \sum_{r<i<k\leq N} \lambda_k\lambda_i \left\| \frac{\partial^4 g}{\partial z_k^2 \partial z_i^2} \right\|_\infty . \tag{3.9}$$

Proof. See end of Chapter 3.3. □

Remark 20. *There are no diagonal quadratic terms with factors λ_k^2 in (3.9). They cancel out with terms in the Taylor expansion of the finite difference scheme due to the specific choice of steps size $\delta\lambda_k = \lambda_k$. This has the important consequence that any solutions which only depend on one of the z_k are integrated exactly, and by superposition any linear combination of such terms. By similar reasoning, $\widehat{u}^\xi$ is zero for initial conditions of the form*

$$g(z) = \sum_{k=r+1}^{N} g_k(z_k; z_1, \dots, z_r) \quad + \sum_{i,k=r+1}^{N} z_i \, g_{i,k}(z_k; z_1, \dots, z_r),$$

for any functions g_k and $g_{i,k}$. If an initial condition can be approximated well by functions of this form, the expansion error $\widehat{u}^\xi$ will be small irrespective of smoothness or smallness of the eigenvalues and time.

The smoothness requirement $g \in C^{2,2,mixed}$ can be considerable weakened using the following observation: We can replace the initial condition g in the estimate (3.9) with a function G which has been smoothed by the heat kernel in the first r directions, namely

$$G(z,t) = \int_{\mathbb{R}^r} \prod_{k\in\nu} \frac{\exp(-y_k^2/4\lambda_k t)}{\sqrt{4\pi\lambda_k t}} g(z_1 - y_1, \dots, z_r - y_r, z_{r+1}, \dots, z_n) \, dy_1 \dots dy_r \tag{3.10}$$

for $\nu = \{1, \dots, r\}$. See also (2.15). The (non-)smootheness of G is all what is relevant for the expansion error. Thus, even if g is only, e.g., piecewise smooth, it is only in degenerate cases that G is not smooth everywhere. We analyse this in Chapter 3.2 and also give illustrative examples.

Corollary 21. *For the heat equation (2.13), (2.14), assume $G(\cdot, t) \in C^{2,2,mixed}$ for*

$$G(z,t) = u^{\xi'}, \quad \xi' = \{(1, \{1, \dots, r\})\}.$$

Then the expansion error $\hat{u}^\xi$ from (2.20), for u^ξ from (2.18) with ξ from (2.19), satisfies

$$\left|\hat{u}^\xi(z,t)\right| \leq t^2 \sum_{r<i<k\leq N} \lambda_k\lambda_i \left\| \frac{\partial^4 G}{\partial z_k^2 \partial z_i^2}(\cdot,t) \right\|_\infty .$$

Proof. Application of Theorem 19 with $r = 0$ to the heat equation in $\mathbb{R}^{n-r}$ with coefficients $\lambda_{r+1}, \ldots, \lambda_n$, and initial condition $G(z,t)$ with $z_1, \ldots, z_r$ fixed. □

Anchored ANOVA Decompositions

The expansion in Example 17 is a special case of the following type of approximations. Let $\alpha_0 = (1, \ldots, 1, 0, \ldots, 0)$ such that the first r entries are 1 and the remaining ones 0, and α with $\alpha_i \in \{0, 1\}$ such that $\alpha \geq \alpha_0$ element-wise, then define

$$\lambda^\alpha \Delta_\lambda^\alpha u(z,t,\lambda,\lambda_0) = \sum_{\alpha_0\leq\beta\leq\alpha} (-1)^{|\alpha-\beta|} u(z,t,\lambda_0 + \delta\lambda\cdot\beta),$$

where '$\cdot$' is element-wise multiplication. This is not only a consistent approximation to the mixed derivative of order α, but the approximation

$$\hat{u}^\xi(z,t) = \sum_{r<j\leq k} \sum_{|\alpha|=j} \lambda^\alpha \Delta_\lambda^\alpha u(z,t,\lambda,\lambda_0)$$

is seen to be exact for u which depend only on $\lambda_1, \ldots, \lambda_r$ and $k-r-1$ additional λ_j, as well as any linear combination thereof. This is in particular the case if g only depends on $x_1, \ldots, x_r$ and $k-r-1$ additional x_j. It is reflected in the error bound (3.9)[1]. Overall, this represents an anchored ANOVA decomposition of the function u, for a particular choice of anchor. See, for instance, [19] and Chapter 2.7.

While published error bounds usually focus on L^2-type errors expressed in terms of the decay and variance of ANOVA terms (see, e.g., [45] or [47]), we use PDE theory to derive bounds on the approximation error at the anchor point itself and in terms of the smoothness of not captured components.

[1] See also Remark 20.

3.2 Existence of Partial Derivatives $D^{\alpha}_{\lambda}u$

In this section, we establish conditions that guarantee the existence of partial derivatives $D^{\alpha}_{\lambda}u$. We also derive bounds on the size of those derivatives, so that an explicit upper bound for the remainder term can be given. This allows us to turn the Taylor series expansion from a heuristically motivated approach into a rigorously justified method.

Does $D^{\alpha}_{\lambda}u$ exist for non-smooth initial conditions? The existence and size of a partial derivative in direction of an eigenvalue can be related to properties of the boundary condition at $t = 0$ via convolution with the heat kernel. As it turns out, this has a very strong smoothing effect on the solution and allows us to answer this highly non-trivial question positively for a wide class of practically relevant problems. This includes boundary functions which are not differentiable everywhere — for example ones with kinks or jumps — as long as the non-differentiabilities are local and isolated. Examples 25 and 26 illustrate this effect.

It is a well-known phenomenon that the heat equation and the heat kernel, resp., have a smoothing effect on solutions for non-smooth initial functions. In fact, observations very similar to the subsequent ones have been made in [20] from an ANOVA decomposition perspective, where the authors are likewise motivated by option pricing problems.

The details in this section are rather technical and provide a theoretical justification for the the method's observed accuracy. They can be skipped at first reading without diminishing the understanding of the rest of the thesis.

The Smooth Case

Let us now consider the question of existence of partial derivatives $D^\alpha_\lambda u$. For simplicity of notation define

$$\sqrt{\lambda} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_N} \end{pmatrix}.$$

Recall that u is the solution of a homogeneous heat equation with diffusion coefficients λ and inital condition g. We can thus try to write the partial derivative $D^\alpha_\lambda u$ explicitly as

$$\begin{aligned} D^\alpha_\lambda u(z,t,\lambda) &= D^\alpha_\lambda \int_{\mathbb{R}^N} \Phi(y,t,\lambda) g(z-y)dy \\ &= D^\alpha_\lambda \int_{\mathbb{R}^N} \Phi(x,t,1) g(z-\sqrt{\lambda}\cdot x)dx \\ &= \int_{\mathbb{R}^N} \Phi(x,t,1) D^\alpha_\lambda g(z-\sqrt{\lambda}\cdot x)dx \end{aligned}$$

In the last line we have pulled the differential operator under the integral sign. This is only correct for certain integrands.

Example 22. *For a simple smooth function differentiation under the integral is justified. For example,*

$$\begin{aligned} \frac{\sqrt{\pi}}{2} &= \int_{\mathbb{R}} x^2 \exp(-x^2)dx = \int_{\mathbb{R}} \frac{\partial}{\partial\lambda} \lambda x^2 \exp(-x^2)dx \\ &= \frac{\partial}{\partial\lambda} \int_{\mathbb{R}} \lambda x^2 \exp(-x^2)dx = \frac{\partial}{\partial\lambda} \frac{\lambda\sqrt{\pi}}{2} = \frac{\sqrt{\pi}}{2} \end{aligned}$$

The same is not necessarily true for functions that diverge, are non-smooth or discontinuous. Gelbaum and Olmsted [15] give an illustrative example

$$\begin{aligned} f(x,\lambda) &= \begin{cases} \frac{\lambda^3}{x^2}\exp(-\lambda^2/x) & x>0 \\ 0 & x=0 \end{cases} \\ \Rightarrow &\frac{\partial}{\partial\lambda}\int_0^1 f(x,\lambda)dx = \frac{\partial}{\partial\lambda}\lambda\exp(-\lambda^2) = (1-2\lambda^2)\exp(-\lambda^2) \\ &\int_0^1 \frac{\partial}{\partial\lambda} f(x,\lambda)dx = \begin{cases} (1-2\lambda^2)\exp(-\lambda^2) & \lambda\neq 0 \\ 0 & \lambda = 0 \end{cases} \\ \Rightarrow &\frac{\partial}{\partial\lambda}\int_0^1 f(x,\lambda)dx\Big|_{\lambda=0} = 1 \neq 0 = \int_0^1 \frac{\partial}{\partial\lambda} f(x,\lambda)dx\Big|_{\lambda=0} \end{aligned}$$

However, the exchange of differentiation and integration is valid for certain classes of such functions. Consider the following calculation:

$$\begin{aligned} 1 &= \int_{\mathbb{R}} |x| \frac{x^2}{\lambda^2} \exp(-x^2/\lambda) dx = \int_{\mathbb{R}} \frac{\partial}{\partial \lambda} |x| \exp(-x^2/\lambda) dx \\ &= \frac{\partial}{\partial \lambda} \int_{\mathbb{R}} |x| \exp(-x^2/\lambda) dx = \frac{\partial}{\partial \lambda} \lambda = 1 \end{aligned}$$

Clearly, if g is sufficiently smooth, e.g. if $D_\lambda^\alpha g(z - \sqrt{\lambda} \cdot x)$ is continuous, and if it grows slowly enough such that the right-hand side exists, this exchange of differentiation and integration is justified. In particular,

$$\begin{aligned} \frac{\partial}{\partial \lambda_j} u(z,t,\lambda) &= \int_{\mathbb{R}^N} \Phi(x,t,1) \frac{\partial}{\partial \lambda_j} g(z - \sqrt{\lambda} \cdot x) dx \\ &= \int_{\mathbb{R}^N} \Phi(x,t,1) \frac{-x_j}{2\sqrt{\lambda_j}} \frac{\partial}{\partial z_j} g(z - \sqrt{\lambda} \cdot x) dx \\ &= \int_{\mathbb{R}^N} \Phi(y,t,\lambda) \frac{-y_j}{2\lambda_j} \frac{\partial}{\partial z_j} g(z-y) dy \\ &= t \int_{\mathbb{R}^N} \Phi(y,t,\lambda) \frac{\partial^2}{\partial z_j^2} g(z-y) dy \end{aligned}$$

using partial integration.

We can alternatively derive the partial derivatives as solutions of PDEs. Consider for ease of notation the two-dimensional case first.

$$\frac{\partial u}{\partial t} = \mathcal{L} u \;=\; \lambda_1 \frac{\partial^2 u}{\partial z_1^2} + \lambda_2 \frac{\partial^2 u}{\partial z_2^2} \tag{3.11}$$

$$u(z,0) = g(z). \tag{3.12}$$

Lemma 23. *1. If $g \in C^{2,1,mixed}$, then for all i*

$$\frac{\partial u}{\partial \lambda_i} = t \frac{\partial^2 u}{\partial z_i^2}$$

exists and is the solution to

$$\frac{\partial v}{\partial t} - \mathcal{L} v = \frac{\partial^2 u}{\partial z_i^2}.$$

2. *If* $g \in C^{2,2,mixed}$, *then if* $i \neq j$

$$\frac{\partial^2 u}{\partial \lambda_i \partial \lambda_j} = t^2 \frac{\partial^4 u}{\partial z_i^2 \partial z_j^2}$$

exists and is the solution to

$$\frac{\partial w}{\partial t} - \mathcal{L} w = 2t \frac{\partial^4 u}{\partial z_i^2 \partial z_j^2}.$$

For $i = j$, *one needs* $g \in C^{4,1,mixed}$, *and the same equations hold.*

Proof. The PDEs are obtained by differentiation of (3.11), and the solutions are verified by substitution. □

Consider now the N-dimensional case.

Corollary 24. *Let* u *be a solution to (2.13). If* $g \in C^{2,1,mixed}$, *then*

$$\frac{\partial}{\partial \lambda_j} u(z_0, t, \lambda) = t^2 \int_{\mathbb{R}^N} \Phi(y, t, \lambda) \frac{\partial^2 g}{\partial z_j^2}(z_0 - y)\, dy.$$

If, moreover, $g \in C^{4,1,mixed}$, *then*

$$\frac{\partial^2}{\partial \lambda_j^2} u(z_0, t, \lambda) = t^2 \int_{\mathbb{R}^N} \Phi(y, t, \lambda) \frac{\partial^4 g}{\partial z_j^4}(z_0 - y)\, dy,$$

and if $g \in C^{2,2,mixed}$, *then*

$$\frac{\partial^2}{\partial \lambda_j \partial \lambda_k} u(z_0, t, \lambda) = t^2 \int_{\mathbb{R}^N} \Phi(y, t, \lambda) \frac{\partial^4 g}{\partial z_j^2 \partial z_k^2}(z_0 - y)\, dy.$$

Analytical Examples with Non-Smooth Data

The primary motivation for the research in this thesis are applications in derivative pricing, where the boundary data are typically non-smooth. We will see from the examples in this section, that this does not necessarily mean that the results derived in the previous section are not applicable. In fact, we will see that with the exception of certain degenerate cases the same convergence orders as in the smooth case hold.

For illustration, we study the two-dimensional Black-Scholes model, where the value function of a European option on a pair of assets $S = (S_1, S_2)$, with payout $h(S)$ at time T satisfies

$$\begin{aligned}\frac{\partial V}{\partial t} + \left(\sigma_1^2 S_1^2 \frac{\partial^2 V}{\partial S_1^2} + \rho\sigma_1 S_1 \sigma_2 S_2 \frac{\partial^2 u}{\partial S_1 \partial S_2} + \sigma_2^2 S_2^2 \frac{\partial^2 V}{\partial S_2^2}\right) \\ + r\left(S_1 \frac{\partial V}{\partial S_1} + S_2 \frac{\partial V}{\partial S_2}\right) = 0, \\ V(S,T) = h(S).\end{aligned}$$

Note that, to stay within the usual conventions of mathematical finance, we used the natural time with payout at $t = T$ rather than at $t = 0$.

Instead of applying the transformation (2.6) directly, we use the log-normality of the Black-Scholes model to first use logarithmic coordinates $(x_1, x_2) = (\log S_1, \log S_2)$. Elementary calculation shows that for the specific choice of $\sigma_1^2 = \sigma_2^2 = 2r = \sigma^2$ for some $\sigma > 0$, the rotation

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

leads to a particularly simple form, the heat equation

$$\begin{aligned}\frac{\partial u}{\partial t} &= \lambda_1 \frac{\partial^2 u}{\partial z_1^2} + \lambda_2 \frac{\partial^2 u}{\partial z_2^2}, \\ u(z,0) &= g(z).\end{aligned}$$

For a general choice of parameters a further transformation to eliminate the drift term is needed, but we avoid this complication here as it does not change the essence of the problem. Here,

$$\begin{aligned}\lambda_1 &= \sigma^2(1+\rho), \\ \lambda_2 &= \sigma^2(1-\rho),\end{aligned}$$

so that for ρ close to 1 or -1, i.e., near perfect correlation or anti-correlation, the problem becomes close to one-dimensional and the asymptotic expansion can be expected to work particularly well.

Example 25. *Consider the payout function*

$$h(s_1, s_2) = \mathbb{1}_{[0,\infty)}(s_1^{\mu_1} s_2^{\mu_2} - 2)$$

with real parameters μ_1 and μ_2, which leads to

$$g(z_1, z_2) = \mathbb{1}_{[0,\infty)} \left(\exp \left[\frac{\mu_1 + \mu_2}{2} z_1 + \frac{\mu_1 - \mu_2}{2} z_2 \right] - 2 \right)$$

We will consider three different choices of μ_1 and μ_2.

Case 1: $\mu_1 = 1$, $\mu_2 = 1$

In this case the value of the boundary function depends only on z_1 and correspondingly

$$\begin{aligned} \frac{\partial}{\partial \lambda_2} u(z, t, \lambda) &= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}^2} \Phi(x, t, 1) g(z - \sqrt{\lambda} \cdot x) dx \\ &= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}} \Phi(x_1, t, 1) \mathbb{1}_{[0,\infty)} \left(\exp \left[z_1 - \sqrt{\lambda_1} x_1 \right] - 2 \right) \int_{\mathbb{R}} \Phi(x_2, t, 1) dx_2 dx_1 \\ &= 0. \end{aligned}$$

Case 2: $\mu_1 = 1$, $\mu_2 = -1$

The value of the boundary function depends only on z_2. We find that the derivative in λ_2 exists and, by a lengthy calculation[2]*, is given by*

$$\frac{\partial}{\partial \lambda_2} u(z, t, \lambda) = - \frac{z_2 - \log 2}{2\sqrt{2t}\sqrt{\lambda_2}^3} \cdot \frac{\exp(-(z_2 - \log 2)^2 / 4t\lambda_2)}{\sqrt{2\pi}},$$

which goes to 0 for $\lambda_2 \to 0$, regardless of the value of z_2. In this specific case, the existence of the derivative at the kink is due to the symmetry of the whole PDE problem and boundary condition.

Case 3: $\mu_1 = 2$, $\mu_2 = 0$. Here, again by elementary but lengthy calculation[3]*,*

$$\frac{\partial}{\partial \lambda_2} u(z, t, \lambda) = \frac{\log 2 - z_1 - z_2}{\sqrt{8\pi}} \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1 + \lambda_2}^3} \exp \left(- \frac{(\log 2 - z_1 - z_2)^2 \lambda_2}{4t\lambda_1(\lambda_1 + \lambda_2)} \right).$$

In particular, the derivative vanishes at the kink where $z_1 + z_2 = \log 2$.

[2] See Appendix A.3.
[3] See Appendix A.3.

Example 26. *Consider the same setting as in the previous Example 25, but with payout function*

$$\begin{aligned} h(s_1, s_2) &= \max(s_1^{\mu_1} s_2^{\mu_2} - 2, 0) \\ \Rightarrow g(z_1, z_2) &= \max\left(\exp\left[\frac{\mu_1 + \mu_2}{2} z_1 + \frac{\mu_1 - \mu_2}{2} z_2\right] - 2, 0\right) \end{aligned}$$

For the first case $\mu_1 = 1$, $\mu_2 = 1$, we again have

$$\begin{aligned} \frac{\partial}{\partial \lambda_2} u(z, t, \lambda) &= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}^2} \Phi(x, t, 1) g(z - \sqrt{\lambda} \cdot x) dx \\ &= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}} \Phi(x_1, t, 1) \max(e^{z_1 - \sqrt{\lambda_1} x_1} - 2, 0) \int_{\mathbb{R}} \Phi(x_2, t, 1) dx_2 dx_1 \\ &= 0. \end{aligned}$$

For the second case $\mu_1 = 1$, $\mu_2 = -1$, we find by straightforward calculus[4]*,*

$$\frac{\partial}{\partial \lambda_2} u(z, t, \lambda) = \frac{t e^{z_2 + \lambda_2 t}}{2} \left\{ - \frac{e^{-(z_2 - \log K + 2\lambda_2 t)^2 / 4t\lambda_2}}{\sqrt{\pi t \lambda_2}} + CDF\left(\frac{z_2 - \log K + 2\lambda_2 t}{\sqrt{2t\lambda_2}}\right) \right\}. \tag{3.13}$$

This derivative exists for all $z_2 \in \mathbb{R}$ and all $\lambda_2 > 0$. For $\lambda_2 \to 0$ the limit exists and is well-defined provided $z_2 \neq \log 2$.

If $z_2 = \log 2$ equation (3.13) does not converge for $\lambda_2 \to 0$. In fact, the derivative does not exist. We can however Taylor expand in $\sqrt{\lambda_2}$ instead of λ_2, since $\frac{\partial}{\partial \sqrt{\lambda_2}} u(z_2, t, \lambda_2) = 2\sqrt{\lambda_2} \frac{\partial}{\partial \lambda_2} u(z_2, t, \lambda_2)$ exists. This will lead to an error of size $O(\lambda_2)$ instead of $O(\lambda_2^2)$.

The Non-Smooth Case

The last examples demonstrate the need to understand when differentiation under the integral sign is justified in the general, non-smooth case. The following Lemma 27 states one possible set of conditions. Other characterizations are possible. Our main motivation in choosing this specific lemma was to establish characterization that is precise but sufficiently general to capture the sort of boundary functions that we are interested in.

[4]See Appendix A.4.

Lemma 27. *(Sufficient conditions for differentiation under the integral sign)[7, 27, 43, 28]*

Let $f : \mathbb{R} \times \Lambda \to \mathbb{R}$, $\Lambda \subseteq \mathbb{R}$. Furthermore, let the following conditions be fulfilled

i. $f(x, \lambda)$ *is measurable in* x *and* λ *jointly, and integrable in* x *for almost all* $\lambda \in \Lambda$

ii. $f(x, \lambda)$ *is absolutely continuous in* λ *for almost all* $x \in \mathbb{R}$ *(which implies* $\frac{\partial f}{\partial \lambda}$ *exists almost everywhere)*

iii. $\frac{\partial f}{\partial \lambda}$ *is locally integrable, i.e.,* $\forall [a, b] \subset \Lambda$

$$\int_a^b \int_{\mathbb{R}} \left| \frac{\partial f}{\partial \lambda}(x, \lambda) \right| dx d\lambda < \infty$$

Then $\int_X f(x, \lambda) dx$ *is absolutely continuous in* λ *and for almost all* $\lambda \in \Lambda$

$$\frac{\partial}{\partial \lambda} \int_{\mathbb{R}} f(x, \lambda) dx = \int_{\mathbb{R}} \frac{\partial f}{\partial \lambda}(x, \lambda) dx \tag{3.14}$$

The generalisation to multi-dimensional $x \in \mathbb{R}^N$ *and* $\lambda \in \Lambda \subseteq \mathbb{R}^N$ *is straightforward.*

Take $x \in \mathbb{R}^N$, $\lambda \in \Lambda = [0, \infty)^N$ and

$$f(\lambda, x; t) = \Phi(x, t, 1) g(z - \sqrt{\lambda} \cdot x).$$

We assume that f satisfies (i), which is true for a very wide class of initial conditions. Without it the corresponding heat PDE would have no explicit, classical solution via Green's function. To guarantee (i) it is for example enough that g is piecewise continuous and grows at most like $e^{C\|z\|^{2-\epsilon}}$ for some constants $C, \epsilon > 0$. In the following we will thus focus on (ii) and (iii). We start with the case $\lambda_j > 0$.

Lemma 28. *Let either*

1. $g \in C^m$ *in* z_j *or*

2. $g \in C^{m-1}$ *in* z_j *with Lipschitz continuous partial derivatives of order* $m - 1$.

Then

$$\frac{\partial}{\partial \lambda_j} u(z,t,\lambda) = \int_{\mathbb{R}^N} \Phi(x,t,1) \frac{\partial}{\partial \lambda_j} g(z - \sqrt{\lambda} \cdot x) dx$$

for all $\lambda_j > 0$.

Proof. If $g \in C^1$ then

$$\frac{\partial}{\partial \lambda_j} f(\lambda, x; t) = \frac{-\Phi(x,t,1)}{2\sqrt{\lambda_j}} \frac{\partial g}{\partial z_j}(z - \sqrt{\lambda} \cdot x) \tag{3.15}$$

and thus (ii)–(iii) are clearly fulfilled for all $\lambda_j > 0$. Similarly, we get the existence for all partial derivatives up to order m for $g \in C^m$.

If g is merely Lipschitz continuous, then $g(z - \sqrt{\lambda} \cdot x)$ is Lipschitz continuous in λ_j for all x_j. Thus $g(z - \sqrt{\lambda} \cdot x)$ is absolutely continuous in λ_j for all x_j and fulfils (ii). Since $g(z)$ is Lipschitz, its derivative is essentially bounded by its Lipschitz constant L_g. This gives

$$\begin{aligned}
&\int_a^b \int_{\mathbb{R}^N} \left| \frac{\partial f}{\partial \lambda_j}(\lambda, x) \right| dx d\lambda_j \\
&= \int_a^b \int_{\mathbb{R}^N} \left| \frac{\partial}{\partial \lambda_j} \left(\Phi(x,t,1) g(z - \sqrt{\lambda} \cdot x) \right) \right| dx d\lambda_j \\
&= \int_a^b \int_{\mathbb{R}^N} \left| \Phi(x,t,1) \frac{-x_j}{2\sqrt{\lambda_j}} \frac{\partial g}{\partial z_j}(z - \sqrt{\lambda} \cdot x) \right| dx d\lambda_j \\
&\leq \int_a^b \int_{\mathbb{R}^N} \left| \Phi(x,t,1) \frac{x_j L_g}{2\sqrt{\lambda_j}} \right| dx d\lambda_j = \int_a^b \frac{L_g}{2\sqrt{\lambda_j}} \int_{\mathbb{R}^N} |x_j \Phi(x,t,1)| \, dx d\lambda_j \qquad (3.16) \\
&< \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \infty \quad \forall \lambda_j > 0
\end{aligned}$$

and thus f fulfils (iii). □

In our expansion approach, we want to expand around a λ_0 that has only few non-zero components, because this causes a reduction in dimensionality for the underlying PDE problems. In the Taylor expansion (3.1) we will thus need partial derivatives in directions j at points with $\lambda_j = 0$. In these cases we consider right-sided derivatives, as u is only defined for $\lambda \geq 0$.

Unfortunately, $\lambda_j = 0$ is not covered by the previous lemma, since the right hand side expressions in equations (3.15) and (3.16) diverge for $\lambda_j \to 0$. We will now focus our analysis on this case.

The following Lemma 32 gives sufficient conditions on the initial condition g to guarantee that condition (iii), local integrability of the partial derivative, is fulfilled.

Definition 29. *In the following, let x' be shorthand for $(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_N)$, and dx' be shorthand for $dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_N$. We also write $\sqrt{\lambda} = (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_N})$ and $x \cdot y = (x_1 y_1, \ldots, x_N y_N)$ for the element-wise product.*

Definition 30. *We call a function f interval-wise C^k if there exist finitely many points $a_0 = \infty$, $a_1, \ldots, a_N$, $a_{N+1} = \infty$ such that*

$$f \in C^k\left((a_i, a_{i+1})\right) \ \forall\, 0 \leq i \leq N.$$

Definition 31. *Let f and g be real-valued functions defined almost everywhere on $\mathbb{R}$. We call f g-integrable in x if*

$$\int_{\mathbb{R}} g(x) f(x)\, dx$$

exists and is finite.

Lemma 32. *Let the boundary function g be such that, for*

$$G_j(z', z_j, t, \lambda) = \int_{\mathbb{R}^{N-1}} \Phi(x', t, 1)\, g(z - \sqrt{\lambda} \cdot x')\, dx',$$

and every fixed z', t, λ,

1. *$G_j(z', \cdot, t, \lambda)$ is interval-wise C^2;*
2. *$G_j(z', \cdot, t, \lambda)$ is $\Phi(\cdot, t, 1)$-integrable;*
3. *$\frac{\partial}{\partial z_j} G_j(z', z_j, t, \lambda)$ is $z_j \Phi(z_j, t, 1)$-integrable in z_j.*

Then the following holds:

1. $\frac{\partial}{\partial \lambda_j} u(z,t,\lambda)$ *exists for all* λ *with* $\lambda_j > 0$*;*

2. *if, for a given* z_j*, the derivative* $\frac{\partial^2 G_j}{\partial z_j^2}(z', z_j, t, \lambda)$ *exists, then*

$$\frac{\partial}{\partial \lambda_j} u(z,t,\lambda)\bigg|_{\lambda_j=0} = t \left.\frac{\partial^2 G_j}{\partial z_j^2}(z', z_j, t, \lambda)\right|_{\lambda_j=0}.$$

Proof. See Appendix B.2. □

Remark 33. *Since* $\Phi(x_j, t, 1) = \exp(-x_j^2/4t)/4t\pi$*, the integrability requirements in the preceding lemma are rather weak technical conditions. We will focus more on the differentiability requirements. In the context of derivatives pricing, both the integrability and the differentiability conditions are usually satisfied for realistic payout functions.*

In particular, the additional integrability condition on the absolute value of the first derivative serves primarily to exclude certain pathological examples, such as strongly oscillating functions with accumulation points. The need for this arises somewhat artificially from our definition of interval-wise C^k *functions. One could presumably weaken the requirements here, but this would unnecessarily complicated matters for the kind of applications that we are interested in.*

Corollary 34 (to Corollary 24)**.** *Let* u *be a solution to (2.13) and* G *from (3.10). Assume the conditions imposed in Corollary 24 on* g *on* $\mathbb{R}^N$ *are replaced by the same conditions on* $G(z_1, \ldots, z_r, \cdot, \ldots, \cdot, t)$ *on* $\mathbb{R}^{N-r}$*, for fixed* $z_1, \ldots, z_r$*. Then the result of Corollary 24 still holds for all* $i, j > r$ *and all* $\lambda \geq 0$ *such that* $\lambda_1, \ldots, \lambda_r > 0$*.*

Data Smoothing

Lemma 32 gives a sufficient but indirect condition in terms of G_j. To understand what conditions on g guarantee that G_j satisfies the requirements, it is instructive to consider first the two-dimensional case. The key idea is that integration in one direction, z_1, helps smooth out discontinuities in z_2, as long these discontinuities are not orthogonal to z_1.

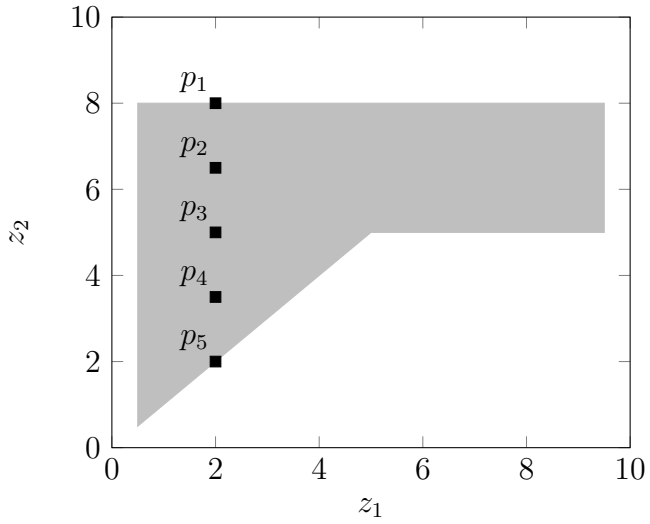


Figure 3.1: Schematic of the payout function for Example 35, to illustrate how integration in z_1 affects the differentiability in z_2. $g \equiv 1$ inside the shaded area and $g \equiv 0$ everywhere else.

Example 35 (Mixed smoothness)**.** *Consider Figure 3.1. Out of the five marked points, g is twice continuously differentiable only in points p_2, p_3 and p_4. Once integrated over z_1, the resulting function G_1 is twice continuously differentiable in p_2, p_4 and p_5. Integration helps to smooth out the discontinuity in p_5. It does not help for p_1, where the kink is orthogonal to the z_2-direction and the discontinuity is orthogonal to z_1. Integration over z_1 decreases the smoothness in p_3, since g is differentiable in p_3 but not differentiable on $\{(z_1, z_2) : 5 \leq z_1 \leq 10, z_2 = 5\}$.*

Example 36 (Single kink or discontinuity)**.** *Assume now that $g : \mathbb{R}^2 \to \mathbb{R}$ is twice continuously differentiable everywhere but on a set*

$$\{(z_1, z_2) : f(z_1, z_2) = 0\}$$

given by the roots of a smooth function $f : \mathbb{R}^2 \to \mathbb{R}$. For simplicity assume a single kink or discontinuity. More specifically, we assume that $\frac{\partial f}{\partial z_1} > 0$ (or < 0) everywhere. This is equivalent to demanding $\nabla f(b(z_2), z_2) \nparallel (0, 1)$, i.e., that the gradient of f is not parallel to the z_2-axis.

Then for every value of z_2 there is at most one value of z_1 such that $f(z_1, z_2) = 0$, and, using the implicit function theorem there exists a smooth function $b : \mathbb{R} \to \mathbb{R}$ describing the location of the kink or discontinuity,

$$f(b(z_2), z_2) = 0 \quad \forall z_2 \in \mathbb{R}.$$

Writing out the partial derivative now leads to

$$\begin{aligned}
\frac{\partial}{\partial z_2} G_1(z, t, x_2, \lambda) &= \frac{\partial}{\partial z_2} \int_{\mathbb{R}} \Phi(x_1, t, 1) g(z - \sqrt{\lambda} \cdot x) dx_1 \\
&= \frac{\partial}{\partial z_2} \left[\int_{-\infty}^{b(z_2)} \Phi(x_1, t, 1) g(z - \sqrt{\lambda} \cdot x) dx_1 \right. \\
&\qquad \left. + \int_{b(z_2)}^{\infty} \Phi(x_1, t, 1) g(z - \sqrt{\lambda} \cdot x) dx_1 \right] \\
&= \frac{\partial b}{\partial z_2}(z_2) \Phi(b(z_2), t, 1) \left[g(z - \sqrt{\lambda} \cdot (b(z_2)^-, x_2)) \right. \\
&\qquad \left. - g(z - \sqrt{\lambda} \cdot (b(z_2)^+, x_2)) \right] \\
&\quad + \int_{-\infty}^{b(z_2)} \Phi(x_1, t, 1) \frac{\partial g}{\partial z_2}(z - \sqrt{\lambda} \cdot x) dx_1 \\
&\quad + \int_{b(z_2)}^{\infty} \Phi(x_1, t, 1) \frac{\partial g}{\partial z_2}(z - \sqrt{\lambda} \cdot x) dx_1
\end{aligned}$$

b is a smooth function and g is C^2 in z_2 everywhere but at the kink. Thus every term in the sum and the sum itself is continuously differentiable in z_2. This can also be seen by explicitly writing out the second derivative

$$
\begin{aligned}
&\frac{\partial^2}{\partial z_2^2} G_1(z,t,x_2,\lambda) \\
=&\frac{\partial^2 b}{\partial z_2^2}(z_2)\Phi(b(z_2),t,1)\left[g(z-\sqrt{\lambda}\cdot(b(z_2)^-,x_2))\right. \\
&\qquad \left.-g(z-\sqrt{\lambda}\cdot(b(z_2)^+,x_2))\right] \\
&+\left(\frac{\partial b}{\partial z_2}(z_2)\right)^2 \frac{\partial \Phi}{\partial z_1}(b(z_2),t,1) \\
&\qquad \cdot\left[g(z-\sqrt{\lambda}\cdot(b(z_2)^-,x_2))-g(z-\sqrt{\lambda}\cdot(b(z_2)^+,x_2))\right] \\
&+\frac{\partial b}{\partial z_2}(z_2)\Phi(b(z_2),t,1)\frac{\partial}{\partial z_2} \\
&\qquad \cdot\left[g(z-\sqrt{\lambda}\cdot(b(z_2)^-,x_2))-g(z-\sqrt{\lambda}\cdot(b(z_2)^+,x_2))\right] \\
&+\frac{\partial b}{\partial z_2}(z_2)\Phi(b(z_2),t,1) \\
&\qquad \cdot\left[\frac{\partial g}{\partial z_2}(z-\sqrt{\lambda}\cdot(b(z_2)^-,x_2))-\frac{\partial g}{\partial z_2}(z-\sqrt{\lambda}\cdot(b(z_2)^+,x_2))\right] \\
&+\int_{-\infty}^{b(z_2)} \Phi(x_1,t,1)\frac{\partial^2 g}{\partial z_2^2}(z-\sqrt{\lambda}\cdot x)dx_1 \\
&+\int_{b(z_2)}^{\infty} \Phi(x_1,t,1)\frac{\partial^2 g}{\partial z_2^2}(z-\sqrt{\lambda}\cdot x)dx_1
\end{aligned}
$$

With the previously mentioned conditions on b and g this is a continuous function. Generalising this argument to finitely many kink points per z_2 by observing that f is then locally invertible is straightforward.

In higher dimensions, for z_j-derivatives of G to exist in a point $z^0 = (z_1^0, \ldots, z_N^0)$, it is sufficient that either

- g is sufficiently z_j-differentiable for all points in a plane through z^0 orthogonal to the z_j-axis, or,
- smoothing of any discontinuities occurs through integration over $z_1, \ldots, z_r$, which happens if the location of the discontinuities is not parallel to all $z_1, \ldots, z_r$. This is guaranteed if the set is locally described by an equation

$$f(z_1, \ldots, z_r; z_{r+1}^0, \ldots, z_{j+1}^0, z_j, z_{j+1}^0, \ldots, z_N^0) = 0,$$

for fixed $z^0_{r+1}, \ldots, z^0_{j-1}, z^0_{j+1}, \ldots, z^0_N$, with smooth f, and

$$\frac{\partial}{\partial z_k} f(z_1, \ldots, z_r; z^0_{r+1}, \ldots, z^0_{j+1}, z_j, z^0_{j+1}, \ldots, z^0_N) \neq 0$$

for at least one $1 \leq k \leq r$, i.e., the kink set is not parallel to the plane and smoothing occurs across it in at least one variable z_k.

Remark 37. *An alternative method of dealing with irregular boundary conditions is to convolute with a mollifier function ξ. This gives a sequence of smoothing functions*

$$g_\epsilon(x) := \int_{\mathbb{R}^N} \epsilon^{-N} \xi(y/\epsilon) g(x-y)\, dy$$

with $g_\epsilon \to g$. While the error in this approximation can be made arbitrarily small, the size of the derivatives of g_ϵ can grow large. Since the latter feature in the error bounds of the expansion method, this would lead to a trade-off between approximation and expansion error. The point is that for $r > 0$ smoothing in the first r directions will take care of this automatically in most circumstances.

3.3 Error Bounds for a First Order Expansion

A good choice of λ^0 and the number of terms to include in the Taylor expansion depends on the problem at hand. However, it is a common feature of processes with strong correlation that there is a dominant eigenvalue which is much larger than the rest of the spectrum. This is also the case for asset models investigated in the Chapters 4, 6, 5.3.

With this in mind, consider again the expansion encoding ξ from Example 17. It corresponds to a first order Taylor expansion of u in λ around the point $(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$, where the first order derivatives are calculated via simple finite differences. As shown, the expansion error $\hat{u}^\xi(z,t,\lambda) = u^\xi(z,t,\lambda) - u(z,t,\lambda)$ fulfils

Proposition 38.

$$\hat{u}^\xi(z,t,\lambda) = O(\lambda^2_{r+1} + \ldots + \lambda^2_N)$$

provided that $u(z,t,\lambda) \in C^2(\lambda)$.

With the results from the last section we can be more specific.

Theorem 39. *Let* $g \in C^{2,2,mix}$. *Then the expansion error* $\hat{u}^\xi(z,t,\lambda)$ *is bound by*

$$|\hat{u}^\xi(z,t,\lambda)| \leq t^2 \sum_{k>l\geq r+1}^{N} \lambda_k \lambda_l \left\| \frac{\partial^4}{\partial z_k^2 \partial z_l^2} g \right\|_\infty$$

Proof. Example 17 together with Section 3.2. □

Remark 40. *The smoothness requirements can be weakened considerably according to the ideas presented in Corollary 21. Furthermore, a structually similar bound is likely to hold even for initial conditions* g *that have isolated kinks after smoothing with the heat kernel in* $\mathbb{R}^r$, *as analysed in Chapter 3.2.*

These bounds follow from the error terms in the Taylor expansion and bounds for the derivatives therein. We can alternatively derive bounds via the PDE for the error itself, as introduced in Chapter 2.4: The relevant differential operators for the heat equation are

$$\begin{aligned}
\mathcal{L} = \mathcal{L}^{\{1,\dots,N\}} &= \sum_{k=1}^{N} \lambda_k \frac{\partial^2}{\partial z_k^2} \\
\mathcal{L}^{\{1,\dots,r\}} &= \sum_{k=1}^{r} \lambda_k \frac{\partial^2}{\partial z_k^2} \\
\mathcal{L}^{\{1,\dots,r,i\}} &= \mathcal{L}^{\{1,\dots,r\}} + \lambda_i \frac{\partial^2}{\partial z_i^2}
\end{aligned}$$

and so on. Using this we can rewrite equation (2.21) as

$$\begin{aligned}
\frac{\partial}{\partial t}\hat{u}^\xi(z,t) = &\mathcal{L}^{\{1,\dots,r\}}\hat{u}^\xi(z,t) \\
&+ (1+r-N)\left[\mathcal{L}^{\{1,\dots,r\}} - \mathcal{L}^{\{1,\dots,r\}}\right] u^{\{1,\dots,r\}}(z,t) \\
&+ \sum_{i=r+1}^{N} \left[\mathcal{L}^{\{1,\dots,r,i\}} - \mathcal{L}^{\{1,\dots,r\}}\right] u^{\{1,\dots,r,i\}}(z,t) \\
&+ \left[\mathcal{L}^{\{1,\dots,r\}} - \mathcal{L}\right] u(z,t) \qquad (3.17) \\
= &\sum_{k=1}^{r} \lambda_k \frac{\partial^2}{\partial z_k^2}\hat{u}^\xi(z,t) + \sum_{k=r+1}^{N} \lambda_k \frac{\partial^2}{\partial z_k^2}\left[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)\right]
\end{aligned}$$

Proposition 41. *For time-constant A_{ij} and the expansion given in equation (3.8), the expansion error $\hat{u}^\xi$ solves the N-dimensional non-homogeneous heat equation*

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}^\xi(z,t) &= \sum_{k=1}^{r}\lambda_k\frac{\partial^2}{\partial z_k^2}\hat{u}^\xi(z,t) + f(z,t,z_0,u,u^\nu) \quad \forall (z,t)\in\mathbb{R}^N\times(0,T]\\ \hat{u}^\xi(z,0) &= 0 \quad \forall z\in\mathbb{R}^N\end{aligned}$$

with source term

$$f(z,t,z_0,u,u^\nu) = \sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}\left[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)\right]$$

Using the heat equation's Green's function we can explicitly give $\hat{u}^\xi(z,t)$ as

$$\hat{u}^\xi(z,t) = \int_0^T\int_{\mathbb{R}^N}\Phi^{\{1,\dots,r\}}(y,s)f(z-y,t-s,z_0,u,u^\nu)dyds$$

This directly leads to bounds on the size of the expansion error, and we can now give a proof of Theorem 19.

Proof (of Theorem 19). Set $\hat{u}^{\{1,\dots,r,k\}}(z,t) = u^{\{1,\dots,r,k\}}(z,t) - u(z,t)$. Then $\hat{u}^{\{1,\dots,r,k\}}(z,t)$ solves the PDE

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}^{\{1,\dots,r,k\}}(z,t) &= \sum_{i\in\{1,\dots,r,k\}}\lambda_i\frac{\partial^2}{\partial z_i^2}u^{\{1,\dots,r,k\}}(z,t) - \sum_{i=1}^{N}\lambda_i\frac{\partial^2}{\partial z_i^2}u(z,t)\\ &= \sum_{i\in\{1,\dots,r,k\}}\lambda_i\frac{\partial^2}{\partial z_i^2}\hat{u}^{\{1,\dots,r,k\}}(z,t) - \sum_{i=r+1,i\neq k}^{N}\lambda_i\frac{\partial^2}{\partial z_i^2}u(z,t)\end{aligned}$$

for all $(z,t)\in\mathbb{R}^N\times(0,T]$ with zero boundary condition. This gives

$$\begin{aligned}\hat{u}^{\{1,\dots,r,k\}}(z,t) = &-\int_0^T\int_{\mathbb{R}^N}\Phi^{\{1,\dots,r,k\}}(y,s)\\ &\cdot\sum_{i=r+1,i\neq k}^{N}\lambda_i\frac{\partial^2}{\partial z_i^2}u(z-y,s)dyds\\ \frac{\partial^2}{\partial z_l^2}\hat{u}^{\{1,\dots,r,k\}}(z,t) = &-\sum_{i=r+1,i\neq k}^{N}\lambda_i\int_0^T\int_{\mathbb{R}^N}\Phi^{\{1,\dots,r,k\}}(y,s)\\ &\cdot\frac{\partial^4}{\partial z_l^2\partial z_i^2}u(z-y,s)dyds\end{aligned}$$

Since it is itself a solution to a PDE, $\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t)$ can be writen as

$$\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t) = \int_{\mathbb{R}^N} \Phi^N(y,t) \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) dy$$

Thus

$$\begin{aligned}
\left| \frac{\partial^2}{\partial z_l^2} \hat{u}^{\{1,\dots,r,k\}}(z,t) \right| \leq & \sum_{i=r+1, i\neq k}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}}(y,s) \\
& \cdot \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty dy ds \\
\leq & t \sum_{i=r+1, i\neq k}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty
\end{aligned}$$

and finally

$$\begin{aligned}
\left| \hat{u}^{\xi}(z,t) \right| = & \left| \sum_{k=r+1}^{N} \lambda_k \int_0^T \int_{\mathbb{R}^r} \prod_{i=1}^{r} \frac{\exp(-y_i^2/4\lambda_i t)}{\sqrt{4\pi\lambda_i t}} \right. \\
& \left. \cdot \frac{\partial^2}{\partial z_k^2} \hat{u}^{\{1,\dots,r,k\}}(z-y',s) dy' ds \right| \\
\leq & \left| \sum_{k=r+1}^{N} \lambda_k \int_0^T \int_{\mathbb{R}^r} \prod_{i=1}^{r} \frac{\exp(-y_i^2/4\lambda_i t)}{\sqrt{4\pi\lambda_i t}} \right. \\
& \left. \cdot \cdot t \sum_{i=r+1, i\neq k}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y') \right\|_\infty dy' ds \right| \\
\leq & t^2 \sum_{k>i\geq r+1}^{N} \lambda_k \lambda_i \left\| \frac{\partial^4}{\partial z_k^2 \partial z_i^2} g(z-y) \right\|_\infty
\end{aligned}$$

□

We can contrast this with the error for a simple zero order PCA approach, i.e., the r-dimensional solution $u^{\{1,\dots,r\}}(z,t)$: Set $\hat{u}^{\{1,\dots,r\}}(z,t) = u^{\{1,\dots,r\}}(z,t) - u(z,t)$. Then $\hat{u}^{\{1,\dots,r\}}(z,t)$ solves the PDE

$$\begin{aligned}
\frac{\partial}{\partial t} \hat{u}^{\{1,\dots,r\}}(z,t) = & \sum_{i\in\{1,\dots,r\}} \lambda_i \frac{\partial^2}{\partial z_i^2} u^{\{1,\dots,r\}}(z,t) - \sum_{i=1}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z,t) \\
= & \sum_{i\in\{1,\dots,r\}} \lambda_i \frac{\partial^2}{\partial z_i^2} \hat{u}^{\{1,\dots,r\}}(z,t) - \sum_{i=r+1}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z,t)
\end{aligned}$$

for all $(z,t) \in \mathbb{R}^N \times (0,T]$ with zero boundary condition. This gives

$$\begin{aligned}
\hat{u}^{\{1,\dots,r\}}(z,t) = &- \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r\}}(y,s) \\
&\cdot \sum_{i=r+1}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z-y,s) dy ds \\
\frac{\partial^2}{\partial z_l^2} \hat{u}^{\{1,\dots,r\}}(z,t) = &- \sum_{i=r+1}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r\}}(y,s) \\
&\cdot \frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z-y,s) dy ds
\end{aligned}$$

Since it is itself a solution to a PDE, $\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t)$ can be written as

$$\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t) = \int_{\mathbb{R}^N} \Phi^N(y,t) \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) dy$$

Thus

$$\begin{aligned}
\left| \frac{\partial^2}{\partial z_l^2} \hat{u}^{\{1,\dots,r\}}(z,t) \right| \leq &\sum_{i=r+1}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r\}}(y,s) \\
&\cdot \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty dy ds \\
\leq &t \sum_{i=r+1}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty .
\end{aligned}$$

We also note that

$$\begin{aligned}
\left| \frac{\partial^2}{\partial z_l^2} \hat{u}^{\{1,\dots,r,k\}}(z,t) \right| \leq &\sum_{i=r+1, i\neq k}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}}(y,s) \\
&\cdot \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty dy ds \\
\leq &t \sum_{i=r+1, i\neq k}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(z-y) \right\|_\infty .
\end{aligned}$$

3.4 Error Bounds for a Second Order Expansion

We can systematically develop a second order expansion. Omitting z and t for readability, we have

$$\begin{aligned}
u(\lambda) =& D^0_\lambda u(\lambda_0) + \sum_{|\alpha|=1,2} \frac{(\lambda-\lambda_0)^\alpha}{\alpha!} D^\alpha_\lambda u(\lambda_0) + R_2(\lambda, \lambda_0, u) \\
=& u(\lambda_0) \\
&+ \sum_{k=r+1}^{N} \left[u(\lambda_0 + \lambda_k e_k) - u(\lambda_0) - \frac{\lambda_k^2}{u} \frac{\partial^2 u}{\partial \lambda_k^2}(\lambda_0) - \lambda_k^3 R^{3\alpha_k}(\lambda, \lambda_0, u) \right] \\
&+ \sum_{k>l\geq r+1}^{N} [u(\lambda_0 + \lambda_k e_k + \lambda_l e_l) - u(\lambda_0 + \lambda_k e_k) \\
&\qquad - u(\lambda_0 + \lambda_l e_l) + u(\lambda_0)] \\
&- \sum_{k>l\geq r+1}^{N} \left[\lambda_k \lambda_k \lambda_l R^{2\alpha_k+\alpha_l}(\lambda, \lambda_0, u) + \lambda_k \lambda_l \lambda_l R^{\alpha_k+2\alpha_l}(\lambda, \lambda_0, u) \right] \\
&+ \sum_{k=r+1}^{N} \frac{\lambda_k^2}{u} \frac{\partial^2 u}{\partial \lambda_k^2}(\lambda_0) + R_2(\lambda, \lambda_0, u) \\
=& \left[1 - (N-r) + ((N-r)^2 - (N-r))/2 \right] u(\lambda_0) \\
&+ \sum_{k=r+1}^{N} \left[1 - (N-r-1+N-r-1)/2 \right] u(\lambda_0 + \lambda_k e_k) \\
&+ \sum_{k>l\geq r+1}^{N} u(\lambda_0 + \lambda_k e_k + \lambda_l e_l) \\
&+ \sum_{k>l>m\geq r+1}^{N} \lambda_k \lambda_l \lambda_m R^{\alpha_k+\alpha_l+\alpha_m}(\lambda, \lambda_0, u)
\end{aligned} \tag{3.18}$$

The coefficients in front of the u terms are the weights of this expansion scheme, as written out explicitly in Proposition 42.

Note that in this expansion, the terms with $\partial^2 u/\partial \lambda_k^2$ from the finite difference approximation for the first derivative and from the second order Taylor sum cancel. This has two advantages: No higher order stencils are necessary for the first order terms, and no finite difference approximation of $\partial^2 u/\partial \lambda_k^2$ is necessary. The latter would have required the computation of the PDE solution at additional points such

as $u(\lambda_0 + 2\lambda_k e_k)$.

Proposition 42. *The expansion*

$$\begin{aligned}\xi = \{&(1 + (N-r)(N-r-3)/2, \{1, \ldots, r\}),\\ &(2-(N-r), \{1, \ldots, r, r+1\}), \ldots, (2-(N-r), \{1, \ldots, r, N\}),\\ &(1, \{1, \ldots, r, r+1, r+2\}), (1, \{1, \ldots, r, r+1, r+3\}), \ldots,\\ &(1, \{1, \ldots, r, N-2, N\}), (1, \{1, \ldots, r, N-1, N\})\}\end{aligned} \tag{3.19}$$

leads to

$$u^{\xi}(z, t, \lambda) = u(z, t, \lambda) + O(\lambda_{r+1}^3 + \ldots + \lambda_N^3) \tag{3.20}$$

provided that $u(z, t, \lambda) \in C^3(\lambda)$

Remark 43. *Computing* u^{ξ} *in (3.20) requires the computation of 1* r*-dimensional,* $N - r$ $(r+1)$*-dimensional and* $(N-r)(N-r-1)/2$ $(r+2)$*-dimensional PDEs. Compared to the first order scheme this adds one dimension to the highest dimensional PDE computations and increases their number by a factor* $O(N-r)$*. In return, the order of approximation increased by one.*

Theorem 44. *For the heat equation (2.13), (2.14), the expansion error* $\hat{u}^{\xi}$ *for* ξ *from (3.19) satisfies*

$$\left|\hat{u}^{\xi}(z, t)\right| \leq t^3 \sum_{i>j>k\geq r+1}^{N} \lambda_i \lambda_j \lambda_k \left\| \frac{\partial^6 g}{\partial z_i^2 \partial z_j^2 \partial z_k^2} \right\|_{\infty} \tag{3.21}$$

provided $g \in C^{2,3,mix}$*.*

Proof. See Appendix B.3. □

Remark 45. *By extension of Remark 20, there are no univariate or bivariate quadratic terms with factors* λ_k^2 *or* $\lambda_j^2 \lambda_k^2$ *in (3.21). Hence, any solutions which only depend on*

one or two of the z_k are integrated exactly, and by superposition any linear combination of such terms. By similar reasoning to before, $\widehat{u}^{\xi}$ is zero for initial conditions of the form

$$g(x) = \sum_{j,k=r+1}^{N} g_{j,k}(x_j, x_k; x_1, \ldots, x_r) \;\; + \sum_{i,j,k=r+1}^{N} x_i \, g_{i,j,k}(x_j, x_k; x_1, \ldots, x_r),$$

for any functions $g_{j,k}$ and $g_{i,j,k}$.

Remark 46. *One can again weaken the smoothness requirements on g, as discussed in Remark 40 for the first order case.*

Chapter 4

Application: European Options in the Black-Scholes Model

4.1 Model and Setup

As a first numerical test case, consider the Black-Scholes model with stocks $S_1, \dots, S_N$, whose dynamics under the risk-neutral measure are given by

$$\mathrm{d}S_i = r_{int} S_i \,\mathrm{d}t + \sigma_i S_i \,\mathrm{d}W_i \quad 1 \leq i \leq N,$$

where $r_{int} \in \mathbb{R}$ is the risk-free interest rate and $\sigma \in \mathbb{R}^N$ describes the stock volatilities. The standard Brownian motions W_i are correlated according to the correlation matrix $(\rho_{ij})_{1 \leq i,j \leq N}$. Here, we use a simple constant correlation structure

$$\rho = \begin{pmatrix} 1 & \gamma & \gamma & \cdots & \gamma \\ \gamma & 1 & \gamma & \cdots & \gamma \\ \vdots & & \ddots & \vdots & \\ \gamma & \gamma & \gamma & \cdots & 1 \end{pmatrix}$$

for some $\gamma \in (-1, 1)$. The covariance matrix Σ is fully characterized by $\Sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$. For identical volatilities $\sigma_1 = \ldots = \sigma_N$ this implies that there are at most two distinct eigenvalues $\lambda_1 \geq \lambda_2$ – with $\lambda_1 = \lambda_2$ if and only if $\gamma = 0$ – such that $\lambda = (\lambda_1, \lambda_2, , \lambda_2, \ldots, , \lambda_2)$. The normalized eigenvector corresponding to λ_1 is $q_1 = (1, \ldots, 1)/\sqrt{N}$ and the other eigenvectors can be chosen as any orthogonal basis of the orthogonal complement of $\{q_1\}$.

For simplicity, we choose $r_{int} = 0$ for the risk-free interest rate. The arbitrage-free price $V(s_1, s_2, 0)$ at time 0 for a financial derivative with payout $h(S_{1,T}, S_{2,T})$ at time T and initial values $S_0 = (S_{1,0}, S_{2,0}) = (s_1, s_2)$ is then given by

$$V(s_1, s_2, 0) = \mathbb{E}\left[\, V(S, T) \mid S_0 = (s_1, s_2) \,\right]. \tag{4.1}$$

We can estimate this value with desired accuracy using, for example, a Monte Carlo simulation with a sufficiently large number of paths. This provides a reference to compare the PDE solution to.

Using the Feyman-Kac theorem, the problem of calculating the expectation (4.1) can be transformed into the problem of computing the solution $u(\log(S_0), T)$ of

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2} \sum_{i,j=1}^{N} \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} - \sum_{i=1}^{N} \frac{\sigma_i^2}{2} \frac{\partial u}{\partial x_i}, \\ u(x, 0) &= g(\exp(x)), \end{aligned}$$

where we have switched to logarithmic coordinates $x = \log(S)$ and reversed time $t \to T - t$.

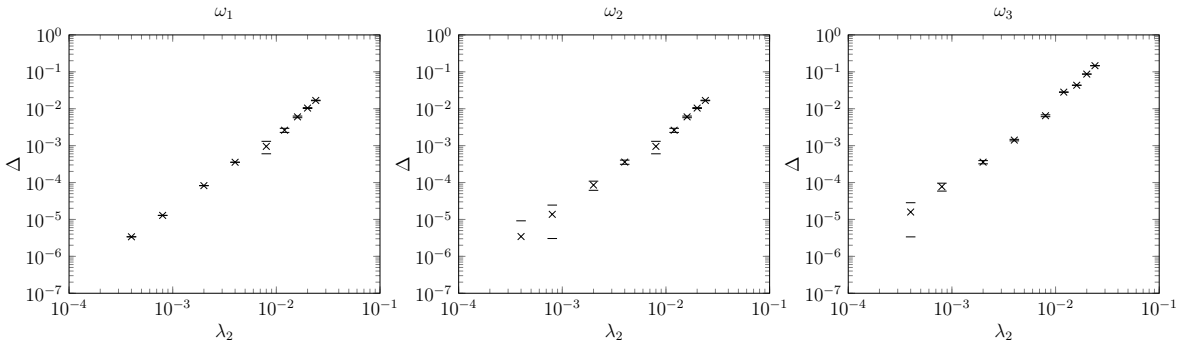
4.2 Convergence of the First Order Expansion

We consider first a European arithmetic basket option with maturity $T = 1$ and payout

$$g(S) = \max\left(\sum_{i=1}^{N} \omega_i S_i - K, 0 \right),$$

where $\omega \in \mathbb{R}^N$ is a vector of weights and K is the strike, which we set to $K = 100$. We are interested in the point $S_{0,i} = 100$ for all i. The numerical tests are for $N = 10$ and $\sigma_i = 0.2\ \forall i$.

Because of the structure of Σ, we expand around the the point $\lambda_0 = (\lambda_1, 0, \ldots, 0)$,



γ	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.98	0.99
λ_1	0.104	0.120	0.136	0.152	0.168	0.184	0.192	0.1968	0.1984
λ_2	0.024	0.020	0.016	0.012	0.008	0.004	0.002	0.0008	0.0004

Figure 4.1: Absolute ($\times$) difference between PDE and MC results with 3σ error bounds (–) for varying λ_2 and payout ω_1, ω_2 and ω_3 using the first order expansion method. The best fit exponents for the differences are 2.04 ± 0.11, 2.03 ± 0.11 and 2.21 ± 0.17 (95% confidence bounds from a linear regression on a log-log scale).

i.e., choose $r = 1$ in the notation of Chapter 3. Figure 4.1 shows results for

$$\omega_1 = (1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10)$$

$$\omega_2 = (4/30, 4/30, 4/30, 4/30, 4/30, 2/30, 2/30, 2/30, 2/30, 2/30)$$

$$\omega_3 = (1/4, 1/4, 1/4, 1/4, 1/4, 1/4, 1/4, -1/4, -1/4, -1/4).$$

The values were computed with $J = 800$ grid points in every direction, with 50 time steps for the PDEs and 10^{10} paths with one time step for the MC simulations. For details of the implementation see Chapter 7 and [40].

For the four smallest values of λ_2, the variance of straightforward MC is too large to allow comparisons to the PDE solutions in sensible run time. We thus chose a control variate approach in which the difference between the full and approximated solution was calculated via MC using the same paths for both, instead of calculating the values separately and than subtracting them.

For example, instead of calculating $E[u^{\{1,\dots,r,k\}}] - E[u^{\{1,\dots,r\}}]$ we directly calculated $E[u^{\{1,\dots,r,k\}} - u^{\{1,\dots,r\}}]$ using on set of MC paths. The variance of the latter expectation can be expected to be a factor of λ_k^2 smaller. To achieve an overall error of size ϵ, one

thus needs $O(\lambda_k^2 \epsilon^{-2})$ MC paths, instead of $O(\epsilon^{-2})$. In our case, λ_k was of the order of 10^{-2}, which resulted in a significant run time reduction by factor of about $10,000$. While we have here employed this control variate techniques to test the numerical accuracy of expansion schemes, we believe that conversely, it would be possible to use PDE expansions to speed up MC computations.

From Figure 4.1 it is evident that the difference Δ between approximate PDE and true MC solution seems to follow a power law of the form

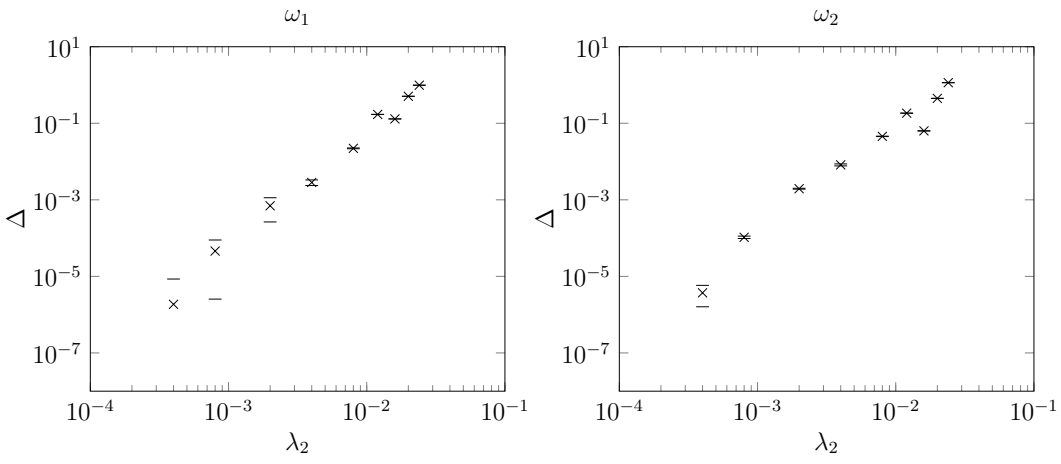
$$\Delta \sim \lambda_2^d = \frac{1}{N-1} \sum_{i=2}^{N} \lambda_i^d.$$

The right hand equality is true because in the chosen dynamics all $\lambda_i, i \geq 2$, have the same value. The least square error best fits for the exponent d are 2.04 ± 0.11, 2.03 ± 0.11 and 2.21 ± 0.17, resp., for ω_1, ω_2 and ω_3. This is consistent with the theoretical lower bound of 2 established in the previous section.

4.3 Convergence of the Second Order Expansion

We consider the same arithmetic basket option model as in Section 4.2 but with $N = 5$ to speed up computations, as $O(N^2)$ three-dimensional PDEs need to be solved here. The values were computed with $J = 500$ grid points in every direction with 50 time steps for the PDEs and 10^{10} paths with one time step for the MC simulations. We use Brian's ADI method [5, 6] for the numerical solution of the three-dimensional PDE problems. For the four smallest values of λ_2 we again directly compute the difference between full and approximate solution via MC, as described in the previous section. The matrix of eigenvectors is

$$Q = \begin{pmatrix} 0.4472 & \star & \star & \star & \star \\ 0.4472 & \star & \star & \star & \star \\ 0.4472 & \star & \star & \star & \star \\ 0.4472 & \star & \star & \star & \star \\ 0.4472 & \star & \star & \star & \star \end{pmatrix},$$



γ	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.98	0.99
λ_1	0.104	0.120	0.136	0.152	0.168	0.184	0.192	0.1968	0.1984
λ_2	0.024	0.020	0.016	0.012	0.008	0.004	0.002	0.0008	0.0004

Figure 4.2: Absolute ($\times$) difference between PDE and MC results with 3σ error bounds (–) for varying λ_2 and payout weights ω_1 and ω_2, using the second order expansion method. The best fit exponents for the absolute differences are 3.04 ± 0.28 and 2.76 ± 0.45.

where, since $\lambda_2 = \ldots = \lambda_N$, the $\star$ entries can be any set of normalized, orthogonal vectors that span the $N-1$ dimensional subspace orthogonal to the first eigenvector. For payout vectors closely aligned with the first eigenvector, such as $\omega = (1/5, 1/5, 1/5, 1/5, 1/5)$, the second order scheme is extremely accurate and the resulting absolute error is considerably less than 10^{-5} even for larger values of λ_2 considered here. We instead choose payout vectors

$$\omega_1 = (1, -1, 1, -1, 1)$$

$$\omega_2 = (3/2, 3/2, -1/2, -1/2, -1)$$

for which the error is relatively large, since they are not aligned with any of the eigenvectors. This allows us to show convergence across a wide range of γ values (Figure 4.2). The data points again appear to follow a power law $\Delta \sim \lambda_2^d = \frac{1}{N-1}\sum_{i=2}^{N} \lambda_i^d$. The best fits for the exponent d are 3.04 ± 0.28 and 2.76 ± 0.45 (95% confidence bounds from a linear fit), compared to the theoretical lower bound of 3.

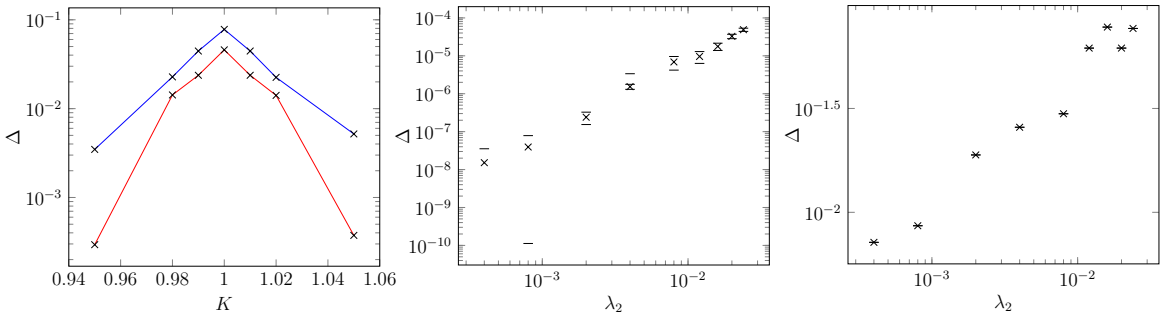


Figure 4.3: Left: Absolute error between PDE and MC solution for a geometric basket with payout vector ω_1' (red/bottom) and ω_2' (blue/top) around the kink point at $K = 1$ for $\gamma = 0.6$ for fixed spot price and varying strike. The option value is of the order of $5 \cdot 10^{-2}$. Centre/Right: Absolute errors with 3σ error bounds (–) for ω_1' with different values for λ_2 at $K = 0.5$ (centre) and $K = 1$ (right). The best fit exponents for the errors are 2.01 ± 0.13 and 0.61 ± 0.10 (95% confidence bounds from a linear regression on a log-log scale).

4.4 The Effect Of Kinks

To demonstrate how the solution behaves around kink points, we switch back to the first order scheme and to a geometric basket option with payout

$$g(S) = \max\left(\prod_{i=1}^{N} S_i^{\omega_i} - K, 0\right) = \max\left(\exp\left(\sum_{i=1}^{N} \omega_i \log S_i\right) - K, 0\right).$$

Due to the log-normality this results in a linear kink set, for which we can easily choose payout vectors that are either exactly parallel or orthogonal to the eigenvectors of Σ. For payout vectors that are not orthogonal to q_1, such as $\omega = q_1$ or $\omega = q_1 + q_2$, this expansion scheme is extremely accurate, even at the kink point $K = 100$. This was expected from the previous section. We thus focus on the orthogonal case where no smoothing occurs.

Consider payout vectors which are orthogonal to q_1, namely[1]

$$\omega_1' = \begin{pmatrix} -0.1160 \\ 0.0929 \\ -0.6527 \\ -0.1121 \\ 0.6986 \\ 0.2091 \\ -0.0438 \\ -0.0758 \\ 0.0000 \\ 0.000 \end{pmatrix} \text{ and } \omega_2' = \begin{pmatrix} 0.1130 \\ -0.0607 \\ -0.1708 \\ -0.2057 \\ 0.8971 \\ -0.2467 \\ -0.1831 \\ -0.1085 \\ -0.0345 \\ -0.0001 \end{pmatrix}.$$

The kink point is now located at strike price $K = 1$. Figure 4.3 shows that there is indeed a marked increase in the error size around the kink and that the scheme is very accurate away from it. The estimated convergence order away from the kink is 2.01 ± 0.13, which is consistent with the theoretically predicted order of 2. At the kink, the order of convergence is reduced to 0.61 ± 0.10.

It is worth emphasising that a high order of convergence is achieved in all but the at-kink, orthogonal case, despite the highly non-smooth initial data. This is in line with and empirically validates our theoretical analysis in the previous chapter.

[1]These two eigenvectors were taken from a MATLAB decomposition for Σ.

Chapter 5

Variable Coefficient PDEs

5.1 General Approach and Parametrix

In the previous two chapters we have investigated how to apply expansion methods — in theory and in practice — to PDE problems with constant coefficients. We now turn our attention to the case of non-constant PDE coefficients. In the context of our particular interest in mathematical finance and derivative pricing, these correspond to more realistic asset models that capture time- and asset value-dependency of drifts, correlations and volatilities.

In principle, there are three ways to proceed:

- Approximate the non-constant coefficient PDE with one containing only constant coefficients and apply expansion methods to the latter. This corresponds to the localization discussed in Chapter 2.5.

  When it can be utilised successfully, this approach is very convenient. Its applicability depends on how good a constant coefficient approximation can be found for the particular problem. For popular models, such as the LIBOR market model, approximations have often been developed and tested in practice already [31]. We investigate this approach in Chapter 6 for financial derivatives in the LIBOR market model.

- Adapt the expansion methods developed for the constant coefficient case to this more general setting, while keeping the transformation matrix Q constant. We can use the expansion encodings derived in Chapter 3 if we incorporate non-diagonal terms and keep track of the changing PDE coefficients during the computation. We investigate this approach in Chapter 5.3 using stock models that are natural generalisations of the log-normal model studied in Chapter 4.

- Develop entirely new expansion methods for general or specific non-constant coefficient PDEs. One example would be to make the initial coordinate transformation dynamic by using $Q(t)$ instead of $Q(0)$, provided the coefficients are at most time-dependent. This would keep the matrix $\lambda(t)$ diagonalised at all times t. Since we are now rotating the coordinate system during the computation, it would require for each time step and each lower-dimensional problem to extrapolate the solution within the new low-dimensional space from the solutions in the previous low-dimensional spaces — a highly non-trivial problem. We do not pursue this direction further in this thesis.

We are interested in applying the expansion method developed for the constant case and with constant transformation to the full PDE problem (2.7)-(2.9) and its restricted versions (2.16)-(2.17) in the case of non-constant coefficients. These PDEs can no longer be transformed into the heat equation for all times and asset values. Instead, after the coordinate transformation (2.6) we are faced with problems of the form

$$\mathcal{L}(z,t)u(z,t) = f(z,t) \tag{5.1}$$

$$\mathcal{L}(z,t) = \frac{\partial}{\partial t} - \sum_{k,l=1}^{N} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k \partial z_l} - \sum_{k=1}^{N} \kappa_k(z,t)\frac{\partial}{\partial z_k} \tag{5.2}$$

$$u(z,0) = g(z) \tag{5.3}$$

on $\Omega = \mathbb{R}^N \times [0, T]$. The transformation (2.6) assures that

$$\lambda_{kl}(z_0, 0) = 0 \quad \forall k \neq l$$

$$\kappa_k(z_0, 0) = 0 \quad \forall k$$

but this is only true at $t = 0$ and $x = z_0$. The difference to the localized problem in the previous section is that since the PDE coefficients $\lambda(z, t)$ and $\kappa(z, t)$ change with spatial and time coordinates, the PDE will in general contain first order and non-diagonal second order terms.

A priori, it is not clear whether this Cauchy problem has a solution, and if so, whether it is unique and can be given explicitly. The existing mathematical literature tends to focus either on special, restricted cases relevant to problems in engineering and physics or on establishing very general existence results. The latter is valuable to us insofar as it allows us to establish existence of the original and expansion solutions, but this is not sufficient to understand the accuracy of any specific expansion approach.

Recall the approach in Chapters 2.5 and 3. Dealing with the heat equation allowed us to construct explicit solutions via the familiar Green's method approach. The properties of the solutions — in particular their size — could thus be linked to the initial conditions and the PDE coefficients. This made it possible to give error bounds in terms of the size of the eigenvalues λ, both via Taylor expansion and via considering the PDEs satisfied by the expansion error itself.

There exists a less well-known fundamental solution approach for variable coefficient PDEs of the form (5.1)-(5.3), that is in a sense a generalisation of the Green's function method. It is called the parametrix method [13] and originally due to E. E. Levi.

Theorem 47. *(Parametrix method [13])*

Let $\mathcal{L}$ be uniformly parabolic on Ω, i.e., let there exist constants $c_0 > 0$ and c_1 such that $c_0|z|^2 \leq \sum_{i,j=1}^{N} \lambda_{kl}(z, t) z_i z_j \leq c_1|z|^2$ for all $(z, t) \in \Omega$. Let λ and κ be bounded

continuous and let there be constants A, B, α and β such that

$$|\lambda_{ij}(z,t) - \lambda_{ij}(z_0,t_0)| \leq A(|z-z_0|^\alpha + |t-t_0|^{\alpha/2})$$
$$|\kappa_i(z,t) - \kappa_i(z_0,t_0)| \leq A|z-z_0|^\alpha$$
$$|g(z)| \leq Be^{\beta|z|^2}$$

Then $u(z,t)$ given by

$$u(z,t) = \int_{\mathbb{R}^N} \Gamma(z,t,x,0)g(x)\,dx$$

is a solution of the problem (5.1)-(5.3) for $f(z,t) \equiv 0$. Here

$$\Gamma(z,t,x,t') = Z(z,t,x,t') + \int_{t'}^{t}\int_{\mathbb{R}^N} Z(z,t,y,\tau)\phi(y,\tau,x,t')\,dyd\tau,$$

where the parametrix Z is given by

$$Z(z,t,x,t') = (2\pi\sqrt{t-t'})^{-N}\sqrt{\det(\lambda(x,t'))}^{-1}\cdot$$
$$\cdot\exp\left(-\frac{\sum_{i,j=1}^N \lambda_{ij}^{-1}(x,t')(z_i-x_i)(z_j-x_j)}{4(t-t')}\right)$$

— i.e., $Z(\cdot,\cdot,x,t')$ is the fundamental solution of the PDE with $\mathcal{L}(x,t')$ fixed at a given (x,t') and vanishing κ — and ϕ is implictly given as the solution of

$$\phi(y,\tau,x,t') = \mathcal{L}(y,\tau)Z(y,\tau,x,t') + \int_{t'}^{\tau}\int_{\mathbb{R}^N} \mathcal{L}(y,\tau)Z(y,\tau,\xi,s)\phi(\xi,s,x,t')d\xi ds.$$

In particular, ϕ exists and can be written as a convergent infinite series.
Let now $f(z,t)$ be continuous with $|f(z,t)| \leq h_1 e^{h_2|z|^2}$, where h_1 and h_2 are constants with $h_2 \leq \bar{c}/(4T)$ and $\bar{c}$ is a constant depending on c_0 and c_1. Let $f(z,t)$ be Hölder continuous locally in z and uniformly in t. Then the function u given by

$$u(z,t) = -\int_0^t\int_{\mathbb{R}^N} \Gamma(z,t,x,t')f(x,t')dxdt'$$

is a solution of the non-homogeneous problem (5.1)-(5.3) with zero boundary condition $g \equiv 0$.

Proof. See [13]. □

The fundamental solution $\Gamma(z,t,x,t')$ is only defined implicitly. It can be constructed via a convergent infinite series, but it cannot be written down explicitly as is the case for the Green's function for constant coefficients.

Theorem 48. *Let the conditions of Theorem 47 be satisfied. Then the solution u of the Cauchy problem (5.1)-(5.3) is bounded by*

$$|u(z,t)| \leq k_1 e^{k_2|z|^2},$$

where k_1 and k_2 are constants depending on the PDE coefficients and initial condition[1]*. Furthermore, u is unique with respect to all solutions that satisfy the boundedness condition*

$$\int_0^T \int_{\mathbb{R}^N} |u(z,t)| e^{-k|z|^2} < \infty.$$

Proof. See [13]. □

With these two theorems, we can express and analyse the expansion error as the unique bounded solution of a PDE. This is similar to the approach used in Chapter 3.3-3.4.

5.2 First Order, First r Eigenvalues Expansion

Consider again the first order expansion around the first r eigenvalues that we discussed in Chapter 3.3 for the case of constant coefficients. It is described by

$$\xi = \{(1+r-N, \{1,\ldots,r\}), (1, \{1,\ldots,r,r+1\}), (1, \{1,,\ldots,r,r+2\}),$$
$$(1, \{1,,\ldots,r,r+3\}), \ldots, (1, \{1,,\ldots,r,N\})\}$$

[1]The dependence is limited via inequalities satisfied due to $\mathcal{L}$ being uniformly parabolic and the growth bounds assumed for g and f in the previous theorem. For details see [13].

We can apply it to non-constant coefficient PDEs as well. This yields an expression similar to equation (3.17), namely

$$
\begin{aligned}
&\frac{\partial}{\partial t}\hat{u}^{\xi}(z,t) \\
&= \mathcal{L}^{\{1,\dots,r\}}(z,t)\hat{u}^{\xi}(z,t) \\
&\quad + \sum_{i=r+1}^{N} \left[\mathcal{L}^{\{1,\dots,r,i\}}(z,t) - \mathcal{L}^{\{1,\dots,r\}}(z,t)\right] u^{\{1,\dots,r,i\}}(z,t) \\
&\quad + \left[\mathcal{L}^{\{1,\dots,r\}}(z,t) - \mathcal{L}(z,t)\right] u(z,t) \\
&= \sum_{k,l=1}^{r} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k z_l}\hat{u}^{\xi}(z,t) \\
&\quad + \sum_{k=r+1}^{N} \left[\lambda_{kk}(z,t)\frac{\partial^2}{\partial z_k^2} + \sum_{l=1}^{r} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k z_l}\right] \left[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)\right] \qquad (5.4) \\
&\quad - \sum_{k,l=r+1,k\neq l}^{N} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k z_l}u(z,t) \qquad (5.5) \\
&\quad + \sum_{k=r+1}^{N} \kappa_k(z,t)\frac{\partial}{\partial z_k}\left[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)\right] \qquad (5.6)
\end{aligned}
$$

This equation contains three source terms, which determine the error size:

- The first term, (5.4), is similar to the source term appearing in the constant coefficient case. It is essentially a restricted differential operator applied to the difference between full and partial solution.
- The second term, (5.5), consists of the non-diagonal terms not captured at all in the expansion applied to the full solution. It contains the full solution rather than the difference between full and partial ones, but the λ_{kl} involved are zero for $t = 0$ and $z = z_0$.
- The third term, (5.6), captures the changes in κ and again acts on the differences between partial and full solutions.

At $t = 0$ and $z = z_0$ all three source terms are zeros, because

$$u^{\{1,\dots,r,k\}}(z,0) - u(z,0) = 0 \,\forall z \in \mathbb{R}^N \text{ and } \lambda_{kl}(z_0,0) = 0.$$

Away from these initial coordinates the terms grow slowly and drive a non-zero error: Let $\Gamma(z,t,x,t')$ be the fundamental solution for the differential operator

$$\mathcal{L}(z,t) = \sum_{k,l=1}^{r} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k z_l}$$

and define

$$u(z,t) = \int_{\mathbb{R}^N} \Gamma(z,t,x,0)g(x)\ dx,$$
$$u^{\{1,\ldots,r,k\}}(z,t) = \int_{\mathbb{R}^N}, \Gamma^{\{1,\ldots,r,k\}}(z,t,x,0)g(x)\ dx$$

where $\Gamma^{\{1,\ldots,r,k\}}$ relates to $\mathcal{L}^{\{1,\ldots,r,k\}}$.

From Theorem 47 we get

$$\begin{aligned}
&\hat{u}^{\xi}(z,t)\\
&= \int_0^t \int_{\mathbb{R}^N} \Gamma(z,t,x,t') \sum_{k=r+1}^{N} \lambda_{kk}(x,t')\frac{\partial^2 \left[u - u^{\{1,\ldots,r,k\}}\right]}{\partial x_k^2}(x,t')dxdt'\\
&\quad + \int_0^t \int_{\mathbb{R}^N} \Gamma(z,t,x,t') \sum_{k=r+1}^{N}\sum_{l=1}^{r} \lambda_{kl}(x,t')\frac{\partial^2 \left[u - u^{\{1,\ldots,r,k\}}\right]}{\partial x_k \partial x_l}(x,t')dxdt'\\
&\quad + \int_0^t \int_{\mathbb{R}^N} \Gamma(z,t,x,t') \sum_{k,l=r+1,k\neq l}^{N} \lambda_{kl}(x,t')\frac{\partial^2 u}{\partial x_k \partial x_l}(x,t')dxdt'\\
&\quad + \int_0^t \int_{\mathbb{R}^N} \Gamma(z,t,x,t') \sum_{k=r+1}^{N} \kappa_{k}(x,t')\frac{\partial \left[u - u^{\{1,\ldots,r,k\}}\right]}{\partial x_k}(x,t')dxdt'
\end{aligned}$$

This shows that the error size is related to the integral of the PDE coefficients as well as the initial boundary conditions. Note that $\Gamma(z,t,\cdot,\cdot)$, u and $u^{\{1,\ldots,r,k\}}$ depend on λ as well. The latter two can be expressed as solutions of the homogeneous PDE problem, which relates their size to the boundary condition g.

The point of this analysis was to sufficiently motivate the use of expansion methods by demonstrating that the error growth is related to properties of the PDE in a similar way as before. For further details on the properties of $\Gamma(z,t,x,t')$ and bounds on its size see [13].

There is little hope for establishing very tight bounds for general λ and κ, contrary to the situation in the constant coefficient case in Chapter 3. If one desires such bounds for a specific problem, it is likely that taking into account the particular coefficients of the corresponding PDE will lead to better results than relying on general upper bounds.

There are six base cases of how the PDE coefficients can be non-constant:

Non-constant component	Parameter	Example
Time-dependent drift	$\mu = \mu(t)$	Exactly described through (2.5)–(2.6)
Time-dependent volatilities	$\sigma = \sigma(t)$	Stock model in Chapter 5.3
Time-dependent correlation	$\rho = \rho(t)$	Stock model in Chapter 5.3
Asset-dependent drift	$\mu = \mu(S)$	LIBOR market model in Chapter 6
Asset-dependent volatilities	$\sigma = \sigma(S)$	Local volatility models
Asset-dependent correlation	$\rho = \rho(S)$	Stock model in Chapter 5.3

5.3 Application: European Options in a Variable Coefficient Stock Model

Consider a stock model with asset dynamics $S_1, \ldots, S_N$ given by

$$dS_i = \sigma_i(S,t)S_i\,dW_i \quad 1 \leq i \leq N,$$

where $\sigma_i : [0,\infty)^N \times [0,T]$ are the volatilities. The Brownian motions W_i are correlated according to the correlation matrix

$$(\rho_{ij})_{1\leq i,j\leq N}(S,t) = (Corr(dW_i(S,t), dW_j(S,t)))_{1\leq i,j\leq N}.$$

We consider two possible correlation structures:

$$\rho_{simple}(\gamma) = \begin{pmatrix} 1 & \gamma & \gamma & \cdots & \gamma \\ \gamma & 1 & \gamma & \cdots & \gamma \\ \vdots & & \ddots & \vdots & \\ \gamma & \gamma & \gamma & \cdots & 1 \end{pmatrix}$$

and

$$\rho_{exp,ij}(\gamma) = \exp(-\gamma|i-j|)$$

with $\gamma(S,t) : \mathbb{R}^N \times [0,T] \to \mathbb{R}$ possibly being asset- and time-dependent. The covariance matrix $\Sigma(S,t)$ is then fully characterised via $\Sigma_{ij}(S,t) = \sigma_i(S,t)\sigma_j(S,t)\rho_{ij}(S,t)$. Due to the asset- and time-dependency of correlations and volatilities, the asset dynamics are no longer log-normal like the one considered in Chapter 4.

As a test case we again choose the European arithmetic basket option considered in Chapter 4 with $N = 10$. The payout at maturity $T = 1$ is

$$g(S) = \max\left(\sum_{i=1}^{N} \omega_i S_i - K, 0\right),$$

with strike $K = 100$. We are interested in the value at the point $S_{0,i} = 100$ for all i. As payout vectors ω we consider again[2]

$$\begin{aligned}
\omega_1 &= (1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10, 1/10) \\
\omega_2 &= (4/30, 4/30, 4/30, 4/30, 4/30, 2/30, 2/30, 2/30, 2/30, 2/30) \\
\omega_3 &= (1/4, 1/4, 1/4, 1/4, 1/4, 1/4, 1/4, -1/4, -1/4, -1/4)
\end{aligned}$$

and expect that the accuracy will be best for ω_1 and worst for ω_3.

The numerical parameters chosen were $J = 800$ and $M = 1000$, corresponding to a time step of size $dt = 0.001$. For the reference Monte Carlo simulation we used 10^8 paths. This time step reduces the time discretization error sufficiently for us to determine a good estimate of the method accuracy.

We implemented and tested two numerical algorithms for the solution of the PDE problems: One is a modification of the diagonal ADI method used for the constant coefficient case, where we now updated the diffusion coefficient values at every time step. Here this does not include any of the off-diagonal terms, though of course we

[2]See Chapter 4.

can include them explicitly[3]. The second method is based on the Hundsdorfer-Verwer (HV) scheme, which does incorporate the off-diagonal terms in the lower dimensional problems. All computations are done in log space. For details see Chapter 7.

We also contrast the results against the diagonal ADI results for a fully localized model, i.e. with covariance matrix fixed at $\Sigma(S_0, 0)$. This allows us to understand how much of an impact the variability of the coefficients has. One might pick a different point than (S_0, t) to localize, of course, but it suffices for our case. We are interested in the general picture rather than the model-dependent best choice.

Our primary intention here is to give a proof of concept, rather than an in-depth study of the performance and convergence as in Chapters 4 and 6. We want to demonstrate that and how expansion methods can be used for variable coefficients. Below we thus considered a number of different model behaviours and calculated the corresponding results via MC, PDE/ADI and PDE/HV.

Time-dependent simple correlation

For time-dependent simple correlation $\rho(t) = \rho_{simple}(t)$ the eigenvalues change over time. However, the lower $N - 1$ eigenvalues are identical and the subspace spanned by their eigenvectors does not change.

The following table shows results for $\sigma_i = 0.2$ and

$$\rho(t) = \rho_{simple}(0.8 - 0.8 \cdot (t/T - 0.5)^2) \in [\rho_{simple}(0.6), \rho_{simple}(0.8)].$$

[3]We have done this for a few test cases and observed that full ADI and HV results were extremely close.

		V_{MC}	$V_{PDE}^{diagADI}$	V_{PDE}^{HV}	V_{MC}^{loc}	V_{PDE}^{loc}
ω_1		6.9463	6.9451	6.9451	6.3784	6.3715
	σ_{MC}	0.0011			0.0010	
	Δ_{abs}		-0.0012	-0.0012		-0.0069
	Δ_{rel}		-0.02%	-0.02%		-0.11%
	Δ_{abs}/σ_{MC}		-1.06	-1.06		-6.73
ω_2		6.9602	6.9584	6.9584	6.3991	6.3932
	σ_{MC}	0.0011			0.0010	
	Δ_{abs}		-0.0018	-0.0018		-0.0059
	Δ_{rel}		-0.03%	-0.03%		-0.09%
	Δ_{abs}/σ_{MC}		-1.57	-1.57		-5.75
ω_3		7.5631	7.5816	7.5816	7.3585	7.4069
	σ_{MC}	0.0012			0.0012	
	Δ_{abs}		0.0185	0.0185		0.0484
	Δ_{rel}		0.24%	0.24%		-0.66%
	Δ_{abs}/σ_{MC}		14.96	14.96		-40.58

PDE/ADI and PDE/HV results were almost identical and very close to the MC results. Only in the third case – ω_3 – did they even differ in a statistically significant way, i.e., relative to the variance σ_{MC}, from the MC computation. It is worth noting that the errors are slightly larger in the fully localized case, implying that the variable coefficients present no particular problem in this model.

Time-dependent exponential correlation

For a time-dependent exponential correlation $\rho(t) = \rho_{exp}(t)$ the eigenvalues and Q change substantially over time, resulting in a significant contribution from non-zero off-diagonal elements in $\lambda(t)$.

The following table shows results for $\sigma_i = 0.2$ and

$$\rho(t) = \rho_{exp}(0.25 - 0.6 \cdot (t/T - 0.5)^2)) \in [\rho_{exp}(0.1), \rho_{exp}(0.25)].$$

		V_{MC}	$V_{PDE}^{diagADI}$	V_{PDE}^{HV}	V_{MC}^{loc}	V_{PDE}^{loc}
ω_1		6.0662	6.0738	6.0590	6.8534	6.8477
	σ_{MC}	0.0010			0.0011	
	Δ_{abs}		0.0076	0.0885		-0.0057
	Δ_{rel}		0.13%	1.46%		-0.08%
	Δ_{abs}/σ_{MC}		7.82	90.88		-5.11
ω_2		6.1646	6.1695	6.1547	6.9109	6.9085
	σ_{MC}	0.0010			0.0011	
	Δ_{abs}		0.0049	-0.0099		-0.0024
	Δ_{rel}		0.08%	-0.16%		-0.03%
	Δ_{abs}/σ_{MC}		4.92	-10.00		-2.15
ω_3		9.6062	9.5346	9.7786	9.2907	9.3279
	σ_{MC}	0.0015			0.0015	
	Δ_{abs}		-0.0716	0.1724		0.0372
	Δ_{rel}		-0.75%	1.80%		-0.40%
	Δ_{abs}/σ_{MC}		-46.34	111.54		24.44

PDE/ADI results are again close to the MC results. The PDE/HV results differ slightly more. The third case, ω_3, is again the most challenging one for our method.

Time-dependent volatilities

For time-dependent $\sigma_i = \sigma(t)$, i.e., the case where all volatilities are time-dependent but equal, the eigenvalues $\lambda_1, \ldots, \lambda_N$ of Σ are simply scaled up or down over time and the matrix Q of eigenvectors stays constant. This means that all non-diagonal terms of λ vanish and the transformation to the heat equation is exact. This case is simple: It merely requires to solve the heat equation with time-dependent diffusion coefficients.

For time-dependent $\sigma_i = \sigma_i(t)$, i.e., the case where the volatilities vary differently over time, the eigenvectors change with t. This in general leads to the appearance of non-zero off-diagonal terms. With no dependency on the asset values S, the initial PDE transformation means that those terms vanish at time $t = 0$ and then grow over time for $t > 0$.

The following table shows results for $\rho = \rho_{simple}(0.7)$ and

$$\sigma_i(t) = 0.1(1 + t/T)\left(1 + \frac{i-1}{N-1}\right) \in [0.1, 0.2]\left(1 + \frac{i-1}{N-1}\right).$$

		V_{MC}	$V_{PDE}^{diagADI}$	V_{PDE}^{HV}	V_{MC}^{loc}	V_{PDE}^{loc}
ω_1		7.7987	7.7947	7.8234	5.1128	5.1123
	σ_{MC}	0.0013			0.0008	
	Δ_{abs}		-0.0040	0.0248		-0.0005
	Δ_{rel}		-0.05%	0.32%		-0.01%
	Δ_{abs}/σ_{MC}		-3.10	19.19		-0.57
ω_2		7.3183	7.3151	7.3416	4.7972	4.7961
	σ_{MC}	0.0012			0.0008	
	Δ_{abs}		-0.0032	0.0233		-0.0011
	Δ_{rel}		-0.04%	0.32%		-0.02%
	Δ_{abs}/σ_{MC}		-2.67	19.39		-1.41
ω_3		6.2074	6.3579	6.3723	4.0555	4.1658
	σ_{MC}	0.0010			0.0006	
	Δ_{abs}		0.1504	0.1649		0.1103
	Δ_{rel}		2.42%	2.66%		-2.72%
	Δ_{abs}/σ_{MC}		150.26	164.72		174.38

Both the PDE/diagonal ADI and PDE/HV results are fairly accurate for the first two test cases. They both struggle with the third one, producing errors of 2.42% and 2.66%. Given that a similar error is present in the fully localized case, i.e., for the model with constant coefficients, we conclude that this error is primarily due to the expansion method be applied to the challenging payout direction ω_3, rather than the non-constant nature of the coefficients.

Time-dependent volatilities and exponential correlation

The following table shows results for

$$\sigma_i(t) = 0.1(1 + t/T)\left(1 + \frac{i-1}{N-1}\right) \in [0.1, 0.2]\left(1 + \frac{i-1}{N-1}\right)$$

and

$$\rho(t) = \rho_{exp}(0.25 - 0.6 \cdot (t/T - 0.5)^2)) \in [\rho_{exp}(0.1), \rho_{exp}(0.25)].$$

		V_{MC}	$V_{PDE}^{diagADI}$	V_{PDE}^{HV}	V_{MC}^{loc}	V_{PDE}^{loc}
ω_1		6.9951	7.0905	7.1454	5.1602	5.1595
	σ_{MC}	0.0012			0.0008	
	Δ_{abs}		0.0955	0.1503		-0.0007
	Δ_{rel}		1.36%	2.15%		-0.01%
	Δ_{abs}/σ_{MC}		83.12	130.87		-0.86
ω_2		6.5570	6.7953	6.7047	4.8380	4.8382
	σ_{MC}	0.0011			0.0008	
	Δ_{abs}		0.2383	0.1477		0.0002
	Δ_{rel}		3.63%	2.25%		0.00%
	Δ_{abs}/σ_{MC}		223.82	138.71		0.27
ω_3		9.6494	10.0537	9.8383	5.6868	5.7252
	σ_{MC}	0.0015			0.0009	
	Δ_{abs}		0.4042	0.1889		0.0384
	Δ_{rel}		4.19%	1.96%		0.67%
	Δ_{abs}/σ_{MC}		265.62	124.13		43.29

By combining time-dependent volatilites with time-dependent correlation we have created a challenging scenario for our method. The PDE/diagonal ADI approach starts to be insufficient for the more complicated cases, differing by more than 4% for ω_3. The PDE/HV algorithm produces are relatively constant error of about 2% in all three test cases.

Contrasting with the fully localized it is evident that this is the first scenario in which the variability of the coefficients creates a major contribution to the overall error.

Asset-dependent correlation

The following table shows results for $\sigma_i = 0.2$ and

$$\rho(S) = \rho_{simple}\left(0.6 + 0.2\exp\left(-\frac{1}{N}\sum_{i}^{N}\frac{|S_i - 100|}{10}\right)\right) \in [\rho_{simple}(0.6), \rho_{simple}(0.8)].$$

Because of the added computational complexity of having to calculate the correlation for every vector of asset values encountered, these calculations were done with 10^6 Monte-Carlo paths, $J = 400$ grid points and $M = 400$ time steps.

		V_{MC}	$V_{PDE}^{diagADI}$	V_{PDE}^{HV}	V_{MC}^{loc}	V_{PDE}^{loc}
ω_1		6.7937	1.4032	6.7393	7.2138	7.2147
	σ_{MC}	0.0108			0.0010	
	Δ_{abs}		-5.3905	-0.0544		-0.0009
	Δ_{rel}		-79.35%	-0.80%		-0.01%
	Δ_{abs}/σ_{MC}		-497.65	-5.02		-0.75
ω_2		6.7910	6.7910	6.7534	7.2232	7.2239
	σ_{MC}	0.0109			0.0010	
	Δ_{abs}		-5.3660	-0.0376		-0.0008
	Δ_{rel}		-79.02%	-0.55%		-0.01%
	Δ_{abs}/σ_{MC}		-494.23	-3.47		-0.64
ω_3		7.4977	2.4238	7.3838	7.6708	7.6663
	σ_{MC}	0.0122			0.0010	
	Δ_{abs}		-5.0739	-0.1139		0.0045
	Δ_{rel}		-67.67%	-1.52%		0.06%
	Δ_{abs}/σ_{MC}		-416.90	-9.36		3.56

Clearly, the PDE/diagonal ADI approach is insufficient and the non-diagonal PDE terms are necessary for the solution. The PDE/HV approach, which incorporates them, correspondingly gives a fairly accurate result for ω_1 and ω_2, relative to the MC variance. As before, the accuracy decreases for the ω_3 case, which coincidentally depends only weakly on the chosen correlation dynamics.

Conclusion

Across all cases, the ω_1 case, where the payout vector is parallel to the first eigenvector of Σ[4], showed the best accuracy, while ω_3 case showed the worst. This was expected from the theoretical analysis and the results for constant coefficients, see Chapter 3. Overall, our computations demonstrate that expansion methods can in principle be applied in this fashion to some variable coefficient asset models. However, higher order methods or other extensions might be necessary to reduce the error sufficiently for real-world financial applications, see Chapter 8.2. More care than in the constant

[4]For ρ_{simple} and ρ_{exp}, the first eigenvector is up to normalisation equal to ω_1 and is neither time- nor asset value-dependent.

coefficient case has to be taken to understand the behaviour and contribution of the error due to time discretisation. Of course, the latter is true for the MC-based calculation as well. In practice, one can find a suitable step size dt by empirically determining the point where both methods remain sufficiently constant under further decreases of dt.

Chapter 6

Application: Financial Derivatives in the LIBOR Market Model

6.1 Overview

In this chapter, which is based on "Numerical Valuation of Derivatives in High-Dimensional Settings via PDE Expansions" [40], we demonstrate the wider applicability of PDE expansion methods in situations where no closed-form solution is known and accurate Monte Carlo estimates are difficult to obtain. A prime candidate for using our methods in practice is the LIBOR market model for the joint evolution of LIBOR rates with different tenors. To value path-dependent products such as TARNs (Targeted Accrual Redemption Notes), Snowballs or Ratchets, and early-exercise options such as Bermudan swaptions, indeed the whole yield curve has to be taken into consideration, which makes the problem genuinely high-dimensional for long enough maturities. The PCA-ANOVA-based PDE expansion approach presented here is very well suited to this setting even in high dimensions, because LIBORs with similar tenors are strongly correlated, such that one observes a fast decay of the eigenvalues, as is seen from Fig. 6.1. On the example of Bermudan swaptions, only a mild loss of accuracy is observed as the dimensionality, determined by the number of LIBORs considered, ranges up to 50–60. This deterioration appears to be an effect mostly of the time to maturity rather than the dimension increase alone. For longer

running contracts, the higher order terms in (1.7) become more relevant.

A similar decay of accuracy for longer maturities is observed with the commonly used Monte Carlo method presented in [2]. There, the necessary restriction of the class of exercise strategies there produces a gap between lower and (dual) upper bounds which widens as the maturity increases. The accuracy of these Monte Carlo results is comparable with the PDE ones, which are obtained in a small fraction of the computational time. Additionally, the expansion (1.7) implicitly defines a systematic accuracy improvement and is relatively straightforward to implement. We study this in Chapter 6.4.

6.2 Model and Setup

We now apply the PCA-ANOVA approach to practically relevant examples from interest rate markets: LIBOR market derivatives. Forward rates will be assumed to follow the LIBOR market model (LMM), which is one of the most widely used models [4, 14, 34] and the basis for a variety of extensions. The methods studied here have the potential to be applied to those as well. Our notation and definition of the LMM follows [14].

As traded product at time t consider a (zero-coupon) bond $P(t,T)$, $T \geq 0$, that pays 1 at time $T \geq t$. The forward LIBOR with fixing date T and payout date T' is then defined on the same probability space $(\Omega, \mathcal{F}, \mathcal{P})$, as the stochastic process $L(\cdot, T, T') : [0,T] \times \Omega \to \mathbb{R}$ given by

$$L(t,T,T') := \frac{1}{T'-T} \frac{P(t,T) - P(t,T')}{P(t,T')}.$$

For a fixed tenor structure

$$0 = T_1 < \ldots < T_N < T_{N+1} = T',$$

the LMM now describes a finite number of forward rates

$$L_i(t) := L(t, T_i, T_{i+1})$$

for $1 \leq i \leq N, t \in [0, T_i]$. Let $T_j - T_{j-1} = \alpha$ for all $2 \leq j \leq N+1$. For example, two practically important values for α are 0.25 and 0.5 for 3-month and 6-month LIBOR. The full dynamics for each $L_i(t)$, $t \leq T_i$, under the equivalent martingale measure $\mathcal{Q}^{n+1}$, $1 \leq n \leq N$, associated with choosing the bond $P(T_{n+1})$ as numéraire, are

$$\mathrm{d}L_i(L,t) = \mu_i(L,t)L_i(t)\,\mathrm{d}t + \sigma_i(t)L_i(t)\,\mathrm{d}W_i^{\mathcal{Q}^{n+1}}, \tag{6.1}$$

where

$$\mu_i(L,t) = \begin{cases} -\sum_{j=i+1}^{n} \frac{\alpha L_j(t)}{1+\alpha L_j(t)} \sigma_i(t)\sigma_j(t)\rho_{ij}(t) & i < n \\ 0 & i = n \\ \sum_{j=i+1}^{N} \frac{\alpha L_j(t)}{1+\alpha L_j(t)} \sigma_i(t)\sigma_j(t)\rho_{ij}(t) & i > n \end{cases} \tag{6.2}$$

for $t \leq T_{n+1}$ and similarly

$$\mu_i(L,t) = \sum_{j=\max\{k:T_k \leq t\}}^{N} \frac{\alpha L_j(t)}{1+\alpha L_j(t)} \sigma_i(t)\sigma_j(t)\rho_{ij}(t) \tag{6.3}$$

for $t > T_{n+1}$. From here on we will use the terminal measure with $n = N$.

Our model for the correlation structure is taken from [29] with

$$\rho_{ij} = \exp(-\phi|i-j|)$$

for $1 \leq i, j \leq N$, $d = N-1$, $\phi > 0$ and a constant volatility $\sigma_i(t) = c = 0.2$ for $1 \leq i \leq N, t \in [0, T_i]$. The eigenvalues of the covariance matrix Σ decrease rapidly. Figure 6.1 demonstrates this for $N = 21$ and $N = 41$ and different values of ϕ.

The first eigenvalue λ_1 is significantly larger than the second and following eigenvalues. For example, for $\phi = 0.0413$ we have $\lambda_1 = 0.64$, $\lambda_2 = 0.11$, $\lambda_3 = 0.03$ ($N = 21$) and $\lambda_1 = 1.01$, $\lambda_2 = 0.30$, $\lambda_3 = 0.11$ ($N = 41$). This motivates the use of the first order, first eigenvalue approach from Chapter 3.3, i.e., based on a Taylor expansion with $r = 1$ eigenvalues to order $s = 1$. In the case $N = 41$, one might consider going to $r = 2$ based on the eigenvalues alone, but we will see in the numerical tests for the Bermudan swaption that even with $r = 1$ the result lies within the Monte Carlo bounds.

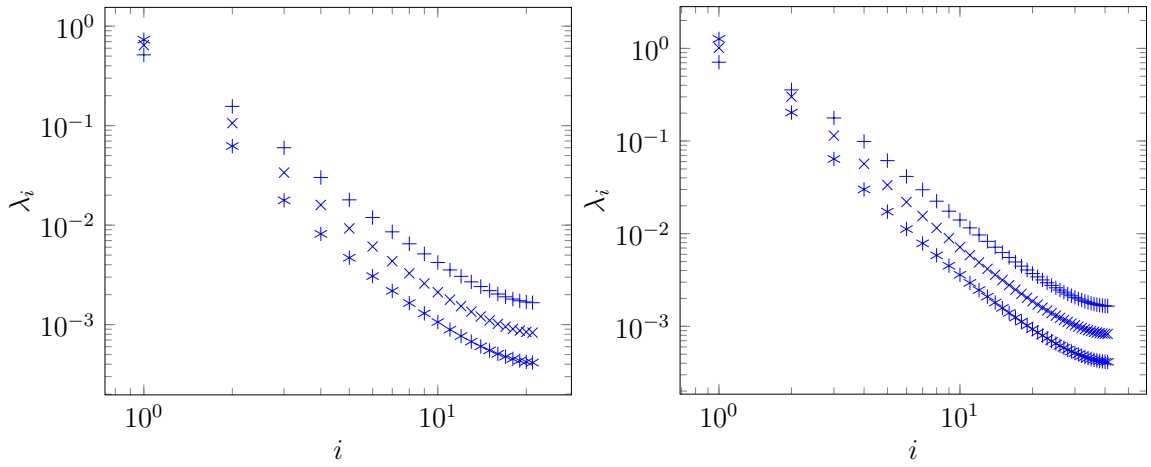


Figure 6.1: Eigenvalues λ_i, $1 \leq i \leq N$, of Σ for $N = 21$ (left) and $N = 41$ (right), constant volatility $c = 0.2$, and $\phi = 0.02065$ $(*)$, $\phi = 0.0413$ $(\times)$ and $\phi = 0.0826$ $(+)$. For each value of N and ϕ the eigenvalues approximately lie on a line with slope -2.

We now choose the terminal bond $P(T_{N+1})$ as numéraire and combine the LIBOR dynamics in equations (6.1)–(6.3) and our covariance structure with the Feyman-Kac PDE to obtain a PDE satisfied by the value function of derivatives on the LIBOR curve. A complication arises in the transformation (2.22) to the heat equation (2.13), as the drift term β_i in (2.23) was assumed to depend only on τ whereas with μ_i as in equation (6.2) it also depends on L_j, $j > i$.

To make the PCA approach directly applicable, we first approximate the drift term. A common approach in practice is to "freeze" the drift at its initial value by setting

$$\mu_i(t) = -\sum_{j=i+1}^{N} \frac{\alpha L_j(0)}{1 + \alpha L_j(0)} \sigma_i \sigma_j \rho_{ij}.$$

This introduces an error in the drift that grows with $L_j(t) - L_j(0)$, and the approximation can be expected to be reasonably accurate for moderate values of T_N and σ. We will confirm this numerically by comparing the PDE results to Monte Carlo estimates with and without drift approximation.

A second point of consideration is that $L_i(t)$ is only financially meaningful for $t \leq T_i$. In order not to have to change the underlying set of arguments of the value function, and hence the PCA, at every tenor time T_i, we consider "extended" LIBORs which

are also defined for $T_i < t \leq T_N$. In the case of constant ρ and σ, a possible extension is obtained by demanding that $L_i(t)$ follows (6.1) for all $0 \leq t \leq T_N$. Note that the exact option value does not depend on $L_i(t)$ for $t > T_i$ and is thus not affected by this extension.

Applying the first order, one-dimensional PCA-ANOVA approach now leads to the approximate solution

$$\begin{aligned} u^{(1,1)}(z,\tau;\lambda) &= u(z,\tau;\lambda^0) + \sum_{i=2}^{N} \left(u(z,\tau;\lambda^0 + \lambda_i e_i) - u(z,\tau;\lambda^0) \right) \\ &= (2-N) \cdot u(z,\tau;\lambda^0) + \sum_{i=2}^{N} u(z,\tau;\lambda^0 + \lambda_i e_i), \end{aligned} \tag{6.4}$$

where $\lambda^0 = (\lambda_1, 0, \ldots, 0)$,

$$\tau = T - t, \quad z = Q^T \ln(L) + \beta(\tau)$$

and

$$\beta_i(\tau) = -\tau \sum_{j=1}^{N} Q_{ji} \left(\frac{\sigma_j^2}{2} - \sum_{k=j+1}^{N} \frac{\alpha L_k(0)}{1 + \alpha L_k(0)} \sigma_j \sigma_k \rho_{jk} \right), \quad 1 \leq i \leq N.$$

Here, Q is an orthogonal matrix of eigenvectors of Σ, and λ the vector of corresponding eigenvalues.

The initial condition for all PDEs is given by $g(L)$ where g is the payoff at time T. The quantity of interest is $u^{(1,1)}(Z(0), T; \lambda)$, where $Z(0) = Q^T \ln(L(0)) + \beta(T)$.

We study two types of derivatives to test the flexibility and accuracy of the approach and benchmark against Monte Carlo results:

- short- to long-running Bermudan swaptions, where the combination of high-dimensionality and early-exercise presents challenges for PDE and MC methods;
- a Ratchet floor, where the path-dependency is conceptually straightforward to include in a MC solver and needs adaptation of the PDE solver.

All prices reported are relative to the bond at time $t = T_1 = 0$, i.e., in units of $P(T_1)$. Details of the implementation can be found in Chapter 7.

6.3 Bermudan Swaptions

A Bermudan (payer) swaption with strike price K can be exercised at any one of a set of exercise dates $\{T_{e_1}, \ldots, T_{e_{N'}}\} \subseteq \{T_1, \ldots, T_N\}$. Here, we consider as an example 3-month LIBOR, i.e., $\alpha = 0.25$ and $T_i = \alpha\,(i-1) = 0.25\,(i-1)$ for $1 \leq i \leq N$, and Bermudan swaptions which can be exercised yearly, i.e., $\{T_{e_1} \ldots, T_{e_{N'}}\} = \{T_1, T_5, T_9, \ldots, T_N\}$, assuming that $N-1$ can be divided by 4. For example, for $N = 41$ the final expiration date is in 10 years.

If the Bermudan swaption is exercised at T_i then the holder receives a (payer) swaption with payout

$$V_{Swaption,i} = \max\left(\alpha \cdot \sum_{j=i}^{N} [L_j(T_i) - K]/P(T_i, T_{j+1}), 0\right).$$

The value $V_{BS,1}$ of a Bermudan swaption is thus determined by backward induction through $V_{BS,N} = V_{Swaption,N}$ and

$$V_{BS,i} = \max(V_{BS,i+4}, V_{Swaption,i}) \text{ at } i \in \{1, 5, \ldots, N-4\}. \tag{6.5}$$

Between T_i and T_{i+4}, $i \in \{1, 5, \ldots, N-4\}$, the value function $V_{BS,i}$ satisfies the Feyman-Kac PDE for the LIBOR market model dynamics (6.1)–(6.3), which we transform and solve approximately with our expansion methods. The interface condition (6.5) can easily be incorporated in the PDE discretisation by evaluating (6.5) on the computational grid.[1]

As reference solutions for the PDE results we use Monte Carlo (MC) estimates. Bermudan options require specialised MC methods, because a straightforward MC solution only gives the option value at one point in time. In general, the numerical approximation of multi-dimensional American and Bermudan options by Monte Carlo methods is an area of active current research. We mention recent work on the computation of tight bounds via iteration approaches (e.g., [29]) and via path-wise

[1] See Chapter 7.2.

Parameter	Value	Description
J	601	Number of grid points in each direction
M_{PDE}	10	Time steps per interval of length α in the PDE computation
N_1	10^6	Number of MC paths for exercise policy learning
N_2	10^7	Number of MC paths to calculate the lower bound
N_{outer}	5000	Number of outer MC paths to calculate the upper bound
N_{inner}	1000	Number of inner MC paths to calculate the upper bound
M_{MC}	5	Time steps per interval of length α in the MC computation

Table 6.1: Numerical parameters for the Bermudan swaption PDE and MC computations.

optimisation (see [9]). Here, we use the well-established and popular primal-dual approach for exercise policy learning due to Andersen and Broadie (see [2]), which provides Monte Carlo estimates for a lower bound V_{MC}^- and upper bound $V_{MC}^- + \Delta_0$ to the true option value.

In the simulations, we used $N_1 = 10^6$ paths for learning the exercise policy (of type 'exercise strategy 1' in the notation of [2]), $N_2 = 10^7$ paths to calculate the lower bound and $N_{outer} = 5000$ and $N_{inner} = 1000$ paths for the outer and inner MC runs to compute the upper bound. For the time discretization of the LMM SDEs we used the log-Euler scheme with $M_{MC} = 5$ time steps per interval of length α. In the tests with localized drift (i.e., lognormal LIBORs), the discretisation is exact for $M \geq 1$. For the PDE we used $J = 601$ grid points in every direction and the Crank-Nicolson scheme with $M_{PDE} = 10$ time steps per time interval of length α. The numerical parameters, summarised in Table 6.1, were chosen such that the numerical error is small compared to the difference between PDE and MC solution and is typically of order 0.1% or less of the derivative value.

Numerical results

Results for Bermudan swaptions at-the-money (ATM, $K = 0.1$) are shown in Table 6.2. The PDE results for frozen drift and the previously introduced second order scheme around $r = 1$ eigenvalues are compared to the values calculated by MC simu-

N	V_{MC}^-	σ	$V_{MC}^- + \Delta_0$	σ_{Δ_0}	$V_{MC}^- + \Delta_0/2$	V_{PDE}	Δ_{abs}	Δ_{rel}
5	1.75E-03	9.02E-07	1.75E-03	0.00E-00	1.75E-03	1.76E-03	1.18E-05	0.68%
11	1.21E-02	5.42E-06	1.22E-02	5.20E-06	1.21E-02	1.24E-02	2.61E-04	2.15%
21	3.05E-02	1.14E-05	3.15E-02	4.65E-05	3.10E-02	3.14E-02	4.03E-04	1.30%
41	6.17E-02	1.94E-05	6.68E-02	1.46E-04	6.42E-02	6.57E-02	1.44E-03	2.24%
61	8.23E-02	2.29E-05	9.10E-02	2.16E-04	8.67E-02	9.04E-02	3.77E-03	4.35%
81	9.45E-02	2.39E-05	1.06E-01	2.69E-04	1.00E-01	1.07E-01	6.84E-03	6.83%
101	1.01E-01	2.38E-05	1.14E-01	3.82E-04	1.08E-01	1.18E-01	1.02E-02	9.45%
5	1.75E-03	9.01E-07	1.75E-03	0.00E+00	1.75E-03		1.33E-05	0.76%
11	1.21E-02	5.43E-06	1.22E-02	5.16E-06	1.21E-02		2.57E-04	2.11%
21	3.06E-02	1.15E-05	3.16E-02	4.66E-05	3.11E-02		3.17E-04	1.02%
41	6.19E-02	1.98E-05	6.77E-02	1.63E-04	6.48E-02		8.62E-04	1.33%
61	8.26E-02	2.35E-05	9.27E-02	2.44E-04	8.77E-02		2.74E-03	3.13%
81	9.49E-02	2.45E-05	1.07E-01	2.85E-04	1.01E-01		6.13E-03	6.07%
101	1.02E-01	2.45E-05	1.16E-01	3.65E-04	1.09E-01		9.23E-02	8.48%

Table 6.2: PDE results for ATM ($K = 0.10$) Bermudan swaptions compared to frozen (top) and full (bottom) drift MC results. V_{MC}^- and σ are the lower MC bound and its estimated standard error. $V_{MC}^- + \Delta_0$ and σ_{Δ_0} are the upper MC bound and the estimated standard error of the MC spread Δ_0. V_{PDE} is the PDE result and the columns Δ_{abs} and Δ_{rel} show the absolute and relative difference to the best MC estimate $V_{MC}^- + \Delta_0/2$.

lation with frozen and full drift, to disentangle the effects of the drift approximation on the one hand and the dimension reduction on the other. The model parameters chosen were $\phi = 0.0413$ and $c = 0.2$, with a flat initial LIBOR curve with $L_i(0) = 0.1$, all identical to [29].

In these tests, the PDE method shows very good accuracy for up to $N = 41$: V_{PDE} is above V_{MC}^- and below the upper MC bound in almost all cases for $N = 21$ or $N = 41$. For lower N it is often slightly higher than the (in these cases fairly tight) upper bound, but the difference Δ_{abs} to the middle value $V_{MC}^- + \Delta_0/2$ – which is considered to be a better estimate for the true price than both the lower or upper bound in [29] – is always less than 2.6 basis points. For $N = 61 - 101$, the PDE values are still close to the MC values. They are above the upper MC bound for $N = 81$ and $N = 101$, though. Taking into account the relatively large difference between lower and upper MC bounds, this seems to indicate that the applicability of the first order, one-dimensional version of the PCA-ANOVA approach reaches its limits (in

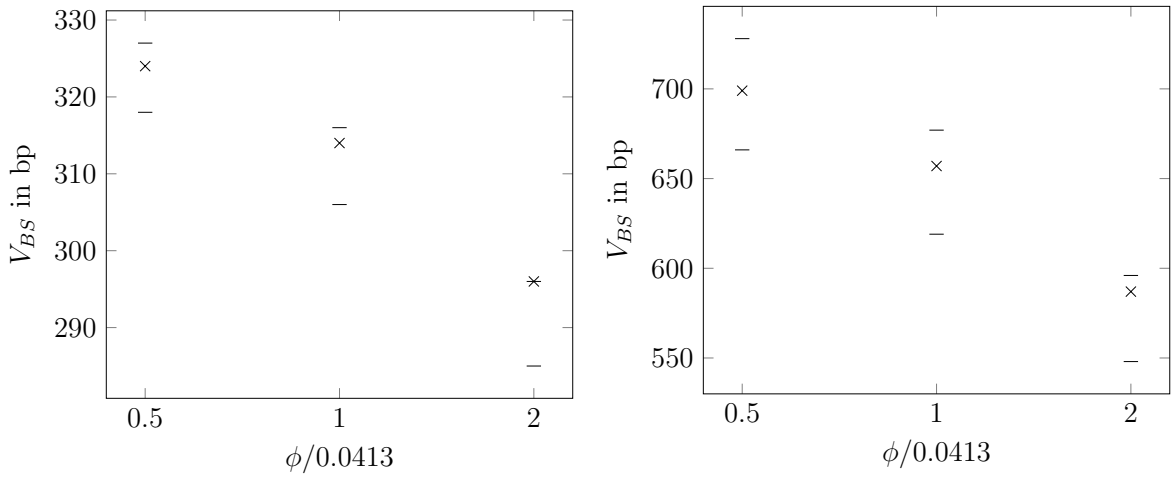


Figure 6.2: PDE ($\times$) and MC ($-$) results for ATM Bermudan swaption values V_{BS} for varying correlation strength ϕ (where $\phi = 0.0413$ is the value used in [29]), given $c = 0.2$ and $N = 21$ (left) and 41 (right). The difference between lower and upper MC bound is about $0.04\,V_{BS}$ (left) and $0.10\,V_{BS}$ (right). The standard deviation for the MC results is $\leq 0.0005\,V_{BS}$ for the lower and $\leq 0.0025\,V_{BS}$ for the upper bound. The corresponding eigenvalues are those shown in Figure 6.1.

this setting) for problems with N higher than $50-60$ (as does the MC approach used here).

In-the-money (ITM, $K = 0.09$) and out-of-the-money (OTM, $K = 0.11$) results are similar and are shown in Tables C.1 and C.2 in Appendix C.1. As the overall value is largest for ITM and lowest for OTM options, the relative difference $\Delta_{rel} = \Delta_{abs}/(V^{-}_{MC}+\Delta_0/2)$ is typically smallest for ITM options and largest for OTM options. To assess the PCA-ANOVA approach under a range of market conditions, we present simulations for ATM Bermudan swaptions with differing parameters: Figures 6.2 and 6.3 show the results for varying correlation and volatility. For stronger correlation and lower volatility, where one would expect the highest accuracy, the PDE solution lies approximately in the middle between the lower and upper MC bound. For weaker correlation and higher volatility it tends towards and reaches the upper MC bound. Table 6.3 shows that the PCA approach also performs well for a lower initial LIBOR curve with $L_i(0) = 0.02$ for all $1 \leq i \leq N$.

Finally, Figure 6.4 shows how the value of the PDE approximation changes when we

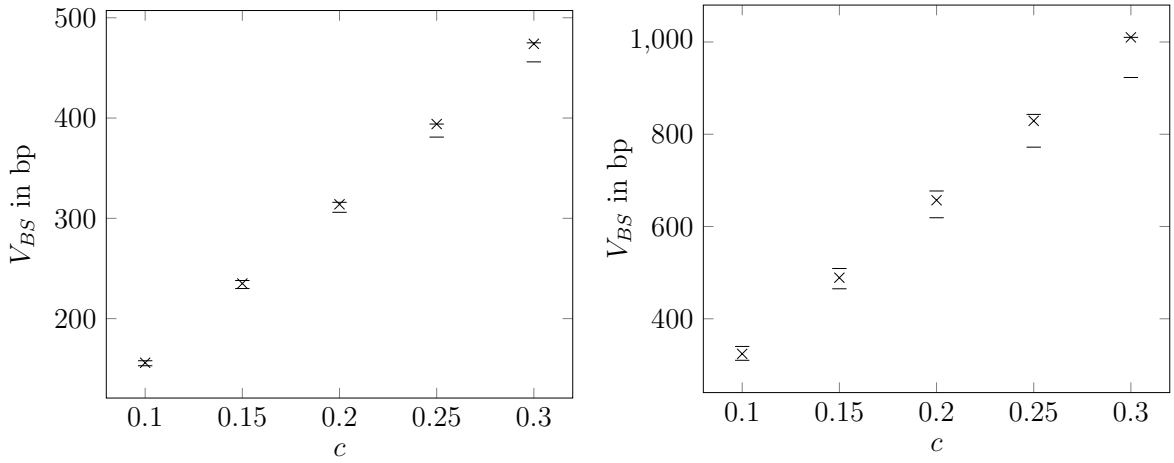
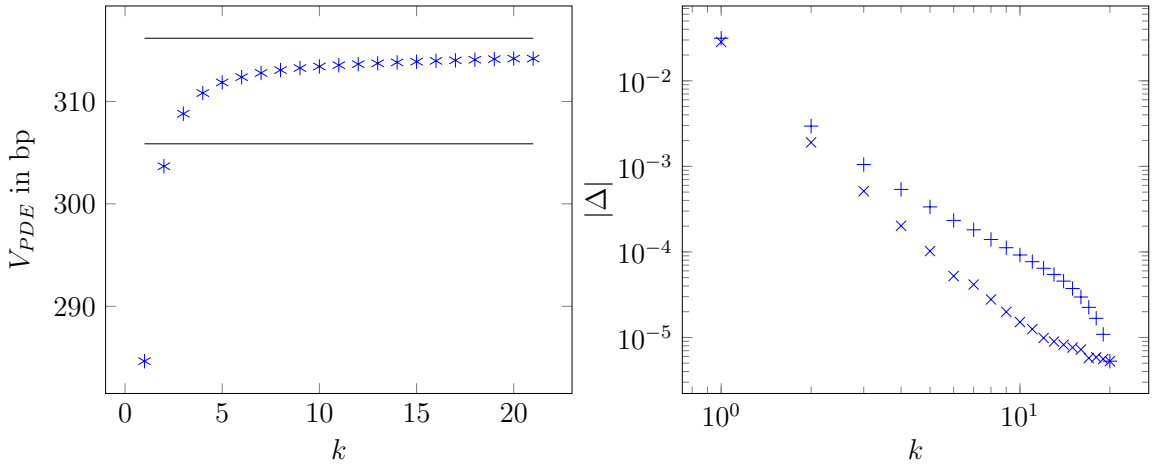


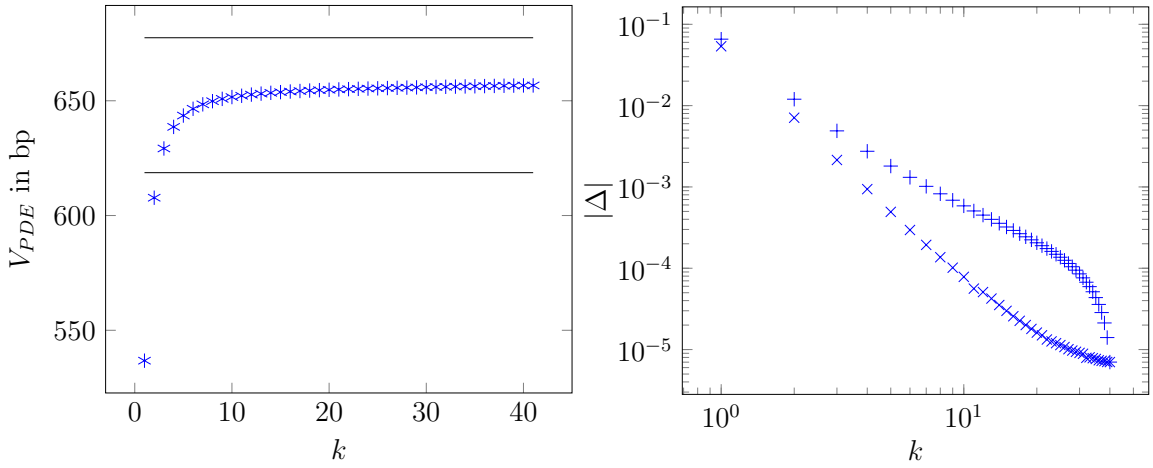
Figure 6.3: PDE ($\times$) and MC ($-$) results for ATM Bermudan swaption values V_{BS} for varying volatility c, given $\phi = 0.0413$ and $N = 21$ (left) and 41 (right). The difference between lower and upper MC bound is about $0.04\,V_{BS}$ (left) and $0.10\,V_{BS}$ (right). The standard deviation for the MC results is $\leq 0.0005\,V_{BS}$ for the lower and $\leq 0.0025\,V_{BS}$ for the upper bound. Note that all eigenvalues are proportional to c^2 and changing c thus does not alter their relative sizes.

N	V_{MC}^-	σ	$V_{MC}^- + \Delta_0$	σ_{Δ_0}	$V_{MC}^- + \Delta_0/2$	V_{PDE}	Δ_{abs}	Δ_{rel}
5	3.88E-04	2.02E-07	3.88E-04	0.00E+00	3.88E-04	3.89E-04	1.04E-06	0.27%
11	2.86E-03	1.31E-06	2.87E-03	1.37E-06	2.87E-03	2.90E-03	3.78E-05	1.32%
21	8.03E-03	3.19E-06	8.39E-03	1.51E-05	8.21E-03	8.17E-03	-3.72E-05	-0.45%
41	2.00E-02	7.34E-06	2.32E-02	7.56E-05	2.16E-02	2.08E-02	-7.83E-04	-3.63%
5	3.88E-04	2.02E-07	3.88E-04	0.00E+00	3.88E-04		6.45E-07	0.17%
11	2.86E-03	1.31E-06	2.88E-03	1.46E-06	2.87E-03		3.74E-05	1.30%
21	8.04E-03	3.18E-06	8.37E-03	1.46E-05	8.21E-03		-3.31E-05	-0.40%
41	2.01E-02	7.36E-06	2.30E-02	7.41E-05	2.15E-02		-7.03E-04	-3.27%

Table 6.3: PDE results for ATM ($K = 0.02$) Bermudan swaptions in a flat initial LIBOR curve setting with $L_i(0) = 0.02\ \forall 1 \leq i \leq N$ compared to frozen (top) and full (bottom) drift MC results.



k	1	2	3	6	11	21
V_{PDE}	285 bp	304 bp	309 bp	312 bp	314 bp	314 bp
Δ	-9.40%	-3.34%	-1.71%	-0.58%	-0.20%	0.00%



k	1	2	3	6	11	21	41
V_{PDE}	537 bp	608 bp	629 bp	647 bp	652 bp	655 bp	657 bp
Δ	-18.26%	-7.45%	-4.18%	-1.55%	-0.69%	-0.27%	0.00%

Figure 6.4: Approximations V_{PDE} to the value of the ATM Bermudan swaption with $N = 21$ (top) and $N = 41$ (bottom) when including only the terms up to $i = k$ in equation (6.4) (left). The solid horizontal lines are the lower and upper MC bounds. The graphs on the right show the terms in (6.4) for each individual i ($\times$) and the sum of the remaining terms $i = k + 1, \ldots, N$ ($+$). The tables give numerical values, with the row Δ showing the relative difference to the PDE result with all N terms.

consider in (6.4) only the 1-dimensional PDE solution and the 2-dimensional PDE solutions associated with the k largest eigenvalues. Specifically, the case $k = 1$ is the one-dimensional approximation using the first principal component, and the case $k = 2$ the standard two-dimensional PCA approximation. The contributions for different eigenvalues approximately lie on a line with slope -2 in the log-log-plots in Figure 6.4, just as the eigenvalues in Figure 6.1, i.e., they show the same decay. Since the i-th contribution is equal to $\lambda_i \frac{\partial u}{\partial \lambda_i}$, this suggests that the partial derivatives $\frac{\partial u}{\partial \lambda_i}$ are all of similar size.

Evidently, it is in fact necessary to include the contributions from several of the largest eigenvalues to compute an accurate solution. At the same time, the solution levels out after including about 10 dimension. This is in line with the decay of the eigenvalues and the fact that the payoff in this case is almost parallel to the eigenvector of the first dimension. For models where the eigenvalues of the covariance matrix decay fast enough – which includes many models in mathematical finance – the PCA-ANOVA PDE approach might be used with a fixed number k for a wide range of values N. This further reduces the computational effort – which in the given case is roughly proportional to k – without sacrificing significant accuracy.

Run times

While considerable effort went into the efficient implementation of both the PDE and MC methods, there is still room for performance improvement, for instance on the algorithmic level and in the optimal trade-offs between numerical parameters. Thus we only want to comment on approximate run times: For the PDE calculations, the computation times for $N = 5, 11, 21$ and 41 were on the order of 10, 60, 240 and 1000 seconds, resp., using MATLAB on a AMD Phenom(tm) II X4 925 Processor (2.8 GHz) with 3.8 GB RAM. The run time is roughly quadratic in N, because the number of PDE solutions required to evaluate (6.4) is N, and the expiry is $T = N/4$,

so the number of (Crank-Nicolson) time steps with fixed step size, for a given PDE, is proportional to N.

For the MC simulations the corresponding computation times were of order $250+65$, $1100+600$, $4000+5200$ and $17000+44000$ seconds, resp., for frozen drift and $250+65$, $1200+700$, $4500+5700$ and $20000+54000$, resp., for full drift. Here the first number is the computation time for the lower bound and the second number is the additional time necessary to compute the upper bound. The run time is roughly quadratic in N for the lower bound, since N processes have to be simulated over N tenor dates, and roughly cubic in N for the upper bound.

Despite the approximate nature of these computation times, it becomes clear that the PDE method is not only competitive time-wise, but indeed faster by a factor of $25-70$ in our implementation. To get an optimal allocation of computational resources, one could also try to further optimise the relative size of the numerical parameters, such that the computation time is optimal for a given size of the combined discretisation error. By, say, halving the mesh size in the two directions of the computational grid one quadruples the computational time, while there is no practical accuracy gain in bringing the discretisation error substantially below the error of the dimension reduction. A similar statement is true for the Monte Carlo estimators. A precise comparison of efficiency is therefore delicate.

Both the PDE and MC methods used in this section are limited in their accuracy: the MC method uses a class of exercise strategies which generally does not include the optimal one; the PDE method employs a drift approximation and asymptotic expansion. For a wide range of N, the errors are comparable. Possible accuracy improvements on the basis of the decomposition are outlined in Chapters 2.7 and 8.2, and will be tested on the following application.

6.4 Ratchet Floors

A Ratchet floor with strike price K_1 and parameters a, b, c is a portfolio (sum) of floorlets with payouts $\max\{K_i - L_i(T_i), 0\}$ at T_{i+1}, where the strike prices K_i are recursively determined from the given initial strike K_1 by

$$K_i = \max\left(aL_{i-1}(T_{i-1}) + bK_{i-1} + c, 0\right), \quad i > 1, \tag{6.6}$$

see [35]. The (relative) price of the Ratchet floor for $t = 0$ is given by

$$V_{RF}(0) = \sum_{i=1}^{N} E\left[V_{Fl,i}(T_i)/\mathcal{N}(T_i)|\mathcal{F}_0\right],$$

where $V_{Fl,i}$ is the value of the i-th floorlet. Due to the linearity in the sum on the right-hand side of this equation it is sufficient to be able to calculate the price of a single floorlet. Without loss of generality we will thus focus on $V_{Fl,N}$.

The ratchet feature (6.6) makes the problem high-dimensional and strongly path-dependent as the payoff depends on the values of all L_i at different points in time. To solve the problem by a backward equation, we need to make it Markovian by including the strike dynamics with the evolution of the LIBORs, and to specify the value function as a function of all the above. As the strike changes discretely in time, the value function satisfies the standard LMM PDE between the T_i. At each tenor time T_i, the jump condition

$$V_{Fl,i}(T_i-, K, L_1, \ldots, L_N) = V_{Fl,i}(T_i+, \max\left(aL_{i-1} + bK + c, 0\right), L_1, \ldots, L_N) \tag{6.7}$$

holds. Here, T_i+ denotes the limit coming from larger t where the solution is already computed, and we use this to compute the solution just prior to T_i before the strike is updated. Details of the complete PDE model and its mathematical analysis can be found in [35].

We approximate the solution on a grid in K-direction, and compute the updated solution for each grid point via cubic spline interpolation for the corresponding value

of K_{i+1}. This adds an extra dimension $N+1$ to the problem, and effectively the one-dimensional ANOVA terms (corresponding to z_1) now live on a two-dimensional grid, and the two-dimensional ANOVA terms (corresponding to z_1 and an additional z_i) on a three-dimensional grid, formulaically (cf. Chapter 2.7),

$$u^{(2,1)}(K,z) = u_0^{(2)}(K_1, Z(0); K, z_1) + \sum_{i=2}^{N} u_i^{(2)}(K_1, Z(0); K, z_1; z_i),$$

where the superscript '2' stands for the variables K and z_1 and the '1' for the extra coordinate z_i in the expansion.

It is conceivable to use an expansion in direction K with anchor K_1 of the form (2.30), which becomes

$$\begin{aligned} u^{(1,1)}(K,z) = u_0^{(1)}(K_1, Z(0); z_1) + \sum_{i=2}^{N} u_i^{(1)}(K_1, Z(0); z_1; z_i) \\ + u_{N+1}^{(1)}(K_1, Z(0); z_1; K), \end{aligned}$$

where the first superscript '1' stands for z_1 and the second '1' for the expansion in one coordinate z_i or K. In other words, we identify the path-dependent quantity K defined in (6.6) with an extra dimension z_{N+1} in (2.30) with N replaced by $N+1$. We do not pursue this further here.

Due to the smoothness of the solution in the K-direction and the higher order of the spline interpolation relative to the finite difference scheme in L directions, a relatively coarse mesh on the K axis is sufficient. Specifically, we use 21, 41 and 21 spline nodes for the intervals $[0, 0.05]$, $[0.05, 0.15]$ and $[0.15, K_{max}]$, resp., with $K_{max} = 0.5$, to give a total of 81 nodes on the K-axis. In comparison, we use $J = 401$ grid points in the z_i-directions and the Crank-Nicolson scheme with $M = 10$ time steps per interval of length $\alpha = 0.25$ between tenors. The numerical parameters are again chosen such that the numerical error is small compared to the difference between PDE and MC solution and is typically around 0.1% or less of the derivative value. The model parameters chosen were $\phi = 0.0413$ and $c = 0.2$, with a flat initial LIBOR curve with $L_i(0) = 0.1$, identical to those for the Bermudan swaptions in Table 6.2.

Parameter	Value	Description
J	401	Number of grid points in each LIBOR direction
I	81	Number of grid points in strike price direction
M_{PDE}	10	Time steps per interval of length α in the PDE computation
N_1	10^7	Number of MC paths
M_{MC}	10	Time steps per per interval of length α in the MC computation

Table 6.4: Numerical parameters for the Ratchet floor PDE and MC computations.

We again use MC estimates as reference solutions. For a path-dependent option without early-exercise features like the Ratchet floor we can use a straightforward MC calculation. In the tests, we sampled $N_1 = 10^7$ paths; for the time discretization, we used the log-Euler scheme with $M = 10$ time steps per interval of length α.

Numerical results

The numerical results for the Ratchet floors are shown in Table 6.5 for $N = 5, 11, 21$. We consider three different configurations (a, b, c) and three different strike prices $K = 0.1, 0.11, 0.09$ (ATM, ITM and OTM, resp.). The absolute difference between the PDE and MC solution is never more than 1.14 basis points and the relative difference is below 1% in all but one case.

Run times

We again report approximate computation times with the same caution as in the previous section. For $N = 5, 11$ and 21 the PDE run times were of order 400, 2400 and 10000 seconds, resp., on a AMD Phenom(tm) II X4 925 Processor (2.8 GHz) with 3.8 GB RAM. For frozen drift the corresponding MC run times were of order 160, 900 and 3500 seconds, while for full drift they werde 190, 1100 and 4600 seconds. The MC computation was faster by a factor of 2 to 3 compared to the PDE computation. Both were roughly quadratic in N. The MC simulation also permits the computation of values for multiple parameters (a, b, c) in parallel, with only a small increase in computation time. Given the fast decay of the correction terms in the ANOVA

(a, b, c)	N	V_{MC}	σ	V_{PDE}	Δ_{abs}	Δ_{rel}
$K_1 = 0.10$						
0/1/0	5	7.08E-003	2.95E-006	7.04E-003	-3.39E-05	-0.48%
	11	9.65E-003	3.87E-006	9.61E-003	-3.87E-05	-0.40%
	21	1.07E-002	4.20E-006	1.07E-002	2.98E-05	0.28%
0.2/0.9/0	5	3.06E-002	4.23E-006	3.06E-002	-6.30E-05	-0.21%
	11	4.94E-002	4.74E-006	4.93E-002	-9.12E-05	-0.18%
	21	5.10E-002	5.02E-006	5.09E-002	-3.33E-05	-0.07%
0.25/0.95/-0.01	5	3.29E-002	4.03E-006	3.29E-002	-5.74E-05	-0.17%
	11	6.06E-002	5.13E-006	6.06E-002	-1.26E-05	-0.02%
	21	7.36E-002	8.13E-006	7.36E-002	-1.13E-05	-0.02%
$K_1 = 0.11$						
0/1/0	5	1.27E-002	3.88E-006	1.26E-002	-5.34E-05	-0.42%
	11	1.44E-002	4.74E-006	1.43E-002	-7.10E-05	-0.49%
	21	1.44E-002	4.94E-006	1.44E-002	-5.28E-05	-0.37%
0.2/0.9/0	5	3.63E-002	4.33E-006	3.63E-002	-6.40E-05	-0.18%
	11	5.20E-002	4.76E-006	5.20E-002	-9.04E-05	-0.17%
	21	5.17E-002	5.02E-006	5.17E-002	-3.94E-05	-0.08%
0.25/0.95/-0.01	5	4.00E-002	4.11E-006	4.00E-002	-5.81E-05	-0.15%
	11	6.52E-002	5.12E-006	6.52E-002	-1.75E-05	-0.03%
	21	7.58E-002	8.08E-006	7.57E-002	-3.69E-05	-0.05%
$K_1 = 0.09$						
0/1/0	5	3.20E-003	1.94E-006	3.17E-003	-2.32E-05	-0.73%
	11	5.83E-003	2.95E-006	5.82E-003	-1.56E-05	-0.27%
	21	7.39E-003	3.42E-006	7.47E-003	7.68E-05	1.04%
0.2/0.9/0	5	2.51E-002	4.07E-006	2.50E-002	-7.27E-05	-0.29%
	11	4.68E-002	4.71E-006	4.67E-002	-1.02E-04	-0.22%
	21	5.03E-002	5.02E-006	5.02E-002	-4.86E-05	-0.10%
0.25/0.95/-0.01	5	2.59E-002	3.88E-006	2.58E-002	-6.73E-05	-0.26%
	11	5.61E-002	5.14E-006	5.60E-002	-2.07E-05	-0.04%
	21	7.15E-002	8.19E-006	7.15E-002	-1.61E-05	-0.02%

Table 6.5: Frozen drift PDE results for Ratchet floors compared to full drift MC results. The column V_{PDE} shows the computed PDE value. Columns Δ_{abs} and Δ_{rel} show the absolute and relative difference to the MC estimate V_{MC}. PDE results compared to frozen drift MC results are shown in Table C.3 in the Appendix.

decomposition, as illustrated in Figure 6.4 for the Bermudan swaption, it would be possible to compute only the first $k \ll N$ ANOVA terms without significant loss of accuracy, which brings the computational times for the PDE in the range of, or below, the MC ones.

Higher Order Taylor terms

In Chapters 3.1 and 2.7 we explained how to include higher order terms in the Taylor and ANOVA expansions. This extension is expected to decrease the size of the error to correspondingly higher orders of lambda. In this section, we present experimental results demonstrating empirically the validity of this assertion. We also investigate the behaviour of the approximation when including fully the first r eigenvalues, i.e., expanding around

$$\lambda^0 = (\lambda_1, \ldots, \lambda_r, 0, \ldots).$$

Previous investigations in [41] and [42] have considered examples with equity baskets, for $(r, s) = (1, 1)$, $(1, 2)$ and $(2, 1)$, where s is the order of the Taylor expansions. See also the discussion at the end of Chapter 2.7. Here, we systematically explore the impact of varying r and s.

Table 6.6 shows estimates for the higher order terms in (3.7) for a Ratchet floor with $(a, b, c) = (0.2, 0.9, 0)$ and $N = 11$. Precisely, for an $N - r$ dimensional multi-index $\omega = (\omega_1, \ldots, \omega_{N-r})$, we study, for different values of r, s and t,

$$\begin{aligned} V_{rst} &= V_{rst}(z_0, 0; \lambda, \lambda^0) \qquad (6.8) \\ &= \sum_{|\omega|=s} \Delta^{r,\omega,t}(u; z_0, 0; \lambda, \lambda^0) \\ &= \sum_{|\omega|=s} \frac{\lambda_{r+1}^{\omega_1} \cdot \ldots \cdot \lambda_N^{\omega_{N-r}}}{\omega!} \frac{\partial^{|\omega|} u}{\partial \lambda_{r+1}^{\omega_1} \ldots \partial \lambda_N^{\omega_r}}(z_0, 0; \lambda^0) \\ &\quad + O(\|\lambda^0 - \lambda\|^{t+1}), \qquad (6.9) \end{aligned}$$

	V	σ		V	σ
V_{100}	498.902	0.002	V_{222}	0.111	0.002
V_{111}	-5.079	0.002	V_{300}	494.610	0.010
V_{112}	-5.208	0.003	V_{311}	-0.528	0.004
V_{113}	-5.179	0.003	V_{322}	0.033	0.001
V_{122}	0.250	0.002	V_{400}	494.233	0.010
V_{123}	0.214	0.003	V_{411}	-0.129	0.003
V_{133}	0.109	0.010	V_{500}	494.071	0.010
V_{200}	495.657	0.009	V_{511}	0.041	0.002
V_{211}	-1.644	0.006	V_{full}	494.107	0.003

Table 6.6: Numerical results (in bp) for the different terms V_{rst} as in (6.8), for a Ratchet floor with $(a, b, c) = (0.2, 0.9, 0)$ and $N = 11$. r is the number of fully included eigenvalues, s is the order of the Taylor expansion and t is the order of the finite difference stencils used to compute the derivatives. The computations used $2 \cdot 10^9$ MC paths for V_{100}, V_{111}, V_{112}, V_{113} and V_{full} and $2 \cdot 10^8$ MC paths for all other values, where σ is the root-mean-squared error.

$u(3\lambda)$	$u(2\lambda)$	$u(\lambda)$	$u(0)$	Taylor term	$t+1$
			$^1/_1$	$u(0)$	$-$
		$^1/_1$	$-^1/_1$	$\lambda u'(0)$	2
	$-^1/_2$	$^4/_2$	$-^3/_2$	$\lambda u'(0)$	3
$^2/_6$	$-^9/_6$	$^{18}/_6$	$-^{11}/_6$	$\lambda u'(0)$	4
	$^1/_2$	$-^2/_2$	$^1/_2$	$\lambda^2 u''(0)/2$	3
$-^1/_2$	$^4/_2$	$-^5/_2$	$^2/_2$	$\lambda^2 u''(0)/2$	4
$^1/_6$	$-^3/_6$	$^3/_6$	$-^1/_6$	$\lambda^3 u'''(0)/6$	4

Table 6.7: Coefficients of the finite difference stencils used to approximate the Taylor expansion terms of a function $u(\lambda)$ at 0, where $t+1$ is the order of the approximation, see (6.9). Multi-dimensional stencils for partial derivatives were constructed by straightforward multiplication of the single-dimensional stencils.

where $\Delta^{r,\omega,t}(u; z_0, 0; \lambda, \lambda^0)$ is an approximation of order $t+1$ to the relevant mixed partial derivative term of u at $(z_0, 0)$ and λ^0 (see Chapter 3.1). We have chosen finite difference stencils with weights as shown in Table 6.7; for instance, the stencil in the second line is the standard right-sided difference and the one in the fifth line a standard second difference shifted to the right. Each of these finite differences requires the solution of $r+s$-dimensional problems. This computation was performed with a Monte Carlo method for illustration purposes and we comment on this at the end of this section.

$r \setminus s$	0	1	2	3	0	1	2	3
1	4.795	-0.284	-0.034	0.075	0.970%	-0.058%	-0.007%	0.015%
2	1.550	-0.093	0.018		0.314%	-0.019%	0.004%	
3	0.502	-0.026	0.007		0.102%	-0.005%	0.001%	
4	0.126	-0.003			0.025%	-0.001%		
5	-0.036	0.005			-0.007%	0.001%		

Table 6.8: Absolute (left, in bp) and relative (right) difference between the PDE expansion approximation values $\sum_{i=0}^{s} V_{rii}$ and the full solution V_{full} for different numbers r of fully included eigenvalues and orders s of the Taylor expansion. The numbers are based on the results in Table 6.6. The standard deviation of these estimates is 0.01 bp and 0.002%, resp.

The magnitude of the V_{rst} decreases significantly with increasing order s, as one would expect from the Taylor expansion. Likewise, the correction terms for higher r are much smaller than for lower r. For example, V_{311} is only about one tenth the size of V_{111}. This is again in line with the theoretical prediction, because V_{300} should be closer to the full solution than V_{100}.

To analyse how the accuracy of the overall approximation changes with increasing r and s, Table 6.8 shows the difference between the sum of all terms up to order s and the full solution, i.e.,

$$\Delta_{abs}^{r,s} = \sum_{i=0}^{s} V_{rii} - V_{full}$$

and

$$\Delta_{rel}^{r,s} = \frac{\sum_{i=0}^{s} V_{rii} - V_{full}}{V_{full}},$$

for a range of values of r and s. Again, the error decreases rapidly with increasing r and s. The absolute error is well below 1 bp for all first order cases and well below 0.1 bp for all second order cases. In general, the results suggest that it is possible to make the error negligibly small even with low to moderate values of r and s. The computational effort is dominated by the term with highest $r+s$, which requires the computation of $O((N-r)^s)$ $r+s$-dimensional partial derivatives with stencils of order s.

A notable outlier is the case $(r, s) = (1, 3)$, which is somewhat less accurate than for $(r, s) = (1, 2)$. One possible explanation is that the stencils used for the 1st and 2nd order Taylor terms have an error of 2nd and 3rd order. This could introduce errors which become larger than the Taylor expansion error. We have therefore recomputed the results for $(r, s) = (1, 2)$ and $(r, s) = (1, 3)$ using stencils that are accurate to 3rd and 4th order, resp. The resulting errors are -0.163 bp / -0.033% and -0.061 bp / -0.012%. The accuracy decreased for $s = 2$ and increased slightly for $s = 3$. However, for $r = 1$ all approximations for $s = 1, 2, 3$ are now all below the full value and seem to converge towards it with increasing s.

We also note that some of the computed terms for large $r + s$ are smaller than their standard deviation of around 0.002. Although the leading digit may not be significant, the point stands that these terms are small.

A general comment is due on the computation of the values in Tables 6.6 and 6.8. For demonstration purposes, instead of solving PDEs as previously, we have written the solutions as expectations (in the obvious way using the Feynman-Kac theorem) and estimated them by Monte Carlo simulation. For each triple of r, s and t values we have used the same Brownian paths for all the terms on the right-hand side of (6.8). This has a two-fold benefit. Firstly, it reduces the variance of the estimator for these terms, by a similar mechanism as for standard finite difference sensitivies (see [16]). Note that a large number of these finite differences are to be computed here, in particular $N - r$ choose s terms of highest order, and these normally have mixed signs. Secondly, recycling the normally distributed samples reduces the computation time.

While this gives us a computationally convenient way to illustrate the behaviour of these terms, this is not how one would solve this problem in practice, since a direct simulation of the full problem is possible in this case. The full benefit of the expansion method is realised in cases where accurate MC solutions are not or not easily available,

but accurate PDE solutions to low-dimensional approximations are feasible — such as the Bermudan swaption investigated earlier. (But since a sufficiently accurate and reliable alternative solution to the full problem is not available in those cases, we have no benchmark to compare against and therefore omitted these computations for the purposes of this study.)

We also anticipate that there are advantages in the use of hybrid methods. In these, the lower order terms can be computed very precisely with PDE methods and the higher order terms, which are much smaller in size and often show a corresponding decline in their variance, are computed with reduced relative accuracy via MC simulation. Due to the decreased requirements on the relative accuracy, the latter may use a comparatively low number of MC paths.

Additionally, MC simulation can be used for those higher-dimensional terms in cases where it only provides a crude approximation, such as the lower bound for the Bermudan Swaptions in the previous section. For example, if the PDE based lower order terms are already correct within 0.1% and the higher order MC based terms that provide a correction of that size are only accurate to within 10%, their inclusion will still reduce the overall error by an order of magnitude.

6.5 Theory: Path-Dependency and Early-Exercise

Our practical motivation for developing expansion methods is the pricing of complex, multi-asset financial derivatives such as the ones considered in this chapter. We have previously explained how this problem - which can be formulated as the calculation of an expectation under a market-dependent probability measure - is connected to solving high-dimensional PDEs in Chapters 1.1 and 2.6.

This approach works for financial derivatives that depend on the value of the underlyings at maturity only and have at most a European-style optionality. Many derivatives traded in practice are more complex. They incorporate path-dependency

and early-exercise features, such as Bermudan (exercise permissible at a fixed, finite number of dates prior to maturity) and American (exercise permissible at any time) option styles. It is also possible to combine both, e.g. by making earlier exercise dependent on the value of assets at specified times.

As explained in the previous sections, these additional features can often be modelled by introducing either additional quantities to capture the path-dependency — for example for Asian options, which depend on the the average value of assets — or through additional conditions at the exercise times. These conditions compare the value of the unexercised option with that of the exercised contract. The optimal strategy — and thus the fair price of the financial derivative — is determined by whichever value is larger.

Simulation-based approaches such as Monte Carlo methods struggle with early-exercise options because of the need to calculate the option value at intermediate times. In contrast, PDE methods are particularly well suited for this task, as the value is inherently known at all times and asset values. This advantage is inherited by PDE-based expansion methods, as we were able to demonstrate experimentally in the preceding sections. However, it is unclear whether the theoretical justification and error analysis are still valid. In particular, it remains to be shown that the partial derivatives of the value functions w.r.t. the eigenvalues still exist.

We have shown in Chapter 3 that under some technical conditions piece-wise smoothness of the payoff is in essence sufficient for convergence, with possible problems (or slower convergence) only at the kinks and there only in degenerate cases. For Bermudan and path-dependent cases, some additional complexity arises due to interval conditions. However, we will see that for typical such conditions, piece-wise smoothness holds and the technical conditions are preserved.

We first sketch a heuristic argument in support of that last statement [41]: Consider a first order method expanded around the first eigenvalue. Virtually all practically

relevant payoffs are piecewise smooth with "kinks" (i.e., gradient discontinuities), or "jumps" (i.e., discontinuities), along curves, surfaces etc. From the discussion before it is clear that the derivatives with respect to λ_i, $i = 2, \ldots, N$ are related to the spatial derivatives of the payoff. We allow for λ-dependence of the payoff in preparation for the multi-period case.

These λ-derivatives thus exist at $\lambda_2 = \ldots = \lambda_N = 0$ as long as the diffusion in the direction of the first coordinate – we leave $\lambda_1 > 0$ fixed and thus the convolution with the heat kernel in direction z_1 remains – provides smoothing, precisely, if the curve, surface, etc. which describes the location of the kink is not locally parallel to the first coordinate axis. Even if this were to happen, it would only happen for isolated coordinate values. We showed in the examples and conjecture for the general case that the leading order error term would then not be $\|\lambda - \lambda_0\|^2$, but $\|\lambda - \lambda_0\|$, and that this only appears at isolated spatial coordinates.

For more complex derivatives such as the ones studied in this Chapter, the diffusion equation holds piecewise in time intervals (T_i, T_{i+1}), while at T_i interface conditions hold. The coefficients in the Taylor expansion are now determined by a recursion over i. The interface conditions, such as Equations (6.5) and (6.7), generate value functions at T_i which are again piecewise smooth – with kinks – and serve as terminal conditions for the preceding time interval. The above ideas apply recursively.

We need to consider two things that could go wrong: First, the interface conditions could create non-smoothness beyond what is permissible for the partial derivative to exist. Second, even if the partial derivatives exist, they could accumulate factors of $1/\lambda_i$ or $1/\sqrt{\lambda_i}$ through the subsequent interface conditions and thus grow fast with the number of time intervals.

To analyse this mathematically – without attempting to give a fully rigorous proof here – consider a finite set of tenor times $0 < T_1 < \ldots < T_n < T$ and functionals $h_1, \ldots, h_n$ such that at $t = T_i$ the derivative's value function u gets updated according

to

$$\begin{aligned} u(z, T_i, \lambda) &= h_i(z, \lambda, \lim_{t \nearrow T_i} u(\cdot, t, \cdot)) \\ &= h_i(z, \lambda, \tilde{u}_i(\cdot, \cdot)), \end{aligned}$$

where we have set $\tilde{u}_i(\cdot, \cdot) = \lim_{t \nearrow T_i} u(\cdot, t, \cdot)$ for ease of notation. This is an interface condition (IC) characterised by the functional h_i. Essentially, the PDE problem separates into n distinct PDE problems for the time intervals between two T_i, where the initial condition for the i+1-th problem depends on the final time value of the i-th. These interface conditions are able to model many path-dependent or early-exercise features, such as the ones in the Bermudan swaptions and Ratchet floors investigated in Chapter 6.

To simplify matters we will limit the functional form of h_i to

$$h_i(z, \lambda, \tilde{u}_i(\cdot, \cdot)) = h_i(z, \lambda, \tilde{u}_i(f(z), \lambda)), \tag{6.10}$$

where $f : \mathbb{R}^N \to \mathbb{R}^N$ smooth function but at isolated kinks[2]. This is sufficiently general for all the applications we are interested. See for example Equation (6.16) for the Bermudan swaption and Equation (6.18) for the Ratchet floor. Dealing with functions rather than general functionals simplifies the following calculations considerably.

Equation (6.10) allows us to write $\tilde{u}_{i+1}$ as

$$\begin{aligned} &\tilde{u}_{i+1}(z, \lambda) \\ &= u(z, \lim_{t \nearrow T_{i+1}} t, \lambda) \\ &= \int_{\mathbb{R}^N} \phi(x, T_{i+1} - T_i, 1) u(z - \sqrt{\lambda} \cdot x, T_i, \lambda) dx \\ &= \int_{\mathbb{R}^N} \phi(x, T_{i+1} - T_i, 1) h_i(z - \sqrt{\lambda} \cdot x, \lambda, \tilde{u}_i(f(z - \sqrt{\lambda} \cdot x), \lambda)) dx. \end{aligned}$$

[2]Essentially, the properties of f and h_i need to assure that $u(z, T_i, \lambda)$ satisfies the conditions discussed in Chapter 3.2.

To calculate $\frac{\partial}{\partial\lambda_j}u(z,\lim_{t\nearrow T_{i+1}}t,\lambda)$, we thus have to consider

$$\begin{aligned}&\frac{\partial}{\partial\lambda_j}\left[h_i(z,\lambda,\tilde{u}_i(f(z),\lambda))\right]\\&=\frac{\partial h_i}{\partial\lambda_j}(z,\lambda,\tilde{u}_i(f(z),\lambda))+\frac{\partial h_i}{\partial\tilde{u}_i}(z,\lambda,\tilde{u}_i(f(z),\lambda))\frac{\partial\tilde{u}_i}{\partial\lambda_j}(f(z),\lambda)\end{aligned}\tag{6.11}$$

and for the convolution

$$\frac{\partial}{\partial\lambda_j}\left[h_i(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x),\lambda))\right]\tag{6.12}$$

$$\begin{aligned}&=\frac{-x_j}{2\sqrt{\lambda_j}}\left[\frac{\partial h_i}{\partial z_j}(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x),\lambda))+\frac{\partial h_i}{\partial\tilde{u}_i}(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x),\lambda))\right.\\&\left.\cdot\frac{\partial\tilde{u}_i}{\partial f}(f(z-\sqrt{\lambda}\cdot x),\lambda)\frac{\partial f}{\partial z}(z-\sqrt{\lambda}\cdot x)\right]\end{aligned}\tag{6.13}$$

$$+\ \frac{\partial h_i}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x),\lambda))\tag{6.14}$$

$$+\ \frac{\partial h_i}{\partial\tilde{u}_i}(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x),\lambda))\frac{\partial\tilde{u}_i}{\partial\lambda_j}(f(z-\sqrt{\lambda}\cdot x),\lambda)\tag{6.15}$$

The first term (6.13) is of a similar form as before, see for example the proof of Lemma 32. We can use this to derive sufficient conditions for the functions in the square brackets.

The second term (6.14) did not appear before, because the initial condition did not depend on λ_j. There is no divergent $\lambda_j^{-1/2}$-factor and so the restrictions on $\frac{\partial h_i}{\partial\lambda_j}$ are less severe than those on $\frac{\partial h_i}{\partial z_j}$. It is sufficient that

$$\lim_{\lambda_j\to 0}\int_{\mathbb{R}^N}\phi(x,T_{i+1}-T_i,1)\frac{\partial h_i}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda,\tilde{u}_i(f(z-\sqrt{\lambda}\cdot x))dx$$

exists. Similar reasoning holds for the third term (6.15).

Let us now explicitly consider a Bermudan option:

$$h_i(z,\lambda,\tilde{u}_i(f(z),\lambda))=\max\left(\tilde{u}_i(z,\lambda),p_i(z,\lambda)\right)\tag{6.16}$$

where $p_i:\mathbb{R}^N\times\mathbb{R}^N\to\mathbb{R}$ is a piecewise C^2-function that satisfies the growth condition in Definition 18. f is the identity and h_i is the maximum between $\tilde{u}_i$ and the second

function p_i at the given point. Equation (6.11) then turns into

$$\begin{aligned}
&\frac{\partial}{\partial\lambda_j}\left[\max\left(\tilde{u}_i(z,\lambda),p_i(z,\lambda)\right)\right] = \frac{\partial\left[\max\left(\tilde{u}_i,p_i\right)\right]}{\partial\lambda_j}(z,\lambda)\\
&= \begin{cases} \frac{\partial\tilde{u}_i}{\partial\lambda_j}(z,\lambda) & \text{if} \quad \tilde{u}_i(z,\lambda) > p_i(z,\lambda)\\ \frac{\partial p_i}{\partial\lambda_j}(z,\lambda) & \text{if} \quad \tilde{u}_i(z,\lambda) < p_i(z,\lambda)\\ \left(\frac{\partial\tilde{u}_i}{\partial\lambda_j}+\frac{\partial p_i}{\partial\lambda_j}\right)(z,\lambda) & \text{if} \quad \tilde{u}_i(z,\lambda) = p_i(z,\lambda) \end{cases}\\
&= H((\tilde{u}_i-p_i)(z,\lambda))\frac{\partial\tilde{u}_i}{\partial\lambda_j}(z,\lambda) + H(-(\tilde{u}_i-p_i)(z,\lambda))\frac{\partial p_i}{\partial\lambda_j}(z,\lambda) \qquad (6.17)
\end{aligned}$$

where $H(\cdot)$ is the Heaviside step function. This derivative exists only in a weak, distributional sense and is only uniquely defined up to set of Lesbegue measure zero. Equation (6.12) becomes

$$\begin{aligned}
&\frac{\partial}{\partial\lambda_j}\left[\max\left(\tilde{u}_i(z-\sqrt{\lambda}\cdot x,\lambda),p_i(z-\sqrt{\lambda}\cdot x,\lambda)\right)\right]\\
&= \frac{\partial\left[\max\left(\tilde{u}_i,p_i\right)\right]}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda)\\
&+ \frac{-x_j}{2\sqrt{\lambda_j}}\frac{\partial\left[\max\left(\tilde{u}_i,p_i\right)\right]}{\partial z_j}(z-\sqrt{\lambda}\cdot x,\lambda).
\end{aligned}$$

The derivative $\frac{\partial[\max(\tilde{u}_i,p_i)]}{\partial z_j}$ is simply equation (6.17) with λ_j replaced by z_j. With $\Delta T_{i+1} = T_{i+1} - T_i$ the integral over the first term is

$$\begin{aligned}
&\int_{\mathbb{R}^N}\phi(x,\Delta T_{i+1},1)\frac{\partial\left[\max\left(\tilde{u}_i,p_i\right)\right]}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda)dx\\
&= \int_{\mathbb{R}^N}\phi(x,\Delta T_{i+1},1)H((\tilde{u}_i-p_i)(z-\sqrt{\lambda}\cdot x,\lambda))\\
&\quad\cdot\frac{\partial\tilde{u}_i}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda)dx\\
&\quad+\int_{\mathbb{R}^N}\phi(x,\Delta T_{i+1},1)H(-(\tilde{u}_i-p_i)(z-\sqrt{\lambda}\cdot x,\lambda))\\
&\quad\cdot\frac{\partial p_i}{\partial\lambda_j}(z-\sqrt{\lambda}\cdot x,\lambda)dx.
\end{aligned}$$

These two integrals exist for all $\lambda_j \geq 0$ if the respective derivatives are integrable functions that do not grow faster than the heat kernel ϕ. For p_i we assured this by assuming that it is sufficiently smooth and bounded. For $\tilde{u}_i$ we can use Lemma 32

to relate this back to properties of $\tilde{u}_{i-1}$, the initial condition of the PDE problem for $\tilde{u}_i$. Iteratively we arrive back at the initial condition g.

Similarly, the integral over the second term is

$$\begin{aligned}
&\int_{\mathbb{R}^N} \phi(x, \Delta T_{i+1}, 1) \frac{-x_j}{2\sqrt{\lambda_j}} \frac{\partial \left[\max\left(\tilde{u}_i, p_i\right)\right]}{\partial z_j}(z - \sqrt{\lambda} \cdot x, \lambda) dx \\
&= \int_{\mathbb{R}^N} \phi(x, \Delta T_{i+1}, 1) \frac{-x_j}{2\sqrt{\lambda_j}} H((\tilde{u}_i - p_i)(z - \sqrt{\lambda} \cdot x, \lambda)) \\
&\quad \cdot \frac{\partial \tilde{u}_i}{\partial z_j}(z - \sqrt{\lambda} \cdot x, \lambda) dx \\
&\quad + \int_{\mathbb{R}^N} \phi(x, \Delta T_{i+1}, 1) \frac{-x_j}{2\sqrt{\lambda_j}} H(-(\tilde{u}_i - p_i)(z - \sqrt{\lambda} \cdot x, \lambda)) \\
&\quad \cdot \frac{\partial p_i}{\partial z_j}(z - \sqrt{\lambda} \cdot x, \lambda) dx
\end{aligned}$$

For $\lambda_j > 0$ the argument is the same as for the first integral. For $\lambda_j \to 0$ we can argue as we did in the proof of Lemma 32 that C^2 but on isolated kinks is sufficient. Overall, the existence of the partial derivative of $\tilde{u}_{i+1}$ is related back to the existence of the partial derivative of $\tilde{u}_{i+1}$ given sufficiently smooth interface conditions. This and the results of Lemma 32 also indicate that we do not accumulate factors in λ_j.

The other option considered in Chapter 6, the Ratchet floor, has interface conditions of the form

$$\begin{aligned}
u(z, T_i, \lambda) &= h_i(z, \lambda, \tilde{u}_i(f(z), \lambda)) \\
&= \tilde{u}_i(f(z), \lambda)
\end{aligned} \tag{6.18}$$

where $f : \mathbb{R}^N \to \mathbb{R}$ is the maximum of a smooth function of z and 0. The spatial coordinates z now include the asset prices and one extra dimension to keep track of the strike price K, which gets updated at the tenor times T_i. The argument for this case is analogous to that for the Bermudan option.

Remark 49. *We can analyse other path-dependent options, such as those incorporating time-averages of asset prices, in a similar way. The details of the model will*

heavily depend on the nature of the path-dependent features. American options require a different type of analysis, as they cannot be modelled with interface conditions.

Chapter 7

EMDP: Expansion Method Derivative Pricer

7.1 Overview and Structure

In this chapter we introduce our Expansion Method Derivative Pricer (EMDP), an object-oriented derivative pricing framework written in MATLAB. It serves as a proof-of-concept for the expansion methods developed in this thesis and as a test suite for their empirical accuracy and computational complexity. The modular design allows to incorporate a wide range of asset models, derivatives and pricing methods. We have provided implementations covering many of the applications and numerical experiments discussed in Chapters 4, 5.3, and 6.[1]

The code behind EMDP is publicly available on GitHub[2]. It is our hope that openly sharing this framework will allow other researches and practitioners to experiment with and investigate expansion methods and their application towards pricing financial derivatives. In an ideal case, it will provide a platform for fruitful future research. In the following, we give a brief description of the overall layout and design of EMDP. We also comment on the implementation of the numerical solvers and the algorithms chosen for PDE and MC computations.

[1]The numerical data in this thesis was for the most part generated with earlier and differently optimised implementations. There are no substantial differences, but minor details of numerical accuracy and performance may vary, in particular regarding run times.

[2]https://github.com/rasmuswissmann/EMDP

The three basic types of classes in EMDP are AssetModels, Derivatives and Pricers. They encode the specifics of the asset and market dynamics, properties of the financial derivative, and the analytical or numerical method for price evaluation. The central operation is

u(S,t) = pricer.price(S,t,assetModel,derivative)

which computes the derivative value $u(S,t)$ under the given asset model using the specified pricer.

- S,t

  S is a vector of starting values at time t for the stochastic assets being modelled. These could be financial assets like stocks, commodities or interest rates, but could also include derived quantities such as volatilities in a stochastic volatility model. For simplicity, we currently assume $t = 0$.

- AssetModel

AssetModel
getVolatility(obj,S,t)
getCorrelation(obj,S,t)
getDrift(obj,S,t)
calculateVolatility(obj,S,t)
calculateCorrelation(obj,S,t)
calculateDrift(obj,S,t)
precompute(obj,S_range,t_range)

  An AssetModel object specifies the dynamics for a given asset model of the form

$$dS_i(S,t) = \mu_i(S,t)dt + \sigma_i(S,t)dW_i, \qquad 1 \leq i \leq N$$

$$Corr(dW_i(S,t), dW_j(S,t)) = \rho_{ij}(S,t), \qquad 1 \leq i,j \leq N$$

  It includes functions to calculate and retrieve model properties such as volatility, correlation, covariance matrix and drift. There are also a number of functions

providing meta-information about the asset model used in internal optimisations of numerical algorithms.

The currently implemented non-abstract classes of asset models are:

- ConstCoeffModel
- LogConstCoeffModel
- VarCoeffModel
- LogExpVolModel
- LogHestonModel
- SemiLogHestonModel

Implemented separately but not yet available as part of EMDP:

- LIBORMarketModel

- Derivative

Derivative
getPayout(obj,S,t) getTerminalTime(obj)

A Derivative object specifies a financial derivative with payout at a given terminal time. This includes European options.

The currently implemented non-abstract classes of derivatives are:

- ArithmeticBasket
- ArithmeticBasketExp
- GeometricBasket
- GeometricBasketExp

Implemented separately but not yet available as part of EMDP:

  - RatchetFloor
  - BermudanSwaption

- Pricer

Pricer
price(obj,S,t,assetModel,derivative)

A Pricer object provides a function to calculate the fair price of a derivative in a given asset model. This can be based on an analytical or numerical method. In the case of a simulation method, the price function also returns an estimate for the standard error of the simulation.

The currently implemented non-abstract classes of pricers are:

  - MCPricer
  - PDEExpansionPricer

Implemented separately but not yet available as part of EMDP:

  - PathDependentMCPricer
  - AndersenBroadieMCPricer
  - PathDependentPDEExpansionPricer
  - BermudanPDEExpansionPricer

The Pricer classes contain the implementation of the mathematical methods behind pricing financial derivatives. We highlight the most important details of these implementations in the following two sections.

7.2 Implementation of PDE-based Pricers

The PDEExpansionPricer is based on the expansion methods developed in this thesis. Computing the approximate solution defined in equation (2.18) requires the solution of low-dimensional PDEs. The current implementation is based on expansion methods constructed via Taylor-expansion of zero, first or second order around $r = 1$ eigenvalues, i.e., we need to numerically solve one-, two- and three-dimensional heat equation-type PDEs with initial conditions given by the derivative's payout function. The PDEExpansionPricer automatically determines the necessary PDEs and sums up their solutions with correct weighting.

For each low-dimensional PDE problem, the spatial domain is unbounded in the z-coordinates. To avoid the introduction of artificial boundary conditions necessary when localising the domain, for each coordinate z_i, we map the interval $(-\infty, \infty)$ to $(0, 1)$ via

$$y_i = \frac{1}{\pi} \arctan\left(\gamma_i z_i + c_i\right) + \frac{1}{2},$$

with parameters γ_i and c_i. This transforms the PDE into

$$\begin{aligned} 0 = &\frac{\partial u}{\partial \tau} - \frac{1}{2} \sum_{k=1}^{N} \lambda_k \frac{\gamma_k^2}{\pi} \cos^3\left[\pi\left(y_k - \tfrac{1}{2}\right)\right] \\ &\cdot \left\{ \frac{1}{\pi} \cos\left[\pi\left(y_k - \tfrac{1}{2}\right)\right] \frac{\partial^2 u}{\partial y_k^2} - \frac{2}{\pi} \sin\left[\pi\left(y_k - \tfrac{1}{2}\right)\right] \frac{\partial u}{\partial y_k} \right\} \end{aligned}$$

on the interval $[0, 1]^N \times [0, T]$, as detailed in Appendix A.2. Initial, boundary and update conditions are transformed accordingly.

Under a standard growth condition on the solution at infinity, the resulting PDE is fully specified without boundary conditions at $z_i \in \{0, 1\}$, because the resulting non-constant coefficients of the transformed diffusion-equation vanish sufficiently fast at the boundaries (see [41, 48]). For options with unbounded option value towards infinity we introduce an artificial cutoff. For example, for the Bermudan swaption

discussed in Chapter 6.3 we apply a payout cutoff at a value $g_{max} = 1000$. This does not significantly impact the computed option value.

We consider an equidistant grid with $J + 1$ grid points along each axis in the y_i coordinates, such that in original coordinates the mesh is denser in the most relevant regions. They can be selected by choosing the parameters γ_i and c_i accordingly. For instance, for interest rates around 0.1 – as considered in Chapter 6 – we chose to map rates between 0.02 and 0.5 to the interval $[0.1, 0.9]$.

For the time discretisation we use the Crank-Nicolson scheme with central spatial differences. In the one-dimensional case, solving the PDE is then straightforward. In the two-dimensional case, we combine this with an Alternating Direction Implicit (ADI) factorisation [36], such that the resulting tridiagonal matrix systems can be solved efficiently in linear time (i.e., proportional to the system size). An initial LU factorisation of the tridiagonal matrices gives significant further speed-up. For non-constant coefficient PDEs, we also use the Hundsdorfer-Verwer scheme [25].

For the efficient solution of the three-dimensional PDEs we have implemented Brian's and Douglas' ADI methods [6, 5, 26, 10]. In our numerical experiments for LIBOR derivatives, the former delivered slightly superior stability and accuracy.

After computing the PDE solution at starting time t we calculate the value for the requested asset values S via cubic spline (one- and two-dimensional case) and linear (three-dimensional case) interpolation.

Depending on the derivative contract, there can be additional parameters and interface conditions to be taken into account. Our numerical experiments dealt with two examples of this, Bermudan swaptions, which offer early-exercise rights at discrete points in time, and Ratchet floors, where a strike parameter is updated depending on interest rates at a pre-defined finite set of tenor dates. The latter makes the payoff strongly path-dependent.

For Bermudan derivatives, update conditions are necessary at every tenor time. We

compare the value of the continuation solution, which we just computed, to the value that would be received in case of immediate exercise. The latter can be calculated analytically.

Path-dependent derivatives can often be dealt with by including one more more additional dimension(s) to model the evolution of path-dependent quantities. Numerically, the value updates at tenor times can then be calculated via interpolation of function values. For example, for the Ratchet floor in Chapter 6.4 this leads to an update condition of the form

$$u(L, t^+, K) = u(L, t^-, \max(a \cdot L_i + b \cdot K + c, 0))$$

In our MATLAB implementation the right hand side is calculated via cubic spline interpolation on the grid for the strike price K, which has to be done separately for every point for the L grid.

7.3 Implementation of MC-based Pricers

Monte Carlo methods are the standard approach for valuing multi-asset financial derivatives. Their application is straightforward for European options with or without path-dependency. We implemented a general MC-based pricer with Euler time-stepping for such derivatives. Our implementation takes into account beneficial properties of the asset model such as constant volatility or drift, and is parallelised using MATLAB's highly efficient functionality for matrix and vector operations. We used advanced techniques such as variance reduction in some of our variance-sensitive numerical experiments, but they are not yet part of the public implementation.

Pricing financial options with early-exercise features via MC methods is significantly more complicated than pricing European options. The reason is that pricing techniques for early-exercise options typically require knowledge of option prices across a range of asset values at intermediate times, whereas basic MC only provides the fair

option value at the initial time and asset price. Specialised algorithms that extended the basic MC method are necessary.

One popular approach is to first try to estimate an exercise policy that is close to optimal. In a second step, this is used in combination with straightforward MC to estimate a lower bound for the fair option value. Using a dual approach one can then calculate an upper bound as well. For a specific example of this see the Bermudan Swaptions in Chapter 6, where we use a primal-dual simulation algorithm due to Andersen and Broadie [2].

7.4 Examples

Example 50. *The following code*[3] *instantiates a log-normal asset model with* $N = 10$ *assets of constant volatility* $\sigma = 0.2$ *and constant asset-asset correlations of* $\rho = 0.7$*. It also defines a European arithmetic basket derivative with equal weights* $1/N$*, maturity* $T = 1$ *and strike* $K = 100$*. It then calculates the derivative price with a MC-based pricer using* 10^7 *paths and with a first order PDE-based pricer using* $J + 1 = 501$ *grid points in each direction.*

```
% European Arithmetic Basket in a Log-Normal Stock Model
N = 10;
sigma = 0.2;
rho = AssetModel.buildSimpleConstantCorrelationMatrix(N,0.7);
assetModel = LogConstCoeffModel(N,sigma,rho,0);

T = 1;
omega = ones(N,1)/N;
K = 100;
derivative = ArithmeticBasketExp(T,omega,K);

S = log(ones(N,1)*100);
t = 0;
dt = 0.1;

J = 500;
order = 1;
pricer = PDEExpansionPricer(J,dt,order,'ADI');
```

[3] See https://github.com/rasmuswissmann/EMDP/ .

```
VPDE = pricer.price(S,t,assetModel,derivative);

numPaths = 10^7;
pricer = MCPricer(numPaths,dt);
[VMC,sMC] = pricer.price(S,t,assetModel,derivative);
```

The calculated output values were $V_{PDE} = 6.8424$, $V_{MC} = 6.8077$ *and* $s_{MC} = 0.0035$. *The difference between the PDE- and MC-based price was* $V_{PDE} - V_{MC} = 0.0347 = 0.0051 \cdot V_{MC} \approx 10 \cdot s_{MC}$.

Example 51. *The following code*[4] *prices a European geometric basket option in an asset model with constant correlation structure and time-dependent volatilities. This is a similar to the ones in Chapter 5.3.*

```
% European Geometric Basket in a Variable Coefficient Stock Model
N = 10;
T = 1;
sigma = @(S,t) (0.1*ones(N,1)+.1*(0:N-1)'/(N-1)) * (1+t/T)/2;
rho = @(S,t) AssetModel.buildSimpleExponentialCorrelationMatrix(N,0.25);
mu = @(S,t) zeros(N,1) - .5*sigma(S,t).^2;
assetModel = VarCoeffModel(N,mu,sigma,rho);
assetModel = assetModel.setIsVarSpaceModel(false);

K = 100;
omega = [2*ones(N/2,1);ones(N/2,1)]/(N*3/2);
derivative = GeometricBasketExp(T,omega,K);

S = log(ones(N,1)*100);
t = 0;
dt = 0.01;

assetModel = assetModel.precompute(zeros(N,1),linspace(0,T,T/dt+1));

J = 300;
order = 1;
pricer = PDEExpansionPricer(J,dt,order,'ADI');
VPDE = pricer.price(S,t,assetModel,derivative);

numPaths = 10^6;
pricer = MCPricer(numPaths,dt);
[VMC,sMC] = pricer.price(S,t,assetModel,derivative);
```

The calculated output values were $V_{PDE} = 2.9300$, $V_{PDEHV} = 3.2706$, $V_{MC} = 2.9196$ *and* $s_{MC} = 0.0046$. *The difference between the PDE- and MC-based price was* $V_{PDE} -$

[4] See https://github.com/rasmuswissmann/EMDP .

$$V_{MC} = 0.0103 = 0.0035 \cdot V_{MC} \approx 2.25 \cdot s_{MC}.$$

Chapter 8

Conclusion

8.1 Results Summary

In this thesis, we developed the idea of a PCA-based Taylor expansion approach to numerically evaluating the pricing equation for basket options [41] into a general framework for calculating point-wise solutions of high-dimensional PDEs via expansions. We defined expansion methods, introduced a general framework, and analysed the resulting approximation errors. We discussed beneficial problem transformations, the connection to financial derivative pricing, and related methods. (Chapter 2)

Focusing on the case of constant coefficient problems first, we explained how to construct methods of different orders from Taylor expansions of the PDE solution. We carefully analysed the existence of expansion solutions and proved rigorous error bounds for first and second order methods through two independent approaches: via the error terms in the Taylor expansion and via directly solving the PDE problem satisfied by the expansion errors themselves. (Chapter 3)

Next we proceeded to apply expansion methods to European options in a standard log-normal stock model. The empirically observed behaviour of the error around kinks of the payout function and its general convergence with the size of the eigenvalues of the covariance matrix of the underlying asset model follow the predicted behaviour. This experimentally demonstrates the accuracy of our theoretical analysis in the preceding two chapters and shows that very precise results can be calculated in practice.

(Chapter 4)

We then considered the general case of non-constant coefficient PDEs and investigated two approaches: Localising coefficients or incorporating varying and non-diagonal contributions in the PDE. With the Parametrix method replacing the Green's function approach we were able to show existence of our approximations and motivate the use of expansion methods by analysing the behaviour of the expansion error. We then conducted numerical experiments for different stock models with varying coefficients, showing that while the error increases and its behaviour becomes harder to predict, the method is still applicable. (Chapter 5)

To analyse the applicability and performance of expansion methods in practically relevant settings we chose to investigate path-dependent and early-exercise derivatives in the LIBOR market model. Using localization in combination with the expansion methods for constant coefficient PDEs, we were able to compute values for Bermudan swaptions and Ratchet floors which very closely tracked the Monte Carlo benchmark for up to $N = 60$ quarterly LIBORs under a range of market conditions. The run times were of the same order of magnitude as the MC run times for the path-dependent example and substantially smaller for the early-exercise case.

In extensions to higher order, we found that both increasing the number of eigenvalues which are fully included in all expansion terms and increasing the order of the Taylor expansion led to a rapid decrease of the relative error, in line with theoretical predictions. While the optimal choice of either order depends on the desired accuracy, including at least the first order Taylor terms seems generally beneficial, given that the corresponding accuracy improvement is significant while keeping the computational cost tractable. A closer look at the raw data going into Figure 6.4 revealed that the correction terms for decreasing eigenvalues are indeed strictly decreasing, which could be used as a heuristic basis for dimension adaptivity. Overall, this application successfully allowed us to demonstrate the practical relevance of expansion

methods for financial derivative pricing problems. (Chapter 6)
Finally, we briefly documented our MATLAB-based implementation, EMDP: Expansion Method Derivative Pricer. This object-oriented framework allows the pricing of financial derivatives in a range of market settings using PDE- and MC-based methods. It serves as a test suite and proof-of-concept for the expansion methods developed in this thesis and is publicly available[1]. (Chapter 7).

Conclusion. Expansion methods offer a widely applicable approach to calculate point-wise solutions of high-dimensional PDEs, in particular in the context of pricing financial derivatives. We have shown how to define, construct and implement expansion methods of various orders. The constant coefficient case allows for a comprehensive theoretical analysis, which includes strong error bounds that are experimentally verified. We have furthermore motivated the use of expansion methods for varying coefficient problems and experimentally shown that these methods can be applied with good accuracy to a range of practically relevant financial market models, such as path-dependent and early-exercise options in the LIBOR market model with up to $60 - 100$ dimensions. Overall, our work has been roughly equally split between theoretical analysis, development of algorithms, implementation, and numerical experiments.

8.2 Future Work

We see four immediate avenues for future work. These are centred around improving the expansion methods developed in this thesis, detailing the theoretical understanding of their behaviour and applying them to additional problems within and outside of mathematical finance.

Variable coefficient methods. The majority of contemporary financial market

[1] https://github.com/rasmuswissmann/EMDP

models include stochastic processes with non-constant volatility and correlation. Expansion methods for varying coefficient PDEs are thus of particular importance for practical applications. A better theoretical understanding of their accuracy and performance would be an important step. The ultimate goal could be a comprehensive set of rigorous theorems regarding applicability and error behaviour – including tight, practically relevant bounds – similar to the results that we gave in Chapter 3 for constant coefficient PDEs.

There is also ample room for method improvements. One could for example consider partial localisation of only a subset of dimensions or coefficients, or a non-constant transformation matrix $Q(t)$. The latter would make it necessary to extrapolate lower-dimensional approximations from calculated results in slightly shifted coordinate systems. Overall, it seems apparent that we have so far only scratched the surface of expansion methods for varying coefficient problems.

In terms of practical applications, more complex asset models such as stochastic volatility or jump-diffusion models would be an obvious next step. We have done some initial work for exponential volatility models of the form

$$\begin{aligned} dS_i &= \exp(Y_i)\, S_i\, dW_i \quad 1 \le i \le N \\ dY_i &= \xi_i \left(\bar{Y}_i - Y_i\right) dt + \theta_i d\tilde{W}_i \quad 1 \le i \le N \end{aligned}$$

and the Heston model [22]

$$\begin{aligned} dS_i &= \sqrt{Y_i}\, S_i\, dW_i \quad 1 \le i \le N \\ dY_i &= \xi_i \left(\bar{Y}_i - Y_i\right) dt + \theta_i \sqrt{Y_i} d\tilde{W}_i \quad 1 \le i \le N. \end{aligned}$$

So far, expansion methods appear applicable in theory and practice, albeit with more effort and complications.

General and problem-specific method variations. It is – due to their quick decay – often possible to include only a subset of the expansion terms without significantly diminishing accuracy. One could also use hybrid methods in which less

accurate MC results are used for smaller high order terms. We demonstrated both approaches in Chapter 6. Beyond this, other dimensionally-adaptive or hybrid PDE-simulation variations are imaginable, as well as further combinations with analytic approximations and simplifications. For example, the control variate approach introduced in Chapter 4 could be used to speed up real world MC computations.

These are examples of how to derive more sophisticated methods from the expansion methods introduced in this thesis. The aim is to reduce the computational complexity or decrease the expansion error – or both. Other modifications along these lines could include different means of evaluating the partial derivatives in the Taylor expansion – i.e., not via finite difference stencils – or different expansion schemes inspired by related approaches, such as Anova, factor and tensor methods.

The original inspiration for expansion methods was a PCA-based coordinate transformation [41], which remains one of the core steps. This coordinate system is optimised to capture as much as possible of the overall system dynamics in subsequent dimension. In terms of the asset prices S and their logarithms $\log(S)$, resp., Q is an orthogonal matrix that first maximises the variance of the first rotated asset $(Q^T \log(S))_1$, then that of the second $(Q^T \log(S))_2$ given the constraints from the first, and so on. We conjecture that other transformations are possible and – depending on the specific problem – might reduce the expansion error. For example, the authors of [24] suggest a modified, problem-specific transformation for MC simulations that optimises for maximal variance in the derivative prices rather than the asset values. In our PDE context, this corresponds to a non-orthogonal factorization rather than an eigendecomposition of Σ and its straightforward application would result in a non-orthogonal coordinate system.

American Options, Greeks, and Calibration. American options, which can be exercised at any time up until maturity, represent a significant percentage of all traded

options. The continuous exercise policy makes pricing them a difficult task. The PDE-based approach involves a free boundary, which gives rise to a linear complimentary problem (LCP) of the form

$$\left\{ \begin{array}{c} \frac{\partial v}{\partial t} - \mathcal{L}v = 0 \\ v - g \leq 0 \end{array} \right\} \text{ or } \left\{ \begin{array}{c} \frac{\partial v}{\partial t} - \mathcal{L}v > 0 \\ v - g = 0 \end{array} \right\},$$

where $\mathcal{L}$ is the Black-Scholes PDE operator and g is the payout function. [8, 46]
One way to attack the LCP are penalty methods. This turns the problem into a non-linear PDE

$$\frac{\partial v}{\partial t} - \mathcal{L}v = \xi \max\{g - v, 0\} \tag{8.1}$$

with a penalty parameter $\xi \to \infty$ [8, 12]. Given again the natural advantages of PDE-based methods over simulation-based methods for options with early-exercise features, applying our expansion methods to American options – e.g. to equations of type (8.1) – seems like an obvious and promising next step.
Another computational challenge in derivatives pricing is the calculation of the Greeks, the derivatives of the value function with regards to asset price, time and model parameters. Their values are crucial for correct risk hedging. We expect our methods to perform well when it comes to estimating the Greeks, though some peculiarities will arise due to the expansion depending on model parameters and the fact that we calculate solutions only point-wise in most of the rotated asset dimensions. Analytical results similar to the ones derived this thesis should hold.
Lastly, pricing of derivatives in practice has to be understood in the context of calibration, i.e., the estimation of model parameters from the values of traded products. It would thus be interesting to understand the interaction of this with calibration methods. For example, how do errors and uncertainties in fitted model parameters propagate through the use of expansion, which depends heavily on (eigenvalues and eigenvectors of) the covariance matrix?

Beyond mathematical finance. Given the advances in solving them, high-dimensional PDEs have become increasingly prominent in science and engineering. This is part of a larger trend of scientific computation having become important in applications that could not previously be tackled with numerical and computational methods. Examples for practically relevant high-dimensional PDEs include "kinetic plasma physics equations, the many body Schrödinger equation, Dirac and Maxwell equations for molecular electronic structures and nuclear dynamic computations, (...) as well as Fokker-Planck and fluid dynamics equations for complex fluids" [3].

It is our hope that researchers in such areas outside of mathematical finance will find the PDE expansion methods developed here useful. We would like to stress again here that our methods are by no means restricted to the kind of second order parabolic equations and boundary conditions primarily considered in this thesis.

Appendix A

Calculations

A.1 PDE Coordinate Transformation

Proof. Denote the old coordinates by (x, τ) and the new by (z, t). The transformation

$$\begin{aligned} t &= \tau \\ z &= Q^T x + \beta(x_0, \tau) \\ \beta_i(x, \tau) &= \sum_j^N Q_{ji} \int_0^\tau B_j(x, t')dt' \end{aligned}$$

leads to

$$\begin{aligned} \frac{\partial}{\partial x_i} &= \sum_{k=1}^N \frac{\partial z_k}{\partial x_i}\frac{\partial}{\partial z_k} + \frac{\partial t}{\partial x_i}\frac{\partial}{\partial t} = \sum_{k=1}^N Q_{ik}\frac{\partial}{\partial z_k} \\ \frac{\partial^2}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_i}\left(\sum_{k=1}^N Q_{jk}\frac{\partial}{\partial z_k}\right) = \sum_{l=1}^N Q_{il}\frac{\partial}{\partial z_l}\sum_{k=1}^N Q_{jk}\frac{\partial}{\partial z_k} = \sum_{k,l=1}^N Q_{jk}Q_{il}\frac{\partial^2}{\partial z_k \partial z_l} \\ \frac{\partial}{\partial \tau} &= \sum_{k=1}^N \frac{\partial z_k}{\partial \tau}\frac{\partial}{\partial z_k} + \frac{\partial t}{\partial \tau}\frac{\partial}{\partial t} = \sum_{k=1}^N \left(\sum_j^N Q_{jk}B_j(x_0, \tau)\right)\frac{\partial}{\partial z_k} + \frac{\partial}{\partial t} \end{aligned}$$

Applying this to the PDE (2.1)-(2.3) gives

$$\begin{aligned}\mathcal{L}(x,\tau) - \frac{\partial}{\partial\tau} &= \sum_{i,j=1}^{N} A_{ij}(x,\tau)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{N} B_i(x,\tau)\frac{\partial}{\partial x_i} - \frac{\partial u}{\partial\tau} \\ &= \sum_{i,j,k,l=1}^{N} A_{ij}(x,\tau)Q_{jk}Q_{il}\frac{\partial^2}{\partial z_k \partial z_l} + \sum_{i,k=1}^{N} B_i(x,\tau)Q_{ik}\frac{\partial}{\partial z_k} \\ &\quad - \sum_{k,j=1}^{N} Q_{jk}B_j(x_0,\tau)\frac{\partial}{\partial z_k} - \frac{\partial}{\partial t} \\ &= \sum_{k,l=1}^{N} \lambda_{kl}(z,t)\frac{\partial^2}{\partial z_k \partial z_l} + \sum_{k=1}^{N} \kappa_k(z,t)\frac{\partial}{\partial z_k} - \frac{\partial}{\partial t}\end{aligned}$$

with

$$\begin{aligned}\lambda_{kl}(z,t) &= \sum_{i,j=1}^{N} Q_{ik}Q_{jl}A_{ij}(x(z,t),t) \\ \kappa_k(z,t) &= \sum_{i=1}^{N} Q_{ik}\left[B_i(x(z,t),t) - B_i(x_0,t)\right]\end{aligned}$$

□

A.2 Infinite to Finite Domain Transformation

We here derive the PDE problem that results after the coordinate transformation

$$y_i = \frac{1}{\pi}\arctan\left(\gamma_i v_i + c_i\right) + \frac{1}{2} \tag{A.1}$$

was introduced in Chapter 7.2. With this transformation, the spatial derivatives can be written as

$$\begin{aligned}\frac{\partial u}{\partial v_i} &= \sum_{k=1}^{N} \frac{\partial y_k}{\partial v_i}\frac{\partial u}{\partial y_k} = \frac{\partial y_i}{\partial v_i}\frac{\partial u}{\partial y_i} \\ &= \frac{{}^{\gamma_i}/_{\pi}}{1+\left(\gamma_i v_i + c_i\right)^2}\frac{\partial u}{\partial y_i} \\ &= \frac{\gamma_i}{\pi}\cos^2\left[\pi\left(y_i - \tfrac{1}{2}\right)\right]\frac{\partial u}{\partial y_i}\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 u}{\partial v_i^2} &= \frac{\partial}{\partial v_i}\left\{\frac{\gamma_i}{\pi}\cos^2\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\frac{\partial u}{\partial y_i}\right\} \\
&= \sum_{k=1}^{N}\frac{\partial y_k}{\partial v_i}\frac{\partial}{\partial y_k}\left\{\frac{\gamma_i}{\pi}\cos^2\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\frac{\partial u}{\partial y_i}\right\} \\
&= \frac{\gamma_i}{\pi}\cos^2\left[\pi\left(y_i-\tfrac{1}{2}\right)\right] \\
&\quad\cdot\left\{\frac{\gamma_i}{\pi}\cos^2\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\frac{\partial^2 u}{\partial y_i^2}+\frac{\gamma_i}{\pi}\left(\frac{\partial}{\partial y_i}\cos^2\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\right)\frac{\partial u}{\partial y_i}\right\} \\
&= \frac{\gamma_i^2}{\pi}\cos^3\left[\pi\left(y_i-\tfrac{1}{2}\right)\right] \\
&\quad\cdot\left\{\frac{1}{\pi}\cos\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\frac{\partial^2 u}{\partial y_i^2}-\frac{2}{\pi}\sin\left[\pi\left(y_i-\tfrac{1}{2}\right)\right]\frac{\partial u}{\partial y_i}\right\}.
\end{aligned}$$

This transforms the N-dimensional heat equation on the infinite spatial domain into

$$\begin{aligned}
0 &= \frac{\partial u}{\partial \tau}-\frac{1}{2}\sum_{k=1}^{N}\lambda_k\frac{\gamma_k^2}{\pi}\cos^3\left[\pi\left(y_k-\tfrac{1}{2}\right)\right] \\
&\quad\cdot\left\{\frac{1}{\pi}\cos\left[\pi\left(y_k-\tfrac{1}{2}\right)\right]\frac{\partial^2 u}{\partial y_k^2}-\frac{2}{\pi}\sin\left[\pi\left(y_k-\tfrac{1}{2}\right)\right]\frac{\partial u}{\partial y_k}\right\}
\end{aligned}$$

on the interval $[0,1]^N \times [0,T]$. The new boundary and update conditions follow directly from expressing v in terms y, i.e. by reversing the transformation in equation (A.1).

A.3 Calculations for Example 25

Case 2: $\mu_1 = 1$, $\mu_2 = -1$

$$\begin{aligned}
\frac{\partial}{\partial\lambda_2}u(z,t,\lambda) &= \frac{\partial}{\partial\lambda_2}\int_{\mathbb{R}}\Phi(x_2,t,1)\mathbb{1}_{[0,\infty)}\left(\exp\left[z_2-\sqrt{\lambda_2}x_2\right]-2\right) \\
&\qquad\cdot\int_{\mathbb{R}}\Phi(x_1,t,1)dx_1dx_2 \\
&= \frac{\partial}{\partial\lambda_2}\int_{\mathbb{R}}\Phi(x_2,t,1)\mathbb{1}_{[0,\infty)}\left(\exp\left[z_2-\sqrt{\lambda_2 x_2}\right]-2\right)dx_2 \\
&= \frac{\partial}{\partial\lambda_2}\frac{\exp(-x_2^2/4t)}{\sqrt{4\pi t}}dx_2 \\
&= \frac{\partial}{\partial\lambda_2}\int_{-\infty}^{(z_2-\log 2)/\sqrt{2t\lambda_2}}\frac{\exp(-y_2^2/2)}{\sqrt{2\pi}}dy_2 \\
&= \frac{\partial}{\partial\lambda_2}\mathrm{CDF}\left(\frac{z_2-\log 2}{\sqrt{2t\lambda_2}}\right) \\
&= -\frac{z_2-\log 2}{2\sqrt{2t}\sqrt{\lambda_2}^3}\cdot\frac{\exp(-(z_2-\log 2)^2/4t\lambda_2)}{\sqrt{2\pi}}
\end{aligned}$$

Case 3: $\mu_1 = 2$, $\mu_2 = 0$

$$
\begin{aligned}
\frac{\partial}{\partial \lambda_2} u(z,t,\lambda) &= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}} \int_{\mathbb{R}} \Phi(x_2,t,1)\Phi(x_1,t,1) \\
&\qquad \cdot \mathbb{1}_{[0,\infty)} \left(\exp\left[z_1 - \sqrt{\lambda_1} x_1 z_2 - \sqrt{\lambda_2} x_2 \right] - 2 \right) dx_1 dx_2 \\
&= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}} \Phi(x_2,t,1) \\
&\qquad \cdot \int_{-\infty}^{(z_1+z_2-\sqrt{\lambda_2}x_2-\log 2)/\sqrt{\lambda_1}} \Phi(x_1,t,1) dx_1 dx_2 \\
&= \frac{\partial}{\partial \lambda_2} \int_{\mathbb{R}} \frac{\exp(-x_2^2/4t)}{\sqrt{4\pi t}} \\
&\qquad \cdot \mathrm{CDF}\left(\frac{z_1 + z_2 - \sqrt{\lambda_2} x_2 - \log 2}{\sqrt{2t\lambda_1}} \right) dx_2 \\
&= \int_{\mathbb{R}} \frac{\exp(-x_2^2/4t)}{\sqrt{4\pi t}} \frac{-x_2}{2\sqrt{\lambda_2}} \\
&\qquad \cdot \frac{\exp(-(z_1 + z_2 - \sqrt{\lambda_2} x_2 - \log 2)^2/4t\lambda_1)}{\sqrt{2\pi}} dx_2 \\
&= -\frac{1}{4\pi\sqrt{2t\lambda_2}} \int_{\mathbb{R}} x_2 \exp(-x_2^2/4t) \\
&\qquad \cdot \exp(-(\sqrt{\lambda_2} x_2 + \alpha)^2/4t\lambda_1) dx_2 \\
&= -\frac{\exp\left(-\frac{\alpha^2\lambda_2}{4t\lambda_1(\lambda_1+\lambda_2)}\right)}{4\pi\sqrt{2t\lambda_2}} \\
&\qquad \cdot \int_{\mathbb{R}} x_2 \exp\left[-\frac{(\sqrt{\lambda_1+\lambda_2} x_2 + \frac{\alpha\sqrt{\lambda_2}}{\sqrt{\lambda_1+\lambda_2}})^2}{4t\lambda_1} \right] dx_2 \\
&= -\frac{\exp\left(-\frac{\alpha^2\lambda_2}{4t\lambda_1(\lambda_1+\lambda_2)}\right)}{4\pi\sqrt{2t\lambda_2}(\lambda_1+\lambda_2)} \\
&\qquad \cdot \int_{\mathbb{R}} \left(y_2 - \frac{\alpha\sqrt{\lambda_2}}{\sqrt{\lambda_1+\lambda_2}} \right) \exp\left[-\frac{y_2^2}{4t\lambda_1} \right] dy_2 \\
&= \frac{\exp\left(-\frac{\alpha^2\lambda_2}{4t\lambda_1(\lambda_1+\lambda_2)}\right)}{4\pi\sqrt{2t\lambda_2}(\lambda_1+\lambda_2)} \frac{\alpha\sqrt{\lambda_2}}{\sqrt{\lambda_1+\lambda_2}} \int_{\mathbb{R}} \exp\left[-\frac{y_2^2}{4t\lambda_1} \right] dy_2 \\
&= \frac{\alpha}{\sqrt{8\pi}} \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1+\lambda_2}^3} \exp\left(-\frac{\alpha^2\lambda_2}{4t\lambda_1(\lambda_1+\lambda_2)} \right) \\
&= \frac{\log 2 - z_1 - z_2}{\sqrt{8\pi}} \frac{\sqrt{\lambda_1}}{\sqrt{\lambda_1+\lambda_2}^3} \exp\left(-\frac{(\log 2 - z_1 - z_2)^2\lambda_2}{4t\lambda_1(\lambda_1+\lambda_2)} \right)
\end{aligned}
$$

A.4 Calculations for Example 26

$$
\begin{aligned}
\frac{\partial}{\partial\lambda_2}u(z,t,\lambda) &= \frac{\partial}{\partial\lambda_2}\int_{\mathbb{R}}\Phi(x_2,t,1)\max(e^{z_2-\sqrt{\lambda_2}x_2}-2,0)dx_2 \\
&= \int_{\mathbb{R}}\Phi(x_2,t,1)\frac{-x_2}{2\sqrt{\lambda_2}}\frac{\partial}{\partial z_2}\max(e^{z_2-\sqrt{\lambda_2}x_2}-2,0)dx_2 \\
&= \int_{-\infty}^{\frac{z_2-\log 2}{\sqrt{\lambda_2}}}\frac{x_2e^{-x_2^2/4t}}{2\sqrt{\lambda_2}\sqrt{4\pi t}}e^{z_2-\sqrt{\lambda_2}x_2}dx_2 \\
&= \frac{e^{z_2+\lambda_2 t}}{2\sqrt{\lambda_2}}\int_{-\infty}^{\frac{z_2-\log K}{\sqrt{\lambda_2}}}\frac{x_2e^{-(x_2+2\sqrt{\lambda_2}t)^2/4t}}{\sqrt{4\pi t}}dx_2 \\
&= \frac{e^{z_2+\lambda_2 t}}{2\sqrt{\lambda_2}}2t\int_{-\infty}^{\frac{z_2-\log K+2\lambda_2 t}{\sqrt{2t\lambda_2}}}\frac{(y-\lambda_2 t)e^{-y^2/2}}{\sqrt{2\pi}}dy \\
&= \frac{e^{z_2+\lambda_2 t}}{2\sqrt{\lambda_2}}\left\{\sqrt{2t}\left[-\frac{e^{-y^2/2}}{\sqrt{2\pi}}\right]_{-\infty}^{\frac{z_2-\log K+2\lambda_2 t}{\sqrt{2t\lambda_2}}}\right. \\
&\qquad \left.+\sqrt{\lambda_2}t\mathrm{CDF}\left(\frac{z_2-\log K+2\lambda_2 t}{\sqrt{2t\lambda_2}}\right)\right\} \\
&= \frac{te^{z_2+\lambda_2 t}}{2}\left\{-\frac{e^{-(z_2-\log K+2\lambda_2 t)^2/4t\lambda_2}}{\sqrt{\pi t\lambda_2}}\right. \\
&\qquad \left.+\mathrm{CDF}\left(\frac{z_2-\log K+2\lambda_2 t}{\sqrt{2t\lambda_2}}\right)\right\}
\end{aligned}
$$

Appendix B

Proofs

B.1 Proof of Theorem 19

Proof. Set $\hat{u}^{\{1,\dots,r,k\}}(z,t) = u^{\{1,\dots,r,k\}}(z,t) - u(z,t)$. Then $\hat{u}^{\{1,\dots,r,k\}}(z,t)$ solves the PDE

$$\begin{aligned}\frac{\partial}{\partial t}\hat{u}^{\{1,\dots,r,k\}}(z,t) &= \sum_{i\in\{1,\dots,r,k\}} \lambda_i \frac{\partial^2}{\partial z_i^2} u^{\{1,\dots,r,k\}}(z,t) - \sum_{i=1}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z,t) \\ &= \sum_{i\in\{1,\dots,r,k\}} \lambda_i \frac{\partial^2}{\partial z_i^2} \hat{u}^{\{1,\dots,r,k\}}(z,t) - \sum_{i=r+1,i\neq k}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z,t)\end{aligned}$$

for all $(z,t) \in \mathbb{R}^N \times (0,T]$ with zero boundary condition. This gives

$$\begin{aligned}\hat{u}^{\{1,\dots,r,k\}}(z,t) = &- \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}}(y,s) \\ &\cdot \sum_{i=r+1,i\neq k}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z-y,s)dyds \\ \frac{\partial^2}{\partial z_l^2}\hat{u}^{\{1,\dots,r,k\}}(z,t) = &- \sum_{i=r+1,i\neq k}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}}(y,s) \\ &\cdot \frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z-y,s)dyds\end{aligned}$$

Since it is itself a solution to a PDE, $\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t)$ can be writen as

$$\frac{\partial^4}{\partial z_l^2 \partial z_i^2} u(z,t) = \int_{\mathbb{R}^N} \Phi^N(y,t) \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(x(z-y,0))dy$$

Thus

$$\begin{aligned}
&\left|\frac{\partial^2}{\partial z_l^2}\hat{u}^{\{1,\dots,r,k\}}(z,t)\right| \\
&\leq \sum_{i=r+1,i\neq k}^{N} \lambda_i \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}}(y,s) \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(x(z-y,0)) \right\|_\infty dyds \\
&\leq t \sum_{i=r+1,i\neq k}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(x(z-y,0)) \right\|_\infty
\end{aligned}$$

and finally

$$\begin{aligned}
\left|\hat{u}^{\xi}(z,t)\right| &= \left| \sum_{k=r+1}^{N} \lambda_k \int_0^T \int_{\mathbb{R}^r} \prod_{i=1}^{r} \frac{\exp(-y_i^2/4\lambda_i t)}{\sqrt{4\pi\lambda_i t}} \right. \\
&\qquad \left. \cdot \frac{\partial^2}{\partial z_k^2} \hat{u}^{\{1,\dots,r,k\}}(z-y',s) dy' ds \right| \\
&\leq \left| \sum_{k=r+1}^{N} \lambda_k \int_0^T \int_{\mathbb{R}^r} \prod_{i=1}^{r} \frac{\exp(-y_i^2/4\lambda_i t)}{\sqrt{4\pi\lambda_i t}} \right. \\
&\qquad \left. \cdot t \sum_{i=r+1,i\neq k}^{N} \lambda_i \left\| \frac{\partial^4}{\partial z_l^2 \partial z_i^2} g(x(z-y',0)) \right\|_\infty dy' ds \right| \\
&\leq \frac{t^2}{2} \sum_{k=r+1}^{N} \sum_{i=r+1,i\neq k}^{N} \lambda_k \lambda_i \left\| \frac{\partial^4}{\partial z_k^2 \partial z_i^2} g(x(z-y,0)) \right\|_\infty
\end{aligned}$$

□

B.2 Proof of Lemma 32

Proof. We assumed that the boundary function g was such that

$$G_j(z,t,x_j,\lambda) = \int_{\mathbb{R}^{N-1}} \Phi(x',t,1) g(z - \sqrt{\lambda}\cdot x) dx'$$

is interval-wise C^2 in z_j and $\Phi(x_j,t,1)$-integrable in x_j. We also assumed that $\frac{\partial}{\partial z_j} G_j(z,t,x_j,\lambda)$ is $x_j\Phi(x_j,t,1)$-integrable in x_j and that for every fixed (z'',t,x_j,λ) its zero set has only finitely many isolated points in z_j. The last condition together with G_j being interval-wise C^2 implies that $|\frac{\partial}{\partial z_j} G_j(z,t,x_j,\lambda)|$ is interval-wise C^1 in z_j, which is a property that we will need later in the proof.

We first show that $f(z,t,x_j,\lambda) = \Phi(x_j,t,1)G_j(z,t,x_j,\lambda)$ satisfies the conditions of Lemma 27, allowing us to bring the differential operator under the integral. Then we use this to calculate the derivative.

Condition (i): $f(z,t,x_j,\lambda)$ is jointly measurable in x and λ for all t and z. It is integrable in x for all $\lambda \in [0,\infty)^N$.

Condition (ii): The derivative in λ_j is formally

$$\begin{aligned}
\frac{\partial}{\partial\lambda_j} f(z,t,x_j,\lambda) &= \Phi(x_j,t,1)\frac{\partial}{\partial\lambda_j} G_j(z,t,x_j,\lambda) \\
&= \Phi(x_j,t,1)\frac{\partial}{\partial\lambda_j}\int_{\mathbb{R}^{N-1}} \Phi(x',t,1)g(z-\sqrt{\lambda}\cdot x)dy' \\
&= \frac{-x_j\Phi(x_j,t,1)}{2\sqrt{\lambda_j}}\frac{\partial}{\partial z_j}\int_{\mathbb{R}^{N-1}} \Phi(x',t,1)g(z-\sqrt{\lambda}\cdot x)dy' \\
&= \frac{-x_j\Phi(x_j,t,1)}{2\sqrt{\lambda_j}}\frac{\partial}{\partial z_j} G_j(z,t,x_j,\lambda)
\end{aligned}$$

which exists almost everywhere for all $\lambda_j > 0$ and f is its Lebesgue integral on $[0,\infty)$. Thus $\frac{\partial}{\partial\lambda_j} f(x,t,\lambda)$ is absolutely continuous on $[0,\infty)$ and satisfies (ii).

Condition (iii): We need to show that $\frac{\partial}{\partial\lambda_j} f$ is locally integrable in λ_j, i.e., that $\forall [a,b] \subset [0,\infty)$

$$\int_a^b \int_{\mathbb{R}} \left| \frac{\partial f}{\partial\lambda_j}(z,t,x_j,\lambda) \right| dx_j d\lambda_j < \infty \tag{B.1}$$

$\frac{\partial}{\partial\lambda_j} f$ is then given by equation (3.14).

Since $|\frac{\partial}{\partial z_j} G_j|$ is piecewise C^1, we can write

$$\begin{aligned}
&\int_{\mathbb{R}} \left| \frac{\partial f}{\partial\lambda_j}(z,t,x_j,\lambda) \right| dx \\
&= \int_{\mathbb{R}} \frac{|x_j|\Phi(x_j,t,1)}{2\sqrt{\lambda_j}} \left| \frac{\partial}{\partial z_j} G_j(z,t,x_j,\lambda) \right| dx_j \\
&= \int_{\mathbb{R}} \frac{|x_j|\Phi(x_j,t,1)}{2\sqrt{\lambda_j}} \sum_{i=1}^{n} \mathbb{1}_{[a_{i-1},a_i]}(z_j - \sqrt{\lambda_j}x_j)\frac{\partial}{\partial z_j} G_j^i(z_j - \sqrt{\lambda_j}x_j)dx_j \\
&= \sum_{i=1}^{n} \int_{\mathbb{R}} \frac{|y_j|\Phi(y_j,t,\lambda_j)}{2\lambda_j} \mathbb{1}_{[a_{i-1},a_i]}(z_j - y_j)\frac{\partial}{\partial z_j} G_j^i(z_j - y_j)dy_j \\
&= \frac{1}{2}\sum_{i=1}^{n} \frac{1}{\lambda_j} \int_{z_j-a_i}^{z_j-a_{i-1}} |y_j|\Phi(y_j,t,\lambda_j)\frac{\partial}{\partial z_j} G_j^i(z_j - y_j)dy_j
\end{aligned} \tag{B.2}$$

where $n \in \mathcal{N} \cup \infty$, $-\infty = a_0 < a_1 < \ldots < a_n = \infty$ and positive functions $G_j^i \in C^1([a_{i-1}, a_i])$ for all i. For simplicity of notation we have not explicitly written out the dependence of G_j^i and a_i on z', t, λ'. Each G_j^i is integrable on its respective interval thus condition (B.1) is satisfied for all $a > 0$.

W.l.o.g. can assume that there exists an l such that $a_l = z_j$, since we can always split the corresponding interval. Equation (B.2) transforms into

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{l}\frac{1}{\lambda_j}\int_{z_j-a_i}^{z_j-a_{i-1}} y_j\Phi(y_j,t,\lambda_j)\frac{\partial}{\partial z_j}G_j^i(z_j-y_j)dy_j \\
&-\frac{1}{2}\sum_{i=l+1}^{n}\frac{1}{\lambda_j}\int_{z_j-a_i}^{z_j-a_{i-1}} y_j\Phi(y_j,t,\lambda_j)\frac{\partial}{\partial z_j}G_j^i(z_j-y_j)dy_j
\end{aligned}$$

Due to symmetry it is sufficient to foucs on the first term. Using partial integration we can rewrite it as

$$\begin{aligned}
&-t\sum_{i=1}^{l}\left[\Phi(y_j,t,\lambda_j)\frac{\partial}{\partial z_j}G_j^i(y_j)\right]_{z_j-a_i}^{z_j-a_{i-1}} \\
&+t\sum_{i=1}^{l}\int_{z_j-a_i}^{z_j-a_{i-1}}\Phi(y_j,t,\lambda_j)\frac{\partial^2}{\partial z_j^2}G_j^i(z_j-y_j)dy_j \\
&=t\sum_{i=1}^{l}\left[\frac{e^{-(z_j-a_i)^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial}{\partial z_j}G_j^i(a_i)-\frac{e^{-(z_j-a_{i-1})^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial}{\partial z_j}G_j^i(a_{i-1})\right] \\
&+t\sum_{i=1}^{l}\int_{z_j-a_i}^{z_j-a_{i-1}}\frac{e^{-y_j^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial^2}{\partial z_j^2}G_j^i(z_j-y_j)dy_j
\end{aligned}$$

This gives

$$\begin{aligned}
&\int_a^b t\sum_{i=1}^{l}\left[\frac{e^{-(z_j-a_i)^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial}{\partial z_j}G_j^i(a_i)\right.\\
&\qquad\left.-\frac{e^{-(z_j-a_{i-1})^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial}{\partial z_j}G_j^i(a_{i-1})\right]d\lambda_j\\
&\quad+\int_a^b t\sum_{i=1}^{l}\int_{z_j-a_i}^{z_j-a_{i-1}}\frac{e^{-y_j^2/4t\lambda_j}}{\sqrt{4\pi t\lambda_j}}\frac{\partial^2}{\partial^2 z_j}G_j^i(z_j-y_j)dy_jd\lambda_j\\
\leq\ & t\sum_{i=1}^{l}\int_a^b\frac{1}{\sqrt{4\pi t\lambda_j}}\left[e^{-(z_j-a_i)^2/4t\lambda_j}\frac{\partial}{\partial z_j}G_j^i(a_i)\right.\\
&\qquad\left.-e^{-(z_j-a_{i-1})^2/4t\lambda_j}\frac{\partial}{\partial z_j}G_j^i(a_{i-1})\right]d\lambda_j\\
&\quad+t\sum_{i=1}^{l}\int_a^b\frac{1}{\sqrt{4\pi t\lambda_j}}\int_{z_j-a_i}^{z_j-a_{i-1}}e^{-y_j^2/4t\lambda_j}\frac{\partial^2}{\partial^2 z_j}G_j^i(z_j-y_j)dy_jd\lambda_j\\
<\ &\infty\quad\forall[a,b]\subset[0,\infty)
\end{aligned}$$

as required.

We can now use Lemma 27, and with a similar decomposition of G into its piecewise components and partial integration calculate $\frac{\partial u}{\partial \lambda_j}$. For $\lambda_j > 0$ we have

$$
\begin{aligned}
&\frac{\partial}{\partial \lambda_j} u(z,t,\lambda) \\
=&\frac{\partial}{\partial \lambda_j} \int_{\mathbb{R}^N} f(z,t,\lambda,x)dx \\
=&\frac{\partial}{\partial \lambda_j} \int_{\mathbb{R}} \Phi(y_j,t,\lambda_j) \int_{\mathbb{R}^{N-1}} \Phi(y',t,\lambda') g(z'-y', z_j - y_j) dy' dy_j \\
=&\frac{\partial}{\partial \lambda_j} \int_{\mathbb{R}} \Phi(x_j,t,1) G_j(z,t,x_j,\lambda) dx_j \\
=&\int_{\mathbb{R}} \Phi(x_j,t,1) \frac{\partial}{\partial \lambda_j} G_j(z,t,x_j,\lambda) dx_j \\
=&\int_{\mathbb{R}} \Phi(x_j,t,1) \frac{-x_j}{2\sqrt{\lambda_j}} \frac{\partial}{\partial z_j} G_j(z,t,x_j,\lambda) dx_j \\
=&\int_{\mathbb{R}} \Phi(x_j,t,1) \frac{-x_j}{2\sqrt{\lambda_j}} \sum_{i=1}^{n} \mathbb{1}_{[a_{i-1},a_i]}(z_j - \sqrt{\lambda_j} x_j) \frac{\partial}{\partial z_j} G_j^i(z,t,x_j,\lambda) dx_j \\
=&\sum_{i=1}^{n} \int_{z_j - a_i}^{z_j - a_{i-1}} \frac{-y_j \Phi(y_j,t,\lambda_j)}{2\lambda_j} \frac{\partial}{\partial z_j} G_j^i(z,t,y_j,\lambda) dy_j \\
=&t \sum_{i=1}^{n} \int_{z_j - a_i}^{z_j - a_{i-1}} \Phi(y,t,\lambda) \frac{\partial^2 G_i}{\partial z_j^2}(z,t,y_j,\lambda) dy_j \\
&+ t \sum_{i=1}^{n} \Phi(z_j - a_i,t,\lambda_j) \frac{\partial G_i}{\partial z_j}(z,t,a_i,\lambda) \\
&- t \sum_{i=1}^{n} \Phi(z_j - a_{i-1},t,\lambda_j) \frac{\partial G}{\partial z_j}(z,t,a_{i-1},\lambda) \\
=&t \int_{\mathbb{R}} \Phi(y_j,t,\lambda_j) \frac{\partial^2 G}{\partial z_j^2}(z,t,y_j,\lambda) dy_j \\
&+ t \sum_{i=1}^{n} \Phi(z_j - a_i,t,\lambda_j) \frac{\partial G_i}{\partial z_j}(z,t,a_i,\lambda) \\
&- t \sum_{i=1}^{n} \Phi(z_j - a_{i-1},t,\lambda_j) \frac{\partial G}{\partial z_j}(z,t,a_{i-1},\lambda)
\end{aligned}
$$

Note that $\frac{\partial^2 G}{\partial z_j^2}$ exists almost everywhere and is integrable, so that the integral on the right hand side is well defined. Assume now that $z_j \notin \{a_0, \ldots, a_n\}$, where the latter only includes the true kink points of g, i.e., not the possibly artificial split used before to prove condition (iii). Then

$$
\begin{aligned}
&\lim_{\lambda_j \to 0} \frac{\partial}{\partial \lambda_j} u(z,t,\lambda) \\
=& t \lim_{\lambda_j \to 0} \int_{\mathbb{R}} \Phi(y_j,t,\lambda_j) \frac{\partial^2 G}{\partial z_j^2}(z,t,y_j,\lambda) dy_j \\
&+ t \lim_{\lambda_j \to 0} \sum_{i=1}^{n} \Phi(z_j - a_i,t,\lambda_j) \frac{\partial G_i}{\partial z_j}(z,t,a_i,\lambda) \\
&- t \lim_{\lambda_j \to 0} \sum_{i=1}^{n} \Phi(z_j - a_{i-1},t,\lambda_j) \frac{\partial G_i}{\partial z_j}(z,t,a_{i-1},\lambda) \\
=& t \int_{\mathbb{R}} \delta(y_j) \frac{\partial^2 G}{\partial z_j^2}(z,t,y_j,\lambda) dy_j \\
&+ t \sum_{i=1}^{n} \delta(z_j - a_i) \frac{\partial G_i}{\partial z_j}(z,t,a_i,\lambda) \\
&- t \sum_{i=1}^{n} \delta(z_j - a_{i-1}) \frac{\partial G_i}{\partial z_j}(z,t,a_{i-1},\lambda) \\
=& t \frac{\partial^2 G}{\partial z_j^2}(z,t,0,\lambda)
\end{aligned}
$$

□

B.3 Proof of Theorem 44

Proof. We start with equation (2.21)

$$
\begin{aligned}
\frac{\partial}{\partial t} \hat{u}^{\xi}(z,t) = \mathcal{L}^{\nu_\xi}(z,t) \hat{u}^{\xi}(z,t) &+ \sum_{(w,\nu) \in \xi} w \left[\mathcal{L}^{\nu}(z,t) - \mathcal{L}^{\nu_\xi}(z,t) \right] u^{\nu}(z,t) \\
&+ \left[\mathcal{L}^{\nu_\xi}(z,t) - \mathcal{L}(z,t) \right] u(z,t)
\end{aligned}
$$

and substitute in the expansion encoding (3.19) leading to

$$
\begin{aligned}
&\frac{\partial}{\partial t}\hat{u}^{\xi}(z,t)\\
&= \mathcal{L}^{\{1,\dots,r\}}\hat{u}^{\xi}(z,t) + (1+(N-r)(N-r-3)/2)\left[\mathcal{L}^{\{1,\dots,r\}} - \mathcal{L}^{\{1,\dots,r\}}\right] u^{\{1,\dots,r\}}(z,t)\\
&\quad + (2-(N-r))\sum_{i=r+1}^{N}\left[\mathcal{L}^{\{1,\dots,r,i\}} - \mathcal{L}^{\{1,\dots,r\}}\right] u^{\{1,\dots,r,i\}}(z,t)\\
&\quad + \sum_{i=r+1}^{N}\sum_{j=i+1}^{N}\left[\mathcal{L}^{\{1,\dots,r,i,j\}} - \mathcal{L}^{\{1,\dots,r\}}\right] u^{\{1,\dots,r,i,j\}}(z,t)\\
&\quad + \left[\mathcal{L}^{\{1,\dots,r\}} - \mathcal{L}^{\{1,\dots,N\}}\right] u(z,t)\\
&= \sum_{k=1}^{r}\lambda_k\frac{\partial^2}{\partial z_k^2}\hat{u}^{\xi}(z,t) + 0 - (N-r-2)\sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}u^{\{1,\dots,r,k\}}(z,t)\\
&\quad + \sum_{k=r+1}^{N}\sum_{l=k+1}^{N}\left(\lambda_k\frac{\partial^2}{\partial z_k^2} + \lambda_l\frac{\partial^2}{\partial z_l^2}\right) u^{\{1,\dots,r,k,l\}}(z,t) - \sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}u(z,t)
\end{aligned}
$$

Observing that the double sum contains exactly $(N-r)(N-r-1)$ terms we can rewrite this as

$$
\begin{aligned}
\frac{\partial}{\partial t}\hat{u}^{\xi}(z,t) &= \sum_{k=1}^{r}\lambda_k\frac{\partial^2}{\partial z_k^2}\hat{u}^{\xi}(z,t) - (N-r-2)\sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}u(z,t)\\
&\quad - (N-r-2)\sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)]\\
&\quad + \sum_{k=r+1}^{N}\sum_{l=k+1}^{N}\left(\lambda_k\frac{\partial^2}{\partial z_k^2} + \lambda_l\frac{\partial^2}{\partial z_l^2}\right)[u^{\{1,\dots,r,k,l\}}(z,t) - u(z,t)]\\
&\quad + (N-r-1)\sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}u(z,t) - \sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}u(z,t)\\
&= \sum_{k=1}^{r}\lambda_k\frac{\partial^2}{\partial z_k^2}\hat{u}^{\xi}(z,t)\\
&\quad - (N-r-2)\sum_{k=r+1}^{N}\lambda_k\frac{\partial^2}{\partial z_k^2}[u^{\{1,\dots,r,k\}}(z,t) - u(z,t)]\\
&\quad + \sum_{k=r+1}^{N}\sum_{l=k+1}^{N}\left(\lambda_k\frac{\partial^2}{\partial z_k^2} + \lambda_l\frac{\partial^2}{\partial z_l^2}\right)[u^{\{1,\dots,r,k,l\}}(z,t) - u(z,t)]
\end{aligned}
$$

In other words, $\hat{u}^{\xi}$ solves a non-homogenous r-dimensional heat equation with source term

$$f(z,t) = \sum_{k=r+1}^{N} \sum_{l=k+1}^{N} \left(\lambda_k \frac{\partial^2}{\partial z_k^2} + \lambda_l \frac{\partial^2}{\partial z_l^2} \right) [u^{\{1,\ldots,r,k,l\}}(z,t) - u(z,t)]$$
$$- (N-r-2) \sum_{k=r+1}^{N} \lambda_k \frac{\partial^2}{\partial z_k^2} [u^{\{1,\ldots,r,k\}}(z,t) - u(z,t)]$$

We know from the proof of the first order bound that

$$\begin{aligned}
\hat{u}^{\{1,\ldots,r,k\}}(z,t) &:= u^{\{1,\ldots,r,k\}}(z,t) - u(z,t) \\
&= -\int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\ldots,r,k\}}(y,s) \\
&\qquad \cdot \sum_{i=r+1, i\neq k}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z-y, t-s) dy ds \\
\hat{u}^{\{1,\ldots,r,k,l\}}(z,t) &:= u^{\{1,\ldots,r,k,l\}}(z,t) - u(z,t) \\
&= -\int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\ldots,r,k,l\}}(y,s) \\
&\qquad \cdot \sum_{i=r+1, i\neq k,l}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u(z-y, t-s) dy ds
\end{aligned}$$

This allows us to expand the expression for f. We leave out the coordinates to increase readability.

$$\begin{aligned}
f = &- \sum_{k=r+1}^{N} \sum_{l=k+1}^{N} \left(\lambda_k \frac{\partial^2}{\partial z_k^2} + \lambda_l \frac{\partial^2}{\partial z_l^2} \right) \\
&\quad \cdot \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k,l\}} \sum_{i=r+1, i\neq k,l}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u \\
&+ (N-r-2) \sum_{k=r+1}^{N} \lambda_k \frac{\partial^2}{\partial z_k^2} \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k\}} \sum_{i=r+1, i\neq k}^{N} \lambda_i \frac{\partial^2}{\partial z_i^2} u \\
= &- \sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \\
&\quad \cdot \int_0^T \int_{\mathbb{R}^N} \lambda_k \lambda_i \Phi^{\{1,\dots,r,k,l\}} \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u + \lambda_l \lambda_i \Phi^{\{1,\dots,r,k,l\}} \frac{\partial^4}{\partial z_l^2 \partial z_i^2} u \\
&+ (N-r-2) \sum_{k=r+1}^{N} \sum_{i=r+1, i\neq k}^{N} \int_0^T \int_{\mathbb{R}^N} \lambda_k \lambda_i \Phi^{\{1,\dots,r,k\}} \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u
\end{aligned}$$

Both lines contain $(N-r)(N-r-1)(N-r-2)$ individual terms. This allows us to combine them into

$$\begin{aligned}
f = &- \sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \int_0^T \int_{\mathbb{R}^N} \lambda_k \lambda_i [\Phi^{\{1,\dots,r,k,l\}} - \Phi^{\{1,\dots,r,k\}}] \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u \\
&+ \lambda_l \lambda_i [\Phi^{\{1,\dots,r,k,l\}} - \Phi^{\{1,\dots,r,l\}}] \frac{\partial^4}{\partial z_l^2 \partial z_i^2} u
\end{aligned}$$

This transformation is guided by the inutition that a difference involving a convolution with $\Phi^{\{1,\dots,r,k,l\}} - \Phi^{\{1,\dots,r,k\}}$ should lead to a factor of order $\sim \lambda_l T$. To derive this formally, note that

$$v_1^l := \int_{\mathbb{R}} \Phi^{\{l\}} \int_{\mathbb{R}^{N-1}} \Phi^{\{1,\dots,r,k\}}] \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u$$

is the solution of the 1-dimensional heat equation

$$\frac{\partial}{\partial t} v_1^l = \lambda_l \frac{\partial^2}{\partial z_l^2} v_1^l$$

with the initial condition given by the inner integral over $\mathbb{R}^{N-1}$. Call v_2^l the corresponding solution of $\frac{\partial}{\partial t} v_2^l = 0$, i.e.,

$$v_2^l := \int_{\mathbb{R}^{N-1}} \Phi^{\{1,\dots,r,k\}}] \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u$$

Then their difference $\hat{v}^l = v_2^l - v_1^l$ solves

$$\frac{\partial}{\partial t} \hat{v}^l = \frac{\partial}{\partial t} \hat{v_1^l} - \frac{\partial}{\partial t} \hat{v_2^l} = \lambda_l \frac{\partial^2}{\partial z_l^2} v_1^l$$

and can thus be written as

$$\begin{aligned}
\hat{v}^l &= \int_0^T \lambda_l \frac{\partial^2}{\partial z_l^2} v_1^l \\
&= \int_0^T \lambda_l \frac{\partial^2}{\partial z_l^2} \int_{\mathbb{R}} \Phi^{\{l\}} \int_{\mathbb{R}^{N-1}} \Phi^{\{1,\dots,r,k\}}] \frac{\partial^4}{\partial z_k^2 \partial z_i^2} u \\
&= \lambda_l \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k,l\}}] \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} u
\end{aligned}$$

Plugging this into the formula for the source term f gives

$$\begin{aligned}
f &= -\sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \int_0^T \lambda_k \lambda_i \hat{v}^l + \lambda_l \lambda_i \hat{v^k} \\
&= -2 \sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \lambda_k \lambda_l \lambda_i \int_0^T \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k,l\}}] \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} u
\end{aligned}$$

Now, as in the first order case, we are finally able to give a bound for the size of the expansion error:

$$\begin{aligned}
\left|\hat{u}^{\xi}(z,t)\right| &= \left|\int_0^T \int_{\mathbb{R}^r} \Phi^{\{1,\dots,r\}} f\right| \\
&= \left|2 \sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \lambda_k\lambda_l\lambda_i \int_0^T \int_{\mathbb{R}^r} \Phi^{\{1,\dots,r\}} \right. \\
&\qquad \left. \cdot \int_0^T \int_0^T \int_{\mathbb{R}^N} \Phi^{\{1,\dots,r,k,l\}}] \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} u\right| \\
&\leq \frac{t^3}{3} \sum_{l>k\geq r+1}^{N} \sum_{i=r+1, i\neq k,l}^{N} \lambda_k\lambda_l\lambda_i \left\| \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} g \right\|_{\infty} \\
&= \frac{t^3}{6} \sum_{l,k\geq r+1, l\neq k}^{N} \sum_{i=r+1, i\neq k,l}^{N} \lambda_k\lambda_l\lambda_i \left\| \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} g \right\|_{\infty} \\
&= t^3 \sum_{i>l>k\geq r+1}^{N} \lambda_k\lambda_l\lambda_i \left\| \frac{\partial^6}{\partial z_k^2 \partial z_l^2 \partial z_i^2} g \right\|_{\infty}
\end{aligned}$$

□

Appendix C

Data

C.1 Further Numerical Results for LIBOR Derivatives

N	V_{MC}^{-}	σ	$V_{MC}^{-}+\Delta_0$	σ_{Δ_0}	$V_{MC}^{-}+\Delta_0/2$	V_{PDE}	Δ_{abs}	Δ_{rel}
5	1.16E-02	2.36E-12	1.16E-02	0.00E+00	1.16E-02	1.16E-02	-8.73E-15	0.00%
11	2.38E-02	1.29E-12	2.38E-02	4.74E-06	2.38E-02	2.38E-02	-1.18E-05	-0.05%
21	4.59E-02	1.34E-05	4.68E-02	4.89E-05	4.63E-02	4.67E-02	4.15E-04	0.90%
41	8.59E-02	2.13E-05	8.93E-02	1.21E-04	8.76E-02	8.93E-02	1.70E-03	1.94%
5	1.16E-02	2.36E-12	1.16E-02	0.00E+00	1.16E-02		-8.73E-15	0.00%
11	2.38E-02	1.29E-12	2.38E-02	4.80E-06	2.38E-02		-1.37E-05	-0.06%
21	4.59E-02	1.34E-05	4.68E-02	5.08E-05	4.64E-02		3.72E-04	0.80%
41	8.59E-02	2.15E-05	8.95E-02	1.29E-04	8.77E-02		1.64E-03	1.87%

Table C.1: PDE results for ITM ($K = 0.09$) Bermudan swaptions compared to frozen (top) and full (bottom) drift MC results.

N	V_{MC}^{-}	σ	$V_{MC}^{-}+\Delta_0$	σ_{Δ_0}	$V_{MC}^{-}+\Delta_0/2$	V_{PDE}	Δ_{abs}	Δ_{rel}
5	9.39E-04	6.78E-07	9.39E-04	0.00E+00	9.39E-04	9.49E-04	9.97E-06	1.06%
11	7.01E-03	4.15E-06	7.08E-03	6.51E-06	7.04E-03	7.30E-03	2.57E-04	3.64%
21	2.02E-02	9.60E-06	2.17E-02	6.29E-05	2.10E-02	2.11E-02	1.01E-04	0.48%
41	4.50E-02	1.74E-05	5.12E-02	1.77E-04	4.81E-02	4.91E-02	9.90E-04	2.06%
5	9.39E-04	6.78E-07	9.39E-04	0.00E+00	9.39E-04		1.02E-05	1.09%
11	7.03E-03	4.16E-06	7.11E-03	6.70E-06	7.07E-03		2.30E-04	3.25%
21	2.03E-02	9.66E-06	2.19E-02	6.31E-05	2.11E-02		-4.85E-05	-0.23%
41	4.55E-02	1.77E-05	5.22E-02	1.85E-04	4.88E-02		2.62E-04	0.54%

Table C.2: PDE results for OTM ($K = 0.11$) Bermudan swaptions compared to frozen (top) and full (bottom) drift MC results.

(a,b,c)	N	V_{MC}	σ	V_{PDE}	Δ_{abs}	Δ_{rel}
$K_1 = 0.10$						
0/1/0	5	7.08E-003	2.95E-006	7.04E-003	-3.25E-05	-0.46%
	11	9.63E-003	3.86E-006	9.61E-003	-2.18E-05	-0.23%
	21	1.06E-002	4.17E-006	1.07E-002	9.78E-05	0.92%
0.2/0.9/0	5	3.06E-002	4.23E-006	3.06E-002	-6.80E-05	-0.22%
	11	4.94E-002	4.73E-006	4.93E-002	-1.04E-04	-0.21%
	21	5.10E-002	5.02E-006	5.09E-002	-1.06E-04	-0.21%
0.25/0.95/-0.01	5	3.29E-002	4.03E-006	3.29E-002	-6.26E-05	-0.19%
	11	6.06E-002	5.13E-006	6.06E-002	-2.65E-05	-0.04%
	21	7.37E-002	8.14E-006	7.36E-002	-1.01E-04	-0.14%
$K_1 = 0.11$						
0/1/0	5	1.27E-002	3.88E-006	1.26E-002	-5.16E-05	-0.41%
	11	1.44E-002	4.73E-006	1.43E-002	-6.55E-05	-0.45%
	21	1.44E-002	4.91E-006	1.44E-002	1.01E-05	0.07%
0.2/0.9/0	5	3.63E-002	4.33E-006	3.63E-002	-6.70E-05	-0.18%
	11	5.21E-002	4.76E-006	5.20E-002	-1.04E-04	-0.20%
	21	5.18E-002	5.02E-006	5.17E-002	-9.75E-05	-0.19%
0.25/0.95/-0.01	5	4.00E-002	4.11E-006	4.00E-002	-6.16E-05	-0.15%
	11	6.52E-002	5.12E-006	6.52E-002	-2.53E-05	-0.04%
	21	7.58E-002	8.09E-006	7.57E-002	-1.03E-04	-0.14%
$K_1 = 0.09$						
0/1/0	5	3.19E-003	1.94E-006	3.17E-003	-1.81E-05	-0.57%
	11	5.82E-003	2.94E-006	5.82E-003	-2.70E-06	-0.05%
	21	7.33E-003	3.40E-006	7.47E-003	1.41E-04	1.92%
0.2/0.9/0	5	2.51E-002	4.07E-006	2.50E-002	-5.99E-05	-0.24%
	11	4.68E-002	4.70E-006	4.67E-002	-1.14E-04	-0.24%
	21	5.03E-002	5.02E-006	5.02E-002	-1.00E-04	-0.20%
0.25/0.95/-0.01	5	2.59E-002	3.88E-006	2.58E-002	-5.57E-05	-0.22%
	11	5.61E-002	5.14E-006	5.60E-002	-3.09E-05	-0.06%
	21	7.15E-002	8.20E-006	7.15E-002	-6.08E-05	-0.09%

Table C.3: PDE results for Ratchet floors compared to frozen drift MC results. The column V_{PDE} shows the computed PDE value. Columns Δ_{abs} and Δ_{rel} show the absolute and relative difference to the MC estimate V_{MC}, resp.

Bibliography

[1] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing*. Philadelphia, 2005.

[2] L. Andersen and M. Broadie. Primal-dual simulation algorithm for pricing multidimensional American options. *Man. Sci.*, 50:1222–1234, 2004.

[3] A. D. Bandrauk, M. C. Delfour, and C. Le Bris. *High-dimensional partial differential equations in science and engineering*, volume 41. 2007.

[4] A. Brace, D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Math. Fin*, 7(2):127–154, 1997.

[5] P. L. T. Brian. A finite-difference method of higher-order accuracy for the solution of three-dimensional transient heat conduction problems. *AIChE J.*, 7:367–370, 1961.

[6] M.J. Chang, L.C. Chow, and W.S. Chang. Improved alternating-direction implicit method for solving transient three-dimensional heat diffusion problems. *Numerical Heat Transfer, Part B: Fundamentals*, 19:1:69–84, 1991.

[7] S. Cheng. Differentiation under the integral sign with weak derivatives, 2010. http://www.gold-saucer.org/math/diff-int/diff-int.pdf and http://planetmath.org/differentiationundertheintegralsign.

[8] C. C. Christara and D. M. Dang. Adaptive and high-order methods for valuing American options. *J. of Comp. Fin.*, 14(4):73, 2011.

[9] V. V. Desai, V. F. Farias, and C. C. Moallemi. Pathwise optimisation for optimal stopping problems. *Man. Sci.*, 58(12):2292–2308, 2012.

[10] Jr. Douglas. Alternating direction method for three space variables. *Numerische Mathematik*, 4:41–63, 1962.

[11] C. Feuersänger. Sparse grid methods for higher dimensional approximation. PhD thesis, Universität Bonn, 2010.

[12] P. A. Forsyth and K. R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6):2095–2122, 2002.

[13] A. Friedman. *Partial Differential Equations Of Parabolic Type*. 1964.

[14] C. Fries. *Mathematical Finance: Theory, Modeling, Implementation*. Hoboken, New Jersey, 2007.

[15] B. R. Gelbaum and J. M. H. Olmsted. *Counterexamples in Analysis*. 1992.

[16] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. New York, 2003.

[17] M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In L. Pardo, A. Pinkus, E. Süli, and M. Todd, editors, *Foundations of Computational Mathematics*, pages 106–161, 2006.

[18] M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *J. Compl.*, 26(5):455–489, 2010.

[19] M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *J. Complex.*, 26(5):455–489, 2010.

[20] M. Griebel, F. Y. Kuo, and I. H. Sloan. The smoothing effect of the anova decomposition. *J. Complex.*, 26(5):523–551, 2010.

[21] A. Heinecke, S. Schraufstetter, and H.-J. Bungartz. A highly parallel Black-Scholes solver based on adaptive sparse grids. *Int. J. Comp. Math.*, 89(9):1212–1238, 2012.

[22] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993.

[23] N. Hilber, S. Kehtari, C. Schwab, and C. Winter. Wavelet finite element method for option pricing in highdimensional diffusion market models. Technical Report 2010–01, SAM, ETH Zürich, 2010.

[24] J. Imai and K. S. Tan. A general dimension reduction technique for derivative pricing. *J. Comp. Fin.*, 10(2):129–155, 2006.

[25] K.J. in't Hout and S. Foulon. ADI finite difference schemes for option pricing in the Heston model with correlation. *International J. of Numerical Analysis and Modeling*, 7:2:303–320, 2010.

[26] Jr. J. Douglas. On the numerical integration of $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = \partial u/\partial t$ by implicit methods. *J. Soc. Indust. Appl. Math.*, 3:42–65, 1955.

[27] B. D. Jones. *The Theory of Generalized Functions*. 1982.

[28] R. P. Kanwal. *Generalized Functions: Theory and Technique*. 1983.

[29] A. Kolodko and J. Schoenmakers. Iterative construction of the optimal Bermudan stopping time. *Fin. Stoch.*, 10:27–49, 2006.

[30] K. Königsberger. *Analysis 2*. 2004.

[31] O. Kurbanmuradov, K. Sabelfeld, and J. Schoenmakers. Lognormal random field approximations to LIBOR market models. *J. of Comp. Fin.*, 2002.

[32] C. C. W. Leentvaar and C. W. Oosterlee. On coordinate transformation and grid stretching for sparse grid pricing of basket options. *J. Comput. Appl. Math.*, 222:193–209, 2008.

[33] S. K. Lele. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.*, 103:16–42, 1992.

[34] K. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with log-normal interest rates. *J. of Finance*, 52(1):409–430, 1997.

[35] A. Pascucci, M. Suárez-Taboada, and C. Vázquez. Mathematical analysis and numerical methods for a PDE model governing a rachet-cap pricing in the Libor market model. *Math. Mod. Meth. Appl. Sci. (M3AS)*, 7(21):1479 – 1498, 2011.

[36] D. W. Peaceman and H. H. Rachford, Jr. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3:28–41, 1955.

[37] H. Rabitz and O. Alis. General foundations of high-dimensional model representations. *J. Math. Chem.*, 25:197–233, 1999.

[38] C. Reisinger. Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben. PhD thesis, Universität Heidelberg, 2004.

[39] C. Reisinger. Asymptotic expansion around principal components and the complexity of dimension adaptive algorithms. In J. Garcke and M. Griebel, editors, *Sparse Grids and Applications*, number 88 in Springer Lectures Notes in Computational Science and Engineering, pages 263–276, 2012.

[40] C. Reisinger and R. Wissmann. Numerical valuation of derivatives in high-dimensional settings via PDE expansions. 2013. To appear in J. of Comp. Fin.

[41] C. Reisinger and G. Wittum. Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM J. Sci. Comp.*, 29(1):440–458, 2007.

[42] P. Schröder, T. Gerstner, and G. Wittum. Taylor-like ANOVA-expansion for high dimensional problems in finance. Working paper, 2012.

[43] L. Schwartz. Theories des distributions. 1950.

[44] D. Tavella and C. Randall. *Pricing Financial Instruments: The Finite Difference Method*. New York, 2000.

[45] X. Wang. On the approximation error in high dimensional model representation. In S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, editors, *Winter Simulation Conference, 2008*, pages 453–462, 2008.

[46] P. Wilmott, S. Howison, and J. Dewynne. *The Mathematics of Financial Derivatives*. 1995.

[47] Z. Zhang, M. Choi, and G. E. M. Karniadakis. Error estimates for the ANOVA method with polynomial chaos interpolation: tensor product functions. *SIAM J. Sci. Comput.*, 34(2):1165–1186, 2012.

[48] Y.-L. Zhu and J. Li. Multi-factor financial derivatives on finite domains. *Comm. Math. Sci.*, 1(2):343–359, 2003.