Diss. ETH No. 20202

Numerical option pricing beyond Lévy

A dissertation submitted to
ETH Zürich

for the degree of
Doctor of Sciences

presented by
Oleg Reichmann
Dipl. Math. WWU Münster
Dipl. Kauf. WWU Münster
born June 21, 1985
citizen of Germany and Russia

accepted on the recommendation of
Prof. Dr. Christoph Schwab, ETH Zürich, examiner
Prof. Dr. Erich Walter Farkas, University of Zürich, co-examiner
Prof. Dr. Fred Espen Benth, University of Oslo, co-examiner

2012

ISBN 978-3-906031-01-9

To my parents

Dank

Die vorliegende Arbeit ist während meiner Assistenzzeit am Seminar für angewandte Mathemathik unter der Leitung meines Doktorvaters Professor Christoph Schwab entstanden. Ich möchte ihm an dieser Stelle für die vielen Dinge, mathematischer und nicht-mathematischer Art, die er mich gelehrt hat, danken und für das große Vertrauen, das er mir immer entgegengebrachte. Ebenfalls danke ich Professor Erich Walter Farkas und Professor Fred Espen Benth für die Übernahme des Koreferats.

Ein besonderer Dank geht an Dr. Norbert Hilber und Dr. Christoph Winter, ehemalige Mitglieder der Computational Finance Gruppe von Professor Schwab, sie haben mir den Einstieg in mein Dissertationsthema durch ihre Vorarbeit und ihre Geduld sehr einfach gemacht.

Ich danke Andrea und Annika für zahlreiche mathematische und sonstige Gespräche, Roman nicht nur für die vielen Diskussionen zu zeitinhomogenen PDEs, Christine für das viele Korrekturlesen, Holger senior für die zahlreichen Unterhaltungen über Advektionsgleichungen, Sid, der weit gereist ist um zu meiner Verteidigung zu kommen, Sohrab, Dan und den Fjordholms.
Ferner danke ich allen ehemaligen und aktuellen Mitarbeitern des Seminars für angewandte Mathematik, die die letzten drei Jahre zu einer sehr schonen Zeit für mich gemacht haben.

Schließlich gilt mein Dank meiner Familie und Franzi, ohne sie wäre all dies nicht möglich.

Abstract

This work considers the numerical approximation of option prices in different market models beyond Lévy processes. The Lévy setup is extended in several directions. The arising partial integrodifferential equations and inequalities are solved with the finite element method. European as well as American type contracts are considered.

Spatially inhomogeneous market models are analyzed, specifically certain Feller processes are considered. The well-posedness of the arising pricing equations is proved using pseudodifferential operator theory. The resulting pricing equations need no longer be parabolic and can exhibit degeneracies under certain conditions. Classical continuous Galerkin methods are therefore inapplicable for the numerical solution of the corresponding pricing equations. Thus we employ a discontinuous Galerkin discretization or alternatively a streamline diffusion approach. Convergence results are shown in both cases.

Besides the spatial inhomogeneity, also the assumption of temporal homogeneity of the coefficients of the partial integrodifferential equations is weakened. The well-posedness for pricing equations with degenerate coefficients in time is shown via a weak space-time formulation. The main problem arising in the discretization of such equations is the non-applicability of classical time-marching schemes due to the possible degeneracy of the coefficients. Therefore two alternative approaches are considered. First, a continuous Galerkin method for the space-time discretization is used, in this case optimality of the solution algorithm can be shown. Second, a discontinuous Galerkin discretization for the temporal domain is studied, in which case exponential convergence of the algorithm can be shown.

Numerical examples are given to confirm the theoretical results. Partial integrodifferential equations with spatially as well as temporally inhomogeneous coefficients are solved numerically. European and American options are priced.

Zusammenfassung

Die vorliegende Arbeit befasst sich mit der numerischen Approximation von Optionspreisen in verschiedenen Marktmodellen, welche über Lévy Prozesse hinausgehen. Lévy Modelle werden in mehrerlei Hinsicht erweitert. Die entsprechenden partiellen Integro-Differentialgleichungen werden mit der Finiten Elemente Methode gelöst. Es werden sowohl europäische als auch amerikanische Optionen betrachtet.

Marktmodelle mit inhomogenen Koeffizienten bezüglich der Zustandsvariablen werden analysiert, wobei insbesondere bestimmte Feller Prozesse behandelt werden. Die Wohlgestelltheit der entsprechenden Bewertungsgleichungen wird mit Hilfe von Pseudodifferentialoperator Theorie gezeigt. Die entsprechenden Bewertungsgleichungen sind in der Regel nicht parabolisch und können unter Umständen entarten, weshalb klassische Diskretisierungsmethoden zur Lösung der Bewertungsgleichungen nicht anwendbar sind. Es werden daher eine unstetige Galerkin-Diskretisierung und ein streamline diffusion Ansatz verwendet. In beiden Fällen werden Konvergenzresultate hergeleitet.

Neben der örtliche Inhomogenität wird auch die Annahme der zeitlichen Homogenität der Koeffizienten der partiellen Integro-Differentialgleichungen abgeschwächt. Die Wohlgestelltheit der Bewertungsgleichungen mit entarteten Koeffizienten bezüglich der Zeit kann mit Hilfe einer schwachen Orts-Zeit-Formulierung bewiesen werden. Die größte Herausforderung bei der Zeitdiskretisierung solcher Gleichungen besteht darin, dass klassische Zeitschrittverfahren aufgrund der Entartungen der Koeffizienten nicht verwendet werden können. Daher werden zwei alternative Ansätze betrachtet. Einerseits wird eine stetige Galerkin-Methode für die Orts-Zeit-Diskretisierung verwendet, in diesem Fall wird die Optimalität des Lösungsalgorithmus bewiesen. Andererseits wird eine unstetige Galerkin-Methode für die Zeitdiskretisierung untersucht und exponentielle Konvergenz des Algorithmus gezeigt.

Numerische Bespiele belegen die theoretischen Resultate. Partielle Integro-Differentialgleichungen mit inhomogenen Koeffizienten sowohl bezüglich der Zeit als auch der Zustandsvariablen werden numerisch gelöst. Es werden europäische und amerikanische Optionen bewertet.

Contents

1 Introduction

According to Webster's dictionary, an option is the power or right to choose. This thesis is concerned with the numerical valuation of options in financial markets.

Options written on a certain underlying, e.g., stocks, indices, coal, gas or electricity, give their holder *the right, but not the obligation* to perform certain transactions, such as buying or selling the underlying, at certain times in the future. The evaluation of options has become an increasingly important field of research and practice since the seminal works of F. Black and M. Scholes (1973) [16], B. Mandelbrot (1963) [82], R.C. Merton (1973) [87], P.A. Samuelson (1969) [102]. The use of models seems unavoidable in order to approximate option prices in many cases, as their payoff depends on the (uncertain) future evolution of the underlying and since pure hedging arguments can often not be employed.

Many models have been proposed to describe the evolution of the underlying. One distinguishes between equilibrium type approaches, which are characterized by the equilibrium between supply and demand of the underlying at every instant in time. An other approach is to directly model the evolution of the underlying. We pursue the latter approach. Stochastic processes, especially semimartingales, have gained huge popularity as models for the evolution of the underlying over the few last decades, due to their flexibility and tractability from the stochastic point of view. Typical examples of market models are the Black-Scholes model (1973) [16], the CEV model (1975) [35] or the CGMY model (2002) [26]. All these models exhibit certain types of homogeneities in the evolution of the underlying, which leads to models that are easier to handle thus simplifying the pricing of options significantly. In the Black-Scholes model, as well as in the CGMY model a spatial and temporal homogeneity is assumed, i.e., the increments of the process are stationary and indepedent, while the CEV model and, for example, many local and stochastic volatility models are only temporally homogeneous.

Despite the huge popularity of these models, real data shows that such homogeneity assumptions may be too restrictive, especially for underlyings more exotic than standard stocks, such as gas, carbon or electricity. Practitioners solve this conflict by frequent recalibration and ad hoc implementation of non-constant parameters. The aim of this work is to contribute to the literature on option pricing beyond Lévy models. The Lévy setup is extended in several directions. Spatially inhomogeneous market models are analyzed, specifically certain Feller processes are considered. The resulting pricing equations need no longer be parabolic, but can exhibit degeneracies and hyperbolic behavior under certain conditions. Therefore, classical continuous Galerkin methods are inapplicable for the numerical solution of the corresponding pricing equations. Besides the spatial inhomogeneity, also the assumption of temporal homogeneity of the process is weakened. Pricing equations with degenerate coefficients in time are solved numerically. The main problem arising for such equations is the non-applicability of classical time-marching schemes due to the possible degeneracy of the coefficients. Therefore, appropriate (weak) space-time formulations are considered. Summarizing, the three main problems tackled in this thesis are well-posedness, discretization and numerical analysis of discretization schemes for

- Partial integro-differential equations with spatially inhomogeneous jump measures,
- Drift dominated partial integro-differential equations,
- Partial (integro-)differential equations with weakly degenerate coefficients in time.

A semimartingale can be well-understood via its semimartingale characteristic, this stochastic triplet describes the law of a semimartingale completely, accounting for the drift component, the diffusion component and the jump component, cf. [72, 88]. We focus on a subclass of semimartingales with time and state-space dependent triplets. The generators of these processes are certain, nonclassical pseudodifferential operators whose symbols are contained in certain symbol classes and the corresponding Dirichlet forms have variable order Sobolev spaces as their domains. The well-posedness of the arising pricing equations can be proved for time-homogeneous models using pseudodifferential operator theory and for time-inhomogeneous equations via a weak space-time formulation. Localization of the pricing equations plays a crucial rule in our algorithm as all equations are solved on a bounded spatial domain and thus an error analysis with respect to the size of the computational domain is crucial. As the pricing equations can exhibit hyperbolic behavior for certain parameters, this has to be considered in the discretization. It is well-known that standard continuous finite elements lead to unstable algorithms for transport dominated equations, therefore we employ a discontinuous Galerkin (DG) discretization or a streamline diffusion approach. As the DG discretization is not applicable to certain integrodifferential operators, we use a small jump approximation in those cases. From the probabilistic point of view, this corresponds to the approximation of an infinite activity pure jump process by a finite activity jump process and an appropriately scaled diffusion. For time-inhomogeneous equations, two approaches are analyzed. First, a continuous Galerkin (CG) method for the space-time discretization is considered, in this case optimality of the solution algorithm can be shown. Second, a discontinuous Galerkin discretization for the temporal domain is studied, in which case exponential convergence of the algorithm can be shown.

Spatially inhomogeneous market models have been studied by several authors, see, e.g., [8, 24]. In [24] a class of local Lévy models is defined, in correspondence to local volatility models in the sense of Dupire, cf. [47]. The pricing is based on a Fourier type approach

for European contracts. In [8] a general state-space and time dependent pricing equation is derived. There, it is shown that such equations arise in option pricing of European options under general discontinuous semimartingales due to a Markovian projection in the sense of Gyöngy, cf. [58]. Following [83, 97, 98] we consider finite element discretizations of weak formulations for a general class of admissible market models. The multidimensional jump measures are constructed via a Lévy copula approach.
For a theoretical foundation of temporally inhomogeneous models, also known as Sato processes in the literature, we refer to [103, 77]. Applications are described in [21, 25]. Following [96] we propose efficient time-discretizations for time-inhomogeneous integro-differential equations with possibly degenerate coefficients.

This thesis is structured as follows. In Chapter 2 the preliminaries are outlined. We define pseudodifferential operators and certain classes of symbols. Variable order Sobolev spaces needed for later analysis are also introduced. Subsequently, examples of market models are presented. Time-homogeneous and inhomogeneous admissible market models are defined and the construction of Feller processes via subordination is discussed. In Chapter 4, the small jump regularization and localization of Markov processes is examined. The small jump regularization is a method for the approximation of infinite activity jump processes by finite activity processes and possibly an appropriately scaled diffusion. This purely probabilistic tool plays an important role in the discretization of the pricing equation, as it enables us to use discontinuous basis functions for the discretization in state space. Additionally, localization estimates for the pricing equations are presented, rigorously justifying the approximation of the pricing equations on a bounded domain. In Chapter 5 different well-posedness results are proved for processes with spatially inhomogeneous jump measures with or without drift dominance. Moreover, an analysis of partial integro-differential equations arising from the small jump regularization is given. In the next chapter the triangulation and the choice of wavelet bases are outlined. It turns out that Riesz bases for the domains of the generators for some market models can be constructed, allowing for efficient solution of the equations and efficient preconditioning. Chapter 7 provides a rigorous error analysis for the various types of time-homogeneous partial integro-differential equations considered in this thesis, CG schemes with and without streamline diffusion are analyzed as well as DG schemes with and without small jump regularization. In the subsequent chapter the numerical quadratures and the time stepping scheme are discussed. For the time-stepping the θ-scheme is proposed. Chapter 9 provides well-posedness results for time-inhomogeneous pricing equations with weakly degenerate coefficients. In the following chapter the results of Chapters 7 and 8 are extended to the time-inhomogeneous setting. Space-time discretizations using CG and DG in time are analyzed. American options for time-homogeneous models are addressed subsequently. Finally, numerical results and an outlook are given.

2 Preliminaries

In this chapter definitions and basic results that are needed throughout the thesis are given. We introduce pseudodifferential operators and variable order Sobolev spaces. Pseudodifferential operators are generators of the market models considered here and variable order Sobolev spaces correspond to the domains of the arising Dirichlet forms. Besides the definition of a Lévy copula is given, which is crucial for parametric constructions of multidimensional jump measures.

2.1 Pseudodifferential operators

The following section introduces a class of stochastic processes that are considered in this work. It turns out that the processes are characterized via the symbol of their generator. Throughout we consider adapted stochastic processes X on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfying the usual assumptions, see [93]. Semimartingales are a well-investigated class of stochastic processes that is sufficiently rich to include most of the stochastic processes commonly employed in financial modeling while still being closed under various operations such as conditional expectations and stopping. Semimartingales can be well understood via their (generally stochastic) semimartingale characteristic, we refer to the standard reference [72] for details. Here, we restrict ourselves to a class of processes with deterministic, but generally time- and state-space dependent characteristic triplets including Lévy processes, affine processes and many local volatility models. The time-homogeneous case is analyzed in the first part of this section, while extensions to certain types of time-inhomogeneity are discussed in a second part.

2.1.1 Time-homogeneous processes

We consider a Markov process X and the corresponding family of linear operators $(T_{s,t})$ for $0 \leq s \leq t < \infty$ given by

$$(T_{s,t}(f))(x) \quad = \quad \mathbb{E}[f(X(t))|X(s) = x],$$

for each $f \in B_b(\mathbb{R}^d)$, $x \in \mathbb{R}^d$. Here $B_b(\mathbb{R}^d)$ denotes the space of bounded Borel measurable functions on $\mathbb{R}^d$. In the following we call Markov processes, whose semigroups satisfy

$$T_{s,t}(B_b(\mathbb{R}^d)) \subset B_b(\mathbb{R}^d), \tag{2.1}$$

normal and recall the following properties:

(i) $T_{s,t}$ is a linear operator on $B_b(\mathbb{R}^d)$ for each $0 \le s \le t < \infty$.

(ii) $T_{s,s} = I$ for each $s \ge 0$.

(iii) $T_{r,s}T_{s,t} = T_{r,t}$, whenever $0 \le r \le s \le t < \infty$.

(iv) $f \ge 0$ implies $T_{s,t}f \ge 0$ for all $0 \le s \le t < \infty$ and $f \in B_b(\mathbb{R}^d)$.

(v) $\|T_{s,t}\| \le 1$ for each $0 \le s \le t < \infty$, i.e. $T_{s,t}$ is a contraction.

(vi) $T_{s,t}(1) = 1$ for all $t \ge 0$.

If we restrict ourselves to time-homogeneous Markov processes satisfying (2.1), we obtain directly from the above properties that the family of operators $T_t := T_{0,t}$ forms a positivity preserving contraction semigroup. The *infinitesimal generator* $\mathcal{A}$ with domain $\mathcal{D}(\mathcal{A})$ of such a process X with semigroup $(T_t)_{t\ge0}$ is defined by the strong pointwise limit

$$\mathcal{A}u := \lim_{t\to0^+} \frac{1}{t}\left(T_t u - u\right) \tag{2.2}$$

for all functions $u \in \mathcal{D}(\mathcal{A}) \subset B_b(\mathbb{R}^d)$ for which the limit (2.2) exists with respect to the sup-norm. We call $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ *generator* of X. Generators of normal Markov processes admit the *positive maximum principle*, i.e.,

$$\text{if} \quad u \in \mathcal{D}(\mathcal{A}) \quad \text{and} \quad \sup_{x\in\mathbb{R}^d} u(x) = u(x_0) > 0, \quad \text{then} \quad (\mathcal{A}u)(x_0) \le 0. \tag{2.3}$$

Furthermore, they admit a pseudodifferential representation (e.g. [17, 34, 69, 70]):

Theorem 2.1.1. *Let $\mathcal{A}$ be an operator with domain $\mathcal{D}(\mathcal{A})$, where $C_0^\infty(\mathbb{R}^d) \subset \mathcal{D}(\mathcal{A})$ and $\mathcal{A}(C_0^\infty(\mathbb{R}^d)) \subset C(\mathbb{R}^d)$, where $C_0^\infty(\mathbb{R}^d)$ denotes the space of smooth functions with support compactly contained in $\mathbb{R}^d$. Then $\mathcal{A}|_{C_0^\infty(\mathbb{R}^d)}$ is a pseudodifferential operator,*

$$(\mathcal{A}u)(x) := -(2\pi)^{-d/2} \int_{\mathbb{R}^d} \psi(x,\xi)\hat{u}(\xi)e^{ix\cdot\xi}d\xi, \tag{2.4}$$

where $u \in C_0^\infty(\mathbb{R}^d)$ and $\hat{u}$ is the Fourier transform of u. The symbol $\psi(x,\xi) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is locally bounded in (x,ξ). The function $\psi(\cdot,\xi)$ is measurable for every ξ and $\psi(x,\cdot)$ is a negative definite function for every x, which admits the Lévy-Khintchine representation

$$\begin{aligned} \psi(x,\xi) \quad &= \quad c(x) - ib(x)\cdot\xi + \frac{1}{2}\xi\cdot Q(x)\xi \\ &\quad + \int_{0\neq z\in\mathbb{R}^d} \left(1 - e^{iz\cdot\xi} + \frac{iz\cdot\xi}{1+|z|^2}\right)\nu(x,dz). \end{aligned} \tag{2.5}$$

Here $c : \mathbb{R}^d \to \mathbb{R}_+$, $b : \mathbb{R}^d \to \mathbb{R}^d$ and $Q : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are functions, $Q(x)$ is symmetric positive semidefinite for all $x \in \mathbb{R}$ and $\nu(x,\cdot)$ is a measure on $\mathbb{R}^d$ for fixed $x \in \mathbb{R}$ with

$$\int_{z\neq0} (1 \wedge |z|^2)\nu(x,dz) < \infty. \tag{2.6}$$

The tuple $(c(x), b(x), Q(x), \nu(x, dz))$ in (2.5) is called *characteristics* of the Markov process X. We sometimes denote $\mathcal{A}$ by $-\psi(x, D)$. In the following we set $c(x) = 0$ for notational convenience and restrict ourselves to a certain kind of normal Markov processes, so called Feller processes ([4, Theorem 3.1.8] states (2.1) for Feller process, see also [99, p.83]). These can be defined via the semigroup $(T_t)_{t\geq 0}$ generated by the corresponding process X. A semigroup $(T_t)_{t\geq 0}$ is called Feller if it satisfies

(i) T_t maps $C_\infty(\mathbb{R}^d)$, the continuous functions on $\mathbb{R}^d$ vanishing at infinity, into itself:

$$T_t : C_\infty(\mathbb{R}^d) \to C_\infty(\mathbb{R}^d) \quad \text{boundedly.}$$

(ii) T_t is strongly continuous, i.e., $\lim_{t\to 0^+} \|u - T_t u\|_{L^\infty(\mathbb{R}^d)} = 0$ for all $u \in C_\infty(\mathbb{R}^d)$.

Spatially homogeneous Feller processes are Lévy-processes (e.g.[12, 103]). Their characteristics, the *Lévy characteristics*, do not depend on x explicitly.

Example 2.1.2. *A standard Brownian motion has the characteristics* $(0, 1, 0)$. *An* $\mathbb{R}$-*valued Lévy process has characteristics* $(b, Q, \nu(\mathrm{d}z))$, *for real numbers* b, $Q \geq 0$ *and a jump measure* ν *with* $\int_{0\neq z\in\mathbb{R}} \min(1, z^2)\nu(\mathrm{d}z) < \infty$.

It is interesting to ask which symbols correspond to PDOs that are generators of Feller processes. This problem is discussed in the following theorem due to [104].

Theorem 2.1.3. *Let* $\psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ *be a negative definite symbol, i.e., a measurable and locally bounded function in both variables* (x, ξ) *that admits for each* $x \in \mathbb{R}^d$ *a Lévy-Khinchine representation* (2.5). *If*

(a) $\sup_{x\in\mathbb{R}^d} |\psi(x, \xi)| \leq \kappa(1 + |\xi|^2)$ *for all* $\xi \in \mathbb{R}^d$,

(b) $\xi \mapsto \psi(x, \xi)$ *is uniformly continuous at* $\xi = 0$,

(c) $x \mapsto \psi(x, \xi)$ *is continuous for all* $\xi \in \mathbb{R}^d$,

then $(-\psi(x, D), C_0^\infty(\mathbb{R}^d))$ *extends to a Feller generator.*

Remark 2.1.4. *Note that Theorem 2.1.3 does not imply the uniqueness of the process. We show well-posedness of the pricing equations in Chapter 5. For the existence of a solution it is sufficient to require (a) from Theorem 2.1.3 and* $\psi(x, 0) = 0$ *for all* $x \in \mathbb{R}^d$, *cf. [62, Theorem 3.15], (b) and (c) are required to obtain a Feller process.*

In the Lévy case existence of a Lévy process can be proven for any Lévy symbol. This does not hold for Feller processes. For (financial) applications it is more convenient to consider the characteristic triplet instead of the symbol. We therefore make the following assumption on the characteristic triplet in the remainder.

Assumption 2.1.5. *The characteristic triplet* $(b(x), Q(x), \nu(x, \mathrm{d}z))$ *of a Feller process in* $\mathbb{R}^d$ *satisfies the following conditions:*

(I) $(b(x), Q(x), \nu(x, \mathrm{d}z))$ *is a Lévy triplet for all fixed* $x \in \mathbb{R}^d$*.*

(II) *The mapping* $x \mapsto \int_{B \cap \mathbb{R}^d \setminus \{0\}} (1 \wedge |z|^2) \nu(x, dz)$ *is continuous for all* $B \in \mathcal{B}(\mathbb{R}^d)$*.*

(III) *There exists a Lévy measure* $\overline{\nu}(z)$ *s.t.*

$$0 \leq \int_{B \cap \mathbb{R}^d \setminus \{0\}} (1 \wedge |z|^2) \nu(x, dz) \quad \leq \quad \int_{B \cap \mathbb{R}^d \setminus \{0\}} (1 \wedge |z|^2) \overline{\nu}(dz) < \infty,$$

for all $x \in \mathbb{R}^d$, $B \in \mathcal{B}(\mathbb{R}^d)$*.*

(IV) *The functions* $x \mapsto b(x)$ *and* $x \mapsto Q(x)$ *are continuous and bounded.*

Our aim is to conclude that there exists a Feller process whose generator is a PDO for symbols that satisfies Assumption 2.1.5. Therefore, it suffices to validate the prerequisites of Theorem 2.1.3.

Lemma 2.1.6. *Let* $(b(x), Q(x), \nu(x, dz))$ *be the characteristic triplet of a process* X *taking values in* $\mathbb{R}^d$ *that satisfies Assumption 2.1.5. Then* $(-\psi(x, D), C_0^\infty(\mathbb{R}^d))$ *extends to a Feller generator, where* $\psi(x, \xi)$ *is given by*

$$\begin{aligned} \psi(x, \xi) \quad &= \quad -ib(x) \cdot \xi + \frac{1}{2} \xi \cdot Q(x) \xi \\ &\quad + \int_{\mathbb{R}^d \setminus \{0\}} \left(1 - e^{iz \cdot \xi} + \frac{iz \cdot \xi}{1 + |z|^2} \right) \nu(x, \mathrm{d}z). \end{aligned} \tag{2.7}$$

Proof. Condition (I) of Assumption 2.1.5 implies that the corresponding Feller symbol is negative definite. Conditions (III) and (IV) imply (a) of Theorem 2.1.3, Conditions (II) and (III) imply (b), and (c) follows from (II) and (IV). $\square$

Remark 2.1.7. *Note that real price market models, as well as Ornstein-Uhlenbeck models do not fit into our modeling framework due to Assumption (a) in Theorem 2.1.3, as they do not admit a uniform estimate in the state space variable. The numerical methods presented in the following can in many cases be straightforwardly extended to this kind of models.*

In order to apply available tools from pseudodifferential calculus we need to impose stronger assumptions on the characteristic triplets of the considered processes. We state the assumptions needed at the end of Chapter 3. In particular smoothness of the characteristic triplet in the state variable x. Numerical experiments indicate strongly that these assumptions can be weakened (see Chapter 12).

2.1.2 Time-inhomogeneous processes

In this section we drop the assumption of time-homogeneity of the processes considered and extend the framework developed above to a time-dependent setting. Using the notation from the last section, we consider a normal Markov process X with the corresponding family of linear operators $T_{s,t}$. The family of generators of such a process is given by

$$\mathcal{A}_s u := \lim_{h\to 0^+} \frac{1}{h}\left(T_{s-h,s}u - u\right) \tag{2.8}$$

for all functions $u \in D(\mathcal{A}_s) \subset B_b(\mathbb{R}^d)$, such that the limit exists in the strong pointwise sense. In analogy to Theorem 2.1.1 we obtain the following result:

Theorem 2.1.8. *Let $(\mathcal{A}_s, \mathcal{D}(\mathcal{A}_s))_{s\in\mathbb{R}^+}$ be a family of operators with $C_0^\infty(\mathbb{R}^d) \subset \mathcal{D}(\mathcal{A}_s)$ and $\mathcal{A}_s(C_0^\infty(\mathbb{R}^d)) \subset C(\mathbb{R})$. Then $\mathcal{A}_s|_{C_0^\infty(\mathbb{R}^d)}$ is a pseudodifferential operator for all $s \in \mathbb{R}^+$ given by*

$$(\mathcal{A}_s u)(x) := -(2\pi)^{-d/2} \int_{\mathbb{R}^d} \psi(s,x,\xi)\hat{u}(\xi)e^{ix\cdot\xi}d\xi$$

for $u \in C_0^\infty(\mathbb{R}^d)$. The symbol $\psi(s,x,\xi) : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ is locally bounded in (x,ξ). Besides, $\psi(s,\cdot,\xi)$ is measurable for every ξ, s and $\psi(s,x,\cdot)$ is a negative definite function for every (s,x), which admits the Lévy-Khintchine representation

$$\begin{aligned}\psi(s,x,\xi) \quad &= \quad c(s,x) - ib(s,x)\cdot\xi + \frac{1}{2}\xi\cdot Q(s,x)\xi \\ &\quad + \int_{0\neq z\in\mathbb{R}^d} \left(1 - e^{iz\cdot\xi} + \frac{iz\cdot\xi}{1+|z|^2}\right)\nu(s,x,dz).\end{aligned}$$

The question arises if we can construct a Markov process with corresponding generator for a given symbol. A general result under mild regularity assumptions on the symbol has been given by [18].

Theorem 2.1.9. *Let $\psi : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ be a negative definite function that satisfies the following conditions for a constant $m \in \mathbb{R}$, $\alpha, \beta \in \mathbb{N}_0^d$*

(i) $\psi(\cdot, x, \xi)$ is a continuous function for all $x, \xi \in \mathbb{R}^d$,

(ii) $\psi(s,x,0) = 0$ holds for all $(s,x) \in \mathbb{R}^+ \times \mathbb{R}^d$,

(iii) $\left|D_x^\beta D_\xi^\alpha \psi(s,x,\xi)\right| \leq c_{\alpha,\beta,J}(1+|\xi|^2)^{(m-|\alpha|\wedge 2)/2}$ holds for all $s \in J \subset \mathbb{R}^+$, $x, \xi \in \mathbb{R}^d$,

(iv) a is elliptic, i.e., on any compact set K it holds uniformly in s that:

$$\begin{aligned}&\textit{there exists } R,\ c > 0, \ \ \textit{such that } \forall x \in \mathbb{R}^d, \\ &|\xi| \geq R : \Re(\psi(s,x,\xi)) \geq c(1+|\xi|^2)^{m/2}.\end{aligned}$$

Then a Markov process whose family of generators are pseudodifferential operators with symbol $\psi(s,x,\xi)$ can be constructed.

Proof. The proof follows from [18, Theorem 4.2, Corollary 4.3]. □

Theorem 2.1.9 can be formulated in a more general setting replacing $|\xi|^2$ in Condition (iii) and (iv) by any element from a certain class of negative definite functions, cf. [18, Definition 1.2].

Remark 2.1.10. *Assumptions (iii) and (iv) in Theorem 2.1.9 significantly restrict the range of market models, as a uniform time-dependence is required and degeneracies in the coefficients are ruled out. Weakly degenerate behavior of the coefficients in the generators requires appropriate weak time formulations in order to obtain a well-posed problem. This is discussed in the following. Besides, due to the independence of the exponent of $(1+|\xi|)$ of β in (iii) in Theorem 2.1.9, the space dependence is to be seen as a local fluctuation. This is significantly more restrictive than the setup in Section 2.1.1. Time-inhomogeneous admissible market models are defined in Section 3.3.*

2.2 Variational formulation of parabolic problems

2.2.1 Variable order Sobolev spaces

For later use we shall introduce anisotropic and variable order Sobolov spaces. We start with the definition of fractional order isotropic spaces and define for a positive non-integer $s \geq 0$ and $u \in \mathcal{S}^*(\mathbb{R}^d)$, where $\mathcal{S}^*(\mathbb{R}^d)$ denotes the space of tempered distributions,

$$\|u\|^2_{H^s(\mathbb{R}^d)} := \int_{\mathbb{R}^d} (1+|\xi|^2)^s \, |\hat{u}(\xi)|^2 \, \mathrm{d}\xi. \tag{2.9}$$

Similarly, we can define anisotropic Sobolev spaces $H^{s}(\mathbb{R}^d)$ with norm $\|\cdot\|_{H^{s}(\mathbb{R}^d)}$ given by

$$\|u\|^2_{H^{s}(\mathbb{R}^d)} := \int_{\mathbb{R}^d} \sum_{j=1}^{d} (1+\xi_j^2)^{s_j} \, |\hat{u}(\xi)|^2 \, \mathrm{d}\xi, \tag{2.10}$$

for any multiindex $s \geq 0$. The consideration of certain symbol classes will be useful for the definition of the variable order Sobolev spaces. We set $\langle\xi\rangle := (1+|\xi|^2)^{1/2}$ for notational convenience.

Definition 2.2.1. *Let $0 \leq \delta < \rho \leq 1$ and let $m(x) \in C^\infty(\mathbb{R}^d)$ be a real-valued function with bounded derivatives on $\mathbb{R}^d$ of arbitrary order. Then, the symbol $\psi(x,\xi)$ belongs to the class $S^{m(x)}_{\rho,\delta}$ of symbols of variable order $m(x)$ if $a(x,\xi) \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ and*

$m(x) = s + \widetilde{m}(x)$ with $\widetilde{m} \in \mathcal{S}(\mathbb{R}^d)$ a tempered function, and if, for every $\alpha, \beta \in \mathbb{N}_0^d$ there is a constant $c_{\alpha,\beta}$ such that

$$\forall x, \xi \in \mathbb{R}^d : \quad |D_x^\beta D_\xi^\alpha \psi(x,\xi)| \leq c_{\alpha,\beta} \langle \xi \rangle^{m(x)-\rho|\alpha|+\delta|\beta|}. \tag{2.11}$$

The variable order pseudodifferential operators $\mathcal{A}(x,D) \in \Psi_{\rho,\delta}^{m(x)}$ correspond to symbols $\psi(x,\xi) \in S_{\rho,\delta}^{m(x)}$ by

$$\mathcal{A}(x,D)u(x) := \frac{1}{2\pi} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{i(x-y)\cdot\xi} \psi(x,\xi) u(y) dy d\xi, \qquad u \in C_0^\infty(\mathbb{R}^d). \tag{2.12}$$

We are now able to define an isotropic Sobolev space of variable order $H^{m(x)}(\mathbb{R}^d)$, $m(x) \geq 0$, using the variable order Riesz potential $\Lambda^{m(x)}$ with symbol $\psi(x,\xi) = \langle \xi \rangle^{m(x)}$. Clearly $\psi(x,\xi)$ is an element of $S_{1,\delta}^{m(x)}$ for $\delta \in (0,1)$. A norm $\|\cdot\|_{H^{m(x)}(\mathbb{R}^d)}$ on $H^{m(x)}(\mathbb{R}^d)$ is given by

$$\|u\|_{H^{m(x)}(\mathbb{R}^d)}^2 := \left\| \Lambda^{2m(x)} u \right\|_{L^2(\mathbb{R}^d)}^2 + \|u\|_{L^2(\mathbb{R}^d)}^2 .$$

Note that for $\psi(x,\xi) = 1$, we obtain the usual $L^2(\mathbb{R}^d)$ norm. For $\psi(x,\xi) = (1 + |\xi|^s)$ we obtain the norm given in (2.9), which follows applying Plancherel's theorem. Now we turn to the definition of anisotropic variable order Sobolev spaces. In analogy to Definition 2.2.1 we start with the definition of an appropriate symbol class.

Definition 2.2.2. *Let $\mathbf{m}(x) = s + \widetilde{\mathbf{m}}(x)$, $\widetilde{\mathbf{m}}(x) : \mathbb{R}^d \to \mathbb{R}^d$ with each component of $\widetilde{\mathbf{m}}(x)$ being a tempered function and $s \in \mathbb{R}_+^d$, $0 \leq \delta < \rho \leq 1$. We define the symbol class $S_{\rho,\delta}^{\mathbf{m}(x)}$ as the set of all $\psi(x,\xi) \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ such that for all multiindices $\alpha, \beta \in \mathbb{N}_0^d$ there exists a constant $C_{\alpha,\beta} > 0$ with*

$$\forall x, \xi \in \mathbb{R}^d \quad : \quad \left| D_x^\beta D_\xi^\alpha \psi(x,\xi) \right| \leq C_{\alpha,\beta} \sum_{i=1}^d (1 + \xi_i^2)^{(m_i(x)-\rho\alpha_i+\delta|\beta|)/2}. \tag{2.13}$$

An anisotropic Sobolev space of variable order can now be defined using the variable order Riesz potential $\Lambda^{\mathbf{m}(x)}$ with symbol $\psi(x,\xi) = \langle \xi \rangle^{\mathbf{m}(x)} := \sum_{i=1}^n (1 + \xi_i^2)^{\frac{1}{2}m_i(x)}$, $m_i(x) \geq 0$, $i = 1, \ldots, d$. Clearly, $\psi(x,\xi)$ is an element of $S_{1,\delta}^{\mathbf{m}(x)}$ for $\delta \in (0,1)$. The norm on $H^{\mathbf{m}(x)}(\mathbb{R}^d)$ is given by

$$\|u\|_{H^{\mathbf{m}(x)}(\mathbb{R}^d)}^2 := \left\| \Lambda^{2\mathbf{m}(x)} u \right\|_{L^2(\mathbb{R}^d)}^2 + \|u\|_{L^2(\mathbb{R}^d)}^2 .$$

There is an alternative representation of the above space, when $\mathbf{m}(x)$ admits the following representation $\mathbf{m}(x) = (m_1(x_1), \ldots, m_d(x_d))$. This will be very useful for the proof of norm equivalences, which play a crucial role in wavelet discretization theory. We consider the anisotropic Sobolev spaces $H_i^{m_i(x_i)}(\mathbb{R}^d)$ of variable order $m_i(x_i)$ in direction x_i, equipped with the norms $\|\cdot\|_{H_i^{m(x)}(\mathbb{R}^d)}$ given by

$$\|u\|^2_{H_i^{m(x)}(\mathbb{R}^d)} := \left\|\Lambda_i^{2m_i(x_i)} u\right\|^2_{L^2(\mathbb{R}^d)} + \|u\|^2_{L^2(\mathbb{R}^d)} ,$$

where $\Lambda_i^{m_i(x_i)}$ is a pseudo-differential operator with symbol $(1+|\xi_i|)^{m_i(x_i)}$. It then follows by the elementary inequality

$$C_1 \left|\sum_{i=1}^d a_i\right|^2 \leq \sum_{i=1}^d a_i^2 \leq C_2 \left|\sum_{i=1}^d a_i\right|^2 ,$$

with $a_i > 0$ and C_1, C_2 only dependent on d, that

$$\|u\|^2_{H^{m(x)}(\mathbb{R}^d)} \sim \sum_{j=1}^d \|u\|^2_{H_j^{m_j(x_j)}(\mathbb{R}^d)} , \tag{2.14}$$

and therefore

$$H^{\mathbf{m}(x)}(\mathbb{R}^d) = \bigcap_{j=1}^d H_j^{m_j(x_j)}(\mathbb{R}^d). \tag{2.15}$$

On the bounded set $D = (\mathbf{a}, \mathbf{b}) = \prod_{i=1}^d (a_i, b_i) \subset \mathbb{R}^d$ we define for a variable order $\mathbf{m}(x)$, $\mathbf{a} \leq x \leq \mathbf{b}$ the space

$$\widetilde{H}^{\mathbf{m}(x)}(D) := \left\{ u|_D \;\middle|\; u \in H^{\mathbf{m}(x)}(\mathbb{R}^d), \quad u|_{\mathbb{R}^d \setminus \overline{D}} = 0 \right\} .$$

This space coincides with the closure of $C_0^\infty(D)$ (the space of smooth functions with support compactly contained in D) with respect to the norm

$$\|u\|_{\widetilde{H}^{\mathbf{m}(x)}(D)} := \|\widetilde{u}\|_{H^{\mathbf{m}(x)}(\mathbb{R}^d)}, \tag{2.16}$$

where $\widetilde{u}$ is the zero extension of u to all of $\mathbb{R}^d$. The spaces of order $\mathbf{m}(x) \leq 0$, $\forall x \in \mathbb{R}^d$, are defined by duality. We have

$$H^{\mathbf{m}(x)}(D) := (\widetilde{H}^{-\mathbf{m}(x)}(D))^*,$$

where duality is understood with respect to the "pivot" space $L^2(D)$, i.e., $L^2(D)^* \cong L^2(D)$.

Remark 2.2.3. *In the BS case $H^1(\mathbb{R}^d)$ is obtained as the domain of the Dirichlet form while $H_0^1(D)$ is the domain in the localized case. In the Lévy case we obtain anisotropic Sobolev spaces as in* (2.10) *and the spaces $\widetilde{H}^{\mathbf{s}}(D)$ in the localized case for $Q = 0$. For $Q > 0$ the domains are equal to those in the BS case, cf. [95, Theorem 4.8].*

2.2.2 Notions of solutions of parabolic PIDEs

In this section we present different notions of solutions of parabolic PIDEs. Classical solutions of certain PDEs, as defined in the following, may not exist or their existence may be hard to show. Therefore different notions of solutions have been developed in the literature, such as variational solutions, weak space-time solutions, ultra weak solutions, mild solution or viscosity solutions. We discuss some of them, as different notions of solutions are used throughout this work. The notion of a classical solution is discussed first. Given some $f \in C^0([0,T] \times \mathbb{R}^d)$ and $g \in C^0(\mathbb{R}^d)$, find some function $u \in C^0([0,T] \times \mathbb{R}^d) \cap C^{1,2}((0,T) \times \mathbb{R}^d)$ such that

$$\begin{aligned}&\partial_t u - \mathcal{A}u = f \text{ in } (0,T) \times \mathbb{R}^d\\ &u(0) = g \text{ in } \mathbb{R}^d,\end{aligned}$$

for some second order linear integrodifferential operator $\mathcal{A}$ and $T > 0$. In the following we briefly outline the standard variational setting for parabolic equations which was applied in, e.g., [95, 113]. Let $\mathcal{V} \subset \mathcal{H}$ be two Hilbert spaces with continuous and dense embedding. We identify $\mathcal{H}$ with its dual $\mathcal{H}^*$ and obtain the Gelfand triple

$$\mathcal{V} \subset \mathcal{H} \equiv \mathcal{H}^* \subset \mathcal{V}^*.$$

The space $\mathcal{V}$ is in this setting the domain of a certain bilinear form $a(\cdot,\cdot)$ associated to an operator $\mathcal{A}$. Let $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ be a densely defined operator on $\mathcal{H} = L^2(\mathbb{R}^d)$ which is negative definite, cf. [69, Definition 4.6.10] and satisfies

$$|(-\mathcal{A}u, v)| \leq C\,(-\mathcal{A}u, u)^{1/2}\,(-\mathcal{A}v, v)^{1/2},$$

where $(\cdot,\cdot)$ denotes the $L^2(\mathbb{R}^d)$ scalar product and $u, v \in \mathcal{D}(\mathcal{A})$. Then we may introduce on $\mathcal{D}(\mathcal{A})$ the bilinear form

$$a(u,v) := (-\mathcal{A}u, v).$$

The bilinear form $\widetilde{a}(\cdot,\cdot)$ given as

$$\widetilde{a}(u,v) := a^{\text{sym}}(u,v) + (u,v) = \frac{1}{2}\left(a(u,v) + a(v,u)\right) + (u,v)$$

defines a scalar product and we may consider the completion of $\mathcal{D}(\mathcal{A})$ with respect to $\widetilde{a}(\cdot,\cdot)$, which is denoted by $\mathcal{D}(a(\cdot,\cdot))$. The following abstract well-posedness result holds.

Theorem 2.2.4. *Let the bilinear form $a(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ satisfy the following properties. There exist some constants $C_1, C_2 > 0$ and $C_3 \geq 0$ such that for all $u, v \in \mathcal{V}$ there holds*

$$\begin{aligned}|a(u,v)| &\leq C_1 \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}},\\ a(u,u) &\geq C_2 \|u\|_{\mathcal{V}}^2 - C_3 \|u\|_{\mathcal{H}}^2.\end{aligned}$$

Then the following abstract parabolic problem admits a unique solution.
Find $u \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$ *such that*

$$(\partial_t u, v)_{\mathcal{V}^*,\mathcal{V}} + a(u,v) = (f,v)_{\mathcal{V}^*,\mathcal{V}}, \forall v \in \mathcal{V}, \text{ a.e. in } (0,T), \tag{2.17}$$
$$u(0) = g, \tag{2.18}$$

with $g \in \mathcal{H}$, $f \in L^2((0,T);\mathcal{V}^*)$ *and* $T > 0$.

Proof. See [80, Theorem 4.1]. □

For an infinitesimal generator $\mathcal{A}$ of a Markov process X the bilinear form $a(\cdot,\cdot)$ is closely linked to its Dirichlet form, cf. [69, Definition 4.7.21]

Remark 2.2.5. *Note that pure transport operators and time-inhomogeneous operators do not fit into this framework and have to be analyzed using different techniques.*

Besides, different notions of solutions are necessary in these cases and therefore weak space-time formulations are used for time-inhomogeneous equations. The idea of a weak space-time formulation is analogous to a variational formulation in space, where the differentiability requirements on the solution are reduced in comparison to a classical solution by an integration over the spatial domain and application of integration by parts. Let us consider functions $u, v \in C^1(I, \mathcal{V})$, then

$$\int_0^T (\partial_t u(t), v(t))_{L^2(D)} dt = -\int_0^T (u(t), \partial_t v(t))_{L^2(D)} dt + (u(t), v(t))_{L^2(D)}|_0^T.$$

A weak space-time formulation reads
Find $u \in L^2((0,T);\mathcal{V})$ such that

$$B(u,v) = f_2(v), \tag{2.19}$$
$$B(u,v) = \int_0^T \left(-(u(t), \partial_t v(t))_{\mathcal{V},\mathcal{V}^*} + a(u(t), v(t))\right) dt, \tag{2.20}$$
$$f_2(v) = \int_0^T (v(t), f(t))_{\mathcal{V},\mathcal{V}^*} dt + (g, v(0))_{\mathcal{H}} \tag{2.21}$$

with $g \in \mathcal{H}$, $f \in L^2((0,T);\mathcal{V}^*)$ and for all $v \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$, $v(T) = 0 \in \mathcal{H}$. The initial condition is enforced weakly and the regularity of the solution in time is weaker than in the variational setting in this formulation. We remark that different space-time formulations are also admissible and refer to Chapter 9 for details, where weighted spaces in time are used to obtain a feasible formulation.
A different notion of solutions are mild solutions which arise in the context of semigroup theory. We consider the following inhomogeneous initial value problem. Find $u(t)$ such that

$$\partial_t u - \mathcal{A}u = f, \quad t \in (0,T] \tag{2.22}$$

$$u(0) = g. \tag{2.23}$$

We assume $\mathcal{A}$ to be the generator of a C_0 semigroup $T(t)$, $f \in L^1((0,T);\mathcal{H})$ and $g \in \mathcal{H}$. Then a function $u \in C^0([0,T];\mathcal{H})$ given as

$$u(t) := T(t)g + \int_0^t T(t-s)f(s)ds$$

is called a mild solution. Under some regularity assumptions on f and g, we obtain $u(t) \in C^0([0,T];\mathcal{D}(\mathcal{A})) \cap C^1((0,T];\mathcal{H})$. We remark that the spatial regularity of a mild solution under some assumptions on the initial data the right hand side is stronger compared to a variational solution as given in (2.17)-(2.18).

2.3 Lévy copulas

In this section some basic definitions and results on Lévy copulas are given. Lévy copulas play a crucial role in the definition of admissible multidimensional market models, as the multivariate jump measure are constructed using a Lévy copula. Alternatively, multidimensional Lévy and Feller processes could be constructed in analogy to the univariate setting via subordination. We refer to [89] for a thorough introduction to copulas and to [75] for details on Lévy copulas. We start with the definition of the F-volume, cf. [89, Definition 2.10.1].

Definition 2.3.1. *The F-volume of $(a,b]$, $a,b \in \overline{\mathbb{R}}^d$ for a function $F : G \to \overline{\mathbb{R}}$, $G \subset \overline{\mathbb{R}}^d$ is defined by*

$$V_F((a,b]) := \sum_{u \in \{a_1,b_1\} \times \cdots \times \{a_d,b_d\}} (-1)^{N(u)} F(u),$$

where $N(u) = |\{k : u_k = a_k\}|$.

For the definition of a Lévy copula we need the concept of a *d-increasing function.*

Definition 2.3.2. *A function $F : G \to \overline{\mathbb{R}}$ is called d-increasing if $V_F((a,b]) \geq 0$ for all $a,b \in G$ with $a \leq b$ and $\overline{(a,b]} \subset G$.*

Marginal distributions play a crucial role in financial modeling and are especially important for the calibration of models.

Definition 2.3.3. *Let $F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}^d$ be a d-increasing function which satisfies $F(u) = 0$ if $u_i = 0$ for at least one $i \in \{1,\dots,d\}$ the I-margin of F is the function $F^I : \overline{\mathbb{R}}^{|I|} \to \overline{\mathbb{R}}$*

$$F^I(u^I) := \lim_{a\to\infty} \sum_{(u_j)_{j\in I^c} \in \{-a,\infty\}^{|I|^c}} \left(\prod_{j \in I^c} \operatorname{sgn} u_j \right) F(u_1,\dots,u_d).$$

We also need the tail integrals of a Feller process.

Definition 2.3.4. *Let X be a Feller process with state space $\mathbb{R}^d$ and jump measure $\nu(x,dz)$. The tail integral of X is the function $U : \mathbb{R}^d \times \mathbb{R}^d \backslash \{0\}$ given by*

$$U(x_1,\dots,x_d,z_1,\dots,z_d) = \prod_{i=1}^{d} \nu(x_1,\dots,x_d,\prod_{j=1}^{d} I(z_j)),$$

where

$$I(z) = \begin{cases} (z,\infty) & \text{for } z \geq 0 \\ (-\infty,z) & \text{for } z < 0 \end{cases}.$$

Furthermore, for $\mathcal{I} \subset \{1,\dots,d\}$ nonempty the $\mathcal{I}$-marginal tail integral $U^{\mathcal{I}}$ of X is the tail integral of the process $X^{\mathcal{I}} = (X^i)_{i\in\mathcal{I}}$.

After these preparations, we define a Lévy copula as follows.

Definition 2.3.5. *A function $F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ is called Lévy copula if*

(i) $F(u_1,\dots,u_d) \neq \infty$ *for* $(u_1,\dots,u_d) \neq (\infty,\dots,\infty)$,

(ii) $F(u_1,\dots,u_d) = 0$ *for* $u_i = 0$ *for at least one* $i \in \{1,\dots,d\}$,

(iii) F *is d-increasing,*

(iv) $F^{\{i\}}(u) = u$ *for any* $i \in \{1,\dots,d\}$, $u \in \mathbb{R}$.

For the definition of an admissible market model some spatial homogeneity of the copula function are required.

Definition 2.3.6. *A Lévy copula is called 1-homogeneous if for any $r > 0$ there holds*

$$F(ru_1,\dots,ru_d) = rF(u_1,\dots,u_d).$$

We conclude with some examples of Lévy copulas, cf. [75].

Example 2.3.7.

(i) Independence Lévy copula

$$F(u_1,\dots,u_d) = \sum_{i=1}^{d} u_i \prod_{j\neq i} \mathbb{1}_{\{\infty\}}(u_j).$$

(ii) Complete dependence Lévy copula

$$F(u_1, \ldots, u_d) = \min\{|u_1|, \ldots, |u_d|\} \mathbb{1}_K(u_1, \ldots, u_d) \prod_{j=1}^{d} \operatorname{sgn} u_j,$$

where $K := \{x \in \mathbb{R}^d : \operatorname{sgn}(x_1) = \ldots = \operatorname{sgn}(x_d)\}$*.*

(iii) Clayton Lévy copulas

$$F(u) = 2^{2-d} \left(\sum_{i=1}^{d} |u_i|^{-\vartheta} \right)^{-\frac{1}{\vartheta}} \left(\eta \mathbb{1}_{\{u_1 \cdots u_d \geq 0\}} - (1-\eta) \mathbb{1}_{\{u_1 \cdots u_d\}} \right),$$

where $\vartheta > 0$ *and* $\eta \in [0, 1]$*.*

3 Examples of market models

Several examples of Feller processes and time-inhomogeneous processes are provided in this chapter. The methods presented in this work are not applicable to all of them. But admissible market models are derived ensuring the well-posedness of the corresponding pricing equations and the applicability of finite element methods. Subordination can be used to construct Feller processes. However, the structure of the symbol is more involved than in the Lévy setting. We also present a construction of Feller processes using Lévy copulas. Finally, time-inhomogeneous Lévy models, which were not covered in Section 2.1.2 due to the possibly degenerate behavior of their symbols in time, are discussed.

3.1 Subordination

Many Lévy models in the context of option pricing are constructed via subordination of a Brownian motion by a corresponding stochastic clock, e.g., an NIG process [6] or a VG process [81]. We describe a similar construction for Feller processes and point out similarities and differences to the Lévy case. Bernstein functions play a crucial role in the representation of subordinators.

Definition 3.1.1. *A function $f(x) \in C^\infty(0,\infty)$ is called a Bernstein function if*

$$f \geq 0, \quad (-1)^k \frac{\partial^k f(x)}{\partial x^k} \leq 0, \quad \forall k \in \mathbb{N}.$$

Example 3.1.2. *The functions $f_1(x) = c$, $f_2(x) = cx$ and $f_3(x) = 1 - e^{-cx}$, $c \geq 0$ are Bernstein functions.*

Bernstein functions admit the following representation.

Theorem 3.1.3. *Let $f(x)$ be a Bernstein function, then there exists a measure μ on $(0,\infty)$, with*

$$\int_{0+}^{\infty} \frac{s}{1+s} \mu(ds) < \infty$$

such that for $x > 0$ and positive constants a, b

$$f(x) = a + bx + \int_{0+}^{\infty} (1 - e^{-xs}) \mu(ds).$$

Proof. See [69, Theorem 3.9.4]. □

If we choose a Lévy process as subordinator, Bernstein functions allow for a nice characterization of the process. Subordinators can be described via their convolution semigroup.

Definition 3.1.4. *A family $(\eta_t)_{t\geq 0}$ of bounded Borel measures on $\mathbb{R}$ is called convolution semigroup on $\mathbb{R}$ if the following conditions are fulfilled*

1. $\eta_t(\mathbb{R}) \leq 1$, *for all* $t \geq 0$,
2. $\eta_s * \eta_t = \eta_{s+t}$, $s, t \geq 0$ *and* $\mu_0 = \delta_0$,
3. $\eta_t \to \delta_0$ *vaguely as* $t \to 0$,

here δ_0 denotes the Dirac measure at 0. By vague convergence of a sequence $(\eta_t)_{t>0}$ of measures to η_0 we mean that for all continuous functions with compact support $u \in C_0(\mathbb{R})$, we have

$$\lim_{t\to 0} \int_{\mathbb{R}} u(x)\eta_t(dx) = \int_{\mathbb{R}} u(x)\eta_0(dx).$$

The relation between convolution semigroups and Bernstein function is given in the following theorem.

Theorem 3.1.5. *Let $f : (0, \infty) \to \mathbb{R}$ be a Bernstein function. Then there exists a unique convolution semigroup $(\eta_t)_{t\geq 0}$ supported on $[0, \infty)$ such that*

$$\mathcal{L}(\eta_t)(x) = e^{-tf(x)}, \quad x > 0 \text{ and } t > 0, \tag{3.1}$$

holds, where $\mathcal{L}$ denotes the Laplace transform, i.e., $\mathcal{L}(\eta_t)(x) := \int_0^\infty e^{-zx}\eta_t(dz)$, for appropriate η_t and $x > 0$. Conversely, for any convolution semigroup $(\eta_t)_{t\geq 0}$ supported by $[0, \infty)$ there exists a unique Bernstein function f such that (3.1) holds.

Proof. See [69, Theorem 3.9.7]. □

We recall the correspondence between convolution semigroups and Lévy processes.

Theorem 3.1.6. *Let X be a Lévy process, where for each $t \geq 0$ $X(t)$ has law η_t, then $(\eta_t)_{t\geq 0}$ is a convolution semigroup.*

Proof. See [4, Proposition 1.4.4]. □

The semigroup of a subordinated Feller processes can now be characterized.

Theorem 3.1.7. *Let $(T_t)_{t\geq 0}$ be a Feller semigroup on a Banach space $\mathcal{H}$ with generator $\mathcal{A}$, let $f : (0,\infty) \to \mathbb{R}$ be a Bernstein function and $(\eta_t)_{t\geq 0}$ the associated convolution semigroup on $\mathbb{R}$ supported on $[0,\infty)$. Define $T_t^f u$ for $u \in \mathcal{H}$ by the Bochner integral*

$$T_t^f u = \int_0^\infty T_s u \, \eta_t(ds). \tag{3.2}$$

Then the integral is well-defined and $(T_t^f)_{t\geq 0}$ is a Feller semigroup on $\mathcal{H}$.

Proof. See [69, Theorem 4.3.1 and Corollary 4.3.4]. □

The representation (3.2) of $T_t^f u$ can be used for numerical methods, but it involves the approximation of an integral over a possibly semi-infinite interval, which can be very costly if the integrand is not well-behaved. The generator $\mathcal{A}^f$ of the semigroup $(T_t^f)_{t\geq 0}$ is a PDO and its symbol is given as follows in the case of a spatially homogeneous semigroup $(T_t)_{t\geq 0}$ corresponding to a Lévy process X.

Theorem 3.1.8. *Let $(T_t)_{t\geq 0}$ be a Feller semigroup with generator $\mathcal{A}$ with constant symbol $\psi(\xi)$ and $f(x)$ as in the previous theorem, then the symbol of $\psi^f(\xi)$ of the generator $\mathcal{A}^f$ of the semigroup $(T_t^f)_{t\geq 0}$ is given as*

$$\psi^f(\xi) = f(\psi(\xi)).$$

Proof. See [4, Proposition 1.3.27]. □

The same characterization does not hold in the case of a more general subordinated process. We obtain the following representation for $\mathcal{A}^f$.

Theorem 3.1.9. *Let f and $(T_t)_{t\geq 0}$ be as in Theorem 3.1.7. For all $u \in D(\mathcal{A})$ we have $u \in D(\mathcal{A}^f)$ and*

$$\mathcal{A}^f u = au + b\mathcal{A}u + \int_0^\infty R_\lambda \mathcal{A} u \mu(d\lambda),$$

where R_λ denotes the resolvent of $\mathcal{A}$, i.e., $R_\lambda u = (\lambda - \mathcal{A})^{-1} u$ and a, b and $\mu(ds)$ are defined in Theorem 3.1.3.

Proof. See [71, Theorem 2.15]. □

This representation is of limited use for computation and we therefore aim at a characterization of the symbol of the PDO $\mathcal{A}^f$. We consider a certain symbol class as in [71]. Let $L(x, D)$ be the differential operator given by

$$L(x, D) = -\sum_{i,j=1}^{d} a_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + c(x),$$

where $a_{i,j}(x)$, $1 \leq i, j \leq d$ are continuously differentiable functions such that $a_{i,j}(x) = a_{j,i}(x)$ and

$$\kappa_1 |\xi|^2 \leq \sum_{i,j=1}^{d} a_{i,j}(x)\xi_i\xi_j \leq \kappa_2 |\xi|^2$$

holds for all $x \in \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$ for some constants $0 < \kappa_1 \leq \kappa_2$. Besides we assume

$$\sum_{i=1}^{d} \frac{\partial a_{i,j}}{\partial x_i} = 0,$$

for any $j = 1, \ldots, d$ and $c(x)$ is a continuous and bounded function satisfying $0 < \underline{c} \leq c(x) \leq \overline{c} < \infty$. In this situation we obtain the following representation of $\mathcal{A}^f = f(\mathcal{A})$ for $u \in D(\mathcal{A})$, $\mathcal{A} = L(x, D)$.

$$f(\mathcal{A})u = f(L(x,\xi))u + \int_0^\infty \lambda R_\lambda K_\lambda(x, D)u\mu(d\lambda), \tag{3.3}$$

where

$$\begin{aligned} K_\lambda(x, D) &= (L(x, D) + \lambda \mathrm{id}) \circ q_\lambda(x, D) - \mathrm{id}, \\ q_\lambda(x, \xi) &= \frac{1}{L(x, \xi) + \lambda\xi} \end{aligned}$$

and R_λ denotes the resolvent of $\mathcal{A}$ at λ. We remark that $K_\lambda \equiv 0$ for constant symbols which proves Theorem 3.1.8. In general both terms in (3.3) have to be considered. We refer to Carr [23] for a generalization of the VG model.

Remark 3.1.10. *The consideration of symbols of the type $a(x, \xi)$, where $a(x, \xi)$ is a Lévy symbol for all $x \in \mathbb{R}$ is therefore in general not equivalent to a construction via subordination. This observation was made by [7], where some asymptotic expansion of the difference in terms of the symbol under certain assumptions on the structure of the process was provided.*

3.2 Multivariate models arising from copulas

Unlike multivariate Lévy processes, cf. [75, Theorem 3.6], not all multivariate Feller processes can be constructed in terms of univariate Feller processes using a homogeneous copula construction. However, parametric constructions of multidimensional Feller processes from the univariate margins of certain Feller processes and certain Lévy copulas are still possible, provided the univariate Feller processes and the copulas meet certain restrictions. The restrictions stem from the fact that smoothness conditions on the characteristic triplet appear to be required in order to prove existence (and uniqueness) of a corresponding Feller process, cf. [104]. Therefore it would be sufficient for the parametric construction of d-dimensional Feller processes to prove that a symbol satisfies Assumption 2.1.5. We only consider here the construction of a d-dimensional jump measure, as the Gaussian part is well-known.

3.2.1 Copula functions

In the following an extension of Sklar's theorem to certain types of Markov processes is proved. Besides, some estimates on the tail behavior of the density of the jump measure stemming from a copula construction are given. Such estimates are needed for the construction of efficient quadratures for the numerical solution of the considered pricing equations.

Theorem 3.2.1. *Let F denote a d-dimensional Lévy copula for which the derivative $\partial_1 \ldots \partial_d F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ exists, is continuous and satisfies the following estimate*

$$|\partial^n F(u)| \leq C^{|n|} |n|! \min\{|u_1|, \ldots, |u_d|\} \prod_{i=1}^{d} |u_i|^{-n_i} \quad \forall u \in \mathbb{R}^d \quad n \in \mathbb{N}^d. \tag{3.4}$$

Further let $U_i(x,z)$, $i = 1, \ldots, d$ denote the tail integrals of real valued Feller processes that satisfy Assumption 2.1.5 and, additionally, the following conditions:

$$\left| \frac{k_i(x,z)}{U_i(x,z)} \right| \leq (C \vee \frac{1}{|z|}) \quad \forall x, z \in \mathbb{R}, \tag{3.5}$$

and

$$\int_{|z|>1} U_i(x,z)\, dz < \infty, \tag{3.6}$$

for $C > 0$ and $i = 1, \ldots, d$. Then there exists an $\mathbb{R}^d$-valued Feller process X whose components have tail integrals $U_1, \ldots, U_d$ and whose marginal tail integrals satisfy

$$U^I\left((x_i)_{i\in I}, (z_i)_{i\in I}\right) = F^I\left((U(x_i, z_i))_{i\in I}\right)$$

for any non-empty $I \subset \{1, \ldots, d\}$, any $(z_i)_{i\in I} \in (\mathbb{R}\backslash\{0\})^{|I|}$ and any $(x_i)_{i\in I} \in \mathbb{R}^{|I|}$. The jump measure is uniquely determined by F and U_i, $i = 1, \ldots, d$.

Proof. We follow closely the argumentation in [75]. As stated there, the argument is not restricted to Lévy models but can be extended to more general processes.
Since F is d-increasing and continuous, we can conclude that there exists a unique measure μ on $\overline{\mathbb{R}}^d \backslash \{\infty, \ldots, \infty\}$ such that $V_F((a,b]) = \mu((a,b])$ for any a, b with $a \leq b$. For the univariate tail integrals $U(x,z)$, we define

$$U^{-1}(x,u) = \begin{cases} \inf\{z > 0 : u \geq U(x,z)\}, & \text{for } u \geq 0 \\ \inf\{z < 0 : u \geq U(x,z)\} \wedge 0, & \text{for } u < 0. \end{cases}$$

Let $\nu' = f(\mu)$ be the image of μ under

$$f : (x, u_1, \ldots, u_d) \mapsto (U_1^{-1}(x_1, u_1), \ldots, U_d^{-1}(x_d, u_d))$$

and let ν be the restriction of ν' to $\mathbb{R}^d \times \mathbb{R}^d \backslash\{0\}$. We need to prove that ν is a Lévy measure for all x and that the marginal tail integrals U_ν^I satisfy

$$U_\nu^I((x)_{i\in I},(z_i)_{i\in I}) = F^I((U_i(x_i,z_i))_{i\in I}).$$

This implies (I) of Assumption 2.1.5. Furthermore, we must prove continuity of the Lévy kernel in x (II) as well as boundedness in the sense of (III).
For ease of notation we assume that $z_i > 0$, $i \in I$. Then

$$\begin{aligned}
U_\nu^I\left((x_i)_{i\in I},(z_i)_{i\in I}\right) &= \nu\Big((x_i)_{i\in I},\{\xi\in\mathbb{R}^d\backslash\{0\} : \xi_i \in (z_i,\infty), i\in I\}\Big)\\
&= \mu\Big(\{u\in\overline{\mathbb{R}}^d : U_i^{-1}(x_i,u_i)\in(z_i,\infty), i\in I\}\Big)\\
&= \mu\Big(\{u\in\overline{\mathbb{R}}^d : 0<u_i<U_i(x_i,z_i), i\in I\}\Big)\\
&= \mu\Big(\{u\in\overline{\mathbb{R}}^d : 0<u_i\le U_i(x_i,z_i), i\in I\}\Big)\\
&= F^I\left((U_i(x_i,z_i))_{i\in I}\right).
\end{aligned}$$

This proves in particular that the univariate marginal tail integrals of ν equal $U_1,\dots,U_d$. Since the margins of $\nu(x,z)$ are Lévy measures on $\mathbb{R}\backslash\{0\}$ for all $x\in\mathbb{R}^d$ we obtain for every $x\in\mathbb{R}^d$:

$$\begin{aligned}
\int_{z\in\mathbb{R}^d}\Big(|z|^2\wedge 1\Big)\,\nu(x,dz) &\le \int_{z\in\mathbb{R}^d}\sum_{i=1}^d\left(z_i^2\wedge 1\right)\nu(x,dz)\\
&= \sum_{i=1}^d\int_{z_i\in\mathbb{R}}\left(z_i^2\wedge 1\right)\nu_i(x_i,dz_i) < \infty.
\end{aligned}$$

Hence, for $x\in\mathbb{R}^d$, $\nu(x,\cdot)$ is a Lévy measure on $\mathbb{R}^d$. For the second part of the proof we use Remark 2.7 in [95] which leads to:

$$k(x,z_1,\dots,z_d) = \partial_1\dots\partial_d F|_{\xi_1=U_1(x_1,z_1),\dots,\xi_d=U_d(x_d,z_d)}k_1(x_1,z_1)\dots k_d(x_d,z_d). \tag{3.7}$$

Using the properties of F and the margins we conclude (II) from Assumption 2.1.5. It remains to prove (III) and (IV). Due to (3.4) we have the following estimate with $g:=\partial_1\dots\partial_d F$:

$$\begin{aligned}
& k(x,z_1,\dots,z_d) &(3.8)\\
=\ & g(U_1(x_1,z_1),\dots,U_d(x_d,z_d))k_1(x_1,z_1)\dots k_d(x_d,z_d)\\
\le\ & C\min\{|U_1(x_1,z_1)|,\dots,|U_d(x_d,z_d)|\}\prod_{i=1}^d|U_i(x_i,z_i)|^{-1}\prod_{i=1}^d k_i(x_i,z_i)\\
\overset{(3.5)}{\le}\ & C\min\{\left|\overline{U}_1(z_1)\right|,\dots,\left|\overline{U}_d(z_d)\right|\}\prod_{i=1}^d\left(C\vee\frac{1}{|z_i|}\right). &(3.9)
\end{aligned}$$

Using the properties of the $\overline{\nu}_i(dz)$, for $i = 1, \dots, d$, we can conclude that (3.9) is a Lévy measure and therefore (IV) is valid for $k(x, z)$. Uniqueness of the jump measure follows from the fact that it is uniquely determined by its marginal tail integrals (cf. [75, Lemma 3.5]). $\square$

We prove the following decay property of the jump density constructed according to the above theorem. We need these estimates later to prove exponential convergence of the numerical quadrature rules employed to approximate the discretized generator of the Feller process.

Lemma 3.2.2. *Let $k(x, z)$ be constructed according to Theorem 3.2.1. Besides, we require the following estimate on the derivatives of $k_i(x, z)$: there exists $C > 0$ s.t. $\forall x \in \mathbb{R}$, $z \in \mathbb{R}\backslash\{0\}$*

$$\begin{aligned} |\partial_x^n k_i(x,z)| &\leq C^{n+1} n! \, |z|^{-Y_i(x)-\delta n-1}\,, && (3.10)\\ |\partial_z^n k_i(x,z)| &\leq C^{n+1} n! \, |z|^{-Y_i(x)-n-1}\,, && (3.11) \end{aligned}$$

for some $\delta \in (0, 1)$ and

$$\max_{i=1,\dots,d} \sup_{x_i \in \mathbb{R}} Y_i(x_i) = \overline{Y} < 2 \text{ as well as } \min_{i=1,\dots,d} \inf_{x_i \in \mathbb{R}} Y_i(x_i) = \underline{Y} > 0.$$

Then there exists $C > 0$ such that for all $x \in \mathbb{R}^d$, $\forall z_i \neq 0$, $\forall n, m \in \mathbb{N}_0^d$,

$$|\partial_x^m \partial_z^n k(x,z)| \leq C^{|n|+1} \, |m|! \, |n|! \, \|z\|_\infty^{-\overline{Y}} \prod_{i=1}^d |z_i|^{-n_i - \delta m_i - 1}\,, \quad \forall z_i \neq 0,$$

for multiindices $n, m \in \mathbb{N}_0^d$

Proof. Using the formula of Faà di Bruno [101] it can be shown that

$$\begin{aligned} &\left|\partial_{x_i}^n \left(\partial_1 \dots \partial_d F(U(x,z))\right)\right| \\ = &\left| \sum \frac{n!}{m_1! \dots m_n!} (\partial_{x_i}^m \partial_1 \dots \partial_d F)(U(x,z)) \left(\frac{\partial_{x_i} U_i(x,z)}{1!}\right)^{m_1} \dots \left(\frac{\partial_{x_i}^n U_i(x,z)}{n!}\right)^{m_n} \right| \\ &\leq \sum C_1^{n+1} \frac{n! m!}{m_1! \dots m_n!} \|z\|_\infty^{-\overline{Y}} \prod_j^d |z_j|^{\overline{Y}} |z_i|^{\overline{Y} m} |z_i|^{-\overline{Y} m_1 - \delta m_1} \dots |z_i|^{-\overline{Y} m_n - \delta n m_n} \\ &\leq C_2^{n+1} n! \, \|z\|_\infty^{-\overline{Y}} |z_i|^{-\delta n} \prod_{j=1}^d |z_j|^{\overline{Y}}\,, \end{aligned}$$

where we sum over all multiindices $(m_1, \dots, m_n)$, $m = \sum_i m_i$ with $n = \sum_{i=1}^n i m_i$. An similar calculation leads to

$$|\partial_{z_i} \left(\partial_1 \dots \partial_d F(U(x,z))\right)| \leq C_2^{n+1} n! \, \|z\|_\infty^{-\overline{Y}} |z_i|^{-n} \prod_{j=1}^d |z_j|^{\overline{Y}}\,.$$

Using the Leibniz rule we obtain

$$\begin{aligned}
& \left|\partial_{x_i}^n k(x,z)\right| \\
= \; & \left|\partial_{x_i}^n \left(\partial_1 \dots \partial_d F(U(x,z)) k_1(x_1,z_1) \dots k_d(x_d,z_d)\right)\right| \\
= \; & \left|\sum_{j=1}^{n} \frac{n!}{j!(n-j)!} \partial_{x_i}^j \left(\partial_1 \dots \partial_d F(U(x,z))\right) \partial_{x_i}^{n-j} k_i(x_i,z_i) \prod_{m=1, m\neq d}^{d} k_m(x_m,z_m)\right| \\
\leq \; & C_3^{n+1} n! \sum_{j=1}^{n} \|z\|_\infty^{-\overline{Y}} |z_i|^{-j\delta} \prod_{j=1}^{d} |z_j|^{\overline{Y}} |z_i|^{-\overline{Y}-1+\delta(-n+j)} \prod_{m=1, m\neq i}^{d} |z_m|^{-\overline{Y}-1} \\
\leq \; & C_4^{n+1} n! \|z\|_\infty^{-\overline{Y}} |z_i|^{-n\delta} \prod_{m=1}^{d} |z_m|^{-1} .
\end{aligned}$$

Analogously it is shown for all $n \in \mathbb{N}$, $0 \neq y \in \mathbb{R}^d$ that

$$\left|\partial_{z_i}^n k(x,z)\right| \leq C_4^{n+1} n! \|z\|_\infty^{-\overline{Y}} |z_i|^{-n} \prod_{m=1}^{d} |z_m|^{-1} , \tag{3.12}$$

which completes the proof. □

3.2.2 A class of admissible market models

We now formulate the requirements for market models which are admissible for our pricing schemes in terms of the marginals and the copula function. These requirements not only ensure existence and uniqueness of a solution of the corresponding pricing problem, but also ensure that the presented FEM based algorithms are feasible.

Definition 3.2.3. *We call a d-dimensional Feller process with characteristic triplet $(\gamma(x), Q(x), \nu(x,dz))$ a time-homogeneous* admissible market model *if it satisfies the following properties.*

1. *The function $x \mapsto b(x) \in \mathbb{R}^d$ is smooth and bounded.*
2. *The function $x \mapsto Q(x) \in \mathbb{R}^{d\times d}_{sym}$ is smooth and bounded and $Q(x)$ is positive semidefinite for all $x \in \mathbb{R}^d$.*
3. *The jump measure $\nu(x,dz)$ is constructed from d independent, univariate jump measures with a 1-homogeneous copula function F that fulfills the following estimate: there is a constant $C > 0$ such that for all $u \in (\mathbb{R}\backslash\{0\})^d$ and all $n \in \mathbb{N}_0^d$ it holds*

$$|\partial^n F(u)| \leq C^{|n|+1} |n|! \min\{|u_1|, \dots, |u_d|\} \prod_{i=1}^{d} |u_i|^{-n_i}$$

4. *For the marginal densities* $\nu_i(x_i, dz_i) = k_i(x,z)dz$ *the mapping* $x_i \mapsto \nu_i(x_i, B)$ *is smooth for all* $B \in \mathcal{B}(\mathbb{R})$.

5. *There exist univariate Lévy kernels* $\overline{k}_i(z)$, $i = 1, ..., d$, *with semiheavy tails, i.e., which satisfy*

$$\overline{k}_i(z) \leq C \begin{cases} e^{-\beta^-|z|}, & z < -1 \\ e^{-\beta^+ z}, & z > 1, \end{cases} \tag{3.13}$$

for some constants $C > 0$, $\beta^- > 0$ *and* $\beta^+ > 1$. *These Lévy kernels satisfy the following estimates*

$$0 \leq \nu_i(x_i, B) \leq \int_B \overline{k}_i(z)\, dz \quad \forall x_i \in \mathbb{R}, \quad B \in \mathcal{B}(\mathbb{R}),\ i = 1, \ldots, d.$$

6. *Besides, we require the following estimate on the derivatives of* $k_i(x,z)$

$$\begin{aligned} |\partial_x^n k_i(x,z)| &\leq C^{n+1} n!\, |z|^{-Y_i(x) - \delta n - 1}, \\ |\partial_z^n k_i(x,z)| &\leq C^{n+1} n!\, |z|^{-Y_i(x) - n - 1}, \end{aligned}$$

for any $\delta \in (0,1)$, *for all* $0 \neq z, x \in \mathbb{R}$ *and* $\overline{Y} < 2$ *as well as* $\underline{Y} > 0$, $Y_i(x) = \widehat{Y}_i + \widetilde{Y}_i(x)$, $\widehat{Y}_i \in \mathbb{R}_+$ *and* $\widetilde{Y}_i(x) \in \mathcal{S}(\mathbb{R})$, $i = 1, \ldots, d$.

7. *Finally, we require* F^0 *to be a* 1*-homogeneous Lévy copula and* $k_i^0(x_i, z_i)$ *to be* $Y_i(x_i)$*-stable densities with tail integrals* $U_i^0(x_i, z_i)$, $i = 1, \ldots, d$ *such that*

$$\begin{aligned} k_i(x,z) &\geq C k_i^0(x,z), \quad \forall 0 < |z| < 1, \forall x \in \mathbb{R}, \quad i = 1, \ldots, d \\ \partial_1 \ldots \partial_n F(U(x,z)) &\geq C \partial_1 \ldots \partial_n F^0(U^0(x,z)) \quad \forall 0 < |z| < 1, \end{aligned}$$

for some constant $C > 0$.

Remark 3.2.4. *An admissible time-homogeneous market model as given in Definition 3.2.3 satisfies the requirements of Theorem 3.2.1, as* (3.4) *follows from (3),* (3.5) *from (5)-(7) and* (3.6) *holds due to (5).*

Remark 3.2.5. *Note that the smoothness assumptions on* $Y_i(x_i)$, $i = 1, \ldots, d$, *are necessary in order to obtain symbols as given in Definition 2.2.2. Such symbols are considered as therefore the results of [97] can be used, that rely on pseudodifferential calculus for symbols of variable order, cf. [63, 76]. The derivation of similar results for symbols with lower regularity is open to our knowledge.*

In the following lemma we characterize the symbol classes of admissible market models. This is crucial for the well-posedness of the pricing equation as discussed in Chapter 5.

Lemma 3.2.6. *The symbol $\psi(x,\xi)$ of a time-homogeneous admissible market model in the sense of Definition 3.2.3 given as*

$$\begin{aligned}\psi(x,\xi) &= -ib(x)\cdot\xi + \frac{1}{2}\xi\cdot Q(x)\xi \\ &\quad + \int_{0\neq z\in\mathbb{R}^d}\left(1-e^{iz\cdot\xi}+iz\cdot\xi\right)\nu(x,dz)\end{aligned}$$

is contained in the following symbol classes.

$$\begin{cases}\psi(x,\xi)\in S^2_{1,\delta} & \text{for } Q(x)\geq Q_0>0,\\ \psi(x,\xi)\in S^{\mathbf{Y}(x)}_{1,\delta} & \text{for } Q=0,\ \gamma=0,\\ \psi(x,\xi)\in S^{2\widetilde{\mathbf{m}}(x)}_{1,\delta} & \text{for } Q=0,\ \gamma\neq 0,\end{cases}$$

where $\delta\in(0,1)$ and $\widetilde{m}_i(x_i)=\frac{\max(Y_i(x_i),1)}{2}$, $i=1,\dots,d$.

Proof. We have, analogously to [95, Proposition 3.5],

$$\forall\xi\in\mathbb{R}^d:\qquad \int_{0\neq z\in\mathbb{R}^d}\left(1-e^{i\xi\cdot z}+iz\cdot\xi\right)\nu(x,dz)\leq C_1\sum_{i=1}^d|\xi_i|^{Y_i(x_i)}$$

for some positive constant $C_1,C_2,C_3>0$. The following estimate holds for the diffusion and the drift component:

$$\forall\xi\in\mathbb{R}^d:\qquad \left|\frac{1}{2}\xi\cdot Q\xi\right|\leq C_2\sum_{i=1}^d|\xi_i|^2 \text{ and } |ib(x)\xi|\leq C_3\sum_{i=1}^d|\xi_i|\,.$$

□

The removal of the drift is discussed in Chapter 5.

Remark 3.2.7. *The partially degenerate case $Q\neq 0$, but $Q\not>0$ can be analysed as in [95, Remark 4.9]. Note that in the case $\gamma\neq 0$ and $Q=0$ additional assumptions on the behavior of $Y_i(x_i)$ at 1 are necessary in order to ensure the smoothness of $\widetilde{m}_i(x_i)$, $i=1,\dots,d$.*

The infinitesimal generator $\mathcal{A}$ of a time-homogeneous admissible market model X reads

$$\begin{aligned}\mathcal{A}\varphi(x) &= \mathcal{A}_{\mathrm{Tr}}\varphi(x)+\mathcal{A}_{\mathrm{BS}}\varphi(x)+\mathcal{A}_{\mathrm{J}}\varphi(x)\\ \mathcal{A}_{\mathrm{Tr}}\varphi(x) &= b(x)\cdot\nabla\varphi,\\ \mathcal{A}_{\mathrm{BS}}\varphi(x) &= \frac{1}{2}\mathrm{tr}(Q(x)D^2\varphi(x)),\\ \mathcal{A}_{\mathrm{J}}\varphi(x) &= \int_{\mathbb{R}^d}\left(\varphi(x+z)-\varphi(x)-z\cdot\nabla\varphi(x)\right)\nu(x,dz),\end{aligned}\tag{3.14}$$

for $\varphi\in C_0^\infty(\mathbb{R}^d)$. Note that the (non-constant) symbol of the infinitesimal generator of X does generally not coincide with the characteristic exponent of the process X, due to the spatial inhomogeneity of X.

Remark 3.2.8. *Note that the jump operator $\mathcal{A}_{\mathrm{J}}$ admits a different representation for finite variation processes, i.e., for $\overline{Y} < 1$. In this situation $\int_{\mathbb{R}^d} z\nu(x, dz) < \infty$ holds and therefore the integral in* (3.14) *can be simplified to*

$$\begin{aligned} \mathcal{A}_{\mathrm{J}}\varphi(x) &= \mathcal{A}_{\mathrm{J-FV}}\varphi(x) + \mathcal{A}_{\mathrm{Tr-FV}}\varphi(x) \\ &= \int_{\mathbb{R}^d} (\varphi(x+z) - \varphi(x))\, \nu(x, dz) - \widetilde{b}(x) \cdot \nabla\varphi(x), \end{aligned} \tag{3.15}$$

where $\widetilde{b}(x) = \int_{\mathbb{R}^d} z\, \nu(x, dz)$.

The corresponding bilinear forms read $a(\varphi, \psi) = (\mathcal{A}\varphi, \psi) = a_{\mathrm{Tr}}(\varphi, \psi) + a_{\mathrm{BS}}(\varphi, \psi) + a_{\mathrm{J}}(\varphi, \psi)$, where

$$a_{\mathrm{J}}(\varphi, \psi) \;=\; (\mathcal{A}_{\mathrm{J}}\varphi, \psi), \quad a_{\mathrm{BS}}(\varphi, \psi) = (\mathcal{A}_{\mathrm{BS}}\varphi, \psi), \quad a_{\mathrm{Tr}}(\varphi, \psi) = (\mathcal{A}_{\mathrm{Tr}}\varphi, \psi). \tag{3.16}$$

The domain $D(a_{\mathrm{J}}(\cdot, \cdot))$ of the bilinear form $a_{\mathrm{J}}(\cdot, \cdot)$ is a variable order Sobolev space as introduced in Section 2.2.1, while $D(a_{\mathrm{BS}}(\cdot, \cdot)) = H^1(\mathbb{R}^d)$. This is shown by proving Gårding inequalities and continuity of the bilinear forms on the corresponding function spaces. Such an approach is not feasible for the transport bilinear form $a_{\mathrm{Tr}}(\cdot, \cdot)$. Therefore one has to take special care of the transport dominated case, i.e., a setup with vanishing diffusion and $\underline{Y} < 1$. All these cases are treated in Chapter 5. We illustrate the preceding, abstract developments with an example related to the so-called *tempered-stable* class of Lévy processes which were advocated in recent years in the context of financial modeling and a process of Ornstein-Uhlenbeck type.

Example 3.2.9 (Feller-CGMY). *We consider a d-dimensional Feller process with Clayton Lévy copula*

$$F(u_1, \ldots, u_d) = 2^{2-d} \left(\sum_{i=1}^{d} |u_i|^{\vartheta} \right)^{-\frac{1}{\vartheta}} \left(\rho \mathbb{1}_{\{u_1, \ldots, u_d \geq 0\}} - (1-\rho) \mathbb{1}_{\{u_1, \ldots, u_d \leq 0\}} \right),$$

where $\vartheta > 0$, $\rho \in [0, 1]$ together with CGMY-type densities

$$k_i(x, z) = C(x) \left(\frac{e^{-\beta_i^-(x)|z|}}{|z|^{1+Y_i(x)}} \mathbb{1}_{\{z<0\}} + \frac{e^{-\beta_i^+(x)|z|}}{|z|^{1+Y_i(x)}} \mathbb{1}_{\{z>0\}} \right),$$

with smooth and bounded functions $C(x) > 0$, $\beta_i^-(x) > 0$, $\beta_i^+(x) > 1$ and $0 < \underline{Y}_i < Y_i(x) \leq \overline{Y}_i < 2$, $Y_i(x)$ sufficiently smooth, for $i = 1, \ldots, d$. We assume the Gaussian component $Q(x)$ to be positive semidefinite, smooth and bounded. The drift $b(x)$ is assumed to be smooth and bounded. It is easy to see that this market model satisfies properties (1), (2), (4)-(6) of the above definition. (3) and (7) follow analogously to the proof of [113, Proposition 2.3.7].

Example 3.2.10. *We consider an Ornstein-Uhlenbeck process $X(t)$ with a pure jump Lévy subordinator without drift as driver, i.e.,*

$$dX(t) = -\lambda(X(t))dt + dL(t),$$

for some globally Lipschitz function $\lambda(x) : \mathbb{R} \to \mathbb{R}$. The Lévy process $L(t)$ has characteristic triplet $(0, 0, \nu(dz))$, where $\nu(dz)$ is supported on $\mathbb{R}_+$. The corresponding infinitesimal generator reads

$$\mathcal{A}\varphi(x) = \lambda(x)\varphi'(x) - \int_{\mathbb{R}_+} (\varphi(x+z) - \varphi(x))\,\nu(dz).$$

Due to the fact that $L(t)$ is a subordinator we obtain a finite variation process $X(t)$. Note that this process does not satisfy property (1) and (7) of Definition 3.2.3, the methods developed in the following are also applicable to such kind of equations. The main difficulties in the treatment of the corresponding Kolmogorov equation resides in the non-constant drift term and the support of the jump measure $\nu(dz)$. Both aspects are addressed in subsequent chapters.

3.3 Time-inhomogeneous processes

In this section we consider time-inhomogeneous models with possibly degenerate coefficients in time. First, an extension of Lévy models is discussed. The theoretical foundation is given in [77, 103], for applications we refer to [25]. We also outline the fractional Brownian motion framework. This type of models has gained popularity over the last decade, see [14], but does not fit into the semimartingale framework.

3.3.1 Time-inhomogeneous Lévy processes

Definition 3.3.1. *An adapted stochastic process X on a complete filtered probability space with values in $\mathbb{R}^d$ is called a time-inhomogeneous Lévy process, if the following conditions are satisfied:*

(i) X has independent increments, i.e., $X(t) - X(s)$ is independent of $\mathcal{F}_s$, for $0 \le s < t \le T$.

(ii) For every $t \in [0, T]$ the law of $X(t)$ is characterized by the characteristic function

$$\begin{aligned} \mathbb{E}[e^{iu\cdot X(t)}] &= \exp \int_0^t \left(iu \cdot b(s)) - \frac{1}{2} u \cdot Q(s) u \right) ds \\ &\quad \times \exp \int_0^t \left(\int_{\mathbb{R}^d} (e^{iu\cdot z} - 1 - iu \cdot z \mathbb{1}_{|z|\le 1} \nu(s, dz)) \right) ds. \end{aligned} \tag{3.17}$$

Here $b(s) \in \mathbb{R}^d$, $Q(s)$ is a symmetric nonnegative-definite $d \times d$ matrix and $\nu(s, dz)$ is a measure on $\mathbb{R}^d$ that integrates $\min\left(|z|^2, 1\right)$ and satisfies $\nu(s, \{0\}) = 0$. Besides, the coefficients satisfy

$$\int_0^T \left(|b(s)| + \|Q(s)\| + \int_{\mathbb{R}^d} (\min\left(|z|^2, 1\right))\nu(s, dz) \right) ds < \infty. \tag{3.18}$$

We call $(b(s), Q(s), \nu(s, dz))$ the characteristics of X. In the following we recall some useful results for time-inhomogeneous Lévy processes, the proofs of the following lemmas can be found in [77].

Lemma 3.3.2. *The process X as given in Definition 3.3.1 is a semimartingale.*

Proof. We repeat the argument of [77, Lemma 1.4]. We consider for $u \in \mathbb{R}^d$ the function $t \mapsto \psi(t, u)$, where $\psi(t, u)$ is given as

$$\begin{aligned} \psi(t, u) &:= \log \mathbb{E}[e^{u \cdot X(t)}] = \\ &= \exp \int_0^t \left(iu \cdot b(s) - \frac{1}{2} u \cdot Q(s) u \right) ds \\ &\quad \times \exp \int_0^t \left(\int_{\mathbb{R}^d} (e^{iu \cdot z} - 1 - iu \cdot z \mathbb{1}_{|z| \leq 1} \nu(t, dz)) \right) ds, \end{aligned}$$

with $X(t)$ as in Definition 3.3.1. Property (3.18) implies that $t \mapsto \psi(t, u)$ has finite variation over finite intervals, therefore the same applies to $t \mapsto \exp(\psi(t, u))$ and [72, Chapter II, Theorem 4.14] implies that X is a semimartingale. □

The distribution of an admissible time-inhomogeneous market model is infinitely divisible and the characteristic triplet can be characterized explicitly, this is used in Chapter 4 for the localization of the pricing problem.

Lemma 3.3.3. *For fix $t \in [0, T]$ the distribution of a time-inhomogeneous Lévy process X is infinite divisible with triplet $(b, Q, \nu(dz))$, where*

$$b := \int_0^t b(s) ds, \quad Q := \int_0^t Q(s) ds, \quad \nu(dz) := \int_0^t \nu(s, dz) ds. \tag{3.19}$$

The integrals are to be understood componentwise.

Proof. For completeness we repeat the argument of [77, Lemma 1.2]. From (3.19) we conclude that $b \in \mathbb{R}^d$ and that Q is an nonnegative definite $d \times d$ matrix. A monotone convergence argument yields that $\nu(dz)$ is a measure on the Borel sets of $\mathbb{R}^d$ and we get

$$\int_{\mathbb{R}^d} f(z) \nu(dz) = \int_0^t \int_{\mathbb{R}^d} f(z) \nu(s, dz) ds, \tag{3.20}$$

for any integrable function f. Therefore, $\int_{\mathbb{R}^d} \min\left(|z|^2, 1\right) \nu(dz) < \infty$ and $\nu(\{0\}) = 0$. The claim follows from (3.17) and the Lévy-Khintchine formula. □

The following assumption on the jump measure of the process X is made.

Assumption 3.3.4. *There exists a constant $R > 0$ such that for $u \in [-R, R]^d$,*

$$\int_0^T \int_{|z|>1} \exp{(u \cdot z)} \nu(s, dz) ds < \infty. \tag{3.21}$$

As in the Lévy case, Assumption 3.3.4 implies the existence of exponential moments, i.e., the process has semi-heavy tails.

Lemma 3.3.5. *Assumption 3.3.4 holds if and only if there exists a constant $R > 0$, such that $\mathbb{E}[\exp{(u \cdot X(t))}] < \infty$ for all $t \in [0, T]$ and $u \in [-R, R]^d$.*

Proof. See [77, Lemma 1.6]. Assume that (3.21) holds and fix $u \in [-R, R]^d$ and $t \in [0, T]$. Let Y be a Lévy process with $Y(1) \overset{(d)}{=} X(t)$. Then the generating triplet of Y is given by (3.19) due to Lemma 3.3.3. Therefore we have

$$\int_{|z|>1} \exp{(u \cdot z)} \nu(dz) < \infty.$$

It follows from [103, Theorem 25.3] that $\mathbb{E}[\exp{(u, Y(1))}] < \infty$ holds and since $Y(1) \overset{(d)}{=} X(t)$, we conclude $\mathbb{E}[\exp u \cdot X(t)] < \infty$. The converse results follows similarly. $\square$

We are now concerned with the derivation of a drift condition in order to ensure that the pricing process is a discounted exponential martingale. Let us denote by $\Theta(s, y)$ the cumulant function of X, i.e.,

$$\Theta(s, y) = y \cdot b(s) + \frac{1}{2} y \cdot Q(s) y + \int_{\mathbb{R}^d} (e^{y \cdot z} - 1 - y \cdot z) \nu(s, dz).$$

Lemma 3.3.6. *Fix $t \in [0, T]$. For $y \in \mathbb{C}^d$ with $\Re(y) \in [-R, R]$ we have $\mathbb{E}[|e^{y \cdot X(t)}|] < \infty$ and*

$$\mathbb{E}[e^{y \cdot X(t)}] = \exp\left(\int_0^t \Theta(s, y) ds \right).$$

Proof. For completeness we repeat the proof of [77, Lemma 1.8]. Lemma 3.3.5 implies $\mathbb{E}[|e^{y \cdot X(t)}|] = \mathbb{E}[e^{\Re y \cdot X(t)}]$. Let Y be the Lévy process with $Y(1) \overset{(d)}{=} X(t)$, then $\mathbb{E}[e^{\Re y \cdot Y(1)}] = \mathbb{E}[e^{\Re y \cdot X(t)}] < \infty$, where the characteristic triplet of Y is given in (3.19). We obtain from [103, Lemma 25.17] $\mathbb{E}[e^{y \cdot Y(1)}] = e^{\psi(y)}$ with

$$\psi(y) := y \cdot b + \frac{1}{2} y \cdot Q y + \int_{\mathbb{R}} (e^{y \cdot z} - 1 - y \cdot z) \nu(dz).$$

We conclude $\psi(y) = \int_0^t \Theta(s, y) ds$ using (3.20) and thus $\mathbb{E}[e^{y \cdot X(t)}] = \exp\left(\int_0^t \Theta(s, y) ds \right)$. $\square$

Therefore, a sufficient condition in order to ensure that the processes X_i, $i = 1, \dots, d$, are exponential martingales reads:

$$\int_0^t \Theta(s, e_i)\, ds = 0 \quad i = 1, \dots, d, \quad t > 0.$$

In addition to Assumption 3.3.4, we need the following requirements for an admissible market model.

Definition 3.3.7. *We call a process X as in Definition 3.3.1 an admissible time-inhomogeneous market model if it satisfies Assumption 3.3.4 and if*

(i) the diffusion coefficient is $Q(s) = \alpha(s)\widetilde{Q}(s)$, where $\alpha(s) = s^\gamma$, $\gamma \in (-1, 1)$ and $\overline{Q}\,\|\xi\|^2 \geq \xi^\top \widetilde{Q}(s)\xi \geq \underline{Q}\,\|\xi\|^2$, $\overline{Q}, \underline{Q} > 0$, $s \in [0, T]$.

(ii) the drift coefficient is $b(s) = \alpha(s)\widetilde{b}(s)$ and $\left|\widetilde{b}(s)\right| \leq \overline{b}$, $\overline{b} > 0$, $s \in [0, T]$.

(iii) for the jump measure $\nu(s, dz)$, we assume the decomposition $\nu(s, dz) = k(s, z)dz$, where

$$0 \leq k(s, z) = \alpha(s)\widetilde{k}(s, z)$$

and $\overline{k}(z) \geq \widetilde{k}(s, z)$, for a Lévy measure $\overline{k}(x)$, satisfying

$$\overline{k}(z) \leq k^0(z), \text{ for } |z| \in (0, 1),$$

and an α-stable Lévy measure $k^0(z)$, $\alpha \in [0, 2)$.

(iv) the density $k(s, z)$, for $s \in (0, T]$ is real analytic outside $z_i = 0$, $i = 1, \dots, d$,

$$|\partial^{\mathbf{n}} k(s, z)| \leq C^{|n|}(s)\, |n|!\, \|z\|_\infty^{-\alpha} \prod_{i=1}^d |z_i|^{-n_i - 1}, \quad \forall z_i \neq 0, i = 1, \dots, d.$$

Remark 3.3.8. *The assumptions (iii) and (iv) in Definition 3.3.7 can be reduced to requirements on the marginals of the process and the copula function, if such a construction is chosen for the process. The constant $C(s)$ in assumption (iv) of Definition 3.3.7 may degenerate as $s \downarrow 0$. The decompositions of the diffusion coefficient, the drift coefficient and the density of the jump measure can be extended to different functions $\alpha(s)$ such as $\alpha(s) = \sum_{i=1}^m s^{\gamma_i}$, for some $m \in \mathbb{N}$ and $\gamma_i \in (-1, 1)$, $i = 1, \dots, d$.*

The infinitesimal generator of X is given as

$$\begin{aligned}
\mathcal{A}(t) &= \mathcal{A}_{\mathrm{BS}}(t) + \mathcal{A}_{\mathrm{J}}(t) + \mathcal{A}_{\mathrm{Tr}}(t), \\
\mathcal{A}_{\mathrm{Tr}}(t)\varphi(x) &= b(t) \cdot \nabla\varphi(x), \\
\mathcal{A}_{\mathrm{BS}}(t)\varphi(x) &= \frac{1}{2}\mathrm{tr}(Q(t)D^2\varphi(x)), \\
\mathcal{A}_{\mathrm{J}}(t)\varphi(x) &= \int_{\mathbb{R}^d} (\varphi(x+z) - \varphi(x) - z \cdot \nabla\varphi(x))\, \nu(t, z)dz.
\end{aligned}$$

Remark 3.3.9. *Note that the functions $Q(t)$, $b(t)$ and $\nu(t,z)$ may degenerate as $t \downarrow 0$, therefore special care has to be taken to achieve an appropriate time discretization. We consider a weak formulation of the Kolmogorov equation in time, therefore resolving the singularity.*

We define the bilinear forms $a_{\mathrm{Tr}}(t,\cdot,\cdot)$, $a_{\mathrm{BS}}(t,\cdot,\cdot)$ and $a_{\mathrm{J}}(t,\cdot,\cdot)$ for $\phi(x), \psi(x) \in C_0^\infty(\mathbb{R}^d)$ as follows

$$\begin{aligned} a_{\mathrm{Tr}}(t;\phi,\psi) &= (\mathcal{A}_{\mathrm{Tr}}(t)\phi,\psi), \\ a_{\mathrm{J}}(t;\phi,\psi) &= (\mathcal{A}_{\mathrm{J}}(t)\phi,\psi), \\ a_{\mathrm{BS}}(t;\phi,\psi) &= (\mathcal{A}_{\mathrm{BS}}(t)\phi,\psi). \end{aligned}$$

The domain of the bilinear form $a_{\mathrm{J}}(t;\cdot,\cdot)$ is the fractional order Sobolev space $\mathcal{D}(a_{\mathrm{J}}(t;\cdot,\cdot)) = H^{\alpha/2}(\mathbb{R})$, for $t > 0$. However, continuity and the Gårding inequality cannot be proved with a uniform constant C, independent of t, for all $t \in [0,T]$. Similarly we obtain for $a_{\mathrm{BS}}(t;\cdot,\cdot)$ the domain $\mathcal{D}(a_{\mathrm{BS}}(t;\cdot,\cdot)) = H^1(\mathbb{R})$, for $t > 0$, and no uniform continuity and Gårding inequality bound. In order to account for the degenerate behavior of the solution in time we use weighted Sobolev spaces. To illustrate the above definitions, we present the following example similar to [25].

Example 3.3.10 (Time-inhomegeous tempered-stable model). *We consider the univariate CGMY-type model X with characteristic $(b(t), Q(t), k(t,y))$ given as*

$$\begin{aligned} b(t) &= t^{Y\gamma-1} b \\ Q(t) &= Q t^{Y\gamma-1} \\ k(t,y) &= \begin{cases} C\dfrac{\exp\left(-\frac{M|y|}{t^\gamma}\right)t^{\gamma Y-1}}{|y|^{Y+1}} & y > 0 \\ C\dfrac{\exp\left(-\frac{G|y|}{t^\gamma}\right)t^{\gamma Y-1}}{|y|^{Y+1}} & y < 0 \end{cases}, \end{aligned}$$

for $Y \in (0,2)$, $\gamma \in (-1,1)$, $C, Q > 0$, $M, G > 1$, $b \in \mathbb{R}$. X is an admissible time-inhomogeneous market model. Similar models arise in the context of self-similar processes as proposed by [25], where it was empirically shown that such models provide a better fit to data, especially over different maturities, than Lévy type models.

3.3.2 Fractional Brownian motion

Similar pricing equations to those considered in the previous section arise in the context of fractional Brownian motion models. We hasten to point out that the derivation of pricing equations for such type of market models involves the use of Wick calculus, which is controversial in the financial context, cf. [15]. We briefly outline the main steps in the derivation of the pricing problems. Let $(\Omega, \mathcal{F}, \mathbb{F}^H, \mathbb{P})$ be a complete probability space supporting a real-valued fractional Brownian motion (FBM) $B_H(t)$ with Hurst parameter $H \in (0,1)$ and let $\mathcal{F}_t^H$ be the σ-algebra generated by $B_H(s)$, $s \leq t$.

Definition 3.3.11. *For $H \in (0,1)$, a fractional Brownian motion B_H is a Gaussian process with mean zero, i.e.,*

$$\mathbb{E}[B_H(t)] = 0$$

for all $t \geq 0$ and covariance:

$$\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\},$$

for all $s, t \geq 0$. We assume $B_H(0) = 0$. For $H = \frac{1}{2}$ we obtain a standard Brownian motion.

Our market model reads as follows. If $S(t)$ denotes the spot price of the risky asset, then its dynamics under the real world measure $\mathbb{P}$ is given as:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB_H(t), \quad t \geq 0, \quad S(0) = s > 0. \tag{3.22}$$

For the notion of a stochastic integral with respect to a fractional Brownian motion $B_H(t)$ we refer to [50] and [66]. Besides, we assume the existence of a risk free bank account $P(t)$ with risk free interest rate $r > 0$. With the Girsanov theorem for FBM, cf. [9, Theorem 2.8] or [66, Theorem 3.18], we obtain the risk adjusted dynamics of the stock $S(t)$ under the equivalent measure $\mathbb{Q}$:

$$dS(t) = rS(t)dt + \sigma S(t)d\widetilde{B}_H(t), \quad t \geq 0, \quad S(0) = s > 0,$$

where $\widetilde{B}_H(t)$ is a fractional Brownian motion under $\mathbb{Q}$ and the discounted stock is a quasi-martingale under $\mathbb{Q}$. We refer to [9, Definition 2.3] for the definition of quasi-conditional expectation and quasi-martingales. Note that $\mathbb{Q}$ is not a martingale measure as the stock is not a martingale under $\mathbb{Q}$. Let $g(S)$ be the payoff of a European type contingent claim V, for sufficiently smooth g. Its value at time t before maturity is given as the discounted quasi-conditional expectation:

$$V(t) = e^{-r(T-t)}\widetilde{\mathbb{E}}_{\mathbb{Q}}[g(S_T)|\mathcal{F}_t^H], \tag{3.23}$$

cf. [9, Theorem 4.2] and [49, Proposition 1]. The option price $V(t)$ admits a PDE representation.

Theorem 3.3.12. *Let $v \in C^{1,2}((0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$ such that $v : [0,T]\times\mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following PDE:*

$$\partial_t v(t,S) + rS\partial_S v(t,S) + H\sigma^2 t^{2H-1}S^2\partial_{SS}v(t,S) - rv(t,S) = 0 \tag{3.24}$$

with terminal condition $v(T,S) = g(S)$, then

$$v(t,S) = V(t,S) \quad \text{for all } t \in [0,T],\ S \in \mathbb{R}_+.$$

Proof. The result follows from [49, Proposition 2] and [9, Proposition 6.1]. □

Remark 3.3.13. *The PDE* (3.24) *can be simplified via a change of variable to a PDE of the following type for* $u(t,x) \in C^{1,2}((0,T)\times\mathbb{R}) \cap C^0([0,T]\times\mathbb{R})$

$$\begin{aligned} \partial_t u(t,x) - Q(T-t)^\gamma \partial_{xx} u(t,x) &= f(t,x) \ \textit{on} \ (0,T)\times\mathbb{R}, \\ u(0,x) &= 0 \ \textit{on} \ \mathbb{R}. \end{aligned}$$

For some constant $Q > 0$ *and an appropriate right hand side* $f(t,x)$*. We refer to [96, Section 5.2.1] for details.*

4 Small jump regularization and localization

In this chapter probabilistic results for the small jump regularization and the localization are presented. These are not based on the parabolic integro-differential equation (PIDE) representation of the option price, but are useful for the analysis of the PIDE, since the probabilistic estimates can be used to obtain error bounds for the numerical solution of the equation. This is done in two steps of the discretization. First, an infinite activity Markov process is approximated by a finite activity process adding an appropriately scaled diffusion. Second, the PIDE formulated on an unbounded domain is localized to a bounded domain. The rigorous justification of both steps using purely numerical analysis methods without any probabilistic tools is much more tedious and technical.

4.1 Small jump approximation

4.1.1 General results

We consider time-homogeneous and time-inhomogeneous Markov processes X as defined in Sections 3.2 and 3.3.1, with characteristic triplets that satisfy Assumptions 3.2.3 and 3.3.7. The easiest approach to the approximation of the jump measure consists in a truncation of $\nu(dz)$ in a small ball around the origin, i.e., we consider the jump measure $\nu^\varepsilon(dz) := \mathbb{1}_{|z|>\varepsilon}\nu(dz)$, $\nu_\varepsilon := \nu - \nu^\varepsilon$, with $\varepsilon > 0$. We denote the process with characteristic triplet $(b, 0, \nu^\varepsilon(dz))$ by Y^ε. We can also approximate the small jumps by an appropriately scaled Brownian motion, i.e., we consider the process Z^ε with characteristic triplet $(b, Q_\varepsilon, \nu^\varepsilon(dz))$, where $Q_\varepsilon = \int_{\mathbb{R}^d} zz^\top\,\nu_\varepsilon(dz)$. The following approximation result for Lévy processes is well known, cf. [32, Theorem 3.1].

Theorem 4.1.1. *Let X be a Lévy process in $\mathbb{R}^d$ with characteristic triplet $(b, 0, \nu(dz))$ and let the decomposition $\nu = \nu^\varepsilon + \nu_\varepsilon$ be given. Assume that Q_ε is non singular for every $\varepsilon > 0$ and that for every $\delta > 0$ there holds*

$$\int_{(Q_\varepsilon^{-1}z,z)>\delta} (Q_\varepsilon^{-1}z, z)\nu_\varepsilon(dz) \to 0, \quad \textit{as } \varepsilon \to 0.$$

Assume further that for some family of non-singular matrices $\{\Sigma_\varepsilon\}_{\varepsilon\in(0,1]}$ there holds

$$\Sigma_\varepsilon^{-1} Q_\varepsilon \Sigma_\varepsilon^{-\top} \to I, \quad \textit{as } \varepsilon \to 0,$$

where I denotes the identity matrix in $\mathbb{R}^d$. Then for all $\varepsilon \in (0,1]$ there exists an $\mathbb{R}^d$-valued càdlàg process R^ε and a process $Z^\varepsilon = (Z^{\varepsilon,1}, \dots, Z^{\varepsilon,d})$ with characteristic triplet $(b, Q_\varepsilon, \nu^\varepsilon)$ such that

$$X(t) \overset{(d)}{=} Z^\varepsilon(t) + R^\varepsilon(t),$$

in the sense of equality of finite dimensional distributions.
Furthermore, we have for all $T > 0$, $\sup_{t\in[0,T]} \left|\Sigma_\varepsilon^{-1} R^\varepsilon(t)\right| \overset{(\mathbb{P})}{\to} 0$, as $\varepsilon \to 0$.

Remark 4.1.2. *Note that the assumption on the matrices Σ_ε can be expressed in terms of the jump measure ν, cf. [32, Theorem 2.4]. An analogous result can be shown for time-inhomogeneous Lévy processes using the fact that for any X as in Section 3.3.1 and any $t > 0$ there exists a Lévy process Y such that $X(t) \overset{(d)}{=} Y(1)$ holds, cf. Lemma 3.3.3.*

Throughout, X is generally not a Lévy process due to the non-constant drift, the spatially non-homogeneous jump measure and the possible temporal inhomogeneity, therefore a more general result is needed. A weaker convergence result in mean square sense also holds for more general Markov processes, cf. [10, Proposition 3.3].

Theorem 4.1.3. *Let X be an $\mathbb{R}^d$-valued time-homogeneous Markov process with characteristic triplet $(b(x), 0, \nu(x, dz))$ such that the following properties hold*

(i) $b(x)$ is continuous and satisfies the linear growth condition, i.e., for some constant $C > 0$

$$|b(x) - b(y)|^2 \le C\,|x-y|^2\,, \quad x, y \in \mathbb{R}^d.$$

(ii) The jump measure can be decomposed as follows $\nu(x, dz) = \gamma(x)\widetilde{\nu}(dz)$, where $\widetilde{\nu}(dz)$ and $\gamma(x)$ satisfy

$$\int_{\mathbb{R}^d} |\gamma(x) - \gamma(y)|^2\, |z|^2\, \widetilde{\nu}(z) < C\,|x-y|^2\,, \quad x, y \in \mathbb{R}^d.$$

Then there holds

$$\mathbb{E}\left[\int_0^T |X(t) - Z^\varepsilon(t)|^2\, dt\right] \le C \sum_{i=1}^d \int_{|z_i|<\varepsilon} z_i^2 \widetilde{\nu}_i(dz_i),$$

for sufficiently small $\varepsilon > 0$ and a constant C independent of ε. We denote by Z^ε the Markov process with characteristic triplet $(b(x), Q_\varepsilon(x), \nu^\varepsilon(x, dz))$, where $\nu^\varepsilon(x, dz) := \mathbb{1}_{|z|>\varepsilon}\nu(x, dz)$, $\nu_\varepsilon(x, dz) := \nu(x, dz) - \nu^\varepsilon(x, dz)$ and $Q_\varepsilon(x) = \int_{\mathbb{R}^d} zz^\top\, \nu_\varepsilon(x, dz)$.

Remark 4.1.4. *Note that requirement (ii) of Theorem 4.1.3 does generally not hold for time-homogeneous admissible market models. The decomposition of the jump measure into the speed function $\gamma(x)$ and the Lévy measure $\widetilde{\nu}(dz)$ is a strong requirement excluding the Feller-CGMY process as in Example 3.2.9. Estimates for the small jump approximation seem not to be available for the general case, when the jump measure does not separate.*

For applications to option pricing we are mainly interested in weak convergence estimates.

4.1.2 Estimates for (time-inhomogeneous) Lévy processes

Let $X = (X^1, \dots, X^d)$ be a Lévy process with characteristic triplet $(b, 0, \nu(dz))$, such that $\nu(dz)$ satisfies (5) – (6) of Definition 3.2.3, where b is chosen such that $e^{X^1}, \dots, e^{X^d}$ are martingales. Now we consider the process $\widetilde{Z}^\varepsilon = (\widetilde{Z}^{\varepsilon,1}, \dots, \widetilde{Z}^{\varepsilon,d})$ with characteristic triplet $(b^\varepsilon, Q_\varepsilon, \nu^\varepsilon(dz))$, where $\nu^\varepsilon(dz)$ and Q_ε are chosen as above and b^ε is chosen such that $e^{\widetilde{Z}^{\varepsilon,1}}, \dots, e^{\widetilde{Z}^{\varepsilon,d}}$ are martingales. Convergence of $\widetilde{Z}^\varepsilon$ to X in an appropriate sense follows from Theorem 4.1.1.

Lemma 4.1.5. *Let the payoff function P be globally Lipschitz, then we obtain the following estimate for $t \in [0, T]$ using $U^\varepsilon(t) = X(t) + (b^\varepsilon - b)t$, where X is a Lévy process with characteristic triplet $(b, 0, \nu(dz))$.*

$$|\mathbb{E}[P(x + X(T))] - \mathbb{E}[P(x + U^\varepsilon(T))]| \leq C \sum_{j=1}^{d} \int_{-\varepsilon}^{\varepsilon} |z_j|^3 \, \nu_j(dz_j), \quad \forall x \in \mathbb{R}^d. \tag{4.1}$$

Proof. This estimate can be obtained by Taylor expansion of e^x around 0 and is given in [113, Proposition 8.2.1]. For the univariate case we refer to [33, Theorem 5.1]. □

Lemma 4.1.6. *If $P \in C^4(\mathbb{R}^d)$ and $\overline{Y} := \max_{i=1,\dots,d} Y_i < 1$, there holds:*

$$\left|\mathbb{E}[P(x + U^\varepsilon(T))] - \mathbb{E}[P(x + \widetilde{Z}^\varepsilon(T))]\right| \leq C \sum_{j=1}^{d} \int_{-\varepsilon}^{\varepsilon} |z_j|^3 \, \nu_j(dz_j), \quad \forall x \in \mathbb{R}^d. \tag{4.2}$$

If $\overline{Y} < 2$ is assumed, then the following estimate holds

$$\left|\mathbb{E}[P(x + U^\varepsilon(T))] - \mathbb{E}[P(x + \widetilde{Z}^\varepsilon(T))]\right| \leq C \sum_{j=1}^{d} \int_{-\varepsilon}^{\varepsilon} |z_j|^2 \, \nu_j(dz_j), \quad \forall x \in \mathbb{R}^d. \tag{4.3}$$

Proof. The estimates (4.2)–(4.3) follow using Taylor expansion of P and Jensen's inequality. Note that the existence of first moments of the jump measure (which is a consequence of $\overline{Y} < 1$) is essentially used in the first part of the proof, cf. [113, Proposition 8.2.3.]. □

Remark 4.1.7. *Intermediate cases, i.e., $\overline{Y} \in (1, 2)$, lead to analogous estimates.*

Finally we obtain the following result from (4.1) - (4.3).

Theorem 4.1.8. *Let X and $\widetilde{Z}^\varepsilon$ be as above and $P \in C^4(\mathbb{R}^d)$, further let $u(t,x) = \mathbb{E}[P(x+X(T-t))]$ and $u^\varepsilon(t,x) = \mathbb{E}[P(x+\widetilde{Z}^\varepsilon(T-t))]$ be the solutions of the corresponding Kolmogorov equations, then the following estimate can be obtained*

$$|u(t,x) - u^\varepsilon(t,x)| \le C \begin{cases} \varepsilon^{3-\overline{Y}}, & \overline{Y} \in (0,1) \\ \varepsilon^{2-\overline{Y}}, & \overline{Y} \in (0,2) \end{cases}, \forall x \in \mathbb{R}^d.$$

Remark 4.1.9. *Note that Theorem 4.1.8 yields at least quadratic convergence with respect to ε for the L^∞-error for processes of finite variation and payoffs $P \in C^4(\mathbb{R}^d)$. Using merely a small jump truncation to approximate the process X without an artificial diffusion would lead to an approximation rate of $\varepsilon^{2-\overline{Y}}$, for all $\overline{Y} \in (0,2)$, cf. [113, Corollary 8.2.5].*

Remark 4.1.10. *The constant C in Theorem 4.1.8 depends on the tail behavior and the moments of the jump measure as well as the time to maturity.*

Remark 4.1.11. *Analogous results to Theorem 4.1.8 can be obtained for time-inhomogeneous Lévy processes X as in Section 3.3.1.*

Due to Lemma 3.3.3, there exists for any time-inhomogeneous Lévy process X and any $t \in [0,T]$ a Lévy process Y, such that $X(t) \overset{(d)}{=} Y(1)$. In the context of option pricing we are interested in the computation of moments and therefore an equality in distribution suffices for our purposes.

Theorem 4.1.12. *Let $g \in C^4(\mathbb{R}^d)$ and let X be an admissible time-inhomogeneous market model with characteristic triplet $(b(s), 0, \nu(s,dz))$, then*

$$\begin{aligned} |\mathbb{E}[g(x+X(t))] - \mathbb{E}[g(x+Y_2^\varepsilon(t))]| &\le C(t)\varepsilon^{2-\max\{Y_1,\dots,Y_d\}}, \quad \forall x \in \mathbb{R}^d, \quad Y_i \in (0,2), \\ |\mathbb{E}[g(x+X(t))] - \mathbb{E}[g(x+Z_2^\varepsilon(t))]| &\le C(t)\varepsilon^{3-\max\{Y_1,\dots,Y_d\}}, \quad \forall x \in \mathbb{R}^d, \quad Y_i \in (0,1), \end{aligned}$$

with $\varepsilon > 0$, where we denote the process with characteristic triplet $(b(s), 0, \nu^\varepsilon(s,dz))$, $\nu^\varepsilon(s,dz) := \mathbb{1}_{|z|>\varepsilon}\nu(s,dz)$, by Y_2^ε and by Z_2^ε the process with characteristic triplet $(b(s), Q_\varepsilon(s), \nu^\varepsilon(s,dz))$, $Q_\varepsilon(s) = \int_{\mathbb{R}^d} zz^\top\,\nu_\varepsilon(s,dz)$.

Proof. The result is a direct consequence of Theorem 4.1.8 in conjunction with Lemma 3.3.3. $\square$

4.1.3 Estimates for time-homogeneous Markov processes

The described procedure is not directly applicable to general Feller processes. We can use Theorem 4.1.3 to obtain a weaker error bound.

Lemma 4.1.13. *Let P be globally Lipschitz and let X and Z^ε be as in Theorem 4.1.3, then the following estimate holds:*

$$|\mathbb{E}[P(X(T))] - \mathbb{E}[P(Z^\varepsilon(T))]| \leq C \sum_{i=1}^{d} \int_{|z_i|<\varepsilon} z_i^2 \widetilde{\nu}_i(dz).$$

Proof. Using the Lipschitz continuity of P and Jensen's inequality, we obtain

$$|\mathbb{E}[P(X(T))] - \mathbb{E}[P(Z^\varepsilon(T))]| \leq K \sum_{i=1}^{d} \mathbb{E}\left[\left|Z^{\varepsilon,i}(T) - X^i(T)\right|\right].$$

The result follows from the Cauchy–Schwarz inequality and Theorem 4.1.3. □

Theorem 4.1.14. *Let X and Z^ε be as above and let P be globally Lipschitz, let further $u(t,x) = \mathbb{E}[P(X(T))|X(t) = x]$ and $u^\varepsilon(t,x) = \mathbb{E}[P(Z^\varepsilon(T))|Z^\varepsilon(t) = x]$. Then, as $\varepsilon \to 0$ the following estimate can be obtained*

$$|u(t,x) - u^\varepsilon(t,x)| \leq C\varepsilon^{2-\overline{Y}}, \quad \forall x \in \mathbb{R}^d, \quad \forall \overline{Y} \in (0,2).$$

Proof. This is a direct consequence of Lemma 4.1.13. □

Remark 4.1.15. *Note that Theorem 4.1.14 yields, in contrast to Theorem 4.1.8, at least linear convergence with respect to ε for the L^∞-error for processes of finite variation with globally Lipschitz payoffs. An analogous estimate can be obtained if merely a small jump truncation, without regularization, is employed.*

4.2 Localization

In the following we estimate the error due to localization of the Kolmogorov equation. This is necessary as the Galerkin discretization is performed on the localized problem. It turns out that the localization error decays exponentially with increasing domain under certain assumptions. We assume the payoff P to satisfy the following polynomial growth condition:

$$P(s) \leq C \left(\sum_{i=1}^{d} |s_i| + 1 \right)^q, \quad \text{for all } s \in \mathbb{R}^d, \tag{4.4}$$

for some constant $C > 0$. The variable s denotes the state variable in a real price model and the exponential of the state variable in a log-price model. The condition is satisfied for all standard multi-asset options like basket, maximum or best-of options. We consider log-price models with $\log(s_i) = x_i$, $i = 1, \ldots, d$, in the following.
The unbounded domain $\mathbb{R}^d$ of x is truncated to a bounded domain $G_R = [-R,R]^d$. In terms of financial modeling, this corresponds to the approximation of an option by the

corresponding double barrier option. In the following we consider two cases. First we derive a localization error estimate for tempered Lévy market models and then extend this to tempered affine market models.

Theorem 4.2.1. *Let the payoff function $P : \mathbb{R}^d \to \mathbb{R}$ satisfy (4.4). Further, let X be a Lévy process with state space $\mathbb{R}^d$ and let the Lévy measure $\nu(dz)$ satisfy (3) – (7) of Definition 3.2.3 with $\beta_i^+, \beta_i^- > q$, where $q > 0$ is as in (4.4). Then*

$$|u(t,x) - u_R(t,x)| \leq Ce^{-\alpha R + \beta \|x\|_\infty},$$

for $0 < \alpha < \min_i \min(\beta_i^+, \beta_i^-) - q$ and $\beta = \alpha + q$, $C > 0$,

$$u_R(t,x) = \mathbb{E}[P(e^{X(T)})\mathbb{1}_{T < \tau_{G_R}} | X(t) = x],$$

and $\tau_{G_R} = \inf\{t \geq 0 | X(t) \in G_R^c\}$, where G_R^c is the complement set of G_R.

Proof. See [95, Theorem 4.14]. □

There holds a corresponding result for affine models.

Theorem 4.2.2. *Let X be a time-homogeneous Markov process with a finite variation jump measure, we set $b(x) = (-b_1 x_1, \ldots, -b_d x_d)$, for some constants $b_1, \ldots, b_d \in \mathbb{R}^+$, $x \in \mathbb{R}^d$. Further let $\nu(dz)$ and P be as in Theorem 4.2.1, then the following estimate holds:*

$$|u(t,x) - u_R(t,x)| \leq Ce^{-\alpha R + \beta \|x\|_\infty}, \tag{4.5}$$

where α, β as in Theorem 4.2.1, $C > 0$.

Proof. The idea of the proof is to reduce this problem to the setting discussed in Theorem 4.2.1. The solution for this SDE is given by for $t \in [0,T]$:

$$X_i(t) = X_{i,0} e^{-tb_i} + \int_0^t e^{-(t-u)b_i} dL_i(u), \quad i = 1, \ldots, d. \tag{4.6}$$

The process $X_i(t)$ can be estimated pathwise as follows:

$$|X_i(t)| \leq |X_{i,0}| + \max\left(\int_0^t dL_i^+(u), -\int_0^t dL_i^-(u)\right). \tag{4.7}$$

Therefore we obtain the following estimate:

$$|u(t,x) - u_R(t,x)| = \mathbb{E}[P(e^{X(T)})\mathbb{1}_{\{T \geq \tau_{G_R}\}} | X(t) = x] \leq \mathbb{E}[e^{qM_T}\mathbb{1}_{\{M_T > R\}} | X(t) = x],$$

where $M_T = \sup_{s \in [t,T]} \|X_s\|_\infty$. It follows using (4.7) :

$$\mathbb{E}[e^{qM_T}\mathbb{1}_{\{M_T > R\}} | X(t) = x] \leq \mathbb{E}[e^{q\widetilde{M}_T}\mathbb{1}_{\{\widetilde{M}_T > R\}} | X(t) = x],$$

for $\widetilde{M}_T = \|X(t)\|_\infty + \sup_{s\in[t,T]} \max\{L^+(s), -L^-(s)\}$.

$$
\begin{aligned}
& \mathbb{E}[e^{q\widetilde{M}_T} \mathbb{1}_{\{\widetilde{M}_T > R\}} | X(t) = x] \\
\leq \; & \mathbb{E}[e^{q\widetilde{M}_T^+} \mathbb{1}_{\{\widetilde{M}_T^+ > R\}} | X(t) = x] + \mathbb{E}[e^{q\widetilde{M}_T^-} \mathbb{1}_{\{\widetilde{M}_T^- > R\}} | X(t) = x]. \qquad (4.8)
\end{aligned}
$$

Both terms in (4.8) can be estimated analogously to Theorem 4.2.1, which yields the claimed result. □

Remark 4.2.3. *Similar results can also be obtained for more general drift functions. E.g. for* $(b_1, \ldots, b_d) \in \mathbb{R}^d$*, an analogous estimate to* (4.5) *under stricter assumptions on* β_i^+*,* β_i^-*,* $i = 1, \ldots, d$*, and for different constants* α *and* β *can be obtained.*

Remark 4.2.4. *Localization estimates for general time-inhomogeneous processes do not seem to be available or easy to obtain. The reason for this resides in the fact that a technical result similar to [103, Theorem 25.3] is not available. The localization error can be easily quantified for time-inhomogeneous diffusions, as in this situation the probability density function is known. Besides, the localization error can be estimated using local times as in [96].*

5 Well-posedness of time-homogeneous PIDEs

A key observation of partial integro-differential equation approaches to deterministic computational pricing of derivative contracts in finance is the observation (going back at least to R. Feynman and M. Kac, cf. [53, 73, 74]) that conditional expectations over all sample paths of a multivariate diffusion process satisfy deterministic, parabolic partial differential equations (PDEs). The most well known representative of these PDEs in financial modeling is the classical Black-Scholes equation. This Feynman–Kac correspondence holds in a much more general context, the deterministic equation being in general nonlinear, and the solution being in general understood as *viscosity solution*. Here, we follow on linear differential equations for which the (unique) solutions are *variational solutions* of suitable *weak formulations* of the deterministic evolution equations. As these formulations form the basis of variational discretizations to be discussed below, we shall present their ingredients (Sobolev spaces and Dirichlet forms, evolution triplets, and the abstract theory of parabolic evolution equations) in some detail here.

We first prove a sector condition for time-homogeneous admissible market models, which is crucial for the proof of well-posedness of the pricing equation. Subsequently, well-posedness results for certain types of pricing equations are given. These equations are of parabolic type, with the highest order operator being the diffusion or jump part of the generator of the market model. In Section 5.3, we address the well-posedness for drift dominated equations. Finally, existence and uniqueness results for pricing problems arising from the small jump regularization, as given in Chapter 4, are described.

5.1 Sector condition

The sector condition for the symbol $\psi(x,\xi)$ of a Feller process X is one of the main ingredients for proving well-posedness of the initial boundary value problems for the PIDEs arising in option pricing problems. The sector condition reads:

$$\exists\, C > 0 \quad \text{s.t. } \forall x, \xi \in \mathbb{R}^d : \qquad \Re\,\psi(x,\xi) + 1 \geq C\langle\xi\rangle^{\mathbf{2m}(x)}. \tag{5.1}$$

We use the following notation $\langle\xi\rangle^{\mathbf{m}(x)} := \sum_{i=1}^{n}(1+\xi_i^2)^{\frac{1}{2}m_i(x)}$. Verification of the sector condition is not straightforward for a general Feller process. Here, we give sufficient conditions for the sector condition to hold in terms of appropriate conditions on the marginals of the Feller process and the copula function.

Definition 5.1.1. *Let the function $F : \overline{\mathbb{R}}^d \to \overline{\mathbb{R}}$ be a homogeneous Lévy copula of order 1 and the functions $k_1^0, \dots, k_d^0$ be jump measures of univariate Feller processes of order $-1 - Y_1(x_1), \dots, -1 - Y_d(x_d)$, i.e.,*

$$k_j^0(x_j, rz_j) \;=\; r^{-1-Y_j(x_j)} k_j^0(x_j, z_j), \quad \forall r > 0,\ \forall x_j \in \mathbb{R}, z_j \in \mathbb{R}\backslash\{0\}$$

for any $j = 1, \dots, d$. Let F and $k_j^0(x_j, z_j)$, $j = 1, \dots, d$, satisfy the assumptions of Theorem 3.2.1. Due to Theorem 3.2.1 there exists a Feller process with corresponding margins. We call such a d-variate Feller process $\mathbf{Y}(x)$-stable, for $\mathbf{Y}(x) = (Y_1(x_1), \dots, Y_d(x_d))$.

For the pure jump case we need the following additional property in order to prove a simple equivalence for the sector condition. We assume that the symmetric part of the jump measure $k^{\text{sym}}(x, z) = \frac{1}{2}(k(x, z) + k(x, -z))$ admits the following estimate:

$$k^{\text{sym}}(x, z) \geq C k^{0,\text{sym}}(x, z), \quad \forall 0 < |z| < 1, \forall x \in \mathbb{R}^d, \tag{5.2}$$

where k^0 is the jump measure of a $\mathbf{Y}(x)$-stable Feller process and C denotes some positive constant. We now prove an anisotropic homogeneity property of the Feller density k^0.

Remark 5.1.2. *Condition* (5.2) *is satisfied for a time-homogeneous admissible market model in the sense of Definition 3.2.3.*

Theorem 5.1.3. *Let the copula F and the marginal densities be as in Definition 5.1.1. Then the function k^0 given by (3.7) is $\mathbf{Y}(x)$*-homogeneous *in the sense that*

$$k^0\left(x, t^{-\frac{1}{Y_1(x_1)}} z_1, \dots, t^{-\frac{1}{Y_d(x_d)}} z_n\right) \;=\; t^{1+\frac{1}{Y_1(x_1)}+\dots+\frac{1}{Y_d(x_d)}} k^0(x, z_1, \dots, z_n) \quad \forall t > 0.$$

Proof. The proof follows analogously to that of [52, Theorem 3.2], where the case $Y_i(x_i) \equiv Y_i$ was treated. □

Theorem 5.1.4. *Let $k^0(x, z_1, \dots, z_d)$ be as in the previous theorem. Then the symbol $\psi^0(x, \xi)$ of the Feller process X with characteristic triplet $(0, 0, k^0(x, z_1, \dots, z_n))$ is a real-valued anisotropic homogeneous function of type $\left(\frac{1}{Y_1(x_1)}, \dots, \frac{1}{Y_d(x_d)}\right)$ and order 1 for all $x \in \mathbb{R}^d$, i.e., it satisfies*

$$\psi^0(x, t^{\frac{1}{Y_1(x_1)}} \xi_1, \dots, t^{\frac{1}{Y_d(x_d)}} \xi_n) = t\psi^0(x, \xi_1, \dots, \xi_d) \quad \forall t > 0, \xi \in \mathbb{R}^d.$$

Proof. The proof follows analogously to that of [52, Theorem 3.3], where the case $Y_i(x_i) \equiv Y_i$ was treated, using Theorem 5.1.3. □

We need the following lemma, which is a modification of [39, Lemma 2.2].

Lemma 5.1.5. *Let $\rho_1(x,z)$, $\rho_2(x,z)$ with $\underline{\rho}_2(z) \le \rho_2(x,z) \le \overline{\rho}_2(z)$ be two anisotropic distance functions of order 1 and type $\mathbf{Y}(x) = (Y_1(x_1), \dots, Y_d(x_d))$ for all $x, z \in \mathbb{R}^d$, further let $\underline{\rho}_2(z)$ and $\overline{\rho}_2(z)$ be continuous. Furthermore, let $\Sigma := \cup_{x\in\mathbb{R}^d}\Sigma_1(x)$, where*

$$\Sigma_1(x) := \{z : \rho_1(x,z) = 1\},$$

be contained in a compact set. Then the following inequalities hold with constants C_1, $C_2 > 0$ independent of $x \in \mathbb{R}^d$ and $z \in \mathbb{R}^d$:

$$C_1\rho_1(x,z) \le \rho_2(x,z) \le C_2\rho_1(x,z).$$

Proof. Let $z \in \mathbb{R}^d\backslash\{0\}$. Then $t(x) = \frac{1}{\rho_1(x,z)}$ is well-defined. Moreover,

$$\left(t(x)^{Y_1(x)}z_1, \dots, t(x)^{Y_d(x)}z_d\right) \in \Sigma_1(x)$$

holds. As Σ is contained in a compact set and $\overline{\rho}_2(z), \underline{\rho}_2(z)$ are continuous, we obtain

$$C_1 \le \rho_2(x,z) \le C_2 \qquad \forall x \in \mathbb{R}^d,\ \forall z \in \Sigma_1.$$

Hence, there exist $C_1, C_2 \in \mathbb{R}^+$ such that for $x \in \mathbb{R}^d$, $z \in \mathbb{R}^d\backslash\{0\}$

$$C_1 \le \frac{1}{\rho_1(x,z)}\rho_2(x,z) = t(x)\rho_2(x,z) = \rho_2(x, t(x)^{Y_1(x)}z_1, \dots, t(x)^{Y_d(x)}z_d) \le C_2.$$

□

Theorem 5.1.6. *Let X be an admissible time-homogeneous market model in the sense of Definition 3.2.3 taking values in $\mathbb{R}^d$ with characteristic triplet $(b(x), Q(x), k(x,z)\mathrm{d}z)$ with $k(x,z)$ being the density of the jump-measure constructed parametrically as in Theorem 3.2.1. Then, there exists a constant $C > 0$ such that for all $x \in \mathbb{R}^d$ and $\|\xi\|_\infty$ sufficiently large*

$$\Re\psi(x,\xi) \ge C\sum_{j=1}^{d} |\xi|^{Y_j(x_j)}, \tag{5.3}$$

where $Y_j(x_j) = 2$ in the case $Q_0 \ge Q > 0$.

Proof. The proof mainly follows the arguments of [113, Proposition 2.4.3]. First consider $Q = 0$. Due to Theorem 5.1.4 one obtains that $\Re\psi^0(x,\xi)$ is an anisotropic distance function of type $(1/Y_1(x_1), \dots, 1/Y_d(x_d))$ and order 1 for all $x \in \mathbb{R}^d$. We obtain from Lemma 5.1.5

$$\psi^0(x,\xi) \ge C_1\sum_{i=1}^{d} |\xi_i|^{Y_i(x_i)}, \quad \forall\xi \in \mathbb{R}^d,$$

where we set $\rho_1(x,\xi) = \sum_{i=1}^d |\xi_i|^{Y_i(x_i)}$ and $\rho_2(x,\xi) = \psi^0(x,\xi)$. Hence,

$$\begin{aligned}
\Re\psi(x,\xi) &= \int_{\mathbb{R}^d} (1-\cos(\xi\cdot z))\, k^{\mathrm{sym}}(x,z)\mathrm{d}y \\
&\geq C_2 \int_{B_1(0)} (1-\cos(\xi\cdot z)) k^{0,\mathrm{sym}}(x,z)\mathrm{d}y \\
&\geq C_2 C_1 \sum_{i=1}^d |\xi_i|^{Y_i(x_i)} - C_3.
\end{aligned}$$

Therefore, the sector condition (5.1) follows from (5.2) for appropriate symbols. The case $Q \geq Q_0 > 0$ is trivial. $\square$

5.2 European options in market models without drift dominance

We consider a European option with maturity $T < \infty$ and payoff $g(S(T))$, where $S^i(t) = S^i(0)e^{rt+X^i(t)}$ and where X is a semimartingale. We treat several special cases of this general setup. By the general theory of asset pricing (as, e.g., in [42]), an arbitrage free value $V(t,s)$ of this option is given by

$$V(t,s) = \mathbb{E}\left(e^{-r(T-t)} g(S(T)) | S(t) = s\right),$$

where the expectation is taken under the measure $\mathbb{Q}$ which is equivalent to the real world measure and under which $S(T)$ is a discounted σ-martingale, cf.[42] and $r > 0$. If X is an admissible time-homogeneous market model, we can derive a PDO and PIDE representation and prove well-posedness of the weak formulation of the problem on a bounded domain. In the following we focus on time-homogeneous admissible market models and return to the time-inhomogeneous setup later in Chapter 9. Due to no arbitrage considerations we require the considered processes to be discounted martingales under a pricing measure $\mathbb{Q}$. This requirement can be expressed in terms of the characteristic triplet:

Lemma 5.2.1. *Let X be an admissible time-homogeneous market model with characteristic triplet $(b(x), Q(x), \nu(x,dz))$ and semigroup $(T_t)_{t\geq 0}$ further let $T_t(e^{x_j}) < \infty$ hold for $t \geq 0$, $j = 1,\dots,d$. Then e^{X_j} is a $\mathbb{Q}$-martingale with respect to the canonical filtration of X if and only if*

$$\frac{Q_{jj}(x)^2}{2} + b_j(x) + \int_{0\neq z_j\in\mathbb{R}} (e^{z_j} - 1 - z_j)\nu_j(x,\,\mathrm{d}z_j) = 0 \quad \forall x \in \mathbb{R}. \tag{5.4}$$

Proof. This is a direct consequence of [48, Section 3]. $\square$

Remark 5.2.2. *Note that without the assumption of finiteness of exponential moments of the processes X_j, the processes e^{X_j}, $j = 1, \dots, d$ would generally only be local martingales. For Lévy processes, exponential decay of the jump measure implies the existence of exponential moments, cf. [103, Theorem 25.3]. This is not obvious for general Feller processes. Recently, Knopova and Schilling have proved in [78] the finiteness of exponential moments for a certain class of Feller processes assuming exponential decay of the density of the jump measure.*

We are now able to derive a PDO and PIDE representation for option prices. Let the stochastic process X be an admissible time-homogeneous market model with generator $\mathcal{A}$ and let be g be sufficiently smooth, $\mathcal{V} = H^1(\mathbb{R}^d)$ for diffusion market models, $\mathcal{V} = H^{\mathbf{m}}(\mathbb{R}^d)$, $\mathbf{m} = [Y_1/2, \dots, Y_d/2] \in (0,1)^d$ for general space and time-homogeneous models and $\mathcal{V} = H^{\mathbf{m}(x)}(\mathbb{R}^d)$ as in Definition 2.2.2, with $\mathbf{m}(x) = [Y_1(x_1)/2, \dots, Y_d(x_d)/2]$, for time-homogeneous admissible market models considered here. Then we obtain formally due to semigroup theory for $u(t,x) = T_t(g) = \mathbb{E}[g(X_t)|X_0 = x]$ by differentiation in t the following PIDE

$$\begin{aligned} \partial_t u + \mathcal{A}u - ru &= 0 \quad \text{in } (0,T) \times \mathbb{R}^d, && (5.5)\\ u(T) &= g \quad \text{in } \mathbb{R}^d. && (5.6) \end{aligned}$$

Testing with a function $v \in \mathcal{V}$ and transforming to time-to-maturity, we end up with the following parabolic evolution problem: Find $u \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$ s.t. for all $v \in \mathcal{V}$ and a.e. $t \in [0,T]$ holds

$$(\partial_t u, v)_{\mathcal{V}^*,\mathcal{V}} + a(u,v) = 0, \qquad u(0) = g, \tag{5.7}$$

where *the bilinear form $a(\varphi,\phi) = (-\mathcal{A}\varphi,\phi)_{\mathcal{V}^*,\mathcal{V}} + r(\varphi,\phi)_{L^2(\mathbb{R}^d)}$ is closely related to the Dirichlet form of the stochastic process X*. Although in option pricing, only the homogeneous parabolic problem (5.7) arises, the inhomogeneous equation (5.8) is useful in many applications. We mention only the computation of the time-value of an option, or the computation of quadratic hedging strategies and the corresponding hedging error. Thus, we consider the non-homogeneuos analogon of the above equation. The general problem reads: Find $u \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$ s.t.

$$(\partial_t u, v)_{\mathcal{V}^*,\mathcal{V}} + a(u,v) = (f,v)_{\mathcal{V}^*,\mathcal{V}} \quad \text{in } (0,T),\ \forall v \in \mathcal{V} \qquad u(0) = g \tag{5.8}$$

for some $f \in L^2((0,T);\mathcal{V}^*)$. Now we consider the localization of the unbounded problem to a bounded domain D. For any function u with support in a bounded domain $D \subset \mathbb{R}^d$ we denote by $\widetilde{u}$ the zero extensions of u to $D^c = \mathbb{R}^d \backslash \overline{D}$ and define $\mathcal{A}_D(u) = \mathcal{A}(\widetilde{u})$. The variational formulation of the pricing equation on a bounded domain $D \subset \mathbb{R}^d$ reads: Find $u \in L^2((0,T);\mathcal{V}_D) \cap H^1((0,T);(\mathcal{V}_D)^*)$ s.t. for all $v \in \mathcal{V}_D$ and a.e. $t \in [0,T]$ holds:

$$\begin{aligned} (\partial_t u, v)_{\mathcal{V}_D^*,\mathcal{V}_D} + a_D(u,v) &= (f,v)_{\mathcal{V}_D^*,\mathcal{V}_D} && (5.9)\\ u(0) &= g|_D, && (5.10) \end{aligned}$$

where $a_D(u,v) := a(\widetilde{u},\widetilde{v})$ and $\mathcal{V}_D =: \{v \in L^2(D) : \widetilde{v} \in \mathcal{V}\}$. Under condition (4.4) pointwise convergence of the solution of the localized problem to the solution of the original problem can be shown for certain admissible market models using [103, Theorem 25.18] and the semiheavy tail property. We refer to Section 4.2 for details. A comparable result for general Feller processes does not appear to be available yet.

Remark 5.2.3. *Formulation* (5.9)-(5.10) *naturally arises for payoffs with finite support such as digital or (double) barrier options. The truncation to a bounded domain can thus be interpreted economically as the approximation of a standard derivative contract by a corresponding barrier option on the same market model. Note also that the variational framework* (5.9)-(5.10) *naturally allows for more general initial conditions, in particular $g \in \mathcal{H} = L^2(\mathbb{R}^d)$. Therefore, discontinuous g are admissible in the variational framework* (5.9)-(5.10). *This is essential for the pricing of exotic contracts such as digital options, for example.*

Existence and uniqueness of weak solutions of (5.9)-(5.10) follows from continuity of the bilinear form $a_D(\cdot,\cdot)$ and a Gårding inequality which follows from Theorem 5.2.4.

Theorem 5.2.4. *Let the generator $\mathcal{A}(x,D) \in \Psi^{\mathbf{2m}(x)}_{\rho,\delta}$ be a pseudo-differential operator of variable order $\mathbf{2m}(x)$, $0 < m_i(x_i) < 1$, $i = 1,\dots,d$ with $\mathbf{m}(x) = (m_1(x_1),\dots,m_d(x_d))$ and symbol $\psi(x,\xi) \in S^{\mathbf{2m}(x)}_{\rho,\delta}$ for some $0 < \delta < \rho \le 1$ for which there exists $C > 0$ with*

$$\Re\,\psi(x,\xi) + 1 \ge C\langle\xi\rangle^{\mathbf{2m}(x)} \quad \forall x,\xi \in \mathbb{R}^d. \tag{5.11}$$

Then $\mathcal{A}(x,D) \in \Psi^{\mathbf{2m}(x)}_{\rho,\delta}$ satisfies a Gårding inequality in the variable order space $\widetilde{H}^{\mathbf{m}(x)}(D)$: There are constants $C_1 > 0$ and $C_2 \ge 0$ such that

$$\forall u \in \widetilde{H}^{\mathbf{m}(x)}(D): \quad \Re a(u,u) \ge C_1\|u\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} - C_2\|u\|^2_{L^2(D)}, \tag{5.12}$$

and

$$\exists\lambda > 0 \quad \text{such that } \mathcal{A}(x,D) + \lambda I : \widetilde{H}^{\mathbf{m}(x)}(D) \to H^{-\mathbf{m}(x)}(D) \tag{5.13}$$

is boundedly invertible, for $a(u,v) = (\mathcal{A}u,v)_{H^{-\mathbf{m}(x)}(D),\widetilde{H}^{\mathbf{m}(x)}(D)}$, $u,v \in \widetilde{H}^{\mathbf{m}(x)}(D)$.

Proof. The proof follows along the lines of the proof of [97, Theorem 5], where the case $d = 1$ was treated. It relies on results from pseudodifferential operator theory which are also available for $d > 1$. □

Note that in the case of an admissible time-homogeneous market model $\Re a_D(u,u) = a_D(u,u)$ holds, for $u \in \mathcal{V}_D$ and $a_D(\cdot,\cdot)$ as in (5.9).

Theorem 5.2.5. *The problem (5.9)-*(5.10) *for an admissible market model X with symbol $\psi(x,\xi)$ with initial condition $g \in \mathcal{H} = L^2(\mathbb{R}^d)$ and $\underline{Y} \ge 1$, $Q = 0$ or $Q \ge Q_0 > 0$ has a unique solution.*

Proof. We obtain from Lemma 3.2.6

$$\psi(x,\xi) \in S_{1,\delta}^{\mathbf{Y}(x)} \quad \text{for} \quad \underline{Y} \geq 1, Q = 0,$$
$$\psi(x,\xi) \in S_{1,\delta}^{2} \quad \text{for} \quad Q \geq Q_0 > 0.$$

Theorem 5.1.6 implies

$$\Re\psi(x,\xi) + 1 \geq C\langle\xi\rangle^{\mathbf{Y}(x)} \quad \text{for } \underline{Y} \geq 1, Q = 0, \tag{5.14}$$
$$\Re\psi(x,\xi) + 1 \geq C\langle\xi\rangle^{2} \quad \text{for } Q \geq Q_0 > 0, \tag{5.15}$$

for all $x, \xi \in \mathbb{R}^d$. The application of Theorems 5.2.4 and 2.2.4 implies the claimed result. □

Remark 5.2.6. *We obtain for admissible market models X drift dominated equations for $\underline{Y} < 1$, therefore Theorem 5.1.6 holds with $\mathbf{Y}(x)$ as in Definition 3.2.3, but we obtain for the symbol $\psi(x,\xi)$ of the infinitesimal generator of X due to Lemma 3.2.6 $\psi(x,\xi) \in S_{1,\delta}^{\widetilde{\mathbf{m}}(x)}$, with $\widetilde{m}_i(x_i) = \max\left(Y_i(x_i), 1\right)$, $i = 1, \dots, d$. Therefore Theorem 5.2.4 is not applicable and we have to remove the drift for standard algorithms to be feasible. The removal of the drift is straightforward in the Lévy case as the drift coefficients in the equation are constant, cf. [95, Corollary 4.3], but more involved in the Feller case, cf. [60].*

In the classical setting of advection-diffusion equations, the drift dominance in a given discretization is measured by the Péclet number.

5.3 European options in market models with drift dominance

If the pricing equation is drift dominated in the sense that the requirements of Theorem 5.2.5 are not satisfied, different techniques have to be employed to prove well-posedness of (5.9)-(5.10). This is due to the dominance of the asymmetric part of the generator. Generally, a change of variable can be used to remove the drift from the pricing equation, but finding an appropriate change of a variable is non trivial. The following result can be obtained.

Theorem 5.3.1. *For $h \in C^{1,2}(I \times \mathbb{R})$, $I = [0,T]$, $T > 0$, with $\partial_x h(t,x) \neq 0$, consider the change of variable $v(t,x) := u(t, h(t,x))$, where $u(t,x)$ is the solution of the following PDE*

$$\partial_t u - \Sigma(x)\partial_{xx} u + b(x)\partial_x u + c(x)u \quad = \quad 0 \quad \textit{in} \quad I \times \mathbb{R},$$

for $\Sigma(x) \geq 0$ sufficiently smooth as well as smooth and bounded functions $b(x)$, $c(x)$. Let h solve the (nonlinear) PDE

$$\partial_t h - \Sigma(h(t,x))\frac{\partial_{xx} h}{\partial_x^2 h} - b(h(t,x)) \quad = \quad 0. \tag{5.16}$$

Then v satisfies the PDE

$$\partial_t v - \frac{\Sigma(h(t,x))}{\partial_x^2 h}\partial_{xx} v + c(h(t,x))v \quad = \quad 0 \quad \text{in} \quad I \times \mathbb{R}.$$

Proof. The claim follows by an application of the chain rule. See [60, Lemma 3.17]. □

Solving the PDE (5.16) is non trivial in general.

Example 5.3.2. *Let X be a univariate Lévy market model, then $b(x) \equiv b$, $\Sigma(x) \equiv Q$, $b \in \mathbb{R}$, $Q \geq 0$ and a particular solution of (5.16) is given by $f(t,x) = bt + x$.*

If the drift cannot be removed as in Theorem 5.3.1, then the results from Section 5.2 are not applicable. We therefore describe a different approach in the following. One typical example of a drift dominated problem arises in option pricing under subordinators, see Example 3.2.10.

Theorem 5.3.3. *We consider the bilinear form $a_{\mathrm{J}} = (\mathcal{A}_J u, v)$, with $\mathcal{A}_J$ given as*

$$\mathcal{A}_{\mathrm{J}}\varphi(x) \quad = \quad -\int_{\mathbb{R}_+} (\varphi(x+z) - \varphi(x))\, \nu(dz), \tag{5.17}$$

for sufficiently smooth $\varphi(x)$ and

$$\nu(dz) = k(z)dz = C\frac{e^{-\beta z}}{z^{1+Y}}dz, \quad C, \beta > 0,\ Y \in [0,1).$$

Then the domain $D(a_J(\cdot,\cdot))$ of $a_J(\cdot,\cdot)$ is given by $H^{Y/2}(\mathbb{R})$, i.e., $a_J(\cdot,\cdot)$ is continuous and satisfies a Gårding inequality on $H^{Y/2}(\mathbb{R})$.

Proof. The proof follows from [60, Proposition 3.13 and Corollary 3.15] and relies on the use of properties of anisotropic distance functions. □

The generator of a subordinator as in Example 3.2.10 reads

$$\mathcal{A} = \mathcal{A}_{\mathrm{J}} + \mathcal{A}_{\mathrm{Tr}},$$

with $\mathcal{A}_{\mathrm{J}}$ as in (5.17) and

$$\mathcal{A}_{\mathrm{Tr}}\varphi = \lambda(x)\varphi'(x),$$

for $\lambda(x) \in W^{1,\infty}(\mathbb{R})$.

Remark 5.3.4. *Due to the fact that the antisymmetric part of the bilinear form* $a_{\mathrm{Tr}}(u,v) = (\mathcal{A}_{\mathrm{Tr}}u, v)$ *is the leading term of the bilinear form* $a(\cdot,\cdot) = a_{\mathrm{Tr}}(\cdot,\cdot) + a_{\mathrm{J}}(\cdot,\cdot)$*, as* $\underline{Y} < 1$*, standard well-posedness results are not applicable. We therefore prove the inf-sup conditions. Note that in the special case of a constant drift coefficient, i.e.,* $\lambda(x) \equiv \lambda \in \mathbb{R}$*, we could apply Theorem 5.3.1 and find the function* $h(t,x)$ *explicitly. Numerically, this is not feasible, as large values for* λ *imply strong restrictions on the space discretization of the transformed system. We therefore present a well-posedness result for the original system.*

The following problem on a bounded Lipschitz domain D is analyzed. Let $a(u,v)$ be given by

$$
\begin{aligned}
a(u,v) &= a_{\mathrm{Tr}}(u,v) + a_{\mathrm{J}}(u,v) \text{ for } u \in \mathcal{V},\ v \in L^2(D), && (5.18)\\
a_{\mathrm{J}}(u,v) &= (\mathcal{A}_J \widetilde{u}, v)_{L^2(D)}, \quad a_{\mathrm{Tr}}(u,v) = (ru + b(x)\cdot\nabla u, v)_{L^2(D)}, && (5.19)\\
\mathcal{A}u &= ru + b(x)\cdot\nabla u + \mathcal{A}_J u, && (5.20)
\end{aligned}
$$

where $\mathcal{A}_J : \widetilde{H}^{2\mathbf{m}(x)}(D) \to L^2(D)$ is the generator of an admissible time-homogeneous market model with characteristic triplet $(0, 0, \nu(x, dz))$ such that

$$
0 \leq a_{\mathrm{J}}(v,v), \quad v \in \mathcal{V} \tag{5.21}
$$

and $\mathcal{V} = \left\{w \in L^2(D) | b\cdot\nabla w + \mathcal{A}_{\mathrm{J}} w \in L^2(D), w|_{\Gamma_-} = 0\right\}$. We consider the norm $\|\cdot\|_{\mathcal{V}}$ on $\mathcal{V}$ given as

$$
\|u\|_{\mathcal{V}}^2 = \|b\cdot\nabla u + \mathcal{A}_{\mathrm{J}} u\|_{L^2(D)}^2 + \|u\|_{L^2(D)}^2 .
$$

The order $2\mathbf{m}(x)$ of the operator $\mathcal{A}_{\mathrm{J}}$ is given by $\mathbf{m} = [Y_1(x)/2, \ldots, Y_d(x)/2]$, as in Definition 3.2.3, for $Q > 0$ the order of the operator $\mathcal{A}_{\mathrm{BS}} + \mathcal{A}_{\mathrm{J}}$ is 2. We make the following standard assumption on the drift:

$$
r - \frac{1}{2}\nabla\cdot b(x) \geq r_{\min} > 0 \tag{5.22}
$$

and assume that r is constant on D. The consideration of a uniformly bounded r is also possible. The Property (5.21) follows from the sector condition, cf. Section 5.1, after a possible rescaling of the time variable.

Lemma 5.3.5. *Let* $a(u,v) : \mathcal{V} \times L^2(D) \to \mathbb{R}$ *be as in* (5.18)*, then*

$$
\inf_{v\in L^2(D)} \sup_{u\in\mathcal{V}} \frac{a(u,v)}{\|u\|_{\mathcal{V}} \|v\|_{L^2(D)}} \geq C,
$$

for a positive constant C*.*

Proof. For any $v \in \mathcal{V}$

$$
\sup_{\phi\in L^2(D)} \frac{a(v,\phi)}{\|\phi\|_{L^2(D)}} \overset{(\phi=v)}{\geq} r_{\min} \|v\|_{L^2(D)} . \tag{5.23}
$$

Therefore

$$\begin{aligned}\sup_{\phi\in L^2(D)}\frac{a(v,\phi)}{\|\phi\|_{L^2(D)}} &\geq \sup_{\phi\in L^2(D)}\frac{\int_D(b\cdot\nabla v+\mathcal{A}_{\mathrm{J}}v)\phi}{\|\phi\|_{L^2(D)}} - r\,\|v\|_{L^2(D)}\\ &\geq \sup_{\phi\in L^2(D)}\frac{\int_D(b\cdot\nabla v+\mathcal{A}_{\mathrm{J}}v)\phi}{\|\phi\|_{L^2(D)}} - \frac{r}{r_{\min}}\sup_{\phi\in L^2(D)}\frac{a(v,\phi)}{\|\phi\|_{L^2(D)}}.\end{aligned}$$

This implies

$$\left(1+\frac{r}{r_{\min}}\right)\sup_{\phi\in L^2(D)}\frac{a(v,\phi)}{\|\phi\|_{L^2(D)}} \quad\geq\quad \|b\cdot\nabla v+\mathcal{A}_{\mathrm{J}}v\|_{L^2(D)}$$

and we conclude using (5.23)

$$\sup_{\phi\in L^2(D)}\frac{a(v,\phi)}{\|v\|_{\mathcal{V}}\,\|\phi\|_{L^2(D)}} \geq \frac{1}{\sqrt{\frac{1}{r_{\min}^2}+\left(1+\frac{r}{r_{\min}}\right)^2}}.$$

□

Lemma 5.3.6. *The bilinear form $a(\cdot,\cdot)$ as in Lemma 5.3.5 is continuous on $\mathcal{V}\times L^2(D)$ and surjective, i.e.,*

$$\forall 0\neq\phi\in L^2(D),\ \sup_{v\in\mathcal{V}}|a(v,\phi)|>0.$$

Proof. The continuity follows from the properties of $\mathcal{A}_J$ and the triangle inequality. The surjectivity follows as in [51, Lemma 6.2.9]. □

We are now able to prove the well-posedness of the pricing equation. We consider the following problem: for $f\in C^1([0,T],L^2(D))$ and $g\in\mathcal{V}$ find $u\in C^1([0,T],L^2(D))\cap C^0([0,T],\mathcal{V})$ such that

$$\partial_t u+\mathcal{A}u=f,\quad u(0)=g, \tag{5.24}$$

where $\mathcal{A}$ is given by (5.20).

Theorem 5.3.7. *The problem* (5.24) *has a unique solution.*

Proof. The proof is a consequence of the Hille-Yosida theorem [51, Theorem 7.3.3]. It remains to show the monotonicity and maximality of the operator $\mathcal{A}:\mathcal{V}\to L^2(D)$. We obtain

$$(\mathcal{A}\widetilde{v},\widetilde{v})\geq 0,\quad \forall v\in\mathcal{V}$$

from (5.21). The maximality follows from Lemmas 5.3.5 and 5.3.6, i.e., for all $f\in L^2(D)$ there exists $v\in\mathcal{V}$ such that

$$v+\mathcal{A}v=f. \tag{5.25}$$

□

Remark 5.3.8. *Note that the assumptions on the regularity of the initial condition g and the right hand side f in* (5.24) *are restrictive. Theorem 5.3.7 implies the well-posedness of transport dominated pricing equations. It covers the cases excluded in Theorem 5.2.5. Note that the bilinear forms arising in the transport dominated setting are non-symmetric.*

5.4 European options in regularized market models

After the small jump regularization of an infinite activity time-homogeneous admissible market model, cf. Section 4.1, the corresponding pricing equation for the value of a European option $u(t,x)$ reads:

$$\begin{aligned} \partial_t u - \mathcal{A}u + ru &= 0 \text{ in } I \times \mathbb{R}^d, && (5.26)\\ u(0) &= g \text{ in } \mathbb{R}^d, && (5.27) \end{aligned}$$

where

$$\begin{aligned} \mathcal{A} &= \mathcal{A}_{\mathrm{Tr}} + \mathcal{A}_{\mathrm{BS}}(\varepsilon) + \mathcal{A}_J(\varepsilon) \\ \mathcal{A}_{\mathrm{Tr}}\varphi(x) &= b(x)\cdot\nabla\varphi(x) \\ \mathcal{A}_{\mathrm{BS}}(\varepsilon)\varphi(x) &= \frac{1}{2}\mathrm{tr}[Q(\varepsilon,x)D^2\varphi(x)] \\ \mathcal{A}_J(\varepsilon)\varphi(x) &= \int_{\mathbb{R}^d} \left(\varphi(x+z)-\varphi(x)-z\cdot\nabla\varphi(x)\right) k_\varepsilon(x,z)\,dz, && (5.28) \end{aligned}$$

with $Q(\varepsilon,x)$ and $k_\varepsilon(x,z)$ as in Section 4.1. The weak formulation of (5.26) is given as: find $u \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$ s.t. for all $v \in \mathcal{V}$ and a.e. $t \in I$ holds

$$(\partial_t u, v)_{\mathcal{V}^*,\mathcal{V}} + a(u,v) = 0, \qquad u(0) = g, \tag{5.29}$$

where the bilinear form $a(u,v)$ reads

$$a(u,v) = (-\mathcal{A}u, v)_{\mathcal{V}^*,\mathcal{V}} + r(u,v)_{L^2(\mathbb{R}^d)}. \tag{5.30}$$

The domain of the bilinear form is $\mathcal{V} = H^1(\mathbb{R}^d)$, continuity and a Gårding inequality can be proved with constants that depend explicitly on ε.

Theorem 5.4.1. *Let $a(\cdot,\cdot)$ be given by* (5.30)*, then*

$$\begin{aligned} |a(u,v)| &\le C_1(\varepsilon)\,\|u\|_{H^1(\mathbb{R}^d)}\,\|v\|_{H^1(\mathbb{R}^d)}\,, && (5.31)\\ a(u,v) &\ge C_2(\varepsilon)\,\|u\|^2_{H^1(\mathbb{R}^d)} - C_3(\varepsilon)\,\|u\|^2_{L^2(\mathbb{R}^d)}\,, && (5.32) \end{aligned}$$

where $C_1(\varepsilon) \le \widetilde{C}_1\varepsilon^{-\overline{Y}-d}$, $C_2(\varepsilon) \ge \widetilde{C}_2\varepsilon^{2-\underline{Y}}$ and $C_3(\varepsilon) \ge \widetilde{C}_3\varepsilon^{2-\underline{Y}}$, for positive constants $\widetilde{C}_1$, $\widetilde{C}_2$, $\widetilde{C}_3$ independent of ε, sufficiently small.

Proof. The proof is a direct consequence of the following estimates on $k_\varepsilon(x,z)$ and $Q(\varepsilon,x)$

$$k_\varepsilon(x,z) \le \widetilde{C}_1 \varepsilon^{-d-\overline{Y}}, \quad Q_{i,j}(\varepsilon,x) \ge \widetilde{C}_2 \varepsilon^{2-\underline{Y}}.$$

□

Theorem 5.4.1 implies the well-posedness of the pricing problem.

Lemma 5.4.2. *The problem* (5.29) *is well-posed. There exists for every* $g \in L^2(\mathbb{R}^d)$ *a unique function* $u \in L^2((0,T);\mathcal{V}) \cap H^1((0,T);\mathcal{V}^*)$, *such that* (5.29) *holds.*

We have the following corollary for the localized pricing problem.

Corollary 5.4.3. *Let the localized problem be given as:*
find $u \in L^2((0,T);\mathcal{V}_D) \cap H^1((0,T);\mathcal{V}_D^*)$ *s.t. for all* $v \in \mathcal{V}_D$ *and a.e.* $t \in I$ *holds*

$$(\partial_t u, v)_{\mathcal{V}_D^*,\mathcal{V}_D} + a_D(u,v) = (f,v), \qquad u(0) = g|_D, \tag{5.33}$$

where the bilinear form $a_D(u,v) = (-\mathcal{A}\widetilde{u}, \widetilde{v})_{\mathcal{V}_D^*,\mathcal{V}} + r(u,v)_{L^2(D)}$ *and* $f \in L^2(I;\mathcal{V}_D^*)$. *The domain of the bilinear form is* $\mathcal{V}_D = H_0^1(D)$. *Then* (5.33) *admits a unique solution.*

Remark 5.4.4. *Note that the constants in the error estimates explicitly depend on* C_1 *and* C_2 *in Theorem 5.4.1 and therefore on* ε. *Hence, a judicious choice of the mesh width in terms of the small jump truncation parameter is essential for a rigorous error analysis of the FEM discretization.*

6 Wavelets and triangulations

For the numerical solution, we discretize the parabolic equation (5.9)-(5.10) in $(0,T)\times D$ in the spatial variable with spline wavelet bases for $\mathcal{V} = \widetilde{H}^{\mathbf{m}(x)}(D)$ and in the time parameter by the θ-scheme or the more sophisticated discontinuous Galerkin timestepping which allows to exploit the time-analycity of the processes' semigroups. To present the spatial discretizations, we briefly recapitulate basic definitions and results on wavelets from, e.g., [29] and the references there. For specific spline wavelet constructions on a bounded interval I, we refer to e.g. [43], [92] and [112].
Since for all infinitesimal generators arising in connection with Markov processes the Sobolev order $2m(x)$ of the generator satisfies $0 \le m(x) \le 1$, the full machinery of multiresolution analyses in Sobolev spaces of arbitrary order is not required; we confine ourselves therefore to *continuous, piecewise polynomial multiresolution systems in* $\mathbb{R}^1$. For wavelet discretizations of Kolmogorov equations for multivariate models, we shall employ tensor products of these univariate, piecewise polynomials multiresolution systems. For wavelet constructions on general domains we refer to [90]. Note that in this case the classical construction as described in the following by means of shifts and dilates from one "mother" wavelet is not applicable. General domains may arise in the pricing of exotic basket options.
Our use of compactly supported, piecewise polynomial multiresolution systems (rather than the more commonly employed B-spline Finite Element spaces) for the Galerkin discretization of Kolmogorov equations is motivated by the following key properties of these spline wavelet systems: a) the approximation properties of the multiresolution sytems equal those of the B-spline systems, b) the spline wavelet systems form *Riesz bases* of the domains of the infinitesimal generators of the Markov processes, thereby allowing for simple and efficient *preconditioning of the matrices arising in wavelet representations of the processes' Dirichlet forms*, c) the spline wavelet systems can be designed to have a large number of *vanishing moments*, thereby allowing for a *compression* of the wavelet matrix for the jump measure.

6.1 Triangulation

In the following we briefly summarize the requirements that have to be imposed on the triangulation. Let $\mathcal{T}_h$ be a partition of D into disjoint open element domains K such that $\overline{D} = \bigcup_{K\in\mathcal{T}_h} \overline{K}$ and each $K \in \mathcal{T}_h$ is an affine image of a fixed master element $\widehat{K}$, i.e.

$K = F(\widehat{K})$, where $\widehat{K}$ is the unit simplex and $F(\hat{x}) = A_K\hat{x} + b_K$, $A_K \in \mathbb{R}^{d\times d}$, $b_K \in \mathbb{R}^d$. We assume the simplicial family $\{\mathcal{T}_h\}_{h\in(0,1]}$ to be shape regular and quasi-uniform:

$$\exists\ C_1,\, C_2 > 0 \,\text{independent of } h \text{ such that for all } h:\ \sup_{K\in\mathcal{T}_h} \frac{h_K}{r_K} < C_1 < +\infty \tag{6.1}$$

$$\text{and} \sup_{K\in\mathcal{T}_h} h_K < C_2 h_{K'}\ \forall\ K' \in \mathcal{T}_h. \tag{6.2}$$

Here h_K and r_K denote the diameter of the element K and the maximum radius of a ball contained in K, $K \in \mathcal{T}_h$, respectively. Besides, we assume that the family of triangulations is regular, in the sense that the intersection $\overline{K} \cap \overline{K'}$ of two elements $K, K' \in \mathcal{T}_h$ for $h \in (0,1]$, is either empty, a single vertex or an entire side for $d = 2$ and analogously for higher dimensions, i.e., there are no hanging nodes. Moreover, we set $h = \max_{K\in\mathcal{T}_h} h_K$.
We denote by $S^{p,0}(D,\mathcal{T}_h)$ the space of discontinuous piecewise polynomial functions, i.e., $v_h \in S^{p,0}(D,\mathcal{T}_h)$ if and only if $v_h|_K \in \mathbb{P}^p(K)$, $\forall K \in \mathcal{T}_h$, where $\mathbb{P}^p(K)$ is the space of polynomials of total degree p in K. Space of continuous piecewise polynomials is denoted by $S^{p,1}(D,\mathcal{T}_h) := S^{p,0}(D,\mathcal{T}_h) \cap H^1(D)$. If the choice of $\mathcal{T}_h$ and D is clear from the context we omit them and write $S^{p,1}$ and $S^{p,0}$. Finally, we assign to the subdivision $\mathcal{T}_h$ the broken Sobolev space of composite order s, where $s_K \in \mathbb{N}_0$ are non-negative integers,

$$H^s(D,\mathcal{T}_h) = \big\{u \in L^2(D)\ :\ u|_K \in H^{s_K}(K)\ \forall K \in \mathcal{T}_h\big\} \tag{6.3}$$

equipped with the norm

$$\|u\|_{H^s(D,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} \|u\|^2_{H^{s_K}(K)}\right)^{1/2}.$$

Furthermore we use the following notations: $\Gamma_h = \bigcup_{K\in\mathcal{T}_h} \partial K$, $\mathcal{T}_h$ being the considered triangulation and $\Gamma_h^0 = \bigcup_{K\in\mathcal{T}_h} \partial K \cap \partial D$. The average and jump operators are defined as follows: if $e \in \Gamma$ is an edge shared by two elements K_1 and K_2 of $\mathcal{T}_h$ and n is the unit vector normal oriented from K_1 to K_2, then

$$\{v\} = \frac{v|_{K_1} + v|_{K_2}}{2} \quad [v] = v|_{K_1} - v|_{K_2},$$

otherwise (i.e., $e \in \Gamma \cap \partial D$) we set

$$\{v\} = [v] = v.$$

6.2 Spline wavelets on an interval

Our Galerkin discretizations of Kolmogorov equations for Feller processes are based on biorthogonal wavelet bases on a bounded interval $I \subset \mathbb{R}$. We write $x \lesssim y$ in the following to express that x is bounded by a constant multiple of y, uniformly with respect to all parameters on which x and y may depend and $x \sim y$ if $x \lesssim y$ and $y \lesssim x$ holds.

We recapitulate the basic definitions from, e.g., [29, 112] to which we also refer for further references and additional details, such as the construction of higher order wavelets.

Our wavelet systems are two-parameter systems $\{\psi_{l,k}\}_{l=0,\dots,\infty,k\in\nabla_l}$ of compactly supported functions $\psi_{l,k}$. Here the first index, l, denotes "level" of refinement resp. resolution: wavelet functions $\psi_{l,k}$ with large values of the level index are well-localized in the sense that $\mathrm{diam}(\mathrm{supp}\psi_{l,k}) = O(2^{-l})$. The second index, $k \in \nabla_l$, measures the *localization* of wavelet $\psi_{l,k}$ within the interval I at scale l and ranges in the index set ∇_l. In order to achieve maximal flexibility in the construction of wavelet systems (which can be used to satisfy other requirements, such as minimizing their support size or to minimize the size of constants in norm equivalences), we consider wavelet systems which are *biorthogonal* in $L^2(I)$, consisting of a primal wavelet system $\{\psi_{l,k}\}_{l=0,\dots,\infty,k\in\nabla_l}$ which is a Riesz basis of $L^2(I)$ (and which enter explicitly in the Galerkin discretizations of the Markov processes) and a corresponding dual wavelet system $\{\widetilde{\psi}_{l,k}\}_{l=0,\dots,\infty,k\in\nabla_l}$ (which are never used explicitly in our algorithms). Notice that construction of *fully* $L^2(I)$ *orthonormal wavelet systems* is feasible, but results in function systems which are either nonpolynomial or have larger supports or fewer vanishing moments.

The primal wavelet bases $\psi_{l,k}$ span finite dimensional spaces

$$\mathcal{W}^l := \mathrm{span}\ \{\psi_{l,k} : k \in \nabla_l\}\ , \quad \mathcal{V}^L := \bigoplus_{l=0}^{L-1} \mathcal{W}^l \quad l = 0, 1 \dots,$$

and the dual spaces are defined analogously in terms of the dual wavelets $\widetilde{\psi}_{l,k}$ by

$$\widetilde{\mathcal{W}}^l := \mathrm{span}\ \{\widetilde{\psi}_{l,k} : k \in \nabla_l\}\ , \quad \widetilde{\mathcal{V}}^L := \bigoplus_{l=0}^{L-1} \widetilde{\mathcal{W}}^l \quad l = 0, 1 \dots, .$$

In the sequel we require the following properties of the wavelet functions to be used on our Galerkin discretization schemes, we assume without loss of generality $I = (0, 1)$.

(i) Biorthogonality: the basis functions $\psi_{l,k}$, $\widetilde{\psi}_{l,k}$ satisfy

$$\langle \psi_{l,k}, \widetilde{\psi}_{l',k'} \rangle = \delta_{l,l'}\delta_{k,k'}. \tag{6.4}$$

(ii) Local support: the diameter of the support is proportional to the meshsize 2^{-l},

$$\mathrm{diam\ supp}\ \psi_{l,k} \sim 2^{-l} \quad , \quad \mathrm{diam\ supp}\ \widetilde{\psi}_{l,k} \sim 2^{-l}. \tag{6.5}$$

(iii) Conformity: the basis functions should be sufficiently regular, i.e.

$$\mathcal{W}^l \subset \widetilde{H}^1(I)\ ,\ \widetilde{\mathcal{W}}^l \subset H^\delta(I) \ \text{ for some } \delta > 0\ , \quad l \geq 0. \tag{6.6}$$

Furthermore $\bigoplus_{l=0}^\infty \mathcal{W}^l$, $\bigoplus_{l=0}^\infty \widetilde{\mathcal{W}}^l$ are supposed to be dense in $L_2(I)$.

(iv) Vanishing moments: The primal basis functions $\psi_{l,k}$ are assumed to satisfy vanishing moment conditions up to order $p^* + 1 \geq p$

$$\langle \psi_{l,k}, x^\alpha \rangle = 0 \ , \ \alpha = 0, \ldots, d = p^* + 1, l \geq 0, \tag{6.7}$$

and for all dual wavelets, except the ones at each end point, one has

$$\langle \widetilde{\psi}_{l,k}, x^\alpha \rangle = 0 \ , \ \alpha = 0, \ldots, d = p + 1, l \geq 0. \tag{6.8}$$

At the end points the dual wavelets satisfy only

$$\langle \widetilde{\psi}_{l,k}, x^\alpha \rangle = 0 \ , \ \alpha = 1, \ldots, d = p + 1, l \geq 0. \tag{6.9}$$

We remark that the third condition implies that the wavelets satisfy the zero Dirichlet condition, namely $\psi_{l,k}(0) = \psi_{l,k}(1) = 0$; the representation of this boundary condition by the subspace is important in the pricing of barrier contracts. To satisfy the homogeneous Dirichlet condition by the wavelet basis, we sacrifice the vanishing moment property of those wavelets whose supports include the endpoints of I, i.e. $x = 0$ or $x = 1$. For example, $\psi_{l,0}$, $l = 0, \ldots$, at the end point $x = 0$ (assuming that the localization index $k \in \nabla_l$ enumerates the wavelets in the direction of increasing values of x).

A systematic and general construction for arbitrary order biorthogonal spline wavelets is presented in [38]. Sufficiently far apart from the end points of $(0,1)$, biorthogonal wavelet (e.g. [29] and the references there) bases are used in this approach. In the recent paper [59] a wavelet bases was constructed with slightly smaller support at the end points. Using biorthogonal wavelets in the case $p = 1$, piecewise linear spline wavelets vanishing outside $I = (0,1)$ are obtained by simple scaling. The interior wavelets have two vanishing moments and are obtained from the mother wavelet $\psi(x)$ which takes the values $(0, -\frac{1}{6}, -\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}, -\frac{1}{6}, 0, 0, 0)$ at the points $(0, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{7}{8}, 1)$ by scaling and translations: $\psi_{l,k}(x) := 2^{l/2}\psi(2^{l-3}x - k + 2)$ for $2 \leq k \leq 2^l - 3$ and $l \geq 3$. At the left boundary $k = 1$, we use the piecewise linear function ψ_{left} defined by the nodal values $(0, \frac{5}{8}, \frac{-3}{4}, \frac{-1}{4}, \frac{1}{4}, \frac{1}{8}, 0, 0, 0)$ and $\psi_{right}(x) = \psi_{left}(1 - x)$. For additional details we refer to [59].

The following particular system of biorthogonal spline wavelet basis functions are Riesz bases for all constant or variable order Sobolev spaces of order $s \in [0,1]$ (and only these spaces arise as domains of the infinitesimal generators admissible time-homogeneous market models) and have proved efficient for our present applications [85]. They are a biorthogonal system of piecewise linear, continuos polynomial spline wavelets which were optimized for having small support. Their dual wavelets do not permit compact support, but they are nevertheless exponentially decaying. Any function $v \in \widetilde{H}^s(I)$, $0 \leq s \leq p + 1$, and, due to the embeddings $\widetilde{H}^{\underline{m}}(I) \subset \widetilde{H}^{m(x)}(I) \subset \widetilde{H}^{\overline{m}}(I)$, in particular any function $v \in \widetilde{H}^{m(x)}(I)$ can be represented in the wavelet series

$$v = \sum_{l=0}^{\infty} \sum_{k=1}^{M^l} v_{l,k} \psi_{l,k} \ = \ \sum_{\lambda \in \mathcal{I}} v_\lambda \psi_\lambda, \qquad v_\lambda = \int_I v \widetilde{\psi_\lambda} dx. \tag{6.10}$$

Here, we used the symbol $\lambda = (l, k)$ to denote a generic index in the index set

$$\mathcal{I} := \{\lambda = (l, k) : l = 0, 1, 2 \dots\ ,\ k = 1, \dots M_l\}.$$

Approximations v_h of functions $v \in \widetilde{H}^{m(x)}(I)$ can be obtained by truncating the wavelet expansion (6.10). In this way, a "quasi-interpolating" approximation operator Q_h : $\widetilde{H}^{m(x)}(I) \to V_h$, can be defined by truncating the wavelet expansion, i.e. by

$$Q_h v = \sum_{l=0}^{L-1} \sum_{k=1}^{M^l} v_{l,k} \psi_{l,k}. \tag{6.11}$$

For all $v_h = \sum_{l=0}^{L-1} \sum_{k=1}^{M^l} v_{l,k} \psi_{l,k} \in V_h = \mathcal{V}^L$, $h \sim 2^{-L}$, there holds the norm equivalence

$$\|v_h\|^2_{\widetilde{H}^s(I)} \sim \sum_{l=0}^{L-1} \sum_{k=1}^{M^l} |v_{l,k}|^2 2^{2ls}, \tag{6.12}$$

for all $0 \leq s < \frac{3}{2}$. This result is sharp in the sense that the norm equivalence fails in the upper limit $s = 3/2$; spline-wavelet systems consisting of higher order, piecewise polynomials with higher regularity across interval boundaries are known, but are not required in the present context, as the arguments in Dirichlet forms of Feller processes must belong locally to $H^1(\mathbb{R}^d)$, at best.

Validity of (6.12) in the variable order spaces $\widetilde{H}^{m(x)}(I)$ was shown in [97, Theorem 3], where $m(x) \in C^\infty(\mathbb{R})$, $m(x) \in [0, 1)$ was considered. There, it was in particular shown that for $u \in \widetilde{H}^{m(x)}$ it holds

$$\|u\|^2_{\widetilde{H}^{m(x)}(I)} \sim \sum_{l=0}^{\infty} \sum_{k=1}^{M^l} |u_{l,k}|^2\, 2^{2\underline{m}_\lambda l}, \tag{6.13}$$

where we recall the notation $\lambda = (l, k) \in \mathcal{I}$ and $\underline{m}_\lambda$ which is defined as

$$\underline{m}_\lambda := \inf\{m(x) : x \in \Omega_\lambda\} \quad \text{and} \quad \overline{m}_\lambda := \sup\{m(x) : x \in \Omega_\lambda\} \tag{6.14}$$

for the extended support Ω_λ of a wavelet basis function ψ_λ defined by

$$\Omega_\lambda := \Omega_{l,k} = \bigcup_{\lambda' \in \mathcal{I} : l' \geq l} \{\operatorname{supp} \psi_{\lambda'} : \operatorname{supp} \psi_\lambda \cap \operatorname{supp} \psi_{\lambda'} \neq \emptyset\}. \tag{6.15}$$

For $0 \leq s < \frac{3}{2} \leq t \leq p + 1$, we have the approximation property (e.g. [29])

$$\|v - Q_h v\|_{\widetilde{H}^s(I)} \leq C h^{t-s} \|v\|_{H^t(I)}. \tag{6.16}$$

Remark 6.2.1. *The pricing equations can also be considered in real price variables and not in log-price variables as described above, this leads to pseudodifferential operators, whose domains are weighted Sobolev spaces with possibly degenerated weights. An example for such kind of equations is given by the CEV model, cf. Remark 2.1.7. Norm*

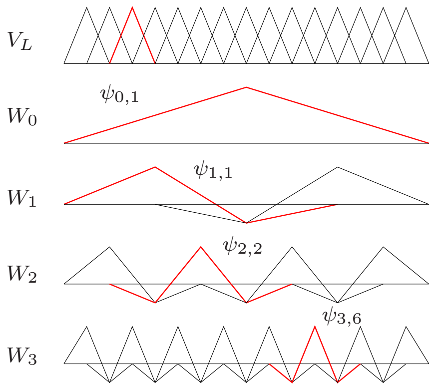


Figure 6.1: Single-scale space V_L and its decomposition into multiscale wavelet spaces W_ℓ for $L = 3$ and $p = 1$.

equivalences and efficient preconditioning for this kind of equations has been considered by [13]. The corresponding norm equivalence for a weighted space $L^2_w(0,1)$ with norm $\|u\|^2_{L^2_w} = \int_0^1 (w(x))^2 \, |u(x)|^2 \, dx$ reads:

$$\|v\|^2_{L^2_w(0,1)} \sim \sum_{l=0}^{\infty} \sum_{k=1}^{M^l} |v_{l,k}|^2 w^2(2^{-l}k), \tag{6.17}$$

for $v \in L^2_w$ and w being a possibly singular weight function fulfilling weak smoothness assumptions, cf. [13, Assumption 3.1]. Note that we obtain a variable weight in the exponent in (6.13) *for the variable order Sobolev space, while we obtain a variable weight in the case of the weighted space* (6.17). *It is also possible to combine the variable order Sobolev spaces and the weighted Sobolev spaces, leading to weighted variable order spaces with analogous norm equivalences to* (6.13) *and* (6.17).

6.3 Tensor product spaces

On $D = (0,1)^d$, $d > 1$, we define the subspace V_L of $\widetilde{H}^{\mathbf{m}(x)}(D)$ as the full tensor product of d univariate approximation spaces, i.e. $V_L := \bigotimes_{1\le i\le d} \mathcal{V}^{l_i}$, which can be written as

$$V_L = \quad \{\psi_{\mathbf{l},\mathbf{k}} : 0 \le l_i \le L-1, k_i \in \nabla_{l_i}, i = 1, \ldots, d\},$$

with basis functions $\psi_{\mathbf{l},\mathbf{k}} = \psi_{l_1,k_1} \cdots \psi_{l_d,k_d}$, $0 \le l_i \le L-1$, $k_i \in \nabla_{l_i}$, $i = 1, \ldots, d$. We can write V_L in terms of increment spaces

$$V_L = \bigoplus_{0\le l_i\le L-1} \mathcal{W}^{l_1} \otimes \ldots \otimes \mathcal{W}^{l_d}.$$

Therefore, we have for any function $u \in L^2(D)$ the series representation

$$u \;=\; \sum_{l_i=0}^{\infty} \sum_{k_i \in \nabla_{l_i}} u_{\mathbf{l},\mathbf{k}} \psi_{\mathbf{l},\mathbf{k}}.$$

Using the univariate norm equivalences and the intersection structure, cf. (2.14)-(2.15), we obtain

$$\|u\|^2_{H^{\mathbf{m}(x)}} \sim \sum_{\lambda} \left(2^{2\underline{m}^1_{\lambda_1} l_1} + \ldots + 2^{2\underline{m}^d_{\lambda_d} l_d} + 1\right) |u_\lambda|^2 . \tag{6.18}$$

Corollary 6.3.1. *Let $u \in H^s(D) \cap \widetilde{H}^1(D)$ for some $1 < s \le p+1$. Then for the quasi-interpolant $u_h = Q_h u = \sum_{l_i=0}^{L-1} \sum_{k=1}^{M^{l_i}} u_{\mathbf{l},\mathbf{k}} \psi_{\mathbf{l},\mathbf{k}}$ there holds for $0 < \overline{m} < 1 \le s \le p+1$ the Jackson estimate*

$$\begin{aligned}
\|u - u_h\|^2_{\widetilde{H}^{m(x)}(D)} &\lesssim \int_I \left(2^{2L(m_1(x_1)-s)} + \ldots + 2^{2L(m_d(x_d)-s)}\right) (|D^s u(x)|^2 + |u(x)|^2) dx \\
&\lesssim 2^{2L(\overline{m}-s)} \|u\|^2_{H^s(D)},
\end{aligned}$$

where $\overline{m} = \max_{i=1,\ldots,d} \overline{m}_i$.

Proof. For multi-indices $\lambda = (l,k), \mu = (L,k') \in \mathcal{I}$, we introduce the notation $\lambda \succeq \mu$ if $l_i \ge L_i$ and $\operatorname{supp} \psi_{\lambda_i} \cap \operatorname{supp} \psi_{\mu_i} \neq \emptyset$ for all $i = 1, \ldots, d$. For $s \ge \frac{3}{2}$ we choose $s' < s$ with $1 \le s' < \frac{3}{2}$, otherwise we set $s' = s$. We observe that $\underline{m}_{\lambda_i} - s' \le \overline{m}_{\mu_i} - s' < 0$ holds for all $\lambda_i \succeq \mu_i$. Therefore we conclude from the norm equivalence (6.18)

$$\begin{aligned}
\|u - u_h\|^2_{\widetilde{H}^{m(x)}(D)} &\sim \sum_{l_i \ge L} \sum_{k_i=1}^{M_{l_i}} \left(2^{2l_1 \underline{m}^1_{\lambda_1}} + \ldots + 2^{2l_d \underline{m}^d_{\lambda_d}}\right) |u_\lambda|^2 \\
&= \sum_{l_i \ge L} \sum_{k_i=1}^{M_{l_i}} \left(2^{2l_1(\underline{m}^1_{\lambda_1} - s')} 2^{2l_1 s'} + \ldots + 2^{2l_d(\underline{m}^d_{\lambda_d} - s')} 2^{2l_d s'}\right) |u_\lambda|^2 \\
&\lesssim \sum_{\mu \in \nabla_L} \left(2^{2L(\overline{m}^1_{\mu_1} - s')} + \ldots + 2^{2L(\overline{m}^d_{\mu_d} - s')}\right) \\
&\quad \times \sum_{\lambda \succeq \mu} \left(2^{2s' l_1} + \ldots + 2^{2s' l_d}\right) |u_\lambda|^2,
\end{aligned}$$

where $\nabla_L = \{\mu = (L,k') : k'_i = 1, \ldots, M_L \, , i = 1, \ldots, d\}$.

Let $\mu = (L,k')$, $L = |\mu|$ and $\square_\mu := \Pi_{i=1}^d [2^{-L} k'_i, 2^{-L}(k'_i + 1)]$. Then, by the norm equivalence (6.18) and the approximation property (6.16), we have

$$\sum_{\mu \in \nabla_L} \sum_{\lambda \succeq \mu} \left(2^{2s' l_1} + \ldots + 2^{2s' l_d}\right) |u_\lambda|^2 \lesssim \sum_{\mu \in \nabla_L} 2^{2L(s'-s)} \|u\|^2_{H^s(\square_\mu)} .$$

Recalling that $2^{L\underline{m}^i_\mu} \sim 2^{Lm^i(x_i)} \sim 2^{L\overline{m}^i_\mu}$ holds for all $x \in \Box_\mu$, we obtain the final result

$$\begin{aligned} \|u - u_h\|^2_{\widetilde{H}^{m(x)}(D)} &\lesssim \int_I \left(2^{2L(m_1(x_1)-s)} + \ldots + 2^{2L(m_d(x_d)-s)}\right)(|D^s u(x)|^2 + |u(x)|^2)dx \\ &\lesssim 2^{2L(\overline{m}-s)} \|u\|^2_{H^s(D)} \,. \end{aligned}$$

$\Box$

7 FE Discretization for time-homogeneous PIDEs

We consider the discretization of problem (5.9)-(5.10), i.e., time-homogeneous market models are studied in this section. We develop and analyze stable discretization schemes. First, results for continuous Galerkin (CG) Finite Element Method (FEM) for pricing equations without drift dominance are presented. Then, two different methods for drift dominated equations are analyzed, a discontinuous Galerkin (DG) Finite Element Method and a CG Finite Element Method with streamline diffusion (SD) stabilization. In the DG-FEM, the small jump regularization of the hypersingular integrals in the Dirichlet form of the pure jump part of infinite variation processes has to be employed in order to obtain a well-defined scheme. This is due to the fact that the basis functions are not globally Lipschitz and therefore the bilinear form $a_{\mathrm{J}}(\cdot,\cdot)$ is not well-defined for such functions. Robustness of the stabilized discretization with respect to various degeneracies in the characteristic triplet of the stochastic process is proved. In the CG-FEM, the SD method is used in the case of a drift dominated equation, as standard methods are known to be unstable in such a situation. Since continuous, piecewise polynomial functions on regular simplicial partitions are Lipschitz, for such functions the Dirichlet form $a_{\mathrm{J}}(\cdot,\cdot)$ in (3.16) which corresponds to the jump part of the process X is finite, even if $\overline{Y} > 1$, i.e., if X has sample paths of infinite variation (see, e.g. [52, 95]). Therefore, in contrast to the DG-FEM we do not regularize the jump measure in this case.

7.1 Continuous Galerkin discretization

7.1.1 A priori error estimate

We first discretize the problem (5.9)-(5.10) with respect to the space variable. The finite element space $V_L \subset \mathcal{V}_D$, as in Section 6.3, is the space of all continuous piecewise polynomials of degree $p \geq 1$ which vanish at the boundary ∂D. The semi-discrete problem corresponding to (5.9)-(5.10) reads: Given $g \in L^2(\mathbb{R}^d)$, find $u_L \in H^1(I, V_L)$ such that

$$(\partial_t u_L, v_L) + a_D(u_L, v_L) = (f, v_L), \quad \forall v_L \in V_L, \tag{7.1}$$

$$u_L(0) = P_h g|_D, \tag{7.2}$$

where P_h is the L^2 projection onto V_L. We first consider a setup without drift dominance as in Section 5.2. In such a setup the a piori error estimate follows along the lines of [86, Section 5], i.e., the following result holds.

Theorem 7.1.1. *Let $\mathcal{A}(x,D) \in \Psi^{2\mathbf{m}(x)}_{\rho,\delta}$ be the generator of a time-homogeneous admissible market with $0 < \delta \leq \rho \leq 1$ with symbol $a(x,\xi)$ satisfying* (5.11)*, then for all $t \in I$, there holds*

$$\|u(t) - u_L(t)\|_{L^2(D)} \leq C \min\left\{1, h^{p+1} t^{-\frac{p+1}{p}}\right\} \left(\|g\|_{L^2(D)} + \|f\|_{L^2(I,H^{-\mathbf{m}(x)}(D))}\right).$$

Here, C is a positive constant, independent of h and t while u and u_L are the solutions of (5.9)-(5.10) *and* (7.1)-(7.2)*.*

Note that Theorem 7.1.1 does not cover the drift dominated case, cf. Theorem 5.2.5. We present two approaches for this case. An SD-FEM approach is discussed in the following section and DG-FEM is analysed in Section 7.3.

7.2 Streamline diffusion discretization

In order to obtain a formulation that is numerically feasible, we need to stabilize the transport operator, since the application of a standard Finite Element discretization generally leads to unstable numerical solutions. In the following we consider the discretization on a general triangulation as introduced in Section 6.1. Convergence results for tensor product spaces are special cases of the more general results.

7.2.1 Streamline diffusion formulation

We consider problem (5.24). Let the space $S^{p,1}(D,\mathcal{T}_h)$ of continuous piecewise polynomial functions v_h on a triangulation $\mathcal{T}_h$ be given as in Section 6.1. Moreover, we assume that the following condition holds.

Assumption 7.2.1. *Let X be a time-homogeneous admissible market model, then we make the following regularity assumption on the drift $b(x)$*

$$b(x) \cdot \nabla_h v_h \in S^{p,1} \quad \forall v_h \in S^{p,1}. \tag{7.3}$$

This condition is further discussed in Remark 7.3.10. The SD-FEM formulation reads as follows: find $u_h^\delta \in H^1\left((0,T); S_0^{1,p}\right)$ such that

$$\begin{aligned} \left(\partial_t u_h^\delta, v_h\right) + a^\delta_{\mathrm{SD}}(u_h^\delta, v_h) &= l^\delta_{\mathrm{SD}}(v_h) \quad \forall v_h \in S_0^{p,1}, && (7.4)\\ u_h^\delta(0,x) &= P_h g(x)|_D, && (7.5) \end{aligned}$$

with

$$a^\delta_{\mathrm{SD}}(u,v) := a_{\mathrm{Tr}}(u,v) + a_{\mathrm{J}}(u,v) + a^\delta_{\mathrm{Tr}}(u,v) \text{ and } l^\delta_{\mathrm{SD}}(v) := (f, v + \delta b\cdot\nabla v), \quad (7.6)$$

where $a^\delta_{\mathrm{Tr}}(u,v) = \delta a_{\mathrm{Tr}}(u, b\cdot\nabla v)$ and P_h denotes the L^2-projection onto $S_0^{p,1}$. We define $S_0^{p,1} = \{v_h \in S^{p,1} : v_h = 0 \text{ on } \partial D\}$. Note that an analogous formulation can be obtained for non-homogeneous Dirichlet boundary conditions.

7.2.2 *A priori* bound

We introduce the streamline-diffusion norm $\|\cdot\|_{\mathrm{SD}(\delta)}$, defined by

$$\|w\|^2_{SD(\delta)} := \|w\|^2_{L^2(D)} + \|b\cdot nw\|^2_{L^2(\Gamma)} + \|w\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} + \delta\,\|b\cdot\nabla w\|^2_{L^2(D)}\,, \quad (7.7)$$

where n denotes the exterior unit normal vector on D.

Theorem 7.2.2. *Let $a^\delta_{\mathrm{SD}}(\cdot,\cdot)$ be as in (7.6) and $\|\cdot\|_{SD(\delta)}$ be given by (7.7), further assume the bilinear form $a_{\mathrm{J}}(\cdot,\cdot)$ is coercive, then the following estimate holds if $\delta < 4r_{\min} r^{-2}$*

$$a^\delta_{\mathrm{SD}}(w,w) \geq C_{\mathrm{co}}\,\|w\|^2_{SD(\delta)} \quad \forall w \in H^1(D), \quad (7.8)$$

for $C_{\mathrm{co}} > 0$ and $r_{\min}$ as in (5.22).

Proof. By integration by parts and using properties of the operators we obtain the following estimate: for any $w \in H^1(D)$

$$\begin{aligned}
a^\delta_{\mathrm{SD}}(w,w) &\geq r_{\min}\,\|w\|^2_{L^2(D)} + \frac{1}{2}\,\|b\cdot nw\|^2_{L^2(\Gamma)} + C\,\|w\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} \\
&\quad +\delta\,\|b\cdot\nabla w\|^2_{L^2(D)} - \delta\,\|rw\|_{L^2(D)}\,\|b\cdot\nabla w\|_{L^2(D)} \\
&\geq r_{\min}\,\|w\|^2_{L^2(D)} + \frac{1}{2}\,\|b\cdot nw\|^2_{L^2(\Gamma)} + C\,\|w\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} \\
&\quad +\delta\,\|b\cdot\nabla w\|^2_{L^2(D)} - r^2\frac{\delta}{2\widehat{C}}\,\|w\|^2_{L^2(D)} - \frac{\delta\widehat{C}}{2}\,\|b\cdot\nabla w\|^2_{L^2(D)} \\
&= \left(r_{\min} - r^2\frac{\delta}{2\widehat{C}}\right)\|w\|^2_{L^2(D)} + \frac{1}{2}\,\|b\cdot nw\|^2_{L^2(\Gamma)} \\
&\quad +C\,\|w\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} + \delta\left(1 - \frac{\widehat{C}}{2}\right)\|b\cdot\nabla w\|^2_{L^2(D)}\,.
\end{aligned}$$

Choosing $\widehat{C}$ such that $\frac{r^2\delta}{2r_{\min}} < \widehat{C} < 2$ yields the required result with
$C_{\mathrm{co}} = \min\left\{r_{\min} - r^2\frac{\delta}{2\widehat{C}},\, \frac{1}{2},\, C,\, 1 - \frac{\widehat{C}}{2}\right\}$. □

Remark 7.2.3. *Note that the coercivity assumption on $a_{\mathrm{J}}(\cdot,\cdot)$ is not restrictive, as we may use the transformation $v(t,x) = e^{-tC^-}u(t,x)$ to obtain a coercive bilinear form in the equation satisfied by v, if $a_{\mathrm{J}}(u,u) \geq C\|u\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} - C^-\,\|u\|^2_{L^2(D)}$ holds.*

The coercivity of the bilinear form $a^{\delta}_{\mathrm{SD}}(\cdot,\cdot)$ implies the following stability result.

Theorem 7.2.4. *Let the assumptions of Theorem 7.2.2 hold and u_h^δ be the solution of (7.4), then the following* a priori *bound holds*

$$\left\|u_h^\delta(T)\right\|^2_{L^2(D)} + \int_0^T \left\|u_h^\delta(t)\right\|^2_{SD(\delta)} dt \le \frac{1}{C_{\mathrm{co}}^2}\left(1+\sqrt{\delta}\right)^2 \|f\|^2_{L^2((0,T);L^2(D))} + \|P_h g\|^2_{L^2(D)},$$

where C_{co} is the coercivity constant in Theorem 7.2.2.

Proof. Considering (7.8) and the right-hand side (7.6) it follows

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\left\|u_h^\delta(t)\right\|^2_{L^2(D)} + \left\|u_h^\delta(t)\right\|^2_{SD(\delta)} &\le \frac{1}{C_{\mathrm{co}}}\|f(t)\|_{L^2(D)}\left\|u_h^\delta(t)+\delta b\cdot\nabla u_h^\delta(t)\right\|_{L^2(D)}\\
&\le \frac{1}{C_{\mathrm{co}}}\|f(t)\|_{L^2(D)}(1+\sqrt{\delta})\left\|u_h^\delta(t)\right\|_{SD(\delta)}\\
&\le \frac{(1+\sqrt{\delta})^2}{2C_{\mathrm{co}}^2}\|f(t)\|^2_{L^2(D)} + \frac{1}{2}\left\|u_h^\delta(t)\right\|^2_{SD(\delta)}
\end{aligned}$$

and the result follows integrating in time. □

7.2.3 *A priori* error estimate

In order to prove an *a priori* error estimate, we need the following lemma.

Lemma 7.2.5. *Suppose that $K \in \mathcal{T}_h$ is a shape regular d-simplex or a shape regular d-parallelepiped of diameter h_K. Suppose further that $u|_K \in H^r(K)$, $r \ge 2$. Then there exists a projection $\Pi_{(p,K)}$ on the space of the polynomials of degree p in K such that, for $s \ge 1$, $p \ge 1$, s integer*

$$\begin{aligned}
||w-\Pi_{(p,K)}w||_{L^2(K)} &\le C\frac{h_K^{\min(p+1,s)}}{p^s}||w||_{H^s(K)},\\
||\nabla\left(w-\Pi_{(p,K)}w\right)||_{L^2(K)} &\le C\frac{h_K^{\min(p+1,s)-1}}{p^{s-1}}||w||_{H^s(K)},\\
||w-\Pi_{(p,K)}w||_{L^2(\partial K)} &\le C\frac{h_K^{\min(p+1,s)-\frac{1}{2}}}{p^{s-\frac{1}{2}}}||w||_{H^s(K)},\\
||\nabla\left(w-\Pi_{(p,K)}w\right)||_{L^2(\partial K)} &\le C\frac{h_K^{\min(p+1,s)-\frac{3}{2}}}{p^{s-\frac{3}{2}}}||w||_{H^s(K)}.
\end{aligned}$$

Moreover, for $0 < l < 1$, it holds

$$||w-\Pi_{(p,K)}w||_{H^l(K)} \le C\frac{h_K^{\min(p+1,s)-l}}{p^{s-l}}||w||_{H^s(K)}.$$

Proof. See [27], [65, Lemma 4.4] and [109, Section 4.5]. □

Remark 7.2.6. *In each estimate, s can be chosen differently, if w is sufficiently smooth.*

Remark 7.2.7. *Lemma 7.2.5 can be extended to the case $s \in (0, 2)$. In this case, the right-hand side of the interpolation estimates contains the term $||w||_{H^s(\delta_K)}$, where δ_K is given as the union of the element K with its neighbor elements, i.e.*
$\delta_K = \left\{K' \in \mathcal{T}_h : \overline{K'} \cap \overline{K} \neq \emptyset\right\}$. For example, all the elements of the triangulation $\mathcal{T}_h$ that share an edge $(d = 2)$ or vertex $(d = 3)$ with K.

The following result is needed in order to prove an *a priori* error estimate.

Lemma 7.2.8. *Let $w \in S^{p,1}$ and $K \in \mathcal{T}_h$, then for any $s > 0$ there exist a constant C_I independent on h and $p \geq 1$, but depending on C_1 and C_2 in (6.1), i.e., on the mesh regularity, such that*

$$||w||_{H^s(K)} \leq C_I p^{2s} h^{-s} ||w||_{L^2(K)}.$$

Proof. Let us consider the case $s = 1$ and $d = 2$. We obtain from [109, Theorem 4.76]

$$\|\widehat{w}\|_{H^1(\widehat{K})} \leq Cp^2 \|\widehat{w}\|_{L^2(\widehat{K})},$$

where $\widehat{w} = F(w)$ and with $F : K \to \widehat{K}$ denoting the element mapping inverse function. Scaling arguments (as in [22, Section 4.5]) show that

$$\begin{aligned} \|w\|^2_{H^1(K)} &= \|w\|^2_{L^2(K)} + |w|^2_{H^1(K)} \leq C\left(\|\widehat{w}\|^2_{L^2(\widehat{K})} + \frac{1}{h_K^2}|\widehat{w}|^2_{H^1(\widehat{K})}\right) \\ &\leq C\left(1 + \frac{p^4}{h_K^2}\right)\|\widehat{w}\|^2_{L^2(\widehat{K})} \leq C\left(1 + \frac{C_2^2 p^4}{h^2}\right)\|w\|^2_{L^2(K)}. \end{aligned}$$

The other cases can be proved similarly. □

We define the norm $|||\cdot|||_{SD(\delta)}$ by:

$$|||w(t)|||^2_{SD(\delta)} = ||w(t)||^2_{L^2(D)} + \int_0^t \|w(\tau)\|^2_{SD(\delta)}\, d\tau \quad \forall t \in (0, T).$$

Theorem 7.2.9. *Let u be the solution of problem (5.24) and u_h^δ be the solution of (7.4)-(7.5), then the following estimate holds for $s \geq 2$, $0 < \overline{m} < 1$, $t \in (0, T)$ and $C > 0$ independent of p and h*

$$\begin{aligned} |||(u - u_h^\delta)(t)|||_{SD(\delta)} &\leq C \max\left(\sqrt{\delta} h^{-\overline{m}} p^{2\overline{m}}, \frac{h^{\min(p+1,s) - \max(2\overline{m},1)/2}}{p^{s - \max(2\overline{m},1)/2}}\right) \\ &\quad \times \left(\|u(t)\|_{H^s(D)} + \|u\|_{H^1((0,T);\, H^s(D))}\right). \end{aligned}$$

Proof. We set

$$\eta := u^{\varepsilon} - \Pi_p u^{\varepsilon} \text{ and } \xi := u_h^{\varepsilon} - \Pi_p u^{\varepsilon}, \tag{7.9}$$

where obviously $\xi \in S^{p,1}$ holds. Notice that a Galerkin orthogonality is not available in this case, but a quasi orthogonality can be used, i.e.,

$$B_\delta(\xi,\xi) = B_\delta(\eta,\xi) - B_\delta(u - u_h^\delta, \xi), \tag{7.10}$$

where we set $B_\delta(w,v) := (\partial_t w,\, v) + a_{\mathrm{SD}}^\delta(w,v)$ and

$$B_\delta(u - u_h^\delta, \xi) = -\delta(\partial_t u, b\cdot\nabla\xi) - \delta a_{\mathrm{J}}(u, b\cdot\nabla\xi). \tag{7.11}$$

Let us examine the terms in (7.10) separately. Since $B_\delta(\eta,\xi) = (\partial_t\eta,\, \xi) + a_{\mathrm{SD}}^\delta(\eta,\xi)$ holds, we obtain, using Assumption 7.2.1 and the interpolation error estimates of Lemma 7.2.5, for $\delta = Ch^\alpha p^{-\beta}$, $\alpha,\beta > 0$

$$\begin{aligned}
\left|a_{\mathrm{SD}}^\delta(\eta,\xi)\right| &= \left|a\left(\eta,\xi\right) + a_{\mathrm{Tr}}^\delta\left(\eta,\xi\right)\right| \\
&\leq \left(r + ||\mathrm{div}(b)||_{L^\infty(D)}\right) \|\eta\|_{L^2(D)} \|\xi\|_{L^2(D)} \\
&+ \|b\cdot n\,\eta\|_{L^2(\Gamma)} \|b\cdot n\,\xi\|_{L^2(\Gamma)} + C\|\eta\|_{\widetilde{H}^{\mathbf{m}(x)}(D)} \|\xi\|_{\widetilde{H}^{\mathbf{m}(x)}(D)} \\
&+ \delta\left(r\|\eta\|_{L^2(D)} \|b\cdot\nabla\xi\|_{L^2(D)} + \|b\cdot\nabla\eta\|_{L^2(D)} \|b\cdot\nabla\xi\|_{L^2(D)}\right) \\
&\leq C\left((1+\sqrt{\delta})\|\eta\|_{L^2(D)} + \|\eta\|_{\widetilde{H}^{\mathbf{m}(x)}(D)} + \|b\cdot n\,\eta\|_{L^2(\Gamma)} + \sqrt{\delta}\|b\cdot\nabla\eta\|_{L^2(D)}\right) \\
&\quad \times\|\xi\|_{SD(\delta)} \\
&\leq C\frac{h^{\min(p+1,s)}}{p^s}\left(1 + \frac{h^{\alpha/2}}{p^{\beta/2}} + \frac{h^{-\overline{m}}}{p^{-\overline{m}}} + ||b||_{L^\infty(D)}\left(\frac{h^{-1/2}}{p^{-1/2}} + \frac{h^{\alpha/2-1}}{p^{\beta/2-1}}\right)\right) \\
&\quad \times||u||_{H^s(D)}\|\xi\|_{SD(\delta)} \\
&\leq C\frac{h^{\min(p+1,s)-\max(2\overline{m},\,1,\,2-\alpha)/2}}{p^{s-\max(2\overline{m},\,1,\,2-\beta)/2}}\left(1 + ||b||_{L^\infty(D)}\right)||u||_{H^s(D)}\|\xi\|_{SD(\delta)}\,.
\end{aligned} \tag{7.12}$$

Moreover, it holds

$$\begin{aligned}
\left|B_\delta(u-u_h^\delta,\xi)\right| &= |\delta(\partial_t u, b\cdot\nabla\xi) + \delta a_{\mathrm{J}}(u, b\cdot\nabla\xi)| \\
&\leq \delta\|\partial_t u\|_{L^2(D)}\|b\cdot\nabla\xi\|_{L^2(D)} + C\delta\|u\|_{\widetilde{H}^{\mathbf{m}(x)}(D)}\|b\cdot\nabla\xi\|_{\widetilde{H}^{\mathbf{m}(x)}(D)}\,.
\end{aligned}$$

Considering Lemma 7.2.8, we have

$$\|b\cdot\nabla\xi\|_{\widetilde{H}^{\mathbf{m}(x)}(D)} \leq C_I h^{-\overline{m}} p^{2\overline{m}} \|b\cdot\nabla\xi\|_{L^2(D)}$$

and thus

$$\left|B_\delta(u-u_h^\delta,\xi)\right| \leq \sqrt{\delta}\left(\|\partial_t u\|_{L^2(D)} + Ch^{-\overline{m}}p^{2\overline{m}}\|u\|_{\widetilde{H}^{\mathbf{m}(x)}(D)}\right)\|\xi\|_{SD(\delta)}\,,$$

where C depends on the continuity constant of the bilinear form $a_J(\cdot,\cdot)$ and on C_I in Lemma 7.2.8. Thus it holds

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}||\xi||^2_{L^2(D)} + ||\xi||^2_{\mathrm{SD}(\delta)} \le \frac{1}{C_{\mathrm{co}}} B_\delta(\xi,\xi)\\
= \ & \frac{1}{C_{\mathrm{co}}}\Big((\partial_t\eta,\xi) + a^\delta_{\mathrm{SD}}(\eta,\xi) - B_\delta(u-u^\delta_h,\xi)\Big)\\
\le \ & \frac{\widehat{C_1}}{2}\frac{h^{2\min(p+1,s)}}{p^{2s}}||\partial_t u||^2_{H^s(D)} + \frac{1}{2\widehat{C_1}}||\xi||^2_{\mathrm{SD}(\delta)} + \frac{1}{2\widehat{C_2}}\|\xi\|^2_{SD(\delta)} + \frac{1}{2\widehat{C_3}}\|\xi\|^2_{SD(\delta)}\\
&+\frac{\widehat{C_2}C^2}{2}\frac{h^{2\min(p+1,s)-\max(2\overline{m},1,2-\alpha)}}{p^{2s-\max(2\overline{m},1,2-\beta)}}\left(1+||b||_{L^\infty(D)}\right)^2||u||^2_{H^s(D)}\\
&+\frac{\widehat{C_3}}{2}\delta\left(\|\partial_t u\|_{L^2(D)} + Ch^{-\overline{m}}p^{2\overline{m}}\|u\|_{\widetilde{H}^{\mathbf{m}(x)}(D)}\right)^2.
\end{aligned}$$

Therefore, choosing positive constants $\widehat{C_1}$, $\widehat{C_2}$ and $\widehat{C_3}$ such that $\frac{1}{2\widehat{C_1}}+\frac{1}{2\widehat{C_2}}+\frac{1}{2\widehat{C_3}}=\frac{1}{2}$, we obtain

$$\begin{aligned}
|||\xi|||^2_{SD(\delta)} = \frac{d}{dt}||\xi||^2_{L^2(D)} + ||\xi||^2_{\mathrm{SD}(\delta)} \ &\le \ C\max\left(\frac{h^{(\alpha-2\overline{m})/2}}{p^{\beta/2-2\overline{m}}},\frac{h^{\min(p+1,s)-\max(2\overline{m},1)/2}}{p^{s-\max(2\overline{m},1)/2}}\right)^2\\
&\left(\|\partial_t u\|_{L^2(D)} + \left(1+||b||_{L^\infty(D)}\right)||u||_{H^s(D)}\right)^2
\end{aligned}$$

Since $|||u-u^\delta_h|||_{SD(\delta)} \le |||\xi|||_{SD(\delta)} + |||\eta|||_{SD(\delta)}$ holds and since the estimate for $|||\eta|||_{SD(\delta)}$ follows from Lemma 7.2.5, we conclude the claimed result. □

Remark 7.2.10. *As the scheme is only asymptotically consistent, this has to be considered in the convergence analysis and leads to strong restrictions on* δ. *A fully consistent discretization scheme reads: find* $u^\delta_h \in S^{p,1}$ *s.t.* $\forall v\in S^{p,1}$

$$\begin{aligned}
\left(\partial_t u^\delta_h,\, v+\delta b\cdot\nabla v\right) + a(u^\delta_h, v+\delta b\cdot\nabla v) \ &= \ (f, v+\delta b\cdot\nabla v)\\
u^\delta_h(0,x) \ &= \ P_h g(x).
\end{aligned}$$

The following error estimate can be obtained along the lines of Theorem 7.2.9 using Galerkin orthogonality

$$\begin{aligned}
|||(u-u^\delta_h)(t)|||_{SD(\delta)} \ &\le \ C\,\frac{h^{\min(p+1,s)-\max(2\overline{m},1,2-\alpha)/2}}{p^{s-\max(2\overline{m},1,2-\beta)/2}}\\
&\times\left(\|u(t)\|_{H^s(D)} + \|u\|_{H^1((0,T);\,H^s(D))}\right),
\end{aligned}$$

for $\delta = Ch^\alpha p^{-\beta}$, $\alpha,\beta>0$, $C>0$ *sufficiently large and independent of* h *and* p. *Choosing* $\alpha=\beta=1$ *would lead to a convergence rate of* $\min(p+1,s)-\max(2\overline{m},1)/2$ *in* h *and* $s-\max(2\overline{m},1)/2$ *in* p.

Remark 7.2.11. *We refer to [60, Chapter 5] for an analysis of SD-FEM for PIDEs on sparse tensor product spaces.*

7.3 Discontinuous Galerkin discretization

A DG-discretization scheme for the forward equation is described in this section. After introducing the necessary notations, the numerical scheme is presented and analyzed. An error analysis in multiple space dimensions is performed.

7.3.1 Discontinuous Galerkin formulation

The DG semidiscrete formulation of (5.33) reads as follows: for a (sufficiently small) jump regularization parameter $\varepsilon > 0$, find $u_h^\varepsilon \in H^1((0,T);S^{p,0})$ such that for all $v_h \in S^{p,0}$ it holds

$$\begin{aligned}(\partial_t u_h^\varepsilon(x), v_h(x)) + a_{\mathrm{DG}}^\varepsilon(u_h^\varepsilon(x), v_h(x)) &= l_{\mathrm{DG}}^\varepsilon(v_h(x)), && (7.13)\\ u_h^\varepsilon(0,x) &= \Pi_p g(x), && (7.14)\end{aligned}$$

where $\Pi_p g$ is the L^2-projection of the initial condition function g in $S^{p,0}$ as in Section 6.1, and

$$a_{\mathrm{DG}}^\varepsilon(w,v) := d_{\mathrm{DG}}^\varepsilon(w,v) + t_{\mathrm{DG}}(w,v) + r_{\mathrm{DG}}(w,v) + j_{\mathrm{DG}}^\varepsilon(w,v), \quad (7.15)$$

$$l_{\mathrm{DG}}^\varepsilon(v) = \int_D f v dx + bc_{\mathrm{DG}}(v). \quad (7.16)$$

The bilinear forms $d_{\mathrm{DG}}^\varepsilon(\cdot,\cdot)$, $t_{\mathrm{DG}}(\cdot,\cdot)$, $r_{\mathrm{DG}}(\cdot,\cdot)$, $j_{\mathrm{DG}}^\varepsilon(\cdot,\cdot)$ and the boundary term $bc_{\mathrm{DG}}(\cdot)$ are defined as follows for any $v, w \in S^{p,0}$.

(i) Diffusion term $d_{\mathrm{DG}}^\varepsilon(\cdot,\cdot)$: for ease of notation we drop the dependency on time t and space x. It holds

$$\begin{aligned}d_{\mathrm{DG}}^\varepsilon(w,v) \;:=\;& \sum_{K\in\mathcal{T}_h} \frac{1}{2}\int_K \nabla w^\top Q(x,\varepsilon)\nabla v dx - \sum_{e\in\Gamma_h}\frac{1}{2}\int_e \left\{\nabla w^\top Q(s,\varepsilon) n\right\}[v] ds\\ +\;& \frac{\beta}{2}\sum_{e\in\Gamma_h}\int_e [w]\left\{\nabla v^\top Q(s,\varepsilon) n\right\} ds + \sum_{e\in\Gamma_h}\frac{\alpha}{|e|}\int_e [w][v] ds, \end{aligned} \quad (7.17)$$

where $\alpha > 0$ is independent of h and ε; $\beta = -1$ yields the Symmetric Interior Penalty Galerkin (SIPG) method (which converges only if α is sufficiently large), while $\beta = 1$ gives the Non-Symmetric Interior Penalty Galerkin (NIPG) method. See [100, Chapter 2] for further details. From now on we set $\beta = 1$, i.e., we discretize the diffusion term with the NIPG method.

(ii) Transport term $t_{\mathrm{DG}}(\cdot,\cdot)$: following [65], we obtain

$$t_{\mathrm{DG}}(w,v) := \sum_{K\in\mathcal{T}_h}\int_K b\cdot\nabla w\, v dx - \int_{\partial_- K}(b\cdot n_K)[w] v_I ds, \quad (7.18)$$

where n_K is the normal unit vector exterior to K, v_I (v_O) is the inner (outer) trace of v relative to K and, according to the above definition, $[v] = v_I - v_O$. Moreover we set

$$\partial_- K := \{x \in \partial K \ : \ b \cdot n_K < 0\} \quad \text{and} \quad \partial_+ K := \{x \in \partial K \ : \ b \cdot n_K > 0\}.$$

Note that the drift vector $b(x)$ might be ε-dependent if a martingale condition on the regularized market model is enforced, cf. Section 4.1.3. In this case ∂K also depends on ε.

(iii) Reaction term $r_{\mathrm{DG}}(\cdot,\cdot)$: it holds

$$r_{\mathrm{DG}}(w,v) := \int_D r\, w\, v dx. \tag{7.19}$$

(iv) Integrodifferential term $j^{\varepsilon}_{\mathrm{DG}}(\cdot,\cdot)$: since the jump operator $\mathcal{A}_{\mathrm{J}}(\varepsilon)$ as in (5.28) can be rewritten as

$$\mathcal{A}_{\mathrm{J}}(\varepsilon)\varphi(x) = \int_{\mathbb{R}^d} (\varphi(y) - \varphi(x)) k_\varepsilon(x, y-x) dy, \tag{7.20}$$

the integrodifferential term is given as

$$j^{\varepsilon}_{\mathrm{DG}}(w,v) := -\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (\widetilde{w}(y) - \widetilde{w}(x)) k_\varepsilon(x, y-x) dy\, \widetilde{v}(x) dx. \tag{7.21}$$

(v) The boundary term $bc_{\mathrm{DG}}(\cdot)$ in this case reads

$$bc_{\mathrm{DG}}(v) = 0.$$

Remark 7.3.1. *Note that if we impose nonhomogeneous Dirichlet conditions $\eta \neq 0$ as in* (5.33), *then the boundary term would read*

$$bc^{\varepsilon}_{DG} = \sum_{e \in \Gamma^0_h} \int_e \left(\frac{1}{2} \nabla v^\top Q(s,\varepsilon) n + \frac{\alpha}{|e|} v \right) \eta ds - \sum_{K \in \mathcal{T}_h} \int_{\partial_- K \cap \partial G} (b \cdot n_K) \eta v_I ds.$$

The first term stems from the discretization of the diffusion part, while the second term originates from the transport term.

Remark 7.3.2. *In [64] the authors deal with a DG discretization for the hyperbolic part $b \cdot \nabla u + cu = f$. More precisely, they discretize the bilinear form $t_{DG}(\cdot,\cdot) + r_{DG}(\cdot,\cdot)$ as in (7.18) and (7.19) adding the following stabilization term*

$$\delta \sum_{K \in \mathcal{T}_h} \int_K (b \cdot \nabla w + rw)\,(b \cdot \nabla v) dx,$$

for $\delta > 0$. The consistency of the method is ensured by adding the term $\delta \int_D f(b \cdot \nabla v) dx$ to the right-hand side of the equation. Consistency, stability and an error analysis is provided in [64] with $\delta = Chp^{-1}$, with C independent of h and p. However, numerical results (see [64, Section 5]) show that the scheme without stabilization, i.e., $\delta = 0$, is marginally more accurate for $p = 1$ and $p = 2$. For larger p the stabilized scheme is slightly more accurate.

Remark 7.3.3. *The above DG formulation is written with integrals over faces of the elements of the mesh, and thus for the case $d > 1$. In the univariate case, this is to be interpreted as follows: if $K = [a, b]$, then $\partial K = \{a, b\}$ and we set for $v \in \mathbb{P}^p(K)$*

$$\int_a v_I(x)dx = v(a), \quad \int_b v_I(x)dx = v(b), \quad \int_a v_I(x)ndx = -v(a), \quad \int_b v_I(x)ndx = v(b),$$

where $n = 1$ (-1) in b (a). Moreover, if h is the length of the interval K, i.e., $h := b - a$, we replace $|e|$ by h in (7.17).

Remark 7.3.4. *Let us consider, for simplicity, the univariate case, i.e., $d = 1$ in* (5.33), *and the integrodifferential term (7.21): if we denote by $k_\varepsilon^{(-1)}$ an antiderivative (to be specified below) of k_ε, then (7.21) can be rewritten as follows:*

$$\begin{aligned} j_{DG}^{\varepsilon}(w, v) &= \int_{\mathbb{R}} \sum_{K \in \mathcal{T}_h} \int_K \widetilde{w}'(y) k_\varepsilon^{(-1)}(x, y - x) dy\, \widetilde{v}(x) dx \\ &- \int_{\mathbb{R}} \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\widetilde{w}_I(s) - \widetilde{w}(x)) n k_\varepsilon^{(-1)}(x, s - x) ds\, \widetilde{v}(x) dx, \end{aligned}$$

where the antiderivatives of $k_\varepsilon(x, y)$ are given by

$$k_\varepsilon^{(-1)}(x, z) = \begin{cases} \int_{-\infty}^{z} k_\varepsilon(x, y) dy & \text{if } z < 0, \\ -\int_{z}^{+\infty} k_\varepsilon(x, y) dy & \text{if } z > 0 \end{cases}. \tag{7.22}$$

In the following sections we analyze the stability and derive error estimates for the DG semidiscrete formulation (7.13)-(7.14) of (5.33). We denote by u the smooth solution of problem (5.9)-(5.10), u^ε is the smooth solution of problem (5.33) with $\varepsilon > 0$ and $u_h^\varepsilon \in S^{p,0}$ is the solution of problem (7.13)-(7.14) according to the DG discretization. We prove the consistency of the considered DG scheme in Section 7.3.2, while in Sections 7.3.3 and 7.3.4 we deal with an *a priori* bound and error estimates of the DG solution.

7.3.2 Consistency

Theorem 7.3.5. *If u^ε is the solution of* (5.33), *then it satisfies* (7.13).

Proof. Let $v_h \in S^{p,0}$ be a test function. We obtain from (5.33)

$$(\partial_t u^\varepsilon, v_h) + a_D(u^\varepsilon, v_h) = (f, v_h).$$

Since

$$(\mathcal{A}_{\mathrm{J}}(\varepsilon)\widetilde{u}^\varepsilon, v_h) \equiv j_{\mathrm{DG}}^{\varepsilon}(u^\varepsilon, v_h), \quad (ru^\varepsilon, v_h) \equiv r_{\mathrm{DG}}(u^\varepsilon, v_h) \quad \text{and} \quad (f, v_h) \equiv l_{\mathrm{DG}}(v_h)$$

holds, we have to deal with the diffusion and transport terms in order to prove consistency of the method. However, the regularity of u^ε implies $[u^\varepsilon] = 0$ on Γ_h, thus $(b \cdot \nabla u^\varepsilon, v_h) = t_{\mathrm{DG}}(u^\varepsilon, v_h)$. Finally, the consistency of the diffusive part (and thus of the whole formulation) follows from [100, Proposition 2.9]. □

7.3.3 *A priori* bound

Let us assume that b and r satisfy

$$(r_0)^2(x) := r - \frac{1}{2}\nabla \cdot b(x) \geq r_{\min} > 0. \tag{7.23}$$

Note that this condition follows from (5.22). Following [65], it holds for all $w \in H^1(D, \mathcal{T}_h)$

$$\begin{aligned} d_{\mathrm{DG}}^{\varepsilon}(w,w) + t_{\mathrm{DG}}(w,w) + r_{\mathrm{DG}}(w,w) \; &= \sum_{K\in\mathcal{T}_h}\Big(|w|^2_{H^{1,\varepsilon}(K)} + ||r_0 w||^2_{L^2(K)} \\ &\quad + \frac{1}{2}||[w]||^2_{L^2(\partial_- K)} + \frac{1}{2}||[w]||^2_{L^2(\partial_+ K\cap\Gamma_h^0)}\Big) \\ &\quad + \sum_{e\in\Gamma_h}\frac{\alpha}{|e|}||[w]||^2_{L^2(e)}, \end{aligned} \tag{7.24}$$

with

$$|w|^2_{H^{1,\varepsilon}(K)} := \frac{1}{2}\int_K \nabla w^\top Q(x,\varepsilon)\nabla w dx.$$

The DG norm $\|\cdot\|_{\mathrm{DG}(\varepsilon)}$ for sufficiently small $\varepsilon > 0$ and $w \in H^1(D, \mathcal{T}_h)$ is defined as:

$$\begin{aligned} ||w||^2_{\mathrm{DG}(\varepsilon)} \; := & \sum_{K\in\mathcal{T}_h}\Big(|w|^2_{H^{1,\varepsilon}(K)} + ||r_0 w||^2_{L^2(K)} + \frac{1}{2}||[w]||^2_{L^2(\partial_- K)} + \frac{1}{2}||[w]||^2_{L^2(\partial_+ K\cap\Gamma_h^0)}\Big) \\ + & \sum_{e\in\Gamma_h}\frac{\alpha}{|e|}||[w]||^2_{L^2(e)}. \end{aligned} \tag{7.25}$$

Let us now consider the term $j_{\mathrm{DG}}^{\varepsilon}(\cdot,\cdot)$. Without loss of generality it can be assumed that $j_{\mathrm{DG}}^{\varepsilon}(w,w) \geq 0$ holds for $\varepsilon > 0$ sufficiently small. This situation can always be achieved after a change of variable. Therefore

$$a_{\mathrm{DG}}^{\varepsilon}(w,w) \geq ||w||^2_{\mathrm{DG}(\varepsilon)}. \tag{7.26}$$

Considering (7.21), it holds for all $w \in H^1(D, \mathcal{T}_h)$ and all sufficiently small $\varepsilon > 0$

$$\begin{aligned} j_{\mathrm{DG}}^{\varepsilon}(w,w) \; &\leq \left|\int_D\int_D (w(y) - w(x))k_\varepsilon(x, y-x)dy\, w(x)dx\right| \\ &\leq C(\varepsilon)||w||^2_{\mathrm{DG}(\varepsilon)}, \end{aligned} \tag{7.27}$$

and we conclude,

$$a_{\mathrm{DG}}^{\varepsilon}(w,w) \leq C(\varepsilon)||w||^2_{\mathrm{DG}(\varepsilon)}. \tag{7.28}$$

Remark 7.3.6. *From* (7.27)*, it is clear that it is not necessary to add an additional term in the definition of the norm* $||\cdot||_{DG(\varepsilon)}$ *to control* $j_{DG}(\cdot,\cdot)$ *once* $\varepsilon > 0$ *is fixed. However,* $C(\varepsilon) \to +\infty$ *in (7.28) as* $\varepsilon \to 0$*. See Section 7.3.5 for details on the case* $\varepsilon \equiv 0$*.*

To prove the *a priori* bound, we need the following result.

Lemma 7.3.7. *Let $w \in S^{p,0} \subset H^2(D, \mathcal{T}_h)$ and $K \in \mathcal{T}_h$, then there exists a constant $\mathcal{C} > 0$, independent of h, p and dependent on the shape regularity of $\mathcal{T}_h$, such that,*

$$\left|\left|\nabla_h w^\top Q(x,\varepsilon) n\right|\right|_{L^2(e)} \le \mathcal{C} p h^{-1/2} \left|\left|\sqrt{\nabla_h w^\top Q(x,\varepsilon)\nabla_h w}\right|\right|_{L^2(K)} \qquad \forall e \in \partial K.$$

Proof. The result follows from trace inequalities, i.e.,

$$\begin{aligned} \left|\left|\nabla_h w^\top Q(x,\varepsilon) n\right|\right|^2_{L^2(e)} &\le \mathcal{C} \sup_{x\in D} \{|Q(x,\varepsilon)|\} \int_e \nabla_h w^\top Q(x,\varepsilon)\nabla_h w \, dx \\ &\le \frac{\mathcal{C}p^2}{h_K} \sup_{x\in D} \{|Q(x,\varepsilon)|\} \int_K \nabla_h w^\top Q(x,\varepsilon)\nabla_h w \, dx, \end{aligned}$$

where the constant $\mathcal{C}$ is independent of h_K, i.e., the diameter of the element K, the polynomial degree p and $|e|$ (see for example [100, Section 2.1.3] and [65, Section 4.2]). □

Theorem 7.3.8. *Let u_h^ε be the solution of (7.13), then the following* a priori *bound holds:*

$$||u_h^\varepsilon(T)||^2_{L^2(D)} + \int_0^T ||u_h^\varepsilon(t)||^2_{DG(\varepsilon)} dt \le C \|f\|^2_{L^2((0,T);\, L^2(D))} + ||\Pi_p g||^2_{L^2(D)}.$$

Proof. Considering (7.26), it holds

$$\begin{aligned} & \frac{1}{2}\frac{d}{dt}||u_h^\varepsilon(t)||^2_{L^2(D)} + ||u_h^\varepsilon(t)||^2_{\mathrm{DG}(\varepsilon)} \\ \le\ & (\partial_t u_h^\varepsilon(t), u_h^\varepsilon(t)) + d^\varepsilon_{\mathrm{DG}}(u_h^\varepsilon(t), u_h^\varepsilon(t)) + t_{\mathrm{DG}}(u_h^\varepsilon(t), u_h^\varepsilon(t)) + j^\varepsilon_{\mathrm{DG}}(u_h^\varepsilon(t), u_h^\varepsilon(t)) \\ & + r_{\mathrm{DG}}(u_h^\varepsilon(t), u_h^\varepsilon(t)) \\ =\ & l_{\mathrm{DG}}(u_h^\varepsilon(t)) \le ||f(t)||_{L^2(D)} ||u_h^\varepsilon(t)||_{L^2(D)} \end{aligned}$$

where n is the exterior normal unit vector to ∂D. We obtain

$$\frac{1}{2}\frac{d}{dt}||u_h^\varepsilon(t)||^2_{L^2(D)} + ||u_h^\varepsilon(t)||^2_{\mathrm{DG}(\varepsilon)} \le \sqrt{\frac{1}{(r_0)^2}} ||f(t)||_{L^2(D)} ||u_h^\varepsilon(t)||_{\mathrm{DG}(\varepsilon)}.$$

Thus

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}||u_h^\varepsilon(t)||^2_{L^2(D)} + ||u_h^\varepsilon(t)||^2_{\mathrm{DG}(\varepsilon)} &\le C \|f\|_{L^2(D)} \|u_h^\varepsilon(t)\|_{\mathrm{DG}(\varepsilon)} \\ &\le C^2 \|f(t)\|^2_{L^2(D)} + \frac{1}{2}||u_h^\varepsilon(t)||^2_{\mathrm{DG}(\varepsilon)}, \end{aligned}$$

where $C^2 = \frac{1}{(r_0)^2}$. We notice that C is bounded independently of ε and of h, provided that $h > 0$ and $\varepsilon > 0$ are sufficiently small. Thus

$$\frac{d}{dt} \|u_h^\varepsilon(t)\|^2_{L^2(D)} + \|u_h^\varepsilon(t)\|^2_{\mathrm{DG}(\varepsilon)} \le 2C^2 \left(||f(t)||^2_{L^2(D)} \right)$$

and the claimed result is obtained integrating in time and setting $u_h^\varepsilon(0,\cdot) = \Pi_p g(\cdot)$. □

7.3.4 *A priori* error estimate

In order to estimate $||u - u_h^\varepsilon||$ in a suitable norm we apply the triangle inequality and estimate the terms separately:

$$||u - u_h^\varepsilon|| \leq \underbrace{||u - u^\varepsilon||}_{(a)} + \underbrace{||u^\varepsilon - u_h^\varepsilon||}_{(b)}.$$

The term (a) stems from the small jump approximation and can be estimated using Theorem 4.1.14, while the term (b) depends on the DG approximation. In order to prove an *a priori* error estimate for $||u^\varepsilon - u_h^\varepsilon||$, we need the Lemma 7.2.5. For ease of notation, we set $w(t) := w(t, \cdot)$ and define the DG norm $|||\cdot|||_{\mathrm{DG}(\varepsilon)}$ as follows:

$$|||w(t)|||^2_{\mathrm{DG}(\varepsilon)} = ||w(t)||^2_{L^2(D)} + \int_0^t ||w(s)||^2_{\mathrm{DG}(\varepsilon)} ds \quad \forall t \in (0, T). \tag{7.29}$$

We are now able to prove the following result.

Theorem 7.3.9. *Let u^ε and u_h^ε be the solutions of* (5.33) *and* (7.13)*, then* $\forall t \in [0, T]$

$$|||\left(u^\varepsilon - u_h^\varepsilon\right)(t)|||_{DG(\varepsilon)} \leq C \frac{h^{\min(p+1,s)-1}}{p^{s-\frac{3}{2}}} \Big(||u^\varepsilon(t)||_{H^s(D,\mathcal{T}_h)} + ||u^\varepsilon||_{H^1((0,t);\, H^s(D,\mathcal{T}_h))} \Big) \tag{7.30}$$

Proof. Since the scheme is consistent, the DG formulation (7.13) satisfies the orthogonality property

$$\forall t \in (0, T),\ \forall v \in S^{p,0} \quad \int_D \partial_t (u^\varepsilon - u_h^\varepsilon) v dx + a^\varepsilon_{\mathrm{DG}}(u^\varepsilon - u_h^\varepsilon, v) = 0.$$

Let us consider a suitable projection Π_p on the space of discontinuous piecewise polynomial functions. We require that Π_p is such that

$$\forall K \in \mathcal{T}_h \quad (\Pi_p v)\,|_K = \Pi_{(p,K)}(v|_K).$$

As in [65] we use for $\Pi_{(p,K)}$ the L^2-orthogonal projector, i.e., given $w \in L^2(D)$, $(w - \Pi_p w, v_h) = 0 \ \ \forall v_h \in S^{p,0}$. We set as in Theorem 7.2.9

$$\eta := u^\varepsilon - \Pi_p u^\varepsilon \text{ and } \xi := u_h^\varepsilon - \Pi_p u^\varepsilon,$$

where obviously $\xi \in S^{p,0}$ holds. Using the Galerkin orthogonality and the equality $u^\varepsilon - u_h^\varepsilon = \eta - \xi$, we obtain

$$\int_D \partial_t \xi\, v dx + a^\varepsilon_{\mathrm{DG}}(\xi, v) = \int_D \partial_t \eta v dx + a^\varepsilon_{\mathrm{DG}}(\eta, v) \ \ \forall v \in S^{p,0}.$$

Thus, setting $v = \xi$ and applying (7.26), we obtain

$$\frac{1}{2}\frac{d}{dt}||\xi||^2_{L^2(D)} + ||\xi||^2_{\mathrm{DG}(\varepsilon)} \leq \int_D \partial_t \xi\, \xi dx + a^\varepsilon_{\mathrm{DG}}(\xi, \xi) = \int_D \partial_t \eta\, \xi dx + a^\varepsilon_{\mathrm{DG}}(\eta, \xi). \tag{7.31}$$

Let us examine the terms in (7.31) in more detail. For any $\widehat{C_1} > 0$, it holds

$$\begin{aligned}\int_D \partial_t \eta\, \xi dx &\leq ||\partial_t \eta||_{L^2(D)} ||\xi||_{L^2(D)} \leq \frac{\widehat{C_1}}{2} ||\xi||^2_{L^2(D)} + \frac{1}{2\widehat{C_1}} ||\partial_t \eta||^2_{L^2(D)} \\ &\leq \frac{\widehat{C_1}}{2 r_{\min}} ||\xi||^2_{\mathrm{DG}(\varepsilon)} + \frac{1}{2\widehat{C_1}} ||\partial_t \eta||^2_{L^2(D)}.\end{aligned}$$

It follows from [65, Lemma 4.3] and Lemma 7.3.7 that for any $\widehat{C_2} > 0$

$$d^{\varepsilon}_{\mathrm{DG}}(\eta, \xi) \leq ||\xi||_{\mathrm{DG}(\varepsilon)} \sqrt{\delta_{\mathrm{DG}}(\eta)} \leq \frac{\widehat{C_2}}{2} ||\xi||^2_{\mathrm{DG}(\varepsilon)} + \frac{1}{2\widehat{C_2}} \delta_{\mathrm{DG}}(\eta)$$

with

$$\begin{aligned}\delta_{\mathrm{DG}}(\eta) &:= ||\sqrt{\alpha_J}[\eta]||^2_{L^2(\Gamma_h)} + \sum_{K \in \mathcal{T}} |\eta|^2_{H^{1,\varepsilon}(K)} \\ &\quad + \mathcal{C} q_{\varepsilon} p^2 h^{-1} ||[\eta]||^2_{L^2(\partial K)} + q_{\varepsilon}^2 \left|\left| \frac{1}{\sqrt{\alpha_J}} \nabla \eta \right|\right|^2_{L^2(\partial K)},\end{aligned}$$

where we recall that α_J is the penalization parameter, i.e., $\alpha_J|_e = \frac{\alpha}{|e|}$ $\forall\, e \in \Gamma_h$ and $q_{\varepsilon} = \sup_{x \in D} \left| \sqrt{\frac{1}{2} Q(x, \varepsilon)} \right|^2_2$ with $|\cdot|_2$ denoting the matrix norm subordinated to the l^2 vector norm on $\mathbb{R}^d$.
Using [65, Lemma 3.2], we obtain for any $\widehat{C_3} > 0$

$$t_{\mathrm{DG}}(\eta, \xi) + r_{\mathrm{DG}}(\eta, \xi) \leq ||\xi||_{\mathrm{DG}(\varepsilon)} \sqrt{\tau_{\mathrm{DG}}(\eta)} \leq \frac{\widehat{C_3}}{2} ||\xi||^2_{\mathrm{DG}(\varepsilon)} + \frac{1}{2\widehat{C_3}} \tau_{\mathrm{DG}}(\eta), \tag{7.32}$$

where

$$\tau_{\mathrm{DG}}(\eta) := \sum_{K \in \mathcal{T}} ||r_0 \eta||^2_{L^2(K)} + 2||\eta||^2_{L^2(\partial_+ K \cap \Gamma^0_h)} + 2||\eta_O||^2_{L^2(\partial_- K \setminus \Gamma^0_h)}.$$

For any $\widehat{C_4} > 0$, reasoning as in (7.27), it holds

$$\begin{aligned} j^{\varepsilon}_{\mathrm{DG}}(\eta, \xi) &\leq C(\varepsilon) \left(\frac{\widehat{C_4}}{2} ||\xi||^2_{L^2(D)} + \frac{1}{2\widehat{C_4}} ||\eta||^2_{L^2(D)} \right) \\ &\leq C(\varepsilon) \left(\frac{\widehat{C_4}}{2 r_{\min}} ||\xi||^2_{\mathrm{DG}(\varepsilon)} + \frac{1}{2 r_{\min} \widehat{C_4}} ||\eta||^2_{\mathrm{DG}(\varepsilon)} \right).\end{aligned}$$

Choosing positive constants $\widehat{C_1}$, $\widehat{C_2}$, $\widehat{C_3}$ and $\widehat{C_4}$ sufficiently small, i.e., such that

$$\frac{\widehat{C_1}}{2 r_{\min}} + \frac{\widehat{C_2}}{2} + \frac{\widehat{C_3}}{2} + \frac{C(\varepsilon)\widehat{C_4}}{2 r_{\min}} = \frac{1}{2}, \tag{7.33}$$

the following stability result holds

$$\frac{1}{2}\frac{d}{dt}||\xi||^2_{L^2(D)} + \frac{1}{2}||\xi||_{\mathrm{DG}(\varepsilon)} \le \underbrace{C\left(||\partial_t\eta||^2_{L^2(D)} + \delta_{\mathrm{DG}}(\eta) + \tau_{\mathrm{DG}}(\eta) + ||\eta||^2_{\mathrm{DG}(\varepsilon)}\right)}_{=:\widehat{\xi}[\eta]}, \quad (7.34)$$

with $C > 0$ independent of h and p, but depending on ε

$$C = \max\left(\frac{1}{2\widehat{C_1}}, \frac{1}{2\widehat{C_2}}, \frac{1}{2\widehat{C_3}}, \frac{C(\varepsilon)}{2r_{\min}\widehat{C_4}}\right). \quad (7.35)$$

Thus, integrating (7.34), since all the above constants are time-independent, we obtain

$$|||u^\varepsilon - u^\varepsilon_h|||^2_{\mathrm{DG}(\varepsilon)} \;\le\; |||\xi|||^2_{\mathrm{DG}(\varepsilon)} + |||\eta|||^2_{\mathrm{DG}(\varepsilon)} \le \int_0^t \widehat{\xi}[\eta](s)ds + |||\eta|||^2_{\mathrm{DG}(\varepsilon)}.$$

Therefore the interpolation error estimates in Lemma 7.2.5 give the claimed result, since $\forall t \in (0,T)$ it holds

$$\begin{aligned}
& \int_0^t \widehat{\xi}[\eta](s)ds + |||\eta|||^2_{\mathrm{DG}(\varepsilon)} \\
\le\; & C\int_0^t \Big(\sum_{K\in\mathcal{T}_h}\left(1+\frac{p^2}{h}+\frac{1}{h}\right)\int_{\partial K}\eta^2(\tau,s)ds + \int_K \eta^2(\tau,x)dx \\
+\; & \int_K |\nabla\eta(\tau,x)|^2\,dx + h\int_{\partial K}|\nabla\eta(\tau,s)|^2\,ds + \int_K|\partial_t\eta(\tau,x)|^2\,dx\Big)d\tau + ||\eta(t)||^2_{L^2(D)} \\
\le\; & C\frac{h^{2\min(p+1,s)-2}}{p^{2s-3}}\left(\int_0^t ||u^\varepsilon(\tau)||^2_{H^s(D,\mathcal{T}_h)} + ||\partial_t u^\varepsilon(\tau)||^2_{H^s(D,\mathcal{T}_h)}d\tau + ||\eta(t)||^2_{L^2(D)}\right).
\end{aligned}$$

□

Remark 7.3.10. *Inequality (7.32) depends on Assumption 7.2.1. In [65, Remark 3.13], the authors comment on this condition: if it is violated, then the presented analysis yields an error bound that is still optimal with respect to h but is p-suboptimal. A possible remedy is to supplement the definition of the scheme with a streamline-diffusion term, this restores the hp optimality. However, numerical results suggest that the DG scheme is hp-optimal even if Assumption 7.2.1 is violated and no streamline-diffusion stabilization term is added. Assumption 7.2.1 has been removed in [54, Remark 5.9], replacing b by a suitable projection on the space of discontinuous piecewise polynomial functions.*

Remark 7.3.11. *We choose the stabilization parameter independent of p, i.e.,* $\alpha_J|_e = \alpha|e|^{-1}$ *for any* $e \in \Gamma_h$, *with* α *independent of h and p. From the above proof it is clear that setting* $\alpha_J|_e = \alpha p|e|^{-1}$ *does not affect the hp-convergence order of the error estimate.*

Remark 7.3.12. *Lemma 3.2.2 implies* $||k_\varepsilon||_\infty \lesssim \varepsilon^{-(2\overline{m}+d)}$ *for sufficiently small* ε, *and thus* $C(\varepsilon) \lesssim \varepsilon^{-2\overline{m}-d}$, *we obtain from condition (7.33)* $C_4 \gtrsim \varepsilon^{2\overline{m}+d}$. *Therefore constant C in (7.35) satisfies* $C \lesssim \left(\varepsilon^{2\overline{m}+d}\right)^{-2}$, *and thus the constant C in (7.30), i.e., in the* a priori *error estimate, satisfies* $C \lesssim \varepsilon^{-(2\overline{m}+d)}$ *as* $\varepsilon \downarrow 0$.

Remark 7.3.13. *The norm $||\cdot||_{DG(\varepsilon)}$, and thus the norm $|||\cdot|||_{DG(\varepsilon)}$, depend explicitly on ε. In fact, if $\varepsilon \to 0$ (and thus $Q(x,\varepsilon) \to 0$), the H^1-part of the norm $||\cdot||_{DG(\varepsilon)}$ tends to zero, and therefore the considered norm becomes weaker. Moreover, as stated in Remark 7.3.6, when $\varepsilon \to 0$ we lose control of the jump term.*
In fact, if we consider error estimates in the DG norm $||\cdot||_{DG}$ given as:

$$\begin{aligned} ||u||^2_{DG} \quad := \quad & \sum_{e\in\Gamma_h} \frac{\alpha}{|e|}||[u]||^2_{L^2(e)}, \\ & + \sum_{K\in\mathcal{T}_h} \left(|u|^2_{H^1(K)} + ||r_0 u||^2_{L^2(K)} + \frac{1}{2}||[u]||^2_{L^2(\partial_- K)} + \frac{1}{2}||[u]||^2_{L^2(\partial_+ K\cap\Gamma^0_h)} \right) \end{aligned} \tag{7.36}$$

and

$$|||u(t)|||^2_{DG} := ||u(t)||^2_{L^2(D)} + \int_0^t ||u(s)||^2_{DG}\, ds \quad \forall t \in [0,T],$$

and we assume that $\sigma^\varepsilon = \sup_{x\in D} \max_{1\le i,j\le d} \left|\sqrt{Q_{ij}(x,\varepsilon)}\right|$, then Theorem 7.3.9 implies

$$\begin{aligned} |||\left(u^\varepsilon - u^\varepsilon_h\right)(t)|||_{DG(\varepsilon)} \quad &\le \quad \frac{1}{\sigma^\varepsilon}|||\left(u^\varepsilon - u^\varepsilon_h\right)(t)|||_{DG(\varepsilon)} \\ &\le \quad \frac{C}{\sigma^\varepsilon}\frac{h^{\min(p+1,s)-1}}{p^{s-\frac{3}{2}}} \left(||u^\varepsilon(t)||_{H^s(D,\mathcal{T}_h)} + ||u^\varepsilon||_{H^1((0,t);\, H^s(D,\mathcal{T}_h))}\right). \end{aligned}$$

Remark 7.3.14. *The diffusion term in (7.13) has been discretized according to the so-called NIPG-DG method (see [100] and Section 7.3.1 for this terminology). The results stated in this section also hold for the SIPG method, i.e., setting $\beta = -1$ in the DG formulation of the diffusion term. In this case, (7.24) does not hold, since we have to deal with the additional term $-\int_{\Gamma_h}\{\nabla w^\top Q(\varepsilon)n\}[w]ds$. However, using the Cauchy-Schwarz inequality (see [100, Section 2.7.1]), we obtain a lower bound for this extra-term.*

7.3.5 Finite variation processes

To approximate problem (5.9)-(5.10) when $\overline{Y} \ge 1$, we have to consider that the integral operator $\mathcal{A}_J$ in (3.14) is only well-defined for Lipschitz u, because of the singularity of k in 0. Thus a Discontinuous Galerkin (DG) discretization is not directly applicable. However, due to the small jump regularization, cf. Theorem 4.1.14, we can consider a DG discretization of the regularized problem (5.33).
Note that for processes with finite variation, e.g. for admissible market models with vanishing diffusion component and $0 < \overline{Y} < 1$, the small jump regularization is not necessary to obtain a formulation which allows for the application of a DG discretization. In fact, the jump term $\int_{\mathbb{R}^d} (\phi(x+y) - \phi(x))\, k(y)\, dy$ is not pointwise well-defined for a discontinuous basis function ϕ. This is not necessary for the present algorithm, as a Galerkin formulation with the Dirichlet form of the process is applied and therefore existence of the integral in a weaker sense is sufficient.

Theorem 7.3.15. *Let $\mathcal{A}_{J-FV}$, as in* (3.15)*, be an operator of order* $2\mathbf{m}(x)$ *and let* $\overline{m} < 0.5$ *hold. Then the following estimate can be proved for* $\phi, \psi \in S^{p,0}$*:*

$$a_{\mathrm{J}}(\widetilde{\phi}, \widetilde{\psi}) = (\mathcal{A}_{\mathrm{J-FV}} \widetilde{\phi}, \widetilde{\psi}) < C \left\| \widetilde{\phi} \right\|_{H^{\mathbf{m}(x)}(\mathbb{R}^d)} \left\| \widetilde{\psi} \right\|_{H^{\mathbf{m}(x)}(\mathbb{R}^d)} < \infty.$$

Proof. This follows directly from the continuity of the bilinear form and the embedding $S^{p,0} \subset \widetilde{H}^{\mathbf{m}(x)}(D)$. □

So, the small jump regularization is not necessary when finite variation Lévy processes are considered. Note that this argument does not hold if a finite difference discretization is applied. In this case a pointwise definition of the jump term is necessary and a regularization has to be performed even for finite variation processes, cf. [33].

DG Formulation

The variational form (7.13) when $\varepsilon = 0$, i.e., when no small jump approximation is considered, reads

$$(\partial_t u_h(x), v_h(x)) + a_{\mathrm{DGFV}}(u_h(x), v_h(x)) = l_{\mathrm{DGFV}}(v_h(x)), \tag{7.37}$$

$$u_h(0,x) = \Pi_p g(x), \tag{7.38}$$

where for $v, w \in S^{p,0}$

$$a_{\mathrm{DGFV}}(w,v) := t_{\mathrm{DGFV}}(w,v) + r_{\mathrm{DGFV}}(w,v) + j_{\mathrm{DGFV}}(w,v), \tag{7.39}$$

$$t_{\mathrm{DGFV}}(w,v) := \sum_{K \in \mathcal{T}_h} \int_K b \cdot \nabla w\, v dx - \int_{\partial_- K} (b \cdot n_K)[w] v_I ds, \tag{7.40}$$

$$r_{\mathrm{DGFV}}(w,v) := \int_D r w\, v dx,$$

$$j_{\mathrm{DGFV}}(w,v) := -\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (\widetilde{w}(t,y) - \widetilde{w}(t,x)) k(x, y-x) dy\, \widetilde{v}(x) dx, \tag{7.41}$$

$$l_{\mathrm{DGFV}}(v) = \int_D f v_h dx. \tag{7.42}$$

Notice that (7.39), (7.40) and (7.41) correspond to (7.15), (7.18) and (7.21), respectively, when $\varepsilon = 0$.

A priori bound and error estimate

The above formulation is consistent, i.e., the following result holds.

Theorem 7.3.16. *If u is the solution of* (5.9)-(5.10) *and* $\overline{m} < 0.5$*, then it satisfies* (7.37)*.*

Proof. The proof follows along the lines of the proof of Theorem 7.3.5. □

Let us now assume that b and r satisfy

$$(r_0)^2(x) := r - \frac{1}{2}\nabla \cdot (b(x)) \geq r_{\min} > 0 \tag{7.43}$$

(see Condition (5.22)). Reasoning as in Section 7.3.3, we define the norm $||\cdot||_{\mathrm{DGFV}}$, for sufficiently smooth w,

$$||w||^2_{\mathrm{DGFV}} := ||r_0 w||^2_{L^2(D)} + ||w||^2_{\widetilde{H}^{\mathbf{m}(x)}(D)} + \frac{1}{2}\sum_{K\in\mathcal{T}_h} ||[w]||^2_{L^2(\partial_- K)} + ||[w]||^2_{L^2(\partial_+ K\cap\Gamma_h^0)}, \tag{7.44}$$

and we have $t_{\mathrm{DGFV}}(w,w) + r_{\mathrm{DGFV}}(w,w) = ||w||^2_{\mathrm{DGFV}} - \|v\|^2_{\widetilde{H}^{\mathbf{m}}(D)}$. Further, we assume $j_{\mathrm{DGFV}}(w,w) \geq C\,\|u\|^2_{\widetilde{H}^{\mathbf{m}(x)}(D)}$, then $a_{\mathrm{DGFV}}(w,w) \geq C||w||_{\mathrm{DGFV}}$ holds, for some $C>0$ and all w, i.e. the bilinear form $a_{\mathrm{DGFV}}(\cdot,\cdot)$ is coercive.

Remark 7.3.17. *Note that the norm $\|\cdot\|_{DG(\varepsilon)}$ is stronger than the norm $\|\cdot\|_{DGFV}$, for arbitrary $\varepsilon > 0$, as $H^1(D,\mathcal{T}_h) \subset \widetilde{H}^{\mathbf{m}(x)}(D)$, for $\overline{m} < 0.5$.*

The following *a priori* bound and error estimate hold.

Theorem 7.3.18. *Let u_h be the solution of (7.37), then*

$$||u_h(T,\cdot)||^2_{L^2(D)} + \int_0^T ||u_h(t,\cdot)||^2_{DGFV}dt \quad\leq\quad C\,\|f\|^2_{L^2((0,T);\,L^2(D))} + ||\Pi_p g||^2_{L^2(D)}.$$

Proof. We consider (7.37) and use the coercivity of $a_{\mathrm{DGFV}}(\cdot,\cdot)$. The result follows estimating $l_{\mathrm{DGFV}}(u_h(t))$ like $l_{\mathrm{DG}}(u_h^\varepsilon(t))$ in Theorem 7.3.8. □

Theorem 7.3.19. *Let us consider the DG norm $|||\cdot|||_{DGFV}$ given by*

$$|||w(t)|||^2_{DGFV} = ||w(t)||^2_{L^2(D)} + \int_0^t ||w(s)||^2_{DGFV}ds \quad \forall t\in(0,T), \tag{7.45}$$

where, for ease of notation, we set $w(t) := w(t,\cdot)$, and u and u_h be the solution of (5.24) *and (7.37), respectively, then*

$$|||\,(u-u_h)\,(t)|||_{DGFV} \leq C\frac{h^{\min(p+1,s)-1/2}}{p^{s-1/2}}\Big(||u(t)||_{H^s(D,\mathcal{T}_h)} + ||u||_{H^1((0,t);\,H^s(D,\mathcal{T}_h))}\Big) \tag{7.46}$$

$\forall t \in [0,T]$.

Proof. Reasoning as in the proof of Theorem 7.3.9, we obtain

$$|||u-u_h|||^2_{\mathrm{DGFV}} \quad\leq\quad \int_0^t |\partial_t\eta|^2 + \tau_{\mathrm{DG}}(\eta) + ||\eta||^2_{\mathrm{DGFV}}ds$$

and the result follows using Lemma 7.2.5. □

Remark 7.3.20. *In the case of vanishing reaction and transport terms, we obtain an analogous result to Theorem 7.3.19, considering the norm* $||\cdot||_{DGFV'} := ||\cdot||_{\widetilde{H}^{\mathbf{m}(x)}(D)}$*. The following estimate holds in this case*

$$||u(t)-u_h(t)||^2_{L^2(D)} + \int_0^t ||u(s)-u_h(s)||^2_{DGFV'} ds \leq C\frac{h^{\min(p+1,s)-\overline{m}}}{p^{s-\overline{m}}}$$
$$\times \left(|||u(t)||_{H^s(D,\mathcal{T}_h)} + \|u\|_{H^1((0,t);\, H^s(D,\mathcal{T}_h))} \right).$$

8 Computational scheme

We have seen in the previous chapters that certain types of weakly singular functions have to be integrated in order to obtain the stiffness matrices in the solution algorithm. Due to the singularity of the integrand standard Gauss integration is not applicable. Therefore, composite Gauss quadratures are discussed in this chapter. The combination of Gauss quadrature rules of different order on subdomains leads to exponential convergence of the quadrature rule in the number of quadrature points, if the subdomains and the orders are chosen adequately. Besides the time-discretization is addressed.

8.1 Numerical quadratures

We shall frequently write $x \lesssim y$ in this chapter to express that x is bounded by a constant multiple of y, uniformly with respect to all parameters on which x and y may depend. As seen in the previous chapters we need to compute matrix entries of the form:

$$\mathbf{A}_{(\mathbf{l}',\mathbf{k}'),(\mathbf{l},\mathbf{k})} := \int_{\mathbb{R}^d}\int_{D_R} \partial_1 \dots \partial_d \psi_{\mathbf{l},\mathbf{k}}(x+y)\psi_{\mathbf{l}',\mathbf{k}'}(x)\kappa(x,y)\mathrm{d}x\mathrm{d}y. \tag{8.1}$$

We consider the following class of function. The kernels we consider fall into this class due to Theorem 3.2.1 and Lemma 3.2.2.

Assumption 8.1.1. *Let* $f \in L^1([0,1]^d \times [0,1]^d)$. *There exist* $0 < \alpha < d$, $\alpha \notin \mathbb{N}$, $C, C_1 > 0$, $\delta \in (0,1)$, *such that for* $k, m \in \mathbb{N}_0$, $i = 1, \dots, d$, $j = 1, \dots, d$

$$\left|\partial_{\xi_i}^k \partial_{x_j}^m f(x,\xi)\right| \leq C_1 k! m! C^{k+m} \|\xi\|_\infty^{-\alpha} \xi_i^{-k} \xi_j^{-\delta}, \quad \forall \xi, x \in (0,1)^d. \tag{8.2}$$

We are now able to prove the exponential convergence in the number of quadrature points of a quadrature rule for the matrix entries $\mathbf{A}_{(\mathbf{l}',\mathbf{k}'),(\mathbf{l},\mathbf{k})}$. We denote the Gauss-Legendre integration rule on $[0,1]$ by $Q_g^{[0,1]} f = \sum_{j=1}^{g} \omega_{g,j} f(\xi_{g,j})$ and set $I^{[0,1]} f := \int_0^1 f(x)dx$. The following error estimate for $f \in C^{2g}([0,1])$ using Stirling's formula:

$$\left|E_g^{[0,1]} f\right| := \left|I^{[0,1]} f - Q_g^{[0,1]} f\right| \leq C \frac{2^{-4g}}{(2g)!} \max_{\xi \in [0,1]} \left|f^{(2g)}(\xi)\right|,$$

for some constant $C > 0$, can be obtained. In the multidimensional case we have a similar error estimate for $f \in C^{2g}([0,1]^d)$ using [114, Lemma 4.1].

$$\left| E_g^{[0,1]^d} f \right| \lesssim \frac{2^{-4g}}{(2g)!} \sum_{i=1}^{d} \max_{\xi \in [0,1]^d} \left| \partial_i^{(2g)} f(\xi) \right|. \tag{8.3}$$

We are now able to define a composite quadrature rule as in [108]. Let a geometric partition on $[0,1]$ be given by $0 < \mu^n < \mu^{n-1} < \ldots < \mu < 1$, for $n \in \mathbb{N}$, $\mu \in (0,1)$. We denote the subdomains by $\Lambda_j := [\mu^{n+1-j}, \mu^{n-j}]$, with $j = 1, \ldots, n$ and $\Lambda_0 = [0, \mu^n]$. Given a linear degree vector $q \in \mathbb{N}^d$ and $q_j = \lceil \lambda j \rceil$ with slope $\lambda > 0$, we use on each subdomain Λ_j, $j = 1, \ldots, n$ a Gauss quadrature with degree q_j and no quadrature points in Λ_0. The composite Gauss quadrature rule is defined by

$$Q_\mu^{n,q} f = \sum_{j=1}^{n} Q_{q_j}^{\Lambda_j} f$$

and its exponential convergence can be proven.

Theorem 8.1.2. *Let f satisfy Assumption 8.1.1. Consider*

$$\mu \in (0,1) \quad \textit{such that } w = \frac{C(1-\mu)}{4\mu} < 1, \tag{8.4}$$

and linear degree vectors $(q^{(1)}, \ldots, q^{(d)})$, $q^{(j)} = (q_1^{(j)}, \ldots, q_n^{(j)})$,

$$q_j^{(i)} = \lceil \lambda^{(i)} j \rceil, \quad \textit{with slopes } \lambda^{(i)} > \frac{(1 - \frac{\alpha}{d}) \ln \mu}{2 \ln w}. \tag{8.5}$$

Then we obtain for any fixed $x \in [0,1]^d$

$$\left| I^{[0,1]^d} f(x) - Q_\mu^{n,(q^1,\ldots,q^d)} f(x) \right| \leq C e^{-\gamma \sqrt[2d]{N}}, \tag{8.6}$$

where N denotes the number of Gauss points and $C > 0$ some constant.

Proof. The proof can be found in [114, Theorem 4.6]. $\square$

We use composite Gauss quadrature rules in the ξ-variable and standard Gauss quadratures in the x-variable.

Theorem 8.1.3. *We consider the following quadrature rule for a function f satisfying Assumption 8.1.1:*

$$Q = Q_\mu^{n,(q^1,\ldots,q^d)} \otimes Q_g$$

and prove the following estimate for the error defined as

$$E[f] = \left| I^{[0,1]^d} f - Qf \right| \leq C e^{\sqrt[2d]{N}},$$

for $g = \lceil \sqrt[8d]{N} \rceil$ and some constant $C > 0$.

Proof.

$$
\begin{aligned}
E[f] &= I_\xi^{[0,1]^d} \otimes \left(I_x^{[0,1]^d} - Q_g\right) f + \left(I_\xi^{[0,1]^d} - Q_\mu^{n,(q^1,\ldots,q^d)}\right) \otimes Q_g f && (8.7)\\
&\lesssim \int_{[0,1]^d} \frac{2^{-4g}}{(2g)!} \max_{x\in[0,1]^d} \left| \frac{\partial^{(2g)} f}{\partial x}(x,\xi) \right| d\xi + e^{-\gamma n} \sum_{k=1}^{d} \sum_{j=1}^{g} \omega_{g,j,k} && (8.8)\\
&\lesssim e^{-4g^2} + e^{-\gamma n}. && (8.9)
\end{aligned}
$$

The number of quadrature points N can be bounded by $N \lesssim n^{2d}$. Therefore we obtain exponential convergence in the number of quadrature points choosing $g = \lceil \sqrt[8d]{N} \rceil$. $\square$

Remark 8.1.4. *Even under weaker assumptions on the integration kernel, we can obtain exponentially converging integration schemes using [28].*

8.2 Time discretization

In order to obtain a fully discrete approximation (in space and time) to the parabolic problem (5.9), we have to discretize the semi-discrete formulation in time. This can be done for example via discontinuous Galerkin time stepping as in [107] or by the θ-scheme. We will present the preconditioning for the θ-scheme for the semidiscrete formulation in Section 7.1 in more detail. Multilevel preconditioning in the implementation of DG-time stepping is analogous to [97, Section 6.3.2].
At each time step, we need to solve a linear system

$$(\mathbf{M} + \theta\Delta t\mathbf{A})\underline{u}_L^{m+1} = (\mathbf{M} - (1-\theta)\Delta t\mathbf{A})\underline{u}_L^m,$$

at each time step $m = 0, \ldots, M-1$, with $\underline{u}_L^0 = \underline{u}_{L,0}$, where $\underline{u}_L^m$ denotes the coefficient vector of $u_L(t_m, \cdot)$, $\mathbf{M}$ the mass matrix and $\mathbf{A}$ the stiffness matrix in the corresponding basis. For the iterative solution of these systems we use multilevel preconditioning obtained through the wavelet norm equivalences. We obtain for $u \in V_{L+1}$ with coefficient vector $\underline{u}$

$$|\underline{u}|^2 \lesssim (\underline{u}, \mathbf{M}\underline{u}) \lesssim |\underline{u}|^2,$$

due to (6.18). We denote by $\mathbf{D}_A$ the diagonal matrix with entries $2^{2\underline{m}_{\lambda_1}^1 l_1} + \ldots + 2^{2\underline{m}_{\lambda_d}^d l_d}$. Then we obtain, from (6.18) and the well-posedness:

$$(\underline{u}, \mathbf{D}_A\underline{u}) \lesssim (\underline{u}, \mathbf{A}\underline{u}) \lesssim (\underline{u}, \mathbf{D}_A\underline{u}).$$

Thus, we have

$$(\underline{u}, \mathbf{D}\underline{u}) \lesssim (\underline{u}, \mathbf{B}\underline{u}) \lesssim (\underline{u}, \mathbf{D}\underline{u}),$$

with $\mathbf{D} = \mathbf{I} + \theta\Delta t\mathbf{D}_A$ and $\mathbf{B} = \mathbf{M} + \theta\Delta t\mathbf{A}$. Finally we obtain for $\widehat{\underline{u}} = \mathbf{D}^{1/2}\underline{u}$:

$$|\widehat{\underline{u}}|^2 \lesssim \left(\widehat{\underline{u}}, \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}\widehat{\underline{u}}\right) \lesssim |\widehat{\underline{u}}|^2.$$

Therefore, we can iteratively solve the linear system $\widehat{\mathbf{B}}\widehat{\underline{u}} = \widehat{b}$ with GMRES in a number of steps that is independent of the level index L, where $\widehat{\mathbf{B}} = \mathbf{D}^{-1/2}\mathbf{B}\mathbf{D}^{-1/2}$ and $\widehat{\underline{b}} = \mathbf{D}^{-1/2}\widehat{\underline{b}}$. An analysis of time-stepping schemes for DG-discretizations as in Section 7.3 is given in [100, Chapter 3.4].

9 Well-posedness of singular time-inhomogeneous PIDEs

This chapter aims at the analysis of certain type of degenerate linear parabolic integro-differential equations. The arising PIDE reads as follows:

$$\begin{aligned}\partial_t u - t^\gamma \mathcal{A} u &= f \text{ on } I \times D, &(9.1)\\ u(0) &= g, &(9.2)\end{aligned}$$

where $\mathcal{A}$ denotes a possibly non-selfadjoint operator, as discussed in Section 3.3, g the sufficiently smooth initial data, γ a constant with $\gamma \in (-1, 1)$, $I = (0, T)$ and a Lipschitz domain $D \subset \mathbb{R}^d$ for $d \geq 1$. Note that negative exponents γ lead to an explosion at $t = 0$, while positive γ lead to a degeneracy of the diffusion coefficients. Therefore, the initial condition has to be imposed in an appropriate sense.
Such equations arise naturally in the context of option pricing under certain self-similar processes as considered by [25]. The processes defined in [25] extend the class of Lévy processes introducing a possibly degenerate time-dependence in the coefficients, it was shown empirically that such processes admit a good fit to option prices over several maturities for various strike prices.
We consider a weak space-time formulation in the sense of [2, 111], as a possible singularity or degeneracy of the diffusion coefficients impedes the application of classical parabolic theory, cf. [1, 91]. The use of appropriate wavelet bases in the space-time domain leads to Riesz bases for the ansatz and test spaces, cf. [13, 111].
We derive three weak space-time formulations for degenerate parabolic equations such as (9.1)-(9.2) in arbitrary space dimensions. The main difference between the three formulations described lies in the enforcement of the initial condition. First we assume $\mathcal{A}$ to be self-adjoint and obtain well-posedness results as well as a priori estimates based on an eigenfunction expansion of the operator $\mathcal{A}$. Then we describe an alternative approach based on a singular change of the temporal variable.

9.1 Essential initial condition

We consider the following degenerate parabolic problem for sufficiently smooth $u(t, x)$:

$$\begin{aligned}\partial_t u - t^\gamma \mathcal{A} u &= f \text{ on } I \times D, &(9.3)\\ u(0) &= g, &(9.4)\end{aligned}$$

for $\gamma = 2H - 1$, $H \in (0,1)$, a bounded Lipschitz domain $D \subset \mathbb{R}^d$ and a finite time interval $I := (0,T)$, $T > 0$. The operator $\mathcal{A} \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ is assumed to be self-adjoint and satisfy

$$a(u,u) \geq C \|u\|_{\mathcal{V}}^2 ,$$

where $C > 0$ and the associated bilinear form $a(\cdot,\cdot)$ reads

$$a(u,v) : \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \qquad a(u,v) := (-\mathcal{A}u, v)_{\mathcal{V}^*,\mathcal{V}} , \quad \forall u,v \in \mathcal{V} \tag{9.5}$$

and $\mathcal{V} := H_0^1(D)$, $\mathcal{V}^* := H^{-1}(D)$. To state the variational formulation of (9.3)-(9.4) we define the following spaces

$$\begin{aligned}
\mathcal{X} &:= H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*) \cap L^2_{t^{\gamma/2}}(I;\mathcal{V}) && (9.6)\\
&\cong \left(H^1_{t^{-\gamma/2}}(I) \otimes \mathcal{V}^*\right) \cap \left(L^2_{t^{\gamma/2}}(I) \otimes \mathcal{V}\right) ,\\
\mathcal{Y} &:= L^2_{t^{\gamma/2}}(I;\mathcal{V}) \cong L^2_{t^{\gamma/2}}(I) \otimes \mathcal{V}, && (9.7)\\
\mathcal{X}_{(0} &:= \{w \in \mathcal{X} : w(0,\cdot) = 0 \text{ in } \mathcal{V}^*\}, && (9.8)\\
\mathcal{X}_{0)} &:= \{w \in \mathcal{X} : w(T,\cdot) = 0 \text{ in } \mathcal{V}^*\}, && (9.9)
\end{aligned}$$

$L^2_{t^{\gamma/2}}(I) := \overline{C^\infty(0,T)}^{\|\cdot\|_{L^2_{t^{\gamma/2}}(I)}}$ and $H^1_{t^{\gamma/2}}(I) := \overline{C^\infty(0,T)}^{\|\cdot\|_{H^1_{t^{\gamma/2}}(I)}}$. We refer to [94, Chapter II.4] for proofs of the isomorphisms given in (9.6) and (9.7) for $\mathcal{X}$ and $\mathcal{Y}$. The weighted norms are defined by

$$\|u\|^2_{L^2_{t^{\gamma/2}}(I)} := \int_I u^2 t^\gamma \, dt, \qquad \|u\|^2_{H^1_{t^{\gamma/2}}(I)} := \int_I u^2 t^\gamma \, dt + \int_I \dot{u}^2 t^\gamma \, dt.$$

We use the following norms on $\mathcal{X}$ and $\mathcal{Y}$

$$\begin{aligned}
\|u\|^2_{\mathcal{X}} &:= \|\dot{u}\|^2_{L^2_{t^{-\gamma/2}}(I;\mathcal{V}^*)} + \|u\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} ,\\
\|u\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} &:= \int_I t^\gamma \|u\|^2_{\mathcal{V}} \, dt, \quad \|u\|^2_{L^2_{t^{-\gamma/2}}(I;\mathcal{V}^*)} := \int_I t^{-\gamma} \|u\|^2_{\mathcal{V}^*} \, dt,\\
\|u\|^2_{H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*)} &:= \int_I t^{-\gamma} \|u\|^2_{\mathcal{V}^*} + t^{-\gamma} \|\dot{u}\|^2_{\mathcal{V}^*} \, dt,
\end{aligned}$$

where we denote by $\|\cdot\|_{\mathcal{V}}$ the energy norm on $\mathcal{V}$, i.e.,

$$\|u\|^2_{\mathcal{V}} = a(u,u).$$

The family of eigenfunctions of the self-adjoint operator $\mathcal{A}$ in (9.3) is denote by $(\phi_\lambda)_{\lambda\in\sigma}$ for $\sigma \subset \mathbb{R}_+$ and is assumed to form an orthonormal basis of $L^2(D)$. Therefore, any element in $v \in \mathcal{V}$ admits the following representation $v = \sum_{\lambda\in\sigma} v_\lambda \phi_\lambda$, $v_\lambda \in \mathbb{R}$, $\lambda \in \sigma$. Due to Parseval's theorem we obtain $\|v\|^2_{L^2(D)} = \sum_{\lambda\in\sigma} |v_\lambda|^2$. Besides,

$$\|v\|^2_{\mathcal{V}} = a(u,u) = \sum_{\lambda\in\sigma} \lambda \, |v_\lambda|^2$$

holds. Any element $h \in \mathcal{V}^*$ admits the following representation

$$h = \sum_{\lambda \in \sigma} h_\lambda \phi_\lambda, \text{ where } h_\lambda := (h, \phi_\lambda)_{\mathcal{V}^*,\mathcal{V}}$$

and it easy to see that

$$\|h\|^2_{\mathcal{V}^*} = \sum_{\lambda \in \sigma} \lambda^{-1} |h_\lambda|^2 .$$

We now show the following result.

Theorem 9.1.1. *For every $f \in \mathcal{Y}^*$, $g = 0$ (9.3)-(9.4) admits a unique solution $u \in \mathcal{X}_{(0}$ and there holds the a priori error estimate*

$$\|u\|_{\mathcal{X}} \leq \sqrt{2}\, \|f\|_{\mathcal{Y}^*} ,$$

where $\|f\|_{\mathcal{Y}^*} = \|f\|_{L^2_{t^{-\gamma/2}}(I;\mathcal{V}^*)}$.

The proof follows from the inf-sup condition (9.10), the surjectivity (9.11) and the continuity (9.12) of the corresponding bilinear form using, eg. [5] or [20, III, Theorem 4.3]. These properties are proved in the following. We need the spaces $X := \left\{u \in L^2_{t^{\gamma/2}}(I) \cap H^1_{t^{-\gamma/2}}(I) : u(0) = 0\right\}$ and $Y := L^2_{t^{\gamma/2}}(I)$ and remark that $H^1_{t^{-\gamma/2}}(I) \subset C_0(I)$ holds, this follows as in Lemma 9.2.2. For $u \in X$ we define the seminorm:

$$\|u\|_{X^\lambda} := \left\| \lambda^{-\frac{1}{2}} t^{-\gamma/2} \dot{u} + \lambda^{\frac{1}{2}} t^{\gamma/2} u \right\|_{L^2(I)} .$$

Lemma 9.1.2. *For $\lambda > 0$ and $u \in X$, define the norm $\|u\|_\lambda$ by*

$$\|u\|^2_\lambda := \lambda^{-1} \left\| t^{-\gamma/2} \dot{u} \right\|^2_{L^2(I)} + \lambda \left\| t^{\gamma/2} u \right\|^2_{L^2(I)} .$$

Then, for all $u \in X$ holds:

$$\|u\|_\lambda \leq \|u\|_{X^\lambda} \leq \sqrt{2}\, \|u\|_\lambda .$$

Proof. Let $u \in X$, then

$$\begin{aligned} \|u\|^2_{X^\lambda} &= \lambda^{-1} \left\| t^{-\gamma/2} \dot{u} \right\|^2_{L^2(I)} + \lambda \left\| t^{\gamma/2} u \right\|^2_{L^2(I)} + 2 \int_I u \dot{u}\, dt \\ &= \|u\|^2_\lambda + |u(T)|^2 \geq \|u\|^2_\lambda . \end{aligned}$$

Further,

$$\begin{aligned} 2 \left| \int_I u \dot{u}\, dt \right| &\leq 2\lambda^{1/2} \left\| t^{\gamma/2} u \right\|_{L^2(I)} \lambda^{-1/2} \left\| t^{-\gamma/2} \dot{u} \right\|_{L^2(I)} \\ &\leq \lambda \left\| t^{\gamma/2} u \right\|^2_{L^2(I)} + \lambda^{-1} \left\| t^{-\gamma/2} \dot{u} \right\|^2_{L^2(I)} \end{aligned}$$

and therefore $\|u\|^2_{X^\lambda} \leq 2\, \|u\|^2_\lambda$. □

Lemma 9.1.3. *We have*

$$\inf_{0\neq u\in\mathcal{X}_{(0}} \sup_{0\neq v\in\mathcal{Y}} \frac{B(u,v)}{\|u\|_{\mathcal{X}}\|v\|_{\mathcal{Y}}} \geq \frac{1}{\sqrt{2}}, \tag{9.10}$$

$$\forall\, 0\neq v\in\mathcal{Y}: \sup_{u\in\mathcal{X}_{(0}} B(u,v) > 0 \tag{9.11}$$

and

$$\sup_{0\neq u\in\mathcal{X}_{(0},0\neq v\in\mathcal{Y}} \frac{|B(u,v)|}{\|u\|_{\mathcal{X}}\|v\|_{\mathcal{Y}}} \leq \sqrt{2}, \tag{9.12}$$

where

$$B(u,v) := \int_0^T \Big((v(t),\dot{u}(t))_{\mathcal{V},\mathcal{V}^*} + t^\gamma a(u(t),v(t)) \Big)\, dt, \tag{9.13}$$

for $u\in\mathcal{X}_{(0}$, $v\in\mathcal{Y}$ *and* $a(\cdot,\cdot)$ *as in* (9.5).

Proof. Let $u\in\mathcal{X}$. Then $u=\sum_{\lambda\in\sigma}u_\lambda(t)\phi_\lambda$, $v\in\mathcal{Y}$, $v=\sum_{\lambda\in\sigma}v_\lambda(t)\phi_\lambda$, where ϕ_λ are the eigenfunctions of the self-adjoint operator $\mathcal{A}$. Since the family of functions $(\phi_\lambda)_{\lambda\in\sigma}$ is assumed to form an orthonormal basis of $L^2(D)$ and $u_\lambda(t)\in L^2_{t^{\gamma/2}}(I)\cap H^1_{t^{-\gamma/2}}(I)$, $v_\lambda\in L^2_{t^{\gamma/2}}(I)$,

$$\begin{aligned} B(u,v) &= \int_0^T \Big((v(t),\dot{u}(t))_{\mathcal{V},\mathcal{V}^*} + t^\gamma a(u(t),v(t)) \Big)\, dt \\ &= \sum_{\lambda\in\sigma}\int_0^T \lambda^{1/2}v_\lambda(t)t^{\gamma/2}\Big(\lambda^{-1/2}t^{-\gamma/2}\dot{u}_\lambda(t)+\lambda^{1/2}t^{\gamma/2}u_\lambda(t)\Big)\, dt. \end{aligned}$$

Therefore,

$$\begin{aligned} |B(u,v)| &\leq \left(\sum_{\lambda\in\sigma}\lambda\int_0^T t^\gamma\,|v_\lambda(t)|^2\,dt\right)^{1/2} \\ &\quad\times\left(\sum_{\lambda\in\sigma}\int_0^T \left|\lambda^{-1/2}t^{-\gamma/2}\dot{u}(t)_\lambda+\lambda^{1/2}t^{\gamma/2}u_\lambda(t)\right|^2 dt\right)^{1/2} \\ &= \|v\|_{L^2_{t^{\gamma/2}}(I;\mathcal{V})}\left(\sum_{\lambda\in\sigma}\|u_\lambda(t)\|^2_{X^\lambda}\right)^{1/2} \\ &\leq \|v\|_{L^2_{t^{\gamma/2}}(I;\mathcal{V})}\sqrt{2}\left(\sum_{\lambda\in\sigma}\|u_\lambda(t)\|^2_\lambda\right)^{1/2} = \sqrt{2}\,\|u\|_{\mathcal{X}}\,\|v\|_{\mathcal{Y}}\,. \end{aligned}$$

This implies (9.12). Next, given $u\in\mathcal{X}_{(0}$, we define $\mathcal{Y}\ni v_u=\sum_{\lambda\in\sigma}\phi_\lambda v_\lambda(t)$ by

$$v_\lambda(t)=\lambda^{-1}t^{-\gamma}\dot{u}_\lambda(t)+u_\lambda(t),$$

then

$$\begin{aligned}
\|v_u\|_{\mathcal{Y}}^2 &= \sum_{\lambda\in\sigma}\lambda\int_0^T t^\gamma\left(\lambda^{-1}t^{-\gamma}\dot{u}_\lambda(t)+u_\lambda(t)\right)^2\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T\left(\lambda^{-1/2}\dot{u}_\lambda(t)t^{-\gamma/2}+\lambda^{1/2}u_\lambda(t)t^{\gamma/2}\right)^2\,dt \\
&= \sum_{\lambda\in\sigma}\|u_\lambda(t)\|_{X^\lambda}^2\le 2\,\|u\|_{\mathcal{X}}^2\,. \qquad (9.14)
\end{aligned}$$

$$\begin{aligned}
B(u,v_u) &= \int_0^T (v_u(t),\dot{u}(t))_{\mathcal{V}^*,\mathcal{V}}+t^\gamma a(u,v_u)\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T\left(\lambda^{-1}t^{-\gamma}\dot{u}_\lambda(t)+u_\lambda(t)\dot{u}_\lambda(t)\right)+\lambda t^\gamma\left(\lambda^{-1}t^{-\gamma}\dot{u}_\lambda(t)+u_\lambda(t)u_\lambda(t)\right)\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T\left(\lambda^{-1}t^{-\gamma}\left|\dot{u}_\lambda(t)\right|^2+\frac{d}{dt}\left|u_\lambda(t)\right|^2+\lambda t^\gamma\left|u_\lambda(t)\right|^2\,dt\right) \\
&= \|u\|_{\mathcal{X}}^2+\|u(T)\|_{L^2(D)}^2-\|u(0)\|_{L^2(D)}^2\,.
\end{aligned}$$

This implies (9.10) using (9.14). Let now $v(t)=\sum_{\lambda\in\sigma}v_\lambda(t)\phi_\lambda$ be given, we define $u_v(t)=\sum_\lambda u_\lambda(t)\phi_\lambda$, where $(u_\lambda(t))_{\lambda\in\sigma}$ is given as solutions of the following sequence of initial value problems for $\lambda\in\sigma$:

$$\lambda^{-1}t^{-\gamma}\dot{u}_\lambda(t)+u_\lambda(t)=v_\lambda(t)\text{ for }t\in(0,T),\quad u_\lambda(0)=0.$$

In the following it will be shown that $v\in\mathcal{Y}$ implies $u_v\in\mathcal{X}$. We have

$$\begin{aligned}
\|v\|_{\mathcal{Y}}^2 &= \sum_{\lambda\in\sigma}\int_0^T t^\gamma\lambda\left|v_\lambda(t)\right|^2\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T\lambda\left|\lambda^{-1/2}t^{-\gamma}\dot{u}_\lambda(t)+\lambda^{1/2}u_\lambda(t)\right|^2 \\
&= \sum_{\lambda\in\sigma}\|u_\lambda(t)\|_{X^\lambda}^2\ge\sum_{\lambda\in\sigma}\|u_\lambda(t)\|_\lambda^2=\|u_v\|_{\mathcal{X}}^2\,.
\end{aligned}$$

We are now able to prove statement (9.11).

$$\begin{aligned}
B(u_v,v) &= \int_0^T (v(t),\dot{u}_v(t))+t^\gamma a(u_v(t),v(t))\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T v_\lambda(t)\dot{u}_\lambda(t)+\lambda u_\lambda(t)v_\lambda(t)t^\gamma\,dt \\
&= \sum_{\lambda\in\sigma}\int_0^T\lambda t^\gamma\left|v_\lambda(t)\right|^2\,dt=\|v\|_{\mathcal{Y}}^2>0.
\end{aligned}$$

□

Remark 9.1.4. *For every $f \in \mathcal{Y}^*$ the problem* (9.3)-(9.4) *with $g = 0$ admits a unique solution $u \in \mathcal{X}_{(0}$ satisfying*

$$B(u, v) = (f, v)_{\mathcal{Y}^*,\mathcal{Y}}\,, \quad \forall v \in \mathcal{Y}.$$

With $\mathcal{X}$ and $\mathcal{Y}$ as in (9.6)-(9.7) *and $B(\cdot, \cdot)$ as in Lemma 9.1.3, we have the a priori estimate*

$$\|u\|^2_{\mathcal{X}} \leq 2\,\|f\|^2_{\mathcal{Y}^*}\,.$$

The existence of a unique weak solution for non-homogeneous initial data follows via the following change of variable $\widetilde{v}(t, x) = v(t, x) - g$, for $g \in \mathcal{V}$. The function $\widetilde{v}(t, x)$ satisfies the same PDE as $v(t, x)$ with homogeneous initial conditions and a different right hand side.

9.2 Natural initial condition

As we assume non-homogeneous initial conditions, we can either transform the problem into a homogeneous setting as described in Section 9.1 or impose natural conditions as follows:

$$\int_0^T (v(t), \dot{u}(t))_{L^2(D)} dt = -\int_0^T (\dot{v}(t), u(t))_{L^2(D)} dt + (u(t), v(t))_{L^2(D)}|_0^T\,,$$

for $v, u \in C^\infty(I; \mathcal{V})$. For $u(0) \neq 0$ we impose homogeneous Dirichlet conditions on v, i.e. we require $v(T) = 0$. The variational formulation with weak enforcement of the initial conditions then reads: given $f \in \mathcal{X}^*_{0)}$, $g \in \mathcal{V}$, find $u \in \mathcal{Y}$:

$$B^*(u, v) = \int_0^T (v(t), f(t))_{\mathcal{V},\mathcal{V}^*}\, dt + (g, v(0))_{\mathcal{V},\mathcal{V}^*}\,,\ \forall v \in \mathcal{X}_{0)}, \tag{9.15}$$

where $B^*(\cdot, \cdot)$ is given by

$$B^*(u, v) := \int_0^T \Big(-(u(t), \dot{v}(t))_{\mathcal{V},\mathcal{V}^*} + t^\gamma a(u(t), v(t)) \Big)\, dt, \tag{9.16}$$

for $u \in \mathcal{Y}$, $v \in \mathcal{X}_{0)}$, with $a(\cdot, \cdot)$ given in (9.5). We define the functional l^* on $\mathcal{X}$ as follows:

$$l^*(v) := \int_0^T (v(t), f(t))_{\mathcal{V},\mathcal{V}^*}\, dt + (g, v(0))_{\mathcal{V},\mathcal{V}^*}\,.$$

Lemma 9.2.1. *For $f \in \mathcal{X}^*_{0)}$ and for $g \in \mathcal{V}$, l^* is a continuous, linear functional on $\mathcal{X}_{0)}$, i.e., there exists a $C > 0$ s.t.*

$$\forall v \in \mathcal{X}_{0)} : \quad |l^*(v)| \leq C \left(\|f\|_{\mathcal{X}^*_{0)}} + \|g\|_{\mathcal{V}} \right) \|v\|_{\mathcal{X}_{0)}}\,.$$

Proof. For $f \in \mathcal{X}_{0)}^*$ we have:

$$\left| \int_0^T (v(t), f(t))_{\mathcal{V},\mathcal{V}^*} \, dt \right| \le \|v\|_{\mathcal{X}_{0)}} \|f\|_{\mathcal{X}_{0)}^*} .$$

By the embedding given in (9.17) we obtain for $v \in \mathcal{X}_{0)}$

$$\|v(0)\|_{\mathcal{V}^*} \le \|v\|_{C^0(\overline{I},\mathcal{V}^*)} \le C \|v\|_{\mathcal{X}} ,$$

which implies,

$$|(v(0), g)|_{\mathcal{V}^*,\mathcal{V}} \le \|g\|_{\mathcal{V}} \|v(0)\|_{\mathcal{V}^*} \le C \|g\|_{\mathcal{V}} \|v\|_{\mathcal{X}} .$$

This implies the claimed result. □

We need the following embedding result.

Lemma 9.2.2. *For $\mathcal{X} := H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*) \cap L^2_{t^{\gamma/2}}(I;\mathcal{V})$ the following continuous embedding holds:*

$$\mathcal{X} \subset C^0(\overline{I}, D(\Lambda^{\frac{1}{2}-\frac{|\gamma|}{2}})), \tag{9.17}$$

where Λ denotes the operator $\Lambda = L^{1/2}$, as defined in [40, Chapter VIII, §3, Definition 8]. The operator Λ^θ denotes the holomorphic interpolant between $\mathcal{V}$ and $\mathcal{V}^*$.

Proof. Consider first $\gamma \in (-1, 0)$, then $L^2_{t^{\gamma/2}}(I;\mathcal{V}) \subset L^2_{t^{-\gamma/2}}(I;\mathcal{V})$. The claimed result follows from [41, Chapter XVIII, §1, Remark 6] for the space $H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*) \cap L^2_{t^{-\gamma/2}}(I;\mathcal{V})$. Let now $\gamma \in (0, 1)$. Then $H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*) \subset H^1_{t^{\gamma/2}}(I;\mathcal{V}^*)$, therefore we can again apply [41, Chapter XVIII, §1, Remark 6] and conclude. □

Remark 9.2.3. *(i) The space $H^1_{t^{-\gamma/2}}(I;\mathcal{V}^*) \cap L^2_{t^{-\gamma/2}}(I;\mathcal{V})$, for $\gamma \in (0, 1)$, is continuously embedded in $C^0(\overline{I}, D(\Lambda^{\frac{1}{2}+\frac{\gamma}{2}}))$, cf.[41, Chapter XVIII, §1, Remark 6].*

(ii) The elementary embedding of $\mathcal{X}$ in $C^0(\overline{I}, \mathcal{V}^*)$ can be shown as follows, cf. [57, Proposition 1.1],

$$\int_0^T \|v(t)\|_{\mathcal{V}^*} \, dt \le \left(\int_0^T \|v(t)\|^2_{\mathcal{V}^*} \, t^{-\gamma} dt \right)^{1/2} \left(\int_0^T t^{\gamma} dt \right)^{1/2} .$$

Therefore the mapping $K : u \to u'$, $K : \mathcal{X} \to L^1_{loc}(I;\mathcal{V}^*)$ is continuous. This implies that v is absolutely continuous on $\overline{I}$ with values in $\mathcal{V}^*$. Note that this does not imply the continuity of the embedding.

(iii) We obtain an analogous result for the weight function $(T - t)^\gamma$ instead of t^γ.

Theorem 9.2.4. *Let $B^*(\cdot,\cdot)$ be given as in (9.16) and $\mathcal{X}$, $\mathcal{Y}$ as in (9.6)-(9.7). Then the following estimates hold*

$$\inf_{0\neq u\in\mathcal{Y}} \sup_{0\neq v\in\mathcal{X}_{0)}} \frac{B^*(u,v)}{\|u\|_{\mathcal{Y}}\,\|v\|_{\mathcal{X}_{0)}}} \geq \frac{1}{\sqrt{2}},$$
$$\forall 0\neq v\in\mathcal{X}_{0)} : \quad \sup_{u\in\mathcal{Y}} B^*(u,v) > 0,$$
$$\sup_{0\neq v\in\mathcal{X}_{0)},0\neq u\in\mathcal{Y}} \frac{|B^*(u,v)|}{\|u\|_{\mathcal{Y}}\,\|v\|_{\mathcal{X}}} < \infty.$$

Proof. The proof is analogous to the proof of Lemma 9.1.3. □

Corollary 9.2.5. *For every $g\in\mathcal{V}$ and $f\in\mathcal{X}^*_{0)}$, there exists a unique weak solution $u\in\mathcal{Y}$ in the sense that u satisfies (9.15).*

Remark 9.2.6. *Note that for this formulation smoothness of the initial data is required, i.e. $g\in\mathcal{V}$. This is stronger than in the standard parabolic setting, as in this situation $g\in L^2(D)$ is sufficient in order to prove well-posedness of the corresponding weak formulation. This stronger condition stems from the fact that in the setup only the continuous embedding $\mathcal{X}\subset C^0(\overline{I},\Lambda^{\frac{1}{2}-\frac{|\gamma|}{2}})$ can be proved, while in the standard parabolic case $\big(L^2(I;\mathcal{V})\cap H^1(I;\mathcal{V}^*)\big)\subset C^0(\overline{I},L^2(D))$ holds.*

Remark 9.2.7. *Alternatively, the following formulation with natural initial conditions could also be considered. Find $w\in\mathcal{X}$ such that*

$$\begin{aligned} B^\dagger(w,v) &= f^\dagger(v) \text{ for all } v:=(v_1,v_2)\in\mathcal{Y}\times\mathcal{V}, \text{ where} \qquad (9.18)\\ B^\dagger(w,v) &= \int_0^T \Big((\dot w(t),v_1(t))_{\mathcal{V}^*,\mathcal{V}} + t^\gamma a(w(t),v_1(t))\Big)\,dt + (w(0),v_2)_{\mathcal{V}^*,\mathcal{V}}\,,\\ f^\dagger(v) &= \int_0^T (v_1(t),f(t))_{\mathcal{V},\mathcal{V}^*}\,dt + (g,v_2)_{\mathcal{V}^*,\mathcal{V}}\,. \end{aligned}$$

The well-posedness of (9.18) follows as in Lemma 9.1.3. The advantage of formulation (9.18) is the absence of any boundary conditions in the temporal domain, therefore the bases presented in the next chapter can be used for the discretization without any additional considerations.

9.3 Transformation approach

In this section we describe a different approach to the proof of well-posedness using a singular change of variable. We consider a slightly more general setup than in the previous sections. The formulation reads: find a sufficiently smooth function $u(t,x)$ such that

$$\partial_t u - t^\gamma \mathcal{A}(t)u = f \text{ on } I\times D, \qquad (9.19)$$

$$u(0) = g \tag{9.20}$$

holds in an appropriate sense to be specified below, where $t \to \mathcal{A}(t) \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ is measurable on I and $I = [0,1]$ for simplicity. Besides we assume that $a(t; u, v) = (-\mathcal{A}(t)u, v)$ satisfies for some constants $C_1, C_2 > 0$ and $C_3 \geq 0$, $\forall u, v \in \mathcal{V}$

$$\begin{aligned} a(t; u, u) &\geq C_1 \|u\|_{\mathcal{V}}^2 - C_2 \|u\|_{L^2(D)}^2, \\ |a(t; u, v)| &\leq C_3 \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}. \end{aligned}$$

We can apply a change of variable to remove the degeneracy in time in (9.19), cf. [3]. The transformation reads as follows:

$$\widetilde{t} = W(t) := \int_0^t \omega(s)\, ds, \tag{9.21}$$

where $\omega(t) = t^\gamma$. The problem (9.19)-(9.20) in new coordinates reads

$$\partial_{\widetilde{t}} v - \widetilde{\mathcal{A}}(\widetilde{t})v = (q(\widetilde{t}))^{-\gamma} \widetilde{f} \text{ on } (0,1) \times D, \tag{9.22}$$

$$v(0) = g, \tag{9.23}$$

where $v(W(t)) = u(t)$, $\widetilde{\mathcal{A}}(W(t)) = \mathcal{A}(t)$, $q(\widetilde{t}) = W^{-1}(\widetilde{t})$ and $\widetilde{f}(W(t)) = f(t)$. Therefore, $\widetilde{a}(t; u, v) = (\widetilde{\mathcal{A}}u, v)_{\mathcal{V}^*, \mathcal{V}}$ satisfies $\forall u, v \in \mathcal{V}$ and C_i, $i = 1, \ldots, 3$ as above

$$\widetilde{a}(t; u, u) \geq C_1 \|u\|_{\mathcal{V}}^2 - C_2 \|u\|_{L^2(D)}^2, \tag{9.24}$$

$$|\widetilde{a}(t; u, v)| \leq C_3 \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}. \tag{9.25}$$

For $\widetilde{f}$ we have

$$\begin{aligned} \|f\|_{L^2_{t^{-\gamma/2}}((0,1);\mathcal{V}^*)}^2 &= \int_0^1 \|f(t)\|_{\mathcal{V}^*}^2 t^{-\gamma}\, dt \\ &= \int_0^1 \left\|\widetilde{f}(\widetilde{t})\right\|_{\mathcal{V}^*}^2 (q(\widetilde{t}))^{-2\gamma}\, d\widetilde{t} = \left\|(q(\widetilde{t}))^{-\gamma} \widetilde{f}\right\|_{L^2((0,1);\mathcal{V}^*)}^2. \end{aligned} \tag{9.26}$$

The source term in (9.22) therefore satisfies $(q(\widetilde{t}))^{-\gamma} \widetilde{f} \in L^2((0,1); \mathcal{V}^*)$ if and only if $f \in L^2_{t^{-\gamma}}((0,1); \mathcal{V}^*)$. Using (9.24), (9.25) and (9.26) well-posedness can be shown for (9.22)-(9.23) using [111, Theorem 5.1].

Theorem 9.3.1. *Consider the operator $B : \mathcal{X}_1 \to \mathcal{Z}^*$ given by*

$$B(u)(v) = \int_0^1 \left(\partial_{\widetilde{t}} u - \widetilde{\mathcal{A}} u, v_1\right) d\widetilde{t} + (u(0), v_2)_{L^2(D)},$$

where $\mathcal{Z} = L^2((0,1); \mathcal{V}) \times L^2(D)$, $\mathcal{X}_1 = L^2((0,1); \mathcal{V}) \cap H^1((0,1); \mathcal{V}^*)$. Then $B : \mathcal{X}_1 \to \mathcal{Z}^*$ is an isomorphism.

Theorem 9.3.1 implies the existence of a unique solution of (9.19)-(9.20) in a weak space-time sense with $u \in \mathcal{X}$.

Remark 9.3.2. *The change of variable can be applied for a wide class of weight functions* $\omega(t)$ *in* (9.21).

As a corollary of Theorem 9.3.1 we obtain the well-posedness for pricing equations corresponding to admissible time-inhomogeneous market models.

Corollary 9.3.3. *Let X be an admissible time-inhomogeneous market model with generator $\mathcal{A}(t)$. Then the following problem is well-posed.*
Find $u \in L^2_{(T-t)^\gamma}(I;\mathcal{V}) \cap H^1_{(T-t)^{-\gamma}}(I;\mathcal{V}^*)$, $u(0)=0$ such that

$$\int_0^T ((v(t), \partial_t u(t))_{\mathcal{V},\mathcal{V}^*} - (T-t)^\gamma (\mathcal{A}(T-t)u, v)_{\mathcal{V}^*,\mathcal{V}})\, dt = \int_0^T (f(t), v(t))_{\mathcal{V}^*,\mathcal{V}}$$

holds for all $v \in L^2_{(T-t)^\gamma}(I;\mathcal{V})$ and $f \in L^2_{(T-t)^{-\gamma}}(I;\mathcal{V}^*)$.

A formulation with natural enforcement of the initial condition can also be considered.

Corollary 9.3.4. *Let X be an admissible time-inhomogeneous market model with infinitesimal generator $\mathcal{A}(t)$ and initial condition $g \in \mathcal{V}$. We consider the following formulation. Find $u \in L^2_{(T-t)^\gamma}(I;\mathcal{V}) \cap H^1_{(T-t)^{-\gamma}}(I;\mathcal{V}^*)$ such that*

$$\begin{aligned}
B^\dagger(w,v) &= f^\dagger(v) \text{ for all } v := (v_1,v_2) \in L^2_{(T-t)^\gamma}(I;\mathcal{V}) \times L^2(D), \text{ where} \qquad (9.27)\\
B^\dagger(w,v) &= \int_0^T \Big((\dot{w}(t), v_1(t))_{\mathcal{V}^*,\mathcal{V}} + (T-t)^\gamma (\mathcal{A}(T-t)w(t), v_1(t))_{\mathcal{V}^*,\mathcal{V}} \Big)\, dt \\
&\quad + (w(0), v_2)_{L^2(D)}\,, \\
f^\dagger(v) &= \int_0^T (v_1(t), f(t))_{\mathcal{V},\mathcal{V}^*}\, dt + (g, v_2)_{L^2(D)}\,.
\end{aligned}$$

The well-posedness follows from Theorem 9.3.1 using (9.21)*, we set $T=1$ for notational convenience. It remains to show that $q(\widetilde{t})^{-\gamma}\widetilde{f} \in L^2((0,1);\mathcal{V}^*)$ holds. This follows directly from* (9.26)*. Therefore we have the following result. The problem*

$$\begin{aligned}
\widetilde{B}^\dagger(u,v) &= \widetilde{f}^\dagger(v), \\
\widetilde{B}^\dagger(u,v) &= \int_0^1 \Big(\partial_{\widetilde{t}} u(\widetilde{t}) - \widetilde{\mathcal{A}} u(\widetilde{t}), v_1(\widetilde{t}) \Big)_{\mathcal{V}^*,\mathcal{V}}\, d\widetilde{t} + (u(0), v_2)_{L^2(D)}\,, \\
\widetilde{f}^\dagger(v) &= \int_0^1 \Big(v_1(\widetilde{t}), (q(\widetilde{t}))^{-\gamma} \widetilde{f}(\widetilde{t}) \Big)_{\mathcal{V},\mathcal{V}^*}\, d\widetilde{t} + (g, v_2)_{L^2(D)}\,,
\end{aligned}$$

has a unique solution $u \in \mathcal{X}_1$, where $\widetilde{B}^\dagger : \mathcal{X}_1 \times \mathcal{Z} \to \mathbb{R}$, $\widetilde{f}^\dagger \in \mathcal{Z}^*$, $\mathcal{Z} = L^2((0,1);\mathcal{V}) \times L^2(D)$ and $\mathcal{X}_1 = L^2((0,1);\mathcal{V}) \cap H^1((0,1);\mathcal{V}^*)$. Therefore the problem (9.27) is well-posed, i.e., there exists a unique function $u \in \mathcal{X}$ such that $B^\dagger(w,v) = f^\dagger(v)$ for all $v := (v_1,v_2) \in \mathcal{Y} \times L^2(D)$.

10 FE Discretization of time-inhomogeneous PIDEs

In this chapter we are concerned with the discretization of integro-differential equations of the type (9.3)-(9.4). The main difficulty in the time-discretization resides in the degeneracy of the coefficients of the PIDE. Two different approaches are presented. The first approach is CG discretization in the space-time domain. Certain basis functions are needed in this case and are introduced below. Optimality of the solution algorithm can be shown in an appropriate sense. The second approach is a DG discretization in time. In this case exponential convergence of the semidiscrete equation of the scheme can be proved for a judicious combination of $h-$ and $p-$refinement.

10.1 Discretization

For the space-time discretization of the degenerate parabolic PIDE, given by (9.3)-(9.4), we follow [111]. The use of tensor product Riesz bases on the space-time domain is crucial for the efficient discretization. We construct appropriate bases in the following and prove the necessary norm equivalences.

10.1.1 Wavelets

In the sequel we require the following properties of the wavelet functions to be used on our Galerkin discretization schemes, we assume without loss of generality $I = (0, 1)$ for the time interval and $D = (0, 1)^d$ for the physical domain. The use of a hypercube as the spatial domain enables us to construct the basis functions for the discretization of the physical space as tensor products of univariate basis functions. Besides, we could also use sparse tensor products to overcome the curse of dimension, cf. [44] for the elliptic case. Domains of this form arise naturally in the discretization of pricing equations due to localization. We now state the requirements for the temporal wavelet basis $\Theta = \{\theta_\lambda : \lambda \in \nabla_\Theta\}$, where ∇_Θ denotes the set of all wavelet indices. Apart from the requirements (i)-(iv) from Section 6.2, we need the following assumptions in order to obtain a Riesz basis for the weighted spaces. We assume the following norm equivalences, for all $0 \leq s \leq \kappa$ and a $\kappa \geq 1$

$$\|u\|_s^2 \sim \sum_{l=0}^{\infty} \sum_{k \in \nabla_l} 2^{2ls} \left| u_k^l \right|^2, \quad u_k^l = \left(\widetilde{\theta}_{k,l}, u \right), \quad u \in H^s(0,1),$$

where $\|\cdot\|_s$ denotes the $H^s(0,1)$-norm and by $x \sim y$ we denote $x \lesssim y$ and $y \lesssim x$. Further we require that the wavelets and the dual wavelets for the time domain belong to $W^{1,\infty}(0,1)$ and the boundary wavelets for the time discretization satisfy:

$$\begin{aligned}
\left|\theta_k^l(t)\right| &\leq C_\theta 2^{l/2}(2^l t)^\beta, && (10.1)\\
\left|(\theta_k^l)'(t)\right| &\leq C_\theta 2^{3l/2}(2^l t)^{\beta-1},\ t \in [0, 2^{-l}],\ \beta \in \mathbb{N}_0,\ k \in \nabla_l^L, && (10.2)\\
\left|\widetilde{\theta}_k^l(t)\right| &\leq C_\theta C_\theta 2^{l/2}(2^l t)^{\widetilde{\beta}}, && (10.3)\\
\left|(\widetilde{\theta}_k^l)'(t)\right| &\leq C_\theta 2^{3l/2}(2^l t)^{\widetilde{\beta}-1},\ t \in [0, 2^{-l}],\ \widetilde{\beta} \in \mathbb{N}_0,\ k \in \widetilde{\nabla}_l^L, && (10.4)
\end{aligned}$$

where $\gamma/2 + \beta > -\frac{1}{2}$ and $-\gamma/2 + \widetilde{\beta} > -\frac{1}{2}$ with γ as in (9.3). The sets ∇_l^L and $\widetilde{\nabla_l^L}$ are given as follows, $\nabla_l^L := \{k \in \nabla_l : 0 \in \text{supp } \theta_k^l\}$ and $\widetilde{\nabla}_l^L := \{k \in \widetilde{\nabla}_l : 0 \in \text{supp } \widetilde{\theta}_k^l\}$. The estimates (10.1)-(10.4) play a crucial role in the proof of the norm equivalences for the weighted spaces, cf. [13, Section 3]. We refer to [37] for explicit constructions.
The spatial basis is constructed as follows: we define the subspace V_L of $H_0^1(D)$, for $D = (0,1)^d$, as the full tensor product of d univariate approximation spaces, i.e. $V_L := \bigotimes_{1 \leq i \leq d} \mathcal{V}^{l_i}$, which can be written as

$$V_L = \quad \{\sigma_{\mathbf{l},\mathbf{k}} : 0 \leq l_i \leq L-1, k_i \in \nabla_{l_i}, i = 1, \ldots, d\},$$

with basis functions $\sigma_{\mathbf{l},\mathbf{k}} = \sigma_{l_1,k_1} \cdots \sigma_{l_d,k_d}$, $0 \leq l_i \leq L-1$, $k_i \in \nabla_{l_i}$, $i = 1, \ldots, d$, where ∇_{l_i} denotes the set of wavelet coefficients in the i-th coordinate on level l_i. We can write V_L in terms of increment spaces

$$V_L = \bigoplus_{-1 \leq l_i \leq L-1} \mathcal{W}^{l_1} \otimes \ldots \otimes \mathcal{W}^{l_d}.$$

We denote by $\Sigma = \{\sigma_\mu : \mu \in \nabla_\Sigma\} = \bigotimes_{i=1}^d \Sigma_i$, $\Sigma_i = \{\sigma_{\mu_i} : \mu_i \in \nabla_{\Sigma_i}\}$. The tensor product spatial basis satisfies the following assumptions, where ∇_Σ is the set of all wavelet multi-indices and ∇_{Σ_i} denotes the set of all wavelet indices in the i-th coordinate.

(i) Local support: the diameter of the support is proportional to the meshsize 2^{-l},

$$\text{diam supp } \sigma_{l,k} \sim 2^{-l}. \tag{10.5}$$

(ii) Continuity: the primal basis function are assumed to be elements in $C^{r_x}(0,1)$, with $r_x \leq p_x - 2$.

(iii) Piecewise polynomial of order p_x, where piecewise means that the singular support consists of a uniformly bounded number of points.

(iv) Vanishing moments: the primal basis functions $\sigma_{l,k}$ are assumed to satisfy vanishing moment conditions up to order for $p_x > 1$

$$(\sigma_{l,k}, x^{\alpha}) = 0 \ , \ \alpha = 0, \ldots, d = p_x, \ l \geq 0. \tag{10.6}$$

(v) Orthonormality in $L^2(0,1)$.

(vi) Riesz basis property in $L^2(0,1)$ and renormalized in $H_0^1(0,1)$ and $H^{-1}(0,1)$.

We refer to [45, Section 5] and [46, Sections 5-7] for explicit constructions of wavelets.

10.1.2 Continuous Galerkin discretization in time

Using the wavelet constructions of the previous section we are now able to obtain Riesz bases for the spaces $L^2_{t^{\gamma/2}}(0,1)$ and $H^1_{t^{\gamma/2}}(0,1)$

Theorem 10.1.1. *The norm $|||\cdot|||_{L^2_{t^{\gamma/2}}(0,1)}$ is given as*

$$|||u|||^2_{L^2_{t^{\gamma/2}}(0,1)} := \sum_{l=0}^{\infty} \sum_{k \in \nabla_l} (2^{-l}k)^{\gamma} \left| u_k^l \right|^2, \tag{10.7}$$

where $u \in L^2_{t^{\gamma/2}}(0,1)$ admits the unique representation

$$u = \sum_{l=0}^{\infty} \sum_{k \in \nabla_l} u_k^l \theta_k^l, u_k^l = \left(\widetilde{\theta}_{k,l}, u \right).$$

Then the following norm equivalence holds for all functions $u \in L^2_{t^{\gamma/2}}(0,1)$:

$$\|u\|^2_{L^2_{t^{\gamma/2}}(0,1)} \sim |||u\,|||^2_{L^2_{t^{\gamma/2}}(0,1)} \,. \tag{10.8}$$

Proof. The result follows from [13, Theorem 3.3] setting $\omega = t^{\gamma/2}$ and checking Assumption 3.1 and 3.2 in [13]. Assumption 3.1 refers to the singularity of ω and Assumption 3.2 to the behavior of the wavelets, i.e., (10.1)-(10.4). □

A similar result can be obtained for $H^1_{t^{\gamma/2}}(0,1)$ using the following theorem:

Theorem 10.1.2. *Let Θ be as above and let $u \in H^1_{t^{\gamma/2}}(0,1)$, then*

$$\left\| u' \right\|^2_{L^2_{t^{\gamma/2}}(0,1)} \sim \sum_{l=0}^{\infty} 2^{2l} \sum_{k \in \nabla_l} (2^{-l}k)^{\gamma} \left| u_k^l \right|^2.$$

Proof. See [13, Theorem 5.1]. □

Therefore Θ forms, after rescaling, a Riesz basis of $H^1_{t^{\gamma/2}}(0,1)$.

Remark 10.1.3. *Note that analogous results can be obtained for the weight function $w(t) = \prod_{j=1}^{k} (t_k - t)^{\gamma_j}$.*

10.1.3 Space-time discretization

We are now able to construct a Riesz basis for the spaces $\mathcal{X}$ and $\mathcal{Y}$ in the case of a bounded spatial domain. The spaces have the following tensor product structure:

$$\mathcal{X} = (L^2_{t^{\gamma/2}}(I) \otimes \mathcal{V}) \cap (H^1_{t^{-\gamma/2}}(I) \otimes \mathcal{V}^*) \text{ and } \mathcal{Y} = L^2_{t^{\gamma/2}} \otimes \mathcal{V},$$

where $\mathcal{V} = H^1_0(D)$. Let Σ and Θ be given as above, then we obtain from [56, Proposition 1 and 2] that the collection $\Theta \otimes \Sigma$ normalized in $\mathcal{X}$, i.e.,

$$\left\{ (t,x) \to \frac{\theta_\lambda(t)\sigma_\mu(x)}{\sqrt{\|\sigma_\mu\|^2_{\mathcal{V}} + \|\theta_\lambda\|^2_{H^1_{t^{-\gamma/2}}(I)} \|\sigma_\mu\|^2_{\mathcal{V}^*}}} : (\lambda,\mu) \in \nabla_{\mathcal{X}} := \nabla_\Theta \times \nabla_\Sigma \right\}$$

is a Riesz basis for $\mathcal{X}$, denoted by $[\Theta \otimes \Sigma]_{\mathcal{X}}$ and that $\Theta \otimes \Sigma$ normalized in $\mathcal{Y}$, i.e.,

$$\left\{ (t,x) \to \frac{\theta_\lambda(t)\sigma_\mu(x)}{\|\sigma_\mu\|_{\mathcal{V}}} : (\lambda,\mu) \in \nabla_{\mathcal{X}} \right\}$$

is a Riesz basis for $\mathcal{Y}$, denoted by $[\Theta \otimes \Sigma]_{\mathcal{Y}}$.

10.2 Optimality

We are interested in optimality of the approximation of the solution process of the bi-infinite linear system, which arises from the discretization of (9.3)-(9.4) using the bases as described in the previous section. We derive estimates for the work requiered to solve the arising linear systems, under the assumption that the best N-term approximation of the solution vector $\mathbf{u}$ converges with a certain rate s. This class of elements in $l^2(\nabla_{\mathcal{X}})$ is formalized in the following definition.

Definition 10.2.1. *For $s > 0$ the approximation class $\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))$ is defined as follows:*

$$\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}})) := \{\mathbf{v} \in l^2(\nabla_{\mathcal{X}}) : \|\mathbf{v}\|_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))} < \infty\},$$

where $\|\mathbf{v}\|_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))} := \sup_{\varepsilon>0} \left(\varepsilon \times [\min\{N \in \mathbb{N}_0 : \|\mathbf{v} - \mathbf{v}_N\|_{l^2(\nabla_{\mathcal{X}})} \leq \varepsilon\}]^s \right)$ and $\mathbf{v}_N$ denotes the best N-term approximation of $\mathbf{v}$.

Let $s > 0$ be such that $\mathbf{u} \in \mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))$. In order to be able to bound the complexity of an iterative solution method for the bi-infinite system $\mathbf{Bu} = \mathbf{f}$, with appropriate $\mathbf{B}$ and $\mathbf{f}$, one needs a suitable bound on the complexity of an approximate matrix-vector product in terms of the prescribed tolerance. We formalize this in the notion of s^*-admissibility.

Definition 10.2.2. $\mathbf{B} \in \mathcal{L}(l^2(\nabla_{\mathcal{X}}), l^2(\nabla_{\mathcal{Y}}))$ *is s^*-admissible if there exists a routine which yields, for any $\varepsilon > 0$ and any finitely supported $\mathbf{w} \in l^2(\nabla_{\mathcal{X}})$, a finitely supported $\mathbf{z} \in l^2(\nabla_{\mathcal{Y}})$ with $\|\mathbf{Bw} - \mathbf{z}\| < \varepsilon$. Further, for any $\overline{s} \in (0, s^*)$, there exists an admissibility constant $a_{\mathbf{B},\overline{s}}$ such that $\#supp\mathbf{z} \leq a_{\mathbf{B},\overline{s}}\varepsilon^{-1/\overline{s}} \|\mathbf{w}\|^{1/\overline{s}}_{\mathcal{A}^{1/\overline{s}}_\infty(l^2(\nabla\mathcal{X}))}$ and the number of arithmetic operations and storage locations used by the call of the routine is bounded by some absolute multiple of*

$$a_{\mathbf{B},\overline{s}}\varepsilon^{-1/\overline{s}} \|\mathbf{w}\|^{1/\overline{s}}_{\mathcal{A}^{1/\overline{s}}_\infty(l^2(\nabla\mathcal{X}))} + \#supp\mathbf{w} + 1.$$

Next we introduce the concept of s^*-computability.

Definition 10.2.3. *The mapping $\mathbf{B} \in \mathcal{L}(l^2(\nabla_{\mathcal{X}}), l^2(\nabla_{\mathcal{Y}}))$ is s^*-computable if, for each $N \in \mathbb{N}$ there exists a $\mathbf{B}_N \in \mathcal{L}(l^2(\nabla_{\mathcal{X}}), l^2(\nabla_{\mathcal{Y}}))$ having in each column at most N nonzero entries whose joint computation takes an absolute multiple of N operations, such that the computability constants*

$$c_{\mathbf{B},\overline{s}} := \sup_{N\in\mathbb{N}} \|\mathbf{B} - \mathbf{B}_N\|^{1/\overline{s}}_{l^2(\nabla_{\mathcal{X}})\to l^2(\nabla_{\mathcal{Y}})}$$

are finite for any $\overline{s} \in (0, s^*)$.

In the following we assume that for $f \in \mathcal{Y}$ and any $\varepsilon > 0$ we can compute $\mathbf{f}_\varepsilon \in l^2(\nabla_{\mathcal{Y}})$ with

$$\|\mathbf{f} - \mathbf{f}_\varepsilon\|_{l^2(\nabla_{\mathcal{Y}})} \leq \varepsilon \text{ and } \#\text{supp } \mathbf{f}_\varepsilon \lesssim \min\{N : \|\mathbf{f} - \mathbf{f}_N\| \leq \varepsilon\},$$

with the number of arithmetic operations and storage locations used by the computation of $\mathbf{f}_\varepsilon$ bounded by some absolute multiple of $\#\text{supp}\mathbf{f}_\varepsilon + 1$. The following theorem links the two concepts of s^*-admissibility and s^*-computability, cf. [111, Theorem 4.10].

Theorem 10.2.4. *An s^*-computable $\mathbf{B}$ is s^*-admissible.*

We use the following result from [111, Corollary 4.6].

Corollary 10.2.5. *If $\mathbf{B} \in \mathcal{L}(l^2(\nabla_{\mathcal{X}}), l^2(\nabla_{\mathcal{Y}}))$ and $\mathbf{C} \in \mathcal{L}(l^2(\nabla_{\mathcal{Y}}), l^2(\nabla_{\mathcal{Z}}))$, then $\mathbf{CB} \in \mathcal{L}(l^2(\nabla_{\mathcal{X}}), l^2(\nabla_{\mathcal{Z}}))$*

The adaptive wavelet methods from [30] and [31] can be shown to be optimal for an s^*-admissible $\mathbf{B}$ and $\mathbf{u} \in \mathcal{A}^{1/\overline{s}}_\infty(l^2(\nabla\mathcal{X}))$.

Theorem 10.2.6. *Consider the bi-infinite system $\mathbf{Bu} = \mathbf{f}$ and let $\mathbf{B}$ be s^*-admissible, then for any $\varepsilon > 0$, both adaptive wavelet methods from [30, 31] produce an approximation $\mathbf{u}_\varepsilon$ to $\mathbf{u}$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{l^2(\nabla_{\mathcal{X}})} \leq \varepsilon$. If $\mathbf{u} \in \mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))$, then $\#supp\mathbf{u}_\varepsilon \lesssim$*

$\varepsilon^{-1/s} \|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))}$ and if, moreover, $s < s^*$, then the number of arithmetic operations and storage locations required by a call of either of these adaptive wavelet solvers with tolerance ε is bounded by some multiple of

$$\varepsilon^{-1/s}(1 + a_{\mathbf{B},s}) \|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))} + 1.$$

The multiples depend only on s when it tends to 0 or ∞, and on $\|\mathbf{B}\|$ and $\left\|\mathbf{B}^{-1}\right\|$ when they tend to infinity.

The following proposition is very useful, as the coefficients in the PDE (9.3)-(9.4) separate, i.e., using appropriate bases for the discretization leads to linear systems that possess a tensor product structure, cf. [111, Proposition 8.1].

Proposition 10.2.7. *For some $s^* > 0$, let $\mathbf{C}$, $\mathbf{D}$ be s^*-computable. Then*

(a) $\mathbf{C} \otimes \mathbf{D}$ is s^*-computable with computability constant satisfying, for $0 < \overline{s} < \tilde{s} < s^*$, $c_{\mathbf{C}\otimes\mathbf{D},\overline{s}} \lesssim (c_{\mathbf{C},\tilde{s}} c_{\mathbf{D},\tilde{s}})^{\tilde{s}/\overline{s}}$ and

(b) for any $\varepsilon \in (0, s^*)$, $\mathbf{C} \otimes \mathbf{D}$ is $(s^* - \varepsilon)$-computable, with computability constant $c_{\mathbf{C}\otimes\mathbf{D},\overline{s}}$ satisfying, for $0 < \overline{s} < s^* - \varepsilon < \tilde{s} < s^*$, $c_{\mathbf{C}\otimes\mathbf{D},\overline{s}} \lesssim \max(c_{\mathbf{C},\tilde{s}}) \max(c_{\mathbf{D},\tilde{s}})$.

Let $[\Theta \otimes \Sigma]_{\mathcal{X}}$ and $[\Theta \otimes \Sigma]_{\mathcal{Y}}$ be the Riesz bases of $\mathcal{X}$ and $\mathcal{Y}$ defined in Section 10.1.3, further let $\dot{\Theta} := \{\dot{\theta}_\lambda, \lambda \in \nabla_\Theta\}$. Denoting by $\|\Sigma\|_{\mathcal{V}}$ the diagonal matrix with entries σ_μ, $\mu \in \nabla_\Sigma$ and by $[\Sigma]_{\mathcal{V}}$ the Riesz basis of V consisting of the collection Σ normalized in V, similarly for other spaces and collections, we obtain the following representation of the bi-infinite system arising from the bilinear form $B(\cdot,\cdot)$ as in Lemma 9.1.3

$$\begin{aligned}
\mathbf{B} &:= B([\Theta \otimes \Sigma]_{\mathcal{X}}, [\Theta \otimes \Sigma]_{\mathcal{Y}}) \qquad (10.9)\\
&= \left[{}_{L^2_{t^{-\gamma/2}}(I)} \left(\Theta', \Theta\right)_{L^2_{t^{\gamma/2}}(I)} \otimes (\Sigma, \Sigma)_{L^2(D)} + \int_I t^\gamma a(\Theta \otimes \Sigma, \Theta \otimes \Sigma)\, dt \right] \\
&\quad \times \left(Id_t \otimes \|\Sigma\|^{-1}_{\mathcal{V}} \right) \|\Theta \otimes \Sigma\|^{-1}_{\mathcal{X}} \\
&= \left[{}_{L^2_{t^{-\gamma/2}}(I)} \left([\Theta']_{H^1_{t^{-\gamma/2}}(I)}, \Theta \right)_{L^2_{t^{\gamma/2}}(I)} \otimes (\Sigma, \Sigma)_{L^2(D)} \right] \left(\|\Theta\|_{H^1_{t^{-\gamma/2}}(I)} \otimes \|\Sigma\|_{\mathcal{V}} \right) \\
&\quad \times \|\Theta \otimes \Sigma\|^{-1}_{\mathcal{X}} + \int_I t^\gamma a(\Theta \otimes [\Sigma]_{\mathcal{V}}, \Theta \otimes [\Sigma]_{\mathcal{V}})\, dt \,(Id_t \otimes \|\Sigma\|_{\mathcal{V}}) \|\Theta \otimes \Sigma\|^{-1}_{\mathcal{X}}.
\end{aligned}$$

The load vector reads:

$$\mathbf{f} := \int_I (f, \Theta \otimes [\Sigma]_{\mathcal{V}})_{\mathcal{V}^*,\mathcal{V}}\, dt. \qquad (10.10)$$

We remark that the solution algorithms of [30] and [31] are only applicable to symmetric system matrices $\mathbf{B}$, we therefore consider the normal equations

$$\mathbf{B}^*\mathbf{B}\mathbf{u} = \mathbf{B}^*\mathbf{f} \qquad (10.11)$$

instead, cf. [111, Section 4]. We now show the s^*-computability of $\mathbf{B}$ and $\mathbf{B}^*$. The term $([\dot{\Theta}]_{H^1_{t^{-\gamma/2}}(I)}, \Theta)_{L^2(I)}$ is considered first. The ∞-computability of the bi-infinite matrix and its adjoint follows as in [111, Section 8.2] using the properties of the temporal basis. Next we consider $([\Sigma]_{\mathcal{V}^*}, [\Sigma]_{\mathcal{V}})_{L^2(D)}$. The ∞-computability follows from [111, Section 8.3]. We now consider the s^*-computability of $\int_I t^\gamma a(\Theta \otimes [\Sigma]_{\mathcal{V}}, \Theta \otimes [\Sigma]_{\mathcal{V}})dt$. Due to the properties of the bilinear form, we get:

$$\int_I t^\gamma a(\Theta \otimes [\Sigma]_{\mathcal{V}}, \Theta \otimes [\Sigma]_{\mathcal{V}})dt = (\Theta, \Theta)_{L^2_{t^{\gamma/2}}(I)} \otimes a([\Sigma]_{\mathcal{V}}, [\Sigma]_{\mathcal{V}}).$$

Therefore, it suffices to investigate the s^*-computability of both factors. The ∞-computability of $(\Theta, \Theta)_{L^2_{t^{\gamma/2}}(I)}$ follows from [13, Theorem 3.1] as in [111, Section 8.3]. For $a([\Sigma]_{\mathcal{V}}, [\Sigma]_{\mathcal{V}})$ we can deduce from [110] that it is s^*-computable with $s^* = p_x + 1$. We arrive at the following theorem.

Theorem 10.2.8. *Consider the weak form of the parabolic problem* (9.3) *on* $\mathcal{X}$, $\mathcal{Y}$ *as in* (9.6)-(9.7) *with bilinear form* $B(\cdot, \cdot)$ *as in* (9.13) *and the right hand side* $\int_I \langle f, \cdot \rangle$ *with* f *as* (9.3). *Its representation using space-time wavelets as in Section 10.1.3 with appropriate boundary conditions reads* $\mathbf{Bu} = \mathbf{f}$ *with* $\mathbf{B}$ *as in* (10.9) *and* $\mathbf{f}$ *as in* (10.10). *Then for any* $\varepsilon > 0$, *the adaptive wavelet methods from [30] and [31] applied to the normal equations* (10.11) *produce an approximation* $\mathbf{u}_\varepsilon$ *with*

$$\|\mathbf{u} - \mathbf{u}_\varepsilon\|_{l^2(\nabla_{\mathcal{X}})} \leq \varepsilon.$$

If for some $s > 0$, $\mathbf{u} \in \mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))$, *then* $\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s} \|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))}$. *The constant only depends on* s *when it tends to* 0 *or* ∞. *If for arbitrary* $s^* > 0$ *it holds that* $s < s^*$, *then the number of operations and storage locations required by one call of the space-time adaptive algorithm with tolerance* $\varepsilon > 0$ *is bounded by some multiple of*

$$\varepsilon^{-1/s} d^2 \|\mathbf{u}\|^{1/s}_{\mathcal{A}^s_\infty(l^2(\nabla_{\mathcal{X}}))} + 1,$$

where this multiple is uniformly bounded in d *and depends only on* $s \downarrow 0$ *and* $s \to \infty$.

Remark 10.2.9. *The complexity estimates in Theorems 10.2.6-10.2.8 apply if any entry in any vector that is generated inside the routine used in the Theorems can be stored in or fetched from memory in* $\mathcal{O}(1)$ *operations. This assumption is valid if an unlimited amount of memory is available, where each element can be accessed in* $\mathcal{O}(1)$ *operations, as this is not the case an additional log-term seems a priori unavoidable in the complexity estimate. We refer to [44, Section 6] for a detailed discussion of this issue.*

Remark 10.2.10. *Instead of applying the methods of [30] and [31] to the normal equations as in Theorem 10.2.8, we could use a GMRES-scheme applied to the original linear system. The author is not aware of theoretical results on such an approach.*

10.3 Discontinuous Galerkin discretization in time

In this section we describe an alternative approach to the time-discretization. A DG-scheme with geometric refinement of the grid towards the initial condition is used, while the polynomial is increased away from the singularity. The judicious combination of h- and p-refinement enables us to prove exponential convergence of the timestepping scheme in an appropriate sense.

Definition 10.3.1. *Let $I = (0,1)$. For a function $u \in L^2(I;\mathcal{V})$ which is continuous at $t = 1$ we define $\Pi^r u \in \mathcal{P}^r(I;\mathcal{V})$, $r \geq 1$, via the $r+1$ conditions*

$$\int_I (\Pi^r u - u, q)_{\mathcal{H}}\, dt = 0, \quad \forall q \in \mathcal{P}^{r-1}(I;\mathcal{V}) \tag{10.12}$$

and

$$\Pi^r u(+1) = u(+1) \in \mathcal{V}. \tag{10.13}$$

For $r = 0$ we use only (10.13) *to define Π^r, $\mathcal{H} = L^2(D)$.*

To describe the smoothness of the initial data we define intermediate spaces between $\mathcal{H}$ and $\mathcal{V}$ by the real method of function space interpolation: specifically,

$$\mathcal{H}_\theta = (\mathcal{H}, \mathcal{V})_{\theta,2}, \quad 0 < \theta < 1\ ,$$

where we define $\mathcal{H}_0 := \mathcal{H}$ and $\mathcal{H}_1 := \mathcal{V}$. We consider the following DG-formulation:

Definition 10.3.2. *Let $\mathcal{M} = \{I_m\}_{m=1}^M$, $M \in \mathbb{N}$ be a partition of $I = (0,T)$, $\underline{r} \in \mathbb{N}_0^M$, then the DGFEM for* (9.3)-(9.4) *reads as follows: find $U \in \mathcal{V}^{\underline{r}}(\mathcal{M};\mathcal{V}) := \{u : I \to \mathcal{V} : u|_{I_m} \in \mathcal{P}^{r_m}(I_m, \mathcal{V}), 1 \leq m \leq M\}$ such that*

$$\begin{aligned}
B_{DG}(U,V) &= F_{DG}(V), \ \textit{where} \qquad\qquad (10.14)\\
B_{DG}(U,V) &= \sum_{m=1}^M \int_{I_m} (U', V)_{\mathcal{H}} dt + \sum_{m=1}^M \int_{I_m} t^\gamma a(U,V) dt + \sum_{m=2}^M ([U]_{m-1}, V_{m-1}^+)_{\mathcal{H}}\\
&\quad + (U_0^+, V_0^+)_{\mathcal{H}},\\
F_{DG}(V) &= (g, V_0^+)_{\mathcal{H}} + \sum_{m=1}^M \int_{I_m} (f(t), V)_{\mathcal{V}^* \times \mathcal{V}} dt
\end{aligned}$$

for all $V \in \mathcal{V}^{\underline{r}}(\mathcal{M};\mathcal{V})$. Here $I_m = (t_{m-1}, t_m)$, $1 \leq m \leq M$,

$$U_m^+ = \lim_{s \downarrow 0} u(t_m + s) \quad \textit{and} \quad U_m^+ = \lim_{s \downarrow 0} u(t_m - s)$$

and we set $[U]_m := U_m^+ - U_m^-$.

The following result holds due to [105, Lemma 1.8] for $B_{DG}(\cdot,\cdot)$.

Lemma 10.3.3. *Let $B_{DG}(\cdot,\cdot)$ be as in Definition 10.3.2, then for all $V,W \in \mathcal{V}^{\underline{r}}(\mathcal{M};\mathcal{V})$*

$$\begin{aligned} B_{DG}(V,W) &= \sum_{m=1}^{M} \int_{I_m} (-V,W')_{\mathcal{H}} + t^{\gamma} a(V,W) dt - \sum_{m=1}^{M} (V_m^-,[W]_m)_{\mathcal{H}} \\ &\quad + (V_M^-, W_M^-)_{\mathcal{H}}, \\ B_{DG}(V-W,V-W) &= \sum_{m=1}^{M} \int_{I_m} t^{\gamma} a(V-W,V-W) dt + \frac{1}{2} \left\|(V-W)_0^+\right\|_{\mathcal{H}}^2 \\ &\quad + \frac{1}{2} \sum_{m=1}^{M-1} \|[V-W]_m\|_{\mathcal{H}}^2 + \frac{1}{2} \sum_{m=1}^{M-1} \left\|(V-W)_M^-\right\|_{\mathcal{H}}^2 . \end{aligned}$$

Theorem 10.3.4. *Problem* (10.14) *has a unique solution U. If u is the solution of* (9.3)-(9.4)*, then we have the Galerkin orthogonality*

$$B_{DG}(u-U,V) = 0 \text{ for all } V \in \mathcal{V}^{\underline{r}}(\mathcal{M};\mathcal{V}).$$

Proof. The proof follows as in [105, Proposition 1.7], where the case $\gamma = 0$ was treated. □

This implies the following quasi-optimality result.

Theorem 10.3.5. *Let u be the exact solution of* (9.3)-(9.4) *and U the semidiscrete solution of* (10.14) *in $\mathcal{V}^{\underline{r}}(\mathcal{M},\mathcal{V})$. Besides, assume $u \in C([\varepsilon,T],\mathcal{V})$, for arbitrary $\varepsilon > 0$. Let $Iu \in \mathcal{V}^{\underline{r}}(\mathcal{M},\mathcal{V})$ be the interpolant of u which is defined on each time interval I_m as $Iu|_{I_m} = \Pi_{I_m}^{r_m}(u|_{I_m})$. Then there holds*

$$\|u-U\|_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \leq C \|u-Iu\|_{L^2_{t^{\gamma/2}}(I;\mathcal{V})}, \tag{10.15}$$

for some $C > 0$.

Proof. Using Lemma 10.3.3, $\|u\|_{\mathcal{V}}^2 = a(\cdot,\cdot)$ and the Galerkin orthogonality we obtain

$$\int_I t^{\gamma} \|U-Iu\|_{\mathcal{V}}^2 dt \leq B_{DG}(U-Iu,U-Iu) = |B_{DG}(u-Iu,U-Iu)|. \tag{10.16}$$

This implies

$$\int_I t^{\gamma} \|U-Iu\|_{\mathcal{V}}^2 dt \leq \int_I t^{\gamma} \|u-Iu\|_{\mathcal{V}}^2 dt \tag{10.17}$$

and therefore the claim follows using triangle inequality. □

Thus, it suffices to estimate the projection error to conclude the a priori error analysis. We have the following approximation result.

Lemma 10.3.6. *Let $I = (a, b)$, $k = b - a$, $r \in \mathbb{N}_0$ and $u \in H^{s_0+1}(I; \mathcal{V})$ for some $s_0 \in \mathbb{N}_0$. Then*

$$\|u - \Pi_I^r u\|_{L^2(I;\mathcal{V})}^2 \leq \frac{C}{\max\{1, r\}^2} \frac{\Gamma(r+1-s)}{\Gamma(r+1+s)} \left(\frac{k}{2}\right)^{2(s+1)} \|u\|_{H^{s+1}(I;\mathcal{V})}^2,$$

for any $0 \leq s \leq \min\{r, s_0\}$, s real.

Proof. See [105, Corollary 1.20] □

In order to complete the error analysis we need bounds on the growth of the solution of (9.3)-(9.4) u and its derivatives. In the following an infinite series representation of u in terms of the eigenfunctions of the operator $\mathcal{A}$ is used for this purpose. For data $g \in \mathcal{H}$ and $f \in L^2_{t^{-\gamma/2}}(I; \mathcal{H})$ the solution of (9.3)-(9.4) can be represented as follows:

$$u(t) = \sum_{i=1}^{\infty} u_{\lambda_i}(t)(g, \varphi_i)_{\mathcal{H}} \varphi_i + \sum_{i=1}^{\infty} \left(\int_0^t u_{\lambda_i}(t)(f(s), \varphi_i)_{\mathcal{H}} ds \right) \varphi_i,$$

where $u_{\lambda_i}(t)$ is given by

$$u_{\lambda_i}(t) = e^{-\lambda_i \frac{t^{\gamma+1}}{\gamma+1}}$$

and $\{\varphi_i\}_{i\in\mathbb{N}}$ denotes the family of eigenfunctions of $\mathcal{A}$, with $\varphi_i \in \mathcal{V}$ for $i \in \mathbb{N}$. We assume that the eigenvalues $\{\lambda_i\}_{i\in\mathbb{N}}$ are non-decreasing and that the eigenfunctions $\{\varphi_i\}_{i\in\mathbb{N}}$ form an orthonormal basis of $\mathcal{H}$.

Theorem 10.3.7. *Let the operator $T(t)$ for $u \in \mathcal{H}$, $t \geq 0$ be given by*

$$T(t)u = \sum_{i=1}^{\infty} u_{\lambda_i}(t)(u, \varphi_i)_{\mathcal{H}} \varphi_i.$$

Then the following estimates hold for $\min(T, 1) > t > 0$, independent of $l \geq 1$,

$$\begin{aligned} \left\|T^{(l)}(t)\right\|_{\mathcal{L}(\mathcal{H},\mathcal{V})}^2 &\leq Ct^{-\gamma-3} \left(\frac{2}{t}\right)^{2(l-1)} (2l)!, \\ \left\|T^{(l)}(t)\right\|_{\mathcal{L}(\mathcal{V},\mathcal{V})}^2 &\leq Ct^{-\gamma-2} \left(\frac{2}{t}\right)^{2(l-1)} (2l)!. \end{aligned}$$

For $l = 0$ and for $\min(T, 1) > t > 0$ we obtain

$$\|T(t)\|_{\mathcal{L}(\mathcal{H},\mathcal{V})}^2 \leq Ct^{-\gamma-1}, \quad \|T(t)\|_{\mathcal{L}(\mathcal{V},\mathcal{V})}^2 \leq C.$$

Proof. We assume without loss of generality $T = 1$. We have for $t \in (0, 1]$

$$\left\| T^{(l)}(t)u \right\|_{\mathcal{V}}^2 \leq C \sum_{i=1}^{\infty} \lambda_i^3 2^{2(l-1)} t^{2\gamma - 2(l-1)} \, (l!)^2 \, u_{\lambda_i}^2(t) \, |u_i|^2 \, ,$$

where we recall $\|u\|_{\mathcal{V}}^2 = a(u, u)$. This estimate holds, as the number of terms for the ℓth derivative is bounded by $2^{2(l-1)}$. The function $h(\lambda) = \lambda^3 e^{-\lambda \frac{t^{\gamma+1}}{\gamma+1}}$ attains its maximum at $\lambda_{\max} = \frac{3(\gamma+1)}{2(t^{\gamma+1})}$, therefore

$$\begin{aligned} \left\| T^{(l)}(t)u \right\|_{\mathcal{V}}^2 &\leq Ch(\lambda_{\max}) t^{2\gamma} \left(\frac{2}{t} \right)^{2(l-1)} (2l)! \, \|u\|_{\mathcal{H}}^2 \\ &\leq Ct^{-\gamma-3} \left(\frac{2}{t} \right)^{2(l-1)} (2l)! \, \|u\|_{\mathcal{H}}^2 \, . \end{aligned}$$

This implies

$$\left\| T^{(l)}(t) \right\|_{\mathcal{L}(\mathcal{H},\mathcal{V})} \leq Ct^{-\gamma-3} \left(\frac{2}{t} \right)^{2(l-1)} (2l)!.$$

The $\mathcal{L}(\mathcal{V}, \mathcal{V})$-norm can be estimated similarly. Using the fact that the maximum of $h(\lambda) = \lambda^2 e^{-\lambda \frac{t^{\gamma+1}}{\gamma+1}}$ is attained at $\lambda_{\max} = \frac{(\gamma+1)}{(t^{\gamma+1})}$, we obtain for all $t \in (0, 1]$

$$\left\| T^{(l)}(t)u \right\|_{\mathcal{V}}^2 \leq Ct^{-\gamma-2} \left(\tfrac{2}{t} \right)^{2(l-1)} (2l)! \, \|u\|_{\mathcal{V}}^2 \, .$$

For $l = 0$, there exists $C > 0$ such that for all $t \in (0, 1]$:

$$\|T(t)u\|_{\mathcal{V}}^2 \leq C \|u\|_{\mathcal{V}}^2 \, , \quad \|T(t)u\|_{\mathcal{V}}^2 \leq Ct^{-\gamma-1} \|u\|_{\mathcal{H}}^2 \, .$$

□

Remark 10.3.8. *Note that the estimates in the previous theorem coincide with the results of [105, Section 2.1],[106, 107], for $\gamma = 0$.*

In the following we split the solution u of (9.3)-(9.4) into its homogeneous and inhomogeneous part, i.e., $u = u_1 + u_2$, where

$$u_1' + t^{\gamma} \mathcal{A} u_1 = 0, \quad u_1(0) = g, \tag{10.18}$$

$$u_2' + t^{\gamma} \mathcal{A} u_2 = f, \quad u_2(0) = 0. \tag{10.19}$$

The behavior of both terms will be studied separately. The function $u_1(t)$, for $t \in [0, T]$ can be represented as

$$u_1(t) = T(t)g. \tag{10.20}$$

Theorem 10.3.9. *Let $g \in \mathcal{H}_\theta$ for $0 \leq \theta \leq 1$. Let u_1 be the solution of* (10.18)*. Then there holds for $l \geq 1$ and for* $\min(T,1) > t > 0$

$$\left\| u_1^{(l)} \right\|_{\mathcal{V}}^2 \leq C t^{-\gamma-3+\theta} \left(\frac{2}{t}\right)^{2(l-1)} (2l)! \, \|g\|_{\mathcal{H}_\theta}^2 , \ \textit{and} \ \|u_1\|_{\mathcal{V}}^2 \leq C t^{(-\gamma-1)(1-\theta)} \|g\|_{\mathcal{H}_\theta}^2 .$$

Proof. The proof follows from (10.20) and Theorem 10.3.7. □

The solution u_2 of (10.19) can be represented as

$$u_2(t) = \int_0^t T(t-s) f(s) ds, \quad 0 \leq t \leq T. \tag{10.21}$$

In the following we assume f in (10.19) to satisfy

$$\left\| f^{(l)}(t) \right\|_{\mathcal{H}} \leq C l! d^l, \quad t \in [0,T], l \in \mathbb{N}_0, \tag{10.22}$$

with some positive constants C and d, independent of l and t.

Lemma 10.3.10. *Under Assumption* (10.22)*, we get for any $t > 0$*

(i) $u_2(t) = \int_0^t T(s) f(t-s) ds$ *in* $\mathcal{H}$.

(ii) $u_2^{(l)}(t) = \sum_{i=0}^{l-1} T^{(i)}(t) f^{(l-i-1)}(0) + \int_0^t T(s) f^{(l)}(t-s) ds$ *for $l \geq 1$ in* $\mathcal{H}$.

Proof. The first claim follows from (10.21) via a change of variable and (ii) follows from (i) by induction. □

Lemma 10.3.11. *Assume* (10.22) *and let u_2 solve* (10.19)*. Then there exist constants C, d such that for* $\min(T,1) > t > 0$

$$\left\| u_2^{(l)}(t) \right\|_{\mathcal{V}} \leq C d^l l! \left(t^{1/2-\gamma/2} + \sum_{i=0}^{l-1} t^{-i-1/2-\gamma/2} \right).$$

Proof. Applying triangle inequality we have

$$\begin{aligned} \left\| u_2^{(l)}(t) \right\|_{\mathcal{V}} \leq & \sum_{i=0}^{l-1} \left\| T^{(i)}(t) \right\|_{\mathcal{L}(\mathcal{H},\mathcal{V})} \left\| f^{(l-1-i)}(0) \right\|_{\mathcal{H}} \\ & + \int_0^t \|T(s)\|_{\mathcal{L}(\mathcal{H},\mathcal{V})} \left\| f^{(l)}(t-s) \right\|_{\mathcal{H}} ds := S_1 + S_2. \end{aligned}$$

The two terms S_1 and S_2 are estimated separately. We first bound S_1. From Theorem 10.3.9 we have

$$\left\| T^{(l)}(t) \right\|_{\mathcal{L}(\mathcal{H},\mathcal{V})} \leq C t^{-(\gamma+3)/2} \left(\frac{2}{t}\right)^{(l-1)} ((2l)!)^{1/2} \tag{10.23}$$

$$\leq \quad Cd^{l-1}t^{-(\gamma+1)/2-l}l!. \tag{10.24}$$

Using (10.22) and (10.24) we conclude

$$\begin{aligned}
S_1 &\leq Cd^{l-1}\sum_{i=0}^{l-1} i!(l-i-1)!t^{-i-1/2-\gamma/2} \\
&\leq Cd^{l-1}(l-1)!\sum_{i=0}^{l-1}\binom{l-1}{i}^{-1} t^{-i-1/2-\gamma/2} \\
&\leq Cd^{l-1}(l-1)!\sum_{i=0}^{l-1} t^{-i-1/2-\gamma/2}.
\end{aligned}$$

The bound on S_2 follows similarly

$$S_2 \leq Cd^l l! \int_0^t s^{-1/2-\gamma/2}ds = Cd^l l! t^{1/2-\gamma/2}.$$

□

Using Lemma 10.3.11, we are now able to derive a bound for $u_2^{(l)}$, $l \geq 0$.

Theorem 10.3.12. *Assume* (10.22) *and let* u_2 *solve* (10.19). *Then there exist constants* C, d *such that for* $\min(T, 1) > t > 0$

$$\left\| u_2^{(l)}(t) \right\|_{\mathcal{V}}^2 \leq Cd^{2l}(2l)!t^{-2l+1-\gamma}$$

for $l \in \mathbb{N}_0$.

Proof. From Lemma 10.3.11 we have

$$\left\| u_2^{(l)}(t) \right\|_{\mathcal{V}} \leq Cd_1^l l!(l+1)t^{-l+1/2-\gamma/2},$$

for some $d_1 > 0$. The claim follows using the properties of the Gamma function, i.e.,

$$\Gamma(c+1)^2 \leq \Gamma(c+1)\Gamma(c+3/2) = C\Gamma(2c+2)2^{-2(c+1)},$$

for $c \in \mathbb{C}$ and $C > 0$. □

The following theorem gives an estimate for the short time behavior of the solution. This is crucial for the proof of exponential convergence of the DG scheme as constant basis functions are used on the first interval I_1.

Theorem 10.3.13. *Let* g *be in* $\mathcal{H}_\theta$, $I = (0, k)$ *and let* f *satisfy* (10.22), *then*

$$\int_0^k \|u(t) - u(k)\|_{\mathcal{V}}^2 t^\gamma dt \leq Ck^{(\gamma+1)\theta}$$

Proof.

$$\begin{aligned} \|u(t)-u(k)\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} &\leq C\|u_1(t)\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} + C\|u_1(k)\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \\ &\quad + C\|u_2(t)\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} + C\|u_2(k)\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \\ &:= T_1 + T_2 + T_3 + T_4. \end{aligned}$$

We first bound T_1 using Theorem 10.3.9

$$T_1 \leq t^{\gamma+1}\|g\|^2_{\mathcal{V}} \text{ and } T_1 \leq C\|g\|^2_{\mathcal{H}}.$$

Therefore by interpolation

$$T_1 \leq Ck^{(\gamma+1)\theta}.$$

We can also bound T_2 with Theorem 10.3.9

$$T_2 \leq Ck^{(1+\gamma)\theta}.$$

T_3 and T_4 can be bounded by Theorem 10.3.12:

$$T_3 \leq Ck^2, \quad T_4 \leq Ck^2.$$

□

Theorem 10.3.14. *Let $g \in \mathcal{H}_\theta$ for $0 \leq \theta \leq 1$ and let f satisfy (10.22), then u satisfies for $0 < a \leq b < \min(1, T)$*

$$\left\|u^{(l)}(t)\right\|^2_{\mathcal{V}} \leq Cd^{2l}(2l)!t^{-(2l+1)-\gamma+\theta}, \tag{10.25}$$

$$\int_a^b \left\|u^{(l)}(t)\right\|^2_{\mathcal{V}} t^\gamma dt \leq Cd^{2l}(2l)!a^{-2l+\theta}. \tag{10.26}$$

Proof. We split the solution u into u_1 and u_2. The estimate (10.25) follows directly from Theorem 10.3.9 and 10.3.12. The estimate (10.26) can be obtained from (10.25):

$$\begin{aligned} \int_a^b \left\|u^{(l)}(t)\right\|^2_{\mathcal{V}} t^\gamma dt &\leq Cd^{2l}(2l)! \int_a^\infty t^{-(2l+1)+\theta} dt \\ &\leq Cd^{2l}(2l)!a^{-2l+\theta}. \end{aligned}$$

□

Lemma 10.3.15. *Let $g \in \mathcal{H}_\theta$ for $0 \leq \theta \leq 1$ and let f satisfy (10.22), then u satisfies for $0 < a \leq b < \min(1, T)$*

$$\left\|u^{(l)}(t)\right\|^2_{H^s_{t^{\gamma/2}}((a,b),\mathcal{V})} \leq Cd^{2s}\Gamma(2l+3)a^{-2s+\theta}.$$

Proof. The result follows by interpolation of the statement in Theorem 10.3.14. □

Definition 10.3.16. *A geometric mesh partition $\mathcal{M}_{n,q} = \{I_m\}_{i=1}^{n+1}$ with grading factor $q \in (0,1)$ and $n+1$ time steps I_m is given by the nodes*

$$t_0 = 0, \quad t_m = q^{n-m+1}, \quad 1 \le m \le n+1.$$

The time steps $k_m = t_m - t_{m-1}$ satisfy

$$k_m = \lambda t_{m-1}, \quad \lambda = \frac{1-q}{q},$$

for $2 \le m \le n+1$.

Definition 10.3.17. *A polynomial degree vector $\underline{r} = \{r_m\}_{m=1}^{n+1}$ is called linear with slope $\nu \ge 0$ on the geometric partition $\mathcal{M}_{n,q}$ if $r_1 = 0$ and $r_m = \lfloor \nu m \rfloor$ for $2 \le m \le n+1$.*

Lemma 10.3.18. *Fix an interval $I_m \in \mathcal{M}_{n,q}$, for $2 \le m \le n+1$ and set $s_m = \alpha_m r_m$ with $\alpha_m \in (0,1)$. Then there exist constants C, d such that*

$$\left\|u - \Pi_{I_m}^{r_m} u\right\|^2_{L^2_{t^{\gamma/2}}(I_m,\mathcal{V})} \le C q^{(n-m+2)\theta - |\gamma|} \left((\mu d)^{2\alpha_m} \frac{(1-\alpha_m)^{1-\alpha_m}}{(1+\alpha_m)^{1+\alpha_m}} \right)^{r_m},$$

where $\mu = \max\{1, \lambda\}$ and $\lambda = \frac{1-q}{q}$. The constants C and d only depend on $g \in \mathcal{H}_\theta$, $\theta \in (0,1]$, and γ, f satisfying (10.22).

Proof. We omit for simplicity the dependence of I, r, α, k and s on m in the following and set $t = t_{m-1}$.

$$\begin{aligned}
\|u - \Pi_l^r u\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} &\le \max(a^\gamma, b^\gamma)\, \|u - \Pi_l^r u\|^2_{L^2(I;\mathcal{V})} \\
&\le C \max(a^\gamma, b^\gamma) \frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)} \left(\frac{k}{2}\right)^{2(s+1)} \|u\|^2_{H^{s+1}(I;\mathcal{V})} \\
&\le C \max\left(\left(\frac{a}{b}\right)^\gamma, \left(\frac{b}{a}\right)^\gamma\right) \frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)} \left(\frac{k}{2}\right)^{2(s+1)} \|u\|^2_{H^{s+1}_{t^{\gamma/2}}(I;\mathcal{V})} \\
&\le C \max\left(\left(\frac{a}{b}\right)^\gamma, \left(\frac{b}{a}\right)^\gamma\right) \frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)} \left(\frac{\mu d}{2}\right)^{2s+2} \\
&\quad \times \Gamma(2s+5) t^{2(s+1)} t^{-2s-2+\theta} \\
&\le C \left(\frac{\mu d_2}{2}\right)^{2s} q^{-|\gamma|} \frac{\Gamma(r+1-s)}{r^2\Gamma(r+1+s)} \Gamma(2s+1) t^\theta.
\end{aligned}$$

Using Sterling formula we obtain

$$\frac{\Gamma(r+1-s)}{\Gamma(r+1+s)} \Gamma(2s+1) \le r^{1/2} 2^{2s} \left(\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1-\alpha}}\right)^r.$$

□

Theorem 10.3.19. *Consider the parabolic problem* (9.3)-(9.4) *on* $I = (0,1)$ *with initial data* $g \in \mathcal{H}_\theta$ *for some* $\theta \in (0,1]$ *and right hand side* f *satisfying* (10.22). *The weak formulation is discretized in time using the DGFEM as given in Definition 10.3.2 on a geometric partition* $\mathcal{M}_{n,q}$. *Then there exists* $\nu_0 > 0$ *such that for all linear polynomial degree vectors* $\underline{r} = \{r_m\}_{m=1}^{n+1}$ *with slope* $\nu \geq \nu_0$ *the semidiscrete DGFEM solution* U *obtained in* $\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,q}, \mathcal{V})$ *satisfies*

$$\|u - U\|_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \leq C \exp\left(-bN^{1/2}\right).$$

Proof. Let

$$\nu > \max\left\{1, \frac{\theta \ln(q)}{\ln(h_{\min})}\right\}, \tag{10.27}$$

where $h_{\min}$ will be defined below. Set $r_1 = 0$ and $r_m = \lfloor \nu m \rfloor \geq 1$ for $2 \leq m \leq n+1$. As before $s_m = \alpha_m r_m$, for $\alpha_m \in (0,1)$ to be selected. We start from (10.15) and use Theorem 10.3.13 to estimate the error on the first interval I_1 near the origin and Lemma 10.3.18 to estimate the error on $I_2, \ldots, I_{n+1}$. This yields

$$\begin{aligned}
\|u - U\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} &\leq Cq^{n\theta(1+\gamma)} + C\sum_{m=2}^{n+1} q^{(n-m+2)-|\gamma|} h_{\mu,d}(\alpha_m)^{r_m} \\
&\leq Cq^{n\theta}\left(q^{n\theta\gamma} + q^{-|\gamma|}\sum_{m=2}^{n+1} q^{(2-m)^\theta} h_{\mu,d}(\alpha_m)^{r_m}\right),
\end{aligned}$$

where $h_{\mu,d}(\alpha) = (\mu d)^{2\alpha}\left(\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1-\alpha}}\right)$. The function $h_{\mu,d}$ satisfies

$$0 < \inf_{0<\alpha<1} h_{\mu,d}(\alpha) = h_{\mu,d}(\alpha_{\min}) < 1 \text{ with } \alpha_{\min} = \frac{1}{\sqrt{1+\mu^2 d^2}}.$$

Set $h_{\min} = h_{\mu,d}(\alpha_{\min})$ and select $\alpha_m = \alpha_{\min}$ for $2 \leq m \leq n+1$. Hence

$$\|u - U\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \leq Cq^{n\theta}\left(q^{n\theta\gamma} + q^{-|\gamma|}\sum_{m=2}^{n+1} q^{(2-m)^\theta} h_{\min}^{r_m}\right). \tag{10.28}$$

Since

$$q^{(2-m)\theta} h_{\min}^{r_m} \leq Cq^{2\theta}\left(\frac{f_{\min}^{\nu}}{q^\theta}\right)^m$$

and $f_{\min}^\nu < q^\theta$ by (10.27), we conclude that the sum in (10.28) can be bounded by

$$\sum_{m=2}^{n+1} q^{(2-m)^\theta} h_{\min}^{r_m} \leq Cq^{2\theta}\sum_{m=2}^{n+1} q^m$$

with $q = f^\nu/q^\theta < 1$. Therefore $\sum_{m=2}^{\infty} q^m < \infty$ holds and we conclude

$$\|u - U\|^2_{L^2_{t^{\gamma/2}}(I;\mathcal{V})} \leq \begin{cases} C_q q^{n\theta} & \text{for } \gamma \geq 0 \\ C_q q^{n\theta(1+\gamma)} & \text{for } \gamma < 0 \end{cases}.$$

Taking into account $N = \mathrm{nrdof}(\mathcal{V}^{\underline{r}}(\mathcal{M}_{n,q}, \mathcal{V})) \leq \mathcal{O}(n^2)$, as $n \to \infty$ with $\mathcal{O}$ dependent on ν, concludes the proof. □

Remark 10.3.20. *The extension of Theorem 10.3.19 to more general operators, as described in Section 9.3 is not straightforward, as a detailed analysis of the smoothness of the solution $u(t,x)$ is needed, as given in Theorems 10.3.9 and 10.3.12.*

11 American options

In this chapter pricing of early exercise contracts is discussed. We consider the corresponding linear complementarity problem (LCP) and show well-posedness for a class of time-homogeneous admissible market models. We discretize the LCP using the implicit Euler scheme in time and linear finite elements in space. The arising sequence of matrix problems is solved employing the primal-dual active set algorithm which converges locally superlinearly for appropriate system matrices.

11.1 Variational formulation

In the following we consider the pricing of American options in admissible time-homogeneous market models. Its value $V(t,s)$ for a Lipschitz continuous payoff functions g is given as

$$V(t,s) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}[e^{-r(T-\tau)} g(S(\tau)) | S(t) = s],$$

where $\mathcal{T}_{t,T}$ denotes the set of all stopping times between t and T. For sufficiently smooth solutions $V(t,s)$, we obtain the following linear complementarity problem for $u(t,x) = V(T-t, e^x)$

$$\begin{aligned} \partial_t u - \mathcal{A}u + ru &\geq 0 \text{ in } I \times \mathbb{R}, && (11.1)\\ u(t,x) &\geq g(e^x) \text{ in } I \times \mathbb{R}, && (11.2)\\ (\partial_t u - \mathcal{A}u + ru)(u-g) &= 0 \text{ in } I \times \mathbb{R} && (11.3)\\ u(0,x) &= g(e^x) \text{ in } \mathbb{R}, && (11.4) \end{aligned}$$

where $\mathcal{A}$ is the infinitesimal generator of the process X.

Remark 11.1.1. *Note that the derivation of this formulation is formal for time-homogeneous admissible market in general. For Lévy market models* (11.1)-(11.4) *can be justified rigorously, cf. [79].*

We assume the pricing problem to be not drift dominated in the following, cf. Remark 5.2.6. In the case of a drift dominated market model we can either remove the drift, cf. Theorem 5.3.1, or use the methods described in Sections 7.2 and 7.3. A continuous Galerkin approach with linear finite elements is used here. The set of admissible solutions for the variational form of (11.1)-(11.4) is the closed, non-empty and convex set $\mathcal{K}_0$ given as

$$\mathcal{K}_0 := \{v \in \mathcal{V} | v \geq 0 \text{ a.e. } x \in \mathbb{R}\},$$

where $\mathcal{V}$ is the domain of the corresponding bilinear form. We denote by $\phi_0(v)$ the indicator function on $\mathcal{K}_0$ given as

$$\phi_0(v) := I_{\mathcal{K}_0}(v) = \begin{cases} 0, & \text{if } v \in \mathcal{K}_0 \\ +\infty, & \text{else.} \end{cases}$$

The variational formulation of (11.1)-(11.4) reads:

Find $u \in L^2(I;\mathcal{V}) \cap H^1(I;\mathcal{V}^*)$ such that $u(t,\cdot) \in D(\phi_0)$ a.e. in I, $u(0) = 0$ and

$$(\partial_t u, v-u)_{\mathcal{V}^*,\mathcal{V}} + a(u, v-u) + \phi_0(v) - \phi_0(u) \geq -a(g, v-u), \ \forall v \in D(\phi_0), \tag{11.5}$$

where $a(\varphi, \phi) = (\mathcal{A}\varphi, \phi)_{\mathcal{V}^*,\mathcal{V}}$ and $\mathcal{V} = \mathcal{D}(a(\cdot,\cdot))$. The unique solvability of (11.5) follows from [55, Chapter 6, Theorem 2.1] using the continuity and the Gårding inequality satisfied by $a(\cdot,\cdot)$ on $\mathcal{V}$ for g such that $\mathcal{A}g \in L^2(\mathbb{R})$. Additionally, $u \in L^\infty(I;\mathcal{H})$ is shown in [55, Chapter 6, Theorem 2.1].

11.2 Discretization

We use P- and M-matrices in the convergence analysis of our algorithm.

Definition 11.2.1. *A $d \times d$ matrix is called P-matrix if all its principal minors are positive.*

We are able to show convergence of the solution algorithm for the linear complementarity system if the system matrix is an M-matrix.

Definition 11.2.2. *A $d \times d$ matrix is called M-matrix if it is a P-matrix and all its non-diagonal entries are non-positive.*

The localization of the pricing equation can be rigorously justified for Lévy market models as in Section 4.2, we refer to [95, Theorem 4.14] for details. The formulation of the pricing problem on a bounded domain $D = (-R, R)$, for some $R > 0$, reads

Find $u \in L^2(I;\mathcal{V}) \cap H^1(I;\mathcal{V}^*)$ such that $u(t,\cdot) \in D(\phi_{0,R})$ and a.e. in I

$$(\partial_t u, v-u)_{\mathcal{V}_D^*,\mathcal{V}_D} + a_D(u, v-u) + \phi_0(v) - \phi_0(u) \geq -a_D(g, v-u), \ \forall v \in D(\phi_{0,R}), \tag{11.6}$$

$u(0) = 0,$

where $\mathcal{K}_{0,R}$ is given as

$$\mathcal{K}_{0,R} := \{v \in \mathcal{V}_D | v \geq 0 \text{ a.e. } x \in D\},$$

$\phi_{0,R}(v) = I_{\mathcal{K}_{0,R}}(v)$ and $\mathcal{V}_D = \{v \in L^2(D) : \widetilde{v} \in \mathcal{V}\}$. Let $k = T/M$ with $M \in \mathbb{N}$ be the time step and denote by u^m, $m = 0, \ldots, M$ the solution of the following backward Euler discretization of (11.6):

Find $u^{m+1} \in \mathcal{K}_{0,R}$, $m = 0, \ldots, M-1$ such that

$$\left(\frac{u^{m+1} - u^m}{k}, v - u^{m+1}\right) + a_D(u^{m+1}, v - u^{m+1}) \geq -a_D(g, v - u^{m+1}),\ \forall v \in \mathcal{K}_{0,R}, \tag{11.7}$$

$u^0 = 0.$

The sequence (11.7) of elliptic variational inequalities can be reduced to a sequence of finite dimensional LCPs by restricting $\mathcal{V}_D$ to a finite dimensional subspace V_h as in Section 6.2. We use the space V_h of continuous piecewise linear functions with respect to the equidistant subdivision $\mathcal{T}_h$. Then the FE discretization of (11.7) reads:

Find $u_h^{m+1} \in V_h \cap \mathcal{K}_{0,R}$, $m = 0, \ldots, M-1$ such that

$$\left(\frac{u_h^{m+1} - u_h^m}{k}, v - u^{m+1}\right) + a_D(u_h^{m+1}, v - u_h^{m+1}) \geq -a_D(g, v - u_h^{m+1}),\ \forall v \in V_h \cap \mathcal{K}_{0,R}, \tag{11.8}$$

$u_h^0 = 0.$

The sequence of LCPs (11.7) can equivalently be written as follows, where we use the standard 'hat' basis of V_h. Given $\underline{u}_h^0 = \underline{0}$, find $\underline{u}_h^{m+1} \in \mathbb{R}^N$, $N = \dim V_h$, such that for $m = 0, \ldots, M-1$,

$$\begin{aligned} &\mathbf{B}\underline{u}_h^{m+1} \geq \underline{F}^m, \\ &\underline{u}_h^{m+1} \geq \underline{0}, \\ &(\underline{u}_h^{m+1})^\top \big(\mathbf{B}\underline{u}_h^{m+1} - \underline{F}^m\big) = 0, \end{aligned} \tag{11.9}$$

where $\mathbf{B} := \mathbf{M} + k\mathbf{A}$, $\underline{F}^m := k\underline{f} + \mathbf{M}\underline{u}_h^m$ and $\underline{f}_i = -a_D(g, b_i)$. We denote by $\mathbf{M}$ and $\mathbf{A}$ the stiffness and mass matrices in 'hat' basis. We denote by $u_h^m(x) = \sum_{k=1}^N (\underline{u}_h^m)_j b_k(x) \in V_h$. A rigorous error analysis was performed in [84], where the Lévy setup was considered.

Theorem 11.2.3. *Let $u(t,x)$ be the solution of* (11.6) *for a Lévy process X, which is an admissible market model with generating triplet $(0, Q, \nu(dz))$ and $u^m := u(t_m, \cdot)$, further let u_h^m be given as above and $g(e^x) = \max(K - e^x, 0)$, $K > 0$, then the following error estimate holds*

$$\begin{aligned} &\max_m \|u^m - u_h^m\|_{L^2(D)} + \left(\sum_{m=1}^M k \|u^m - u_h^m\|_{\mathcal{V}_D}^2\right)^{1/2} \\ &\leq C(k^\gamma + Qh^{s-1} + h^{\min(s/2, s-Y/2)}), \end{aligned}$$

for some $\gamma \in (0,1]$.

Choose an initial guess $\underline{u}^0, \underline{\lambda}^0$ and $\varepsilon > 0$. Set $k = 0$.
Set $I_k = \{i : \lambda_i^k - cu_i^k < 0\}$, $A_k = \{i : \lambda_i^k - cu_i^k \geq 0\}$
Solve $\mathbf{B}\underline{u}^{k+1} + \underline{\lambda}^{k+1} = \underline{F}$,
$\underline{u}^{k+1} = \underline{0}$ on A_k,
$\underline{\lambda}^{k+1} = 0$ on I_k.
    If $|\underline{u}^{k+1} - \underline{u}^k| < \varepsilon$ stop else
Next k

Table 11.1: Description of the primal-dual active set algorithm

Proof. The proof is given in [84, Appendix B]. □

Remark 11.2.4. *Examining the proof in [84, Appendix B] it becomes clear that the statement can be generalized to more general setups, i.e., to admissible time-homogeneous market models as given in Definition 3.2.3 without drift dominance.*

We are now concerned with the solution of the sequence of linear complementarity problems. A popular method is the PSOR algorithm, cf. [36]. The main drawback of this method is the slow convergence speed. We therefore follow a different approach and use a semismooth Newton method for the solution of (11.9), cf. [61, 67]. One step of (11.9) can equivalently be written as

$$\mathbf{B}\underline{u} - \underline{\lambda} = \underline{F}, \quad \underline{u} \geq \underline{0}, \quad \underline{\lambda} \geq \underline{0}, \quad \underline{u}^\top \underline{\lambda} = 0, \tag{11.10}$$

for some $\underline{F} \in \mathbb{R}^d$. The system (11.10) has a unique solution $(\underline{u}^*, \underline{\lambda}^*)$ if $\mathbf{B}$ is a P-matrix, cf. [11, Theorem 10.2.14]. Note that the complementarity system in (11.10) can be equivalently expressed as

$$\mathcal{C}(\underline{u}, \underline{\lambda}) = 0, \text{ where } \mathcal{C}(\underline{u}, \underline{\lambda}) = \underline{\lambda} - \max(0, \underline{\lambda} - c\underline{u}) \tag{11.11}$$

for each $c > 0$. The primal-dual active set method is based on using (11.11) as a prediction strategy. Given a current primal-dual pair $(\underline{u}, \underline{\lambda})$, the choice of the next active and inactive sets I and A is given by

$$I = \{i : \lambda_i - cu_i < 0\}, \quad A = \{i : \lambda_i - cu_i \geq 0\}.$$

This leads to the primal-dual active set algorithm as given in Table 11.1. As described in [61], the primal-dual active set method can also be understood as a semi-smooth Newton method. We have the following local convergence result.

Theorem 11.2.5. *Let $\mathbf{B}$ be a P-matrix, then the primal-dual active set algorithm as given in Table 11.1 converges superlinearly to $(\underline{u}^*, \underline{\lambda}^*)$, provided that $\|\underline{u}^0 - \underline{u}^*\| + \|\underline{\lambda}^0 - \underline{\lambda}^*\|$ is sufficiently small.*

Proof. The proof is given in [61, Theorem 3.1]. It relies on the representation of the algorithm as a semi-smooth Newton method. □

A global convergence result can also be shown.

Theorem 11.2.6. *Assume that* $\mathbf{B}$ *is an* M*-matrix. Then* $(\underline{u}^k, \underline{\lambda}^k) \to (\underline{u}^*, \underline{\lambda}^*)$ *for arbitrary initial data.*

Proof. We refer to [61, Theorem 3.2 and Appendix A] for the proof. □

In order to complete the analysis of American options, it remains to show that the system matrices arising in the discretization of time-homogeneous market models are M-matrices.

Lemma 11.2.7. *Let* $\mathbf{A}_{\mathrm{BS}} = (a_{\mathrm{BS}}(b_i, b_j))_{j,i}$ *and* $\mathbf{A}_{\mathrm{J}} = (a_{\mathrm{J}}(b_i, b_j))_{j,i}$ *be given. For*

$$\begin{aligned} a_{\mathrm{BS}}(u,v) &= \int_D \frac{1}{2} Q(x) \partial_x u(x) \partial_x v(x) dx, \quad Q(x) > Q_0 \geq 0, \\ a_{\mathrm{J}}(u,v) &= \int_D \int_D \partial_z u(z) \partial_x v(x) k(z-x) dz dx, \end{aligned}$$

for a jump measure $k(y)$ *and a diffusion coefficient* $Q(x)$ *that satisfy the assumptions of an admissible time-homogeneous market model. Additionally we assume that* $k(y)$ *is non-increasing for* $y > 0$ *and non-decreasing* $y < 0$*, besides let* $a_J(u,u) \geq C \|u\|^2_{L^2(D)}$ *hold. Then* $\mathbf{A}_{\mathrm{BS}}$ *and* $\mathbf{A}_{\mathrm{J}}$ *are* M*-matrices.*

Proof. It follows directly from the definition of $\mathbf{A}_{\mathrm{BS}}$ that its off-diagonal entries are non-positive . The positivity of all real eigenvalues follows from the coercivity of the bilinear form $a_{\mathrm{BS}}(\cdot,\cdot)$. For $\mathbf{A}_{\mathrm{J}}$ we obtain that all off-diagonal entries are non-positive due to the monotonicity of $k(z)$ and therefore of $k^{(-2)}(z)$. The positivity of all real eigenvalues follows from the coercivity of $a_{\mathrm{J}}(\cdot,\cdot)$. □

Remark 11.2.8. *The coercivity of the* $a_J(\cdot,\cdot)$ *follows from the Gårding inequality after an appropriate change of variable. Note that the* M*-matrix property can also be proved for stiffness matrices for general time-homogeneous market models under certain assumptions on the jump measure* $k(x,z)$.

Remark 11.2.9. *Note that wavelet basis can also be used for the solution of the linear system in the primal-dual active set algorithm, therefore leading to a well-conditioned system matrix* $\mathbf{B}$*. The* M*-matrix property for this matrix is not obvious, while the* P*-matrix property is clear, ensuring local convergence. Additionally, a transformation to standard hat basis is necessary in every step to preserve the sign condition.*

12 Numerical Examples and Outlook

12.1 Drift dominated PIDEs

In the following we present numerical examples in one space dimension confirming the analytical results obtained in the previous chapters.

Example 12.1.1. *We consider the* tempered stable process *(for $c = c_+ = c_-$ also called* CGMY *process in [26] or* KoBoL *in [19]) which has parametric a Lévy density of the form*

$$\nu(dz) = \left(c_+ \frac{e^{-\beta_+|z|}}{|z|^{1+Y}} 1_{\{z>0\}} + c_- \frac{e^{-\beta_-|z|}}{|z|^{1+Y}} 1_{\{z<0\}} \right) dz, \tag{12.1}$$

for $c_+, c_- > 0$, $\beta_+, \beta_- > 1$ and $0 \leq Y < 2$.

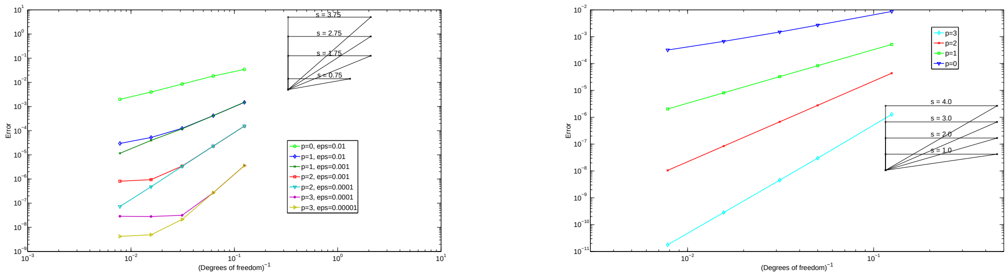
Remark 12.1.2. *Note that we obtain a finite variation process for $Y < 1$ and an infinite variation process for $Y \geq 1$.*

Different choices for the drift $b(x)$ are considered in the following test problem:

$$\mathcal{A}u = f \quad \text{for } x \in G = (0,1), \qquad u = 0 \quad \text{for } x \in \partial G = \{0,1\},$$

where $\mathcal{A}$ will be the corresponding Lévy operator and we choose f such that $u = x^2(x-1)^2$.

Let $\mathcal{A}$ be the pure jump operator with a CGMY jump density and parameters given as $Y \in \{0, 0.5\}$, $\beta_- = \beta_+ = 5$, $C = 1$. The convergence rates can be observed in Figure 12.1, where the error has been measured in the $\widetilde{H}^{Y/2}((0,1))$-norm. Note that in Figure 12.1.(a) we additionally employ the small jump approximation as described in Chapter 4 to approximate the Lévy process X. The figure supports the theoretical results (see Remark 7.3.20) and it can be observed that the truncation error dominates the discretization error for fine discretization levels and large truncation parameters. In Figure 12.2 we consider the same equation with a drift term and observe the convergence rates in the DG-norm $||\cdot||_{\mathrm{DGFV}}$, defined in (7.44), when artificial diffusion is not considered. We choose $b(x) = 20 - 20x$. The results are analogous to the driftless case and confirm the error estimates of Theorem 7.3.19 when $\varepsilon \equiv 0$, i.e., the truncation of the small jumps does not affect the solution. In Figure 12.3 we consider the same problem adding artificial diffusion. The convergence rates in the DG-norm $||\cdot||_{\mathrm{DG}(\varepsilon)}$, defined in (7.25), obtained numerically agree with the theoretical results of Theorem 7.3.9.



(a) CGMY model with $Y = 0.5$ and small jump truncation. Shown is the discretization error measured in the $\widetilde{H}^{0.25}$-norm.

(b) CGMY model with $Y = 0$ (Variance Gamma). Shown is the discretization error measured in the L^2-norm

Figure 12.1: Convergence rates for different Lévy measures

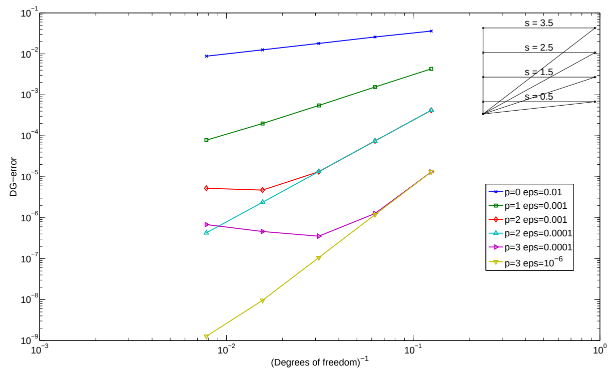


Figure 12.2: Convergence rates for CGMY jump measure with $Y = 0.5$ (drift dominance). Shown is the discretization error measured in the DGFV-norm as given in (7.44) with penalty parameter $\alpha = 5$.

Remark 12.1.3. *We show the order of convergence of the time-independent problem in Figures 12.1-12.3, since for this case an exact solution is known. Theorems 7.3.9 and 7.3.19 present* a priori *error estimates for the time-dependent case in the norm* $|||\cdot|||_{DG(\varepsilon)}$ *given in* (7.29) *and the norm* $|||\cdot|||_{DGFV}$ *given in* (7.45)*, respectively. However, the error measured in the norm* $||\cdot||_{DG(\varepsilon)}$ *(*$||\cdot||_{DGFV}$*, see* (7.44)*) for the time-independent problem has the same order of convergence as the error measured in the norm* $|||\cdot|||_{DG(\varepsilon)}$ *(*$|||\cdot|||_{DGFV}$*) of the corresponding time-dependent problem. This can be easily shown along the lines of the proofs of Theorems 7.3.9 and 7.3.19.*

Now we consider the dependence of the solution on the regularization parameter ε. In the driftless case we observe the behavior presented in Figure 12.4, which confirms the results of Theorem 4.1.8 and Remark 4.1.9. We either only truncate the jump measure on the interval $(-\varepsilon, \varepsilon)$ or add an appropriately scaled diffusion as described in Theorem 4.1.1.

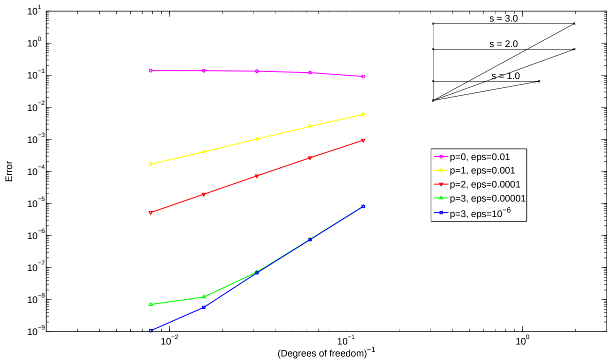
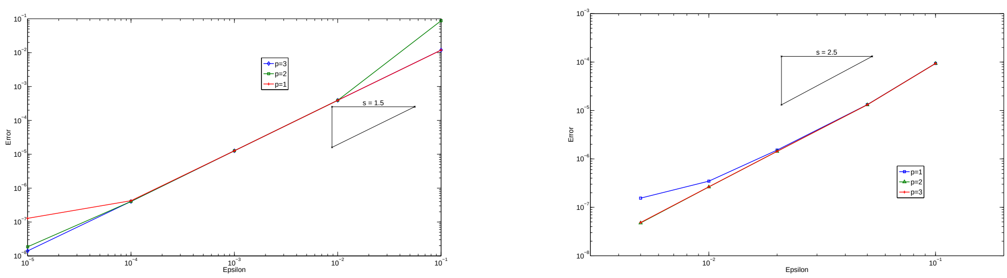


Figure 12.3: Convergence rates for CGMY jump measure with $Y = 0.5$ and artificial diffusion (drift dominance). Shown is the discretization error measured in the DG-norm as given in (7.25) with penalty parameter $\alpha = 5$.

Note that in order to observe a convergence behavior in ε, we have to choose a sufficiently fine discretization, such that the discretization error is negligible in comparison with the truncation error. For the general case we refer to the result in Theorem 4.1.14. The

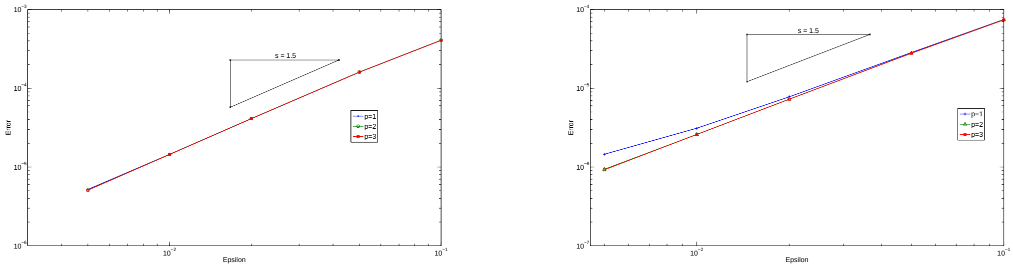


(a) CGMY model with small jump truncation without artificial diffusion.

(b) CGMY model with small jump truncation with artificial diffusion.

Figure 12.4: Convergence rates with respect to ε for CGMY jump measure with $Y = 0.5$. Shown is the discretization error measured in the L^2 norm.

same Lévy kernel as above with the drift component $b(x) = 20 - 20x$ is considered. The numerical results are depicted in Figure 12.5 and confirm the estimate in Theorem 4.1.14. The results suggest that the estimates are optimal. Finally we present a parabolic test case. We consider a pure transport operator with drift $b(x) = 10 - 10x$ and a Lévy operator with the same drift and the Lévy kernel chosen as above. We observe a diffusive behavior of the Lévy operator (Figure 12.6).



(a) CGMY model with small jump truncation without artificial diffusion.

(b) CGMY model with small jump truncation with artificial diffusion.

Figure 12.5: Convergence rates with respect to ε for CGMY jump measure with $Y = 0.5$ (drift dominance). Shown is the discretization error measured in the L^2 norm.

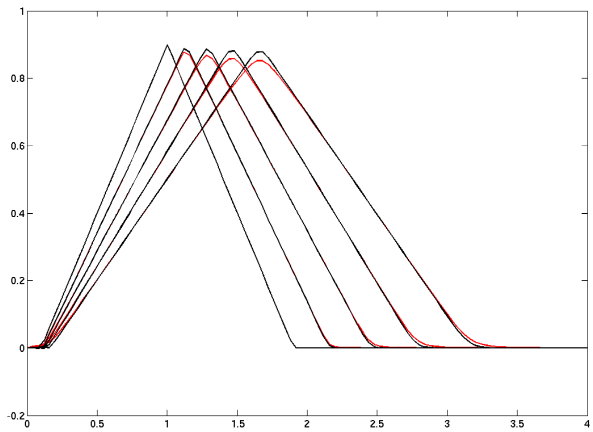


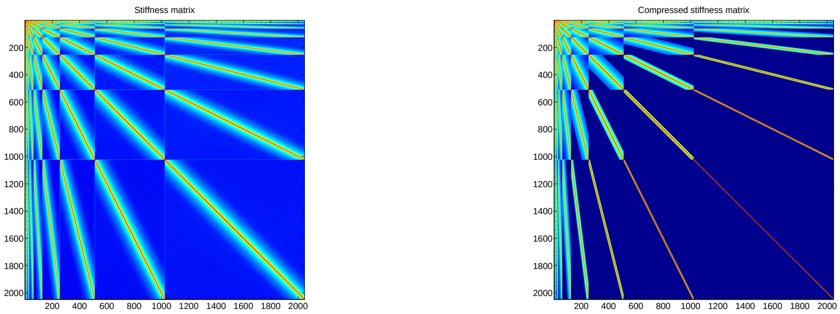
Figure 12.6: Parabolic test problem. Transport operator (black line). Transport dominated Lévy operator (red line).

12.2 PIDEs with inhomogeneous jump measures

In this section the implementation of numerical solution methods for the Kolmogorov equations for time-homogeneous market models with inhomogeneous jump measures using the techniques described above is discussed. We assume the risk-neutral dynamics of the underlying asset to be given by

$$S(t) = S(0)e^{rt+X(t)},$$

where X is a Feller process with characteristic triple $(b(x), Q(x), k(x,z)dz)$ under a risk neutral measure $\mathbb{Q}$ such that e^X is a martingale with respect to the canonical filtration of X. In the following we set $r = 0$ for notational convenience. We consider Feller processes X that are admissible time-homogeneous market models. In the following we consider a special family of Feller processes to confirm the theoretical results of the previous chapters.



(a) Stiffness matrix (b) Compressed stiffness matrix

Figure 12.7: Stiffness matrices for the pure jump case with CGMY-type Lévy kernel $\left(Y(x) = 1.25e^{-x^2} + 0.5\right)$.

(a) Stiffness matrix (b) Mass matrix

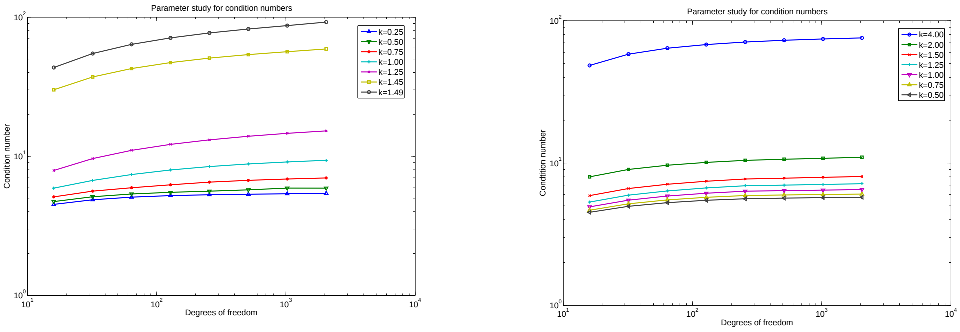
Figure 12.8: Stiffness and Mass matrices for the Black-Scholes model with $\sigma = 0.3$ and $r = 0$.

Example 12.2.1. *We consider a CGMY-type Feller process with jump kernel*

$$k(x,z) \quad = \quad C \begin{cases} e^{-\beta^+ z} y^{-1-Y(x)}, & z > 0 \\ e^{-\beta^- |z|} \, |y|^{-1-Y(x)}, & z < 0, \end{cases} \qquad Y(x) = ke^{-x^2} + 0.5.$$

This process has no Gaussian component and the drift $b(x)$ is chosen according to (5.4).

We also consider the following family of processes that do not satisfy the conditions of the theory developed above, since the variable order is assumed to be Lipschitz continuous only.



(a) Example 12.2.1 (b) Example 12.2.2

Figure 12.9: Condition numbers for different levels and choices of k.

Example 12.2.2. *We consider again a CGMY-type Feller process with jump kernel*

$$\begin{aligned} k(x,z) &= C \begin{cases} e^{-\beta^+ z} y^{-1-Y(x)}, & z > 0 \\ e^{-\beta^- |z|} |z|^{-1-Y(x)}, & z < 0, \end{cases} \\ Y(x) &= 0.5 + k \begin{cases} 0.4x, & 0.25 > x > 0 \\ 0.8x - 0.1, & 0.5 > x \geq 0.25 \\ -0.4x + 0.5, & 0.75 > x \geq 0.5 \\ -0.8x + 0.8, & 1 > x \geq 0.75 \\ 0, & \text{else} \end{cases} . \end{aligned}$$

This process has no Gaussian component and the drift $b(x)$ is chosen according to (5.4).

In Figure 12.7 the stiffness matrix for the process in Example 12.2.1 is depicted. Note that the uncompressed stiffness matrix is densely populated, but structurally very similar to the matrix in the Black-Scholes model. The condition numbers of the preconditioned stiffness matrices have to be uniformly bounded in the number of levels due to Section 8.2. A parameter study for various choices of k in Example 12.2.1 and Example 12.2.2 is shown in Figure 12.9. The condition numbers are uniformly bounded and of order 10^1 in most cases, although the norm equivalences (6.18) only apply to Example 12.2.1. For variable orders with $1.95 \leq \overline{Y}$ we obtain condition numbers of order 10^2. Note that the condition numbers are not only influenced by the order of the singularity of the jump kernel at $z = 0$, but also by the rates of exponential decay β^+ and β^-. Fast decaying tails, i.e., large β^+ and β^- may lead to larger constants. Figure 12.10 shows the price of a European put option for several Lévy processes and one Feller process. In the Feller case we choose $Y(x) = 0.8e^{-x^2} + 0.1$ in Example 12.2.1 and for the Lévy models we set $Y \in \{0.1, 0.5, 0.7, 0.8, 0.9\}$. In all cases we set $C = 1$, $\beta^+ = \beta^- = 10$ and use truncation parameters $a = -3$, $b = 3$ in log-moneyness coordinates. The prices in the Feller model are significantly different from the prices in the different Lévy models. This can be

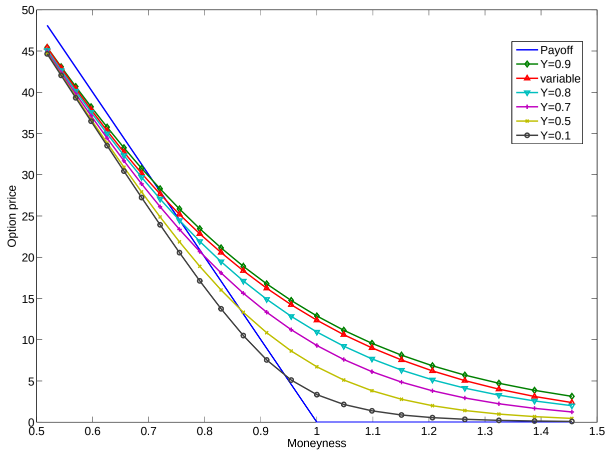


Figure 12.10: Option prices for several models for a European put option with $T = 1$ and $K = 100$.

explained by the ability of the Feller model to account for different tail behavior for different states of the process, which is not possible using Lévy processes. Figure 12.11 shows the prices of an American put option for a Feller process and several Lévy models. We use a Lagrangian multiplier approach as described in [67, 68], cf. Chapter 11. The parameters were chosen as above.

12.3 Time-inhomogeneous PIDEs

In the following we present numerical results for time-degenerate parabolic equations. As a test problem we consider a time-degenerate heat equation

$$\begin{aligned} \partial_t u - t^\gamma \partial_{xx} u &= f(t,x) \text{ on } [0,T] \times (0,1), \\ u(0,x) &= g(x), \end{aligned} \tag{12.2}$$

$T = 1$. For numerical testing purposes, the exact solution $u(t,x)$ is selected equal to $u(t,x) = e^{-t}\sin(\pi x)$ and the right hand side f is chosen such that (12.2) holds. Linear finite elements are used for the spatial discretization. The numerical results are depicted in Figure 12.12. The decrease of the convergence rate for large N stems from the fact the error of the time-stepping scheme is dominated by the discretization error in space. We observe exponential convergence for various parameters γ and σ.

Figure 12.13 shows European option prices for different Hurst parameters $H \in (0,1)$ in an FBM market model. Significant differences in the option price can be observed even for plain vanilla contracts.

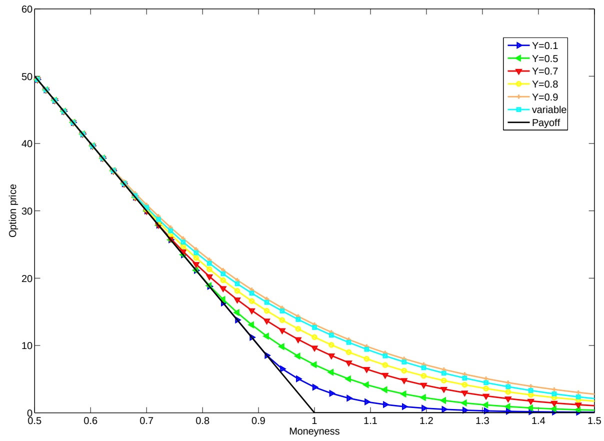
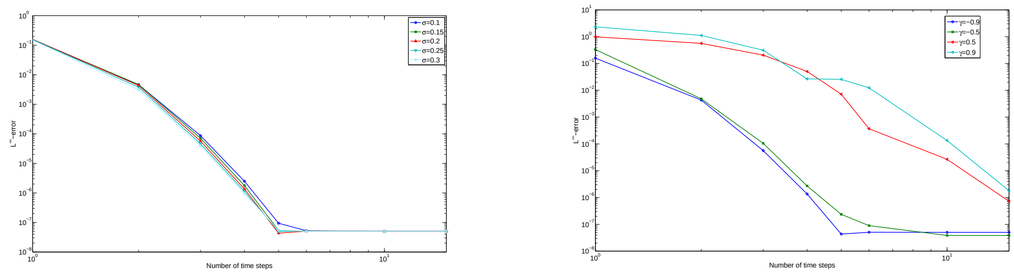


Figure 12.11: Option prices for several models for an American put option with $T = 1$ and $K = 100$.



(a) Exponential convergence rates for different parameters σ and $\gamma = -0.9$.

(b) Exponential convergence rates for different parameters γ and $\sigma = 0.2$.

Figure 12.12: Exponential convergence rates for the test problem (12.2) with slope $\nu = 1$, measured in the L^2-norm at $t = 1$.

12.4 Outlook

We have analyzed integro-differential equations and inequalities arising in option pricing for extensions of Lévy processes. Several types of extensions have been considered. Spatially inhomogeneous market models have been analyzed. The well-posedness of the corresponding pricing equations in a setup with and without drift dominance has been shown. The discretization was performed using discontinuous Galerkin methods with small jump approximation and continuous Galerkin methods with streamline diffusion. Time-inhomogeneous models have also been considered. Due to the possible degeneracy of the coefficients of the arising equation, appropriate weak space-time formulations have been used. The discretization was performed using a continuous Galerkin method

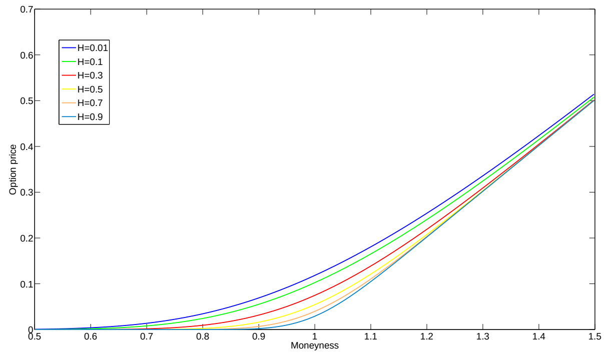


Figure 12.13: Prices of plain vanilla European options in the FBM market model for different Hurst parameters H

in space-time and alternatively a discontinuous Galerkin discretization in time. We close by listing several directions of future research related to the results of this thesis.

- The smoothness requirements on the symbol of an admissible time-homogeneous market model stem from the use of pseudodifferential calculus. Numerically, the derived results, in particular for the preconditioning, seem to hold even for Lipschitz continuous variable orders. An analysis for symbols with non-smooth orders is desirable.
- The pricing of exotic contracts, such as contracts of Asian or Parisian type, under the market models could be considered.
- The results on localization and small jump approximation in Chapter 4 could be extended to more general market models.
- The application of the derived results to the pricing of options in commodity markets, such as the gas or the electricity market, is an interesting topic, as standard models usually do not produce satisfactory results in such markets, cf. [21].

References

[1] H. Amann. *Linear and Quasilinear Parabolic Problems.* Birkhäuser, Basel, 1995.

[2] R. Andreev. Space-time wavelet FEM for parabolic equations. Technical Report 20, SAM, ETH Zurich, 2010.

[3] R. Andreev. Private communication. 2011.

[4] D. Applebaum. *Lévy processes and Stochastic Calculus, 2nd Edition.* Cambridge Univ. Press, New York, 2009.

[5] I. Babuška. Error-bounds for finite element methods. *Numer. Math.*, 16:322–333, 1971.

[6] O. Barndorff-Nielsen. Processes of Normal Inverse Gaussian type. *Finance and Stochastics*, 2:41–68, 1998.

[7] O. Barndorff-Nielsen and S.Z. Levendorskiĭ. Feller Processes of Normal Inverse Gaussian type. *Quantitative Finance*, 1:318–331, 2001.

[8] A. Bentata and R. Cont. Matching marginal distributions of a semimartingale with a Markov process. *C. R. Acad. Sci. Paris*, (Ser. I 347), 2009.

[9] F.E. Benth. On the arbitrage-free pricing of weather derivatives based on fractional Brownian Motion. *Appl. Mathematical Finance*, 10:303–324, 2003.

[10] F.E. Benth, G. Di Nunno, and A. Khedher. Robustness of option prices and their deltas in markets modelled by jump-diffusions. *Communications on Stochastic Analysis*, 5(2):285–307, 2011.

[11] A. Berman and Plemmons R.J. *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York, 1979.

[12] J. Bertoin. *Lévy processes.* Cambridge Univ. Press, New York, 1996.

[13] S. Beuchler, R. Schneider, and Ch. Schwab. Multiresolution weighted norm equivalences and applications. *Numer. Math.*, 98(1):67–97, 2004.

[14] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications.* Springer, New York, 2008.

[15] T. Björk and H. Hult. A note on Wick products and the fractional Black-Scholes model. *Finance and Stochastics*, 9(2):197–209, 2005.

[16] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Pol. Econ.*, 81(3):637–59, 1973.

[17] J.M. Bony, Ph. Courrege, and P. Priouret. Semi-groupes de feller sur une variété à bord compacte et problèmes aux limites intégro-différentiels du second ordre donnant lieu au principe du maximum. *Ann. Inst. Fourier Grenoble*, 18(2):369–521, 1968.

[18] B. Böttcher. Construction of time-inhomogeneous Markov processes via evolution equations using pseudo-differential operators. *Journal of London Mathematical Society*, 1:1–17, 2008.

[19] S.I. Boyarchenko and S.Z. Levendorskiĭ. *Non-Gaussian Merton-Black-Scholes theory*, volume 9 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ, 2002.

[20] D. Braess. *Finite Elements*. Cambridge Univ. Press, New York, 2001.

[21] N. Branger, O. Reichmann, and M. Wobben. Pricing electricity derivatives on an hourly basis. *The Journal of Energy Markets*, 3:51–89, 2010.

[22] S.C. Brenner and L.R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, 2002.

[23] P. Carr. Local variance gamma. Technical report, Private communication, 2009.

[24] P. Carr, H. Geman, D. Madan, and M. Yor. From local volatility to local Lévy models. *Quantitative Finance*, 4:581–588, 2004.

[25] P. Carr, H. Geman, D. Madan, and M. Yor. Self-decomposability and option pricing. *Mathematical Finance*, 17:31–57, 2007.

[26] P. Carr, H. Geman, D.B. Madan, and M. Yor. The fine structure of assets returns: An empirical investigation. *J. Business*, 75(2):305–332, 2002.

[27] A. Chernov. Optimal convergence estimates for the trace of the polynomial L^2-projection operator on a simplex. Technical report, Institute for Numerical Simulation, University of Bonn, 2010. `http://chernov.ins.uni-bonn.de/research/pub/1002.pdf`, to appear in Math. Comp.

[28] A. Chernov, T. von Petersdorff, and Ch. Schwab. Exponential convergence of *hp* quadrature for integral operators with Gevrey kernels. *ESAIM Math. Model. Numer. Anal.*, 45(3):387–422, 2011.

[29] A. Cohen. *Numerical Analysis of Wavelet Methods*, volume 32. North Holland-Elsevier, Amsterdam, 2003.

[30] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations - convergence rates. *Math. Comp.*, 70:27–75, 2001.

[31] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods II - beyond the elliptic case. *Found. Comput. Math.*, 2:203–245, 2002.

[32] S. Cohen and J. Rosiński. Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes. *Bernoulli*, 13(1):195–210, 2007.

[33] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics*, 9(3):299–325, 2005.

[34] P. Courrège. Sur la forme intégro-differentielle des opérateurs de C_k^∞ dans C satisfaisant du principe du maximum, sém. *Théorie du Potentiel*, 2:38pp, 1965/66.

[35] J.C. Cox. Notes on option pricing I: constant elasticity of variance diffusions. Technical report, Stanford University, Stanford, CA., 1975.

[36] C.W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM J. Control*, 9:385–392, 1971.

[37] W. Dahmen, A. Kunoth, and K. Urban. Biorthogonal spline wavelets on the interval - stability and moment conditions. *Applied and Computational Harmonic Analysis*, 6:132–196, 1999.

[38] W. Dahmen and R. Schneider. Wavelets with complementary boundary conditions, function spaces on the cube. *Result. Math.*, 34:255–293, 1998.

[39] H Dappa. *Quasiradiale Fouriermultiplikatoren*. PhD thesis, University of Darmstadt, 1982.

[40] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Spectral Theory and Applications*, volume 3. Springer, Heidelberg, 1990.

[41] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Evolution Problems I*, volume 5. Springer, Heidelberg, 1992.

[42] F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.*, 312:215–250, 1998.

[43] T.J. Dijkema. *Adaptive tensor product wavelet methods for solving PDEs*. PhD thesis, University of Utrecht, 2009.

[44] T.J. Dijkema, Ch. Schwab, and R. Stevenson. An adaptive wavelet method for solving high-dimensional elliptic PDEs. *Constr. Approx.*, 30:423–455, 2009.

[45] G.C. Donovan, J.S. Geronimo, and D.P. Hardin. Interwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM J. Math. Anal.*, 27:1791–1815, 1996.

[46] G.C. Donovan, J.S. Geronimo, and D.P. Hardin. Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets. *SIAM J. Math. Anal.*, 30:1029–1056, 1999.

[47] B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[48] E. Eberlein, A. Papapantoleon, and A. N. Shiryaev. On the duality principle in option pricing: semimartingale setting. *Finance and Stochastics*, 12:265–292, 2008.

[49] R.J. Elliott and L. Chan. Perpetual American options with fractional Brownian motion. *Quantitative Finance*, 4:123–128, 2004.

[50] R.J. Elliott and J. van der Hoek. Fractional Brownian Motion and financial modelling. In M. Kohlmann and Tang S., editors, *Mathematical Finance, Workshop on the Mathematical Finance research project*, pages 140–151. Birkhäuser Boston, Boston, MA, 2001.

[51] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*. Applied mathematical sciences. Springer, New York, 2002.

[52] W. Farkas, N. Reich, and Ch. Schwab. Anisotropic Stable Lévy Copula Processes-Analytical and Numerical Aspects. *Mathematical Models and Methods in Applied Sciences*, 17:1405–1443, 2007.

[53] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20:367–387, 1948.

[54] E.H. Georgoulis. *Discontinuous Galerkin Methods on Shape-Regular and Anisotropic Meshes*. PhD thesis, Computing Laboratory, University of Oxford, 2003.

[55] R. Glowinski, J.-L. Lions, and R. Trémolières. *Numerical Analysis of Variational Inequalities*, volume 8 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1981.

[56] M. Griebel and P. Oswald. Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, 4(1-2):171–206, 1995.

[57] P. Grisvard. Espaces intermédiares entre espaces de Sobolev aves poids. *Annali della Scuola Normale Superiore di Pisa*, 17:255–296, 1963.

[58] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theo. Rel. Fields*, 71:501–516, 1986.

[59] H. Harbrecht and R. Stevenson. Wavelets with patchwise cancellation properties. *Math. Comp.*, 75:1871–1889, 2006.

[60] N. Hilber. *Stabilized Wavelet Method for Pricing in High Dimensional Stochastic Volatility Models*. PhD thesis, SAM, ETH Dissertation No. 18176, 2009. `http://e-collection.ethbib.ethz.ch/view/eth:41687`.

[61] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13:865–888, 2003.

[62] W. Hoh. Pseudodifferential operators generating Markov processes. Habilitationsschrift, University of Bielefeld, 1998.

[63] W. Hoh. Pseudodifferential operators with negative definite symbols of variable order. *Rev. Math. Iberoamericana*, 16:219–241, 2000.

[64] P. Houston, Ch. Schwab, and E. Süli. Stabilized *hp*-finite element methods for first-order hyperbolic problems. *SIAM Journal on Numerical Analysis*, 37:1618–1643, 2000.

[65] P. Houston, Ch. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM Journal on Numerical Analysis*, 39:2133–2163, 2002.

[66] Y. Hu and B. Øksendal. Fractional white noise calculus and applications to finance. *Infinite Dimensional Analysis, Quantum Probability and related topics*, 6:1–32, 2003.

[67] K. Ito and K. Kunisch. Semi-smooth Newton Methods for variational inequalities of the first kind. *Mathematical Modelling and Numerical Analysis*, 37:41–62, 2003.

[68] K. Ito and K. Kunisch. Parabolic variational inequalities: The Lagrange multiplier approach. *J. Mat. Pures Appl.*, 85:415–449, 2005.

[69] N. Jacob. *Pseudo-Differential Operators and Markov Processes Vol. 1: Fourier Analysis and Semigroups*. Imperial College Press, London, 2001.

[70] N. Jacob. *Pseudo-Differential Operators and Markov Processes Vol. 2: Generators and their Potential Theory*. Imperial College Press, London, 2002.

[71] N. Jacob and R. Schilling. Subordination in the sense of S. Bochner - an approach through Pseudo Differential Operators. *Math. Nachr.*, 178:199–231, 1996.

[72] J. Jacod and A. Shiriyaev. *Limit Theorems for Stochastic Processes*. Springer, Heidelberg, 2003.

[73] M. Kac. On the distribution of certain Wiener functionals. *Trans. Amer. Math. Soc.*, 65:1–13, 1949.

[74] M. Kac. On some connections between probability theory and differential and integral equations. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 189–215, 1951.

[75] J. Kallsen and P. Tankov. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *J. Mult. Anal.*, 97:1551–1572, 2006.

[76] K. Kikuchi and A. Negoro. On Markov process generated by pseudodifferential operator of variable order. *Osaka J. Math.*, 34:319–335, 1997.

[77] W. Kluge. *Time-inhomogeneous Lévy processes in interest rate and credit risk models*. PhD thesis, University of Freiburg, 2005.

[78] V. Knopova and R.L. Schilling. Transition density estimates for a class of Lévy and Lévy-type processes. Technical report, arXive, 2009. `http://arxiv.org/abs/0912.1482`, to appear in J. Theor. Probab.

[79] D. Lamberton and M. Mikou. The critical price for the American put in an exponential Lévy model. *Finance and Stochastics*, 12(4):561–581, 2008.

[80] J.-L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications*, volume 1 of *Travaux et Recherches Mathématiques*. Dunod, Paris, 1968.

[81] D.B. Madan and E. Seneta. The variance gamma model for share market returns. *J. Business*, 63:511–524, 1990.

[82] B. Mandelbrot. The variation of certain speculative prices. *J. Business*, 36:394–419, 1963.

[83] D. Marazzina, O. Reichmann, and Ch. Schwab. *hp*-DGFEM for Kolmogorov equations of multivariate Lévy processes. Technical Report 17, SAM, ETH Zurich, 2011. (to appear in M3AS).

[84] A.-M. Matache, P.-A. Nitsche, and Ch. Schwab. Wavelet Galerkin pricing of American options on Lévy driven assets. *Quantitative Finance*, 5(4):403–424, 2005.

[85] A.-M. Matache, Ch. Schwab, and T. P. Wihler. Linear complexity solution of parabolic integrodifferential equations. *Numerische Mathematik*, 104:69–102, 2004.

[86] A.-M. Matache, Ch. Schwab, and T. P. Wihler. Fast numerical solution of parabolic integro-differential equations with applications in finance. *SIAM J. Sci. Comput.*, 27:369–393, 2005.

[87] R.C. Merton. Theory of rational option pricing. *Bell J. Econ. Manag. Sci.*, 4:141–183, 1973.

[88] M. Métivier. *Semimartingales*. W. de Gruyter, Berlin, 1982. a Course on Stochastic Processes.

[89] R.B. Nelsen. *An introduction to copulas*, volume 139 of *Lecture Notes in Statistics*. Springer, New York, 1999.

[90] H. Nguyen. *Finite element wavelets for solving partial differential equations*. PhD thesis, Universiteit Utrecht, 2005.

[91] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, Berlin, 1983.

[92] M. Primbs. *Stabile biorthogonale Spline-Waveletbasen auf dem Intervall*. PhD thesis, University of Duisburg-Essen, 2006.

[93] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer, Berlin, 2005.

[94] M. Reed and B. Simon. *Methods of modern mathematical physics - Functional analysis*. Academic Press, London, 1980.

[95] N. Reich, Ch. Schwab, and Ch. Winter. On Kolmogorov equations for anisotropic multivariate Lévy processes. *Finance and Stochastics*, 14:527–567, 2010.

[96] O. Reichmann. Optimal space-time adaptive wavelet methods for degenerate parabolic PDEs. Technical Report 03, SAM, ETH, 2011. (to appear in Numerische Mathematik).

[97] O. Reichmann, R. Schneider, and Ch. Schwab. Wavelet solution of variable order pseudodifferential equations. *Calcolo*, 47(2):65–101, 2010.

[98] O. Reichmann and Ch. Schwab. Numerical analysis of additive, Lévy and Feller processes with applications to option pricing. *Lévy Matters I, Lecture Notes in Mathematics*, 2001:137–196, 2010.

[99] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, New york, third edition, 1990.

[100] B. Rivière. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*. SIAM, Philadelphia, 2008.

[101] S. Roman. The formula of Faà di Bruno. *Amer. Math. Monthly*, 87:805–809, 1980.

[102] P.A. Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6:13–31, 1965.

[103] K. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge Studies in Advanced Mathematics, Cambridge, 1999.

[104] R. Schilling. On the existence of Feller processes with a given generator. Technical report, University of Dresden, 2004.

[105] D. Schötzau. *hp-DGFEM for parabolic evolution problems*. PhD thesis, SAM, ETH Dissertation No. 13041, 1999. `http://e-collection.library.ethz.ch/view/eth:22928`.

[106] D. Schötzau and Ch. Schwab. An *hp* a priori error analysis of the DG time-stepping method for initial value problems. *Calcolo*, 37(4):207–232, 2000.

[107] D. Schötzau and Ch. Schwab. *hp*-discontinuous Galerkin time-stepping for parabolic problems. *C.R.Acad. Sci. Paris*, 333:1121–1126, 2001.

[108] Ch. Schwab. Variable order composite quadrature of singular and nearly singular integrals. *Computing*, 53:173–194, 1994.

[109] Ch. Schwab. *p- and hp- Finite Element Methods. Theory and Applications in solid and fluid mechanics*. Oxford Science Publications, Oxford, 1998.

[110] Ch. Schwab and R. Stevenson. Adaptive wavelet algorithms for elliptic PDEs on product domains. *Math. Comp.*, 77(261):71–92, 2008.

[111] Ch. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comp.*, 78(267):1293–1318, 2009.

[112] K. Urban. *Wavelet Methods for Elliptic Partial Differential Equations*. Oxford University Press, Oxford, 2009.

[113] Ch. Winter. *Wavelet Galerkin schemes for option pricing in multidimensional Lévy models*. PhD thesis, SAM, ETH Dissertation No. 18221, 2009. `http://e-collection.ethbib.ethz.ch/view/eth:41555`.

[114] Ch. Winter. Wavelet Galerkin schemes for multidimensional anisotropic integrodifferential operators. *SIAM Journal on Scientific Computing*, 32(3):1545–1566, 2010.

Curriculum Vitae

Personal details

Name	Oleg Reichmann
Date of birth	June 21, 1985
Place of birth	St.Petersburg, Russia
Citizenship	Germany and Russia

Education

02/2009–01/2012	PhD studies in Mathematics at ETH Zürich Zürich, Switzerland
12/2008	Diploma in Business Managament at WWU Münster
08/2008	Diploma in Mathematics at WWU Münster
10/2004–12/2008	Studies in Business Management at WWU Münster Münster, Germany
10/2003–08/2008	Studies in Mathematics at WWU Münster Münster, Germany
10/2005–06/2006	Erasmus student at Complutense University Madrid, Spain
06/2003	Abitur at Goethe-Gymnasium
08/1995–06/2003	Secondary school at Goethe-Gymnasium Düsseldorf, Germany