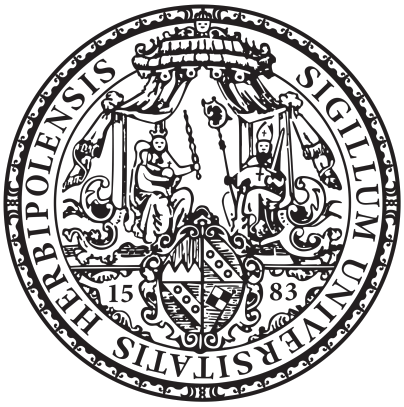
Theoretical and numerical analysis of Fokker-Planck optimal control problems for jump-diffusion processes



Dissertation zur Erlangung

des naturwissenschaftlichen Doktorgrades

der Julius-Maximilians-Universität Würzburg

vorgelegt von

Beatrice Gaviraghi

Würzburg, 7. März 2017

Eingereicht am: December 2016

1. Betreuer: Prof. Dr. Alfio Borzì, Julius-Maximilians-Universität Würzburg

2. Betreuer: Dr. Mario Annunziato, Università degli Studi di Salerno

Abstract

The topic of this thesis is the theoretical and numerical analysis of optimal control problems, whose differential constraints are given by Fokker-Planck models related to jump-diffusion processes. We tackle the issue of controlling a stochastic process by formulating a deterministic optimization problem. The key idea of our approach is to focus on the probability density function of the process, whose time evolution is modeled by the Fokker-Planck equation. Our control framework is advantageous since it allows to model the action of the control over the entire range of the process, whose statistics are characterized by the shape of its probability density function.

We first investigate jump-diffusion processes, illustrating their main properties. We define stochastic initial-value problems and present results on the existence and uniqueness of their solutions. We then discuss how numerical solutions of stochastic problems are computed, focusing on the Euler-Maruyama method.

We put our attention to jump-diffusion models with time- and space-dependent coefficients and jumps given by a compound Poisson process. We derive the related Fokker-Planck equations, which take the form of partial integro-differential equations. Their differential term is governed by a parabolic operator, while the nonlocal integral operator is due to the presence of the jumps. The derivation is carried out in two cases. On the one hand, we consider a process with unbounded range. On the other hand, we confine the dynamic of the sample paths to a bounded domain, and thus the behavior of the process in proximity of the boundaries has to be specified. Throughout this thesis, we set the barriers of the domain to be reflecting.

The Fokker-Planck equation, endowed with initial and boundary conditions, gives rise to Fokker-Planck problems. Their solvability is discussed in suitable functional spaces. The properties of their solutions are examined, namely their regularity, positivity and probability mass conservation. Since closed-form solutions to Fokker-Planck problems are usually not available, one has to resort to numerical methods.

The first main achievement of this thesis is the definition and analysis of conservative and positive-preserving numerical methods for Fokker-Planck problems. Our SIMEX1 and SIMEX2 (Splitting-Implicit-Explicit) schemes are defined within the framework given by the method of lines. The differential operator is discretized by a finite volume scheme given by the Chang-Cooper method, while the integral operator is approximated by a mid-point rule. This leads to a large system of ordinary differential equations, that we approximate with the Strang-Marchuk splitting method. This technique decomposes the original problem in a sequence of different subproblems with simpler structure, which are separately solved and linked to each other through initial conditions and final solutions. After performing the splitting step, we carry out the time integration with first- and second-order time-differencing methods. These steps give rise to the SIMEX1 and SIMEX2 methods, respectively.

A full convergence and stability analysis of our schemes is included. Moreover, we are able to prove that the positivity and the mass conservation of the solution to Fokker-Planck problems are satisfied at the discrete level by the numerical solutions computed with the SIMEX schemes.

The second main achievement of this thesis is the theoretical analysis and the numerical solution of optimal control problems governed by Fokker-Planck models. The field of optimal control deals with finding control functions in such a way that given cost functionals are minimized. Our framework aims at the minimization of the difference between a known sequence of values and the first moment of a jump-diffusion process; therefore, this formulation can also be considered as a parameter estimation problem for stochastic processes. Two cases are discussed, in which the form of the cost functional is continuous-in-time and discrete-in-time, respectively.

The control variable enters the state equation as a coefficient of the Fokker-Planck partial integro-differential operator. We also include in the cost functional a L^1-penalization term, which enhances the sparsity of the solution. Therefore, the resulting optimization problem is nonconvex and nonsmooth. We derive the first-order optimality systems satisfied by the optimal solution. The computation of the optimal solution is carried out by means of proximal iterative schemes in an infinite-dimensional framework.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit der theoretischen und numerischen Analyse von Optimalsteuerungsproblemen, deren Nebenbedingungen die Fokker-Planck-Gleichungen von Sprung-Diffusions-Prozessen sind. Unsere Strategie baut auf der Formulierung eines deterministischen Problems auf, um einen stochastischen Prozess zu steuern. Der Ausgangspunkt ist, die Wahrscheinlichkeitsdichtefunktion des Prozesses zu betrachten, deren zeitliche Entwicklung durch die Fokker-Planck-Gleichung modelliert wird. Dieser Ansatz ist vorteilhaft, da er es ermöglicht, den gesamten Bereich des Prozesses durch die Wirkung der Steuerung zu beeinflussen.

Zuerst beschäftigen wir uns mit Sprung-Diffusions-Prozessen. Wir definieren Ausgangswertprobleme, die durch stochastische Differentialgleichungen beschrieben werden, und präsentieren Ergebnisse zur Existenz und Eindeutigkeit ihrer Lösungen. Danach diskutieren wir, wie numerische Lösungen stochastischer Probleme berechnet werden, wobei wir uns auf die Euler-Maruyama-Methode konzentrieren.

Wir wenden unsere Aufmerksamkeit auf Sprung-Diffusions-Modelle mit zeit- und raumabhängigen Koeffizienten und Sprüngen, die durch einen zusammengesetzten Poisson-Prozess modelliert sind. Wir leiten die zugehörigen Fokker-Planck-Gleichungen her, die die Form von partiellen Integro-Differentialgleichungen haben. Ihr Differentialterm wird durch einen parabolischen Operator beschrieben, während der nichtlokale Integraloperator Sprünge modelliert. Die Ableitung wird auf zwei unterschiedlichen Arten ausgeführt, je nachdem, ob wir einen Prozess mit unbegrenztem oder beschränktem Bereich betrachten. In dem zweiten Fall muss das Verhalten des Prozesses in der Nähe der Grenzen spezifiziert werden; in dieser Arbeit setzen wir reflektierende Grenzen.

Die Fokker-Planck-Gleichung, zusammen mit einem Anfangswert und geeigneten Randbedingungen, erzeugt das Fokker-Planck-Problem. Die Lösbarkeit dieses Problems in geeigneten Funktionenräumen und die Eigenschaften dessen Lösung werden diskutiert, nämlich die Positivität und die Wahrscheinlichkeitsmassenerhaltung. Da analytische Lösungen von Fokker-Planck-Problemen oft nicht verfügbar sind, müssen numerische Methoden verwendet werden.

Die erste bemerkenswerte Leistung dieser Arbeit ist die Definition und Analyse von konservativen numerischen Verfahren, die Fokker-Planck-Probleme lösen. Unsere SIMEX1 und SIMEX2 (Splitting-Implizit-Explizit) Schemen basieren auf der Linienmethode. Der Differentialoperator wird durch das Finite-Volumen-Schema von Chang und Cooper diskretisiert, während der Integraloperator durch eine Mittelpunktregel angenähert wird. Dies führt zu einem großen System von gewöhnlichen Differentialgleichungen, das mit der Strang-Marchuk-Splitting-Methode gelöst wird. Diese Technik teilt das ursprüngliche Problem in eine Folge verschiedener Teilprobleme mit einer einfachen Struktur, die getrennt gelöst werden und danach durch deren Anfangswerte miteinander verbunden werden. Dank der Splitting-Methode kann jedes Teilproblem implizit oder explizit gelöst werden. Schließlich wird die numerische Integration des Anfangswertsproblems mit zwei Verfahren durchgeführt, nämlich dem Euler-Verfahren und dem Predictor-Corrector-Verfahren.

Eine umfassende Konvergenz- und Stabilitätsanalyse unserer Systeme ist enthalten. Darüber hinaus können wir beweisen, dass die Positivität und die Massenerhaltung der Lösung von Fokker-Planck-Problemen auf diskreter Ebene durch die numerischen Lösungen erfüllt werden, die mit den SIMEX-Schemen berechnet wurden.

Die zweite bemerkenswerte Leistung dieser Arbeit ist die theoretische Analyse und die numerische Behandlung von Optimalsteuerungsproblemen, deren Nebenbedingungen die Fokker-Planck-Probleme von Sprung-Diffusions-Prozessen sind. Der Bereich der optimalen Steuerung befasst sich mit der Suche nach einer optimalen Funktion, die eine gegebene Zielfunktion minimiert. Wir zielen auf die Minimierung des Unterschieds zwischen einer bekannten Folge von Werten und dem ersten Moment eines Sprung-Diffusions-Prozesses. Auf diese Weise kann unsere Formulierung auch als ein Parameterschätzungsproblem für stochastische Prozesse angesehen werden. Zwei Fälle sind erläutert, in denen die Zielfunktion zeitstetig beziehungsweise zeitdiskret ist.

Da die Steuerung ein Koeffizient des Integro-Differentialoperators der Zustandsgleichung ist und die Zielfunktion einen L^1-Term beinhaltet, der die dünne Besetzung der Lösung erhöht, ist das Optimierungsproblem nichtkonvex und nichtglatt. Die von der optimalen Lösung erfüllten notwendigen Bedingungen werden hergeleitet, die man mit einem System beschreiben kann. Die Berechnung optimaler Lösungen wird mithilfe von Proximal-Methoden durchgeführt, die entsprechend um den unendlichdimensionalen Fall erweitert wurden.

Contents

Chapter 1

Introduction

1.1 Motivation and review

This thesis is motivated by the need of developing effective and robust numerical schemes for solving problems in quantitative finance. With this spirit, the initial training network STRIKE (January 2013 - December 2016) has been established. Our research project was supported by the European Union in the FP7-PEOPLE-2012-ITN Program under Grant Agreement Number 304617 (FP7 Marie Curie Action, Project Multi-ITN STRIKE - Novel Methods in Computational Finance).

Beyond the Black-Scholes framework, the market behavior can be modeled by stochastic processes with jumps. In fact, empirical data suggest that Lévy processes could be most appropriate for describing the dynamics of stock prices, rathen than an Itô diffusion [2, 3, 7, 23, 24, 25, 47, 55, 57, 67]. The increasing popularity of Lévy processes stems also from the fact that they are Markov jump processes and they include models such as Brownian motions, Poisson processes and subordinators [6, 11, 31, 55, 59, 83]. However, processes with jumps find application in other areas than finance, such as physics and image reconstruction [6, 11, 18, 23, 87, 77, 83].

The research effort in quantitative finance regarding jump processes also considers optimization problems such as option pricing in incomplete markets, portfolio optimization and model calibration. A number of publications aiming at the numerical solution of this class of problems is available, see [16, 24, 25, 52] and references therein. The mathematical modeling of such issues consists of setting stochastic control problems In these references, the expected value of a given cost functional is considered. This averaging step is needed since the random variable is inserted into a deterministic cost functional. The Hamilton-Jacobi-Bellman equation, whose form depends on the chosen cost functional, is derived and then solved with a suitable method [8, 10, 27, 35, 36, 46, 60, 75, 93].

The purpose of this work is to provide an alternative approach for solving optimization problems subject to random perturbations. We consider stochastic processes with jumps. The key observation is to investigate the probability density function (PDF) of the considered process, whose evolution is governed by the Fokker-Planck (FP) equation; this allows one to formulate a deterministic optimal control problem. Up to our knowledge, the literature on FP-based control problems is very recent, and processes containing jumps have been less investigated than Itô diffusion processes. Existing works deal indeed with processes with continuous sample paths, see [4, 5, 54, 82].

In this thesis, we consider a jump-diffusion (JD) Markov process $X = \{X_t\}_{t\in I}$ with range in $\mathbb{R}^d$, evolving in the time interval $I := [0, T] \subset \mathbb{R}^+$. The process X solves the following stochastic initial value problem

$$\begin{cases} dX_t = b(X_{t^-}, t)dt + \sigma(X_{t^-}, t)dW_t + dP_t, \\ X_{\{t=0\}} = X_0, \end{cases} \tag{1.1}$$

where $X_0 \in \mathbb{R}^d$ is a given initial random data. This stochastic differential equation (SDE) relates the infinitesimal increments of the stochastic process X to both deterministic and random increments, given by the multidimensional Wiener process $W \in \mathbb{R}^m$ and the compound Poisson process $P \in \mathbb{R}^d$. We denote with $\lambda \in \mathbb{R}^+$ the rate of the time events of the compound Poisson process and with g the distribution of the jump size. The density g is nonnegative and normalized, $\int_{\mathbb{R}^d} g(y)dy = 1$. The deterministic functions $b : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^{d\times m}$ represent the drift and the diffusion coefficients, respectively. It is always assumed that the matrix σ is full-rank. Due to the compound Poisson process P, the sample paths of X are discontinuous; the notation X_{t^-} stands for the left limit of X_t, $X_{t^-} := \lim_{s\uparrow t} X_s$.

The solvability of (1.1) follows under growth and regularity conditions on b and σ. Classical references dealing with initial value problems such as (1.1) are, e.g., [6, 53, 56, 73]. A number of existing works provides the rate of convergence on existing numerical methods for solving (1.1); see [8, 19, 48, 55, 56, 64, 68, 79] and references therein. However, all these schemes require the simulation of the sample paths of the driving noise; existing algorithms can be found in, e.g., [8, 20, 23, 46, 48, 55, 56].

In contrast to the simulation of (1.1) based on Monte Carlo (MC) techniques [58], we examine the stochastic problem (1.1) by considering its PDF $f(x, t)$. The PDF characterizes the statistics of X over its entire space-time range and its time evolution is modeled by the FP equation, which plays a fundamental role in many problems involving random quantities. This equation has been first applied to problems with randomness given by Brownian motion, i.e. not containing jumps; in this case, this equation is governed by a partial differential equation of parabolic type as follows

$$\partial_t f(x, t) = \mathcal{L} f(x, t). \tag{1.2}$$

The differential operator $\mathcal{L}$ is defined as follows

$$\mathcal{L}f(x,t) := -\sum_{i=1}^{d} \partial_i \left(b_i(x,t)f(x,t)\right) + \sum_{i,j=1}^{d} \partial_{ij}^2 \left(C_{ij}(x,t)f(x,t)\right), \tag{1.3}$$

where $C_{ij}(x) := \frac{1}{2}\sum_{k=1}^{m} \sigma_{ik}(x,t)\sigma_{jk}(x,t)$. Since the diffusion coefficient σ is full rank, the matrix C is positive definite. The derivation of the FP equation, some methods of solution and its application to diffusion models can be found in [26, 32, 38, 62, 63, 77, 80, 87].

In case of a process X solving (1.1), therefore containing jumps, the corresponding FP equation is defined in $\mathbb{R}^d \times I$ and it is governed by the following partial integro-differential equation

$$\partial_t f(x,t) = \mathcal{L}f(x,t) + \mathcal{I}f(x,t), \tag{1.4}$$

where $\mathcal{L}$ is defined as in (1.3). The integral operator in (1.4) depends on the jump rate λ and on the distribution g defining the compound Poisson process in (1.1). For example, in case of an unbounded domain, we have

$$\mathcal{I}f(x,t) = \lambda \int_{\mathbb{R}^d} f(x-y,t)g(y)dy - \lambda f(x,t). \tag{1.5}$$

Equation (1.4), endowed with initial and possibly boundary conditions, gives rise to the FP problem, whose solution represents the PDF of X. The initial condition is chosen to be the PDF f_0 of the initial data X_0 in (1.1).

In this work, we show how the FP problem of a JD process governed by (1.1) can be derived. We focus on two cases. In the first case we consider a process with range in the whole space $\mathbb{R}^d$. The second case limits the dynamic of a process to a bounded domain $\Omega \subset \mathbb{R}^d$ and sets reflecting barriers; the behavior at the boundaries is translated into suitable boundary conditions in the FP problem. In recent works [15] a similar derivation is carried out in case of a discrete random process with bounded jumps, while in [30] and [61] diffusion models with reflecting barriers are investigated.

Since the solution f of a FP problem is a PDF, it must be nonnegative and its integral over the space domain of the process must be equal to 1. In this thesis, we show that these two structural properties are valid when the initial and boundary conditions are suitably chosen.

Classical solutions of initial-boundary value problems containing partial integro-differential equations have been examined in the context of Hölder spaces and uniformly parabolic operators [39, 40, 41]. Numerical schemes for problems governed by partial integro-differential equations are implemented and investigated in [2, 3, 16, 17, 18, 24, 25, 50, 66]. These schemes are mainly motivated by financial applications and a rigorous numerical analysis is often not available.

With the aim of solving a FP equation as (1.2), Chang and Cooper [21] proposed a conservative scheme that computes the density of a diffusion process. A complete

numerical analysis of this method has been carried out in [69]. However, numerical schemes for the FP problem for a JD process have been less investigated; this problem is addressed in this thesis. We discretize (1.4) within the framework of the method of lines [84]. The differential operator $\mathcal{L}$ is discretized with the Chang-Cooper (CC) scheme, and the integral operator $\mathcal{I}$ is approximated by the mid-point rule. This leads to a large system of ordinary differential equations, which we approximate with the Strang-Marchuk (SM) splitting method [34, 45, 50, 65, 90]. This method decomposes the original problem in a sequence of different subproblems with simpler structure, which are separately solved and linked to each other through initial conditions and final solutions. After performing the SM splitting, we carry out the time integration with a first- and a second-order time-differencing method. Our discretization procedure with the two different time-discretization schemes leads to the SIMEX1 and SIMEX2 (Splitting-Implicit-Explicit) schemes, respectively. For clarity, our discretization workflow is summarized in Figure 1.1.

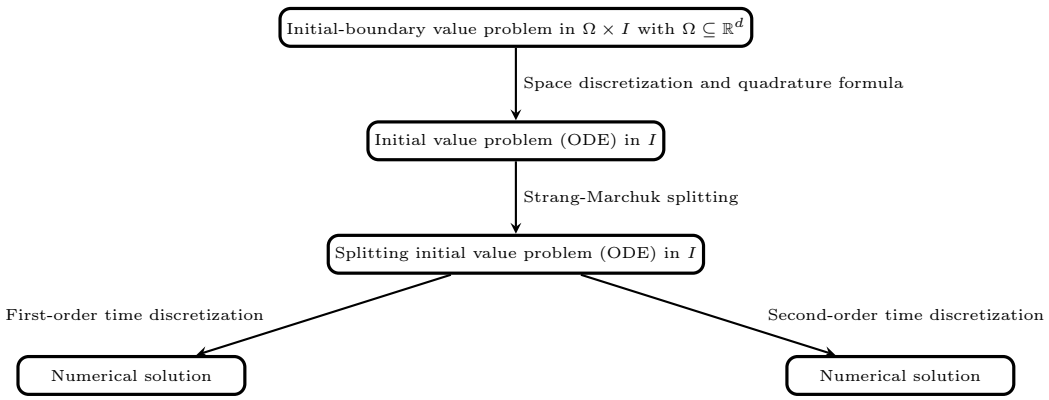


Figure 1.1: Discretization workflow of the SIMEX1 and SIMEX2 schemes.

We remark that splitting methods and finite differences are frequently used by practitioners [31, 51, 52]. However, in many works less attention has been put on positivity and conservation properties, and the numerical analysis has mainly focused on time-approximation properties. In this thesis, we prove that both the SIMEX1 and SIMEX2 schemes guarantees conservativeness of the total probability and nonnegativity of the solutions.

Our stable, convergent and probability-preserving numerical methods allow to numerically address optimal control problems with constraints given by FP problems. In this thesis, we set and investigate optimal control problems to steer the JD process of interest by defining suitable cost functionals. These problems belong to a very active research field, motivated by a broad range of applications ranging from, e.g., fluid flow, space technology, heat phenomena, image reconstruction and finance

[14, 60, 91]. The main focus of this research has been on problems with smooth cost functionals governed by partial differential equations with linear or bilinear control mechanism [4, 12, 37, 60, 85, 86, 91]. In these references, the purpose is often to compute optimal controls such that an appropriate norm of the difference between a given target and the resulting state is minimized.

Very recently, PDE-based optimal control problems with sparsity promoting L^1-cost functionals have been investigated starting with [85, 86, 89]. Such formulation considers nonsmooth cost functionals that give rise to a sparse optimal control; in order to solve such problems, we consider proximal iterative schemes that have been introduced in [70] and [81] and further developed in the framework of finite-dimensional optimization [9, 22, 72]. Recent works extended this framework with the aim of solving infinite-dimensional PDE optimization problems [85, 86].

In this thesis, we set an infinite-dimensional optimal control problem with a non-smooth cost functional. Denote with f the PDF of the process modeled by (1.1) and with u the control variable, which is part of the drift coefficient of the partial integro-differential operator of (1.4). Our optimal control problem reads as follows

$$\begin{aligned} &\min_{f,u} \mathcal{J}(f,u) \\ &\text{s.t. } \mathcal{K}(f,u) = 0, \end{aligned} \tag{1.6}$$

where the differential constraint $\mathcal{K}$ embodies the FP equation (1.4) with initial and boundary conditions. The cost functional $\mathcal{J}$ is defined as follows

$$\mathcal{J}(f,u) := D(f) + \frac{\nu}{2}\|u\|_2^2 + \gamma\|u\|_1. \tag{1.7}$$

The set of admissible controls $\mathcal{U}_{ad}$ is usually a closed and convex set of a Hilbert space where the minimum has to be computed. The L^2-term in (1.7) represents a regularization term that prevents the norm of u from going to infinity. The L^1-term in (1.7) enhances the sparsity of the optimal solution u. The tracking objective $D(f)$ in (1.7) aims at the minimization of the difference between a known sequence of values or a function and the first moment of the JD process (1.1). For this reason, our formulation can also be considered as a parameter estimation problem for stochastic processes. In the discrete-in-time case, the form of the cost functional gives rise to a finite number of discontinuities in time of the adjoint variable and hence of the control. A similar situation has already been considered in [13].

We numerically address the problem (1.6) by applying a method that exploits the additive structure of $\mathcal{J}$ and relies on the subgradient of a nondifferentiable function [33]. The chosen proximal method combines a gradient method [71] and a proximal algorithm proposed in finite-dimension in [72] and extended to infinite-dimensional control problems in [86].

1.2 Outline of the thesis

This thesis is organized as follows.

In Chapter 2, we introduce JD processes, investigate SDEs driven by such processes and discuss how they can be solved numerically. In Section 2.1, we define Lévy processes, which constitute a broad class of jump processes, and outline their characterizing properties. A literature review of results stating existence and uniqueness of solutions to SDEs driven by jump processes is included in Section 2.2. The last part of this chapter, from Section 2.3 on, deals with the numerical solution of SDEs. After outlying the Euler-Maruyama (EM) method, we prove its rate of convergence. We conclude with an application of the EM method for estimating the PDF of a JD process.

Chapter 3 deals with the FP equation related to a JD process, which is the focus of this thesis. We carry out its explicit derivation in two cases, in an unbounded domain and in a bounded domain, respectively. The former is outlined in Section 3.1 and relies on the application of the Itô's formula. The latter is illustrated with a constructive approach in Section 3.2 and takes into account the presence of reflecting barriers. When the dynamics of a stochastic process is restricted to a bounded range, the behavior of the process at the boundaries is of fundamental importance for the definition of appropriate boundary conditions. The reflecting barriers translate into zero-flux boundary conditions in the FP problem. We provide a priori estimates as well as existing results on existence and uniqueness of solutions to our FP problems.

Chapter 4 provides the foundation for the theoretical investigation of optimal control problems, whose constraints are given by JD FP models. In Section 4.1, we define the cost functionals of our optimization problems. We consider two different tracking objectives, discrete-in-time and continuous-in-time, respectively. Both have the purpose to steer the mean value of the stochastic process towards a sequence of given values. The existence of at least an optimal solution is proven. In Section 4.2, we derive the necessary conditions satisfied by local optimal solutions. These conditions take the form of first-order optimality systems, which are outlined in both cases of discrete-in-time and continuous-in-time cost functionals.

In Chapter 5, we address the issue of numerically solving the FP problems. Our SIMEX methods are outlined in the case of a one-dimensional setting in Section 5.1. After discretizing the differential and integral operators with the Chang-Cooper method and the mid-point rule, respectively, the time integration of the resulting large system of ordinary differential equations is carried out. We compute the numerical solution by applying first- and second-order Euler time-discretization schemes. The convergence and stability analysis of the SIMEX scheme is carried out in Section 5.2. We show in Section 5.3 that the chosen discretization approach preserves the two properties of the solution to the FP problem, namely the positivity and the

mass conservation over the space range. Section 5.4 illustrates results of numerical experiments that validate the theoretical statements of the previous sections.

Chapter 6 discusses the numerical solution of our FP-based optimal control problems. Section 6.1 investigates the chosen optimization algorithm. Our approach relies on the definition of a proximity operator of a convex and possibly nonsmooth function. We combine a fixed-point iteration with a gradient method. The convergence analysis of our proximal method is outlined in Section 6.2. In order to investigate the effectiveness of the algorithm, results of numerical experiments are presented in Section 6.3.

Each chapter of this thesis ends with a section of conclusion and some remarks. A final section of conclusions completes this work.

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

- B. Gaviraghi, M. Annunziato and A. Borzì,
  - *Splitting methods for Fokker-Planck equations related to jump-diffusion processes* (Chapter 23)
  - *A Fokker-Planck based approach to control jump processes* (Chapter 24)

  in M. Ehrhardt, Michael Günther and E. J.W. ter Maten (eds.), *STRIKE – Novel Methods in Computational Finance*, Springer, 2017.

- B. Gaviraghi, M. Annunziato and A. Borzì, *Analysis of splitting methods for solving a partial integro-differential Fokker-Planck equation.* Applied Mathematics and Computation (2017) **294**, 1–17.

- B. Gaviraghi, A. Schindele, M. Annunziato and A. Borzì, *On optimal sparse-control problems governed by jump-diffusion processes.* Applied Mathematics (2016) **7**, 1978–2004.

- B. Gaviraghi and A. Borzì, *An operator splitting method for solving a class of Fokker-Planck equations.* European Consortium for Mathematics in Industry Newsletter (2014) **56**, 87–88.

Chapter 2

Formulation and simulation of jump-diffusion processes

The aim of this chapter is to introduce JD processes, investigate SDEs driven by such processes and present how they can be numerically solved. In Section 2.1 we define Lévy processes and outline their main properties. Section 2.2 includes a literature review on the results stating existence and uniqueness of solutions to initial value stochastic models. Section 2.3 deals with the numerical solution of SDEs. We first outline the EM method and we prove its strong rate of convergence. We conclude with an application of this method, which aims at the estimation of the PDF of a JD process.

2.1 Lévy processes

In this section, we introduce the theoretical setting needed in the remainder of the chapter. Additional results and discussion of the material presented in this chapter can be found, for example, in [6, 26, 38, 88].

Throughout this chapter we consider a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, where Ω is the set of the possible outcomes, $\mathfrak{F}$ is a σ-algebra and $\mathbb{P}$ is a probability measure. Let us equip $\mathbb{R}^d$ with the Borel σ-algebra, i.e. the σ-algebra generated by the open sets, denoted by $\mathcal{B}(\mathbb{R}^d)$. A mapping Z from Ω into $\mathbb{R}^d$ is measurable if $Z^{-1}(B) \in \mathfrak{F}$ whenever $B \in \mathcal{B}(\mathbb{R}^d)$. A measurable mapping Z is called a *random variable*. Its associated *probability law* or *distribution* is the probability measure p_Z on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ defined for each $B \in \mathcal{B}(\mathbb{R}^d)$ by

$$p_Z(B) := \mathbb{P}(Z^{-1}(B)).$$

The *expected value* of Z is defined as follows

$$\mathbb{E}[Z] := \int_{\mathbb{R}^d} x\, p_Z(dx).$$

If p_Z is absolutely continuous with respect to Lebesgue measure, the random variable Z admits a density f_Z and the expected value can be written as $\mathbb{E}[Z] = \int_{\mathbb{R}^d} x\, f_Z(x)dx$. The cumulative distribution function F_Z of a real-valued random variable is defined as $F_Z(x) := \mathbb{P}(Z \leq x)$ and it can be written in terms of the density f_Z as follows, $F_Z(x) = \int_{-\infty}^{x} f_Z(y)dy$.

Given the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, a random variable Z with $\mathbb{E}[Z] < \infty$ and a σ-algebra $\mathcal{G} \subseteq \mathfrak{F}$, the *conditional expectation* of Z given $\mathcal{G}$ is the almost surely unique random variable K such that

- K is $\mathcal{G}$-measurable.
- $\mathbb{E}[ZY] = \mathbb{E}[KY]$ for all random variables Y bounded and $\mathcal{G}$-measurable.

In literature, K is denoted with $\mathbb{E}[Z|\mathcal{G}]$. When, in particular, $\mathcal{G}$ is the smallest σ-algebra on Ω that contains all preimages of a random variable J, defined as follows

$$\mathcal{G} := \sigma\{J^{-1}(E) \,|\, E \in \mathcal{B}(\mathbb{R}^d)\},$$

then K is also denoted with $\mathbb{E}[Z|J]$ and is called *conditional expectation* of Z given the random variable J.

Let μ_1 and μ_2 be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. For each set $E \in \mathcal{B}(\mathbb{R}^d)$, the *convolution* of μ_1 and μ_2 is the probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ defined as follows

$$(\mu_1 * \mu_2)(E) = \int_{\mathbb{R}^d} \mu_1(E - x)\mu_2(dx).$$

Two random variables Z_1 and Z_2 are said to be *independent* if for any couple $E_1, E_2 \subseteq \mathcal{B}(\mathbb{R}^d)$, we have $\mathbb{P}(\{Z_1 \in E_1\} \cap \{Z_2 \in E_2\}) = \mathbb{P}(\{Z_1 \in E_1\})\mathbb{P}(\{Z_2 \in E_2\})$. If two random variables Z_1 and Z_2 have the same probability law, they are said to be *identically distributed*, and we write the following equality

$$Z_1 \stackrel{d}{=} Z_2.$$

The characteristic function $\phi_Z : \mathbb{R}^d \to \mathbb{C}$ of a random variable Z with law p_Z is defined for each $u \in \mathbb{R}^d$ as follows

$$\phi_Z(u) := \mathbb{E}[e^{iu\cdot Z}] = \int_{\mathbb{R}^d} e^{iu\cdot y} p_Z(dy),$$

where the scalar product is intended in $\mathbb{R}^d$.

Given the class of random variables, the concept of limit naturally arises. Consider the sequence of random variables $\{Z_n\}_{n\in\mathbb{N}}$. We say that Z_n converges to Z in *mean-square* sense if

$$\lim_{n\to\infty} \mathbb{E}[(Z_n - Z)^2] = 0 \tag{2.1}$$

and we write m.s. $\lim_{n\to\infty} Z_n = Z$. We say that Z_n converges to Z in probability if for each $\varepsilon > 0$ we have that

$$\lim_{n\to\infty} \mathbb{P}(\{|Z_n - Z| > \varepsilon\}) = 0, \tag{2.2}$$

and we write $\lim_{n\to\infty} Z_n \overset{\mathbb{P}}{=} Z$.

The following definition plays a pivotal role in the investigation of Lévy processes.

Definition 1 (Infinite divisibility)**.** A random variable Z is *infinitely divisible* if for all $m \in \mathbb{N}$ there exist m i.i.d. (independent identically distributed) random variables $\{Z_j\}_{j=1}^m$ such that

$$Z \overset{d}{=} \sum_{j=1}^{m} Z_j.$$

Equivalently, we can say that a random variable Z with law p_Z is infinitely divisible if for all $m \in \mathbb{N}$ there exists another law p_{Z_m} of another random variable Z_m such that p_{Z_m} is the m-th convolution root of p_Z. We have

$$p_Z = \underbrace{p_{Z_m} * \cdots * p_{Z_m}}_{m \text{ times}}.$$

Alternatively, it is possible to define an infinitely divisible random variable with its characteristic function. In fact, a random variable is infinitely divisible if and only if for all $m \in \mathbb{N}$ there exists a random variable Z_m such that

$$\phi_Z(u) = [\phi_{Z_m}(u)]^m.$$

Definition 2. (Lévy measure) A measure ν defined on $\mathbb{R}^d\backslash\{0\}$ is a Lévy measure if

$$\int_{\mathbb{R}^d\backslash\{0\}} \min\{||y||^2, 1\}\, \nu(dy) < \infty,$$

where $||y||^2 := \sum_{k=1}^d y_k^2$.

The following theorem gives a necessary and sufficient condition fulfilled by the characteristic function of an infinitely divisible random variable.

Theorem 1. *(Lévy-Khintchine formula) A random variable Z in $\mathbb{R}^d$ is infinitely divisible if and only if there exists a triplet (b, A, ν) with $b \in \mathbb{R}^d$, $A \in \mathbb{R}^{d\times d}$ a positive definite symmetric matrix, and a Lévy measure ν such that*

$$\phi_Z(u) = \mathbb{E}[e^{iu\cdot Z}] = \int_{\mathbb{R}^d} e^{iu\cdot s} p_Z(ds) = \exp\{\psi(u)\} =$$
$$= \exp\left\{ ib\cdot u - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d\backslash\{0\}} [e^{iu\cdot s} - 1 - iu\cdot s \mathbb{1}_{\{||s||<1\}}(s)]\nu(ds) \right\},$$

for each $u \in \mathbb{R}^d$.

The measure ν (and thus the integral) can be extended to the whole set $\mathbb{R}^d$ by defining $\nu(\{0\}) := 0$. The mapping $\psi : \mathbb{R}^d \to \mathbb{C}$ is called the *Lévy* or *characteristic exponent*.

A *stochastic process* X is a parametrized collection of random variables $\{X_t\}_{t\in I}$ defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. For each $t \in I$, the random variable X_t has range in $\mathbb{R}^d$. The parameter $t \in I$ often models the evolution of X over time; in this case, we have that $I \subseteq \mathbb{R}^+$.

Given Ω and $\mathfrak{F}$, a filtration is a sequence of σ-algebras $\{\mathcal{F}_t\}_{t\in I}$ such that if $t_1 \leq t_2$ then $\mathcal{F}_{t_1} \subseteq \mathcal{F}_{t_2} \subseteq \mathfrak{F}$. A filtration represents the time evolution in the set of events that can be measured, according to the growth of the time index t. We always consider the case in which the filtration $\{\mathcal{F}_t\}_{t\in I}$ satisfies the so-called *usual conditions*:

- completeness, which means that $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets;
- right-continuity, which means that $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ for all times t.

A stochastic process X is said to be *adapted* to a filtration $\{\mathcal{F}_t\}_{t\in I}$ if the random variable X_t is $\mathcal{F}_t$-measurable for each $t \in I$. Given a stochastic process, we can define its *natural filtration* $\{\mathcal{F}_t^X\}_{t\in I}$ as follows

$$\mathcal{F}_t^X := \sigma\{X_s^{-1}(E) \,|\, 0 \leq s \leq t,\, E \in \mathcal{B}(\mathbb{R}^d)\}, \tag{2.3}$$

which is the smallest σ-algebra on Ω that contains all preimages of measurable subsets of $\mathbb{R}^d$ for all the times up to t.

The Brownian motion, observed first by Robert Brown in the 1820s and estensively investigated by Albert Einstein in [32], is the most studied example of Lévy process.

Definition 3 (Brownian motion)**.** A stochastic process $W = \{W_t\}_{t\in I}$ defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ is a *Brownian motion* if the following conditions are fulfilled.

- $W_0 = 0$ almost surely, i.e. $\mathbb{P}(\{\omega \in \Omega \text{ such that } W_0(\omega) \neq 0\}) = 0$.
- The mapping $t \mapsto W_t$ is almost surely continuous, in the sense that $\mathbb{P}(\{\omega \text{ such that } W_t(\omega) \text{ is continuous w.r.t. } t\}) = 1$.
- $\{W_t\}_{t\in I}$ has independent increments, which means that the increments corresponding to two nonoverlapping time intervals are independent, as follows

$$t_1 \leq t_2 \leq t_3 \leq t_4 \quad \Rightarrow \quad W_{t_2} - W_{t_1} \text{ and } W_{t_4} - W_{t_3} \text{ are independent.}$$

- For each t and s with $0 \leq s \leq t$, the increment $W_t - W_s$ is distributed as $\mathcal{N}(0, (t-s)^2 I)$, the normal distribution with expected value 0 and covariance matrix $(t-s)^2 I$, where I is the d-dimensional identity matrix.

In one dimension, W is also called *Wiener process*, since Norbert Wiener in the 1920s provided its first rigorous mathematical construction [6].

2.1.1 Definition and properties

In this section, we present the class of Lévy processes, named after the French mathematician Paul Lévy. Roughly speaking, Lévy processes are stochastic processes whose increments in nonoverlapping time intervals are independent and stationary in time.

Definition 4. (Lévy process) A stochastic process $X = \{X_t\}_{t\in I}$ defined on $(\Omega, \mathfrak{F}, \mathbb{P})$ is called a *Lévy process* if the following properties hold.

1. X is right-continuous with left limits (càdlàg, from the French *continue à droite, limite à gauche*). For each t, we have the following.
   - The left limit $\lim_{s\uparrow t} X_s$ exists and is denoted with X_{t^-}.
   - The right limit $\lim_{s\downarrow t} X_s$ exists and is equal to X_t.
2. $\mathbb{P}(\{X_0 = 0\}) = 1$.
3. X is continuous in probability:
   $$\lim_{t\to s} X_t \overset{\mathbb{P}}{=} X_s,$$
   for each $t, s > 0$, in the sense of (2.2).
4. X has stationary increments:
   $$\text{for } 0 \le s \le t, \quad X_t - X_s \overset{d}{=} X_{t-s}.$$
   This means that the distribution of $X_t - X_s$ is invariant under time shifts $(s, t) \to (s + h, t + h)$.
5. X has independent increments:
   $$\text{for } 0 \le s \le t,\ X_t - X_s \text{ is independent of } \{X_u : u \le s\}.$$
   As a consequence, for any finite ordered sequence of times $0 \le t_1 < ... < t_n$ the random variables $X_{t_1} - X_0, ..., X_{t_{n-1}} - X_{t_n}$ are mutually independent.

While the first three properties are usual technical assumptions, the latter ones define the key features of a Lévy process X. For each $t \in \mathbb{R}$ and $m \in \mathbb{N}$, we can write the following equality

$$X_t = Y_t^1 + \ldots + Y_t^m,$$

where each Y_t^k is an increment defined as follows

$$Y_t^k = X_{\frac{kt}{m}} - X_{\frac{(k-1)t}{m}}, \quad k = 1, \ldots, m.$$

Thanks to the properties of a Lévy process as in Definition 4, for each $m \in \mathbb{N}$ the random variables $\{Y_t^k\}_{k=1}^m$ are independent and identical distributed. Therefore, for each arbitrary $t \in \mathbb{R}$, X_t is infinitely divisible and Theorem 1 holds.

In order to introduce the decomposition that characterizes a Lévy process, we focus on the discontinuous component of such process. To this end, it is useful to introduce the general notion of *random measure*.

Definition 5 (Random measure)**.** Given a σ-algebra $\mathcal{S}$ on a set Σ and a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, a *random measure* on $(\Sigma, \mathcal{S})$ is a collection of random variables $\Pi = \{\Pi(S)\}_{S \in \mathcal{S}}$ on $(\Omega, \mathfrak{F}, \mathbb{P})$ such that

1. $\Pi(\emptyset) = 0$;
2. (σ-additivity property) if $\{S_n\}_{n \in \mathbb{N}}$ is a collection of mutually disjoint subsets, then $\Pi(\cup_n S_n) = \sum_n \Pi(S_n)$ almost surely;
3. if $S_1, ..., S_n$ are mutually disjoint, the random variables $\Pi(S_1), ..., \Pi(S_n)$ are independent.

Definition 6 (Poisson random measure)**.** A *Poisson random measure* is a random measure Π on $(\Sigma, \mathcal{S})$ such that each $\Pi(S)$ has a Poisson distribution whenever $\Pi(S) < \infty$. The *intensity* of a Poisson random measure Π is the measure defined on $(S, \mathcal{S})$ defined by $\mathbb{E}[\Pi(S)]$, for each $S \in \mathcal{S}$.

We show that given a Lévy process X, we can construct a Poisson random measure on $(\mathbb{R}^d \backslash \{0\}, \mathcal{B}(\mathbb{R}^d \backslash \{0\}))$. The jump process $\Delta X = \{\Delta X_t\}_{t \in I}$ associated to a Lévy process X is defined as follows

$$\Delta X_t := X_t - X_{t^-}, \text{ whenever } X_{t^-} \neq X_t,$$

for each time $t \in I$. Therefore, given a Lévy process X, we can formally write the following

$$X_t = X_t^c + \sum_{s \leq t} \Delta X_s.$$

The process X_t^c has continuous sample paths, while the second term is discontinuous and gives all the jumps occurred up to time t. In fact, a path of X can be seen as a continuous random walk interspersed by jumps of random size.

Given a time $t \in I$ and a set $E \in \mathcal{B}(\mathbb{R}^d \backslash \{0\})$, the random measure of the jumps of X is defined as follows

$$\mu(t, E) := \sum_{s \leq t} \mathbb{1}_E(\Delta X_s) = \text{card}\{s \in [0, t] \text{ such that } \Delta X_s \in E\}. \tag{2.4}$$

In other words, μ counts the jumps of size belonging to E occurred up to time t. If we fix t and E, then $\mu(t, E)$ is a random variable. If we fix $\omega \in \Omega$ and $t \geq 0$, then $\mu(t, \cdot)$ is a set function $E \mapsto \mu(t, E)$ that defines a σ-finite measure on $\mathbb{R} \backslash \{0\}$. If we fix E, we note that $\mu(\cdot, E)$ has stationary and independent increments, thus it is a Poisson random measure. These premises allow us to state the Lévy-Itô decomposition theorem, which completely describes a Lévy process.

Theorem 2 (Lévy-Itô decomposition)**.** *If X is a Lévy process on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, then there exist $b \in \mathbb{R}^d$, a Brownian motion $W_A(t)$ with covariance matrix $A \in \mathbb{R}^{d\times d}$, and an independent Poisson random measure μ on $(\mathbb{R}^+, \mathbb{R}^d\backslash\{0\})$ as in* (2.4) *with intensity ν such that, for each $t \geq 0$, we have*

$$X_t = bt + W_A(t) + \int_{||x||<1} x\tilde{\mu}(t, dx) + \int_{||x||\geq 1} x\mu(t, dx),$$

where $\tilde{\mu}$ is the compensated Poisson random measure defined as follows

$$\tilde{\mu}(t, E) := \mu(t, E) - t\nu(E). \tag{2.5}$$

The Brownian motion $W_A(t) \in \mathbb{R}^d$ is defined as follows

$$W_A^i(t) := \sum_{j=1}^{m} \sigma_{ij} W^j(t),$$

where $W^1, ..., W^m$ are standard one-dimensional Wiener processes and $\sigma \in \mathbb{R}^{d\times m}$ satisfies $\sigma\,\sigma^T = A$.

By merging Theorem 2 with the Lévy-Khinthine formula valid for any infinitely divisible random variable, we can state that if X is a Lévy process, then for each $u \in \mathbb{R}^d$ and each $t \geq 0$, the characteristic function of X_t has the following form

$$\begin{aligned}\phi_{X_t}(u) &= \mathbb{E}[e^{iu\cdot X_t}] = \\ &= \exp\left\{t\left(ib\cdot u - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d\backslash\{0\}} [e^{iu\cdot y} - 1 - iu\cdot y\mathbb{1}_{||y||\leq 1}(y)]\nu(dy)\right)\right\},\end{aligned}$$

where ν is the intensity measure of the Poisson random measure μ. Therefore, the Lévy triplet (b, A, ν) completely describes a Lévy process X.

Proposition 1. *Let X be a Lévy process with triplet (b, A, ν). We have the following properties.*

1. *If $\nu(\mathbb{R}^d) < \infty$, X_t is said to have a* finite activity*: almost all paths of X have a finite number of jumps on every compact interval.*

2. *If $\nu(\mathbb{R}^d) = \infty$, X_t is said to have an* infinite activity*: almost all paths of X have an infinite number of jumps on every compact interval.*

Processes with infinite activity constitute a very rich class of jump processes; however, in this thesis, we consider processes with finite activity, as they are more easily tractable and find many applications.

2.1.2 Lévy processes as Markov processes

In this section, we investigate Lévy processes in the framework of Markov processes. A complete analysis about the link between Lévy processes and Markov processes can be found in, e.g., [6].

Let $B_b(\mathbb{R}^d)$ be the set of all functions $f : \mathbb{R}^d \to \mathbb{R}$ bounded and measurable, which is a Banach space with the ∞-norm. In what follows, we denote with $C_c^\infty(\mathbb{R}^d)$ all the infinitely differentiable functions with compact support. The Schwartz space of rapidly decreasing functions on $\mathbb{R}^d$ is defined as follows

$$S(\mathbb{R}^d) := \{f \in C^\infty(\mathbb{R}^d) \,|\, \|f\|_{\alpha,\beta} < \infty, \forall \alpha, \beta\}, \tag{2.6}$$

where α, β are multi-indices and

$$\|f\|_{\alpha,\beta} := \sup_{x\in\mathbb{R}^d} \left|x^\alpha \partial^\beta f(x)\right|.$$

Notice that the Fourier transform is a linear isomorphism of $S(\mathbb{R}^d)$ [6].

We say that a stochastic process $M = \{M_t\}_{t\in I}$ has the *Markov property* (or M is a *Markov process*) if for all $f \in B_b(\mathbb{R}^d)$ and each $s \in [0, t]$ the following equality holds:

$$\mathbb{E}[f(M_t) \,|\, \mathcal{F}_s] = \mathbb{E}[f(M_t) \,|\, M_s] \text{ almost surely.}$$

The *transition probability* $p_{s,t}(x, \cdot)$ of an arbitrary Markov process $M = \{M_t\}_{t\geq 0}$ is defined for each s, t with $0 \leq s \leq t$, $x \in \mathbb{R}$ and $E \in \mathcal{B}(\mathbb{R})$ as follows

$$p_{s,t}(x, E) := \mathbb{P}(M_t \in E \,|\, M_s = x).$$

For each Markov process M, we can define on $B_b(\mathbb{R}^d)$ a family of *evolution operators* $\{T_{s,t}\}_{0\leq s\leq t}$ as follows

$$(T_{s,t}f)(x) := \mathbb{E}[f(M_t) \,|\, M_s = x], \tag{2.7}$$

for each $x \in \mathbb{R}^d$. If $T_{s,t}(B_b(\mathbb{R}^d)) \subseteq B_b(\mathbb{R}^d)$, the Markov process X is said to be *normal*. If M is a normal Markov process, then its transitional probabilities are connected by the well-known Chapman-Kolmogorov equation, as follows. For each $0 \leq r \leq s \leq t < \infty$, $x \in \mathbb{R}^d$ and $E \in \mathcal{B}(\mathbb{R}^d)$, we have

$$p_{r,t}(x, E) = \int_{\mathbb{R}^d} p_{s,t}(y, E) p_{r,s}(x, dy). \tag{2.8}$$

If $T_{s,t} = T_{0,t-s}$, the Markov process M is said to be *(time-)homogeneous*. In this case, the operator $T_{0,t}$ is written as T_t and it constitutes a semigroup on $B_b(\mathbb{R}^d)$. The transition probability of a homogeneous Markov process at time t is written as p_t, since

$$p_{s,t}(x, E) = p_{0,t-s}(x, E) =: p_{t-s}(x, E),$$

for each $E \in \mathcal{B}(\mathbb{R}^d)$, and the following relation holds

$$(T_t f)(x) = \int_{\mathbb{R}^d} f(y) p_t(x, dy), \tag{2.9}$$

for each $t \geq 0$, $E \in \mathcal{B}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ and $f \in B_b(\mathbb{R}^d)$.

In general, when an arbitary semigroup $\{T_t\}_{t \in I}$ of operators is defined within a Banach space B with norm $\|\cdot\|_B$, it is possible to define the following linear set

$$D_{\mathcal{A}} := \left\{ \psi \in B \text{ such that } \exists \phi_\psi \in B \text{ that satisfies } \lim_{t \to 0} \left\| \frac{T_t \psi - \psi}{t} - \phi_\psi \right\|_B = 0 \right\}.$$

If such an element ϕ_ψ exists, then it is unique, and thus it is possible to define a linear operator $\mathcal{A} : \mathcal{D}_{\mathcal{A}} \to B$ as follows

$$\mathcal{A}\psi := \phi_\psi = \lim_{t \to 0} \frac{T_t \psi - \psi}{t}. \tag{2.10}$$

The operator $\mathcal{A}$ is called the *infinitesimal generator* of the semigroup $\{T_t\}_{t \in I}$.

In the context of homogeneous Markov processes and bounded and measurable functions belonging to the Banach space $B_b(\mathbb{R}^d)$, the family of operators $\{T_t\}_{t \in I}$ associated to the process M allows us to define the operator $\mathcal{A}$ as in (2.10), which is called the *generator of* M. Moreover, for each $f \in \mathcal{D}(\mathcal{A})$, $x \in \mathbb{R}^d$, $t \geq 0$, we define the following mapping

$$u : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$$

$$(x, t) \mapsto u(x, t) := (T_t f)(x).$$

From (2.10), it is possible to state the following initial-value problem on $\mathbb{R}^d \times \mathbb{R}^+$

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} = \mathcal{A} u(x, t) \\ u(x, 0) = f(x). \end{cases} \tag{2.11}$$

Assuming that for each $y \in \mathbb{R}$ the mapping $t \mapsto p_t(x, y)$ is differentiable and its derivative is uniformly bounded with respect to y, we obtain an equation that governs the time evolution of the transition probability of a homogeneous Markov process, as we show next.

Consider the right-hand side of the equation (2.11). On the one hand, for all $f \in C_c^\infty(\mathbb{R}^d)$, by using (2.9), we have

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial}{\partial t}(T_t f)(x) = \frac{\partial}{\partial t} \int_{\mathbb{R}^d} f(y) p_t(x, y) dy = \int_{\mathbb{R}^d} f(y) \frac{\partial p_t(x, y)}{\partial t} dy. \tag{2.12}$$

On the other hand, the left-hand side of the equation (2.11) can be recast as follows

$$\mathcal{A}u(t,x) = (T_t\mathcal{A}f)(x) = \int_{\mathbb{R}^d} (\mathcal{A}f)(y)p_t(x,y)dy = \int_{\mathbb{R}^d} f(y)\bar{\mathcal{A}}p_t(x,y)dy, \tag{2.13}$$

where $\bar{\mathcal{A}}$ acts on the variable y and denotes the adjoint of the operator $\mathcal{A}$.

Since (2.12) and (2.13) must coincide for each $f \in C_c^\infty(\mathbb{R}^d)$, we conclude that

$$\frac{\partial p_t(x,y)}{\partial t} = \bar{\mathcal{A}}p_t(x,y), \tag{2.14}$$

which is called the *forward Kolmogorov equation* or *Fokker-Planck (FP) equation.* Note that it operates on the "forward variables" t and y. An alternative equation is the so-called *Kolmogorov backward equation*, since it operates on the "backward variables" s and x as follows

$$\frac{\partial p_{t-s}(x,y)}{\partial s} = -\mathcal{A}p_{t-s}(x,y).$$

In [6] it is proven that each Lévy process X is a Markov process with respect to its natural filtration defined in (2.3). Moreover, X is both normal and homogeneous. Using the notion of generators of semigroups outlined above, we obtain a key tool for analyzing X. Its generator is completely determined by the Lévy-Khintchine formula. We have an explicit and compact form, which can be expressed as a pseudo-differential operator in terms of the Lévy triplet, as stated in the next theorem.

Theorem 3. *Let X be a Lévy process with Lévy triplet (b, A, ν), characteristic exponent ψ and evolution operator T_t. Recall the definition* (2.6)*. We denote with $\hat{f}$ the Fourier transform of a function $f \in S(\mathbb{R}^d)$.*

1. *For each $t \in I$, for each $f \in S(\mathbb{R}^d)$ and for each $x \in \mathbb{R}^d$, the evolution operator T_t of X is given by*

$$(T_t f)(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{iu\cdot x}\, e^{t\psi(u)} \hat{f}(u)du.$$

2. *For each $f \in S(\mathbb{R}^d)$ and for each $x \in \mathbb{R}^d$, the generator of X defined by Equation (2.10) is given by*

$$\begin{aligned}(\mathcal{A}f)(x) =& \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} e^{iux}\, \psi(u)\hat{f}(u)du = b\cdot\nabla f(x) + \frac{1}{2}\sum_{i,j} A_{ij}\frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \\ &+ \int_{\mathbb{R}^d} [f(x+y) - f(x) - y\cdot\nabla f(x)\mathbb{1}_{\{||y||\leq 1\}}(y)]\nu(dy).\end{aligned}$$

The form of the generator is completely determined by the Lévy triplet, with the following three correspondences:

$$\textit{drift component} \leftrightarrow \textit{first-order differential operator}$$

$$\textit{diffusion component} \leftrightarrow \textit{second-order differential operator}$$

$$\textit{jump component} \leftrightarrow \textit{integral operator.}$$

2.1.3 Examples of Lévy processes

In this section, we present of the most investigated Lévy processes. For each one, we summarize the properties introduced in the previous sections.

Standard Brownian motion

As introduced above, a Brownian motion $W = \{W_t\}_{t\geq 0}$ is a Lévy process, since it satisfies the properties in Definition (4). Moreover, a Brownian motion is continuous in probability, since it is locally Hölder continuous with exponent γ for every $\gamma \in (0, \frac{1}{2})$. This means that for every $s, t < \infty$ and for every $\omega \in \Omega$, there exists $K = K(\omega, s, t)$ such that

$$|W_t(\omega) - W_s(\omega)| \leq K|t - s|^{\gamma}.$$

The process $W = \{W_t\}_{t\geq 0}$ has the Lévy triplet $(0, I, 0)$, where I is the $d \times d$ identity matrix, and characteristic exponent given by

$$\psi(u) = -\frac{1}{2}|u|^2.$$

Its generator $\mathcal{A}$ takes the form

$$\mathcal{A} = \frac{1}{2}\Delta,$$

where Δ is the Laplacian operator, $\Delta := \sum_{j=1}^{d} \partial_j^2$.

Brownian motion with drift

Let $b \in \mathbb{R}^d$ and $\sigma \in \mathbb{R}^{d\times m}$ be a square root of A, which means that $A = \sigma\sigma^T \in \mathbb{R}^{d\times d}$ is symmetric and positive definite. If W is a standard Wiener process, the process that satisfies

$$dX_t := bdt + \sigma dW_t$$

is a Lévy process where each X_t is distributed as $\mathcal{N}(tb, tA)$. The process X has the Lévy triplet $(b, A, 0)$ and characteristic function and generator given by

$$\phi_{X_t}(u) = \exp\left\{t\left[ib \cdot u - \frac{1}{2}(u, Au)\right]\right\}$$

and

$$\mathcal{A} = b \cdot \nabla + \frac{1}{2} \sum_{i,j} A_{i,j} \frac{\partial^2}{\partial x_i \partial x_j},$$

respectively.

Poisson process

The Poisson process is an $\mathbb{N}$-valued continuous-time counting process $\{P_t\}_{t \in I}$ which is characterized by a rate parameter λ, also known as *intensity*, and possesses the following properties:

- $P_0 = 0$.
- (Independent increments) the numbers of occurrences counted in disjoint intervals are independent of each other.
- (Stationary increments) the probability distribution of the number of occurrences counted in any time interval only depends on the length of the interval.
- The probability distribution of the waiting time until the next occurrence is an exponential distribution.
- No counted occurrences are simultaneous.

The number of events in the time interval $(t, t+\tau]$ follows a Poisson distribution with associated parameter $\lambda\tau$:

$$\mathbb{P}[(P_{t+\tau} - P_t) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}, \qquad k = 0, 1, \ldots. \tag{2.15}$$

The Poisson process is a Lévy process since it satisfies the properties in Definition 4. It has the Lévy triplet $(0, 0, \lambda\delta_1)$, where δ_1 denotes the Dirac delta centered in $x = 1$.

The characteristic exponent of a Poisson process with intensity λ reads as follows

$$\psi(u) = \exp\left\{\lambda(e^{iu} - 1)\right\}.$$

and its generator is a difference operator as follows

$$(\mathcal{A}f)(x) = \lambda[f(x+1) - f(x)], \tag{2.16}$$

for each $f \in S(\mathbb{R}^d)$.

Compound Poisson process

Let $\{F_n\}_{n \in \mathbb{N}}$ be a sequence of independent identically distributed (i.i.d.) random variables with common law F and let P be an independent Poisson process with rate λ. The compound Poisson process Z defined as follows

$$Z_t := \sum_{j=1}^{P_t} F_j \tag{2.17}$$

is a Lévy process. The process Z has triplet $(0, 0, \lambda F)$ and characteristic exponent

$$\psi(u) = \lambda \int_{\mathbb{R}^d} (e^{iuy} - 1) F(dy).$$

Its generator takes the form

$$(\mathcal{A}f)(x) = \lambda \int_{\mathbb{R}^d} [f(x + y) - f(x)] F(dy). \tag{2.18}$$

Note that the Poisson process itself can be seen as a compound Poisson process such that $Y_j = 1$ for all $j \in \mathbb{N}$.

Alpha stable process

A real-valued random variable is called *stable* if it arises as a limit. Given a sequence $\{Y_n\}_{n\in\mathbb{N}}$ of random variables, a sequence of real numbers $\{\zeta_n\}_{n\in\mathbb{N}}$ and one of positive real numbers $\{\theta_n\}_{n\in\mathbb{N}}$, it is of interest the case where there exists a random variable Y such that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{Y_1 + ... + Y_n - \zeta_n}{\theta_n} \leq x\right) \stackrel{d}{=} \mathbb{P}(Y \leq x).$$

It can be shown that the only possible choice for θ_n is $\theta_n := \sigma n^{1/\alpha}$, with $\sigma > 0$ and $\alpha \in (0, 2]$. The parameter α is called *index of stability* and it can be shown that it always belongs to the interval $(0, 2]$. If each $\zeta_n = 0$, Y is called *strictly stable*.

A stable random variable Y is infinitely divisible. Therefore, given Y with index of stability α, thanks to Theorem 1 we can find a Lévy triplet (b, A, ν), which has one of the following forms

1. $\alpha = 2 \Rightarrow \nu = 0 \Rightarrow Y \sim \mathcal{N}(b, A)$;
2. $\alpha \neq 2 \Rightarrow A = 0$ and

$$\nu(x) = \frac{k_1}{x^{1+\alpha}} \mathbb{1}_{\{x>0\}}(x) + \frac{k_2}{|x|^{1+\alpha}} \mathbb{1}_{\{x<0\}}(x)$$

for some constants $k_1, k_2 \geq 0$ such that $k_1 + k_2 > 0$.

For each real-valued α-stable random variable there exist $\sigma > 0$, $\beta \in [-1, 1]$, $\mu \in \mathbb{R}$ such that the characteristic function of Y is given by

$$\Phi_Y(u) = \begin{cases} \exp\{i\mu u - \sigma^\alpha |u|^\alpha (1 - i\beta(\operatorname{sgn} u) \tan \frac{\pi\alpha}{2})\} & \text{if } \alpha \neq 1, 2 \\ \exp\{i\mu u - \sigma|u|(1 + i\beta\frac{2}{\pi}(\operatorname{sgn} u) \log|u|)\} & \text{if } \alpha = 1, \\ \exp\{i\mu u - \frac{\sigma^2}{2}u^2\} & \text{if } \alpha = 2, \end{cases}$$

where $\sigma \geq 0$ is the scale parameter, $\beta \in [-1, 1]$ is the skewness parameter, $\mu \in \mathbb{R}$ is the shift parameter. A stable random variable with this parametrization is denoted by $S_\alpha(\sigma, \beta, \mu)$.

While algorithms for simulating an α-stable random variable exist for all values of the parameters, the probability density function $f(x)$ of an α-stable random variable is not known in closed form except for the following three cases:

1. Normal distribution $S_2(\sigma, 0, \mu) \sim \mathcal{N}(\mu, \sigma^2)$;
2. Cauchy distribution $S_1(\sigma, 0, \mu)$, with density
$$f(x) = \frac{\sigma}{\pi[(x-\mu)^2 + \sigma^2]};$$
3. Lévy distribution $S_{1/2}(\sigma, 1, \mu)$, with density
$$f(x) = \left(\frac{\sigma}{2\pi}\right)^{1/2} \frac{1}{(x-\mu)^{3/2}} \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\} \text{ for } x > \mu.$$

Given an α-stable random variable Y, it is possible to construct a Lévy process $X = \{X_t\}_{t \in I}$ such that $X_{\{t=1\}} \stackrel{d}{=} Y$.

Subordinators

A subordinator is a one-dimensional Lévy process $S = \{S_t\}_{t \in I}$ that is almost surely nondecreasing. S can be thought as a random model for time evolution. In fact, given an arbitrary Lévy process X and an independent subordinator S, the process Z defined as follows

$$Z_t := X_{S_t} \text{ for each } t \in I$$

is also a Lévy process.

The Lévy symbol of a subordinator S takes the following form

$$\psi(u) = ibu + \int_{\mathbb{R}^+} (e^{iuy} - 1)\nu(dy),$$

where b is nonnegative and the Lévy measure ν satisfies

$$\nu(-\infty, 0) = 0 \qquad \text{and} \qquad \int_0^\infty \min(y, 1)\nu(dy) < \infty.$$

An example of subordinator is a compound Poisson process with nonnegative jump distribution. It can be shown [6] that for each $\alpha \in (0, 1)$, there exists an alpha-stable subordinator.

2.2 Stochastic differential equations

In this section, we introduce SDEs. We consider examples driven by Brownian motion and an independent Poisson random process. These processes represent generalizations of Lévy processes, since the drift b and the diffusion σ may depend on the process itself and may explicitely depend on time. The presence of the term governed by the Poisson process makes the sample paths of the solution of the considered initial-value problem discontinuous.

Definition 7 (Itô integral)**.** Given a Wiener process W and an $\mathcal{F}$-adapted process $H = \{H_t\}_{t\in I}$ such that $\int_0^t \mathbb{E}[H_s^2]ds < \infty$, the Itô integral of H with respect to W is the random variable defined as follows

$$\int_0^t H(s,\omega)dW(s,\omega) := \text{m.s.}\lim_{n\to\infty}\sum_{i=1}^{n-1} H(s_i,\omega)(W(s_{i+1},\omega) - W(s_i,\omega)), \tag{2.19}$$

where the limit is intended as in (2.1).

An alternative definition of the integral with respect to a Brownian motion is the Stratonovich integral. A complete discussion about stochastic integration can be found in, e.g., [6].

Consider the time interval $I = [0,T]$. The focus of this section is the following stochastic initial-value problem

$$\begin{cases} dX_t = b(X_{t^-},t)dt + \sigma(X_{t^-},t)dW_t + \int_{||q||<c} F(X_{t^-},t,q)\tilde{\mathcal{P}}(dt,dq) + \\ \qquad\quad + \int_{||q||\geq c} G(X_{t^-},t,q)\mathcal{P}(dt,dq), \\ X_{\{t=0\}} = X_0. \end{cases} \tag{2.20}$$

The functions $b : \mathbb{R}^d\times\mathbb{R}^+ \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d\times\mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$, $F : \mathbb{R}^d\times\mathbb{R}^+\times\mathbb{R}^d \to \mathbb{R}^d$, $G : \mathbb{R}^d \times \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$ are assumed to be measurable. $W = \{W_t\}_{t\in I}$ is an m-dimensional Wiener process and $\mathcal{P}$ is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}^d$ with intensity ν and the process $\tilde{\mathcal{P}}$ is the compensator associated to $\mathcal{P}$. The parameter $c \in [0,\infty]$ is used to specify what is meant by "small" or "big" jumps.

It is natural to ask whether or not (2.20), given a known initial random variable X_0, has a solution, if this solution is unique and which are its properties. When the processes W and $\mathcal{P}$ are given in advance, and therefore used to construct the solution X, then X is called *strong solution*. On the other hand, if only the deterministic functions in (2.20) are given and one is asked to provide $(\{X_t\}_{t\in I}, \{W_t\}_{t\in I}, \mathcal{P})$ such that (2.20) is satisfied, then the process X is called *weak solution*. A strong solution is also a weak solution, but the converse is not true in general [6, 73].

A strong solution $\{X_t\}_{t\in I}$ is said to be *pathwise unique* if for any other stochastic process $\{\bar{X}_t\}_{t\in I}$ satisfying (2.20), we have that

$$\mathbb{P}(\{\omega \in \Omega \,|\, X_t(\omega) = \bar{X}_t(\omega) \text{ for all } t\}) = 1.$$

Let us define

$$a(x, y, t) := \sigma(x, t)\sigma(y, t)^T \in \mathbb{R}^{d\times d} \tag{2.21}$$

and define the matrix seminorm as follows

$$\|a\|_d := \sum_{i=1}^{n} |a_{ii}|.$$

Under specific assumptions on the coefficients b, σ, F and G, the existence and the pathwise uniqueness of a strong solution to (2.20) are ensured. For the details regarding the proof, see [6, 56, 73].

Theorem 4. *Let $(\Omega, \mathfrak{F}, \{\mathcal{F}_t\}_{t\in I}, \mathbb{P})$ be a filtered probability space. Let b, a, F, G be measurable functions satisfying the following conditions.*

1. *(Lipschitz condition) There exists $K_1(t) > 0$ such that, for all $x_1, x_2 \in \mathbb{R}^d$ and $t \in I$, we have that*

$$||b(x_1, t) - b(x_2, t)||^2 + \|a(x_1, x_1, t) - 2a(x_1, x_2, t) + a(x_2, x_2, t)\|_d + \tag{2.22}$$

$$\int_{||q||<c} ||F(x_1, t, q) - F(x_2, t, q)||^2 \nu(dq) \leq K_1(t)||x_1 - x_2||^2. \tag{2.23}$$

2. *(Growth condition) There exists $K_2(t) > 0$ such that, for all $x \in \mathbb{R}^d$ and $t \in I$,*

$$||b(x, t)||^2 + \|a(x, x, t)\|_d + \int_{||q||<c} ||F(x, t, q)||^2 \nu(dq) \leq K_2(t)(1 + ||x||^2).$$

3. *The mapping $q \to G(x, t, q)$ is continuous.*

Let X_0 be a random variable known up to a $\mathbb{P}$-null set and let W and $\mathcal{P}$ be a Brownian motion and a Poisson random measure with intensity ν, respectively. Let us assume that W and $\mathcal{P}$ are independent of $\mathcal{F}_0$. Then the initial-value problem (2.20) *unique right-continuous with left limit $\mathcal{F}$-adapted solution.*

Moreover, if $\mathbb{E}[||X_0||^2] < \infty$, there exists a time-dependent quantity $D(t) > 0$ such that

$$\mathbb{E}[||X_t||^2] \leq D(t)(1 + \mathbb{E}[||X_0||^2]).$$

Definition 8 (Itô diffusion)**.** A stochastic process $X = \{X_t\}_{t\in I}$ is an *Itô diffusion* or *diffusion process* if it satisfies the following initial-value problem

$$\begin{cases} dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t \\ X_{\{t=0\}} = X_0, \end{cases} \tag{2.24}$$

where b and σ are known measurable functions, $W = \{W_t\}_{t\in I}$ is a Wiener process defined on $(\Omega, \mathfrak{F}, \mathbb{P})$ and $X_0 \in \mathbb{R}^d$ is a known random variable.

Definition 9 (Jump-diffusion process)**.** A stochastic process $X = \{X_t\}_{t\in I}$ is a *jump-diffusion process* if it satisfies the following initial-value problem

$$\begin{cases} dX_t = b(X_{t^-}, t)dt + \sigma(X_{t^-}, t)dW_t + \int_{\mathbb{R}^d\setminus\{0\}} G(X_{t^-}, t, q)\mathcal{P}(dt, dq), \\ X_{\{t=0\}} = X_0, \end{cases} \tag{2.25}$$

where the intensity measure ν of the Poisson random measure $\mathcal{P}$ is finite.

We remark that Itô included jumps in his discussion [53], but the process (2.24) named after him does not include any jump. Moreover, there is no universal agreement about the use of the name *jump-diffusion process*. For (2.25), we refer to the definition given by [6], while in literature the same name denotes a process solving the more general model (2.20).

The FP equation, already introduced in (2.14) for models with constant drift and diffusion coefficients, also holds for diffusion processes solving (2.24). In this case, we have the following definition for the infinitesimal generator $\mathcal{A}$ formally defined in (2.10). Set $X_0 = x \in \mathbb{R}^d$ in (2.24). The operator $\mathcal{A}$ takes the following form, see, e.g., [73]

$$\mathcal{A}\phi(x) = \sum_{i=1}^{d} b_i(x,t)\frac{\partial\phi(x)}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^T)_{i,j}(x,t)\frac{\partial^2\phi(x)}{\partial x_i\partial x_j}, \tag{2.26}$$

for each $\phi \in \mathcal{C}_0^2(\mathbb{R}^d)$. Denote with $\langle\cdot,\cdot\rangle$ the L^2 inner product. The adjoint operator $\bar{\mathcal{A}}$ of (2.26) is defined as follows

$$\langle\mathcal{A}\phi, \psi\rangle = \langle\phi, \bar{\mathcal{A}}\psi\rangle,$$

for each $\phi \in \mathcal{C}_0^2(\mathbb{R}^d)$, $\psi \in \mathcal{C}^2(\mathbb{R}^d)$. After integrating by parts, we find the following form for $\bar{\mathcal{A}}$

$$\bar{\mathcal{A}}\psi(x) = -\sum_{i=1}^{d}\frac{\partial}{\partial x_i}(b_i(x,t)\psi(x)) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}((\sigma\sigma^T)_{i,j}(x,t)\psi(x)), \tag{2.27}$$

for each $\psi \in \mathcal{C}^2(\mathbb{R}^d)$. As discussed in Section 2.1.2, the transitional probability $p_t(x,y)$ of (2.24) solves the FP equation

$$\partial_t p_t(x,y) = \bar{\mathcal{A}}p_t(x,y), \tag{2.28}$$

where $\bar{\mathcal{A}}$ is defined in (2.27) and acts on the variable y. The fundamental solution of the FP equation is obtained by solving (2.28) with the following initial condition

$$p_{\{t=0\}}(x, x_0) = \delta(x - x_0),$$

where δ stands for the Dirac delta. However, in order to investigate a stochastic process, we consider the PDF $f(x,t)$, which is obtained from the transitional probability p_t by averaging over the initial conditions as follows

$$f(x,t) = \int_{\mathbb{R}^d} p_{\{t=0\}}(x,x_0)p_t(x,x_0)dx_0 = \int_{\mathbb{R}^d} \delta(x-x_0)p_t(x,x_0)dx_0. \tag{2.29}$$

By taking the time derivative of (2.29), we have the following

$$\partial_t f(x,t) = \int_{\mathbb{R}^d} p_{t=0}(x,x_0)\partial_t p_t(x,x_0)dx_0.$$

Moreover, by using (2.28) we get

$$\partial_t f(x,t) = \int_{\mathbb{R}^d} \delta(x-x_0)\bar{\mathcal{A}}p_t(x,x_0)dx_0 = \bar{\mathcal{A}}f(x,t), \tag{2.30}$$

where $\bar{\mathcal{A}}$ acts on the variable x. Therefore, the time evolution of the PDF of (2.24) is governed by the following FP equation

$$\partial_t f(x,t) = \bar{\mathcal{A}}f(x,t). \tag{2.31}$$

Next, we consider a process doing jumps. We show how the integral term of the FP equation related to (2.25) arises. We take $b=0$, $\sigma=0$, $G(x,t,q)=q$ in (2.25). Therefore, the last addend in (2.25) can be rewritten as follows

$$dP_t = \int_{\mathbb{R}^d\setminus\{0\}} q\,\mathcal{P}(dt,dq),$$

where P is a compound Poisson process (2.17) with rate of jumps $\lambda \in \mathbb{R}^+$ and jump PDF g. We first investigate the case $g(x) := \delta(x-1)$, which corresponds to the Poisson process (2.15). Consider the interval $(t, t+\tau]$. The cumulative distribution function (CDF) F at time t is given, while $F(x,t+\tau)$ can assume the following values

$$F(x,t+\tau) = \begin{cases} \lambda\tau F(x-1,t) \\ \\ (1-\lambda\tau)F(x,t), \end{cases}$$

depending on whether or not a jump occurs in the interval $(t,t+\tau]$, respectively. Therefore, we have

$$F(x,t+\tau) = \lambda\tau F(x-1,t) + (1-\lambda\tau)F(x,t).$$

By computing the time derivative of the CDF, we obtain

$$\partial_t F(x,t) = \lim_{\tau\to 0} \frac{F(x,t+\tau)-F(x,t)}{\tau} = \lambda F(x-1,t) - \lambda F(x,t),$$

and therefore the PDF $f(x,t)$ satisfies the following equation

$$\partial_t f(x,t) = \lambda f(x-1,t) - \lambda f(x,t). \tag{2.32}$$

An analogous computation can be carried out in case of a compound Poisson process (2.17) with rate $\lambda \in \mathbb{R}^+$ and arbitrary jump distribution with PDF g. We have the following

$$F(x,t+\tau) = \lambda\tau \int_{\mathbb{R}} F(x-y,t)g(y)dy + (1-\lambda\tau)F(x,t).$$

Proceeding as before, we obtain the following relation

$$\partial_t F(x,t) = \lambda \int_{\mathbb{R}} F(x-y,t)g(y)dy - \lambda F(x,t).$$

Therefore, the time evolution of the PDF $f(x,t)$ is governed by the following equation

$$\partial_t f(x,t) = \lambda \int_{\mathbb{R}} f(x-y,t)g(y)dy - \lambda f(x,t). \tag{2.33}$$

We remark the fact that the left-hand sides in (2.32) and (2.33) correspond to the adjoint operators of (2.16) and (2.18), respectively.

2.3 Numerical solution of stochastic differential equations

Stochastic models in the general form (2.20) govern the evolution of a state in presence of random noise and arise in a multitude of application areas. Under specific assumptions, the solution of a SDE is unique, as discussed in Section 2.2. However, a closed form of the solution is available only in a limited number of cases, whose list can be found in, e.g., [56]. For this reason, one has to rely on numerical methods that approximate the solution. We focus on the Euler-Maruyama (EM) method and we provide an existing result stating its strong order of accuracy in the spirit of [49, 64].

We consider (2.25) and choose $G(x,t,q) = q$. Assume that the hypotheses of Theorem 4 are fulfilled. Given an initial data X_0, we have the following initial-value problem in I

$$\begin{cases} dX_t = b(X_{t^-},t)dt + \sigma(X_{t^-},t)dW_t + dP_t \\ X_{\{t=0\}} = X_0, \end{cases} \tag{2.34}$$

where P is a compound Poisson process with rate $\lambda \in \mathbb{R}^+$ and jump PDF g. Our aim is to simulate sample paths of the solution to (2.34) on a time discrete grid. For $N \in \mathbb{N}$, define the mesh size $h := T/N$ and the following time discretization

$$I_h := \{t_k = hk,\, k = 0, ..., N\} \subset I.$$

The values of a sample path of (2.34) are specified at the points of the time grid I_h.

Let $X(t_k)$ and $\bar{X}(t_k)$ be the values of the analytic solution to (2.34) and the value of the numerical solution corresponding to the time t_k, respectively. We say that the *strong mean-square rate of convergence* or *strong mean-square order of accuracy* of the numerical method is p if the following holds

$$(\mathbb{E}[(X(t_k) - \bar{X}(t_k))^2])^{1/2} \leq Ch^p,$$

where C is a constant not depending on h.

However, in many problems it is only required to evaluate an expectation such as $\mathbb{E}[\zeta(X(t_k))]$, for a sufficiently large class of functions ζ. Hence, the following definition of *weak rate of convergence* or *weak order of accuracy* is also useful. We say that the weak rate of convergence of a numerical method is p if the following inequality holds

$$|\mathbb{E}[\zeta(X(t_k))] - \mathbb{E}[\zeta(\bar{X}(t_k))]| \leq Ch^p.$$

where C is a constant not depending on h.

In this thesis, we focus on the strong mean-square rate of convergence of the EM method to the initial-value problem (2.34). However, a number of results on the weak convergence of numerical schemes solving (2.20) is available; see, e.g., [56, 76]. The EM method is the generalization to SDEs of the Euler method for deterministic ordinary differential equations. In fact, when a SDE does not contain any randomness, the EM method corresponds to the explicit Euler method for deterministic equations.

Given the uniform stepsize $h := T/N$, the EM method is recursively defined for each $k = 0, ..., N-1$ as follows

$$\bar{X}(t_{k+1}) = \bar{X}(t_k) + b(\bar{X}(t_k), t_k)h + \sigma(\bar{X}(t_k), t_k)\Delta W_{k+1} + \Delta P_{k+1}, \tag{2.35}$$

where ΔW_{k+1} and ΔP_{k+1} denote the increments of the Wiener and Poisson processes over $(t_k, t_{k+1}]$.

Next, we present how to simulate the random increments in (2.35) over the time grid $t_0, ..., t_N$. We also show how an α-stable process introduced in Section 2.1.3 can be simulated.

Discretized Brownian motion

Exploiting the properties of the Brownian motion W, we have that for each time t_k, $W_{t_k} = W_{t_{k-1}} + \Delta W_{t_k}$, where the increment ΔW_{t_k} over the interval $[t_{k-1}, t_k]$ is distributed as $\mathcal{N}(0, (t_k - t_{k-1})^2)$.

We set $W_0 = 0$ and generate the increments ΔW_{t_k} for each k. The value of W in t_k is given by the cumulative sum of the increments up to time t_k as follows

$$W_{t_k} = \sum_{j=1}^{k} \Delta W_j.$$

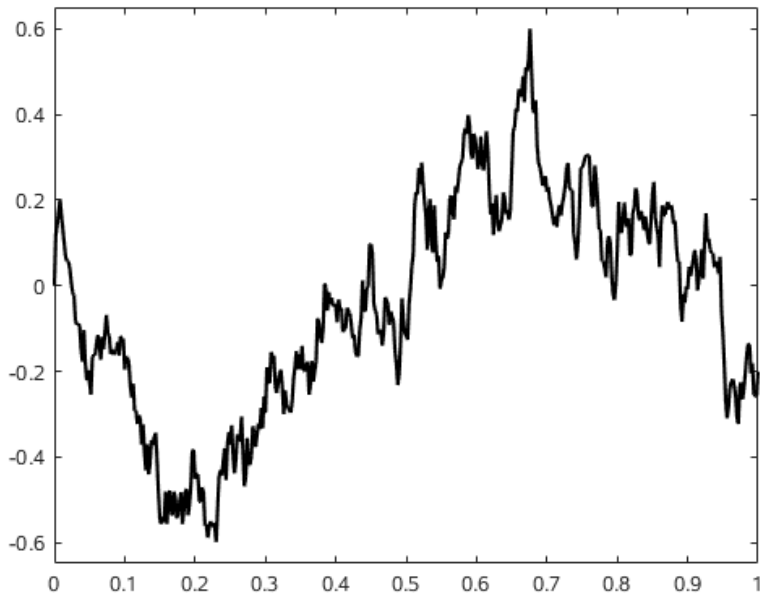


Figure 2.1: Sample path of a standard Brownian motion in the time interval $[0, 1]$ with $N = 500$ time steps.

Discretized compound Poisson process

Recall (2.17). A compound Poisson process X in the time interval $[0, T]$ is defined as follows

$$X_t = \sum_{j=1}^{P_t} F_j,$$

where P is a Poisson process with rate λ and $\{F_j\}_{j\in\mathbb{N}}$ is a sequence of i.i.d. random variables with common distribution F.

Unlike the Brownian motion, it is possible to sample an exact trajectory of X in any closed interval, assuming that it is possible to sample the jump sizes from the distribution F. In fact, the path is piecewise constant between two consecutive jump times and it contains a finite number of jumps distributed as F.

A discretized trajectory of X in $[0, T]$ can be obtained with the following steps:

1. generation of the total number of jumps in $[0, T]$, which is a Poisson random variable P with parameter λT;
2. generation of the P jump times $\tau_1 < ... < \tau_P$ uniformly distributed in $[0, T]$;
3. generation of P jump sizes $F_1, ..., F_P$, which all have common law F.

The discretized trajectory X is computed as follows

$$X_{t_k} = \sum_{j=1}^{P} \mathbb{1}_{\{\tau_j < t_k\}} F_j.$$

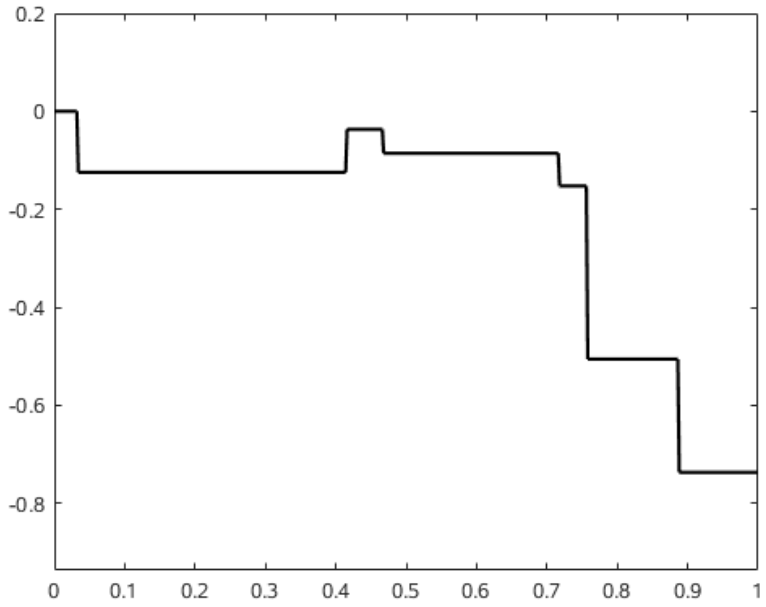


Figure 2.2: Sample path of a compound Poisson process in the time interval $[0, 1]$ with $\lambda = 5$ and $F \sim \mathcal{N}(0, 0.2^2)$.

Discretized α-stable process

We write $Z \sim S_\alpha(\sigma, \beta, \mu)$ when X is an α-stable random variable with skewness parameter β, scale parameter σ and shift parameter μ. We recall some results [8, 20, 23].

- If $Z \sim S_\alpha(\sigma, \beta, \mu)$ and $a \in \mathbb{R}$, then $Z + a \sim S_\alpha(\sigma, \beta, \mu + a)$ and $aZ \sim S_\alpha(|a|\sigma, \text{sgn}(a)\beta, \mu)$.

- If $Z_1, Z_2 \sim S_\alpha(1, 0, 0)$ and independent, then

$$Z := \mu + \sigma \left(\frac{1+\beta}{2}\right)^{1/\alpha} Z_1 - \sigma \left(\frac{1-\beta}{2}\right)^{1/\alpha} Z_2$$

  has an $S_\alpha(\sigma, \beta, \mu)$ distribution.

- If $Z \sim S_\alpha(1, \beta, 0)$, then

$$\sigma Z + \mu \sim S_\alpha(\sigma, \beta, \mu) \qquad \text{if } \alpha \neq 1$$

$$\sigma Z + \mu + \frac{2}{\pi}\beta\sigma \ln \sigma \sim S_\alpha(\sigma, \beta, \mu) \qquad \text{if } \alpha = 1$$

  has an $S_\alpha(\sigma, \beta, \mu)$ distribution.

- In general, if $Z_i \sim S_\alpha(\sigma_i, \beta_i, \mu_i)$ for $i = 1, ..., n$, then $\sum_{i=1}^n Z_i \sim S_\alpha(\sigma, \beta, \mu)$, with

$$\sigma = \left(\sum_{i=1}^n \sigma_i^\alpha\right)^{1/\alpha}, \ \beta = \frac{\sum_{i=1}^n \beta_i \sigma_i^\alpha}{\sum_{i=1}^n \sigma_i^\alpha} \text{ and } \mu = \sum_{i=1}^n \mu_i.$$

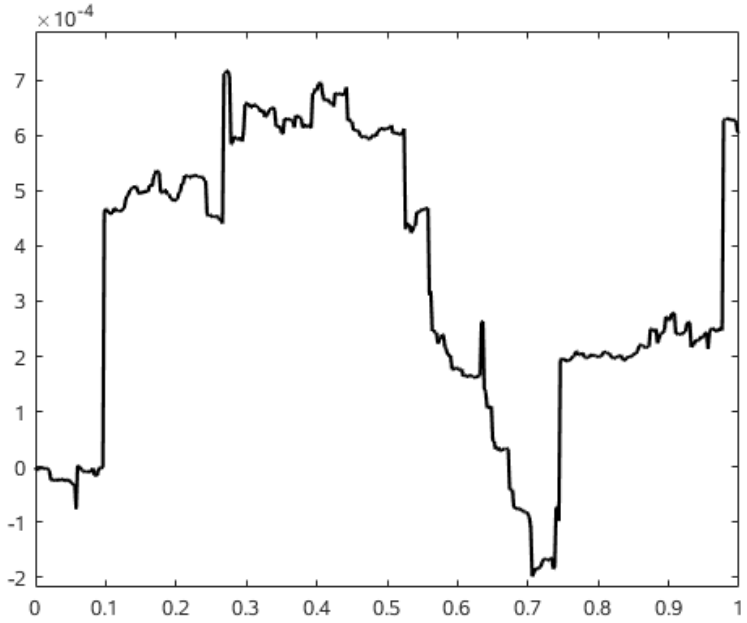


Figure 2.3: Sample path of an α-stable process in the time interval $[0, 1]$ with 500 time steps, with $\alpha = 0.5$, $\beta = 1$, $\sigma = 0$ and $\mu = 0$.

An algorithm for simulating a stable process X such that $X_1 \sim S_\alpha(0, 1, 0)$ is available in [20, 23]. First, we simulate $U_1, ..., U_N$ independent random variables, uniformly distributed in $[-\pi/2, \pi/2]$ and $E_1, ..., E_N$ independent standard exponential random variables. Each increment ΔX_k for $k = 1, ..., N$, is

given by

$$\Delta X_k = (t_k - t_{k-1})^{1/\alpha} \frac{\sin(\alpha U_k)}{(\cos U_k)^{1/\alpha}} \left(\frac{\cos[(1-\alpha)U_k]}{E_k} \right)^{(1-\alpha)/\alpha}.$$

The value of X at time t_k is given by the cumulative sum up the i-th increment as follows

$$X_{t_k} = \sum_{i=1}^{k} \Delta X_i.$$

Next, we state the strong convergence rate of (2.35). We define the step function Z as follows

$$Z(t) := \sum_{k=0}^{N-1} \bar{X}_k \, \mathbb{1}_{[kh,(k+1)h)}(t), \tag{2.36}$$

so that $Z(t_k) = \bar{X}_k$ for each $k = 0, ..., N$. We then define the function Y, which extends the numerical solution to the whole interval I as follows

$$Y(t) := X_0 + \int_0^t b(Z(s), s)ds + \int_0^t \sigma(Z(s), s)dWs + \int_0^t dP_s. \tag{2.37}$$

Theorem 5. *Let the SDE* (2.34) *have Lipschitz coefficients* b *and* σ*, as in* (2.22)*. Then, the EM method defined by* (2.35) *has strong mean-square rate of convergence equal to* $\frac{1}{2}$*. There exists* $C > 0$ *such that the following holds*

$$\mathbb{E}\left[\sup_{t \in I} ||X(t) - Y(t)||^2 \right] \leq C \left(1 + \mathbb{E}[||X_0||^2]\right) h,$$

where X *is the solution to* (2.34) *and* Y *is defined in* (2.37)*.*

Proof. We first prove that

$$\mathbb{E}\left[\sup_{t \in [0,T]} ||Y(t) - Z(t)||^2 \right] \leq C(1 + \mathbb{E}[||X_0||^2])h. \tag{2.38}$$

Consider $t \in [kh, (k+1)h] \subset I$. We have

$$Y(t) - Z(t) = Y(t) - Y_k = \int_{kh}^t b(Z(s), s)ds + \int_{kh}^t \sigma(Z(s), s)dW_s + \int_{kh}^t dP_s.$$

Therefore,

$$\mathbb{E}\left[\sup_{t \in [0,T]} ||Y(t) - Z(t)||^2 \right] \leq$$

$$\max_k \sup_{\tau \in [kh,(k+1)h]} \mathbb{E}\left[C \left(\left\| \int_{kh}^{\tau} b(Z(s), s)ds \right\|^2 + \left\| \int_{kh}^{\tau} \sigma(Z(s), s)dW_s \right\|^2 + \left\| \int_{kh}^{\tau} dP_s \right\|^2 \right) \right]. \tag{2.39}$$

We apply the Cauchy-Schwarz inequality to the first term in (2.39) and the Itô isometry [6] to the second term in (2.39), respectively. Since $\tau \in [kh, (k+1)h]$, we get

$$\mathbb{E}\left\|\int_{kh}^{\tau} b(Z(s), s)ds\right\|^2 \leq h\mathbb{E}\left[\int_{kh}^{(k+1)h} ||b(s, Z(s))||^2 ds\right]$$

and

$$\mathbb{E}\left\|\int_{kh}^{\tau} \sigma(Z(s), s)dW_s\right\|^2 \leq \int_{kh}^{(k+1)h} \mathbb{E}\left[||\sigma(s, Z(s))||^2\right] ds.$$

Note that the Lipschitz condition (2.22) implies

$$||b(x,t)||^2 \leq L(1 + ||x||^2)$$

$$||\sigma(x,t)||^2 \leq L(1 + ||x||^2).$$

Therefore,

$$\mathbb{E}\left\|\int_{kh}^{\tau} b(Z(s), s)ds\right\|^2 \leq h\mathbb{E}\left[\int_{kh}^{(k+1)h} L(1 + ||Z(s)||^2)ds\right] = h\int_{kh}^{(k+1)h} L(1+\mathbb{E}||Z(s)||^2)ds$$

and

$$\mathbb{E}\left\|\int_{kh}^{\tau} \sigma(Z(s), s)dW_s\right\|^2 \leq \int_{kh}^{(k+1)h} \mathbb{E}\left[L(1 + ||Z(s)||^2)\right] ds = \int_{kh}^{(k+1)h} L(1+\mathbb{E}||Z(s)||^2)ds,$$

thanks to Fubini's theorem. Moreover, we decompose P as the sum of its compensator $\tilde{P}_t$ and a deterministic term. We have

$$\left\|\int_{kh}^{\tau} dP_s\right\|^2 \leq C\left(\left\|\int_{kh}^{\tau} d\tilde{P}_s\right\|^2 + \left\|\int_{kh}^{\tau} \lambda ds\right\|^2\right).$$

Taking the expectation, we obtain

$$\mathbb{E}\left[\left\|\int_{kh}^{t} dP_s\right\|^2\right] \leq C\left(\lambda h + \lambda^2 h^2\right).$$

Therefore, we recast (2.39) as follows

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||Y(t) - Z(t)||^2\right] \leq$$
$$\max_k \sup_{\tau\in[kh,(k+1)h]} C\left\{hL\int_{kh}^{(k+1)h} (1 + \mathbb{E}[||Z(s)||^2])ds + \lambda h(1 + \lambda h)\right\}.$$

Note that $Z(s) = \bar{X}_k$ on $[kh, (k+1)h)]$, thus

$$\mathbb{E}\left[\sup_{t\in[0,T]} ||Y(t) - Z(t)||^2\right] \leq$$
$$\max_k \sup_{\tau\in[kh,(k+1)h]} C\left\{hL\int_{kh}^{(k+1)h}(1+\mathbb{E}||\bar{X}_k||^2)ds + \lambda h(1+\lambda h)\right\}.$$

The claim (2.38) follows after noting that

$$\mathbb{E}||\bar{X}_k||^2 \leq C(1+\mathbb{E}||X_0||^2)$$

for each k, which follows from arguments similar to the discussion above.

Next, we consider the term $\mathbb{E}\left[\sup_{t\in I}||X(t) - Y(t)||^2\right]$. By the definition (2.37), we have

$$Y(t) - X(t) = \int_0^t [b(Z(s), s) - b(X(s), s)]ds + \int_0^t [\sigma(Z(s), s) - \sigma(X(s), s)]dW_s.$$

Therefore, we have

$$\mathbb{E}\left[\sup_{t\in I}||X(t) - Y(t)||^2\right] \leq C\left(\mathbb{E}\left[\sup_{t\in I}||\int_0^t [b(Z(s), s) - b(X(s), s)]ds||^2\right] + \mathbb{E}\left[\sup_{t\in I}||\int_0^t [\sigma(Z(s), s) - \sigma(X(s), s)]ds||^2\right]\right).$$

By applying the Cauchy-Schwarz inequality and Fubini's theorem to the first term and the Itô isometry to the second term, respectively, we have

$$\mathbb{E}\left[\sup_{t\in I}||X(t) - Y(t)||^2\right] \leq C\left(T\int_0^{t_1}\mathbb{E}||b(Z(s), s) - b(X(s^-), s)||^2 ds + \int_0^{t_1}\mathbb{E}\left[||\sigma(Z(s), s) - \sigma(X(s^-), s)||^2\right]ds\right).$$

After applying the Lipschitz condition (2.22), up to a redefinition of the constant C, we have

$$\mathbb{E}\left[\sup_{t\in I}||X(t) - Y(t)||^2\right] \leq C\left(T\int_0^{t_1}\mathbb{E}||Z(s) - X(s^-)||^2 ds + \int_0^{t_1}\mathbb{E}\left[||Z(s) - X(s^-)||^2\right]ds\right).$$

After redefining the constant C, we have

$$\mathbb{E}\left[\sup_{t\in I}||X(t) - Y(t)||^2\right] \leq C\int_0^{t_1}\mathbb{E}||Z(s) - Y(s)||^2 ds \leq$$
$$\leq C\int_0^{t_1}(\mathbb{E}||Z(s) - X(s^-)||^2 + \mathbb{E}||X(s^-) - Y(s)||^2)\, ds.$$

We use (2.38) and we get

$$\mathbb{E}\left[\sup_{t\in[0,t_1]}||X(t)-Y(t)||^2\right] \leq C_1 h(1+\mathbb{E}[||Y(0)||^2])+ \\ + C_2\int_0^{t_1}\mathbb{E}\sup_{t\in[0,s]}||Y(t)-X(t^-)||^2 ds.$$

The claim follows from the Gronwall inequality.

□

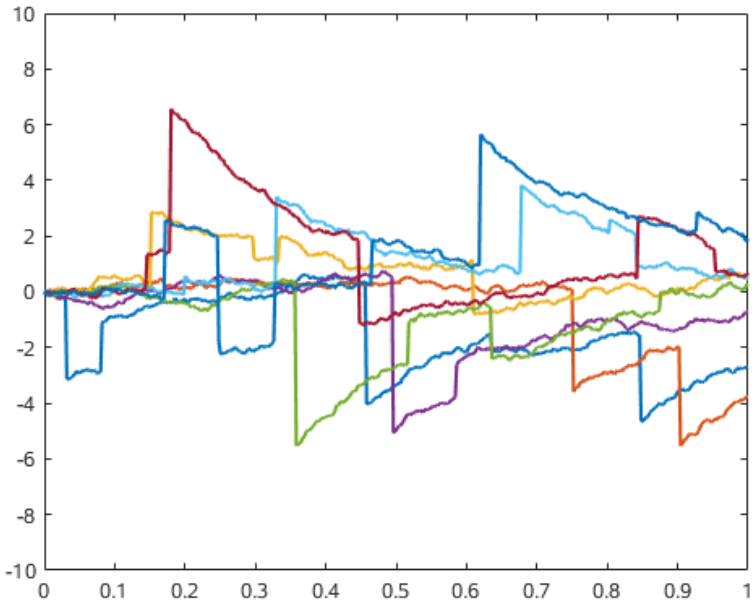


Figure 2.4: Plot of 8 sample paths of (2.34) with $X_0 = 0$, $b(x,t) = \frac{1}{1+x^2} - 4x$, $\sigma(x,t) = 1$, $\lambda = 5$ and $g \sim \mathcal{N}(0, 3^2)$ by applying the EM method in the interval $[0,1]$ with $N = 500$.

2.4 The empirical probability density function

In the last section, we computed sample paths of a JD process by applying the EM method. In this subsection, we illustrate a method that computes the empirical PDF of a one-dimensional JD process as in (2.34) with range in $D \subset \mathbb{R}$, $D := (r, s)$ by making use of the EM method.

The empirical PDF $\hat{f}(x,t)$ approximates the continuous PDF $f(x,t)$ of the solution to our SDE (2.34), as we illustrate next. First, the time t has to be set. Then the number of sample paths $R \in \mathbb{N}$ has to be chosen, with $R >> 1$. By means of a numerical method, e.g. the EM scheme, R samples of X_t, $X_t^1, ..., X_t^R$ are obtained, in the sense that $X_t^j \sim X_t$ for each $j = 1, ..., R$.

Choose $K < R$. The domain D is divided in K uniform intervals such that $K\,|I_k| = |D|$. On each interval I_k, $k = 1, ..., K$ the empirical PDF $\hat{f}(x,t)$ is defined as

$$\hat{f}(x,t) := \{\#j \text{ such that } X_t^j \in I_k\}/(R|I_k|), \qquad x \in I_k.$$

Existing results of convergence of $\hat{f}$ to f in probability as well as in distribution can be found in, e.g., [10, 58]. The following holds

$$\lim_{R\to\infty} |f(x,t) - \hat{f}(x,t)| = 0$$

almost everywhere in D, for each $t \geq 0$.

We consider in $\mathbb{R}$ the following particular case of (2.34), where $u, \sigma \in \mathbb{R}^+$

$$\begin{cases} dX_t = -uX_t dt + \sigma dW_t + dP_t \\ X_{\{t=0\}} = X_0. \end{cases} \tag{2.40}$$

The problem (2.40) models the Ornstein-Uhlenbeck process with jumps. The case without jumps is named after the pioneering work of Ornstein and Uhlenbeck [92], which overcame the limits of the Brownian motion when fitting physical data for small value of time t. Their aim was to model the velocity of a Brownian particle under the influence of friction, moving under the random impacts of neighboring particles. However, the process (2.40) has several application in finance; see [78] and references therein.

In our numerical experiments, we choose $X_0 \sim \mathcal{N}(0, 4^2)$, $u = 1$, $\sigma = 4$, $\lambda = 15$ and $g \sim \mathcal{N}(0, 4^2)$. We apply the EM method in the time interval $[0, T]$ with $N = 200$, $T = 1$ and $h = T/N$. We take $R = 10^5$ sample paths. Figure 2.5 depicts the PDF of X_T solving (2.40) at time $T = 1$.

Monte Carlo (MC) methods constitute a well-established class of numerical methods; however, their accuracy relies on the choice of the number of sample paths R. The central limit theorem ensures that the convergence rate of the MC method to compute the empirical PDF is $\mathcal{O}(1/\sqrt{R})$. Since this convergence speed is considered to be rather slow, it is natural to ask whether the PDF of a JD process can be computed with an alternative method. In Chapter 5, we illustrate and analyze an alternative approach for computing the PDF of a JD process.

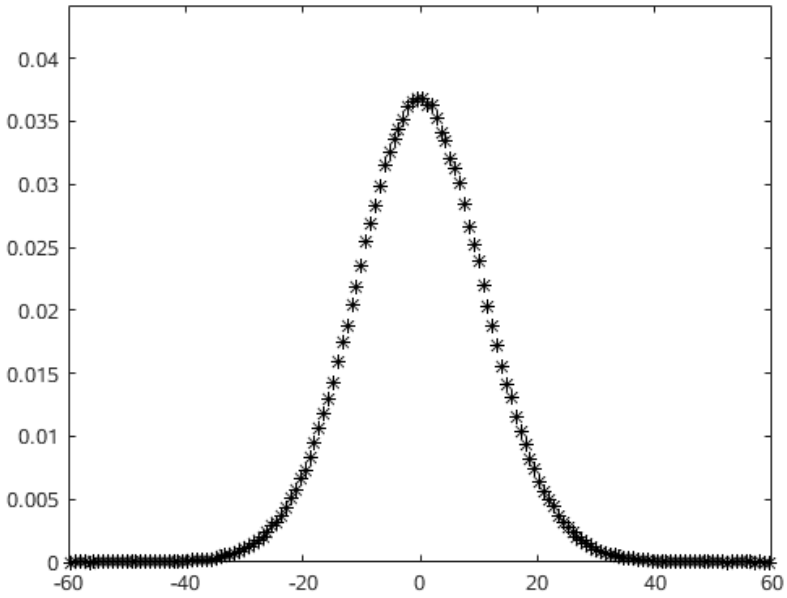


Figure 2.5: The emprirical PDF of the Ornstein-Uhlenbeck process with jumps (2.34) at time $T = 1$ with 10^5 sample paths and 150 histograms.

2.5 Summary and remarks

The aim of this chapter was to discuss initial value problems governed by SDEs with jumps. In Section 2.1 we illustrated the class of Lévy processes. We provided a review of known results stating the existence and uniqueness of solutions to SDEs driven by JD processes in Section 2.2. In Section 2.3 some existing algorithms simulating sample paths of the driving noise of a SDE are illustrated; we also presented the EM method, that generalizes the Euler method for ordinary differential equations, and we proved its strong rate of convergence. We concluded in Section 2.4 with an application of the EM method that computed the empirical PDF of a JD process.

Chapter 3

Analysis of Fokker-Planck problems related to jump-diffusion processes

The FP equation equation plays a fundamental role in modeling systems subject to randomness, since it governs the time evolution of the PDF of Markov processes. Its derivation can be carried out within a multitude of approaches that can be found, among others, in [26, 38, 80, 87]. In Chapter 2, we discussed the case of Lévy processes, which have constant drift and diffusion coefficient. In this chapter, we investigate the FP equation related to JD processes with time- and space-dependent drift and diffusion coefficients. The jump part is taken to be a compound Poisson process introduced in (2.17).

Define the time interval $I := [0, T] \subset \mathbb{R}^+$ and the space range $\Omega \subseteq \mathbb{R}^d$. The boundedness of Ω will be specified later in the chapter. In what follows, we consider the following JD model

$$\begin{cases} dX_t = b(X_{t^-}, t)dt + \sigma(X_{t^-}, t)dW_t + dP_t \\ X_{\{t=0\}} = X_0. \end{cases} \tag{3.1}$$

The functions $b : \Omega \times \mathbb{R}^+ \to \mathbb{R}^d$ and $\sigma : \Omega \times \mathbb{R}^+ \to \mathbb{R}^{d\times m}$ are assumed to be measurable. $W = \{W_t\}_{t\geq 0}$ is an m-dimensional Wiener process and $P = \{P_t\}_{t\geq 0}$ is a compound Poisson process with rate λ. The jump distribution will be specified later. We assume that the coefficients b and σ satisfy the Lipschitz and growth conditions of Theorem 4 in Chapter 2, so that the initial-value problem (3.1) admits a unique solution.

We carry out the explicit derivation of the FP equation related to (3.1) in two cases, where the space domain Ω is unbounded or bounded, respectively. The former case

is outlined in Section 3.1 and relies on the application of the Itô's formula. The latter case is illustrated in Section 3.2; due to the boundedness of the domain Ω, we take into account the presence of reflecting barriers. When the dynamics of a stochastic process is restricted to a bounded range, the behavior of the process at the boundaries is of fundamental importance for the definition of appropriate boundary conditions. Existing results on existence, uniqueness and continuous dependence on data in suitable functional spaces are presented. In Section 3.3, we provide useful a priori estimates for the FP problem, which models the constraint in our optimization problems.

Next, we present the function spaces that are essential to the investigation of the FP problem considered in this chapter. Let $\alpha \in (0,1)$. In the following definitions, $i, j \in \mathbb{N}^d$ denote multi-indeces.

- The space $C^0(\Omega)$ refers to the functions ϕ that are continuous in Ω and it is endowed with the supremum norm

$$\|\phi\|_{C^0(\Omega)} := \sup_{x\in\Omega} |\phi(x)|.$$

- The space $C^k(\Omega)$ refers to the functions ϕ that are continuous in Ω and whose derivatives up to order k belong to $C^0(\Omega)$. It is endowed with the following norm

$$\|\phi\|_{C^k(\Omega)} := \sum_{i=1}^{k} \sum_{|s|=i} \|\partial^i \phi(x)\|_{C^0(\Omega)}.$$

- The space $C^{2,\alpha}(\Omega)$ refers to the functions ϕ that are $\mathcal{C}^2$ and Hölder continuous on Ω, with Hölder exponent α. Define the following seminorm

$$\langle \phi \rangle_{\Omega}^{\alpha} := \inf\{C \geq 0 \,:\, |\phi(x) - \phi(y)| \leq C|x-y|^{\alpha},\, \forall x, y \in \Omega\}. \tag{3.2}$$

  The space $C^{2,\alpha}(\Omega)$ is endowed with the following norm

$$\|\phi\|_{C^{2,\alpha}(\Omega)} = \|\phi\|_{C^2(\Omega)} + \langle \phi \rangle_{\Omega}^{\alpha}.$$

- The space $C^{\alpha;\frac{\alpha}{2}}(\Omega \times I)$ refers to the functions ϕ that are Hölder continuous on $\Omega \times I$, with Hölder exponents α and $\alpha/2$ with respect to the space and time variables, respectively. Define the following seminorm

$$\langle \phi \rangle_{\Omega\times I}^{\alpha} := \langle \phi \rangle_{x,\Omega}^{\alpha} + \langle \phi \rangle_{t,I}^{\frac{\alpha}{2}}, \tag{3.3}$$

  where the last due terms are intended in the sense of (3.2) and the subscript stand for variable related to the Hölder exponent.

  The space $C^{\alpha;\frac{\alpha}{2}}(\Omega \times I)$ is endowed with the following norm

$$\|\phi\|_{C^{\alpha;\frac{\alpha}{2}}(\Omega\times I)} := \|\phi\|_{C^0(\Omega\times I)} + \langle \phi \rangle_{\Omega\times I}^{\alpha}.$$

- The space $\mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}(\Omega \times I)$ refers to the $\mathcal{C}^{2,1}(\Omega \times I)$ functions ϕ that are Hölder continuous on $\Omega \times I$, with Hölder exponents α and $\alpha/2$ with respect to the space and time variables, respectively. Define the following seminorm

$$\langle \phi \rangle_{\Omega}^{2,\alpha} := \langle \partial_t \phi \rangle_{\Omega \times I}^{\alpha} + \sum_{i,j=1}^{d} \langle \partial_{ij} \phi \rangle_{\Omega \times I}^{\alpha} + \sum_{i=1}^{d} \langle \partial_i \phi \rangle_{t,I}^{\frac{\alpha}{2}}.$$

  It is endowed with the following norm

$$\|\phi\|_{\mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}(\Omega \times I)} := \sum_{2j+i=0}^{2} \|\partial_t^j \partial_x^i \phi\|_{C^0(\Omega \times I)} + \langle \phi \rangle_{\Omega \times I}^{2,\alpha}$$

- The space $L^\infty(\Omega \times I)$ denotes all the functions v that are essentially bounded on $\Omega \times I$. The L^∞-norm is defined as follows

$$||v||_{L^\infty(\Omega \times I)} := \inf_{C>0} \{|v(x,t)| \leq C, \text{ a.e. in } \Omega \times I\}.$$

- The spaces $H^1(\Omega)$ is defined as follows

$$H^1(\Omega) := \{v \in L^2(\Omega) \,|\, \partial_i v \in L^2(\Omega), i = 1, ..., d\},$$

  and is endowed with the following norm

$$\|v\|_{H^1(\Omega)} := \sum_{|i|=0}^{1} \|\partial^i v\|_{L^2(\Omega)}.$$

- The space $H^{2,1}(\Omega \times I)$ is defined as follows

$$H^{2,1}(\Omega \times I) := \{v \in L^2(\Omega \times I) \,|\, \partial_t v, \partial_i v, \partial_{ij}^2 v \in L^2(\Omega \times I),\ i, j = 1, ..., d\}$$

  and it is endowed with the following norm

$$\|v\|_{H^{2,1}(\Omega \times I)} := \sum_{2j+|i|=0}^{2} \|\partial_t^j \partial_x^i v\|_{L^2(\Omega \times I)}.$$

In this chapter, we always assume that the PDF g of the jump component in (3.1) satisfies the following condition

$$\int_{\mathbb{R}^d} \frac{\|y\|^2}{(1 + \|y\|)} g(y) dy = C_0 < \infty, \tag{3.4}$$

with $C_0 > 0$ and $\|y\| := (\sum_{j=1}^{d} y_j^2)^{1/2}$ for $y \in \mathbb{R}^d$.

3.1 A Fokker-Planck model in an unbounded domain

In this section, we consider the model (3.1) with range in the whole space $\mathbb{R}^d$. In what follows, we make use of the Itô's formula [6, 23], which plays a fundamental role in stochastic analysis.

Definition 10. (Itô's formula with jumps) Consider (3.1) and denote with g the PDF of the jump distribution of P. For each twice-differentiable function $h : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ with gradient ∇h and Hessian matrix $H(h) \in \mathbb{R}^{d\times d}$ defined by $[H(h)(x,t)]_{ij} := \partial^2_{ij} h(x,t)$, the following formula holds

$$
\begin{aligned}
&h(X_t,t) - h(X_0,0) = \\
&\int_0^t \left[\frac{\partial h(X_s,s)}{\partial t} + \nabla h(X_s,s)^T \cdot b(X_s,s) + \frac{1}{2}\mathrm{Tr}\left(\sigma(X_s,s)^T\, H(h)(X_s,s)\, \sigma(X_s,s)\right)\right] ds + \\
&+ \int_0^t \nabla h(X_s,s)^T \sigma(X_s,s) dW_s + \int_0^t \lambda ds \int_{\mathbb{R}^d} [h(X_{s^-}+y,s) - h(X_{s^-}),s]\, g(y) dy + \\
&+ \int_0^t \int_{\mathbb{R}^d} [h(X_{s^-}+y,s) - h(X_{s^-},s)] \tilde{\mathcal{P}}(dy,ds),
\end{aligned}
\tag{3.5}
$$

where $\mathcal{P}$ is the Poisson random measure introduced in Chapter 2 and $\tilde{\mathcal{P}}$ is the compensated Poisson random measure associated to $\mathcal{P}$ defined in (2.5).

In what follows, we denote with $f(x,t)$ the PDF of the process X solving (3.1). Our goal is to write a deterministic initial-value problem governed by the FP equation associated to (3.1), whose solution is $f(x,t)$. The derivation outlined in this section extends the approach followed in [38].

The starting point of our discussion is the following expected value

$$
\frac{d}{dt}\mathbb{E}[h(X_t)], \tag{3.6}
$$

where $h \in \mathcal{C}^2(\mathbb{R}^d)$ is an arbitrary function and X_t is governed by (3.1).

One one hand, we switch the integration and differentiation steps in (3.6) as follows

$$
\frac{d}{dt}\mathbb{E}[h(X_t)] = \frac{d}{dt}\int_{\mathbb{R}^d} h(x) f(x,t) dx = \int_{\mathbb{R}^d} h(x) \frac{\partial f(x,t)}{\partial t} dx, \tag{3.7}
$$

assuming that the involved functions are smooth enough in order to allow the operator exchange.

On the other hand, we recast (3.6) by making use of the formula in (3.5). After taking the expectation in (3.5), we have

$$
\begin{aligned}
&\mathbb{E}[h(X_t)] - \mathbb{E}[h(X_0)] = \\
&\mathbb{E}\left[\int_0^t \left(\nabla h(X_s)^T \cdot b(X_s, s) + \frac{1}{2}\mathrm{Tr}\left(\sigma(X_s, s)^T H(h)(X_s)\, \sigma(X_s, s)\right)\right) ds + \right. \\
&\left. + \int_0^t \lambda ds \int_{\mathbb{R}^d} [h(X_{s^-} + y) - h(X_{s^-})]\, g(y) dy\right] = \\
&\int_{\mathbb{R}^d} f(x,t) \left[\int_0^t \left(\nabla h(x)^T \cdot b(x, s) + \frac{1}{2}\mathrm{Tr}\left(\sigma(x, s)^T H(h)(x)\, \sigma(x, s)\right)\right) ds + \right. \\
&\left. + \int_0^t \lambda \int_{\mathbb{R}^d} [h(x + y) - h(x)]\, g(y) dy\, ds\right] dx,
\end{aligned}
$$

where we exploited the fact that the expected values with respect to a Brownian motion and to a compensated Poisson random measure are zero; see, e.g., [23]. After differentiating with respect to time, we have

$$
\begin{aligned}
&\frac{d}{dt}\mathbb{E}[h(X_t)] = \\
&\int_{\mathbb{R}^d} f(x,t) \left(\nabla h(x)^T \cdot b(x, t) + \frac{1}{2}\mathrm{Tr}\left(\sigma(x, t)^T H(h)(x)\, \sigma(x, t)\right) + \right. \\
&\left. + \lambda \int_{\mathbb{R}^d} [h(x + y) - h(x)]\, g(y) dy\right) dx.
\end{aligned} \tag{3.8}
$$

Note that the dynamics of (3.1) has range in the unbounded domain $\mathbb{R}^d$, and therefore the PDF f is defined on the whole space $\mathbb{R}^d$. We assume that the following decay conditions hold

$$
\lim_{\|x\| \to \infty} f(x,t) = 0, \quad \lim_{\|x\| \to \infty} \partial_i f(x,t) = 0 \text{ for each } i = 1, ..., d, \tag{3.9}
$$

which are called *natural boundary conditions* [38]. We integrate by parts the right-hand side in (3.8) and obtain the following

$$
\begin{aligned}
&\frac{d}{dt}\mathbb{E}[h(X_t)] = \\
&\int_{\mathbb{R}^d} f(x,t) \left(\nabla h(x)^T \cdot b(x, t) + \frac{1}{2}\sum_{k=1}^{m}\sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_{ik}(x,t)\, [H(h)(x)]_{ij}\, \sigma_{jk}(x,t) + \right. \\
&\left. + \lambda \int_{\mathbb{R}^d} [h(x + y) - h(x)]\, g(y) dy\right) dx = \\
\end{aligned}
$$

$$\int_{\mathbb{R}^d} h(x) \left(-\nabla(f(x,t)b(x,t)) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} \frac{\partial^2}{\partial_i \partial_j} \left(f(x,t) \sum_{k=1}^{m} \sigma_{ik}(x,t)\sigma_{jk}(x,t) \right) \right.$$
$$\left. +\lambda \int_{\mathbb{R}^d} [f(y,t) - f(x,t)]\, g(x-y)dy \right) dx. \tag{3.10}$$

We define the differential operator $\mathcal{L}$ as follows

$$\mathcal{L}f(x,t) := \nabla \cdot F(x,t), \tag{3.11}$$

where

$$\begin{aligned} F_i(x,t) &:= B_i(x,t)f(x,t) + \sum_{j=1}^{d} C_{ij}(x,t)\nabla f(x,t)_j, \\ B_i(x,t) &:= \sum_{i=1}^{d} \partial_j C_{ij}(x,t) - b_i(x,t), \\ C_{ij}(x,t) &:= \frac{1}{2}\sum_{k=1}^{m} \sigma_{ik}(x,t)\sigma_{jk}(x,t). \end{aligned} \tag{3.12}$$

We define the integral operator $\mathcal{I}$ as follows

$$\mathcal{I}f(x,t) := \lambda \int_{\mathbb{R}^d} f(y,t)g(x-y)dy - \lambda f(x,t), \tag{3.13}$$

since g is a PDF and therefore it satisfies $\int_{\mathbb{R}^d} g(x)dx = 1$.

With this notation, (3.10) can be recast into the following equation

$$\frac{d}{dt}\mathbb{E}[h(X_t)] = \int_{\mathbb{R}^d} h(x) \left(\mathcal{L}f(x,t) + \mathcal{I}f(x,t)\right) dx. \tag{3.14}$$

Recall that the function h was arbitrarily chosen. By equating (3.7) to (3.14), we obtain the following equation on $\mathbb{R}^d \times I$

$$\partial_t f(x,t) = \mathcal{L}f(x,t) + \mathcal{I}f(x,t), \tag{3.15}$$

which is the PIDE FP equation governing the time evolution of the PDF f of the process (3.1).

Remark 1. It holds that $\mathcal{L} = \bar{\mathcal{A}}$, where $\bar{\mathcal{A}}$ is defined in (2.27) and corresponds to the adjoint operator of the Itô diffusion (2.24), i.e. the process (3.1) with no jumps.

We discuss the following FP PIDE initial value problem

$$\begin{cases} \partial_t f(x,t) = \mathcal{L}f(x,t) + \mathcal{I}f(x,t) + \psi(x,t) & (x,t) \in \mathbb{R}^d \times I \\ f(x,0) = f_0(x) & x \in \mathbb{R}^d, \end{cases} \tag{3.16}$$

where f_0 is a known initial data. The operators $\mathcal{L}$ and $\mathcal{I}$ are defined in (3.11) and (3.13), respectively, while the source term ψ is included, as this is needed later in this thesis, see Chapter 5.

Next, we present a theorem that ensures existence and uniqueness of classical solutions to (3.16).

Theorem 6. *Let the coefficients b and σ in* (3.1) *belong to the space of Hölder continuous function $\mathcal{C}^{\alpha,\frac{\alpha}{2}}(\mathbb{R}^d \times I)$. Let g satisfy* (3.4). *Then for any $f_0 \in \mathcal{C}^{2,\alpha}(\mathbb{R}^d)$ and for any source term $\psi \in \mathcal{C}^{\alpha;\frac{\alpha}{2}}(\mathbb{R}^d \times I)$, the initial-value problem* (3.16) *admits a unique solution $f \in \mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}(\mathbb{R}^d \times I)$ satisfying*

$$\|f\|_{\mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}(\mathbb{R}^d\times I)} \leq K\left(\|f_0\|_{\mathcal{C}^{2,\alpha}(\mathbb{R}^d)} + \|\psi\|_{\mathcal{C}^{\alpha;\frac{\alpha}{2}}(\mathbb{R}^d\times I)}\right),$$

where the constant K does not depend on ψ and f_0.

Proof. The proof relies on the construction of a sequence $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}(\mathbb{R}^d \times I)$, which admits a limit in the $\mathcal{C}^{2,\alpha;1,\frac{\alpha}{2}}$-norm. A complete proof can be found in [39, Theorem 3.1]. □

The solution f to (3.16) with $\psi = 0$ is the PDF of the process governed by (3.1), provided that f_0 is the PDF of the initial random variable X_0 in (3.1); therefore, f must satisfy the following conditions.

$$\begin{aligned} &1)\ f(x,t) \geq 0 \text{ for each } (x,t) \in \mathbb{R}^d \times I, \\ &2)\ \int_{\mathbb{R}^d} f(x,t)dx = 1 \text{ for each time } t \in I. \end{aligned} \tag{3.17}$$

While the positivity follows from standard arguments for PIDEs [39], the conservation of the total probability is proved in the next theorem.

Theorem 7. *Consider the FP problem* (3.16) *with $\psi = 0$. Let us assume the following condition on F defined in* (3.12)

$$\lim_{\|x\|\to\infty} F(x,t) = 0 \qquad \forall t \in I.$$

Then

$$\int_{\mathbb{R}^d} f(x,t) = \int_{\mathbb{R}^d} f_0(x)dx \qquad \forall t \in I.$$

Proof. Recall that $\int_{\mathbb{R}^d} g(x) = 1$ since g is the PDF of the jump distribution in (3.1). Applying Fubini's theorem, we have that

$$\int_{\mathbb{R}^d} \mathcal{I}f(x,t)dx = \lambda\left[\int_{\mathbb{R}^d} f(y,t)\left(\int_{\mathbb{R}^d} g(x-y)dx\right)dy - \int_{\mathbb{R}^d} f(x,t)dx\right] = 0.$$

Hence, we have

$$\frac{d}{dt}\int_{\mathbb{R}^d} f(x,t)dx = \int_{\mathbb{R}^d} \partial_t f(x,t)dx = \int_{\mathbb{R}^d} \left[\nabla \cdot F(x,t) + \mathcal{I}f(x,t)\right]dx = 0,$$

that proves conservation of the total probability. □

Note that the condition on F in Theorem 7 is feasible, thanks to the hypotheses on b and σ of Theorem 6 together with the so-called natural boundary conditions (3.9) employed in the above derivation of the FP equation.

3.2 A Fokker-Planck model in a bounded domain

In this section, we derive the FP problem related to the one-dimensional process (3.1) with range in a bounded domain $\Omega := (r, s) \subset \mathbb{R}$. The dynamics of (3.1) in proximity of the barrier $\partial\Omega$ has to be specified and suitable modeled in the FP problem. The term $\mathcal{L}$ in the FP equation is given by (3.11). The boundary conditions are specified below. The presence of jumps gives rise to a nonlocal operator. The definition of $\mathcal{I}$ in (3.13) does not take into account the reflecting barriers in $x = r$ and $x = s$. Next, we show how the integral operator has to be defined. We consider a specific model for the jump part of the process.

In this section, we denote with $\hat{g}$ the PDF of the jump amplitude of the compound process P in (3.1). We assume that $\hat{g}$ has support in $(r - s, s - r)$. This means that the size of a jump is smaller than $|\Omega|$ and thus multiple reflections are avoided. Since $\hat{g}$ is a PDF, the following holds

$$\int_{r-s}^{s-r} \hat{g}(z)dz = 1, \tag{3.18}$$

since $\hat{g}$ must be a PDF over its range of definition.

The integral operator is defined as follows

$$\mathcal{I}f(x,t) := \lambda \left[\int_{\Omega} f(y,t)g(x,y)dy - f(x,t) \right], \tag{3.19}$$

where the kernel g of is given by the sum of three components as follows

$$g(x,y) := \hat{g}(x-y) + \hat{g}(2r-x-y) + \hat{g}(2s-x-y).$$

The definition of the kernel g is motivated as follows. Assume that the process is in y at time $t - \delta t$; in order to reach the new position in x, the following scenarios are possible.

1. The boundaries are not encountered and therefore the density $\hat{g}$ is evaluated in $x - y$.

2. A jump with size $z < 0$ in the negative direction occurs, with $z + y < r$. The left boundary point is therefore hit and the PDF $\hat{g}$ is evaluated in $2r - x - y$.

3. A jump with size $z > 0$ in the negative direction occurs, with $z + y > s$. The right boundary point is therefore hit and the PDF $\hat{g}$ is evaluated in $2s - x - y$.

Define $Q := \Omega \times I$ and $\Sigma := \partial\Omega \times I$. We consider the following initial-boundary value problem

$$\begin{cases} \partial_t f(x,t) = \nabla \cdot F(x,t) + \mathcal{I} f(x,t) + \psi(x,t) & \text{for } (x,t) \in Q \\ f(x,0) = f_0(x) & \text{for } x \in \Omega \\ F(x,t) = 0 & \text{for } (x,t) \in \Sigma, \end{cases} \tag{3.20}$$

where F and $\mathcal{I}$ are defined in (3.12) and (3.19), respectively. The zero-flux boundary conditions $F = 0$ are motivated below. The right-hand side term ψ has been added for the purpose of the analysis in Chapter 5. However, the PDF of (3.1) on a bounded domain with reflecting barriers solves (3.20) with $\psi = 0$.

Next, we present a theorem that ensures existence and uniqueness of solutions to (3.20). We focus on the framework given by weak solutions, since it becomes essential in the analysis of FP optimal control problems discussed in Chapter 4.

Theorem 8. *Let the coefficients b and C in* (3.12) *satisfy the following conditions*

$$b \in L^\infty(Q) \cap \mathcal{C}^{\alpha;\frac{\alpha}{2}}(\Sigma), \qquad \partial_x b \in L^\infty(Q), \tag{3.21}$$

$$C \in C^0(\bar{Q}) \cap \mathcal{C}^{\alpha;\frac{\alpha}{2}}(\Sigma), \qquad \partial_x C \in L^\infty(Q) \cap \mathcal{C}^{\alpha;\frac{\alpha}{2}}(\Sigma), \qquad \partial_x^2 C \in L^\infty(Q).$$

Then, for given $f_0 \in H^1(\Omega)$ and $\psi \in L^2(Q)$, the initial-boundary value problem (3.20) *admits a unique solution $f \in H^{2,1}(Q)$, and the following holds*

$$\|f\|_{H^{2,1}(Q)} \leq K \left(\|f_0\|_{H^1(\Omega)} + \|\psi\|_{L^2(Q)} \right),$$

where the constant K does not depend on ψ and f_0.

Proof. The proof relies on the construction of a map Λ within the complete metric space $H^{2,1}(Q)$. After proving that Λ is a contraction, thanks to a priori estimates, the claim follows. A complete proof can be found in [39, Theorem 3.6]. □

Note that b and C must be defined on the closure $\bar{\Omega}$ due to their role in the boundary conditions in (3.20).

The properties in (3.17) most hold for the solution to (3.20). The positivity follows from standard arguments for PIDEs [39]. The next theorem proves the conservation of the total probability.

Theorem 9. *Consider the FP problem as in* (3.20) *with $\psi = 0$.*
Then

$$\int_\Omega f(x,t)dx = \int_\Omega f_0(x)dx \qquad \forall t \in I.$$

Proof. We first prove that

$$\int_\Omega g(x,y)dx = 1 \tag{3.22}$$

for each $y \in \Omega$. After performing a change of variable, we have

$$\begin{aligned}
&\int_\Omega g(x,y)dx = \\
&\int_r^s \left[\hat{g}(x-y) + \hat{g}(2r-x-y) + \hat{g}(2s-x-y)\right] dx = \\
&\int_{r-y}^{s-y} \hat{g}(z)dz + \int_{2r-s-y}^{r-y} \hat{g}(z)dz + \int_{s-y}^{2s-r-y} \hat{g}(z)dz = \int_{2r-s-y}^{2s-r-y} \hat{g}(z)dz.
\end{aligned}$$

Note that $y \in (r,s)$ implies $2r-s-y < r-s$ and $2s-r-y > s-r$, and therefore

$$\int_\Omega g(x,y)dx = \int_{r-s}^{s-r} \hat{g}(z)dz = 1,$$

independently from y. Applying Fubini's theorem, we have that

$$\begin{aligned}
\int_\Omega \mathcal{I}f(x,t)dx &= \lambda \int_\Omega \left[\int_\Omega f(y,t)g(x,y)dy - f(x,t)\right] dx = \\
&= \lambda \left[\int_\Omega f(y,t) \left(\int_\Omega g(x,y)dx\right) dy - \int_\Omega f(x,t)dx\right] = 0.
\end{aligned}$$

Next, we consider $\int_\Omega f(x,t)dx$. After integrating by parts and after exploiting the zero-flux boundary conditions in (3.20), we have

$$\frac{d}{dt}\int_\Omega f(x,t)dx = \int_\Omega \partial_t f(x,t)dx = \int_\Omega \left[\nabla \cdot F(x,t) + \mathcal{I}f(x,t)\right] dx = 0,$$

that proves conservation of the total probability. □

3.3 A priori estimates for the Fokker-Planck problem

In this section, we provide a priori estimates satisfied by the solution to FP problems. In order to do so, we need some definitions. A complete discussion on the topics covered in this section can be found in standard references, e.g. [60, 91].

As discussed in Section 3.1, σ is full rank and therefore C in (3.12) is positive definite. In other words, there exists $k_C > 0$ such that

$$v^T C(x,t)\, v \geq k_C \|v\|^2_{\mathbb{R}^d} \text{ for each } v \in \mathbb{R}^d, \text{ a.e. in } \Omega \times I. \tag{3.23}$$

We consider (3.20) in $\mathbb{R}^d$. We assume that g in (3.19) is essentially bounded and therefore belongs to $L^\infty(\Omega \times \Omega)$ and we define

$$k_g := \inf_{\alpha \geq 0} \{|g(x,y)| \leq \alpha \text{ for almost every } (x,y) \in \Omega \times \Omega\}. \tag{3.24}$$

Consider the following spaces $V := H^1(\Omega)$, $H := L^2(\Omega)$. We denote with V^* the dual space of V and with $\langle \cdot, \cdot \rangle_{V^*V}$ their canonical pairing. We consider the following Gelfand triple $V \subset H \subset V^*$, that exploits the natural isomorphism $H \simeq H^*$ between a Hilbert space with his dual. Each embedding is dense and continuous.

Given the interval I and an arbitrary Banach space Z, we define the following Bochner spaces

$$L^2(I;Z) = \{y : I \to Z \text{ such that } \int_I \|y(t)\|_Z^2 dt < \infty\} \tag{3.25}$$

$$\mathcal{C}(I;Z) = \{y : I \to Z \text{ such that } \lim_{\tau \to t} \|y(\tau) - y(t)\|_Z = 0 \ \forall t \in I\}, \tag{3.26}$$

which are also Banach spaces equipped with the following norms

$$\|y\|_{L^2(I;Z)} := \left(\int_I \|y(t)\|_Z^2 dt \right)^{\frac{1}{2}} \text{ and } \|y\|_{\mathcal{C}(I;Z)} := \max_{t \in I} \|y(t)\|_Z,$$

respectively. We consider the following space

$$W := \{y \in L^2(I;V) \text{ with } \partial_t y \in L^2(I;V^*)\}, \tag{3.27}$$

which is a Hilbert space with respect to the scalar product defined as follows

$$(f,g)_W := \int_I (f,g)_V + \int_I (\partial_t f, \partial_t g)_{V^*}.$$

With this preparation, we can recall the following theorem [91].

Theorem 10. *The embedding $W \hookrightarrow \mathcal{C}(I;H)$ is continuous. Therefore, every $y \in W$ coincides with an element of $\mathcal{C}(I;H)$, up to a set of null measure.*

The following proposition provides a useful a priori estimate of the solution to (3.20).

Proposition 2. *Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$. Assume that the cofficients b and σ satisfy the conditions in Theorem 8. Therefore, each B_i in (3.12) is essentially bounded. Then if f is a solution to (3.20), the following inequality holds*

$$\|f\|_{L^2(I;V)} + \|\partial_t f\|_{L^2(I;V^*)} \leq c \|f_0\|_{L^2(\Omega)}. \tag{3.28}$$

Proof. Consider the H inner product of the equation in (3.20) with f. Exploiting the properties of the Gelfand triple, we have

$$\langle \partial_t f, f \rangle_{V^*V} = \int_\Omega (\nabla \cdot F) f + \int_\Omega \mathcal{I}(f) f. \tag{3.29}$$

We make use of the following fact, $\langle \partial_t f, f \rangle_{V^*V} = \frac{1}{2}\frac{d}{dt}\|f(t)\|^2_{L^2(\Omega)}$. The terms on the right-hand side in (3.29) are recast as follows.

First, we exploit the zero-flux boundary conditions in (3.20) and the coercivity of C as given in (3.23). Moreover, we make use of the following Cauchy inequality

$$\alpha\beta \leq \frac{\alpha^2}{2\varepsilon^2} + \frac{\beta^2\varepsilon^2}{2},$$

which holds for each $\alpha, \beta \in \mathbb{R}$ and $\varepsilon > 0$. Integrating by parts and recalling the definition of F in (3.12), we have

$$\begin{aligned}
&\int_\Omega (\nabla \cdot F) f = \int_{\partial\Omega} fF \cdot \partial\nu - \int_\Omega F \cdot \nabla f = \\
&= \int_\Omega -fB \cdot \nabla f - \int_\Omega \nabla f^T C \nabla f \leq \int_\Omega \frac{B^T B f^2}{2\varepsilon^2} + \int_\Omega \frac{\nabla f^T \nabla f \varepsilon^2}{2} - k_C \|\nabla f\|^2_{L^2(\Omega)}.
\end{aligned}$$

We choose $\varepsilon := \sqrt{2k_C}$, where k_C is defined in (3.23), and define

$$k_B := \sum_{i=1}^n \|B_i\|^2_\infty = \sum_{i=1}^n \|\sum_{j=1}^n \partial_j C_{ij}(x,t) - b_i(x,t)\|^2_\infty$$

We have $k_B < \infty$ thanks to hypotheses on the boundedness of the coefficients in (3.12). Therefore we have

$$\int_\Omega (\nabla \cdot F) f \leq \frac{k_B}{4k_C} \|f(t)\|^2_{L^2(\Omega)}. \tag{3.30}$$

Recalling the definition of $\mathcal{I}$ in (3.19) and the definition of k_g in (3.24)

$$\left| \int_\Omega \mathcal{I}(f) f \right| \leq \lambda k_g \|f(t)\|^2_{L^1(\Omega)} - \lambda \|f(t)\|^2_{L^2(\Omega)}.$$

Since Ω is bounded, we have

$$\|f(t)\|_{L^1(\Omega)} \leq \sqrt{|\Omega|}\|f(t)\|_{L^2(\Omega)}.$$

Therefore,

$$\left| \int_\Omega \mathcal{I}(f) f \right| \leq \lambda |\Omega| |k_g - 1| \|f(t)\|^2_{L^2(\Omega)}. \tag{3.31}$$

Define $c := \frac{k_B}{2k_C} + 2\lambda|\Omega||k_g - 1|$. Note that c is a bounded time-dependent function. The estimates in (3.30) and (3.31) allow us to write (3.29) as follows

$$\frac{d}{dt}\|f(t)\|^2_{L^2(\Omega)} \leq c\|f(t)\|^2_{L^2(\Omega)}.$$

By applying the Gronwall inequality, we have

$$\|f(t)\|_{L^2(\Omega)} \leq \|f_0\|_{L^2(\Omega)} \exp\left\{\frac{1}{2}\int_{t_0}^{t} c\, ds\right\}. \tag{3.32}$$

Next, we outline how to obtain an upper bound of $\|\nabla f(t)\|_{L^2(\Omega)}$. We integrate (3.23) over Ω and then recall the definition of F in (3.12). We have

$$\begin{aligned} k_C\|\nabla f(t)\|^2_{L^2(\Omega)} &\leq \int_\Omega (\nabla f)^T C \nabla f = \int_\Omega (\nabla f)^T (F - Bf) \\ &\leq \int_{\partial\Omega} fF \cdot \partial \mathrm{n} - \int_\Omega f(\nabla \cdot F) - \int_\Omega f(\nabla f)^T B \\ &\leq \int_\Omega f(\partial_t f - \mathcal{I}(f)) - \int_\Omega f(\nabla f)^T B. \end{aligned}$$

Proceeding as above, we obtain

$$\|\nabla f(t)\|_{L^2(\Omega)} \leq \bar{c}\|f_0\|_{L^2(\Omega)},$$

with $\bar{c} > 0$. This last estimate, together with (3.32), proves that

$$\|f\|_{L^2(I;V)} \leq c\|f_0\|_{L^2(\Omega)},$$

up to a redefinition of the constant $c > 0$.

The bound for $\|\partial_t f\|_{L^2(I;V^*)}$ follows by similar arguments by considering the $\langle\cdot,\cdot\rangle_{V^*V}$ product of the equation in (3.20) and the integrating the result over the entire time interval $[0, T]$. □

Proposition 3. *Let $f_0 \in H^1(\Omega)$, $f_0 \geq 0$. Assume that the cofficients b and σ satisfy the conditions in Theorem 8. Then the unique solution to* (3.20) *belongs to $L^2(I;V)$, with $\partial_t f \in L^2(I;V^*)$. Moreover, $f \in \mathcal{C}(I;H)$.*

Proof. The statement follows from the a priori estimates of Proposition 2 and Theorem 10 in Chapter 3. □

With these premises, we define the space where the solution to the FP problem is sought as follows

$$\mathcal{F} := \{f \in W \mid f(0) = f_0\}. \tag{3.33}$$

3.4 Summary and remarks

In this chapter, we investigated the FP problem associated to a JD process. In Section 3.1 we derive the FP equation of process with range in the whole space $\mathbb{R}^d$ by means of the Itô's formula. When considering a bounded domain $\Omega \subset \mathbb{R}^d$, the dynamics of the stochastic process when encountering the barrier $\partial\Omega$ must be

suitable modeled in the FP problem. In Section 3.2 we derived the FP problem of a one-dimensional JD process on a bounded domain with reflecting barriers. The range of the jumps was taken to be a compact set of $\mathbb{R}$, so that multiple reflections were avoided. In Section 3.3 we derived a priori estimates for the FP problem on a bounded domain.

Chapter 4

Optimal control of Fokker-Planck equations for jump-diffusion processes

In this chapter, we set optimal control problems whose constraints are modeled by FP problems. Our purpose is to solve control problems governed by JD processes by following an approach that is alternative with respect to the field of stochastic optimal control. The key observation in our discussion is to investigate the PDF of the considered process, in the same spirit as [4, 5, 14, 54, 82].

Within the theory of optimal control, the aim is to find a control variable for a given system in such a way that an optimality criterion is achieved. This issue is often modeled by means of a cost functional that has to be minimized with respect two variables, the state and the control; together they satisfy a differential constraint [37, 60, 91]. In our work, the control acts as a part of the coefficient defining the JD process, while the state variable is the PDF of the JD process. The constraint is modeled by the FP problem.

We proceed as follows. In Section 4.1, we define the cost functionals that have to be minimized. We consider two different tracking objectives, discrete-in-time and continuous-in-time. Both have the purpose to steer the average state of a JD process towards a sequence of given values in time. The existence of at least an optimal solution is proven. In Section 4.2, we derive the first-order necessary conditions within the Lagrangian approach; in doing so, we take into account the subdifferential of the nonsmooth part of the cost functional.

4.1 Formulation of Fokker-Planck control problems

Consider the JD process in $I = [0, T] \subset \mathbb{R}^+$ solving (3.1) with range in $\Omega := (r, s)$ and reflecting barriers. Define $Q := \Omega \times I$. In Section 3.2 we derived the related FP problem (3.20) on Q. The FP problems models the constraints of our optimal control problems.

We assume the presence of control constraints given by $u_a, u_b \in \mathbb{R}$, with $u_a < 0 < u_b$. We consider the following control space

$$\mathcal{U} := L^2(I) \tag{4.1}$$

$$\mathcal{U}_{ad} := \{u \in \mathcal{U} : u_a \leq u \leq u_b\} \subset \mathcal{U}. \tag{4.2}$$

Note that the admissible set of controls $\mathcal{U}_{ad}$ is nonempty, closed, and convex.

We remark that a time-dependent control function is a natural choice considering that it originates from the stochastic differential model such as (3.1), where the time is the only independent variable.

Our control mechanism acts on the drift function $b = b(x, u)$ in (3.1). We assume that b is Lipschitz in x and smooth, such that $b'(u)$ exists for each $u \in \mathcal{U}$ and the following first-order expansion holds

$$b(u + h) = b(u) + b'(u)h + \mathcal{O}(\|h\|^2_{L^2(I)}). \tag{4.3}$$

for each increment $h \in \mathcal{U}$.

The next step consists in defining the cost functionals that have to be minimized. Denote with f the solution to (3.20). Let ν and γ be positive constants. The cost functional $\mathcal{J}$ to be minimized is defined as follows

$$\mathcal{J}(f, u) := D(f) + \frac{\nu}{2}\|u\|_2^2 + \gamma\|u\|_1, \tag{4.4}$$

with $D(f)$ denotes the tracking term defined below. The norms in (4.4) are defined as follows

$$\|u\|_2 := \left(\int_I |u(t)|^2 dt\right)^{1/2} \text{ and } \|u\|_1 := \int_I |u(t)| dt.$$

Note that the choice of a bounded time interval I ensures that the L^1-norm is finite whenever $u \in \mathcal{U}$. On the one hand, the L^2-term is a standard choice in optimal control theory. On the other hand, the L^1-term enhances the sparsity of the optimal solution [29, 89].

The term $D(f)$ in (4.4) represents a tracking objective that involves the expectation value of X_t, $\mathbb{E}[X_t] := \int_\Omega x f(x, t) dx$, and a desired trajectory or a discrete set of values (e.g. measurements). We investigate the following two cases.

1. Discrete-in-time tracking term. Given a set of values $\{\xi_k\}_{k=1}^K$ at different times $t_k \in (t_0, t_f)$, $\forall k =, 1, ..., K$, we have

$$D(f) := \frac{1}{K} \sum_{k=1}^{K} \left(\xi_k - \int_\Omega x f(x, t_k) dx \right)^2. \tag{4.5}$$

2. Continuous-in-time tracking term. Given a square-integrable function $\xi : I \to \mathbb{R}^d$, we have

$$D(f) := \int_0^T \left(\xi(t) - \int_\Omega x f(x, t) dx \right)^2 dt. \tag{4.6}$$

With these premises, we define the FP optimal control problem investigated in this chapter. We have

$$\begin{aligned} &\min_{u,f} \mathcal{J}(f, u) \\ &\text{such that } \mathcal{K}(f, u) = 0, \end{aligned} \tag{4.7}$$

with $\mathcal{J}$ is defined in (4.4) and $\mathcal{K}(f, u)$ embodies the FP problem (3.20) with the setting given by Theorem 8.

In order to discuss the existence and uniqueness of solutions to (4.7), we consider the control-to-state operator $\mathcal{S} : \mathcal{U} \to \mathcal{F}$, that maps a given $u \in \mathcal{U}$ into $\mathcal{S}(u) := f$, where (f, u) satisfies $\mathcal{K}(f, u) = 0$. Thanks to Theorem 8, we have that $\mathcal{S}$ is well defined.

The constrained optimization problem (4.7) can be transformed into an unconstrained one as follows

$$\min_{u \in \mathcal{U}_{ad}} \hat{\mathcal{J}}(u), \tag{4.8}$$

where $\hat{\mathcal{J}} : u \mapsto \hat{\mathcal{J}}(u) := \mathcal{J}(\mathcal{S}(u), u)$ is the so-called reduced cost functional.

The solvability of (4.8) is ensured by the next theorem, whose proof employs standard techniques in optimal control problems [1, 37, 91] and the a priori estimates in Chapter 3.

Theorem 11. *There exists at least one optimal pair $(\bar{f}, \bar{u})$ that solves (4.7), such that $\bar{u}$ solves (4.8) and $\bar{f} = \mathcal{S}(\bar{u})$ in $\mathcal{C}(I, H)$.*

Proof. The functional $\hat{\mathcal{J}}$ in (4.8) is bounded from below and therefore we can define $i := \inf_{u \in \mathcal{U}_{ad}} \hat{\mathcal{J}}(u)$. Let $\{u_n\}_{n \in \mathbb{N}} \subset \mathcal{U}_{ad}$ be a minimizing sequence, such that $\lim_{n \to \infty} \hat{\mathcal{J}}(u_n) = i$. We have that $\mathcal{U}_{ad}$ is a convex, closed, and bounded subset of the reflexive Banach space $\mathcal{U}$. Hence, $\mathcal{U}_{ad}$ is weakly sequentially compact and we can extract a subsequence $\{u_k\}_{k \in \mathbb{N}} \subset \{u_n\}_{n \in \mathbb{N}}$ such that $u_k \rightharpoonup \bar{u} \in \mathcal{U}_{ad}$.

The weakly convergent sequence $\{u_k\}_{k \in \mathbb{N}}$ gives rise to the sequence $\{f_k\}_{k \in \mathbb{N}} \subset \mathcal{F} \subset W$, defined by $f_k := \mathcal{S}(u_k)$. Since the embedding $W \hookrightarrow L^2(I; H)$ is compact, there exists a subsequence $\{f_j\}_{j \in \mathbb{N}} \subset \{f_k\}_{k \in \mathbb{N}}$ and $\bar{f} \in L^2(I; H)$ such that $\{f_j\}_{j \in \mathbb{N}}$

converges strongly to $\bar{f}$. Note that each couple (f_j, u_j) satisfies $K(f_j, u_j) = 0$ by definition. Next, we want to pass to the limit in $K(f_j, u_j) = 0$.

Thanks to the estimate (3.28) in Proposition 2, the sequence $\{\partial_t f_j\}_{j\in\mathbb{N}}$ is bounded in $L^2(I, V^*)$ and therefore weakly convergent to $\partial_t \bar{f}$. Define $\bar{B} := b(\bar{u})$. The boundedness of B and the smoothness of b with respect to u together with (3.28), ensures that $\{f(\nabla\cdot B)_j\}_{j\in\mathbb{N}}$ and $\{(B\cdot\nabla f)_j\}_{j\in\mathbb{N}}$ converge weakly to $\bar{f}(\nabla\cdot\bar{B})$ and $(\bar{B}\cdot\nabla\bar{f})$, respectively, where the norm in $L^2(I, H)$ has been considered. The weak convergence of $\{\mathcal{I}f_j\}_{j\in\mathbb{N}}$ to $\mathcal{I}\bar{f}$ follows from the estimates in Proposition 2. These observations lead to the conclusion that the limit $\bar{f}$ solves the FP problem and it holds $\bar{f} = \mathcal{S}(\bar{u})$. Therefore, the constraint $\mathcal{K}(\bar{f}, \bar{u}) = 0$ is satisfied. Since $\mathcal{J}$ is weakly sequentially lower semicontinuous, we have

$$i = \liminf_j \mathcal{J}(f_j, u_j) \geq \mathcal{J}(\bar{f}, \bar{u})$$

and therefore the pair $(\bar{f}, \bar{u})$ is a minimizer for the problem (4.7). □

We remark that the uniqueness of the control $\bar{u}$ can not be stated a priori since $\hat{\mathcal{J}}$ is nonconvex.

4.2 First-order optimality systems

In this section, we derive the necessary conditions satisfied by a local optimal solution. We obtain two first-order optimality systems, related to the discrete-in-time and to the continuous-in-time cost functionals, respectively. We follow a standard approach [33, 89, 91]. For preparation, we discuss the Frechét differentiability of the control-to-state operator $\mathcal{S}$.

Proposition 4. *The mapping $\mathcal{S} : \mathcal{U} \to \mathcal{C}(I; H), u \mapsto f = \mathcal{S}(u)$ solution to* (3.20) *is Fréchet differentiable. For each $h \in \mathcal{U}$, the directional derivative $e := \mathcal{S}'(u)\cdot h$ satisfies the following initial-boundary problem*

$$\begin{cases} \partial_t e = \partial_x F(e, u) + \mathcal{I}(e) - h\,\partial_x(fb'(u)) & \text{for } (x,t) \in Q, \\ F(e, u) = fb'(u)h & \text{for } (x,t) \in \Sigma, \\ e(x, 0) = 0 & \text{for } x \in \Omega, \end{cases} \tag{4.9}$$

where b is the drift of the JD process solving (3.1) *and F is defined in* (3.12).

Proof. Recall that $\mathcal{S}(u+\tau h)$ and $\mathcal{S}(u)$ satisfy (3.20) with $u+\tau h$ and u, respectively. By computing the limit

$$\lim_{\tau\to 0} \frac{\mathcal{S}(u+\tau h) - \mathcal{S}(u)}{\tau},$$

we obtain (4.9). The results of Theorem 8, together with the assumption on b defining F in (3.12), allow us to state that the right-hand side in (4.9) is in $L^2(\Omega)$. Therefore, Theorem 8 holds and (4.9) has a unique solution in $H^{2,1}(Q)$.

In order to prove the Fréchet differentiability of $\mathcal{S}$, we consider the difference $\zeta_h := \mathcal{S}(u+h) - \mathcal{S}(u) - \mathcal{S}'(u) \cdot h$. Recall that $\mathcal{K}(\mathcal{S}(u+h), u+h) = 0$ and $\mathcal{K}(\mathcal{S}(u), u) = 0$.

Therefore, ζ solves the following initial-boundary value problem

$$\begin{cases} \partial_t \zeta = \partial_x F(\zeta, u) + \mathcal{I}(\zeta) + \partial_x \eta & \text{for } (x,t) \in Q \\ F(\zeta, u) = -\eta & \text{for } (x,t) \in \Sigma \\ \zeta(x,0) = 0 & \text{for } x \in \Omega, \end{cases} \tag{4.10}$$

where

$$\eta(x,t) := \mathcal{S}(u+h)(b(u+h) - b(u)) - \mathcal{S}(u) b'(u) h.$$

By applying Theorem 8, we have that

$$\|\zeta\|_{H^{2,1}(Q)} \leq K_1 \|\partial_x \eta\|_{L^2(Q)}, \tag{4.11}$$

with $K_1 > 0$. Note that $\mathcal{S}(u+h) - \mathcal{S}(u)$ solves the FP problem (3.20) with $f_0 = 0$ and $\psi(x,t) = -\partial_x(\mathcal{S}(u+h)(b(u+h) - b(u))$, and therefore, thanks to Theorem 8, we have

$$\|\mathcal{S}(u+h) - \mathcal{S}(u)\|_{H^{2,1}(Q)} \leq K \|\partial_x(\mathcal{S}(u+h)(b(u+h) - b(u))\|_{L^2(Q)}, \tag{4.12}$$

with $K_2 > 0$. Next, we compute the following limit

$$\lim_{\|h\|_{L^2(I)} \to 0} \frac{\|\zeta\|_{L^2(Q)}}{\|h\|_{L^2(I)}}.$$

The embedding $H^{1,2}(Q) \subset L^2(Q)$ together with (4.3), (4.11) and (4.12) yields

$$\lim_{\|h\|_{L^2(I)} \to 0} \frac{\|\zeta\|_{L^2(Q)}}{\|h\|_{L^2(I)}} = \lim_{\|h\|_{L^2(I)} \to 0} (K\|h\|_{L^2(I)} + \mathcal{O}(\|h\|_{L^2(I)}),$$

where $K > 0$ does not depends on h. This concludes the proof of the differentiability of $\mathcal{S}$. □

Consider the reduced problem (4.8) and write the reduced functional $\hat{\mathcal{J}}$ as $\hat{\mathcal{J}} := \mathcal{J}_1 + \mathcal{J}_2$, $\mathcal{J}_i : \mathcal{U} \to \mathbb{R}^+$, $i = 1, 2$, where

$$\begin{aligned} \mathcal{J}_1(u) &:= D(\mathcal{S}(u)) + \frac{\nu}{2}\|u\|_2^2, \\ \mathcal{J}_2(u) &:= \gamma \|u\|_1. \end{aligned} \tag{4.13}$$

Note that the functional $\mathcal{J}_1$ is smooth, thanks to Proposition 4, and to the smoothness of (4.5) and (4.6) with respect to f. The gradient $\nabla\mathcal{J}_1$ plays a key role in the definition of the optimality system that follows. On the other hand, the addend $\mathcal{J}_2$ is convex and nonsmooth. With the aim of deriving the optimality conditions, we need the following definitions. We refer to [1, 9, 33, 91].

In the remainder of this chapter, $\mathcal{U}^*$ is the dual space of $\mathcal{U} = L^2(I)$, while $\langle\cdot,\cdot\rangle$ stands for the scalar product between $\mathcal{U}^*$ and $\mathcal{U}$. If $\hat{\mathcal{J}}$ is finite at a point u, the Fréchet subdifferential of $\hat{\mathcal{J}}$ at u is defined as follows [33]

$$\partial\hat{\mathcal{J}}(u) := \left\{ \varphi \in \mathcal{U}^* : \liminf_{v\to u} \frac{\hat{\mathcal{J}}(v) - \hat{\mathcal{J}}(u) - \langle\varphi, v-u\rangle}{\|v-u\|_2} \geq 0 \right\}. \tag{4.14}$$

Any element $\varphi \in \partial\hat{\mathcal{J}}(u)$ is called a subgradient. In our framework, we have

$$\partial\hat{\mathcal{J}}(u) = \nabla\mathcal{J}_1(u) + \partial\mathcal{J}_2(u),$$

since $\mathcal{J}_1$ is Fréchet differentiable at u. Moreover, for each $\alpha > 0$, it holds that $\partial(\alpha\hat{\mathcal{J}}) = \alpha\partial\hat{\mathcal{J}}$.

The following proposition gives a necessary condition for a local minimum of $\hat{\mathcal{J}}$.

Proposition 5. *If $\hat{\mathcal{J}} = \mathcal{J}_1 + \mathcal{J}_2$, with $\mathcal{J}_1$ and $\mathcal{J}_2$ given by (4.13), attains a local minimum in $\bar{u} \in \mathcal{U}_{ad}$, then*

$$0 \in \partial\hat{\mathcal{J}}(\bar{u}),$$

or equivalently

$$-\nabla\mathcal{J}_1(\bar{u}) \in \partial\mathcal{J}_2(\bar{u}).$$

Proof. Since $\mathcal{U}_{ad}$ is convex, $\bar{u} + \theta(v-\bar{u}) \in \mathcal{U}_{ad}$, for each $v \in \mathcal{U}_{ad}$ and $\theta \in (0,1]$. Since $\bar{u}$ is a local minimum, we have

$$\hat{\mathcal{J}}(\bar{u}) \leq \hat{\mathcal{J}}(\bar{u} + \theta(v-\bar{u})),$$

for v sufficiently close to $\bar{u}$. Exploiting the convexity of $\mathcal{J}_2$, we have

$$\mathcal{J}_1(\bar{u} + \theta(v-\bar{u})) - \mathcal{J}_1(\bar{u}) + \theta(\mathcal{J}_2(v) - \mathcal{J}_2(\bar{u})) \geq 0.$$

Dividing by θ and considering the limit $\theta \to 0$, we obtain

$$\mathcal{J}_2(v) - \mathcal{J}_2(\bar{u}) + \langle\nabla\mathcal{J}_1(\bar{u}), v-\bar{u}\rangle \geq 0. \tag{4.15}$$

Dividing by $\|v-\bar{u}\|_2$ and considering the following limit

$$\liminf_{v\to\bar{u}} \frac{\mathcal{J}_2(v) - \mathcal{J}_2(\bar{u}) + \langle\nabla\mathcal{J}_1(\bar{u}), v-\bar{u}\rangle}{\|v-\bar{u}\|_2} \geq 0,$$

we conclude that $-\nabla\mathcal{J}_1(\bar{u}) \in \partial\mathcal{J}_2(\bar{u})$, according to the definiton in (4.14).

□

The variational inequality (4.15) implies that each $\lambda \in \partial \mathcal{J}_2(\bar{u})$, with $\bar{u}$ a local minimum, satisfies the following inequality

$$\langle \nabla \mathcal{J}_1(\bar{u}) + \lambda, v - \bar{u} \rangle \geq 0 \text{ for each } v \in \mathcal{U}_{ad}. \tag{4.16}$$

Moreover, recalling the definition of $\mathcal{J}_2$ in (4.13) and exploiting the isomorphism $\mathcal{U}^* \simeq \mathcal{U}$, the inclusion $\lambda \in \partial \mathcal{J}_2(\bar{u})$ gives the following

$$\lambda \in \Lambda_\gamma := \{l \in \mathcal{L}^2(I) : |l| \leq \gamma \text{ a.e. on } I\}. \tag{4.17}$$

A pointwise analysis of (4.16), which takes into account the definition (4.2) of the admissible controls, ensures the existence of two nonnegative functions $\bar{\lambda}_a, \bar{\lambda}_b \in \mathcal{U}^*$ that play the role of Lagrange multipliers. The previous considerations lead to the following proposition, that states the optimality system for the reduced problem (4.8).

Proposition 6. *A local optimal solution $\bar{u}$ of the minimization problem (4.8) with $\hat{\mathcal{J}} = \mathcal{J}_1 + \mathcal{J}_2$ defined in (4.13), is characterized by the existence of $(\bar{\lambda}, \bar{\lambda}_a, \bar{\lambda}_b) \in \Lambda_\gamma \times \mathcal{U}^* \times \mathcal{U}^*$ such that*

$$\begin{cases} \nabla \mathcal{J}_1(\bar{u}) + \bar{\lambda} - \bar{\lambda}_a + \bar{\lambda}_b = 0 \\ \bar{\lambda}_b \geq 0, \qquad u_b - \bar{u} \geq 0, \qquad \langle \bar{\lambda}_b, u_b - \bar{u} \rangle = 0 \\ \bar{\lambda}_a \geq 0, \qquad \bar{u} - u_a \geq 0, \qquad \langle \bar{\lambda}_a, \bar{u} - u_a \rangle = 0 \\ \bar{\lambda} = \gamma \text{ a.e. on } \{t \in I : \bar{u}(t) > 0\} \\ |\bar{\lambda}| \leq \gamma \text{ a.e. on } \{t \in I : \bar{u}(t) = 0\} \\ \bar{\lambda} = -\gamma \text{ a.e. on } \{t \in I : \bar{u}(t) < 0\} \end{cases} \tag{4.18}$$

We refer to the last three conditions in (4.18) for the pair $(\bar{u}, \bar{\lambda})$ as the complementarity conditions.

Next, we compute $\nabla \mathcal{J}_1(u)$ in (4.18) within the adjoint approach. Recall the definition of $\mathcal{J}_1(u)$ in (4.13). For each $u \in \mathcal{U}$, we have

$$\nabla \mathcal{J}_1(u) = \nu u + (\mathcal{S}'(u))^* D'(\mathcal{S}(u)).$$

Note that $\mathcal{K}$ is smooth in both arguments. By considering the total derivative of $\mathcal{K}(\mathcal{S}(u), u) = 0$, we have

$$\mathcal{K}_f(\mathcal{S}(u), u)\mathcal{S}'(u) + \mathcal{K}_u(\mathcal{S}(u), u) = 0.$$

Therefore, we obtain

$$\nabla \mathcal{J}_1(u) = \nu u - \mathcal{K}_u(\mathcal{S}(u), u)^*(\mathcal{K}_f(\mathcal{S}(u), u)^*)^{-1} D'(\mathcal{S}(u)).$$

Defining the adjoint variable p as the solution to the following adjoint problem

$$\mathcal{K}_f(\mathcal{S}(u), u)^* p = -D'(\mathcal{S}(u)), \tag{4.19}$$

we obtain the following reduced gradient

$$\nabla \mathcal{J}_1(u) = \nu u + \mathcal{K}_u(\mathcal{S}(u), u)^* p. \tag{4.20}$$

Define the operator $\tilde{\mathcal{I}}$ as follows

$$\tilde{\mathcal{I}} p(x,t) := \lambda \left[\int_\Omega p(y,t) g(y,x) dy - p(x,t) \right] \text{ for each } (x,t) \in Q.$$

After some calculation, we have that (4.19) can be rewritten in terms of the following adjoint system

$$\begin{cases} -\partial_t p(x,t) = b(x;u)\partial_x p(x,t) + C(x,t)\partial^2_{xx} p(x,t) + \\ \qquad\qquad + \tilde{\mathcal{I}} p(x,t) + \alpha(x,t) & \text{on } Q \\ p(x,t_f) = 0 & \text{on } \Omega \\ \partial_x p(x,t) = 0 & \text{on } \partial\Omega \times I \\ p(x,t_k^-) = p(x,t_k^+) + \beta(x,k) & \text{on } \Omega,\ \forall k = 1,...,K, \end{cases} \tag{4.21}$$

where α and β depend on the choice of D in (4.5) and (4.6). When D is given by (4.5), $\alpha(x,t) = 0$ and $\beta(x,k) = -2x(\xi_k - \int_\Omega s f(s,t_k) ds)$, for each $k = 1,...,K$. On the other hand, when D is given by (4.6), $\alpha(x,t) = 2x(\xi(t) - \int_\Omega s f(s,t) ds)$ and $\beta(x,k) = 0$.

The terminal boundary-value problem (4.21) in case of D as in (4.6) admits a unique solution $p \in H^{2,1}(Q)$, following the same arguments as in Theorem 8 in Chapter 3. In case of D as in (4.5), we have that $p \in H^{2,1}(\Omega \times (t_k, t_{k+1}))$.

The reduced gradient (4.20), for given u, $f = \mathcal{S}(f)$, and p, takes the following form

$$\nabla \hat{\mathcal{J}}_1(u) = \nu u + \int_\Omega f\, \partial_x p\, b'(u). \tag{4.22}$$

The complementarity conditions in (4.18) can be recast in a more compact form, as follows. We define $\bar{\mu} := \bar{\lambda} - \bar{\lambda}_a + \bar{\lambda}_b$. For each $k \in \mathbb{R}^+$, we define the following quantity

$$\begin{aligned} E(\bar{u}, \bar{\mu}) :=& \bar{u} - \max\{0, \bar{u} + k(\bar{\mu} - \gamma)\} - \min\{0, \bar{u} + k(\bar{\mu} + \gamma)\} + \\ & \max\{0, \bar{u} - u_b + k(\bar{\mu} - \gamma)\} + \min\{0, \bar{u} - u_a + k(\bar{\mu} + \gamma)\}. \end{aligned}$$

The complementarity conditions in (4.18) and the inequalities related to the Lagrange multipliers $\bar{\lambda}_a$ and $\bar{\lambda}_b$, together with the requirement $\bar{\lambda} \in \Lambda_\gamma$, are equivalent to $E(\bar{u}, \bar{\mu}) = 0$.

The previous considerations can be summarized in the following propositions.

Proposition 7. *(Optimality system for the discrete-in-time tracking functional) A local solution $(f, u) \in \mathcal{F} \times \mathcal{U}_{ad}$ of (4.7) with D given by (4.5) is characterized by the existence of $(p, \mu) \in H^{2,1}(Q) \times \mathcal{U}^*$, such that the following system is satisfied*

$$\begin{cases} \nu u + \int_\Omega f\, \partial_x p\, b'(u) + \mu = 0 & \text{a.e. in } I \\ \partial_t f(x,t) = \partial_x F(x,t) + \mathcal{I} f(x,t) & \text{for } (x,t) \in Q \\ f(x,0) = f_0(x) & \text{for } x \in \Omega \\ F(x,t) = 0 & \text{for } (x,t) \in \Sigma \\ -\partial_t p(x,t) = b(x;u)\partial_x p(x,t) + C(x,t)\partial^2_{xx} p(x,t) + \tilde{\mathcal{I}} p(x,t) & \text{for } (x,t) \in Q \\ p(x,T) = 0 & \text{for } x \in \Omega \\ \partial_x p(x,t) = 0 & \text{for } (x,t) \in \Sigma \\ p(x,t_k^-) = p(x,t_k^+) - 2x(\xi_k - \int_\Omega s f(s,t_k) ds) & \text{for } x \in \Omega,\ \forall k = 1, ..., K \\ E(u,\mu) = 0 & \text{a.e. in } I. \end{cases} \tag{4.23}$$

Proposition 8. *(Optimality system for the continuous-in-time tracking functional) A local solution $(f, u) \in \mathcal{F} \times \mathcal{U}_{ad}$ of (4.7) with D given by (4.6) is characterized by*

the existence of $(p, \mu) \in H^{2,1}(Q) \times \mathcal{U}^*$*, such that the following system is satisfied*

$$\begin{cases} \nu u + \int_\Omega f\, \partial_x p\, b'(u) + \mu = 0 & \text{a.e. in } I \\ \partial_t f(x,t) = \partial_x F(x,t) + \mathcal{I} f(x,t) & \text{for } (x,t) \in Q \\ f(x,0) = f_0(x) & \text{for } x \in \Omega \\ F(x,t) = 0 & \text{for } (x,t) \in \Sigma \\ -\partial_t p(x,t) = b(x;u)\partial_x p(x,t) + C(x,t)\partial^2_{xx} p(x,t) + \tilde{\mathcal{I}} p(x,t) + \\ \qquad + 2x(\xi(t) - \int_\Omega s f(s,t) ds) & \text{for } (x,t) \in Q \\ p(x,T) = 0 & \text{for } x \in \Omega \\ \partial_x p(x,t) = 0 & \text{for } (x,t) \in \Sigma \\ E(u,\mu) = 0 & \text{a.e. in } I. \end{cases} \tag{4.24}$$

4.3 Summary and remarks

In this chapter, we presented theoretical results regarding optimal control problems governed by PIDEs of FP type. The issue of controlling a JD process is addressed by means of the action of a control variable in the drift coefficient of the FP equation. In Section 4.1, we defined the cost functionals, considering two different tracking objectives, discrete-in-time and continuous-in-time. The existence of at least an optimal solution was proven. In Section 4.2, we determined the first-order optimality systems within the Lagrangian approach; this derivation was carried out by taking into account the subdifferential of the nondifferentiable part of the cost functional.

Chapter 5

Numerical solution of Fokker-Planck problems for jump-diffusion processes

In this chapter, we discuss the numerical solution of FP problems related to JD processes. For this purpose, we investigate the SIMEX schemes in Section 5.1. After discretizing in space the differential and integral operators with the CC method and the mid-point rule, respectively, the time integration of the resulting ODE system is carried out. We apply a first-order scheme and a predictor-corrector method. The full convergence and stability analysis of the SIMEX scheme is presented in Section 5.2. We show in Section 5.3 that the chosen discretization preserves the two properties of the solution to the FP problem, i.e. the positivity and the probability conservation. Section 5.4 illustrates the numerical experiments that validate the theoretical results of the previous section.

5.1 The SIMEX schemes

In this section, we formulate our SIMEX schemes. We consider a one-dimensional jump-diffusion process modeled by

$$\begin{cases} dX_t = b(X_{t^-})dt + \sigma(X_{t^-})dW_t + dP_t, \\ X_{\{t=0\}} = X_0, \end{cases} \tag{5.1}$$

with range in $\Omega := (r, s)$ and $t \in I := [0, T]$.

The random data X_0 is given. Define $C(x) := \frac{1}{2}\sigma^2(x)$ and $B(x) := \partial_x C(x) - b(x)$. Recall that we have two operators $\mathcal{L}$ and $\mathcal{I}$ defined for each $(x,t) \in \Omega \times I$ as follows

$$\mathcal{L}f(x,t) := \partial_x F(x,t), \tag{5.2}$$

where

$$F(x,t) := B(x)f(x,t) + C(x)\partial_x f(x,t) \tag{5.3}$$

and

$$\mathcal{I}f(x,t) := \lambda \int_{\mathbb{R}} f(y,t)g(x-y)dy - \lambda f(x,t). \tag{5.4}$$

We consider the following initial-boundary value problem

$$\begin{cases} \partial_t f(x,t) = \mathcal{L}f(x,t) + \mathcal{I}f(x,t) + \psi(x,t) & \text{for } (x,t) \in \Omega \times I \\ f(x,0) = f_0(x) & \text{for } x \in \Omega \\ F(x,t) = 0 & \text{for } (x,t) \in \partial\Omega \times I. \end{cases} \tag{5.5}$$

The source term ψ has been added with the purpose of the numerical analysis of Section 5.2, whose theoretical results are tested in Section 5.4.

Our purpose is to numerically solve (5.5). We first set the mesh sizes h and δt as follows

$$h := \frac{s-r}{N-1} \text{ and } \delta t := \frac{T}{M}, \text{ with } N, M \in \mathbb{N}.$$

We consider uniform meshes in space and time. We have

$$\begin{aligned} \Omega_h &:= \{x_j = r + (j-1)h,\ j = 1, ..., N\} \\ I_{\delta t} &:= \{t_n = n\delta t,\ n = 0, ..., M\} \subset I. \end{aligned} \tag{5.6}$$

Note that Ω_h contains the boundaries r, s of Ω. In order to numerically solve (5.5), we first consider the spatial discretization of $\mathcal{L}$ and $\mathcal{I}$, defined in (5.2) and (5.4), respectively. This leads to a large system of ODEs. Exploiting the form (5.2) of the differential operator $\mathcal{L}$, we discretize the spatial derivative of the flux F defined by (5.3) using the CC scheme. This is a cell-centered finite-volume scheme that computes the flux at the j-cell's boundaries, $x_{j\pm\frac{1}{2}}$, for $j = 1, ..., N$, where $x_{j\pm\frac{1}{2}} := \frac{x_j + x_{j\pm1}}{2}$. The unknown variables are computed on the cell-centers x_j, for each cell $j = 1, ..., N$. In what follows, given a generic continuous function $\Phi(x,t)$, we denote with $\Phi_j(t)$ and $\Phi_{j\pm\frac{1}{2}}(t)$ the time-continuous restrictions of Φ to x_j and $x_{j\pm\frac{1}{2}}$, respectively.

For each $j = 1, ..., N$, we have the following discretization formula

$$\partial_x F(x_j,t) \approx \frac{F_{j+\frac{1}{2}}(t) - F_{j-\frac{1}{2}}(t)}{h}, \tag{5.7}$$

where for example

$$F_{j+\frac{1}{2}}(t) = B_{j+\frac{1}{2}}[(1-\delta_{j+\frac{1}{2}})f_{j+1}(t) + \delta_{j+\frac{1}{2}}f_j(t)] + C_{j+\frac{1}{2}}\frac{f_{j+1}(t)-f_j(t)}{h}. \tag{5.8}$$

The linear combination of the values of the unknown f on two consecutive grid points x_j and x_{j+1} (or x_{j-1} and x_j) by means of the weight δ is the distinguishing feature of the CC method and is motivated as follows. The key observation is that the equilibrium solution of the equation in (5.5) with no integral term and no source term is given by

$$f(x) = k\exp\left\{-\int_r^x \frac{B(y)}{C(y)}dy\right\},$$

with k integration factor. Define the following quantity

$$w(x) := h\frac{B(x)}{C(x)}. \tag{5.9}$$

The ratio between the equilibrium solution at two adjacent points is approximated by the midpoint rule [28] as follows

$$\frac{f(x_{j+1})}{f(x_j)} = \exp\left\{-\int_{x_j}^{x_{j+1}} \frac{B(y)}{C(y)}dy\right\} \approx \exp\left\{-w_{j+\frac{1}{2}}\right\}. \tag{5.10}$$

Using (5.8), we impose that the discrete flux F has to vanish in $x_{j+\frac{1}{2}}$, where we use $f_{j+\frac{1}{2}} = \delta_{j+\frac{1}{2}}f_j + (1-\delta_{j+\frac{1}{2}})f_{j+1}$. We obtain

$$\frac{f_{j+1}}{f_j} = \frac{C_{j+\frac{1}{2}} - hB_{j+\frac{1}{2}}\delta_{j+\frac{1}{2}}}{C_{j+\frac{1}{2}} + hB_{j+\frac{1}{2}}(1-\delta_{j+\frac{1}{2}})}. \tag{5.11}$$

By equating (5.10) and (5.11), we find that the space-dependent parameter δ is defined as follows

$$\delta(x) := \frac{1}{w(x)} + \frac{1}{1-e^{w(x)}}. \tag{5.12}$$

From (5.12), it follows that $\delta(x) \in [0,1]$, and thus the linear combinations of the unknown in (5.8) are convex. Notice that $\delta \to \frac{1}{2}$ as $w \to 0$; $\delta \to 0^+$ as $w \to -\infty$; $\delta \to 1^-$ as $w \to \infty$.

The integral in (5.4) is approximated with the midpoint rule [28] on each cell $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$, for $i = 1, ..., N$, as follows

$$\int_{\mathbb{R}} f(y,t)g(x-y)dy = \sum_{i=1}^{N}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(y,t)g(x-y)dy \approx \sum_{i=1}^{N} hf(x_i,t)g(x-x_i).$$

Hence, for each $j = 1, ..., N$, the discretization of the integral operator (5.4) on the line (x_j, t) takes the following form

$$\mathcal{I}f(x_j,t) \approx \lambda\left(h\sum_{i=1}^{N} f(x_i,t)g(x_j-x_i) - f(x_j,t)\right). \tag{5.13}$$

The discretization steps (5.7) and (5.13) give the following MOL approximation

$$\begin{cases} f'_{\mathrm{SD}}(t) = (\mathcal{A} + \mathcal{G}) f_{\mathrm{SD}}(t) + \Psi(t) \\ f_{\mathrm{SD}}(0) = f(0), \end{cases} \tag{5.14}$$

where $\mathcal{A}$ and $\mathcal{G}$ are defined below. Notice that (5.14) is a system of ordinary differential equations parametrized by the space mesh size h in $\mathcal{A}$ and $\mathcal{G}$. In other words, $f_{\mathrm{SD}}(t) = \{f_{SD,1}(t), \ldots, f_{SD,N}(t)\} \in \mathbb{R}^N$ can be viewed as a grid function, where each component describes the time evolution of f_{SD} on the correspondent grid point of Ω_h. The initial value $f(0)$ and the source term Ψ represent the restriction on the grid Ω_h of the sufficiently smooth initial data f_0 and of the source term ψ in (5.5), respectively.

The matrix $\mathcal{A}$ in (5.14) follows from the CC method and it takes into account the zero-flux boundary condition of (5.5), which are implemented at the boundary points $x_{\frac{1}{2}}$ and $x_{N+\frac{1}{2}}$. By setting

$$\alpha(x) := \frac{C(x)}{h} + (1 - \delta(x))B(x), \qquad x \in \Omega \tag{5.15}$$

$$\beta(x) := \frac{C(x)}{h} - \delta(x)B(x), \qquad x \in \Omega \tag{5.16}$$

the tridiagonal matrix $\mathcal{A}$ is given by

$$\mathcal{A}_{ij} = \frac{1}{h} \begin{cases} \alpha_{i+\frac{1}{2}} & 1 \le i \le N-1, \quad j = i+1, \\ -\alpha_{i-\frac{1}{2}} - \beta_{i+\frac{1}{2}} & 1 \le i \le N, \quad j = i, \\ \beta_{i-\frac{1}{2}} & 2 \le i \le N, \quad j = i-1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.17}$$

We set $\alpha_{\frac{1}{2}} = 0$ and $\beta_{N+\frac{1}{2}} = 0$, which correspond to the boundary conditions in (5.5) as follows

$$\begin{aligned} F(x_{\frac{1}{2}}) = 0 \quad &\Leftrightarrow \quad \alpha_{\frac{1}{2}} = 0 \\ F(x_{N+\frac{1}{2}}) = 0 \quad &\Leftrightarrow \quad \beta_{N+\frac{1}{2}} = 0. \end{aligned} \tag{5.18}$$

The matrix $\mathcal{G}$ in (5.14) is defined as follows

$$\mathcal{G} := \lambda(G - I),$$

where I denotes the $N-$dimensional identity matrix and G is the matrix with normalized columns as follows

$$G_{ij} := \frac{h\, g(x_i - x_j)}{\sum_{k=1}^{N} h\, g(x_k - x_j)}. \tag{5.19}$$

The choice of the normalization in (5.19) is discussed in Section 5.3.

The step leading to the full discretization of our FP problem consists of applying a time-discretization method to (5.14). We choose an operator splitting method, exploiting the fact that the semidiscretized system (5.14) is naturally decoupled into two linear operators. The idea behind a splitting method is to divide the evolution problem into simpler sequential sub-problems that are separately solved with different methods. Setting δt also as the splitting time step, we apply the Strang-Marchuk (SM) splitting scheme [45, 50, 65, 90]. In the following, we refer to the time-continuous solution of the splitting scheme as $f_{\mathrm{SP}}(t)$. The initial data is set as follows, $f_{\mathrm{SP}}(t_0) := f_{\mathrm{SD}}(t_0)$, where f_{SP} is the splitting solution and f_{SD} is the solution of the semi-discretized system (5.14) without splitting.

In each time interval $[t_n, t_{n+1}]$, given the splitting solution $f_{\mathrm{SP}}(t_n)$, the following subproblems, connected via the initial conditions, are solved

$$\begin{aligned}
&1.\ \begin{cases} \phi_1'(t) = \mathcal{A}\phi_1(t) \\ \phi_1(t_n) = f_{\mathrm{SP}}(t_n) \end{cases} && t \in [t_n, t_{n+\frac{1}{2}}] \\
&2.\ \begin{cases} \phi_2'(t) = \mathcal{G}\phi_2(t) + \Psi(t) \\ \phi_2(t_n) = \phi_1(t_{n+\frac{1}{2}}) \end{cases} && t \in [t_n, t_{n+1}] \\
&3.\ \begin{cases} \phi_3'(t) = \mathcal{A}\phi_3(t) \\ \phi_3(t_{n+\frac{1}{2}}) = \phi_2(t_{n+1}) \end{cases} && t \in [t_{n+\frac{1}{2}}, t_{n+1}]. \\
&4.\ f_{\mathrm{SP}}(t_{n+1}) := \phi_3(t_{n+1})
\end{aligned} \tag{5.20}$$

This system of continuous-time equations is approximated by time discretization. The fully discrete numerical solution will be referred to as $\hat{f} = (f_j^n)$, $j = 1, ..., N$, $n = 0, ..., M$. We propose to use two different time discretization methods that together with the space discretization give the schemes named SIMEX1 and SIMEX2.

In SIMEX1, the solution of the first and third step of (5.20) is carried out with the implicit Euler method, while the second step is explicit, in order to avoid the drawback of inverting the dense matrix $\mathcal{G}$. Given f^n at time t_n, the three initial

value problems in (5.20) read as follows

$$\begin{aligned} &1.\ \frac{f^{n+\frac{1}{2}} - f^n}{\delta t/2} = \mathcal{A} f^{n+\frac{1}{2}} \\ &2.\ \frac{f^{(n+\frac{1}{2})^*} - f^{n+\frac{1}{2}}}{\delta t} = \mathcal{G} f^{n+\frac{1}{2}} + \Psi(t_n) \\ &3.\ \frac{f^{n+1} - f^{(n+\frac{1}{2})^*}}{\delta t/2} = \mathcal{A} f^{n+1}, \end{aligned} \tag{5.21}$$

where the unknown are sequentially solved: $f^n \to f^{n+\frac{1}{2}} \to f^{(n+\frac{1}{2})^*} \to f^{n+1}$.

The time discretization of (5.20) in SIMEX2 is carried out with the predictor corrector method. Given f^n at time t_n, the discretization of the three initial value problems in (5.20) take the following form

$$\begin{aligned} &1.\ \begin{cases} \bar{f}^{n+\frac{1}{2}} = f^n + \frac{\delta t}{2} \mathcal{A} f^n \\ f^{n+\frac{1}{2}} = f^n + \frac{\delta t}{4} \left[\mathcal{A} f^n + \mathcal{A} \bar{f}^{n+\frac{1}{2}} \right] \end{cases} \\ &2.\ \begin{cases} \bar{f}^{n+\frac{1}{2}^*} = f^{n+\frac{1}{2}} + \delta t [\mathcal{G} f^{n+\frac{1}{2}} + \Psi(t_n)] \\ f^{n+\frac{1}{2}^*} = f^{n+\frac{1}{2}} + \frac{\delta t}{2} \left[\mathcal{G} f^{n+\frac{1}{2}} + \Psi(t_n) + \mathcal{G} \bar{f}^{n+\frac{1}{2}^*} + \Psi(t_{n+1}) \right] \end{cases} \\ &3.\ \begin{cases} \bar{f}^{n+1} = f^{n+\frac{1}{2}^*} + \frac{\delta t}{2} \mathcal{A} f^{n+\frac{1}{2}^*} \\ f^{n+1} = f^{n+\frac{1}{2}^*} + \frac{\delta t}{4} \left[\mathcal{A} f^{n+\frac{1}{2}^*} + \mathcal{A} \bar{f}^{n+1} \right]. \end{cases} \end{aligned} \tag{5.22}$$

5.2 Accuracy analysis

In this section, we investigate stability and accuracy properties of the SIMEX1 and SIMEX2 schemes, given by (5.21) and (5.22), respectively. After determining the order of convergence of the spatial discretization, we focus on the rate of convergence in time.

The discrete L^2-scalar product of two grid functions u and v on $\Omega_h \times I_{\delta t}$ is defined as follows

$$(u, v)_{L^2_{h,\delta t}} := h \delta t \sum_{j=1}^{N} \sum_{n=0}^{M} u_j^n v_j^n,$$

with associated norm $\|u\|_{L^2_{h,\delta t}} := \sqrt{(u,u)_{L^2_{h,\delta t}}}$. In a similar fashion, the discrete L^2_h inner product and norm are defined for functions w, z on the spatial grid Ω_h as follows

$$(w, z)_{L^2_h} := h \sum_{j=1}^{N} w_j z_j \text{ and } \|w\|_{L^2_h} := \sqrt{(w,w)_{L^2_h}}. \tag{5.23}$$

We aim at comparing the continuous PIDE solution f of (5.5) and the numerical solution $\hat{f}$, which is defined on the grid points of $\Omega_h \times I_{\delta t}$. We have the following inequality

$$\|f - \hat{f}\|_{L^2_{h,\delta t}} \leq \|f_h - f_{\mathrm{SD}}\|_{L^2_{h,\delta t}} + \|f_{\mathrm{SD}} - f_{\mathrm{SP}}\|_{L^2_{h,\delta t}} + \|f_{\mathrm{SP}} - \hat{f}\|_{L^2_{h,\delta t}}, \tag{5.24}$$

where $f_h(t) \in \mathbb{R}^N$ is the PIDE solution restricted to $\bar{x} \in \Omega_h$, with $f_h(t)_j := f(x_j, t)$, f_{SD} solves (5.14) and f_{SP} is obtained as in (5.20). In (5.24), the $L^2_{h,\delta t}$ norms are computed after evaluating the continuous functions at the points of the meshes Ω_h and $I_{\delta t}$ defined in (5.6).

In the following, we provide bounds for each of the three norms of (5.24). Specifically, we prove in Proposition 10 that $\|f_h - f_{\mathrm{SD}}\|_{L^2_{h,\delta t}} = \mathcal{O}(h^2)$. In Proposition 11, we obtain $\|f_{\mathrm{SD}} - f_{\mathrm{SP}}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t^2)$. Further, for the SIMEX1 scheme, we prove in Proposition 13 that $\|f_{\mathrm{SP}} - \hat{f}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t)$, while in Proposition 15 we obtain $\|f_{\mathrm{SP}} - \hat{f}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t^2)$ for the SIMEX2 scheme.

For our analysis, we assume the following conditions.

- The PIDE solution f of (5.5) is 4 times continuously differentiable with respect to the space variable.
- The function B that defines the flux F in (5.3) is $\mathcal{C}^1(\Omega)$. Therefore, B is Lipschitz continuous, with Lipschitz constant L.
- The function C defining the flux in (5.3) belongs to $\mathcal{C}^3(\Omega)$.
- The PDF g of the integral operator $\mathcal{I}$ in (5.4) is two times differentiable.
- The source term ψ in (5.5) has continuous first-time derivative.

Next, we aim at a bound for the first addend in (5.24). For each $j = 1, ..., N$, we define the following time-continuous quantities

$$\begin{aligned} \varepsilon_j(t) &:= f_h(t)_j - f_{\mathrm{SD}_j}(t), \\ \alpha_j(t) &:= \partial_t f(x_j, t) - ([\mathcal{A} + \mathcal{G}] f_h(t))_j - \Psi_j(t), \end{aligned} \tag{5.25}$$

called spatial discretization error and spatial truncation error, respectively. The spatial truncation error α is the residual obtained by inserting the exact solution f in the semidiscretized equation (5.14).

Notice that f satisfies both the PIDE in (5.5) restricted to the line (x_j, t) and the following equation

$$\partial_t f_h(t)_j = ([\mathcal{A} + \mathcal{G}] f_h(t))_j + \Psi_j(t) + \alpha_j(t),$$

where the spatial truncation error α is

$$\alpha_j(t) = [\mathcal{L} f(x_j, t) - (\mathcal{A} f_h(t))_j] + [\mathcal{I} f(x_j, t) - (\mathcal{G} f_h(t))_j]. \tag{5.26}$$

Proposition 9. *Assume $B \in \mathcal{C}^1(\Omega)$, $C \in \mathcal{C}^3(\Omega)$, $g \in \mathcal{C}^2(\mathbb{R})$. If the solution f of (5.5) has continuous space derivatives up to the 4th order, the spatial truncation error α in (5.26) is consistent of order 2 as follows*

$$\|\alpha(t)\|_{L_h^2} = \mathcal{O}(h^2).$$

Proof. We first consider the first addend in (5.26). Recalling the definitions of $\mathcal{L}$ and $\mathcal{A}$ in (5.2) and (5.17), respectively, we have

$$\begin{aligned}
&\mathcal{L}f(x_j,t) - (\mathcal{A}f_h(t))_j = \\
&= \partial_x(Bf_h)(x_j) - \frac{1}{h}\left(B_{j+\frac{1}{2}}[(1-\delta_{j+\frac{1}{2}})f_{j+1} + \delta_{j+\frac{1}{2}}f_j] - B_{j-\frac{1}{2}}[(1-\delta_{j-\frac{1}{2}})f_j + \delta_{j-\frac{1}{2}}f_{j-1}]\right) \\
&+ \partial_x(C\partial_x f_h)(x_j) - \frac{1}{h^2}\left(C_{j+\frac{1}{2}}(f_{j+1}-f_j) - C_{j-\frac{1}{2}}(f_j - f_{j-1})\right).
\end{aligned}$$

We consider the following Taylor expansions, holding for $h, w \to 0$

$$\begin{aligned}
f_h(x_{j\pm1}) &= f_h(x_j) \pm h\partial_x f_h(x_j) + \frac{h^2}{2}\partial^2_{xx} f_h(x_j) \pm \frac{h^3}{3!}\partial^3_{xxx} f_h(x_j) + \frac{h^4}{4!}\partial^4_{xxxx} f_h(x_j) + \\
&\quad + \mathcal{O}(h^5) \\
B(x_{j\pm\frac{1}{2}}) &= B(x_j) \pm \frac{h}{2}\partial_x B(x_j) + \mathcal{O}(h^2) \\
C(x_{j\pm\frac{1}{2}}) &= C(x_j) \pm \frac{h}{2}\partial_x C(x_j) + \frac{h^2}{8}\partial^2_{xx} C(x_j) \pm \frac{h^3}{48}\partial^3_{xxx} C(x_j) + \mathcal{O}(h^4) \\
e^w &= \sum_{k=0}^{\infty} \frac{w^k}{k!}.
\end{aligned}$$

Exploiting the Taylor expansion, we have

$$\begin{aligned}
&\partial_x(Bf_h)(x_j) - \frac{1}{h}\left(B_{j+\frac{1}{2}}[(1-\delta_{j+\frac{1}{2}})f_{j+1} + \delta_{j+\frac{1}{2}}f_j] - B_{j-\frac{1}{2}}[(1-\delta_{j-\frac{1}{2}})f_j + \delta_{j-\frac{1}{2}}f_{j-1}]\right) = \\
&= \partial_x f_h(x_j,t)B(x_j)(\delta_{j+\frac{1}{2}} - \delta_{j-\frac{1}{2}}) + \frac{h}{2}\partial_x(\partial_x f_h B)(x_j)(1 - \delta_{j+\frac{1}{2}} - \delta_{j-\frac{1}{2}}) + \mathcal{O}(h^2).
\end{aligned}$$

We also have

$$\begin{aligned}
&\partial_x(C\partial_x f_h)(x_j) - \frac{1}{h^2}\left(C_{j+\frac{1}{2}}(f_{j+1}-f_j) - C_{j-\frac{1}{2}}(f_j - f_{j-1})\right) = \\
&= \frac{h^2}{2}\left(\frac{1}{12}\partial^3_{xxx}C(x_j)\partial_x f_h(x_j) + \frac{1}{4}\partial^2_{xx}C(x_j)\partial^2_{xx}f_h(x_j) + \frac{1}{6}\partial_x C(x_j)\partial^3_{xxx}f_h(x_j) + \right. \\
&\left. + \frac{1}{6}C(x_j)\partial^4_{xxxx}f_h(x_j)\right) + \mathcal{O}(h^4).
\end{aligned}$$

Moreover, exploiting the definition of δ in (5.12) and the Taylor expansion of the exponential function, we obtain the following relations

$$\delta_{j-\frac{1}{2}} - \delta_{j+\frac{1}{2}} = \mathcal{O}(h^2) \quad \text{and} \quad 1 - \delta_{j+\frac{1}{2}} - \delta_{j-\frac{1}{2}} = \mathcal{O}(h),$$

which lead us to the conclusion

$$\mathcal{L}f(x_j,t) - (\mathcal{A}f_h(t))_j = \mathcal{O}(h^2).$$

Further, we consider the second addend in (5.26). We have

$$\begin{aligned}
&\mathcal{I}f(x_j,t) - (\mathcal{G}f_h(t))_j = \lambda\left(\int_\Omega f(y,t)g(x_j-y)dy - (Gf_h(t))_j\right)\\
&= \lambda\left(\int_\Omega f(y,t)g(x_j-y)dy - \sum_{i=1}^N \frac{h\,g(x_j-x_i)}{\sum_{k=1}^N h\,g(x_k-x_i)} f(x_i,t)\right)\\
&= \lambda\left(\int_\Omega f(y,t)g(x_j-y)dy - \sum_{i=1}^N h\,g(x_j-x_i)f(x_i,t)\right)\\
&+ \lambda\left(\sum_{i=1}^N h\,g(x_j-x_i)f(x_i,t) - \sum_{i=1}^N \frac{h\,g(x_j-x_i)}{\sum_{k=1}^N h\,g(x_k-x_i)} f(x_i,t)\right)\\
&= \mathcal{O}(h^2) + \lambda\sum_{i=1}^N \frac{h\,g(x_j-x_i)f(x_i,t)}{\sum_{k=1}^N h\,g(x_k-x_i)}\left(\sum_{k=1}^N h\,g(x_k-x_i) - 1\right) = \mathcal{O}(h^2)
\end{aligned}$$

We notice that the last term consists of the midpoint rule to calculate the integrals over Ω of the following two functions

$$\varphi_1(y) := f(y,t)g(x_j-y) \text{ and } \varphi_2(y) := g(y-x_i), \text{ for each } j = 1,...,N.$$

The error associated to the midpoint rule is given by

$$\int_\Omega \varphi_l(y)dy - \sum_{k=1}^N h\,\varphi_l(x_k) = \mathcal{O}(h^2) \text{ for } l = 1,2.$$

In particular, we have $|\sum_{k=1}^N h\,g(x_k-x_i) - 1| = \mathcal{O}(h^2)$, hence

$$\alpha_j(t) = \mathcal{O}(h^2) \text{ for each } j = 1,...,N.$$

Since $\|\alpha(t)\|_{L_h^2}^2 = \sum_{j=1}^N h|\alpha_j(t)|^2$, we can conclude that $\|\alpha(t)\|_{L_h^2} = \mathcal{O}(h^2)$.

□

Proposition 10. *Let f be the solution to (5.5) and f_{SD} be the solution to (5.14). Then the following holds $\|f_h - f_{SD}\|_{L^2_{h,\delta t}} = \mathcal{O}(h^2)$.*

Proof. By definition of spatial truncation error in (5.25), we have that $\varepsilon(t)$ satisfies the following initial value problem

$$\begin{cases} \varepsilon'(t) = [\mathcal{A}+\mathcal{G}]\varepsilon(t) + \alpha(t) \\ \varepsilon(0) = 0. \end{cases} \tag{5.27}$$

We carry out the proof within the framework given by logarithmic norms [50]. Recall that, given an arbitrary matrix norm $|||\cdot|||$, the logarithmic norm μ is defined for a square matrix M as follows

$$\mu(M) := \lim_{\tau \to 0} \frac{|||I + \tau M||| - 1}{\tau}.$$

If the norm $|||\cdot|||$ stems from an inner product $(\cdot, \cdot)$, the logarithmic norm μ of M satisfies the following relation

$$\mu(M) := \sup_{v \neq 0} \frac{(Mv, v)}{(v, v)}.$$

We exploit the following equivalence, holding for any matrix M with associated norm $|||\cdot|||$ and logarithmic norm μ, for $\omega \in \mathbb{R}$. We have

$$\mu(M) \leq \omega \qquad \Leftrightarrow \qquad |||e^{\delta t M}||| \leq e^{\delta t \omega}. \tag{5.28}$$

In this proof, we consider the scalar product in L_h^2 and its related norm (5.23). We first provide an estimate for the logarithmic norm $\mu_2(\mathcal{A}+\mathcal{G})$ induced by $(\cdot,\cdot)_{L_h^2}$. We consider the scalar product $(\mathcal{A}\varepsilon(t), \varepsilon(t))_{L_h^2}$, showing that it is nonpositive for all $\varepsilon(t) \in L_h^2$. Thanks to the definition of $\mathcal{A}$ in (5.17), we note that there exist two matrices L and U defined as follows

$$L_{ij} = \begin{cases} -1 & i = 2, ..., N, \quad j = i - 1 \\ 1 & i = 1, ...N, \quad j = i \\ 0 & \text{otherwise,} \end{cases}$$

$$U_{ij} = \begin{cases} \alpha_{i+\frac{1}{2}} & i = 1, ..., N-1, \quad j = i + 1 \\ \beta_{i+\frac{1}{2}} & i = 1, ..., N-1, \quad j = i \\ 0 & \text{otherwise,} \end{cases}$$

such that $-h\mathcal{A} = LU$. Notice that $U_{NN} = 0$. Therefore the determinant of U, and thus of $\mathcal{A}$, is zero, which means that $\mathcal{A}$ has a zero eigenvalue. Furthermore, by inspection, we see that all other leading principal minors of U are positive. Then, we conclude that the matrix $\mathcal{A}$ has $N-1$ negative eigenvalues and a zero eigenvalue. Therefore, $(\mathcal{A}\varepsilon(t), \varepsilon(t))_{L_h^2} \leq 0$ for each $\varepsilon(t) \in \mathbb{R}^N$, which implies that $\mu_2(\mathcal{A}) = 0$.

For the matrix $\mathcal{G}$ we have the following

$$\begin{aligned}(\mathcal{G}\varepsilon(t),\varepsilon(t))_{L_h^2} &= \lambda(G\varepsilon(t),\varepsilon(t))_{L_h^2} - \lambda(\varepsilon(t),\varepsilon(t))_{L_h^2} \\ &= \lambda h\sum_{i=1}^{N}\sum_{j=1}^{N} G_{ij}\varepsilon_i(t)\varepsilon_j(t) - \lambda h\sum_{j=1}^{N}\varepsilon_j^2(t) \\ &\leq \lambda\frac{h}{2}\sum_{i=1}^{N}\left(\sum_{j=1}^{N} G_{ij}\right)\varepsilon_i^2(t) + \lambda\frac{h}{2}\sum_{j=1}^{N}\left(\sum_{i=1}^{N} G_{ij}\right)\varepsilon_j^2(t) - \lambda\|\varepsilon(t)\|_{L_h^2}^2 \\ &\leq c\lambda h^2\|\varepsilon(t)\|_{L_h^2}^2,\end{aligned} \tag{5.29}$$

hence $\mu_2(\mathcal{G}) \leq c\lambda h^2$. Therefore, $\mu_2(\mathcal{A}+\mathcal{G}) \leq \mu_2(\mathcal{A})+\mu_2(\mathcal{G}) \leq c\lambda h^2$. This inequality, together with (5.28), allows us to state the following results

$$\|e^{\mathcal{A}\delta t}\|_{L_h^2} \leq 1, \tag{5.30}$$

$$\|e^{\mathcal{G}\delta t}\|_{L_h^2} \leq e^{c\lambda h^2\delta t}, \tag{5.31}$$

$$\|e^{(\mathcal{A}+\mathcal{G})\delta t}\|_{L_h^2} \leq e^{c\lambda h^2\delta t}. \tag{5.32}$$

By integrating (5.27) and exploiting the inequality (5.32), we conclude that

$$\|\varepsilon(t)\|_{L_h^2} \leq e^{c\lambda h^2 t}\|\varepsilon(0)\|_{L_h^2} + \frac{e^{c\lambda h^2 t}-1}{c\lambda h^2}\max_{0\leq s\leq t}\|\alpha(s)\|_{L_h^2},$$

and therefore $\|f_h - f_{\mathrm{SD}}\|_{L_{h,\delta t}^2} = \mathcal{O}(h^2)$.

□

We now aim at proving a bound for the second addend in (5.24) related to the splitting method. Consider the matrices $\mathcal{A}$ and $\mathcal{G}$ as in (5.17) and (5.19), respectively, and define the operator $\mathcal{S}$ as follows

$$\mathcal{S} := e^{\frac{\delta t}{2}\mathcal{A}}e^{\delta t\mathcal{G}}e^{\frac{\delta t}{2}\mathcal{A}}. \tag{5.33}$$

Consider the time interval $[t_n, t_{n+1}]$. We apply four times the variation of constants formula for ODEs to integrate the ODE systems (5.20). Therefore, the splitting solution can be formally written as follows

$$\begin{aligned}f_{\mathrm{SP}}(t_{n+1}) =& \mathcal{S}f_{\mathrm{SP}}(t_n) + e^{\frac{\delta t}{2}\mathcal{A}}e^{\frac{\delta t}{2}\mathcal{G}}\int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)\mathcal{G}}\Psi(t_n+s)ds + \\ &+ e^{\frac{\delta t}{2}\mathcal{A}}\int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)\mathcal{G}}\Psi(t_{n+\frac{1}{2}}+s)ds.\end{aligned} \tag{5.34}$$

We define with d_n the local truncation splitting error for each $n = 0, ..., M-1$, which is the residual obtained at time t_n by inserting the exact solution of the semidiscretized system (5.14) in the formal expression of the splitting solution (5.34) as follows

$$\begin{aligned} d_n :=& f_{\mathrm{SD}}(t_{n+1}) - \mathcal{S} f_{\mathrm{SD}}(t_n) - e^{\frac{\delta t}{2}\mathcal{A}} e^{\frac{\delta t}{2}\mathcal{G}} \int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)\mathcal{G}} \Psi(t_n + s) ds \\ & - e^{\frac{\delta t}{2}\mathcal{A}} \int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)\mathcal{G}} \Psi(t_{n+\frac{1}{2}} + s) ds. \end{aligned} \tag{5.35}$$

Define the global splitting error at time t_n as $E_n := f_{\mathrm{SD}}(t_n) - f_{\mathrm{SP}}(t_n)$. Subtracting (5.34) from (5.35), we obtain the following relation

$$E_{n+1} = \mathcal{S} E_n + d_n. \tag{5.36}$$

Exploiting the linearity of the solution operator $\mathcal{S}$ and the fact that $E_0 = 0$, we recursively apply (5.36), and obtain

$$E_n = \sum_{k=0}^{n} \mathcal{S}^{n-k} d_k. \tag{5.37}$$

In the remainder of this section, we make use of the following two facts. Given $z \in \mathbb{R}$ and a matrix $M \in \mathbb{R}^{N\times N}$, we have the following two properties. The exponential of the matrix zM is defined by the convergent power series

$$e^{zM} = \sum_{k=0}^{+\infty} \frac{(zM)^k}{k!}. \tag{5.38}$$

For $z \to 0$,

$$(I - zM)^{-1} = I + zM + z^2 M^2 + \mathcal{O}(z^3). \tag{5.39}$$

Proposition 11. *Let Ψ in (5.14) be of class $\mathcal{C}^1([0,T])$. Then $\|f_{SD} - f_{SP}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t^2)$.*

Proof. We first determine the order of the splitting error d_n. By applying twice to (5.14) the variation of constant formula, we have that the exact solution f_{SD} of (5.14) can be written as follows

$$\begin{aligned} f_{\mathrm{SD}}(t_{n+1}) =& e^{\delta t(\mathcal{A}+\mathcal{G})} f_{\mathrm{SD}}(t_n) + e^{\frac{\delta t}{2}(\mathcal{A}+\mathcal{G})} \int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)(\mathcal{A}+\mathcal{G})} \Psi(t_n + s) ds \\ & + \int_0^{\frac{\delta t}{2}} e^{(\frac{\delta t}{2}-s)(\mathcal{A}+\mathcal{G})} \Psi(t_{n+\frac{1}{2}} + s) ds. \end{aligned} \tag{5.40}$$

By using of the matrix series (5.38) and exploiting that $\Psi \in \mathcal{C}^1([0,T])$, we rewrite (5.34) and (5.40) as follows

$$f_{\mathrm{SP}}(t_{n+1}) = \mathcal{S} f_{\mathrm{SP}}(t_n) + \delta t \left(I + \frac{\delta t}{2}(\mathcal{A}+\mathcal{G}) + \frac{\delta t^2}{8}\left(\mathcal{A}^2 + 2\mathcal{A}\mathcal{G} + \mathcal{G}^2\right)\right)(\Psi(t_n) + \mathcal{O}(\delta t)) + \mathcal{O}(\delta t^4) \tag{5.41}$$

and

$$f_{\mathrm{SD}}(t_{n+1}) = e^{\delta t(\mathcal{A}+\mathcal{G})} f_{\mathrm{SD}}(t_n) + \delta t \left(I + \frac{\delta t}{2}(\mathcal{A}+\mathcal{G}) + \frac{\delta t^2}{8}(\mathcal{A}+\mathcal{G})^2\right)(\Psi(t_n) + \mathcal{O}(\delta t)) + + \mathcal{O}(\delta t^4). \tag{5.42}$$

Therefore, by using (5.35) the local splitting error can be rewritten as follows

$$d_n = (e^{\delta t(\mathcal{A}+\mathcal{G})} - \mathcal{S}) f_{\mathrm{SD}}(t_n) + \frac{\delta t^3}{8}\left(\mathcal{G}\mathcal{A} - \mathcal{A}\mathcal{G}\right)\Psi(t_n) + \mathcal{O}(\delta t^4).$$

By applying (5.38) to $e^{\frac{\delta t}{2}\mathcal{A}}$ and $e^{\delta t \mathcal{G}}$, we have that

$$\mathcal{S} = I + \delta t(\mathcal{A}+\mathcal{G}) + \frac{\delta t^2}{2}(\mathcal{A}+\mathcal{G})^2 + + \frac{\delta t^3}{6}\left(\mathcal{A}^3 + \mathcal{G}^3 + \frac{3}{4}\mathcal{A}^2\mathcal{G} + \frac{3}{2}\mathcal{A}\mathcal{G}\mathcal{A} + \frac{3}{2}\mathcal{A}\mathcal{G}^2 + \frac{3}{4}\mathcal{G}\mathcal{A}^2 + \frac{3}{2}\mathcal{G}^2\mathcal{A}\right) + \mathcal{O}(\delta t^4). \tag{5.43}$$

Furthermore, we have

$$e^{\delta t(\mathcal{A}+\mathcal{G})} = I + \delta t(\mathcal{A}+\mathcal{G}) + \frac{\delta t^2}{2}(\mathcal{A}+\mathcal{G})^2 + \frac{\delta t^3}{6}(\mathcal{A}+\mathcal{G})^3 + \mathcal{O}(\delta t^4),$$

it is clear that $e^{\delta t(\mathcal{A}+\mathcal{G})} - \mathcal{S} = \mathcal{O}(\delta t^3)$. Hence, $d_n = \mathcal{O}(\delta t^3)$ for each $n = 0, ..., M-1$.

Next, we consider the L^2_h norm of (5.37), hence

$$\|E_n\|_{L^2_h} \leq \sum_{k=0}^{n} e^{c\lambda h^2 \delta t(n-k)} \|d_k\|_{L^2_h} \leq c_1\, \delta t^2,$$

where we used (5.31). Therefore, $\|E_n\|_{L^2_h} = \mathcal{O}(\delta t^2)$ and

$$\|f_{\mathrm{SD}} - f_{\mathrm{SP}}\|_{L^2_{h,\delta t}} = \sqrt{\sum_{n=0}^{M} \delta t \|E_n\|^2_{L^2_h}} \leq c_2\, \delta t^2,$$

that completes the proof. □

In the remainder of this section, we aim to prove a bound for the third term in (5.24), where the full numerical solution $\hat{f} = f_j^n$, $j = 1, ..., N$, $n = 0, ..., M$, is given by either the Euler discretization or by the predictor corrector scheme. By the definition of the operator $\mathcal{S}$ in (5.33), we write the solution of (5.20) in a convenient form as follows

$$f_{\mathrm{SP}}(t_{n+1}) = \mathcal{S} f_{\mathrm{SP}}(t_n) + e^{\frac{\delta t}{2}\mathcal{A}} \int_0^{\delta t} e^{(\delta t - s)\mathcal{G}} \Psi(t_n + s) ds, \tag{5.44}$$

where we applied three times the variation of constants formula [50] to (5.20).

Next, we write (5.21) in a compact form. Note that the nonsingularity of the matrix $I - \frac{\delta t}{2}\mathcal{A}$ is guaranteed also under the condition $\delta t \leq 2/L$, where L is the Lipschitz constant of B. Given f^n, the computation of f^{n+1} is carried out as follows

$$\begin{aligned} f^{n+1} &= \left(I - \frac{\delta t}{2}\mathcal{A}\right)^{-1} \left((I + \delta t \mathcal{G})(I - \frac{\delta t}{2}\mathcal{A})^{-1} f^n + \delta t \Psi(t_n)\right) \\ &= R_1(\mathcal{A}, \mathcal{G}, \delta t) f^n + \delta t \left(I - \frac{\delta t}{2}\mathcal{A}\right)^{-1} \Psi(t_n), \end{aligned} \tag{5.45}$$

where

$$R_1(\mathcal{A}, \mathcal{G}, \delta t) := \left(I - \frac{\delta t}{2}\mathcal{A}\right)^{-1} (I + \delta t \mathcal{G}) \left(I - \frac{\delta t}{2}\mathcal{A}\right)^{-1}$$

is the amplification factor.

For each time window $[t_n, t_{n+1}]$, we define the respective time truncation error T_n, obtained by inserting the formal splitting solution f_{SP} defined in (5.44) in the numerical approximation given by (5.45). We have

$$T_n := f_{\mathrm{SP}}(t_{n+1}) - R_1(\mathcal{A}, \mathcal{G}, \delta t) f_{\mathrm{SP}}(t_n) - \delta t \left(I - \frac{\delta t}{2}\mathcal{A}\right)^{-1} \Psi(t_n). \tag{5.46}$$

Proposition 12. *(Consistency of SIMEX1) Let $\delta t \leq 2/L$. The truncation error (5.46) is of order $\mathcal{O}(\delta t^2)$.*

Proof. Recall the definitions of the splitting solution and truncation error, in (5.44) and in (5.46), respectively. Exploiting the fact that $\Psi \in \mathcal{C}^1(I)$ and making use of (5.38), we have that

$$T_n = (\mathcal{S} - R_1(\mathcal{A}, \mathcal{G}, \delta t)) f_{\mathrm{SP}}(t_n) + \delta t^2 \left(\frac{\mathcal{G}}{2} + \mathcal{O}(\delta t)\right) \Psi(t_n) + \mathcal{O}(\delta t^3). \tag{5.47}$$

We consider the Taylor expansion of $(I - \frac{\delta t}{2}A)^{-1}$ as in (5.39), which exists since $\delta t \leq 2/L$, and note that the amplification factor R_1 can be rewritten as follows

$$R_1(\mathcal{A}, \mathcal{G}, \delta t) = I + \delta t(\mathcal{A} + \mathcal{G}) + \frac{\delta t^2}{2}\left(\frac{3}{2}\mathcal{A}^2 + \mathcal{A}\mathcal{G} + \mathcal{G}\mathcal{A}\right) + \mathcal{O}(\delta t^3).$$

Recalling the expansion of $\mathcal{S}$ in (5.43), we can state that $\mathcal{S} - R_1(\mathcal{A}, \mathcal{G}, \delta t) = \mathcal{O}(\delta t^2)$.

These observations lead to the conclusion that $T_n = \mathcal{O}(\delta t^2)$, hence the proof is completed.

□

We define, for each $n = 1, ..., M$, the time discretization error e_n as follows

$$e_n := f_{\mathrm{SP}}(t_n) - f^n,$$

such that by subtracting (5.46) from (5.45), we obtain the following relation

$$e_{n+1} = R_1(\mathcal{A}, \mathcal{G}, \delta t)e_n + T_n. \tag{5.48}$$

Proposition 13. *(Accuracy of SIMEX1) If δt is chosen such that $\|R_1(\mathcal{A}, \mathcal{G}, \delta t)\|_{L_h^2} \leq 1 + c\lambda h^2 \delta t$, then $\|f_{SP} - \hat{f}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t)$.*

Proof. First, we notice that the bound for $\|R_1\|_{L_h^2}$ in the hypothesis results from the definition of R_1 and from the fact that (5.30) implies

$$\|(I - \delta t\mathcal{A})^{-1}\|_{L_h^2} \leq 1. \tag{5.49}$$

We consider the L_h^2 norm of (5.48), obtaining

$$\|e_{n+1}\|_{L_h^2} \leq \|R_1(\mathcal{A}, \mathcal{G}, \delta t)\|_{L_h^2}\|e_n\|_{L_h^2} + \|T_n\|_{L_h^2} \leq (1 + c\lambda h^2 \delta t)\|e_n\|_{L_h^2} + \|T_n\|_{L_h^2},$$

This recursive relation gives the following

$$\|e_n\|_{L_h^2} \leq e^{c\lambda h^2 T}\left(\|e_0\|_{L_h^2} + \sum_{k=0}^{n-1}\|T_k\|_{L_h^2}\right),$$

that is the stability of the discrete operator R_1. Thus, $\|e_n\|_{L_h^2} = \mathcal{O}(\delta t)$. By noting that $\|f_{\mathrm{SP}} - \hat{f}\|^2_{L^2_{h,\delta t}} = \sum_{n=1}^{M} \delta t\|e_n\|^2_{L_h^2}$, the proof is completed.

□

Next, we write (5.22) in a compact form. Given f^n, the computation of f^{n+1} is carried out as follows

$$f^{n+1} = R\left(\frac{\delta t}{2}\mathcal{A}\right) R(\delta t\mathcal{G}) R\left(\frac{\delta t}{2}\mathcal{A}\right) f^n + R\left(\frac{\delta t}{2}\mathcal{A}\right)\bar{\Psi}_n, \tag{5.50}$$

where

$$\bar{\Psi}_n := \frac{\delta t}{2}\left[(I + \delta t\mathcal{G})\Psi(t_n) + \Psi(t_{n+1})\right]$$

and the function R is the amplification factor; given a matrix M and $z \in \mathbb{R}$, R is defined as

$$R(zM) := I + zM + \frac{z^2}{2}M^2.$$

For each time window $[t_n, t_{n+1}]$ we define the respective time truncation error T_n, obtained by inserting the formal splitting solution f_{SP} defined in (5.44) in the numerical approximation given by (5.50)

$$T_n := f_{\mathrm{SP}}(t_{n+1}) - R\left(\frac{\delta t}{2}\mathcal{A}\right) R(\delta t\mathcal{G}) R\left(\frac{\delta t}{2}\mathcal{A}\right) f_{\mathrm{SP}}(t_n) - R\left(\frac{\delta t}{2}\mathcal{A}\right)\bar{\Psi}_n. \tag{5.51}$$

Proposition 14. *(Consistency of SIMEX2) The truncation error* (5.51) *is of order* $\mathcal{O}(\delta t^3)$.

Proof. Recall the definitions of the splitting solution and of the truncation error, in (5.44) and in (5.51), respectively. Exploiting the fact that $\Psi \in \mathcal{C}^1(I)$ and making use of (5.38), we have

$$T_n = (\mathcal{S} - R_2(\mathcal{A}, \mathcal{G}, \delta t)) f_{\mathrm{SP}}(t_n) + \delta t^3 \frac{\mathcal{G}^3}{6} \Psi(t_n) + \mathcal{O}(\delta t^4), \tag{5.52}$$

where

$$R_2(\mathcal{A}, \mathcal{G}, \delta t) := R\left(\frac{\delta t}{2}\mathcal{A}\right) R(\delta t \mathcal{G}) R\left(\frac{\delta t}{2}\mathcal{A}\right).$$

By considering the Taylor expansion of $e^{\frac{\delta t}{2}\mathcal{A}}$ and $e^{\delta t \mathcal{G}}$ as in (5.38) up to the 4-th order and (5.43), we note that the solution operator $\mathcal{S}$ can be written as follows

$$\mathcal{S} = R_2(\mathcal{A}, \mathcal{G}, \delta t) + \frac{\delta t^3}{6}\left(\frac{\mathcal{A}^3}{2} + \mathcal{G}^3\right) + \mathcal{O}(\delta t^4),$$

and therefore $\mathcal{S} - R_2(\mathcal{A}, \mathcal{G}, \delta t) = \mathcal{O}(\delta t^3)$. These observations lead to the conclusion that $T_n = \mathcal{O}(\delta t^3)$, hence the proof is completed.

□

We define for each $n = 1, ..., M$ the time discretization error e_n as follows

$$e_n := f_{\mathrm{SP}}(t_n) - f^n,$$

such that by subtracting (5.51) from (5.50), we obtain the following relation

$$e_{n+1} = R_2(\mathcal{A}, \mathcal{G}, \delta t) e_n + T_n. \tag{5.53}$$

Proposition 15. *(Accuracy of SIMEX2) If* δt *is chosen such that* $\|R_2(\mathcal{A}, \mathcal{G}, \delta t)\|_{L_h^2} \leq e^{c\lambda h^2 \delta t}$, *then* $\|f_{SP} - \hat{f}\|_{L^2_{h,\delta t}} = \mathcal{O}(\delta t^2)$.

Proof. First, we notice that the bound for $\|R_2\|$ in the hypothesis results from the definition of R_2 and from the inequalities (5.30) and (5.31). We apply the L_h^2 norm to (5.53), obtaining

$$\|e_{n+1}\|_{L_h^2} \leq \|R_2(\mathcal{A}, \mathcal{G}, \delta t)\|_{L_h^2} \|e_n\|_{L_h^2} + \|T_n\|_{L_h^2} \leq e^{c\lambda h^2 \delta t} \|e_n\|_{L_h^2} + \|T_n\|_{L_h^2},$$

This recursive relation gives the following

$$\|e_n\|_{L_h^2} \leq e^{c\lambda h^2 T}\left(\|e_0\|_{L_h^2} + \sum_{k=0}^{n-1} \|T_k\|_{L_h^2}\right),$$

that is the stability of the discrete operator R_2. Thus, $\|e_n\|_{L_h^2} = \mathcal{O}(\delta t^2)$. By noting that $\|f_{\mathrm{SP}} - \hat{f}\|^2_{L^2_{h,\delta t}} = \sum_{n=1}^{M} \delta t \|e_n\|^2_{L_h^2}$, the proof is completed.

□

5.3 Positivity and conservativeness of the SIMEX schemes

In this section, we prove that the SIMEX1 and SIMEX2 schemes are conservative and positive preserving, in the sense that the numerical solution of (5.5) preserves the properties of the continuous PDF of (5.1). In particular, we apply the midpoint rule with the aim to show that the following integral is conserved

$$\int_\Omega f(x,t_n)dx = \sum_{i=1}^{N} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x,t_n)dx \approx h \sum_{i=1}^{N} f_i^n,$$

for each $n = 1, ..., M$.

First, we focus on the SIMEX1 scheme, where the time discretization is given by the Euler scheme. Given the numerical solution f^n at time t_n, we compute f^{n+1} as follows

$$\begin{aligned} &1.\ \frac{f^{n+\frac{1}{2}} - f^n}{\frac{\delta t}{2}} = \mathcal{A} f^{n+\frac{1}{2}} \\ &2.\ \frac{f^{n+\frac{1}{2}^*} - f^{n+\frac{1}{2}}}{\delta t} = \mathcal{G} f^{n+\frac{1}{2}} \\ &3.\ \frac{f^{n+1} - f^{n+\frac{1}{2}^*}}{\frac{\delta t}{2}} = \mathcal{A} f^{n+1}. \end{aligned} \tag{5.54}$$

We have the following result concerning the positivity preserving property of the SIMEX1 scheme.

Proposition 16. *Consider* (5.54) *and assume that* $\delta t \leq \min\{\frac{1}{\lambda}, \frac{2}{L}\}$*, with* λ *rate of jumps of the compound Poisson process* P *in* (5.1)*. Define with* L *the the Lipschitz constant of the function* B *that defines* F *in* (5.3)*. If* $f_j^n \geq 0$ *for each* $j = 1, ...N$*, then* $f_j^{n+1} \geq 0$ *for each* $j = 1, ..., N$*.*

Proof. Let us consider each step of (5.54).

1. Given $f_j^n \geq 0$ for each $j = 1, ..., N$, then $f_j^{n+1/2} \geq 0$ for each $j = 1, ..., N$. In fact, the evolution matrix of Step 1. is given by $I - \frac{\delta t}{2}\mathcal{A}$ and it is a nonsingular M-matrix provided that it is diagonal dominant. According to the definition of $\mathcal{A}$ given in (5.17), this property is satisfied when

$$\frac{\delta t}{2h} < |\alpha_{i+\frac{1}{2}} - \beta_{i+\frac{1}{2}} - (\alpha_{i-\frac{1}{2}} - \beta_{i-\frac{1}{2}})|^{-1}$$

for each $i = 1, ..., N$, that is true for $\delta t \leq 2/L$. Since for nonsingular M-matrix it is $M^{-1} \geq 0$, the assertion is proved.

2. The Step 2 in (5.54) can be recast, for each $j = 1, ..., N$, as follows

$$f_j^{(n+\frac{1}{2})^*} = \lambda\delta t (G f^{n+\frac{1}{2}})_j + (1 - \lambda\delta t) f_j^{n+\frac{1}{2}},$$

where $f_j^{n+\frac{1}{2}} \geq 0$ by hypothesis. Since G has nonnegative components and δt is such that both $\lambda\delta t$ and $(1 - \lambda\delta t)$ are nonnegative, then $f_j^{(n+\frac{1}{2})^*} \geq 0$ for each $j = 1, ..., N$.

3. In Step 3. in (5.54), given $f_j^{(n+\frac{1}{2})^*} \geq 0$ for each $j = 1, ..., N$, then $f_j^{n+1} \geq 0$, according to the argument used for analyzing Step 1.

□

Proposition 17. *Consider* (5.54). *The total probability is conserved, in the sense that*

$$\sum_{j=1}^{N} f_j^n = \sum_{j=1}^{N} f_j^{n+1}.$$

Proof. Let us separately consider each step of (5.54).

1. The definition of $\mathcal{A}$ in (5.17) and (5.18) leads to the following

$$\sum_{j=1}^{N} \mathcal{A}_{jk} = 0,$$

which implies

$$\sum_{j=1}^{N} (\mathcal{A}v)_j = \sum_{k=1}^{N} \left(\sum_{j=1}^{N} \mathcal{A}_{jk} \right) v_k = 0 \quad \text{for each } v \in \mathbb{R}^N. \tag{5.55}$$

By summing over $j = 1, ..., N$ both sides of the first equation in (5.54), we have that

$$\sum_{j=1}^{N} f_j^{n+\frac{1}{2}} = \sum_{j=1}^{N} f_j^{n+1}.$$

2. Let us consider

$$\frac{f_j^{(n+\frac{1}{2})^*} - f_j^{n+\frac{1}{2}}}{\delta t} = (\mathcal{G} f^{n+\frac{1}{2}})_j.$$

Summing over j, we have that

$$\sum_{j=1}^{N} f_j^{(n+\frac{1}{2})^*} = \sum_{j=1}^{N} f_j^{n+\frac{1}{2}} + \delta t \sum_{j=1}^{N} (\mathcal{G} f^{n+\frac{1}{2}})_j,$$

thus it is sufficient to show that $\sum_{j=1}^{N}(\mathcal{G}f^{n+\frac{1}{2}})_j = 0$. This follows from the construction of the matrix G, in fact

$$\sum_{j=1}^{N}(\mathcal{G}f^{n+\frac{1}{2}})_j = \sum_{j=1}^{N}\lambda\left[\sum_{k=1}^{N}\left(G_{jk}f_k^{n+\frac{1}{2}}\right)\right] - \lambda\sum_{j=1}^{N}f_j^{n+\frac{1}{2}} =$$
$$= \lambda\left[\sum_{k=1}^{N}f_k^{n+\frac{1}{2}}\left(\sum_{j=1}^{N}G_{jk}\right) - \sum_{j=1}^{N}f_j^{n+\frac{1}{2}}\right] = 0,$$

since $\sum_{j=1}^{N}G_{jk} = 1$ holds independently of k, as defined in (5.19).

3. In the same fashion as in the first step, we use (5.55) and obtain

$$\sum_{j=1}^{N}f_j^{n+1} = \sum_{j=1}^{N}f_j^{(n+\frac{1}{2})^*}.$$

□

Next, we focus on the SIMEX2 scheme, where the time discretization is given by the predictor-corrector scheme and discuss its positivity-preserving. Given the numerical solution f^n at time t_n, the three steps required to compute f^{n+1} are as follows

$$\begin{aligned}
&1.\begin{cases}\bar{f}^{n+\frac{1}{2}} = f^n + \frac{\delta t}{2}\mathcal{A}f^n\\ f^{n+\frac{1}{2}} = f^n + \frac{\delta t}{4}\left[\mathcal{A}f^n + \mathcal{A}\bar{f}^{n+\frac{1}{2}}\right]\end{cases}\\
&2.\begin{cases}\bar{f}^{n+\frac{1}{2}^*} = f^{n+\frac{1}{2}} + \delta t\mathcal{G}f^{n+\frac{1}{2}}\\ f^{n+\frac{1}{2}^*} = f^{n+\frac{1}{2}} + \frac{\delta t}{2}\left[\mathcal{G}f^{n+\frac{1}{2}} + \mathcal{G}\bar{f}^{n+\frac{1}{2}}\right]\end{cases}\\
&3.\begin{cases}\bar{f}^{n+1} = f^{n+\frac{1}{2}^*} + \frac{\delta t}{2}\mathcal{A}f^{n+\frac{1}{2}^*}\\ f^{n+1} = f^{n+\frac{1}{2}^*} + \frac{\delta t}{4}\left[\mathcal{A}f^{n+\frac{1}{2}^*} + \mathcal{A}\bar{f}^{n+1}\right].\end{cases}
\end{aligned} \tag{5.56}$$

Proposition 18. *Consider* (5.56) *and suppose that* $\delta t \leq \min\{\frac{1}{\lambda}, \frac{2}{\max_j|\mathcal{A}_{jj}|}\}$, *where* λ *is the rate of the compound Poisson process* P *in* (5.1), *and* $\mathcal{A}_{jj}$ *are the diagonal elements of* $\mathcal{A}$ *as defined in* (5.17). *If* $f_j^n \geq 0$ *for each* $j = 1, ..., N$, *then* $f_j^{n+1} \geq 0$ *for each* $j = 1, ..., N$.

Proof. Let use separately consider each step of (5.56).

1. This step can be rewritten as

$$f^{n+\frac{1}{2}} = \bar{\mathcal{A}} f^n,$$

where

$$\bar{\mathcal{A}} := I + \frac{\delta t}{2}\mathcal{A} + \frac{\delta t^2}{8}\mathcal{A}^2 = \frac{1}{8}\left[4I + (2I + \delta t \mathcal{A})^2\right]$$

and I is the N-dimensional identity matrix.

Given $f_j^n \geq 0$ for each $j = 1, ..., N$, then $f_j^{n+1/2} \geq 0$ for each $j = 1, ..., N$, provided that $\bar{A}$ has positive entries. This condition holds if $2I + \delta t \mathcal{A}$ has positive entries.

From the definition of $\mathcal{A}$ in (5.17), we have that $\mathcal{A}_{ii} \leq 0$, while $\mathcal{A}_{ij} \geq 0$ for $i \neq j$. Therefore, it is sufficient to choose $(2 + \delta t \mathcal{A}_{jj}) \geq 0$ for each j, which is satisfied since $\delta t \leq \frac{2}{\max_j |\mathcal{A}_{jj}|}$ by hypothesis.

2. The intermediate step in (5.56) can be rewritten as follows

$$f^{(n+\frac{1}{2})^*} = \frac{1}{2}\left(I + (I + \delta t \mathcal{G})^2\right) f^{n+\frac{1}{2}},$$

where I is the N-dimensional identity matrix. The time step $\delta t \geq 0$ is chosen such that $(1 - \lambda \delta t) \geq 0$ and thus $I + \delta t \mathcal{G} = (1 - \lambda \delta t) I + \lambda \delta t G$ has positive entries. Hence we have that $f_j^{(n+\frac{1}{2})^*} \geq 0$ for each $j = 1, ..., N$.

3. Given $f_j^{(n+\frac{1}{2})^*} \geq 0$ for each $j = 1, ..., N$, then $f_j^{n+1} \geq 0$ for each $j = 1, ..., N$, by the same reasoning of the first step.

□

Recall that w in (5.12) is defined as follows

$$w(x) := h\frac{B(x)}{C(x)},$$

we have the following estimate for $\max_j |\mathcal{A}_{jj}|$

$$\begin{aligned}\max_j |\mathcal{A}_{jj}| &\leq \frac{\max_x\{C(x)\}}{h^2}[\max\{1, 1 - w(x)\} + \max\{1, 1 + w(x)\}] \leq \\ &\leq \frac{\max_x\{C(x)\}}{h^2}(2 + \max_x\{|w(x)|\}).\end{aligned}$$

Hence, the bound to δt for guaranteeing positivity in the previous theorem becomes

$$\delta t \leq \min\left\{\frac{1}{\lambda}, \frac{2h^2}{\max_x\{C(x)\}(2 + h\max_x\{|B(x)|/C(x)\}}\right\}.$$

We observe that for vanishing space step size h, the time step size must vanishes with order 2. In small diffusion regime, i.e. $C(x) \simeq 0$, it scales linearly as $\delta t < 2h/\max_x\{B(x)\}$.

We conclude this analysis proving the conservation of the total probability of the SIMEX2 scheme.

Proposition 19. *Consider* (5.56). *The total probability is conserved, in the sense that*

$$\sum_{j=1}^{N} f_j^n = \sum_{j=1}^{N} f_j^{n+1}.$$

Proof. Let use separately consider each step of (5.56).

1. Let us consider the second equation. Summing over j, we obtain

$$\sum_{j=1}^{n} f_j^{n+\frac{1}{2}} = \sum_{j=1}^{n} f_j^n + \frac{\delta t}{4}\left[\sum_{j=1}^{n}(\mathcal{A}f^n)_j + \sum_{j=1}^{n}(\mathcal{A}\bar{f}^{n+\frac{1}{2}})_j\right].$$

   We use (5.55) and obtain

$$\sum_{j=1}^{N} f_j^n = \sum_{j=1}^{N} f_j^{n+\frac{1}{2}}.$$

2. Let us consider the second equation. Summing over j and reshaping, we have that

$$\sum_{j=1}^{N} f_j^{(n+\frac{1}{2})^*} = \sum_{j=1}^{N} f_j^{n+\frac{1}{2}} + \frac{\delta t}{2}\left[\sum_{j=1}^{N}(\mathcal{G}f^{n+\frac{1}{2}})_j + \sum_{j=1}^{N}(\mathcal{G}\bar{f}^{n+\frac{1}{2}^*})_j\right],$$

   With same arguments as in Proposition 16, we claim that $\sum_{j=1}^{N}(\mathcal{G}f^{n+\frac{1}{2}})_j = 0$ and that $\sum_{j=1}^{N}(\mathcal{G}\bar{f}^{n+\frac{1}{2}^*})_j = 0$ and hence

$$\sum_{j=1}^{N} f_j^{(n+\frac{1}{2})^*} = \sum_{j=1}^{N} f_j^{n+\frac{1}{2}}.$$

3. By the same reasoning for the first step, we use (5.55) and obtain

$$\sum_{j=1}^{N} f_j^{n+1} = \sum_{j=1}^{N} f_j^{(n+\frac{1}{2})^*}.$$

□

5.4 Numerical experiments

In this section, we present results of numerical experiments for the numerical solution of the FP problem. We first consider the solution of FP problem with source term as in (5.5) and validate the space and time convergence rates theoretically proven in Section 5.2. Next, we aim at computing the PDF of the OU process with jumps, introduced in Section 2.4. To this end, we first numerically apply the SIMEX2 scheme to (5.5) with no source term (i.e. $\psi = 0$) and compare the obtained result with the empirical PDF obtained with the method outlined in Section 2.4.

5.4.1 The accuracy of the SIMEX schemes

In order to test the performance of the SIMEX schemes, we set the solution to (5.5) as the following moving Gaussian

$$f(x,t) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} \exp\left\{-\frac{(x-\mu t)^2}{2\tilde{\sigma}^2}\right\},$$

with $\mu = 10$ and $\tilde{\sigma} = 3$. With this choice, the corresponding source term $\psi(x,t)$ in (5.5) can be analytically computed by exploiting the closed form of the product of two Gaussian densities.

We consider a sufficiently large domain $\Omega := (-15, 30)$ and $I := [0,1]$. We set in (5.1) $b(x) = -x$, $\sigma = \sqrt{2}$, i.e. $C(x) = 1$. The parameters of the compound Poisson process are chosen to be $\lambda = 5$, $g \sim \mathcal{N}(3, 0.2^2)$.

In Table 5.1, we report the norm of the SIMEX1 solution error as a function of the mesh size. We see that the scheme is first-order accurate in time and second-order accurate in space, as proved in Section 5.2.

N	M	$\|f - \hat{f}\|_{h,\delta t}$
100	100	$5.39 \cdot 10^{-3}$
200	400	$1.09 \cdot 10^{-3}$
400	800	$2.90 \cdot 10^{-4}$
800	1600	$7.37 \cdot 10^{-5}$

Table 5.1: $L^2_{h,\delta t}$-error of the scheme SIMEX1.

In Table 5.2, we present results for the same test case, obtained with the SIMEX2 method. We have second-order convergence in time and space, as theoretically proven in Section 5.2.

N	M	$\|f - \hat{f}\|_{h,\delta t}$
200	200	$1.10 \cdot 10^{-3}$
400	400	$2.93 \cdot 10^{-4}$
800	800	$7.45 \cdot 10^{-5}$
1600	1600	$1.87 \cdot 10^{-5}$

Table 5.2: $L^2_{h,\delta t}$-error of the scheme SIMEX2.

5.4.2 The range of an Ornstein-Uhlenbeck process with jumps

In this section, we consider the OU process with jumps given by a compound Poisson process [78], whose empirical PDF has already been computed in Section 2.4. In one dimension, the dynamic of this process evolves according to the following stochastic differential equation

$$\begin{cases} dX_t = -\gamma X_t dt + \sigma dW_t + dP_t, \\ X_{\{t=0\}} = X_0, \end{cases} \tag{5.57}$$

where $\gamma, \sigma \in \mathbb{R}^+$, W is Brownian motion and P has rate of jumps λ and jump PDF g. Define with f the probability density function of (5.57) and the convolution $(f * g)(x,t)$ as follows

$$(f * g)(x,t) := \int_{\mathbb{R}} f(y,t) g(x-y) dy,$$

for each $t \in I$. The FP problem related to the Ornstein-Uhlenbeck process (5.57) on an unbounded domain takes the form

$$\begin{cases} \partial_t f(x,t) = \gamma \partial_x (x f(x,t)) + \frac{\sigma^2}{2} \partial_x^2 f(x,t) + \lambda (f * g)(x,t) - \lambda f(x,t) & (x,t) \in \mathbb{R} \times I \\ f(x,0) = f_0(x) & x \in \mathbb{R}. \end{cases} \tag{5.58}$$

Our aim is to apply our SIMEX2 scheme (5.56) to numerically solve (5.58) on a bounded domain (r,s) with zero-flux boundary conditions as in (5.5) with $\psi = 0$. We present a methodology to estimate the size of the computational domain (r,s), such that the dynamics of (5.57) is negligible in $\mathbb{R} \backslash (r,s)$ in some sense that will be investigated below. Using the Fourier transforms of $f(x,t)$ and $g(x)$, given by

$h(\omega,t)=\int_{-\infty}^{\infty}e^{i\omega x}f(x,t)dx$ and $W(\omega)=\int_{-\infty}^{\infty}e^{i\omega x}g(x)dx$, respectively, we have that $h(\omega,t)$ satisfies the following initial-value problem

$$\begin{cases}\partial_t h(\omega,t)=-\gamma\omega\partial_\omega h(\omega,t)-P(\omega)h(\omega,t), & (\omega,t)\in\mathbb{R}\times I\\ h(\omega,0)=h_0(\omega), & \omega\in\mathbb{R},\end{cases}\tag{5.59}$$

where $P(\omega):=\frac{\sigma^2}{2}\omega^2+\lambda(1-W(\omega))$ and $h_0(\omega)$ is the Fourier transform of the initial data $f_0(x)$.

After performing a change of variable for $\omega>0$, and noting that the following calculations can be carried out also for the case $\omega<0$ by defining $\bar{h}(\omega,t):=h(-\omega,t)$, we have that

$$h(\omega,t)=h_0(e^{-\gamma t}\omega)\exp\left\{-\frac{1}{\gamma}\int_{\omega e^{-\gamma t}}^{\omega}\frac{P(\omega')}{\omega'}d\omega'\right\}$$

solves (5.59).

Defining with $M^k[g]=\int_{-\infty}^{\infty}x^k g(x)dx$ the k-th moment of g, the Taylor expansion of P on $\omega=0$ reads as follows

$$P(\omega)=\frac{\sigma^2}{2}\omega^2-\lambda\sum_{k=1}^{\infty}\frac{i^k M^k g}{k!}\omega^k,$$

since

$$W(\omega)=\sum_{k=0}^{\infty}\frac{W^{(k)}(0)}{k!}\omega^k=\sum_{k=0}^{\infty}\frac{i^k M^k[g]}{k!}\omega^k.$$

Hence, the solution of (5.59) can be written as follows

$$h(\omega,t)=h_0(e^{-\gamma t}\omega)\exp\{z(\omega,t)\},$$

where

$$z(\omega,t):=i\frac{\lambda}{\gamma}(M^1[g])\omega(1-e^{-\gamma t})-\frac{\sigma^2+\lambda M^2[g]}{4\gamma}\omega^2(1-e^{-2\gamma t})+\frac{\lambda}{\gamma}\sum_{k=3}^{\infty}\frac{i^k M^k[g]}{k(k!)}\omega^k(1-e^{-k\gamma t}).$$

This form for $h(\omega,t)$ allows us to calculate $M_t^1[f]$ and $M_t^2[f]$, that is, the first- and second-time dependent moments of $f(x,t)$.

We are interested in these two quantities at a time near to equilibrium, i.e. $T>>1/\gamma$. Therefore we suppose that the support of the PDF is vanishing outside the interval of size $[M_T^1[f]-5\sqrt{M_T^2[f]},\ M_T^1[f]+5\sqrt{M_T^2[f]}]$. We exploit the fact that $h(0,t)=1$ for each t to state the following

$$\lim_{t\to\infty}M_t^1 f=\lim_{t\to\infty}-ih'(0,t)=\frac{\lambda}{\gamma}(M^1[g])$$

and that

$$\lim_{t\to\infty}M_t^2 f=-\lim_{t\to\infty}h''(0,t)=\frac{\lambda^2}{\gamma^2}\left(M^1[g]\right)^2+\frac{\sigma^2+\lambda M^2[g]}{2\gamma}.$$

Notice that the moments of the density g of the jump amplitude play a key role in the width of the relevant range of the dynamics of the considered process.

Therefore, the choices $\gamma = 1$ and $\sigma = 1$ in (5.57) allow to set the domain $\Omega = (-20, 50)$. We consider the time interval $I := [0, 1]$ and the initial random variable in (5.57) $X_0 \sim \mathcal{N}(15, 3)$. we choose $\lambda = 5$ and $g \sim \mathcal{N}(3, 0.2^2)$.

Figure 5.1 depicts the PDF of the process X at time $T = 1$. We compare the results obtained with the SIMEX2 scheme with $N = 400$ and $M = 400$ and the empirical PDF with the algorithm illustrated in Section 2.4. To this end, we solved the initial-value problem (5.57) by applying the EM method in the time interval $[0, 1]$ with M intervals, as outlined in Section 2.4.

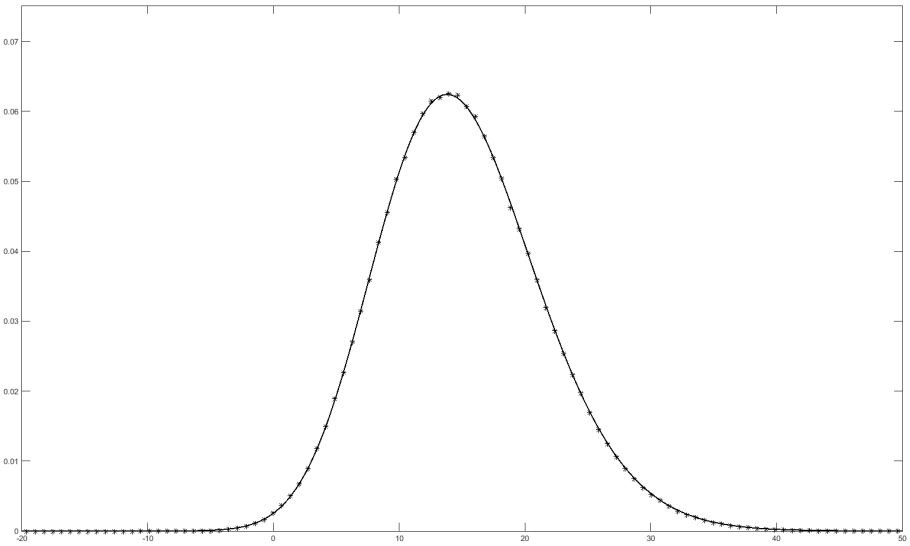


Figure 5.1: The PDF of an Ornstein-Uhlenbeck process with jumps.

Comparison between the solution of the SIMEX2 scheme with $N = 400$ and $M = 400$ (solid line) and the empirical PDF with 150 histogram midpoints with 10^6 sample paths (stars).

5.5 Summary and remarks

In this chapter, we outlined our numerical methods SIMEX1 and SIMEX2 that compute the solution of a FP problem related to a JD process. Our discretization is achieved by applying the method of lines, which consists of performing a spacial discretization of the differential and integral operators and subsequently integrating with respect to time the resulting system of ordinary differential equations. The differential part of the FP equation is discretized through the finite volume scheme given by the CC method, while the integral is approximated with the mid-point rule. The time integration step combines an operator splitting method with the Euler scheme in case of SIMEX1 and with a predictor-corrector scheme in case of SIMEX2. The discretization procedure was outlined in Section 5.1. We carried out the complete convergence and stability analysis in Section 5.2 and showed in Section 5.3 that our methods preserve the two properties of a PDF, namely its positivity and the conservation of the total probability. A section of numerical experiments that validated the theoretical estimates concluded the chapter.

Chapter 6

Numerical solution of Fokker-Planck control problems

In this chapter, we address the numerical solution of the FP control problems defined and investigated in Chapter 4. Some references on optimization problems governed by PDEs are, e.g., [12, 14, 71]. As outlined in Chapter 4, the L^1-term in our cost functional makes the problem nonsmooth and therefore appropriate optimization algorithm are needed. We implement a proximal iterative scheme, introduced in [70] and [81] and first developed in the framework of finite-dimensional optimization [9, 22, 72]. Recent works have adapted the structure of these algorithms for solving infinite-dimensional PDE optimization problems [85, 86, 89]. In this chapter, we apply the results of Chapter 5 and of the recent works [44, 86]. We proceed as follows. Section 6.1 investigates the proposed optimization algorithm. We first introduce the definition of proximity operator of a convex semi-continuous function, which is needed for definition of our method. The chosen algorithm combines a fixed-point iteration with a gradient method.

Section 6.2 is devoted to the convergence of the proximal scheme. In order to investigate the effectiveness of the algorithm, results of numerical experiments are presented in Section 6.3.

6.1 The proximal method

In this section, we discuss a proximal optimization scheme for solving (4.8). This scheme and the related theoretical discussion follow the work in [22, 72, 86]. Proximal methods conveniently exploit the additive structure of the reduced objective; in our framework, we have that the reduced functional $\hat{\mathcal{J}}$ is given by the sum of a

nonconvex smooth function $\mathcal{J}_1$ and a convex nonsmooth function $\mathcal{J}_2$ as in (4.13). For our discussion, we need the following definitions and properties.

Definition 11. Let Z be a Hilbert space and l a convex lower semi continuous function, $l : Z \to \mathbb{R}$. The proximity operator $prox_l : Z \to Z$ of l is defined as follows

$$prox_l(z) := \arg\min_{w \in Z} \left\{ l(w) + \frac{1}{2}\|z - w\|_Z^2 \right\}, \quad z \in Z.$$

The following proposition, whose proof can be found in [22], is of fundamental importance in the remainder of the section.

Proposition 20. *Let Z be a Hilbert space and l a convex lower semi continuous function, $l : Z \to \mathbb{R}$, with proximity operator $prox_l$. The following relation holds*

$$p = prox_l(z) \Leftrightarrow z - p \in \partial l(p), \tag{6.1}$$

where ∂l is the subdifferential of l.

Proposition 21. *For each lower semicontinuous function l defined on a Hilbert space Z, the following holds*

$$l(z) \geq l(w) + \langle s, z - w \rangle, \tag{6.2}$$

for each $z, w \in Z$ and each $s \in \partial l(w)$.

Proposition 22. *The solution $\bar{u}$ of (4.8), i.e. $\min_{u \in \mathcal{U}_{ad}} \hat{\mathcal{J}}(u)$, satisfies*

$$\bar{u} = prox_{\alpha \mathcal{J}_2}(\bar{u} - \alpha \nabla \mathcal{J}_1(\bar{u})). \tag{6.3}$$

for each $\alpha > 0$.

Proof. From Proposition 5 in Chapter 4 and by using (6.1), we have

$$\begin{aligned} \bar{u} \text{ solves (4.8)} &\Rightarrow -\nabla \mathcal{J}_1(\bar{u}) \in \partial \mathcal{J}_2(\bar{u}) \\ &\Leftrightarrow (\bar{u} - \alpha \nabla \mathcal{J}_1(\bar{u})) - \bar{u} \in \alpha \partial \mathcal{J}_2(\bar{u}) \\ &\Leftrightarrow \bar{u} = prox_{\alpha \mathcal{J}_2}(\bar{u} - \alpha \nabla \mathcal{J}_1(\bar{u})). \end{aligned}$$

□

The relation (6.3) suggests that a solution procedure based on a fixed point iteration should be pursued. We discuss how such algorithm can be implemented.

In the following, we assume that $\mathcal{J}_1(u)$ in (4.13) has a locally Lipschitz-continuous gradient $\nabla \mathcal{J}_1$ as follows

$$\|\nabla \mathcal{J}_1(u) - \nabla \mathcal{J}_1(v)\| \leq L\|u - v\|, \tag{6.4}$$

for each $v \in \mathcal{V}$, $\mathcal{V} \subset \mathcal{U}_{ad}$ neighborhood of u, with L a Lipschitz continuity constant.

The condition (6.4) implies the following inequality

$$\mathcal{J}_1(u) \leq \mathcal{J}_1(v) + \langle \nabla \mathcal{J}_1(v), u - v \rangle + \frac{L}{2}\|u - v\|^2, \tag{6.5}$$

for each $v \in \mathcal{V}$. This can easily be seen by expanding $\mathcal{J}_1$ as follows

$$\mathcal{J}_1(u) = \mathcal{J}_1(v) + \langle \nabla \mathcal{J}_1(v), u - v \rangle + \int_0^1 \langle \nabla \mathcal{J}_1(v + t(u - v)) - \nabla \mathcal{J}_1(v), u - v \rangle dt,$$

and the by applying the Cauchy-Schwarz inequality to the scalar product in the last addend. Therefore, we have

$$\begin{aligned} &\min_{u \in \mathcal{U}_{ad}} \{\mathcal{J}_1(u) + \mathcal{J}_2(u)\} \leq \\ &\min_{u \in \mathcal{U}_{ad}} \left\{ \mathcal{J}_1(v) + \mathcal{J}_2(u) + \langle \nabla \mathcal{J}_1(v), u - v \rangle + \frac{L}{2}\|u - v\|^2 \right\}. \end{aligned} \tag{6.6}$$

Inequality (6.6) is the starting point for the formulation of a proximal scheme, whose strategy consists of minimizing the right-hand side in (6.6). The following equality holds

$$\begin{aligned} &\arg\min_{u \in \mathcal{U}_{ad}} \left\{ \mathcal{J}_1(v) + \mathcal{J}_2(u) + \langle \nabla \mathcal{J}_1(v), u - v \rangle + \frac{L}{2}\|u - v\|^2 \right\} = \\ &\arg\min_{u \in \mathcal{U}_{ad}} \left\{ \mathcal{J}_2(u) + \frac{L}{2} \left\| u - \left(v - \frac{1}{L} \nabla \mathcal{J}_1(v) \right) \right\|^2 \right\} \end{aligned} \tag{6.7}$$

and it can be easily proven by exploiting the definition of scalar product. Recall the definition of $\mathcal{J}_2$ in (4.13). The following lemma, whose proof can be found in [85], gives an explicit expression for the right-hand side in (6.7).

Lemma 1. *Let $\mathcal{U}_{ad}$ be as in* (4.2)*. Then*

$$\arg\min_{u \in \mathcal{U}_{ad}} \left\{ \tau \|u\|_1 + \frac{1}{2}\|u - w\|^2 \right\} = \mathbb{S}_\tau^{\mathcal{U}_{ad}}(w) \text{ for each } w \in \mathcal{U}, \text{ for each } \tau > 0,$$

where the projected soft thresholding function $\mathbb{S}_\tau^{\mathcal{U}_{ad}}$ is defined as follows

$$\mathbb{S}_\tau^{\mathcal{U}_{ad}}(w) := \begin{cases} \min\{w - \tau, u_b\} & \text{on } \{t \in I : w(t) > \tau\} \\ 0 & \text{on } \{t \in I : |w(t)| \leq \tau\} \\ \max\{w + \tau, u_a\} & \text{on } \{t \in I : w(t) < -\tau\}. \end{cases}$$

Based on this lemma, we conclude the following

$$\arg\min_{u\in\mathcal{U}_{ad}}\left\{\mathcal{J}_2(u)+\frac{L}{2}\left\|u-\left(v-\frac{1}{L}\nabla\mathcal{J}_1(v)\right)\right\|^2\right\}=\mathbb{S}^{\mathcal{U}_{ad}}_{\frac{\gamma}{L}}\left(v-\frac{1}{L}\nabla\mathcal{J}_1(v)\right),$$

which can be taken as starting point for a fixed-point algorithm as follows

$$u_{k+1}=\mathbb{S}^{\mathcal{U}_{ad}}_{\gamma\alpha_k}\left(u_k-\alpha_k\nabla\mathcal{J}_1(u_k)\right), \tag{6.8}$$

where α_k depends on L_k, the local Lipschitz continuity constant defined in (6.4), as we discuss below. Such method has been investigated in [22, 85, 81]. In this work, we apply an extension of (6.8), which takes for each iteration k the following form

$$u_{k+1}=\mathbb{S}^{\mathcal{U}_{ad}}_{\gamma\alpha_k}\left(u_k-\alpha_k\nabla\mathcal{J}_1(u_k)+\theta_k(u_k-u_{k-1})\right), \tag{6.9}$$

with $\theta_k\in(0,1)$. This method has been proposed in [72] and investigated in [86]. Our inertial proximal method is summarized in the Algorithm 1.

The nonincreasing condition on $\{\delta_k\}_{k\in\mathbb{N}}$ is equivalently reformulated as follows

$$\alpha_k\geq\frac{1-\frac{\theta_k}{2}}{\delta_{k-1}+\frac{L_k}{2}}.$$

The backtracking scheme in Algorithm 1 provides an estimation of the upper bound of the Lipschitz constant in (6.4), since it is not known a priori. The initial guess for L is chosen as follows. Given a small variation ε of u, we have

$$L=\max\left\{\frac{\|\nabla\mathcal{J}_1(u)-\nabla\mathcal{J}_1(u+\varepsilon)\|_2}{\|\varepsilon\|_2},\frac{\|\nabla\mathcal{J}_1(u)-\nabla\mathcal{J}_1(u-\varepsilon)\|_2}{\|\varepsilon\|_2}\right\}.$$

The computation of the gradient requires the discretization of the optimality systems given in (4.23) and (4.24). The space and time grids are defined in (5.6).

The discretization of the forward equation solved by f has been carried out in Chapter 5. Next, we outline how the discretization of the system satisfied by the adjoint variable p in (4.23) and (4.24) is carried out. If we follow the optimize-before-discretize (OBD) approach, the optimality system has already been computed on a continuous level in the systems in (4.23) and (4.24). As a consequence, the OBD approach allows one to discretize the forward and adjoint FP problems according to different numerical schemes. However, the OBD procedure might introduce an inconsistency between the discretized objective and the reduced gradient. For this reason, the DBO (discretize-before-optimize) approach could be preferred and we pursue it in this work. In our framework, for each grid point x_j, $j=1,...,N$, the DBO approach results in the following approximations

$$\begin{aligned}\partial_x p(x_j,t)&\approx(1-\delta_{j-1/2})\frac{p_{i-1}(t)-p_i(t)}{h}+\delta_{j+1/2}\frac{p_i(t)-p_{i+1}(t)}{h},\\ \partial^2_{xx}p(x_j,t)&\approx\frac{p_{j+1}(t)-2p_j(t)+p_{j-1}(t)}{h^2},\end{aligned}$$

Algorithm 1 (Inertial proximal method).

Input: initial guess u_0, $k = 0$, $k_{\max}$, $\theta_k \in (0,1)$, tolerance tol, $c_1, c_2 > 0$, $\eta > 1$.

1. While $k \leq k_{\max}$, do:

   (a) Evaluate $\nabla \mathcal{J}_1(u_k)$ according to Algorithm 2.

   (b) Update $L_k = \eta L_k$ until

$$\mathcal{J}_1(\tilde{u}) \leq \mathcal{J}_1(u_k) + \langle \nabla_u \mathcal{J}_1(u_k), \tilde{u} - u_k \rangle + \frac{L_k}{2} \|\tilde{u} - u_k\|^2$$

   where

$$\begin{aligned} &\tilde{u} = \mathbb{S}_{\gamma\alpha_k}^{\mathcal{U}_{ad}} \left(u_k - \alpha_k \nabla \mathcal{J}_1(u_k) + \theta_k (u_k - u_{k-1})\right) \\ &\theta_k \geq 0 \text{ for each } k \\ &\delta_k := \frac{1}{\alpha_k} - \frac{L_k}{2} - \frac{\theta_k}{2\alpha_k} \text{ is monotonically decreasing.} \\ &\delta_k \geq c_2 \text{ for each } k \end{aligned} \tag{6.10}$$

   (c) Set $u_{k+1} = \tilde{u}$.

   (d) Compute E according to (4.23) or (4.24).

   (e) If $E < tol$, break.

   (f) $k = k + 1$.

---

together with the midpoint formula applied to $\tilde{\mathcal{I}}$. We have the following semi-discretized system

$$-p'(t) = (\mathcal{A}^T + \mathcal{G}^T)p(t), \quad p(t) \in \mathbb{R}^N. \tag{6.11}$$

The time integration of (6.11) is carried out with the combination of the SM splitting with a predictor corrector scheme, as in Chapter 5. We remark that the advantages of the DBO approach are manifold. With this choice, the boundary conditions of the backward equation in (4.23) and (4.24) are automatically implemented.

Algorithm 2 (Evaluation of the gradient).

Input: u_k, initial value f_0, terminal value p_T at time T.

1. Compute f_k, given f_0 and u_k.

2. Compute p_k.

3. Evaluate $\nabla \mathcal{J}_1(u_k)$ according to (4.22).

6.2 Convergence analysis

In this section, we discuss the convergence of the proximal algorithm and report here some results presented in [44, 72, 86].

Definition 12. The proximal residual r is defined as follows

$$r(u) := u - \mathbb{S}_{\gamma}^{\mathcal{U}_{ad}}\left(u - \nabla \mathcal{J}_1(u)\right). \tag{6.12}$$

Proposition 22 tells us that $r(u) = 0$ in $L^2(I)$ whenever u solves (4.8). In what follows, we establish a connection between the condition $r(u) = 0$ and the sequence $\{u_k\}_{k\in\mathbb{N}}$ generated by Algorithm 1. Define the following sequence

$$\Delta_k := ||u_k - u_{k-1}||_2 \tag{6.13}$$

for $k \in \mathbb{N}$.

Proposition 23. *The sequence $\{\Delta_k\}_{k\in\mathbb{N}}$ satisfies the following*

$$\lim_{k\to\infty} \Delta_k = 0.$$

Proof. For $\delta \in \mathbb{R}^+$, we define the following quantity

$$H_\delta(u, v) := \hat{\mathcal{J}}(u) + \delta||u - v||_2^2 \tag{6.14}$$

for each $u, v \in \mathcal{U}_{ad}$. The following inequality holds

$$H_{\delta_{k+1}}(u_{k+1}, u_k) \leq H_{\delta_k}(u_k, u_{k-1}) - \zeta_k \Delta_k^2, \tag{6.15}$$

where

$$\zeta_k := \frac{1}{\alpha_k} - \frac{L_k}{2} - \frac{\theta_k}{\alpha_k}.$$

Note that α_k and θ_k in Algorithm 1 can be chosen such that $\zeta_k \geq c_2$ for each k. In fact, from the definition of u_k in algorithm 1, it holds that

$$\frac{u_k - u_{k+1}}{\alpha_k} - \nabla \mathcal{J}_1(u_k) + \frac{\theta_k}{\alpha_k}(u_k - u_{k-1}) \in \partial(\beta \mathcal{J}_2)(u_{k+1}).$$

From (6.2) and (6.5) it follows that

$$\begin{aligned}\hat{\mathcal{J}}(u_{k+1}) &\leq \hat{\mathcal{J}}(u_k) - \left(\frac{1}{\alpha_k} - \frac{L_k}{2}\right)\Delta_{k+1}^2 + \frac{\theta_k}{\alpha_k}\langle u_{k+1} - u_k, u_k - u_{k-1}\rangle \\ &\leq \hat{\mathcal{J}}(u_k) - \left(\frac{1}{\alpha_k} - \frac{L_k}{2} - \frac{\theta_k}{2\alpha_k}\right)\Delta_{k+1}^2 + \frac{\theta_k}{2\alpha_k}\Delta_k^2.\end{aligned}$$

Therefore, we have

$$\hat{\mathcal{J}}(u_{k+1}) + \delta_k\Delta_{k+1}^2 \leq \hat{\mathcal{J}}(u_k) + \delta_k\Delta_k^2 - \zeta_k\Delta_k^2,$$

which implies (6.15) since δ_k is monotonically decreasing. Moreover, the condition $\zeta_k \geq c_2 \geq 0$ implies that $\{H_{\delta_k}(u_k, u_{k-1})\}_{k\in\mathbb{N}}$ is monotonically decreasing, and thus converging, since $\hat{\mathcal{J}}(u) \geq 0$ for each $u \in \mathcal{U}$.

Summing (6.15) from $k = 0, ..., \bar{k}$ yields

$$\begin{aligned}\sum_{k=0}^{\bar{k}} \zeta_k\Delta_k^2 &\leq \sum_{k=0}^{\bar{k}}(H_{\delta_k}(u_k, u_{k-1}) - H_{\delta_{k+1}}(u_{k+1}, u_k)) = \hat{\mathcal{J}}(u_0) - H_{\delta_{\bar{k}+1}}(u_{\bar{k}+1}, u_{\bar{k}}) \\ &\leq \hat{\mathcal{J}}(u_0) < \infty.\end{aligned}$$

The claim follows by letting $\bar{k} \to \infty$ and remembering that $\zeta_k \geq c_2 > 0$ for each k. □

Next, we can state the desired convergence results. The complete discussion can be found in [86].

Proposition 24. *The following holds.*

- *The sequence $\{\hat{\mathcal{J}}(u_k)\}_{k\in\mathbb{N}}$ converges.*
- *There exists a weakly convergent subsequence $\{u_{k_j}\}_{j\in\mathbb{N}}$.*
- *If, in addition, $\mathcal{J}_1$ is strictly convex in $\mathcal{U}_{ad}$, then any weak limit u^* of $\{u_{k_j}\}_{j\in\mathbb{N}}$ is a critical point of (4.8) and $\hat{\mathcal{J}}(u^*) \leq \liminf_j \hat{\mathcal{J}}(u_{k_j})$.*

Proposition 25. *Let $\{u_k\}_{k\in\mathbb{N}}$ be the sequence generated by Algorithm 1, then the following holds*

$$\min_{0\leq k\leq K} \|r(u_k)\|^2 \leq (c_1c_2)^{-1}\frac{2\hat{\mathcal{J}}(u_0)}{K+2},$$

where r is defined in (6.12).

6.3 Numerical experiments

In this section, we present results of numerical experiments to validate the performance of our optimal control framework. The purpose is to determine a sparse control $u = u(t)$ such that the expected value of a jump-diffusion process minimizes the quantity defined by (4.5) and (4.6).

We implement the discretization scheme and the algorithm described in Section 6.1. We take $\Omega = (-20, 20)$ and $[0, T] = [0, 1]$, and assume that the initial PDF f_0 is given, $f_0 \sim \mathcal{N}(0, 10)$. The compound Poisson process corresponds to the choice $\lambda = 5$ and $g \sim \mathcal{N}(0, 1)$. We take $b(x, u) := u(t) - x/10$ and $C = 0.5$. In case of (4.5), we consider $\xi = [-0.5, 2, 0, 1, -1, 0, -0.5, 2]$. In the case of (4.6), we take $\xi(t) = 2\sin(10t)$. We choose $N = 200$ and $M = 200$. We choose $c_1 = c_2 = 10^{-3}$ in Algorithm 1.

In the first series of experiments, we consider the setting with $\nu = 10^{-8}$ and $\gamma = 0$ in (4.13). Further, we consider constraints on the control. Corresponding to this choice and to the discrete-in-time tracking functional (4.5), we report in the Figures 6.1 and 6.2 the solution for the state and the adjoint variables, respectively. On the other hand, using the continuous-in-time tracking functional (4.6), we obtain the state and the adjoint variables depicted in the Figures 6.4 and 6.5, respectively.

Also for the case $\gamma = 0$ and both tracking functionals, we report in the Tables 6.1 and 6.2 the values of the tracking error for different values of the weight ν. As expected, the tracking improves as the value of this optimization parameter becomes smaller. In the Figures 6.1 and 6.4, we can see that the optimal control u drives the expected mean value of the PDF towards the mean values given by ξ_k and $\xi(t)$, respectively; moreover, Figures 6.3 and 6.6 depict the given values of ξ and the expected value of the PDF with respect to time.

Next, we investigate the behavior of the optimal solution considering the full optimization setting, that is, the case when the L^1-cost actively enters in the optimization process, i.e. $\gamma > 0$, and the control is constrained by the bounds $u_a = -10, u_b = 10$ defining the set of admissible controls. For simplicity, we discuss only the case with $\nu = 10^{-8}$.

In the Figures 6.7, 6.8, 6.9, we depict the optimal controls for three different choices of values of γ and considering the discrete-in-time tracking functional given by (4.5). In the Figures 6.10, 6.11 and 6.12, we show the optimal controls for three different choices of values of γ and considering the continuous-in-time tracking functional given by (4.6). In both cases, we can clearly see that when the value of the parameter γ is increased, the sparsity of the solution is significantly enhanced, as expected.

Figures 6.13 and 6.14 depicts 10 sample paths of the JD process with respect to time. We choose $\gamma = 0.1$. The MC simulations are computed with the optimal control $\bar{u}$ in case of (4.5) and (4.6), together with the given $\{\xi_k\}_{k=1}^{\bar{k}}$ and $\xi(t)$, respectively.

In the Tables 6.1 and 6.2, we also report values of the tracking error when both the L^2- and L^1-costs are considered. For a direct comparison with the first series of experiments, we consider an unconstrained control. We find that already with a small value of γ, the tracking ability of the optimization scheme worsen for both choices the tracking functional.

6.4 Summary and remarks

In this chapter we addressed the numerical resolution of the optimization problems defined and investigated on a continuous level in Chapter 4. We focused on the optimize-before-discretized approach, showing its advantages in the considered problem. In Section 6.1 we introduced some theoretical results needed for the definition of our proximal optimization scheme. A convergence result was included. In the last section 6.3 we tested the proposed method. In the first series of experiments, we considered a smooth functional (i.e. no L^1-term was added). As expected, the tracking error improved as the value of the L^2-weight parameter became smaller. In the second series of experiments, we considered the case when the L^1-cost actively enters in the optimization process. By numerical inspection we clearly saw that the sparsity of the solution was enhanced.

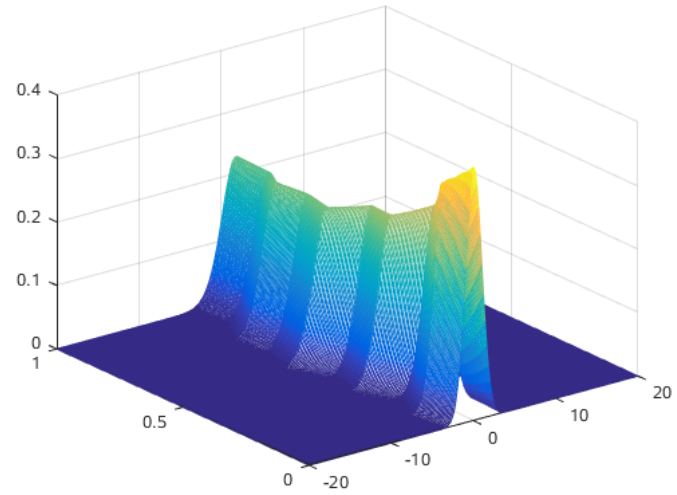


Figure 6.1: State variable in the case of the discrete-in-time tracking functional defined in (4.5), with $\gamma = 0$.

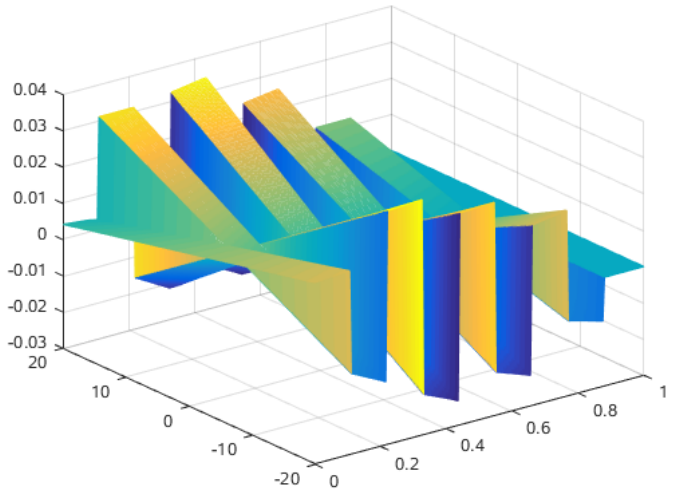


Figure 6.2: Adjoint variable in case of the discrete-in-time tracking functional defined in (4.5), with $\gamma = 0$.

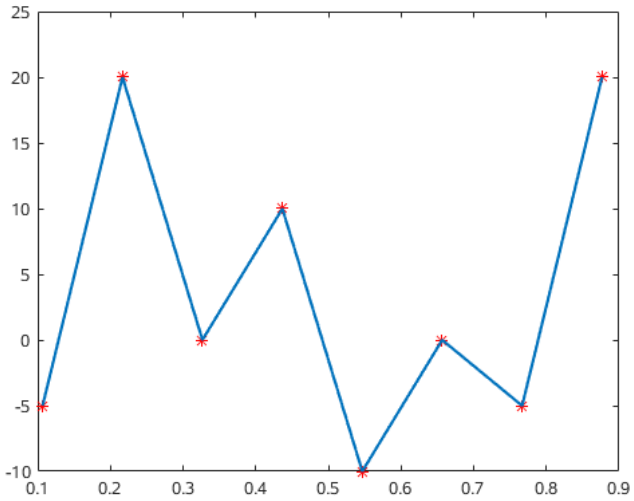


Figure 6.3: Expected value of the state equation (solid line) and given values for ξ (stars) with respect to time $t \in [0, T]$ of the discrete-in-time tracking functional defined in (4.5), with $\gamma = 0$.

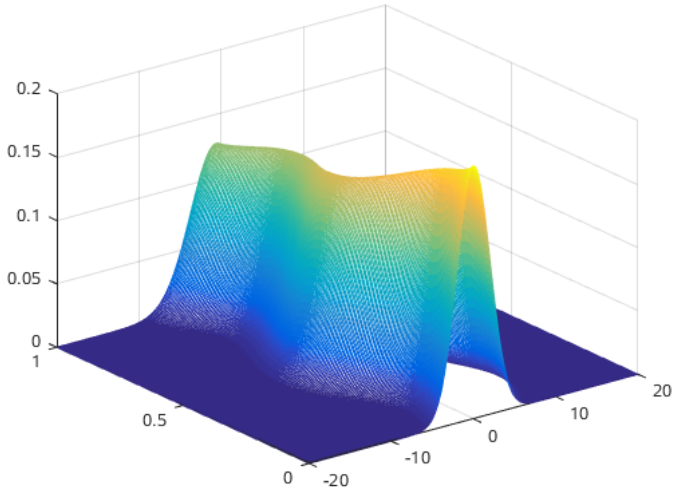


Figure 6.4: State variable in the case of the continuous-in-time tracking functional defined in (4.6), with $\gamma = 0$.

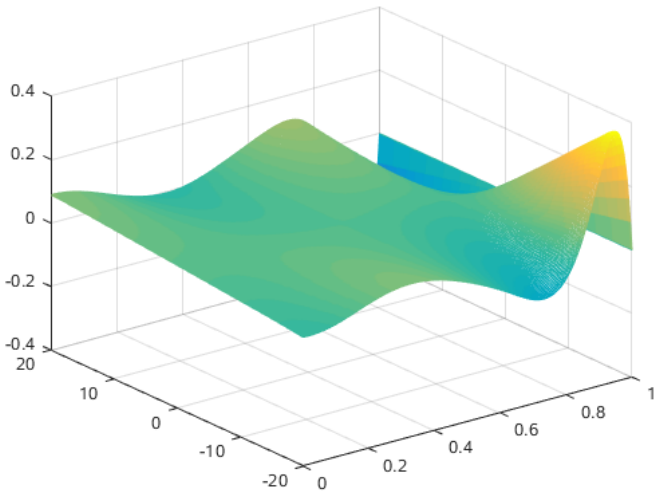


Figure 6.5: Adjoint variable in case of the continuous-in-time tracking functional defined in (4.6), with $\gamma = 0$.

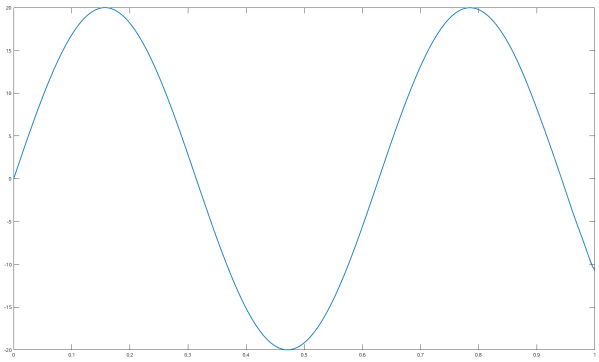


Figure 6.6: Expected value of the state equation (solid line) and given values for ξ (stars) with respect to time $t \in [0, T]$ of the continuous-in-time tracking functional defined in (4.6), with $\gamma = 0$.

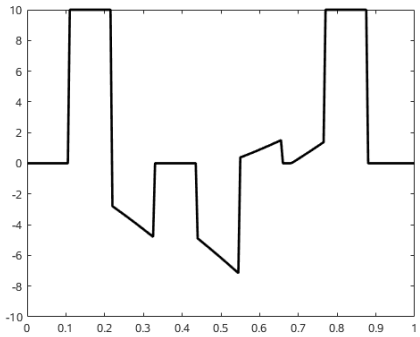


Figure 6.7: Optimal control with $\gamma = 0.1$ and tracking objective given by (4.5).

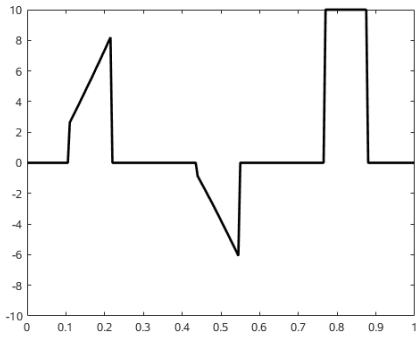


Figure 6.8: Optimal control with $\gamma = 0.2$ and tracking objective given by (4.5).

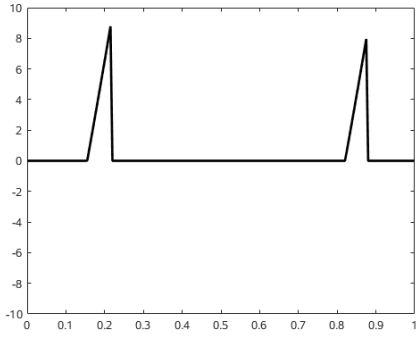


Figure 6.9: Optimal control with $\gamma = 0.4$ and tracking objective given by (4.5).

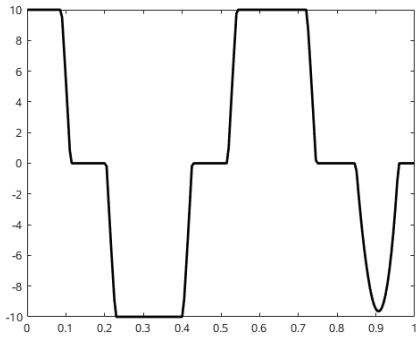


Figure 6.10: Optimal control with $\gamma = 0.1$ and tracking objective given by (4.6).

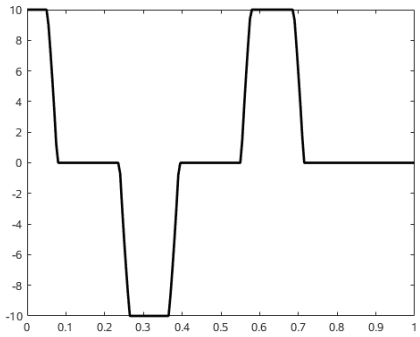


Figure 6.11: Optimal control with $\gamma = 0.2$ and tracking objective given by (4.6).

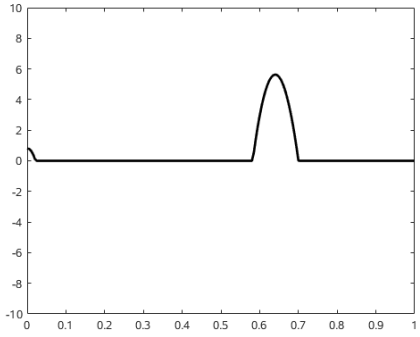


Figure 6.12: Optimal control with $\gamma = 0.4$ and tracking objective given by (4.6).

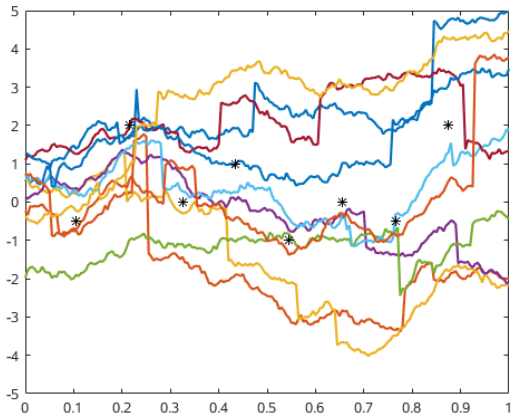


Figure 6.13: Monte Carlo paths (line) of the JD process (3.1) with the optimal control $\bar{u}$ in case of cost functional given by (4.5) and given $\xi_1, ...\xi_K$ (black points).

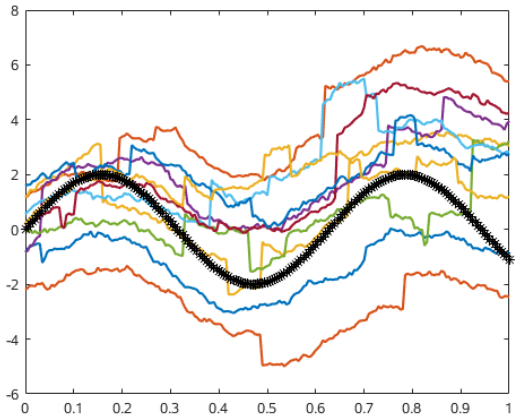


Figure 6.14: Monte Carlo paths (line) of the JD process (3.1) with the optimal control $\bar{u}$ in case of cost functional given by (4.6) and given $\xi(t)$ (black stars).

Table 6.1: Tracking error of the discrete-in-time functional $D(f)$ given by (4.5).

$D(f)$	ν	γ
$8.13 \cdot 10^{-7}$	10^{-10}	0
$1.17 \cdot 10^{-4}$	10^{-6}	0
$1.50 \cdot 10^{-4}$	10^{-4}	0
$2.47 \cdot 10^{-6}$	10^{-10}	10^{-4}
$8.51 \cdot 10^{-5}$	10^{-6}	10^{-4}
$1.68 \cdot 10^{-4}$	10^{-4}	10^{-4}

Table 6.2: Tracking error of the continuous-in-time functional $D(f)$ given by (4.6).

$D(f)$	ν	γ
$7.00 \cdot 10^{-4}$	10^{-10}	0
$7.00 \cdot 10^{-4}$	10^{-6}	0
$8.52 \cdot 10^{-4}$	10^{-4}	0
$1.30 \cdot 10^{-3}$	10^{-10}	10^{-4}
$1.31 \cdot 10^{-3}$	10^{-6}	10^{-4}
$1.48 \cdot 10^{-3}$	10^{-4}	10^{-4}

Chapter 7

Conclusion

In this thesis, infinite-dimensional optimal control problems with constraints given by PIDE FP problems related to JD processes were investigated.

At the beginning of this thesis, we introduced the class of jump processes given by Lévy processes and we investigated stochastic initial value problems governed by such processes. Existing theorems on existence and uniqueness of solutions, as well as numerical methods for SDEs, were presented. We focused on the EM method, which allowed us to compute a MC empirical PDF of a JD process. However, since the rate of convergence of MC methods is quite slow, we computed the PDF of a JD process by analytically deriving the FP equation. This equation takes the form of a PIDE and it governs the time evolution of the PDF of a stochastic process. In this thesis, we derived the FP equation in two cases, where a process has range in the whole space $\mathbb{R}^d$ and where its dynamics is limited to a bounded domain with reflecting barriers, respectively.

The FP equation, endowed with initial and boundary conditions, gives rise to FP problems. The solvability of these problems was discussed in suitable functional spaces. Moreover, we provided a priori estimates for their solutions.

Since closed-form solutions of FP problems are often not available, one has to resort to numerical methods. We addressed this issue by constructing two schemes, namely SIMEX1 and SIMEX2. Our schemes combined the CC method for spatial discretization with the SM splitting and first- and second- order time discretization methods. We chose the Euler method and a predictor-corrector scheme. A full numerical and stability analysis was carried out. Moreover, it was shown that the SIMEX schemes are also conservative, in the sense that the numerical solution to the FP problem satisfies the properties of the PDF of a JD process, namely the positivity and the mass conservation.

Our stable and convergent numerical SIMEX methods allowed us to numerically address optimal control problems with a differential constraint given by the FP problem. First, we defined and investigated the infinite-dimensional optimal control

problems of interest. We considered tracking objectives with the aim of steering the mean value of JD processes towards a known sequence of values. Two different cost functionals were investigated, continuous-in-time and discrete-in-time, respectively. In our formulation, the control variable entered the state equation in the drift coefficient of the FP PIDE, making the optimization problem nonconvex. An L^1-penalization term was added to the cost functionals, with the aim of enhancing the sparsity of the optimal solution; this term made the problem nonsmooth. We were able to prove the existence of at least an optimal solution. By exploiting the additive structure of the cost functionals and considering on the subgradient of the nonsmooth term, we derived the first order-optimality system in the Lagrangian framework. We numerically addressed the nonconvex and nonsmooth FP problem by applying a proximal method.

Appendix A

MATLAB code

This work is completed with a CD-ROM containing the MATLAB codes used throughout the present doctoral thesis.

Listing A.1 contains the code for numerically solving the JD process (2.34) with the EM method defined in (2.35) and computing the empirical PDF of (2.34). The SDE, the space and time range and the time grid are given. The output consists of a sample path of (2.34), whose values are specified in the time grid points. This allows to compute the empirical PDF of (2.34), as outlined in Section 2.4.

Listing A.1: Numerical solution of a SDE with the EM method and computation of the empirical PDF

```
close all, clear vars, clc, format long, clf

LB = -50; % lower bound of the space domain
UB = 50;  % upper bound of the space domain
T = 1; % final time
NT = 500; % # intervals in the time grid
lambda = 5; % jump rate
mu_jumps = 0; % jump mean
sigma_jumps = 3; % jump variance
dt =  T / NT; % mesh size
time_array = 0 : dt :T-dt;
n_MCruns = 1e5; % # Monte Carlo runs
```

```
13
14 %% Monte Carlo runs
15 waitbar_MC = waitbar(0,'Monte Carlo runs');
16 for k = 1:n_MCruns % for each run, W and P are simulated
17     timeArray = dt : dt : T;
18     Xiniz = 0; Xtemp = Xiniz; Xvalues = zeros(1,NT);
19     waitbar(k / n_MCruns)
20     % 1 simulating the Brownian and Poisson increments
21     dW = sqrt(dt) * randn(1, NT);
22     nJumps = poissrnd(lambda*(T)); % number of jumps in the interval
           [0, T]
23
24
25     nJumps = poissrnd(lambda*(T)) % number of jumps in the interval [0,
           T]
26     if nJumps>0
27         jumpTimes = sort ( unifrnd(0, T, 1, nJumps) ) % sorted array
               containing the jump times
28         jumpAmplitudes = normrnd(mu_jumps, sigma_jumps, 1, nJumps)
29         for cont_jumps = 1 : nJumps
30             B(cont_jumps,:) = jumpAmplitudes(cont_jumps) * stepfun(
                   timeArray, jumpTimes(cont_jumps));
31         end
32         B = cumsum(B,1);
33         PoisProc = B(nJumps,:);
34     else
35         PoisProc = zeros(1,NT);
36     end
37     % 2 Euler-Maruyama method
38     Pinc = PoisProc(1);
39     for j = 1:NT
40         if j>1
```

```
41                  Pinc = PoisProc(j) - PoisProc(j-1) ; break;
42              end
43              Xtemp = Xtemp + b_drift(Xtemp) * dt + sigma_diff(Xtemp) * dW(j)
                    + Pinc;
44              % Reflecting boundary condition
45              while (Xtemp < LB || Xtemp > UB)
46                  Xtemp = (2*UB - Xtemp) * (Xtemp > UB) + ...
47                          (2*LB - Xtemp) * (Xtemp < LB);
48              end
49              Xvalues(j) = Xtemp;
50          end
51
52          timeArray = [0, timeArray];
53          Xvalues = [Xiniz, Xvalues];
54
55          figure(1)
56          plot(timeArray, Xvalues,'LineWidth',1.6)
57          %axis([0 T -10 10])
58          hold on
59
60      end
61      close(waitbar_MC)
62
63      %% Empirical PDF
64      delta_hist = (UB-LB)/nHist; % = width of the "histograms" of the x-axis
65      % The x-values of the Monte Carlo pdf are centered in the nHist
            intervals
66      % between LB and UB:
67      x_mc_pdf = LB + delta_hist/2 : delta_hist : UB - delta_hist/2; % length
            (x_pdf) = nHist
68      % Cycle over the number of intervals
69      % (the number of values of x falling in the current interval is stored
```

```
    in
% y_pdf, which contains the height of the "histograms")
y_mc_pdf = zeros(1, nHist); % length(y_pdf) = length(x_pdf)
Xvalues = sort(Xvalues);
x_cont = 1;
waitbar_hist = waitbar(0,'Cycle over the number of histograms');
for y_cont = 1:nHist % y_pdf has to be filled
    waitbar(y_cont / nHist)
    while x_cont<=n_MCruns && ( Xvalues(x_cont) <= (x_mc_pdf(y_cont) +
        delta_hist/2))
        y_mc_pdf(y_cont) = y_mc_pdf(y_cont) + 1;
        x_cont = x_cont+1;
    end
end
close(waitbar_hist)
% Normalization (the integral of the MC PDF over the domain has to be
    1):
y_mc_pdf = y_mc_pdf ./ ( sum(y_mc_pdf).*delta_hist );
% Now, x_pdf and y_pdf contain, respectively, the midpoints and the
    heights
% of the "histograms" needed for the Monte Carlo PDF

%% Truncated normal
function [y] = trunc_norm(mu, sigma, a, b)
    pb = normcdf( (b - mu) / sigma );
    pa = normcdf( (a - mu) / sigma );
    diff_val = pb - pa;
    y = mu + sigma * norminv(diff_val * rand + pa);
end

%% Drift coefficient
function y = b_drift(x)
```

```
98      y = - 4.*x ;
99  end
100
101 %% Diffusion coefficient
102 function y =  sigma_diff(x)
103     y = 1 + x.*0;
104 end
```

Listing A.2 contains the code for numerically solving the FP problem (5.5) related to (5.1), according to the SIMEX2 scheme outlined in (5.22).

Listing A.2: Numerical solution of the FP equation with the SIMEX2 scheme

```
%% Initial data
LB = -20; % lower bound of the space domain
UB = 50;  % upper bound of the space domain
NX = 20; % #intervals in the space grid
T = 1; % final time
NT = 400; % #of intervals in the time grid
% Rate and parameter distribution of the jumps
lambda = 5;
mu_jumps = 3;
sigma_jumps = 10;
% Parameters of the Cauchy data
mu_f0 = 15;
sigma_f0 = 3;
% Spatial mesh
dx = (UB-LB) / NX;
x = LB : dx : UB; % NX + 1
% Temporal mesh
dt =  T / NT;
% Numerical solution
f_num = zeros(NX + 1, NT + 1);
% Initial condition
f_num(:,1) = f0(x, mu_f0, sigma_f0);

%% Numerical solution of the PDF
M_ode = matrix_cc(dx, x);
% G = matrix needed for the integral in FP equation
G = zeros(NX+1,NX+1);
for j=1:NX+1
```

```
29      for k=1:NX+1
30          G(j,k) = normpdf(x(j)-x(k), mu_jumps, sigma_jumps) + ...
31                   normpdf(2*LB -x(j)-x(k), mu_jumps, sigma_jumps) + ...
32                   normpdf(2*UB -x(j)-x(k), mu_jumps, sigma_jumps);
33      end
34  end
35  for cont_col = 1 : NX+1
36      G(:,cont_col) = G(:,cont_col) ./ (sum(G(:,cont_col)));
37  end
38  G =  lambda * ( G - eye(NX+1) );
39  % Cycle over the time steps
40  for n = 1 : NT
41      waitbar(n / NT)
42      % 1st step : dt = dt_scheme/2
43      f_mid = solve_ode_pc(dt*0.5, f_num(:,n), M_ode);
44      % 2nd step: integral part (trapezoidal rule)
45      f_star = solve_ode_pc(dt, f_mid, G);
46      % 3rd step : Chang Cooper with dt = dt_scheme/2
47      f_new = solve_ode_pc(dt*0.5, f_star, M_ode);
48      % The numerical solution at time (n+1) is stored
49      f_num(:,n+1) = f_new';
50  end
51
52  %% Function B(x,t)
53  function y = B(x)
54      y = 1+1.*x;
55  end
56
57  %% Function C(x,t)
58  function y = C(x)
59      y = 3 + x.*0;
60  end
```

```
61
62 %% Function f0(x) (initial condition)
63 function y = f0(x,mu_f, sigma_f)
64     y = normpdf(x,mu_f,sigma_f);
65 end
66
67 %% Chang-Cooper matrix
68 % Calculate the matrix of the Chang-Cooper discretization of the
69 % differential operator
70 function [M] = matrix_cc(dx, x)
71 w =  dx .* B(x) ./ C(x) ;
72 delta_cc = w .* 0; % initialization delta chang chooper
73 exp_w = 0 .* w;    % initialization exponential of w
74 id_w_big = ( w > 700 );    % indexes where w is very big, and its exp
       would explode
75 exp_w(id_w_big) = exp(700); % where w is too big, we set w=exp(700)
76 id_w_small = ~ ( id_w_big ); % w is "ok" where it's not too big
77 exp_w(id_w_small) = exp(w(id_w_small)); % where w is "ok", its exp can
       be computed
78
79 id_w_0 = (abs(w) < 1e-10); % indexes where w is close to zero
80 id_w_non_null = ~ ( id_w_0 ); % where w is not too close to zero
81 delta_cc(id_w_0) = 0.5;  % where w is null, delta = 1/2
82 delta_cc(id_w_non_null) = 1 ./ w(id_w_non_null) - 1 ./ (exp_w(
       id_w_non_null) - 1);
83 % where w is far from zero, delta can be computed
84 % Mian diag
85 bound_cond_left = - C( x(1)+dx/2 ) / (dx^2) + ...
86                     B( x(1)+dx/2 ) * delta_cc(1) / dx; % paper
87 bound_cond_right = (-1/dx) * B( x(end)-dx/2 ) / (1-1/exp_w(end-1));
88 Btilda = - ( (  C( x(2:end-1)+dx/2 ) + C( x(2:end-1)-dx/2)  )/dx +...
89             delta_cc(2:end-1) .* B(x(2:end-1)+dx/2) ...
```

```
90                  - ( 1 - delta_cc(1:end-2)) .* B(x(2:end-1)-dx/2)...
91                  ) / dx;
92  Btilda = [ bound_cond_left, Btilda, bound_cond_right] ;
93  Atilda = ( C( x(1:end-1) + dx/2 ) / dx + ...
94             B( x(1:end-1) + dx/2 ) .* ( 1- delta_cc(1:end-1) ) ) /dx;
95  Ctilda = ( B( x(2:end) - dx/2 ) .* delta_cc(2:end) - ...
96             C( x(2:end) - dx/2 ) / dx ) / dx;
97  M = diag(Btilda) + diag(Atilda, +1) + diag(Ctilda,-1);
98  end
99
100 %% ODE
101 % Predictor corrector scheme to solve the ODE containing the
        approximation
102 function [f_succ] = solve_ode_pc(deltat, f_prec, M_ode)
103     f_succ = ( eye(length(f_prec)) + deltat * M_ode + (deltat)^2 * (
            M_ode^2) ) * f_prec;
104 end
```

Listing A.3 shows the proximal algorithm for solving (4.8), which provides the optimal control $\bar{u}$. The inputs are the FP parameters of (5.5), the time and space parameters, the initial guess u_0, and the tolerance *tol*.

Listing A.3: Proximal algorithm

```
function [u] =...
calc_min_proximal(jmax, dx, x, NX, NT, dt, t, f_iniz, G, u, Nxi, xi,
    time_index_jumps, tol, T)
global low
global up
global beta_L1
global nu
p_final = 0 .* x; % terminal value of the adjoint equation
uold = u;
j = 1;
J_current = zeros(1,jmax); % storage of the values of the reduced
    functional
norm_grad_J = zeros(1,jmax); % storage of the values of the norm of the
     reduced gradient
%%Parameters of the optimization algorithm
eta = 1.5; % multiplication factor for the constant L
par2 = 0.2; % inertial factor
c2 = 1e-3;
L = calc_L(dx, x, NX, NT, dt, t, f_iniz, G, u, xi, Nxi,
    time_index_jumps, T);
while j <= jmax
     disp(['Current iteration :  ', num2str(j)])
     % Compute the gradient in u_j (NB: f and p depend on u!)
     % 2.1 : Compute the state f_j (Fokker-Planck equation)
     f_current = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u);
     % 2.2 : Compute the adjoint p_j (adjoint equation)
     jumps_p = calc_jumps_p(Nxi, x, f_current, time_index_jumps, dx, xi)
         ;
```

```
24      [p_current] = solve_adjoint(t, dx, x, NX, NT, dt, p_final, G, u,
            jumps_p, time_index_jumps, Nxi);
25      % 2.3 : Compute the reduced gradient
26      grad_J_current = calc_grad_J(x, t, dx, dt, f_current, p_current, u,
             nu, time_index_jumps, Nxi, jumps_p);
27      norm_grad_J(j) = sqrt ( dt * sum(grad_J_current.^2) ); % needed
            only for the plot
28      J_current(j) =  calc_J(x, f_current, xi, dx, u, nu,
            time_index_jumps, Nxi, dt); % needed for the update of u
29      disp('Value of the functional')
30      disp(J_current(j))
31      % Compute of the first u_tilde (out of the cycle for computing L)
32      % 1 (out of 2) : Compute u_tilde
33      u_tilde = u - 1/L * grad_J_current + par2 * (u - uold);
34      u_tilde = prox(u_tilde, L); % proximal step
35      u_tilde = max(u_tilde,low); % 1st box constraint
36      u_tilde = min(u_tilde,up);  % 2nd box constraint
37      f_tilde = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_tilde); %
            needed for J_tilde
38       % 2 (out of 2) : Compute J_tilde (u_tilde and f_tilde are needed!)
39      J_tilde =  calc_J(x, f_tilde, xi, dx, u_tilde, nu, time_index_jumps
            , Nxi, dt);
40      right_hand_side = J_current + ...
41                        dot( grad_J_current, u_tilde - u ) * dt + ...
42                        0.5 * L * calc_int_f( (u_tilde - u).^2, dt );
43      while (J_tilde - right_hand_side) > 1e-4
44          L = eta * L;
45          % Compute the new u_tilde
46          u_tilde = u - 1./L * grad_J_current + par2*(u - uold);
47          u_tilde = prox(u_tilde, L);
48          u_tilde = max(u_tilde, low);
49          u_tilde = min(u_tilde, up);
```

```
50          f_tilde = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_tilde); %
                needed for J_tilde
51          % Compute J_tilde (u_tilde and f_tilde are needed!)
52          J_tilde =  calc_J(x, f_tilde, xi, dx, u_tilde, nu,
                time_index_jumps, Nxi, dt);
53          right_hand_side = J_current + ...
54                              dot( grad_J_current, u_tilde - u ) * dt +...
55                              0.5 * L * calc_int_f( (u_tilde - u).^2, dt );
56      end
57      Lt = ( L + 2*c2 ) / ( 1 - par2 ); % update of the constant
58      uoldtemp = u; % storage of the old u (needed for the update)
59      % Update of u : 1st step (gradient step + inertial step)
60      u = u - 1./Lt * grad_J_current + par2 * (u - uold);
61      uold = uoldtemp;
62      % Update of u : 2nd step (proximal map + box constraints)
63      u = prox(u, Lt);
64      u = max(u, low);
65      u = min(u, up);
66      [f_check] = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u);
67      [jumps_p_check] = calc_jumps_p(Nxi, x, f_check, time_index_jumps,
            dx, xi);
68      [p_check] = solve_adjoint(t, dx, x, NX, NT, dt, p_final, G, u,
            jumps_p_check, time_index_jumps, Nxi);
69      grad_check = calc_grad_J(x, t, dx, dt, f_check, p_check, u, nu,
            time_index_jumps, Nxi, jumps_p_check);
70      mu_optsyst = - grad_check;
71      stopping_criteria = u - max( 0, u + mu_optsyst - beta_L1 )...
72                            - min( 0, u + mu_optsyst + beta_L1 )...
73                            + max( 0, u - beta_L1 + mu_optsyst - up  )...
74                            + min( 0, u + beta_L1 + mu_optsyst - low);
75      norm_stopping_criteria = sqrt( dt * sum(stopping_criteria.^2) ) ;
76      if norm_stopping_criteria < tol
```

```
77              break
78          end
79          j = j + 1;
80      end
81  end
```

Listing A.4 shows the initial guess of the Lipschitz constant of $\nabla \mathcal{J}_1(u)$ defined in (6.4) and used in Listing A.3.

Listing A.4: Initial guess of the Lipschitz constant of the gradient $\nabla \mathcal{J}_1(u)$

```
function [L] = calc_L(dx, x, NX, NT, dt, t, f_iniz, G, u, xi, Nxi,
    time_index_jumps, T)

global nu

f_mid = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u);
J_mid = calc_J(x, f_mid, xi, dx, u, nu, time_index_jumps, Nxi, dt);

epsL = 10;
norm_epsL = epsL * sqrt(T);

u_plus = u + epsL;
f_plus = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_plus);
J_plus = calc_J(x, f_plus, xi, dx, u, nu, time_index_jumps, Nxi, dt);

u_minus = u - epsL;
f_minus = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_minus);
J_minus = calc_J(x, f_minus, xi, dx, u_minus, nu, time_index_jumps, Nxi
    , dt);

norm_plus  = sqrt( dt * sum( ( J_mid - J_plus ).^2 ) );
norm_minus = sqrt( dt * sum( ( J_mid - J_minus).^2 ) );

L = min(norm_plus,norm_minus) / norm_epsL;

end
```

Listing A.5 shows how the proximal step in Listing A.3 is performed.

Listing A.5: Proximal step in Listing A.3

```
function [L] = calc_L(dx, x, NX, NT, dt, t, f_iniz, G, u, xi, Nxi,
    time_index_jumps, T)

global nu

f_mid = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u);
J_mid = calc_J(x, f_mid, xi, dx, u, nu, time_index_jumps, Nxi, dt);

epsL = 10;
norm_epsL = epsL * sqrt(T);

u_plus = u + epsL;
f_plus = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_plus);
J_plus = calc_J(x, f_plus, xi, dx, u, nu, time_index_jumps, Nxi, dt);

u_minus = u - epsL;
f_minus = solve_FP(dx, x, NX, NT, dt, t, f_iniz, G, u_minus);
J_minus = calc_J(x, f_minus, xi, dx, u_minus, nu, time_index_jumps, Nxi
    , dt);

norm_plus  = sqrt( dt * sum( ( J_mid - J_plus ).^2 ) );
norm_minus = sqrt( dt * sum( ( J_mid - J_minus).^2 ) );

L = min(norm_plus,norm_minus) / norm_epsL;

end
```

Bibliography

[1] A. Addou and A. Benbrik, *Existence and uniqueness of optimal control for a distributed-parameter bilinear system*. Journal of Dynamical Control Systems (2002) **8**, 141–152.

[2] A. Almendral and C. W. Oosterlee, *Numerical Valuation of Options with Jumps in the Underlying*. Applied Numerical Mathematics (2005) **53**, 1–18.

[3] L. Andersen and J. Andreasen, *Jump-Diffusion Processes: Volatility Smile Fitting and Numerical Methods for Option Pricing*. Review of Derivatives Research (2000) **4**, 231–262.

[4] M. Annunziato and A. Borzì, *A Fokker-Planck control framework for multimensional stochastic processes*. Journal of Computational and Applied Mathematics (2013) **237**, 487–507.

[5] M. Annunziato and A. Borzì, *Optimal control of probability density functions of stochastic processes*. Mathematical Modelling and Analysis (2010) **15**, No. 4, 393–407.

[6] D. Applebaum, *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2004.

[7] D. Applebaum, *Lévy processes - from probability to finance and quantum groups*. Notices of the American Math Society (2004) **51**, 1336-47.

[8] S. Asmussen & P. W. Glynn, *Stochastic Simulation - Algorithms and Analysis*. Springer, 2007.

[9] H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke and H. Wolkowicz. *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer, 2011.

[10] F. E. Benth, *Option Theory with Stochastic Analysis. An Introduction to Mathematical Finance*. Springer, 2004.

[11] J. Bertoin, *Lévy Processes*, Cambridge University Press, 1998.

[12] A. Borzì and S. Gonzalez Andrade, *Second-order approximation and fast multigrid solution of parabolic bilinear optimization problems*. Advances in Computational Mathematics (2015) **41**, No. 2, 457–488.

[13] A. Borzì, K. Ito and K. Kunisch, *Optimal control formulation for determining optical flow*. SIAM Journal on Scientific Computing (2002) **24**, No. 3, 818–847.

[14] A. Borzì and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*. SIAM Philadelphia, 2012.

[15] T. Breitenbach, M. Annunziato and A. Borzì, *On the optimal control of a random walk with jumps*. Submitted to Methodology and Computing in Applied Probability (2016), under review.

[16] M. Briani, C. La Chioma and R. Natalini, *Convergence of Numerical Schemes for Viscosity Solutions to Integro-Differential Degenerate Parabolic Problems Arising in Financial Theory*. Numerische Mathematik (2004) **4**, 607–646.

[17] M. Briani and R. Natalini, *Asymptotic High-Order Schemes for Integro-Differential Problems Arising in Markets with Jumps*. Communications in Mathematical Sciences (2006) **4**, 81–96.

[18] M. Briani, R. Natalini and G. Russo, *Implicit-Explicit Numerical Schemes for Jump-Diffusion Processes*. Calcolo (2007) **44**, 33–57.

[19] E. Buckwar and M. Riedler, *Runge-Kutta methods for jump-diffusion differential equations*. Journal of Computational and Applied Mathematics (2011) **6**, No. 236, 1155–1182.

[20] J. Chambers, C. Mallows and B. Stuck, *A method for simulating stable random variables*. American Statistical Association (1976) **71**, 340–344.

[21] J. S. Chang and G. Cooper, *A Practical Difference Scheme for Fokker-Planck Equations*. Journal of Computational Physics (1970) **6**, 1–16.

[22] P. L. Combettes and V. R. Wajs, *Signal recovery by proximal forward-backward splitting*. Multiscale Modeling & Simulation (2005) **4**, No. 4, 1168–1200.

[23] R. Cont and P. Tankov, *Financial Modeling with Jump Processes*. Chapman & Hall, 2004.

[24] R. Cont and E. Voltchkova, *A Finite Difference Scheme for Option Pricing in Jump Diffusion and Exponential Lévy Models*. SIAM Journal on Numerical Analysis (2005) **43**, No. 4, 1596–1626.

[25] R. Cont and E. Voltchkova, *Integro-Differential Equations for Option Prices in Exponential Lévy Models*. Finance and Stochastics (2005) **9**, 299–325.

[26] D. R. Cox and H. D. Miller, *The Theory of Stochastic Processes*. CRC Press, 1977.

[27] M. G. Crandall, H. Ishi and P. L. Lions, *Users' Guide to Viscosity Solutions of Second Order Partial Differential Equations. Bulletin of the American Mathematical Society* (1992) **27**, No. 1, 1–67.

[28] G. Dahlquist and Å. Björck, *Numerical Methods in Scientific Computing: Volume 1*. SIAM, 2008.

[29] J. C. De Los Reyes, *Numerical PDE-Constrained Optimization*. Springer Briefs in Optimization, 2015.

[30] D. Ding and Y. Ying Zhang *A splitting-step algorithm for reflected stochastic differential equations*, Computers & Mathematics with Applications (2008), **55**, No. 1, 2413–2425.

[31] D. Duffy, *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. The Wiley Finance Series, 2006.

[32] A. Einstein, *Über die von der Molekularkinetischen Theorie der Wärme Geforderte Bewegung von in Ruhenden Flüssigkeiten Suspendierten Teilchen*. Annalen der Physik (1905) **322**, No. 8, 549–560.

[33] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems*. SIAM Philadelphia, 1999.

[34] E. Faou, A. Ostermann and K. Schratz, *Analysis of Exponential Splitting Methods for Inhomogeneuos Parabolic Equations*. IMA Journal of Numerical Analysis (2015) **35**, No. 1, 161–178.

[35] W. H. Fleming and R. Rishel, *Deterministic and stochastic optimal control*, Springer-Verlag, 1975.

[36] W. H. Fleming and M. Soner, *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, 1993.

[37] A. V. Fursikov, *Optimal Control of Distributed Systems. Theory and Applications*. American Mathematical Society, 2000.

[38] C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, 2009.

[39] M. G. Garroni and J. L. Menaldi, *Green Functions for Second-Order Parabolic Integro-Differential Problems*. Longman, 1992.

[40] M. G. Garroni and J. L. Menaldi, *Maximum Principle for Integro-Differential Parabolic Operators*. Differential and Integral Equations (1995) **8**, No. 1, 161–182.

[41] M. G. Garroni and J. L. Menaldi, *Regularizing Effect for Integro-Differential Parabolic Equations*. Communication in Partial Differential Equations (1993) **18**, No. 12, 2023–2025.

[42] B. Gaviraghi, M. Annunziato and A. Borzì, *Analysis of splitting methods for solving a partial-integro Fokker-Planck equation*. Applied Mathematics and Computation (2016), (2017) **294**, 1–17.

[43] B. Gaviraghi and A. Borzì, *An operator splitting method for solving a class of Fokker-Planck equations*. European Consortium for Mathematics in Industry Newsletter (2014) **56**, 87–88.

[44] B. Gaviraghi, A. Schindele, M. Annunziato, A. Borzì, *On optimal sparse-control problems governed by jump-diffusion processes*. Applied Mathematics (2016) **7**, 1978–2004.

[45] J. Geiser, *Decomposition Methods for Differential Equations: Theory and Applications*. Chapman & Hall, 2009.

[46] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation*. SIAM Books: Advances in Design and Control Series, 2007.

[47] J. M. Harrison and S. R. Pliska, *A stochastic calculus model of continuous trading: Complete markets*. Stochastic Processes and their Applications (1983) **15**, 313–316.

[48] D. J. Higham, *An algorithmic introduction to numerical simulation of stochastic differential equations*. SIAM Review, Education Section (2001) **43**, 525–546.

[49] D. J. Higham and P. E. Kloeden *Convergence and stability of implicit methods for jump-diffusion systems*. International Journal of Numerical Analysis and Modeling (2006) **3**, No. 2, 125–140.

[50] W. Hundsdorfer and J. G. Verwer, *Numerical Solutions of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer, 2010.

[51] A. Itkin, *Efficient Solution of Backward Jump-Diffusion PIDEs with Splitting and Matrix Exponentials*. Journal of Computational Finance, to appear.

[52] A. Itkin, *Splitting and Matrix Exponentials Approach for Jump-Diffusion Models with Inverse Normal Gaussian, Hyperbolic and Meixner jumps*. Algorithmic Finance (2014) **3**, 233–250.

[53] K. Itô, *On stochastic differential equations*. Memoirs of the American Mathematical Society (1951) **4**, 1–51.

[54] S. Jäger and E. A. Kostina, *Parameter estimation for forward Kolmogorov equation with application to nonlinear exchange rate dynamics*. PAMM (2005) **5**, No. 1, 745–746.

[55] J. Kienitz and D. Wetterau, *Financial Modelling - Theory, Implementation and Practice (with MATLAB Source)*. Wiley Finance Series, 2012.

[56] P. E. Kloeden and E. Platen, *Numerical Solutions of Stochastic Differential Equations*. Springer, 1999.

[57] S. G. Kou, *A jump-diffusion model for option pricing*. Management Science (2002) **48**, No. 8, 1086–1101.

[58] D. P. Kroese, T. Taimre, Z. I. Botev, *Handbook of Monte Carlo Methods*. Wiley, 2011.

[59] A. E. Kyprianou, *Introductory lectures on fluctuations of Lévy processes with applications*. Wiley Finance Series, 2012.

[60] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*. Springer Berlin, 1971.

[61] P. L. Lions and A. S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*. Communications on Pure and Applied Mathematics, (1984) **37**, No. 4, 511–537.

[62] M. Magdziarz, A. Weron and J. Klafter, *Equivalence of the Fractional Fokker-Planck and Subordinated Langevin Equations: the Case of a Time-Dependent Force*. Physical Review Letters (2008) **101** No. 21, Article Number: 210601.

[63] M. Magdziarz, A. Weron and K. Weron, *Fractional Fokker-Planck Dynamics: Stochastic Representation and Computer Simulation*. Physical Review E (2007) **74**, No. 1, Article Number: 016708.

[64] Y. Maghsoodi, *Mean square efficient numerical solution of jump-diffusion stochastic differential equations*. Indian Journal of Statistics (1996) **58**, 25–47.

[65] G. I. Marchuk, *Methods of Numerical Mathematics*. Springer, 1981.

[66] A. Matache, C. Schwab and T. P. Wihler, *Fast Numerical Solution of Parabolic Integrodifferential Equations with Applications in Finance*. SIAM Journal on Scientific Computing (2005) **27**, No. 5, 369–393.

[67] R. C. Merton, *Option pricing when underlying stock returns are discontinuous.* Journal of Financial Economics (1976) **3**, 125–144.

[68] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics.* Springer, 2003.

[69] M. Mohammadi and A. Borzì, *Analysis of the Chang-Cooper Discretization Scheme for a Class of Fokker-Planck Equations.* Journal of Numerical Mathematics (2015) **23**, 271–288.

[70] Y. E. Nesterov, *A method for solving the convex programming problem with convergence rate* $\mathcal{O}(1/k^2)$. Soviet Mathematics Doklady (1983) **27**, No. 2, 372–376.

[71] J. Nocedal, S. Wright, *Numerical Optimization*, Springer, 2006.

[72] P. Ochs, Y. and G. Cooper, *iPiano: Inertial proximal algorithm for nonconvex optimization.* SIAM Journal on Imaging Sciences (2014) **7**, No. 2, 1388–1419.

[73] B. Øksendal, *Stochastic Differential Equations*, Springer, 2003.

[74] A. Pascucci, *PDE and Martingale Methods in Option Pricing.* Springer, 2011.

[75] H. Pham, *Continuous-time Stochastic Control and Optimization with Financial Applications.* Springer, 2009.

[76] E. Platen and N. Bruti-Liberati *Numerical Solution of Stochastic Differential Equations with Jumps in Finance.* Springer, 2010.

[77] S. Primak, V. Kontorovich and V. Lyandres, *Stochastic Methods and Their Applications to Communications.* John Wiley & Sons, 2004.

[78] E. Priola and J. Zabczyk, *Densities for Ornstein-Uhlenbeck Processes with Jumps.* Bulletin of the London Mathematical Society (2009) **41**, 41–50.

[79] P. Protter and D. Talay, *The Euler scheme for Lévy driven stochastic differential equations*, The Annals of Probability (1997) **25**, No. 1, 393–423.

[80] R. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications.* Springer, 1996.

[81] R. T. Rockafellar, *Monotone operators and the proximal point algorithm.* SIAM Journal on Control and Optimization (1976) **14**, No. 5, 877–898.

[82] S. Roy, M. Annunziato and A. Borzì, *A Fokker−Planck feedback control-constrained approach for modelling crowd motion.* Journal of Computational and Theoretical Transport (2016). http://dx.doi.org/10.1080/23324309.2016.1189435

[83] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[84] W. E. Schiesser, *The Numerical Method of Lines*. Academic Press, 1991.

[85] A. Schindele and A. Borzì, *Proximal methods for elliptic optimal control problems with sparsity cost functional. Applied Mathematics* (2016) **7**, No. 9, 967–992.

[86] A. Schindele and A. Borzì, *Proximal schemes for parabolic optimal control problems with sparsity promoting cost functionals*. International Journal of Control (2016), forthcoming.

[87] Z. Schuss, *Theory and Applications of Stochastic Processes: an Analytical Approach*. Springer, 2010.

[88] A. N. Shiryaev, *Probability*. Springer, 1996.

[89] G. Stadler, *Elliptic optimal control problems with L^1-control costs and applications for the placement of control devices*. Computational Optimization and Applications (2009) **44**, No. 2, 159–181.

[90] G. Strang, *On the Construction and Comparison of Difference Schemes*. SIAM Journal on Numerical Analysis (1968) **5**, 506–517.

[91] F. Tröltzsch, *Optimal control of Partial Differential Equations: Theory, Methods and Applications*. American Mathematical Society, 2010.

[92] G. E. Uhlenbeck and L. S. Ornstein, *On the theory of Brownian Motion*. Physical Review (1930) **36**, 823–841.

[93] J. Yong and X. Y. Zhou, *Stochastic Controls, Hamiltonian Systems and Hamilton-Jacobi-Bellman Equations*. Springer-Verlag, 2000.