Approximation and calibration of stochastic processes in finance

JONAS KIESSLING

Doctoral thesis
Stockholm, Sweden 2010

TRITA-MAT-10-MA-11
ISSN 1401-2278
ISRN KTH/MAT/DA 10/05-SE
ISBN 978-91-7415-741-3

KTH Matematik
SE-100 44 Stockholm
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i matematik måndagen den 18 oktober 2010 klockan 13.00 i sal F3, Kungliga Tekniska högskolan, Lindstedsvägen 26, Stockholm.

Tryck: Universitetsservice US-AB

Abstract

This thesis is a study of approximation and calibration of stochastic processes with applications in finance. It consists of an introduction and four research papers.

The introduction is as an overview of the role of mathematics in certain areas of finance. It contains a brief introduction to the mathematical theory of option pricing, as well as a description of a mathematical model of a financial exchange. The introduction also includes summaries of the four research papers.

In Paper I, Markov decision theory is applied to design algorithmic trading strategies in an order driven market. A high dimensional Markov chain is used to model the state and evolution of the limit order book. Trading strategies are formulated as optimal decision problems. Conditions that guarantee existence of optimal strategies are provided, as well as a value-iterative algorithm that enables numerical construction of optimal strategies. The results are illustrated with numerical experiments on high frequency data from a foreign exchange market.

Paper II focuses on asset pricing with Lévy processes. The expected value $E[g(X_T)]$ is estimated using a Monte Carlo method, when X_t is a d-dimensional Lévy process having infinite jump activity and a smooth density. Approximating jumps smaller then a parameter $\epsilon > 0$ by diffusion results in a weak approximation, $\overline{X}_t$, of X_t. The main result of the paper is an estimate of the resulting model error $E[g(X_T)] - E[g(\overline{X}_T)]$, with a computable leading order term.

Option prices in exponential Lévy models solve certain partial integro-differential equations (PIDEs). A finite difference scheme suitable for solving such PIDEs is studied in Paper III. The main results are estimates of the time and space discretization errors, with leading order terms in computable form. If the underlying Lévy process has infinite jump activity, the jumps smaller than some $\epsilon > 0$ are replaced by diffusion. The size of this diffusion approximation is estimated, as well as its effect on the space and time discretization errors. Combined, the results of the paper are enough to determine how to jointly choose the grid size and the parameter ϵ.

In Paper IV it is demonstrated how optimal control can be used to calibrate a jump-diffusion process to quoted option prices. The calibration problem is formulated as an optimal control problem with the model parameter as a control variable. The corresponding regularized Hamiltonian system is solved with a symplectic Euler method.

Sammanfattning

Denna avhandling behandlar approximation och kalibrering av stokastiska processer. Den består av fyra vetenskapliga artiklar samt en inledning.

I inledningen ges en överblick av matematikens roll inom vissa delar av finansvetenskapen. Den innehåller en introduktion till derivatprissättning och en beskrivning av en matematisk modell av en marknad. I inledningen presenteras även de fyra artiklarna.

I Artikel I används beslutsteori och Markovkedjor för att utforma och utvärdera handelsstrategier i en marknad med orderbok. Tillståndet och dynamiken för orderboken modelleras med en Markovkedja av hög dimension och handelsstrategier omformuleras som optimala beslutsproblem. I artikeln anges villkor som garanterar existens av optimala strategier samt en metod för att konstruera dessa strategier. Resultaten illustreras med ett antal numeriska experiment utförda med högfrekvent data från en valutamarknad.

I Artikel II studeras en metod för att beräkna väntevärden $E[g(X_T)]$ med Monte Carlo, då X_t är en d-dimensionell Lévyprocess med obegränsat Lévymått och slät täthetsfunktion. En svag approximation $\overline{X}_t$ av X_t konstrueras genom att ersätta hopp som är mindre än $\epsilon > 0$ med en diffusionsprocess. Huvudresultatet i artikeln är en uppskattning av modellfelet, $E[g(X_T)] - E[g(\overline{X}_T)]$, med beräkningsbar ledande term.

Optionspriser i en exponentiell Lévymodell löser vissa partiella integral- differentialekvationer (PIDEer). I Artikel III studeras ett finit differansschema som är lämpligt att använda för att lösa denna typ av PIDEer. I artikeln anges uppskattningar av diskretiseringsfelen, i både rum och tid, med ledande termer på beräkningsbar form. I fallet då den underliggande Lévyprocessen har ett obegränsat Lévymått används en diffusionsprocess för att approximera hopp som är mindre än $\epsilon > 0$. Beräkningsfelet som orsakas av denna diffusionsapproximation uppskattas, samt dess effekt på rums- och tidsdiskretiseringen.

Den avslutande artikeln, Artikel IV, behandlar modellkalibrering med optimal styrning. Artikeln innehåller en beskrivning av en metod för att kalibrera en hopp och diffusionsprocess till observerade optionspriser. Genom att formulera kalibreringsproblemet som ett styrproblem, där kontrollvariabeln utgörs av den sökta modellparametern, kan optimal styrning tillämpas. Det tillhörande Hamiltonska systemet regulariseras och löses med en symplektisk Euler-metod.

Contents

List of Papers

Paper I: *Algorithmic trading with Markov chains*, Henrik Hult and Jonas Kiessling

Paper I is the result of a close collaboration between the the author and Henrik Hult. Both contributed equally to the text and the development of the theory. The author was responsible for the numerical experiments.

Paper II: *Diffusion approximation of Lévy processes with a view towards finance*, Jonas Kiessling and Raúl Tempone

Paper II is the result of a close collaboration between the author and Raúl Tempone. Both contributed equally to the development of the theory. The author wrote the report under close contact with his co-worker, and

was responsible for the numerical experiments.

Paper III: *Computable error estimates of a finite difference scheme for option pricing in exponential Lévy models*, Jonas Kiessling and Raúl Tempone

Paper III is the result of a close collaboration between the author and Raúl Tempone. Both contributed equally to the development of the theory. The author wrote the report under close contact with his co-worker.

Paper IV: *Calibration of a jump-diffusion process using optimal control*, Jonas Kiessling

Acknowledgments

The first steps towards this work were taken in the spring of 2008. In June 2008 I defended my Licentiate thesis in algebraic topology and started to work full time on this project. Completing the thesis was possible only through the assistance of many people around me. The support of those individuals needs to be acknowledged:

First of all, I thank my advisor Anders Szpessy. Despite the unusual situation, Anders accepted to become my advisor. Working with Anders is interesting, at times demanding, and always inspiring.

I thank my co-worker Raúl Tempone with whom I have written two papers. Working with Raúl has been an intense and instructive experience, and I hope that at least some of our countless planned projects will come to pass. I am also grateful to Raúl for inviting me to UT Austin and KAUST, where a large part of the research was performed.

I am grateful to my second advisor and co-worker Henrik Hult, with whom I have written one paper and hope to write several more. Henrik made sure I had the chance to present my work at various conferences. I have also benefited greatly from our many discussions.

Being part of an active research group has been most beneficial. I especially thank Eric von Schwerin, Jesper Carlsson and Håkon Hoel for the many stimulating exchanges. I also thank Mattias Sandberg for reading a preliminary version of this thesis, his feedback was most welcomed.

I would like to show my gratitude to Dan Laksov for his support at times when it was most needed.

Michael Benedicks encouraged me to apply to graduate school in the first place and guided me at various moments during my studies. For this I am grateful.

I spent my first three year at graduate school doing research in algebraic topology. During that time I worked closely with my former advisor Wojciech Chachólski from whom I learned much mathematics.

I am indebted to my many colleagues for making the math department at KTH a stimulating and fun place to work.

There is much I owe Helena Enflo. Her interest in my work keeps me motivated. Her annoying habit of winning arguments keeps me alert. Her love keeps me going. She also did a fantastic job proofreading the introduction.

Finally, I thank my family, Ann-Charlotte Sörelius, Leif Kiessling, Lisa

Sörelius, and Helene Sörelius, for their unconditional love and support.

A major part of my graduate studies was funded by a grant from the Knut and Alice Wallenberg Foundation. Part of the studies was funded by CIAM. The support of these institutions is gratefully acknowledged.

Introduction

1 A financial market

We begin this thesis with a brief description of the scene of much financial activity – a financial exchange. It should immediately be pointed out that there is an abundance of financial markets around the world today, all with their own special rules and features. The stylized market described in this section is a so-called 'order driven market'. There are today plenty of order driven markets, including several modern stock exchanges, such as the Paris Bourse and the Stockholm Stock Exchange, as well as the interbank currency market EBS. The key characteristic of an order driven market is the absence of centralized market makers. Rather, all market participants have the option of providing liquidity by placing limit orders in the order book.

The primary role of financial mathematics has traditionally been to price complex derivatives (see Section 3) that are not sold on an exchange. Complex derivatives are typically sold 'over the counter', thus keeping the market makers (usually big financial institutions) from revealing their prices to competitors. However derivatives are priced and hedged with the the help of some 'underlying' financial instruments (such as stocks, bonds and vanilla options), that *are* traded on an exchange. In other words, it is at the exchange that market prices are revealed and it is the belief of at least this author that a basic understanding of how market prices are formed helps to understand asset pricing.

In addition, a quantitative understanding of exchanges can be used to construct algorithmic trading strategies. Today this is a huge and booming industry, with giants such as Citadel reaping huge profits from trading opportunities that last only fractions of a second. In Paper I, Markov chains are used to model the limit order book, and it is shown how Markov decision

theory can be used to find optimal strategies. See also Section 2.

There is today an extensive and growing literature on order book dynamics. Some authors are interested in short-term predictions based on the current and recent state of the order book, see for instance Paper I in this thesis, as well as [14] and [4]. A related but slightly different topic is the study of market impact and its relation to optimal execution strategies of large market orders. See for instance [38] and [1].

The econophysics community has produced several interesting empirical studies of statistical properties of the order book. In the enticingly titled paper 'What really causes large price changes?', [22], the authors claim that large changes in share prices are not due to large market orders. They find that statistical properties of prices depend more on fluctuations in revealed supply and demand than on their mean behavior, highlighting the importance of models taking the whole order book into account. In [11], the authors study certain static properties of the order book. They find that limit order prices follow a power-law around the current price, suggesting that market participants believe in the possibility of very large price variations within a rather short time horizon.

Another approach to study the dynamical aspects of limit order books is by means of game theory. Each agent is thought to take into account the effect of the order placement strategy of other agents when deciding between limit or market orders. Some of the systematic properties of the order book may then be explained as properties of the resulting equilibrium, see e.g. [26] and [39] and the references therein.

1.1 Limit orders and the order book

An order driven market is a continuous double auction where agents can submit limit orders. A limit order, or quote, is a request to buy or sell a certain quantity together with a worst allowable price, the limit. A limit order is executed immediately if there are outstanding quotes of opposite sign with the same or better limit. Limit orders that are not immediately executed are entered into the limit order book, where they await the possibility of being matched in the future. An agent having an outstanding limit order can at any time cancel this order. Limit orders are executed using time priority at a given price and price priority across prices.

Following [22], limit orders that are 'effective market orders' are distinguished from 'effective limit orders'. An order resulting in an immediate

transaction is an effective market order, and an order that is not executed but stored in the order book is an effective limit order. For the rest of this thesis, effective market and effective limit orders will be referred to simply as market orders and limit orders.

The limit order book thus contains all outstanding orders at any given time during the opening hours of the exchange. Exchanges only allow discrete prices, with the size between two consecutive price levels often referred to as a 'tick' in the stock market, or a 'bib' in the currency market. Let $X_t = (X_t^n)$ denote the state of the order book at time t, with X_t^n being the total volume (sell positive, buy negative) of quotes with limit n. A buy limit order with volume V submitted at time t_0 with price n_0 will cause

$$X_{t_0}^n = \begin{cases} X_{t_0-}^n & \text{if } n \neq n_0, \\ X_{t_0-}^{n_0} - V & \text{if } n = n_0, \end{cases}$$

where $X_{t_0-}^n$ denotes the outstanding volume at level n just before the order was submitted. The sign in front of V in the equation above is due to the convention that buy orders are given a negative sign.

Assuming that the order book contains both buy and sell quotes, the bid and the ask price are defined respectively by

$$j_B = \max\{j : X_t^j < 0\},$$
$$j_A = \min\{j : X_t^j > 0\}.$$

A buyer looking for a quick transaction can submit a market order at j_B and be sure that the order is executed immediately. The difference between the bid and the ask, $j_A - j_B$, is called the spread. The spread can be thought of as an ill-liquidity premium that measures how the market values the risk connected to submitting a limit order instead of a market order. Illiquid markets typically exhibit larger spreads than liquid markets since low liquidity entails greater risk for large price swings and hence a higher risk for the liquidity provider.

It should be noted that although the limit order book contains all outstanding orders, this information might not be available to the market participants. At some markets only the first few non-zero levels on each side of the order book are reveled. Many markets also allow so-called 'hidden orders'. When an agent submits a hidden order, only a fraction of the volume is revealed. As soon as this part of the order has been executed, a new

fraction of the order is revealed. This enables agents to submit large orders without giving away too much information.

Figures 1.1–1.4 illustrate the order book and how it changes as different orders arrive.

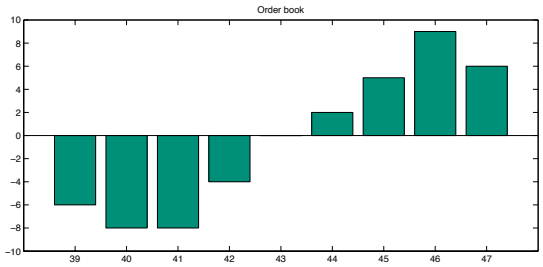

Figure 1.1: State of the order book. The negative volumes to the left represent limit buy orders and the positive volumes represent limit sell orders. In this state $j_A = 44$, $j_B = 42$, and the spread is equal to 2.

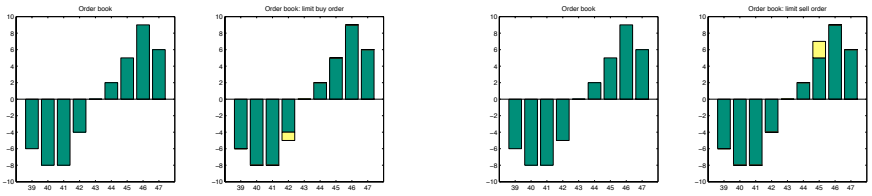

Figure 1.2: Left: Limit buy order of size 1 arrives at level 42. Right: Limit sell order of size 2 arrives at level 45.

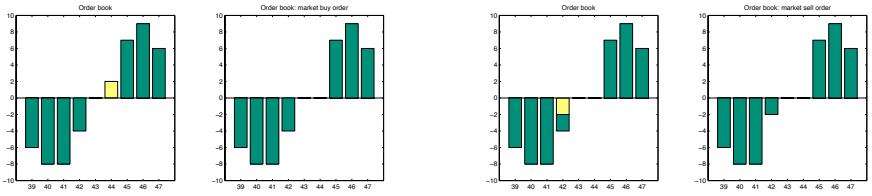

Figure 1.3: Left: Market buy order of size 2 arrives and knocks out level 44. Right: Market sell order of size 2 arrives.

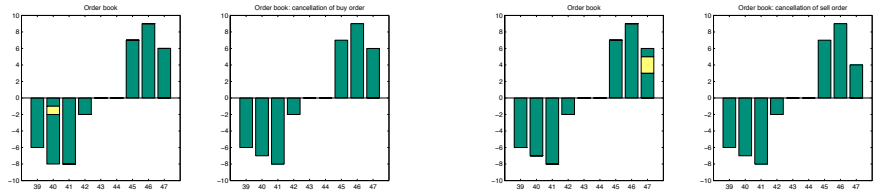

Figure 1.4: Left: A cancellation of a buy order of size 1 arrives at level 40. Right: A cancellation of a sell order of size 2 arrives at level 47.

2 Markov decision theory and Paper I

Suppose an agent wants to buy a unit of a security traded on an order driven market like the one described in Section 1. He is then faced with a range of options. He can submit a market order, obtaining the security at the best available ask price. He also has the option of submitting a limit order hoping that this order eventually is matched against a future market sell order. The answer to what is the best option typically depends on both the agent's view of current market condition, as well as on the state of the order book.

The purpose of Paper I is to set up a framework where one can design and study different trading strategies. The limit order book, see Section 1.1, is modeled as a high dimensional Markov chain, $X = (X_t)$, where each coordinate corresponds to a price level and the state of each coordinate to the number of outstanding limit orders at that level. Doing so opens the door to some powerful tools from probability theory. This is not the first time Markov chains are used to model a limit order book. Paper I was influenced by [14], where the authors model the limit order book using a simple Markov chain model. See also [27] and [43], and the references therein.

It is assumed that there are $d \in \mathbb{N}$ possible price levels in the order book, denoted $\pi^1 < \cdots < \pi^d$. The Markov chain $X_t = (X_t^1, \ldots, X_t^d)$ represents the volume at time $t \geq 0$ of buy orders (negative values) and a sell orders (positive values) at each price level. It is assumed that $X_t^j \in \mathbb{Z}$ for each $j = 1, \ldots, d$. That is, all volumes are integer valued.

Having established a Markovian framework, trading decisions such as the problem of buying one security at the best expected price explained above, can be interpreted as a Markov decision problem. But first some general

results on Markov decision theory.

2.1 Markov decision theory

Let $(X_n)_{n=0}^{\infty}$ be a Markov chain in discrete time on a countable state space $\mathbf{S}$ with transition matrix P. In the applications to come, the discrete Markov chain (X_n) is not the same as, but will be related to, the Markov chain used to model the limit order book described in the previous section. Let A be a finite set of possible actions. Every action can be classified as either a continuation action or a termination action. The set of continutation actions is denoted $\mathbf{C}$ and the set of termination actions $\mathbf{T}$. Then $\mathrm{A} = \mathbf{C} \cup \mathbf{T}$ where $\mathbf{C}$ and $\mathbf{T}$ are disjoint. When a termination action is selected, the Markov chain is terminated. Every action need not be available in every state of the chain. An example of a termination action could be the submission of a market buy order.

For every action there are associated costs. The cost of continuation is denoted $v_C(s, a)$. The cost of termination is denoted $v_T(s, a)$. It is assumed that both v_C and v_T are non-negative and bounded. The termination cost of submitting a market buy order would be the buy price.

A policy $\alpha = (\alpha_0, \alpha_1, \dots)$ is a sequence of functions: $\alpha_n : \mathbf{S}^{n+1} \to \mathrm{A}$ such that $\alpha_n(s_0, \dots, s_n) \in A(s_n)$, where $A(s)$ denotes the actions available at state s and $(s_0, \dots, s_n) \in \mathbf{S}^{n+1}$. If after n transitions the Markov chain has visited $(X_0, \dots, X_n)$, then $\alpha_n(X_0, \dots, X_n)$ is the action to take when following policy α.

The expected total cost starting in $X_0 = s$, and following a policy α until termination, is denoted by $V(s, \alpha)$. This could (and will) be interpreted as the expected buy price. The purpose of Markov decision theory is to analyze optimal policies and optimal (minimal) expected costs. A policy α_* is called optimal if, for all states $s \in \mathbf{S}$ and policies α,

$$V(s, \alpha_*) \leq V(s, \alpha).$$

The optimal expected cost V_* is defined by

$$V_*(s) = \inf_{\alpha} V(s, \alpha).$$

Clearly, if an optimal policy α_* exists, then $V_*(s) = V(s, \alpha_*)$.

A stationary policy is a policy that does not change with time. That is $\alpha_* = (\alpha_*, \alpha_*, \dots)$, with $\alpha_*(X_0, \dots, X_n) = \alpha_*(X_n)$ for $n = 1, 2, \dots$. Here α_* denotes both the policy as well as each individual decision function.

Let V_n denote the minimal incurred cost before time n with termination no later than n. It is defined recursively as

$$\begin{aligned} V_0(s) &= \min_{a\in\mathbf{T}} v_T(s,a), \\ V_{n+1}(s) &= \min\Big(\min_{a\in\mathbf{C}} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_n(s'), \min_{a\in\mathbf{T}} v_T(s,a)\Big), \end{aligned} \tag{2.1}$$

for $n \geq 0$. Let $V_\infty(s)$ denote the limit of the monotone sequence $(V_n)_{n\geq 0}$. It is straightforward, see Lemma 4.2 in Paper I, to verify that V_∞ satisfies a Bellman equation

$$V_\infty(s) = \min\Big(\min_{a\in\mathbf{C}} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_\infty(s'), \min_{a\in\mathbf{T}} v_T(s,a)\Big). \tag{2.2}$$

In fact, under certain extra conditions, V_∞ is the optimal expected cost:

Theorem 2.1 *Let $\aleph$ be the collection of policies α that terminate in finite time, i.e. $\mathbb{P}[\tau_\alpha < \infty \mid X_0 = s] = 1$ for each $s \in \mathbf{S}$. Let $\alpha_\infty = (\alpha_\infty, \alpha_\infty, \dots)$ be a stationary policy where $\alpha_\infty(s)$ is a minimizer to*

$$a \mapsto \min\Big(\min_{a\in\mathbf{C}(s)} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_\infty(s'), \min_{a\in^T(s)} v_T(s,a)\Big). \tag{2.3}$$

The following statements hold.

(a) For each $\alpha \in \aleph$, $V(s,\alpha) \geq V_\infty(s)$.

(b) V_∞ is the optimal expected cost for $\aleph$. That is, $V_\infty = \inf_{\alpha\in\aleph} V(s,\alpha)$.

(c) If $\alpha_\infty \in \aleph$, then $V_\infty(s) = V(s,\alpha_\infty)$.

(d) Suppose that W is a solution to the Bellman equation (2.2) and let α_w denote the minimizer of (2.3) with V_∞ replaced by W. If $\alpha_w, \alpha_\infty \in \aleph$ then $W = V_\infty$.

In particular, if all policies belong to $\aleph$, then V_∞ is the unique solution to the Bellman equation (2.2). Moreover, V_∞ is the optimal expected cost and is attained by the stationary policy α_∞.

Proof: See Theorem 4.3 in Paper I.

In practice one can use Algorithm 2.1 to construct close to optimal policies.

To see that Algorithm 2.1 indeed constructs a close to optimal policy, recall first the $\{V_n\}_{n\geq 0}$ defined in (2.1). Since $\{V_n\}_{n\geq 0}$, defined by (2.1),

Algorithm 2.1 Optimal trading strategies

Input: Tolerance TOL, transition matrix P, state space $\mathbf{S}$, continuation actions $\mathbf{C}$, termination actions $\mathbf{T}$, continuation cost v_C, termination cost v_T.

Output: Upper bound V_n of optimal cost and almost optimal policy α_n.

Let

$$V_0(s) = \min_{a\in\mathbf{T}} v_T(s,a), \text{ for } s \in \mathbf{S}.$$

Let $n = 1$ and $d >$ TOL.

while $d >$ TOL **do**

Put

$$V_n(s) = \min\Big(\min_{a\in\mathbf{C}} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_{n-1}(s'), \min_{a\in\mathbf{T}} v_T(s,a)\Big),$$

and

$$d = \max_{s\in\mathbf{S}} V_{n-1}(s) - V_n(s), \text{ for } s \in \mathbf{S}$$
$$n = n + 1.$$

end while

Define $\alpha : \mathbf{S} \to \mathbf{C} \cup \mathbf{T}$ as a minimizer to

$$\min\Big(\min_{a\in\mathbf{C}(s)} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_{n-1}(s'), \min_{a\in\mathbf{T}} v_T(s,a)\Big).$$

converges to $V_\infty(s)$, a close to optimal policy is obtained by finding one that attains the expected cost at most $V_n(s)$ for some large n. For $s \in \mathbf{S}$. Let $\alpha_0(s)$ be a minimizer of $a \mapsto v_T(s,a)$ and for $n \geq 1$, $\alpha_n(s)$ is a minimizer of

$$a \mapsto \min\Big(\min_{a\in\mathbf{C}(s)} v_C(s,a) + \sum_{s'\in\mathbf{S}} P_{ss'}(a)V_{n-1}(s'), \min_{a\in\mathrm{T}(s)} v_T(s,a)\Big). \tag{2.4}$$

The stationary policy α_n from (2.4) coincides with the policy constructed by Algorithm 2.1. By the following theorem, the expected cost following α_n is bounded by V_n.

Theorem 2.2 *The policy* $\alpha_{n:0} = (\alpha_n, \alpha_{n-1}, \dots, \alpha_0)$ *has expected total cost given by* $V(s, \alpha_{n:0}) = V_n(s)$*. Moreover, if the stationary policy*

$$\alpha_n = (\alpha_n, \alpha_n, \dots)$$

satisfies $\tau_{\alpha_n} \in \aleph$*, then the expected total cost of* α_n *satisfies*

$$V_n(s) \geq V(s, \alpha_n) \geq V_\infty(s).$$

The book [37] contains a very readable brief introduction to Markov decision processes. A more detailed introduction can be found in [19] and [24]. See also [45] for an interesting discussion on algorithms and Markov decision theory.

2.2 Optimal buy strategies

The problem of finding optimal buy strategies can now be phrased as a Markov decision process, and analyzed with the methods presented in Section 2.1. The complexity of the problem depends on the possible actions of the agent. In Paper I three different buy strategies are analyzed, see Sections 3.3, 4.2 and 4.3.

The 'keep-or-cancel' buy strategy, described in Section 4.2, can be summarized as follows:

In this setting the agent is assumed to initially submit a limit order at level j_0. After each transition of the limit order book, she has two options: 1) do nothing, 2) cancel the limit order and submit a market order at the best available market price.

The agent is assumed to have decided on a maximum price level J. If the agent's limit buy order has not been processed, and J equals the best ask price, then the agent will immediately cancel the buy order and place a market buy order at level J. It will be implicitly assumed that there are always are limit sell orders available at level J. Buying at level J can be thought of as a stop-loss. Placing a stop-loss is typically enough to guarantee that all strategies terminate in finite time.

By Theorem 2.1, an optimal keep-or-cancel strategy for buying one unit

is the stationary policy α_∞, with expected buy price V_∞ satisfying

$$V_\infty(s) = \min\Big(\min_{a\in\mathbf{C}} \sum_{s'\in\mathbf{S}} P_{ss'}V_\infty(s'), \min_{a\in\mathbf{T}} v_T(s)\Big)$$
$$= \begin{cases} \pi^{j_0}, & \text{if the limit order is matched,} \\ \pi^J, & \text{stop-loss,} \\ \min\Big(\sum_{s'\in\mathbf{S}} P_{ss'}V_\infty(s'), \pi^{j_A(s)}\Big), & \text{otherwise.} \end{cases}$$

Here $P_{ss'}$ denotes the possibility that a transition takes the Markov chain from state s to state s'. Recall that π^n is the price corresponding to level n, so π^{j_0} is the final buy price if the limit order is executed, π^J is the price at the stop-loss and $\pi^{j_A(x)}$ is the best offer in state s.

The stationary policy α_n in Theorem 2.2 provides a useful numerical approximation of an optimal policy, and $V_n(s)$ in (2.1) provides an upper bound of the expected buy price. Both V_n and α_n can be computed by Algorithm 2.1.

2.3 Parameterization, calibration and numerical experiments

Recall from Section 1.1 the different orders that define the possible transitions of the order book. To completely determine the Markov chain $X = (X_t)$, it is enough to specify the arrival intensities and size distributions of these orders. In Section 5 in Paper I a model that strikes a balance between flexibility and tractability is suggested. It can be summarized as follows:

- Limit buy orders arrive at a distance of i levels from best ask level with intensity $\lambda_L^B(i)$.
- Market buy orders arrive with intensity λ_M^B.
- The size of limit and market orders is exponentially distributed with parameters α_L and α_M respectively. That is, the distributions $(p_k)_{k\geq 1}$ and $(q_k)_{k\geq 1}$ of limit and market order sizes are given by

$$p_k = (e^{\alpha_L} - 1)e^{-\alpha_L k}, \qquad q_k = (e^{\alpha_M} - 1)e^{-\alpha_M k}.$$

- The size of cancellation orders is assumed to be 1. Each individual unit size buy order located at a distance of i levels from the best ask level is

cancelled with a rate $\lambda_C^B(i)$. At the cumulative level the cancellations of buy orders at a distance of i levels from opposite best ask level arrive with a rate proportional to the volume at the level: $\lambda_C^B(i)|x^{j_A-i}|$.

- Sell orders are modeled in the same way: market sell orders arrive with intensity λ_M^S, if i denotes the distance from best bid, then limit sell orders arrive with intensity $\lambda_L^S(i)$, and the cancellations of limit sell orders always have size 1 and arrive with rate $\lambda_C^S(i)|x^{j_B+i}|$.

The model described above is an example of a *zero intelligence model*: Transition probabilities are state independent except for their dependence on the location of the best bid and ask. Zero intelligence models of the market's micro structure were considered already in 1993, in the work of Gode and Sunder [27]. Despite the simplicity of such models, they capture many important aspects of the order driven market. Based on a mean field theory analysis, the authors of [43] and [18] derive laws relating the mean spread and the short term price diffusion rate to the order arrival rates. In [23], the validity of these laws are tested on high frequency data on eleven stocks traded at the London Stock Exchange. The authors find that the model does a good job of predicting the average spread and a decent job of predicting the price diffusion rate.

The model outlined above is straightforward to calibrate to market data. Moreover, the calibration is fast and easy to implement, a fact that could be potentially important in applications. See Section 5.1 in Paper I for details regarding the calibration.

Paper I is concluded by some numerical experiments on data from a foreign exchange market. High frequency data on USD/EUR exchange rates from EBS, a major foreign exchange market, is used for calibration. Different buy strategies are then evaluated using Algorithm 2.1 and compared.

Figure 1.5 illustrates the expected buy price and choice matrix for different market situations following a strategy called the 'ultimate-buy-strategy', explained in Section 4.3 in Paper I. In this particular example, the best sell price is at 1.4344. The agents buy order is currently placed 2 of a total of 10 buy orders at 1.4342. The image to the right illustrates the expected buy price for different number of outstanding orders at 1.4343 and 1.4344. The image to the left illustrates the choice matrix, the color in each square determines which action to take given that precise market situation.

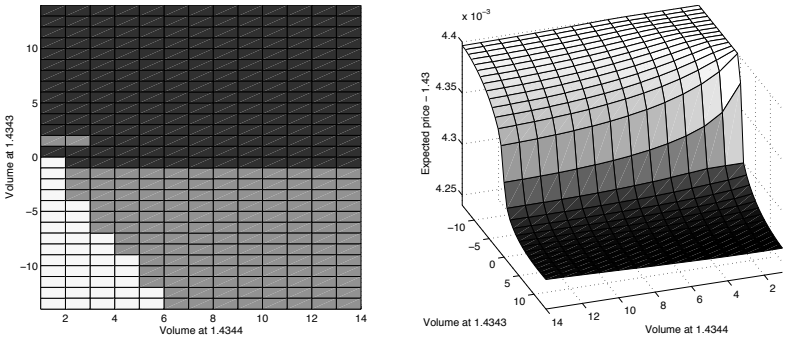

Figure 1.5: Expected buy price and optimal choice of buy order placement in the ultimate buy-one-unit strategy. The buy order currently has place 2 out of a total of 10 limit buy orders at price level 1.4342. The choice matrix to the left shows optimal choice given different volumes at levels 1.4343 and 1.4344. Dark grey indicates that the buy order should be kept in place. Light grey indicates that the buy order should be cancelled and resubmitted at level 1.4343. In the white region, the buy order should be cancelled and replaced by a market order. The plot to the right shows the optimal expected buy price.

3 Derivative pricing

A derivative is an asset that derives its value from an other asset, the underlying. Usually the underlying is some tangible asset, such as a stock or a bond. A derivative can typically be viewed as a contingent claim written on the underlying. A contingent claim is a contract with a payoff contingent on the future state of the underlying, i.e. it is essentially a bet on a future state of the underlying.

A simple example of a derivative is the European call option. A call option gives the owner of the option the right, but not the obligation, to buy a certain quantity of the underlying at a certain price, the strike price. The term 'European' specifies that this option can only be exercised at expiration, also called the maturity of the option. Let S_t denote the value of a stock at time t. A call option with strike K and maturity T can be

considered a contingent claim that at time T pays the owner

$$g(S_T) = \max(S_T - K, 0).$$

A European put option is a contract that instead gives the owner the right to sell the underlying at a predefined price.

More generally, a European style contingent claim written on the stock S, with maturity T and payoff $g : \mathbb{R} \to \mathbb{R}$, is a contract that at expiration pays the owner $g(S_T)$.

The contingent claim is called American if the owner has the possibility of early exercise, i.e. at any time t from the time of issuing until maturity, the owner of the contract can take the payment $g(S_t)$. Examples of American style derivatives include American put and call options.

Both put and call options, of both European and American type, are today widely traded on various exchanges, much like the one outlined in Section 1, around the world.

One of the main objectives of financial mathematics is to price derivatives in terms of some observable market conditions, such as the price and volatility of the underlying. Essentially one would like to determine a function, depending on the time t, the price of the underlying S_t, and possibly on some other market parameters, such that the price of the derivative at time t equals the value of the function. That such a function exists is really hoping for quite a bit. Indeed, it should be clear from the previous section that one cannot simply speak about the value of a financial instrument. The value of one stock is, to a seller of that stock, the highest bid price in the market. This is different from the value to a buyer. It might also be different from the value of several stocks due to the market impact of unloading a large position. If we need to observe the interaction of thousands of market participants on an exchange to value stocks, why should a call option be so much easier to value?

The very existence of exchanges for vanilla options (i.e. put and call options) seems to imply that their valuation is in fact as difficult as the valuation of stocks. Nevertheless, financial mathematics can provide helpful insights and valuable tools, even though derivative pricing is more than just a mathematical exercise.

3.1 The Black and Scholes equation

In the seminal publication 'The Pricing of Options and Corporate Liabilities', [10], Black and Scholes derived an explicit pricing formula for European put and call options. The formula is probably still the most widely used mathematical formula in quantitative finance. In addition to being an apparently useful result, the derivation touches several ideas that are key to understanding how mathematics enter into the world of derivatives.

Suppose that an agent believes that future stock price returns are log-normally distributed, or, put differently, that the stock price S_t satisfies the stochastic differential equation

$$\mathrm{d}S_t = S_t \left(\alpha \mathrm{d}t + \sigma \mathrm{d}W_t\right). \tag{3.1}$$

Here and for the rest of this introduction W_t denotes the Wiener process. Also α denotes the expected rate of return of the asset, and σ is called the volatility of the price process.

The agent is interested in issuing a call option on S_t with strike K and maturity T and is trying to determine a reasonable price.

The market is assumed ideal in the sense that the following assumptions hold:

(a) The interest rate r is known and constant.

(b) The stock pays no dividends.

(c) There are no transaction costs.

(d) There are no penalties for short selling.

(e) Continuous trading in the underlying is possible.

At time $t = 0$ the agent sells the option and invests in δ_0 units of the underlying, and puts β_0 in the bank. The agent has thus established a portfolio that at time t is worth

$$P_t = \delta_t S_t + \beta_t - C_t,$$

where C_t denotes the unknown value of the option. The portfolio is assumed to be 'self-financing', meaning that the agent is allowed to change her investment in the underlying δ_t and the amount in the bank β. However, no money is allowed to be neither withdrawn nor entered into the portfolio. At maturity the value equals $P_T = \delta_T S_T + \beta_T - \max(S_T - K, 0)$. Self-financing is equivalent to stating that the dynamics of the portfolio equal

$$\mathrm{d}P_t = \delta_t \mathrm{d}S_t + r\beta_t \mathrm{d}t - \mathrm{d}C_t.$$

Itô's lemma applied to $\mathrm{d}C_t$ yields

$$\mathrm{d}C = \left(\frac{\partial C}{\partial t} + \alpha S \frac{\partial C}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 C}{\partial S^2} \right) \mathrm{d}t + \sigma S \frac{\partial C}{\partial S} \mathrm{d}W_t.$$

Note that the uncertainty in the option value is captured by the term $\partial C/\partial S \, \mathrm{d}W_t$. The uncertainty in the portfolio is eliminated if the agent trades so that $\partial P/\partial S = 0$. This is achieved by keeping

$$\delta_t = \frac{\partial C}{\partial S}.$$

The portfolio value is now completely deterministic:

$$\mathrm{d}P = \left(\beta r - \frac{\partial C}{\partial t} - \frac{\sigma^2 S^2}{2} \frac{\partial^2 C}{\partial S^2} \right) \mathrm{d}t. \tag{3.2}$$

A portfolio, which future value is completely deterministic, is 'as good as money in the bank', provided, of course, that the returns equal the interest rate. Suppressing any concerns the agent might have with her market view, this implies that she would be indifferent to initiating the transactions provided $\mathrm{d}P = rP\mathrm{d}t$. Inserting this equality into (3.2) results in the famous Black & Scholes differential equation,

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 C}{\partial S^2} = rC, \tag{3.3}$$

that together with the terminal condition $C(T, S) = \max(S - K, 0)$ determines $C = C(t, S)$. In fact, (3.3) can be used to price any European contingency claim. The indifference price $C = C(t, S)$ of a contingency claim with payoff g satisfies (3.3) and the terminal condition $C(T, S) = g(S)$.

A straightforward application of the Feynman–Kac formula to (3.3) establishes the existence of a probability measure Q such that, under Q,

$$\mathrm{d}S_t = S_t \left(r\mathrm{d}t + \sigma \mathrm{d}W_t \right),$$

and the option value can be determined as the expected value

$$C(t, S) = E^Q \left[e^{-(T-t)r} \max(S_T - K, 0) : S_t = S \right]. \tag{3.4}$$

Remark 3.1 *The derivation of* (3.3) *and* (3.4) *depend on the assumption of a perfect market, defined by the statements (a)–(e). In a real market these statements would typically all be, to a varying degree, false.*

Plenty of studies have been devoted to non–perfect markets. Options on stocks paying dividends were studied already in [35]. Markets with non-constant interest rates have also been a popular subject, see for instance [9]. Markets with transaction costs are treated in [32].

The option value, given by (3.3) or (3.4), is the market value of a call option in the eyes of an agent holding the belief that (3.1) is a correct description of the future states of the stock price. In addition to establishing the value of the option, the derivation determine the strategy that the agent should follow to achieve that value. Note that the option price is independent of the expected rate of return α. Hence, in the same way as a stock can be seen as a bet on its future rate of return, a call option is a bet on the future volatility of the underlying. It is, however, also a bet on the belief that future stock price returns are (log-)normally distributed. Considering the turbulent nature of financial markets, there are probably rather few market participants that would subscribe to such a view today. Non-normally distributed price returns are popular also in academia, see for instance Papers I-IV in this thesis. Disproving the assumption of normally distributed returns actually has a long and venerable history predating even the work of Black and Scholes. In his famous study of cotton prices, Mandelbrot suggested already in the 1960th that certain asset returns have no well-defined second moment and are hence very far from being normally distributed, see for instance [34]. An interesting and provoking attack on financial mathematics in general, and the 'paradigm of the normal distribution' in particular, is the book [44].

3.2 No arbitrage

Arbitrage is defined as making risk-free money with zero initial investment. It is easy to understand why the assumption of 'no arbitrage' is popular in finance and economics. As with all ideal assumptions, it is not true. However, in a market where prices are established through supply and demand, arbitrage opportunities tend to be scarce and short lived. The following is an example of a typical arbitrage trade that is quite popular today:

Example 3.2 *Volvo is today traded on two large exchanges in Stockholm: Burgundy and Nasdaq OMX. Since the two exchanges are independent, there is nothing guaranteeing that the Volvo shares trade at exactly the same price all the time at both exchanges. Indeed, every now and then a large order arrives at one of the exchanges and pushes the price in some direction. It then becomes possible to buy Volvo at (say) Burgundy and at the same time sell at Nasdaq OMX, with a buy price that is strictly lower then the sell price. Since the trade is so profitable, the opportunity typically exists only for a fraction of a second.*

Despite its simplicity the assumption of no arbitrage has some surprisingly strong consequences. For instance, if (3.1) was indeed the true dynamics of a certain stock, and everyone knew this, then the price of a call option would truly satisfy the Black & Scholes equation (3.3). Another consequence of the no arbitrage assumption is a relationship between the price of call option, the price of a put option and the stock price, the so-called put-call parity:

Example 3.3 *Consider a call option with strike K and maturity T and a put option on the same underlying with the same strike and maturity. An investor buys one unit of the underlying and the call option and sells the put option. At time t the agent's portfolio is worth*

$$S_t + P_t - C_t,$$

where C_t is the value of the call, P_t the value of the put and S_t the value of the underlying. At maturity the portfolio is worth

$$S_T + max\,(K - S_T, 0) - max\,(S_T - K, 0) = K. \tag{3.5}$$

By no arbitrage, if (3.5) *holds at time T, then at all times $t < T$,*

$$S_t + P_t - C_t = e^{-r(T-t)}K. \tag{3.6}$$

3.3 Replicating portfolio and market completeness

A contingency claim C with payoff g and maturity T is said to be replicated by a portfolio P if P is self-financing and the value of P and C are equal (almost surely) at time T,

$$P_T = C_T = g(S_T).$$

Here 'replicating portfolio' should be understood in a broad sense: a replicating portfolio is a self-financing portfolio and with a trading strategy.

In the text so far we have encountered several examples of replicating portfolios: in the derivation of (3.3), $\delta_t S_t + \beta_t$ replicated the value of the call option. A second replicating portfolio of the call option is given by the put-call parity (3.6). In Example 3.2, a share of Volvo bought at the Nasdaq OMX replicates a share bought at Burgundy. Note the differences between the first example and the rest. Replicating a call option in the underlying was possible only due to the very specific assumptions of the market dynamics (3.1) whereas the other examples are independent of the stock price dynamics.

Replicating portfolios in arbitrage free markets provide a powerful tool in asset pricing: the value of a claim replicated by a portfolio must at all times equal the value of the portfolio, otherwise the difference between the claim and the portfolio would generate a risk-free profit, an arbitrage. Claims that can be replicated thus have a well-defined value in terms of the replicating portfolio.

A (system of) market(s) is said to be complete if any claim can be replicated:

Definition 3.4 *Let $\mathcal{H}$ denote a set of contingency claims. A market is said to be complete with respect to $\mathcal{H}$ if all claims $h \in \mathcal{H}$ can be replicated.*

So for instance, in a stock market where the price of a stock solves (3.1), is, by (3.3) complete with respect to any European style contingency claim.

Any book on derivative pricing will contain a discussion on replicating portfolios and market completenes. See for instance [7, 8, 15]. An interesting discussion of market completness from an economists perspective is provided by Flood in [25]. He argues that exchange traded derivatives make sense from an economic point of view as a way of completing the market.

3.4 The Q measure

At the end of Section 3.1, a probability measure Q on the future states of the stock price process was constructed. Under the measure Q, the price of a European contingency claim was equal to its discounted expected future payoff, (3.4). It is rather surprising that, under quite reasonable assumptions, such pricing measures can be constructed.

Let $\mathcal{H}$ be a collection of European style contingency claims with maturity T on some underlying with price S_t. A pricing rule Π is a function associating a price $\Pi(h)$ to each claim $h \in \mathcal{H}$. Suppose that Π is linear and excludes arbitrage. One can for instance think of Π as the price of these claims at some exchange. If the collection $\mathcal{H}$ is rich enough, and if Π satisfies certain regularity conditions, then the pricing rule determines and is determined by a probability measure.

Suppose that for every measurable set $A \subset \mathbb{R}$, there is a claim 1_A paying 1 at time T in the event of $S_T \in A$, and zero otherwise. For instance, the claim $1_{\mathbb{R}}$ pays 1 at time T independent of S_T, a zero-coupon bond. Such a claim is perfectly replicated by investing e^{-rT} in the bank at time $t = 0$. The exclusion of arbitrage implies that $\Pi(1_{\mathbb{R}}) = e^{-rT}$. Let Q be given by:

$$Q(A) = e^{rT}\Pi(1_A).$$

By linearity, if A and B are disjoint then $Q(A \cup B) = Q(A) + Q(B)$. Extending Q linearly to countable sums thus defines a measure, called the *pricing* or *risk neutral* measure. Note that $Q(\mathbb{R}) = e^{rT}\Pi(1_{\mathbb{R}}) = 1$, so Q is a probability measure.

The expected payoff, under Q, of a contingency claim $f \in \mathcal{H}$ is denoted by $E^Q[f]$. By linearity, if f can be decomposed into a finite sum $f = \sum_{k=1}^n \lambda_k 1_{A_k}$, then

$$\Pi(f) = e^{-rT}E^Q[f]. \tag{3.7}$$

In fact, if the pricing rule Π is determined by its value on claims on the form $f = \sum_{k=1}^n \lambda_k 1_{A_k}$ then (3.7) is valid for all claims $f \in \mathcal{H}$, so in this situation Π can be recovered from its pricing measure Q.

The density of the pricing measure can be determined explicitly in terms of the price of call options. This was first observed by Dupire in [20], and plays an important role e.g. in model calibration (see Paper IV). Let $p_T(S)$ denote the density of the pricing measure. The price of a call option with strike K and maturity T, $C(T, K)$, equals

$$\begin{aligned} C(T,K) &= e^{-rT}E^Q\left[\max\left(S_T - K, 0\right)\right] \\ &= e^{-rT}\int_{\mathbb{R}} \max\left(S - K, 0\right) p_T(S)\mathrm{d}S. \end{aligned}$$

Differentiating twice yields

$$p_T(K) = e^{rT}\frac{\partial^2 C}{\partial K^2}. \tag{3.8}$$

In terms of market completeness discussed in Section 3.3, equation (3.8) implies in particular that a market containing European call options is complete with respect to the collection of *all* European contingency claims.

The absence of arbitrage imposes certain conditions on the pricing measure. For instance, the present value of a claim paying S_T at maturity, a futures, is S_0. Indeed, such a claim is perfectly replicated by a unit of the underlying. Put differently, the discounted price $e^{-rt}S_t$ is a martingale under Q,

$$S_0 = e^{-rT}E^Q[S_T]. \tag{3.9}$$

Equation (3.7) establishes a link between contingency pricing on the one hand, and stochastic calculus on the other. Except for (3.8) and (3.9), there is however very little information regarding how to choose the pricing measure Q. This unfortunate fact has lead to a veritable explosion of literature treating different pricing measures. The standard approach is to parametrize the measure in terms of the dynamics of the underlying, i.e. specifying the so-called risk-neutral dynamics of the underlying. Popular models include jump-diffusion, [2] and Paper IV, stochastic volatility, [28, 5], and pure jump models, [15]. Lévy processes provide a particular class of jump-diffusion and pure jump models, and are discussed in Section 3.5.

Note however that Q does not quantify uncertainty concerning future stock price movements. Unless the underlying market is complete (as in Section 3.1), it is unclear how the dynamics of the underlying affect the pricing measure. Determining how statistical properties of the stock price, and how subjective beliefs of the future, affect the pricing measure remains an important question that deserves more attention from both mathematicians and economists.

Risk-neutral option pricing involves two important steps: 1) estimation of the expected value (3.7), and 2) model calibration. Different models and contracts require different methods: Monte Carlo methods are treated in Paper II and option pricing using integro-differential equations is treated in Paper III. An other popular method is to use the fast Fourier transform, see for instance [13]. Model calibration is the process of determining the pricing measure from observed option prices. That is, the measure Q is determined so that the theoretical prices, as given by (3.7), equal the observed prices. In Paper IV it is demonstrated how optimal control can be used to calibrate a jump-diffusion process.

3.5 Lévy processes

An increasingly popular way to parametrize the pricing measure is to let the underlying asset be driven by a Lévy process. A common motivation for using Lévy processes is that they capture the 'jumpy' nature, as well as the 'heavy tail' of asset returns, see for instance [30] or [15]. However, as was pointed out in the previous section, the pricing measure Q is defined by a pricing rule and is *not* a measure of price dynamics. Since markets driven by Lévy processes are typically not complete, it is not clear how the Lévy structure of the market translates into a pricing rule for contingency claim.

Arguing from an asset pricing perspective, there are still good reasons why Lévy processes are useful. Lévy processes provide a class of processes that is flexible enough to capture observed volatility smiles and skews. They are semimartingales and hence posses good stochastic integration properties. The Lévy-Khinchin formula, (3.13), provides an explicit formula for the Fourier transform. Contingency claims satisfy certain parabolic partial integro-differential equations, similar to (3.3), a fact useful in numerical computations.

Definition 3.5 *A stochastic process X_t is called a Lévy process if the increments of X_t are independent and stationary, and the map*

$$t \mapsto X_t$$

is almost surely right continuos with left limits.

Two standard examples of Lévy processes are Brownian motion and the Poisson process.

Fix a Lévy process X_t in $\mathbb{R}^d$. The processes is said to jump at t_0 with size $s \neq 0$ provided

$$s = X_{t_0} - X_{t_0-} \neq 0.$$

Here X_{t_0-} denotes the left limit at t_0, guaranteed to exist by Definition 3.5. Let ν be the measure on $\mathbb{R}^d - \{0\}$ defined by

$$\nu(A) = \text{expected number of jumps per unit time with size in } A.$$

In general ν can be any non-negative measure on $\mathbb{R}^d - \{0\}$ satisfying

$$\int_{|y|>1} \nu(\mathrm{d}y) < \infty, \quad \text{and} \quad \int_{|y|\leq 1} |y|^2 \nu(\mathrm{d}y) < \infty. \tag{3.10}$$

The process X_t is said to have finite jump activity if $\nu(\mathbb{R}^d) < \infty$. If $\nu(\mathbb{R}^d) = \infty$ then X_t has infinite jump activity. Finite activity processes can be represented as the sum of a Brownian motion with drift and a compound Poisson process,

$$X_t = \sigma W_t + t\gamma + \sum_{k=1}^{N_t} Y_k. \tag{3.11}$$

Here σW_t is d-dimensional Brownian motion with covariance matrix σ^2, $\gamma \in \mathbb{R}^d$, N_t is a counting process with intensity $\nu(\mathbb{R})$ and Y_k's are i.i.d. with law $\nu/\nu(\mathbb{R})$. In general the jump intensity $\nu(\mathbb{R})$ could equal infinity and one needs to compensate for the effect of the small jumps,

$$X_t = \gamma t + \sigma W_t + \lim_{\epsilon \to 0} \left(\sum_{0 \le s \le t} \Delta X_s 1_{|\Delta X_s| > \epsilon} - t \int_{\epsilon < y \le 1} \nu(\mathrm{d}y) \right). \tag{3.12}$$

Here $\Delta X_s := X_s - X_{s-}$. For each $\epsilon > 0$, $\sum_{0 \le s < t} \Delta X_s 1_{|\Delta X_s| > \epsilon}$ is a compound Poisson process with intensity $\nu(\{y : |y| > \epsilon\})$, finite by (3.10), and jump size distribution given by

$$1_{|y|>\epsilon} \nu(\mathrm{d}y) / \nu(\{y : |y| > \epsilon\}).$$

The Lévy measure ν together with σ^2 and γ in (3.12) completely characterizes X_t. The triplet (σ^2, γ, ν) is referred to as the characteristic triplet of the Lévy process X_t.

A fundamental result is the explicit expression of the characteristic function in terms of the characteristic triple, the Lévy-Khinchin formula

$$\begin{aligned} &E\left[e^{iz \cdot X_t}\right] = e^{t\psi(z)}, z \in \mathbb{R}^d \text{ with} \\ &\psi(z) = -\frac{1}{2} z \sigma^2 z + i\gamma \cdot z + \int_{\mathbb{R}^d} \left(e^{iz \cdot y} - 1 - iz \cdot y 1_{|y| \le 1} \right) \nu(\mathrm{d}y). \end{aligned} \tag{3.13}$$

The most common role of Lévy processes in asset pricing is to let the underlying be given as the exponential of a Lévy process,

$$S_t = S_0 e^{X_t}.$$

The universe of all Lévy processes is too vast and, without further restrictions on the Lévy measure, the calibration is potentially very ill-posed. The standard approach when using Lévy processes in asset pricing is to first

chose a particular parameterization of the Lévy triple, then calibrate the parameters to observed market prices. Different parameterizations of the characteristic triple (σ^2, γ, ν) give rise to different Lévy models.

The zoo of Lévy models can broadly be divided into two sections, depending on the activity of the jump term: finite and infinite jump activity.

Probably the first model with jumps to be considered in asset pricing was the 'Merton model'. Merton, [36], considered a jump-diffusion model with log-normal sized jumps arriving with finite intensity. In this setting, the price of European contingency claims can be expressed as a series expansion in their Black & Scholes prices. The 'Kou model' is an example of a finite jump activity model with exponential decay in the jump size distribution, see [31].

Infinite jump activity models are typically 'pure jump', i.e. $\sigma^2 = 0$. Popular examples include the variance Gamma [33] and the tempered stable processes [29].

For a general background on Lévy processes, see the books [6] and [42]. See [15] for applications to finance.

4 Monte Carlo methods, diffusion approximation and Paper II

Let Q denote a pricing measure, see Section 3.4. Suppose that under Q, the price of the underlying is driven by a Lévy process X_t. The value of a European style contingency claim with payoff g and maturity T is given by

$$C(t,x) = E^Q \left[g(X_T) \,|\, X_t = x \right], \tag{4.1}$$

where we assume for simplicity that the risk-free interest rate is zero.

The Monte Carlo approach to estimating (4.1) is to generate N independent samples $x_1, \ldots, x_N$ of X_T and approximate

$$C(t,x) \approx \frac{1}{N} \sum_{k=1}^{N} g(x_k).$$

The difference,

$$\mathcal{E}_S = \frac{1}{N} \sum_{k=1}^{N} g(x_k) - E^Q \left[g(X_T) \,|\, X_t = x \right],$$

is called the *statistical error*. Provided that the variance of $g(X_T)$ is finite, it is a consequence of the Central Limit Theorem that the standard deviation of the statistical error decreases as $N^{-1/2}$,

$$\operatorname{std}(\mathcal{E}_S) \sim \frac{1}{N^{1/2}}. \tag{4.2}$$

If the characteristic triple (see Section 3.5) of X_t is known and $\nu(\mathbb{R}^d) < \infty$, then X_T can be sampled exactly using Algorithm 4.1.

Algorithm 4.1 Sampling a finite jump activity Lévy process

Input: Characteristic triple (σ^2, γ, ν), time T, initial value X_0.
Output: Sample of final value X_T.

Let B_k, $k = 1, \dots d$ be d independent samples from a $N(0,1)$-distribution. Let

$$B = (B_1, \dots, B_d).$$

Sample the number of jumps N_T from an exponential distribution with intensity $T\nu(\mathbb{R}^d)$.
Sample N_T i.i.d. jump sizes $Y_1, \dots, Y_{N_T}$ from a distribution with law

$$\nu(\mathrm{d}y)/\nu\left(\mathbb{R}^d\right).$$

Let X_T be given by

$$X_T = \sqrt{T}B + T\gamma + \sum_{k=1}^{N_T} Y_k.$$

Note that the expected computational work in simulating each value of X_T following Algorithm 4.1 grows like the expected number of jumps $\nu(\mathbb{R}^d)$. If $\nu(\mathbb{R}^d) = \infty$, the algorithm is impossible to apply. A standard technique to deal with infinite jumps measures, dating back at least to the work of Asumussen and Rosińksi in [3], is to introduce a type of viscosity approximation of X_t where the small jumps are replaced by a suitably chosen diffusion term.

The idea of the diffusion approximation is to replace all jumps smaller than some $\epsilon > 0$ by diffusion. The main ambition in Paper II is to understand how to chose ϵ. A large value results in fast simulations since the

number of jumps decreases with ϵ. However, a large value of ϵ also results in a large model error, defined as the difference

$$\mathcal{E}_M = E[g(\overline{X}_T)] - E[g(X_T)],$$

where $\overline{X}_T$ denotes the approximate process. The main result of Paper II is a computable estimate of the model error, see Theorem 4.5 below.

Note that the entire approximation error when using Algorithm 4.1 and $\overline{X}_t$ to compute $E[X_T]$ equals the sum of the statistical error and the model error,

$$\begin{aligned}\mathcal{E} &= \frac{1}{N}\sum_{k=1}^{N} g(x_k) - E^Q\left[g(X_T)\right] \\ &= \left(\frac{1}{N}\sum_{k=1}^{N} g(x_k) - E^Q\left[g(\overline{X}_T)\right]\right) + \left(E^Q[g(\overline{X}_T)] - E^Q[g(X_T)]\right) \\ &= \mathcal{E}_S + \mathcal{E}_M,\end{aligned}$$

where $x_1, \ldots, x_N$ are N independent samples of $\overline{X}_T$. The statistical error is determined by (4.2), so it remains to estimate the model error. But first a more detailed description of how to construct the diffusion approximation $\overline{X}_t$.

4.1 A diffusion approximation of small jumps

Recall that X_t denotes a d-dimensional Lévy process with characteristic triple (σ^2, γ, ν). Let R_t^ϵ denote the compensated sum of jumps not exceeding ϵ in absolute value. The process $X_t - R_t^\epsilon$ can be simulated with Algorithm 4.1 since its jump measure is finite. A higher order weak approximation is constructed by adding a diffusion term approximating R_t^ϵ. For $\epsilon > 0$ define the Lévy measure ν^ϵ, the number γ^ϵ and the covariance matrix $\sigma^2(\epsilon)$ by

$$\nu^\epsilon(x) = 1_{|x|>\epsilon}\nu(x) = \begin{cases} \nu(x) & |x| > \epsilon \\ 0 & |x| < \epsilon. \end{cases} \tag{4.3}$$

$$\gamma^\epsilon = \begin{cases} \gamma & \text{if } \epsilon \leq 1 \\ \gamma + \int_{1<|x|<\epsilon} x\nu(\mathrm{d}x) & \text{if } \epsilon > 1 \end{cases} \tag{4.4}$$

and

$$\sigma_{ij}^2(\epsilon) = \int_{|x|<\epsilon} x_i x_j \nu(\mathrm{d}x). \tag{4.5}$$

Let $\overline{X}_t$ denote the Lévy process with characteristic triple $(\sigma^2+\sigma^2(\epsilon), \gamma^\epsilon, \nu^\epsilon)$. The process $\overline{X}_t$ is referred to as the diffusion approximation of X_t. Note that $\overline{X}_t$ has finite jump activity, the new drift γ^ϵ ensures that $E[X_t] = E[\overline{X}_t]$, and the matrix $\sigma^2(\epsilon)$ is defined so that the variance of $\overline{X}_t$ equals that of X_t.

Remark 4.1 *The precise choices of γ^ϵ and ν^ϵ vary from situation to situation. For instance, in Paper III, γ^ϵ is chosen so that $e^{\overline{X}_t}$ is a martingale.*

4.2 Estimation of the model error

In Paper II the focus is on pure jump Lévy processes. More specifically it is assumed that the driving Lévy process X_t satisfies the following condition:

Condition 4.1 *A Lévy process X_t is said to satisfy* Condition 1 with order *β if,*

$$\int_{|x|<\epsilon} |x|^2 \nu(\dot{x}) = \mathcal{O}(\epsilon^\beta) \quad as \quad \epsilon \to 0,$$

the density $p_t(x)$ of X_t is $C^\infty(\mathbb{R}^d)$ and

$$\partial_\alpha p_t(x) \to 0 \quad as \quad |x| \to 0,$$

for all $|\alpha| \geq 0$.

To guarantee regularity of the value function $u(t,x)$ with respect to x, we need to make certain integrability assumptions on the payoff $g(x)$:

Definition 4.2 *Let X_t be a Lévy process satisfying Condition 4.1 with probability density $p_t(x)$. The payoff $g : \mathbb{R}^d \to \mathbb{R}$ is said to be integrable to the n'th order if g can be written as finite sum $g = \sum_{k=1}^N g_k$, and there are $\lambda_1, \ldots, \lambda_N \in \mathbb{R}^d$, such that*

$$e^{-\lambda_k \cdot x} g_k \in L^1 \cap L^2 \text{ and } e^{\lambda_k \cdot x} \partial_\alpha p_t \in L^1 \cap L^2 \text{ for } k = 1, \ldots N \text{ and } |\alpha| \leq n.$$

Remark 4.3 *The requirement that a contract is integrable is not very restrictive. For instance, the call option, the put option, the digital option, and the basket option, are all examples of contracts that have integrable payoff functions, of any order, when the underlying is a Lévy process with 'stable-like' behavior close to zero,*

$$\nu(\dot{y}) \sim \frac{1}{|y|^{1+\alpha}}, \text{ for } \alpha > 0.$$

Integrability of the payoff guarantees that the value function is smooth:

Remark 4.4 *The value function $u(t,x)$ can be written as a convolution of the payoff $g(x)$ with the density p_{T-t} of X_T, conditioned on $X_t = x$:*

$$u(t,x) = \int_{\mathbb{R}^d} g(y) p_{T-t}(x-y) \dot{y}.$$

Hence if g is integrable to the n'th order, then $u(t,x)$ is C^α in x, for all $|\alpha| < n$, independent of the regularity of the payoff $g(x)$.

The following theorem estimates the difference between X_t and $\overline{X}_t$.

Theorem 4.5 *Let X_t be a d–dimensional Lévy process with characteristic triple (ν, σ^2, γ) satisfying Condition 1 with order β, and let $\overline{X}_t$ be its diffusion approximation defined above for some $\epsilon > 0$. Let $g(x)$ be a real valued function that is integrable to the 6'th order. The model error*

$$\mathcal{E}_M = E[g(X_T)] - E[g(\overline{X}_T)]$$

is then

$$\mathcal{E}_M = \mathcal{O}(\epsilon^{1+\beta}) \quad as \quad \epsilon \to 0.$$

Moreover, the model error can be expressed as

$$\mathcal{E}_M = \frac{T f(\epsilon)_{ijk}}{6} E[g^{(3)}_{ijk}(\overline{X}_T)] + \mathcal{O}(\epsilon^{2+\beta}), \tag{4.6}$$

where

$$f(\epsilon)_{ijk} = \int_{|y|<\epsilon} y_i y_j y_k \nu(\dot{y}).$$

Proof: See Theorem 4.1 in Paper II.

The estimate (4.6) is an example of an a posteriori error estimate, i.e. information about the approximate solution $\overline{X}_T$ is used to determine the size of the error.

Example 4.6 *Consider a contract with maturity $T = 1$ and payoff*

$$g(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0. \end{cases}$$

Suppose that the underlying is a pure jump Lévy process with characteristic triple $(0,0,\nu)$, where $\nu(dy) = 1_{0<y<1} 1/y^2 dy$. This process has infinite jump activity.

Let $\overline{X}_t$ denote the diffusion approximation of X_t, defined for some $\epsilon > 0$. Let $\mathcal{E}$ denote the approximation error in a Monte Carlo calculation of $E[g(X_T)]$, using N independent samples of $\overline{X}_t$. From (4.2) *and Theorem 4.5 it follows that there are constants C_1 and C_2 such that, with a high probability,*

$$|\mathcal{E}| \leq C_1 \frac{1}{N^{1/2}} + C_2 \epsilon^2. \tag{4.7}$$

The computational work of simulating N samples of $\overline{X}_1$ is proportional to the expected number of jumps

$$\textit{Work} \sim N\nu(|y| > \epsilon) \sim N\frac{1}{\epsilon}. \tag{4.8}$$

An optimal choice of (N, ϵ) would minimize (4.8) *while keeping the right hand side of* (4.7) *constant. It follows that ϵ should decrease when N increases as*

$$\epsilon \sim \frac{1}{N^{1/4}}.$$

4.3 A simple adaptive scheme

Recall the situation of the previous section: X_t denotes some Lévy process with infinite jump activity and $g = g(x)$ is the payoff of some European style contingency claim with maturity T. The idea of the adaptive scheme is to identify a region where the payoff g is more sensitive to approximations, this could for instance be near the strike of a call option. One first simulates a coarse approximation of X_t, corresponding to a large ϵ. If the coarse approximation ends up in the sensitive region, the approximation is refined. The details are as follows:

Fix $\epsilon_2 > \epsilon_1 > 0$ and some region $L \subset \mathbb{R}^d$. We call L the critical region, since it should correspond to the area where the contract is most sensitive to approximations. Recall the definition of $\sigma^2(\epsilon)$ in (4.5). Let $\overline{X}_t^{(1)}$ and $\overline{X}_t^{(2)}$ respectively be the viscosity approximations given by ϵ_1 and ϵ_2. Let $X_t^{\epsilon_1}$ and $X_t^{\epsilon_2}$ denote the compensated sum of jumps exceeding ϵ_1 and ϵ_2 respectively. Note that in absolute value,

$$\begin{aligned}
\overline{X}_t^{(1)} &= X_t^{\epsilon_1} + \sigma(\epsilon_1)W_t, \text{ and} \\
\overline{X}_t^{(2)} &= X_t^{\epsilon_2} + \sigma(\epsilon_2)W_t = X_t^{\epsilon_1} + \Delta_t + \sigma(\epsilon_2)W_t,
\end{aligned}$$

where $\Delta_t := X_t^{\epsilon_2} - X_t^{\epsilon_1}$. Note that the three terms in the last expression of $\overline{X}_t^{\epsilon_2}$ are independent. Let the control parameter $\alpha \in \{1, 2\}$ be given by

$$\alpha = \begin{cases} 1 & \text{if } X_T^{\epsilon_1} \notin L \\ 2 & \text{if } X_T^{\epsilon_1} \in L. \end{cases}$$

Finally the *adaptive approximation* is defined by

$$\begin{aligned} \overline{X}_t^{(\alpha)} &= 1_{\alpha=1}\overline{X}_t^{(1)} + 1_{\alpha=2}\overline{X}_t^{(2)} \\ &= X_t^{\epsilon_1} + 1_{\alpha=1}\sigma(\epsilon_1)W_t + 1_{\alpha=2}\left(\Delta_t + \sigma(\epsilon_2)W_t\right). \end{aligned}$$

The model error when using the adaptive approximation to estimate (4.1),

$$\mathcal{E}_A = E[g(X_T)] - E[g(\overline{X}_T^{(\alpha)})],$$

is estimated in the following theorem.

Theorem 4.7 *Let X_t be a Lévy process satisfying Condition 1 with order β and let $\overline{X}_t^{(\alpha)}$ be its adaptive approximation, defined for a particular choice of ϵ_1, ϵ_2 and L. Fix some function g that is integrable to the 6th order. The model error*

$$\mathcal{E}_M = E[g(X_T)] - E[g(\overline{X}_T^{(\alpha)})] \tag{4.9}$$

can be expressed as

$$\mathcal{E}_M = \frac{T}{6}\sum_{n=1}^{2} f_{ijk}(\epsilon_n)E\Big[1_{\alpha=n}g_{ijk}^{(3)}(\overline{X}_T^{(n)})\Big] + \mathcal{O}(\epsilon_n^{2+\beta})E[1_{\alpha=n}].$$

Proof: See Theorem 5.2 in Paper II.

Section 5 in Paper II is concluded with an illustration on the work reduction that can be obtained with the adaptive approximation. The computational work for a given level of tolerance for both the adaptive and the diffusion approximation are compared. In this particular example, the contract is a digital option, and the parameters $(\epsilon_1, \epsilon_2, L)$ of the adaptive approximation are chosen in a simple fashion. It appears (see Figure 4.3 below) that there is a substantial amount to gain from using even such simple adaptive schemes.

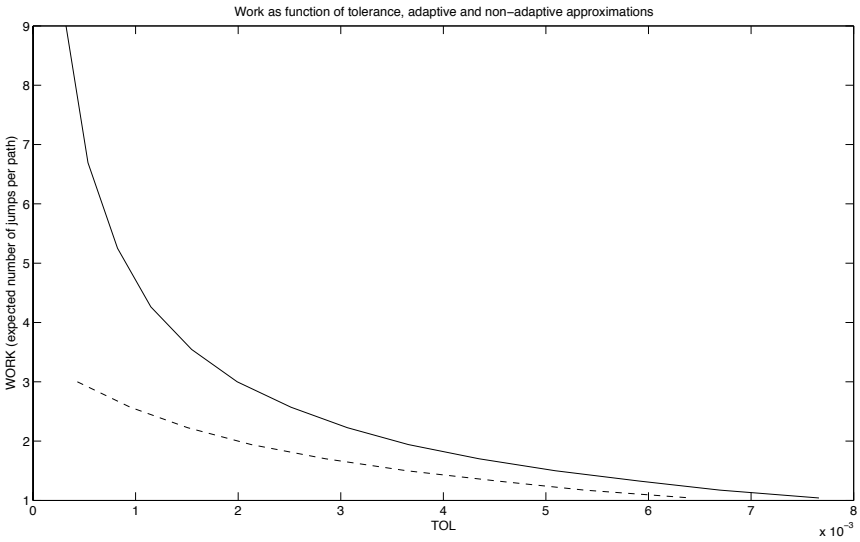

Figure 1.6: Here the work, defined as the expected number of jumps per path, is plotted as a function of the tolerance for the adaptive and non-adaptive approximations. The solid line represents the non-adaptive approximation. The dashed line represents the adaptive approximation. For details on the particular choice of processes X_t, the contract and details on the choice of adaptive approximation, see Example 5.6 in Paper II.

5 Integro-differential equations, computable error estimates and Paper III

In Paper III a finite difference scheme for option pricing in exponential Lévy models is discussed. In particular computable estimates of space and time discretization and diffusion approximation errors are derived under weak assumptions of the payoff.

Let X_t denote a Lévy process with infinitesimal generator $\mathcal{L} = \mathcal{L}^{X_t}$, see 3.13, and let $g : \mathbb{R} \to \mathbb{R}$ be a measurable function satisfying certain exponential growth conditions, see (5.3) below. In Paper III we derive error estimates of a finite difference scheme suitable for constructing weak

solutions to initial value equations on the form

$$\begin{aligned} f'_t(t,x) &= \mathcal{L}f(t,x), \text{ for } (t,x) \in (0,T] \times \mathbb{R}, \\ f(0,x) &= g(x). \end{aligned} \tag{5.1}$$

Here $f'_t(t,x)$ denotes the partial derivative of f with respect to t evaluated at the point (t,x).

Such equations arise in asset pricing when the underlying is an exponential Lévy process. Indeed, let $\Pi = \Pi(t, S_t)$ denote the price of a European style contingency claim with maturity T and payoff G. Recall from Section 3.4 that under certain assumptions there is a probability measure Q such that Π equals the future discounted expected payoff,

$$\Pi(t, S_t) = e^{-r(T-t)} E^Q[G(S_T) \,|\, S_t].$$

Suppose that under Q, S_t is an exponential Lévy process $S_t = S_0 e^{X_t}$, where X_t is a Lévy process with infinitesimal generator $\mathcal{L} = \mathcal{L}^{X_t}$, see 3.13. Let τ denote time to maturity, $\tau = T - t$, and x the 'log-moneyness', $x = \log(S/S_0)$. Introducing the functions $f(\tau, x) = e^{r\tau}\Pi(t, S_0 e^x)$ and $g(x) = G(S_0 e^x)$ leads to

$$f(\tau, x) = E[g(x + X_\tau)]. \tag{5.2}$$

Under certain regularity assumptions, for instance if $f \in C^{1,2}$ and has bounded derivatives, then $f = f(\tau, x)$ is a classical solution to (5.1).

In financial applications, the contract g is typically not C^2. Some contracts, such as the binary, are not even continuous. As a consequence, f will usually not be a classical solution to (5.1). In [17], Cont and Voltchkova show that, if g is Lipschitz then f is a viscosity solution to (5.1). In [16], the same authors suggest a finite difference scheme suitable for solving (5.1). In Paper III, the conditions imposed by Cont and Voltchkova are relaxed, and g is only assumed to be measurable and such that

$$e^{\lambda_1 x} g_- + e^{-\lambda_2 x} g_+ \in \mathrm{L}^1 \cap \mathrm{L}^2, \tag{5.3}$$

for some $\lambda_1, \lambda_2 \geq 0$, where $\mathrm{L}^k = \{\int_{\mathbb{R}} |f|^k \mathrm{d}x < \infty\}$, $g_- = g1_{x<0}$ and $g_+ = g1_{x>0}$. It is interesting to study this wider class of contracts for several reasons. First of all, as mentioned above, some contracts have a payoff that is not Lipschtiz. Second, the Greeks of a contract can be determined with equations similar to (5.1). However, a Greek is typically even less regular than the original contract.

In Paper III we study weak solutions to (5.1). A weak solution is defined in terms of observable $\psi : \mathbb{R} \to \mathbb{R}$ satisfying certain regularity conditions, see Theorem 5.1 below. If $f = f(t,x)$ is a weak solution to (5.1), then the expression $\int_{\mathbb{R}} \psi(x) f(T,x) \mathrm{d}x$ is well defined and equals

$$\int_{\mathbb{R}} \psi(x) f(T,x) \mathrm{d}x = \int_{\mathbb{R}} \varphi(0,x) g(x) \mathrm{d}x,$$

where φ is the solution to the backward equation

$$\begin{aligned} \varphi_t'(t,x) + \mathcal{L}^* \varphi(t,x) = 0, \text{ for } (t,x) \in [0,T) \times \mathbb{R}, \text{ and} \\ \varphi(T,x) = \psi(x). \end{aligned}$$

Recall from Section 3.4 that S_t is a martingale under the pricing measure Q. For an exponential Lévy process $S_t = S_0 e^{X_t}$, where the Lévy process X_t has characteristic triple (σ^2, γ, ν), the martingale condition (3.9) is equivalent to

$$\begin{aligned} &\int_{|y|>1} e^y \nu(\mathrm{d}y) < \infty, \\ &\gamma = r - \frac{1}{2}\sigma^2 - \int_{\mathbb{R}} \left(e^y - 1 - y 1_{|y| \leq 1} \right) \nu(\mathrm{d}y). \end{aligned}$$

Consequently, the infinitesimal generator $\mathcal{L}$ in (5.1) takes the form

$$\begin{aligned} \mathcal{L} f(x) = \left(r - \frac{\sigma^2}{2} \right) f'(x) + \frac{\sigma^2}{2} f''(x) \\ + \int_{\mathbb{R}} \left(f(x+y) - f(x) - (e^y - 1) f'(x) \right) \nu(\mathrm{d}y). \end{aligned}$$

The operator $\mathcal{L}$ is typically non-local, the Lévy measure can even have support on the whole real line. As a consequence, solving (5.1) with an implicit scheme is potentially computationally very costly. An explicit scheme is not suitable for stability reasons. The scheme suggested in [16] is an explicit-implicit scheme, that treats the integral term explicitly and the differential term explicitly. The details are as follows:

Assume at first that $\nu(\mathbb{R}) < \infty$. The operator $\mathcal{L}$ can be divided into a sum of two terms, $\mathcal{L} = \mathcal{D} + \mathcal{J}$, where

$$\begin{aligned} \mathcal{D} f(x) &= \left(r - \frac{\sigma^2}{2} - \int_{\mathbb{R}} (e^y - 1) \nu(\mathrm{d}y) \right) f'(x) + \frac{\sigma^2}{2} f''(x) - f(x) \int_{\mathbb{R}} \nu(\mathrm{d}y), \\ \mathcal{J} f(x) &= \int_{\mathbb{R}} f(x+y) \nu(\mathrm{d}y). \end{aligned}$$

To obtain a finite dimensional problem, the operator $\mathcal{L}$ and the measure ν need to be truncated. Fix $B, K > 0$, let $\mathcal{L}_B$ equal $\mathcal{L}$ inside the compact set $[-B, B]$ and zero outside, and let $\nu_K(\mathrm{d}y) = 1_{|y|<K}\nu(\mathrm{d}y)$. Fix two integers M, N and introduce a finite grid on $[0, T] \times [-B, B]$: let $t_n = n\Delta t$, $x_m = m\Delta x$ for $n = 0, \ldots, N$, $m = -M, \ldots M$, where $\Delta t = T/N$ and $\Delta x = K/M$. Let f_m^n denote the approximate solution $f_m^n \approx f_B(n\Delta t, m\Delta x)$, for $0 \le n \le N$ and $-M \le m \le M$. The explicit-implicit scheme used to solve (5.1) involves the two matrices $\mathcal{D}_\Delta$ and $\mathcal{J}_\Delta$ given by

$$(\mathcal{D}_\Delta f^n)_m = \left(r - \frac{\sigma^2}{2} - \sum_k \left(e^{k\Delta x} - 1\right)\nu_K(k\Delta x)\Delta x\right) Df_m^n \tag{5.4}$$
$$+ \frac{\sigma^2}{2} D^2 f_m^n + f_m^n \sum_k \nu_K(k\Delta - x)\Delta x,$$
$$(\mathcal{J}_\Delta f^n)_m = \sum_k f_{m+k}^n \nu_K(k\Delta x)\Delta x, \text{ where } f_{m+k}^n = 0 \text{ if } |m+k| > M. \tag{5.5}$$

Where D and D^2 are standard finite difference approximations of the first and second derivatives respectively, namely upwind for the first derivative and centered difference quotient for the second derivative.

If X_t has infinite jump activity, i.e. if $\nu(\mathbb{R}) = \infty$, then X_t is replaced by its diffusion approximation, as described in Section 4.1, defined for some $\epsilon > 0$. The explicit-impliticit time stepping method amounts to solving

$$f_m^0 = g(m\Delta x), \text{ and} \tag{5.6}$$
$$f_m^{n+1} = f_m^n + \Delta t(\mathcal{D}_\Delta f^{n+1})_m + \Delta t(\mathcal{J}_\Delta f^n)_m,$$

for $n = 1, \ldots, N$ and $m = -M, \ldots, M$.

5.1 Error estimates and model parameter optimization

Fix an observable ψ. Using Algorithm 5.6 to solve (5.1) results in an approximation error, defined as the difference

$$\mathcal{E} = \sum_k \psi(k\Delta x) f_k^N \Delta x - \int_{\mathbb{R}} \psi(x) f(T, x)\mathrm{d}x.$$

The approximation error can be split into a sum of different error terms: space and time discretization errors, measure truncation error and domain localization error. In addition, if the process has infinite jump activity, there

is a diffusion approximation error. The measure truncation and domain localization errors will, under suitable conditions, decrease exponentially with respect to the truncation parameters, see [40]. In Paper III computable estimates of the time and space discretization errors are determined. The diffusion approximation error was treated in Paper II. The space discretization error, $\mathcal{E}_\Delta$, is by definition the difference

$$\mathcal{E}_\Delta = \langle \psi, f_\Delta(T,\cdot)\rangle - \langle \psi, f(T,\cdot)\rangle\,. \tag{5.7}$$

and the time discretization error, $\mathcal{E}_T$, is the difference

$$\mathcal{E}_T = \langle \psi, f^N\rangle - \langle \psi, f(T,\cdot)\rangle. \tag{5.8}$$

Here $\langle\cdot,\cdot\rangle$ denotes the usual L^2 inner product of functions. The following theorem estimates the space discretization error:

Theorem 5.1 *Assume that there are constants $\lambda_1, \lambda_2 \geq 0$ such that the contract g satisfies*

$$1_{x>0}e^{-\lambda_1 x}g + 1_{x<0}e^{\lambda_2 x}g \in L^2 \cap L^1.$$

Suppose further that the observable ψ is such that $\left(e^{\lambda_1 x} + e^{-\lambda_2 x}\right)\psi^{(n)} \in L^2 \cap L^1 \cap \mathcal{C}$ for $n = 0,\ldots,6$. If the Lévy measure ν satisfies

$$\int_{\mathbb{R}} \left(e^{\lambda_1 x} + e^{-\lambda_2 x}\right)\nu(dx) < \infty,$$

then there is a constant C_1 independent of Δx such that the space discretization error (5.7) *verifies*

$$|\mathcal{E}_\Delta| \leq C_1 \Delta x.$$

Moreover, there is a constant C_2, independent of Δx, such that the leading order term of (5.7) *can be expressed in computable form*

$$\mathcal{E}_\Delta = \frac{\Delta x}{2}\left(r - \frac{\sigma^2}{2} - \int_{\mathbb{R}}(e^y - 1)\,\nu(dy)\right)\int_{\mathbb{R}} \psi''(x) f_\Delta(T,x)\,dx + R, \text{ where} \tag{5.9}$$

$$|R| \leq C_2 \Delta x^2.$$

In particular, if the Lévy measure ν is an approximation of an infinite activity measure for some diffusion approximation parameter $\epsilon < \epsilon_0$, i.e. $\nu = \tilde{\nu}^\epsilon$, where $\tilde{\nu}(\mathbb{R}) = \infty$. Then the constant C_2 in (5.9) *can be chosen independently of ϵ.*

Proof: See Theorem 5.4 in Paper III.

The time discretization error is estimated in the following theorem:

Theorem 5.2 *Assume that there are constants* $\lambda_1, \lambda_2 \geq 0$ *such that the contract* g *satisfies*

$$1_{x>0} e^{-\lambda_1 x} g + 1_{x<0} e^{\lambda_2 x} g \in L^2 \cap L^1.$$

Suppose further that the observable ψ *is such that* $\left(e^{\lambda_1 x} + e^{-\lambda_2 x}\right) (\mathcal{L}^n)^* \psi \in L^2 \cap L^1 \cap \mathcal{C}$ *for* $n = 0, \ldots, 3$. *If the Lévy measure* ν *satisfies*

$$\int_{\mathbb{R}} \left(e^{\lambda_1 x} + e^{-\lambda_2 x}\right) \nu(dx) < \infty$$

and Δt *the constraint*

$$\Delta t \left(\lambda_k \left(\frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^y - 1)\, \nu(dy) - r \right) + \frac{\sigma^2}{2} \lambda_k^2 \right) < 1/2,$$

for $k = 1, 2$, *then there is a constant* C, *independent on* Δt, *such that the time discretization error* (5.8) *satisfies*

$$|\mathcal{E}_T| \leq C \Delta t. \tag{5.10}$$

If the Lévy measure is an approximation of an infinite activity measure for some diffusion approximation parameter ϵ, *i.e.* $\nu = \tilde{\nu}^\epsilon$, *where* $\tilde{\nu}(\mathbb{R}) = \infty$, *and if* Δt *satisfies the relaxed condition*

$$\Delta t \left(\lambda_k \left(\frac{\sigma^2}{2} + \int_{|y|>\epsilon_0} (e^y - 1)\, \nu(dy) - r \right) + \frac{\sigma^2}{2} \lambda_k^2 \right) < 1/2, \text{ for } k = 1, 2,$$

for some ϵ_0, $\epsilon <\epsilon_0 < log\ (1 + 1/\lambda)$. *Then there is a constant* C *satisfying* (5.10), *which is independent of both* Δt *and* ϵ.

Finally, the leading order term of the time discretization error can be expressed in computable form

$$\mathcal{E}_T = \frac{T}{2} \int_{\mathbb{R}} \psi(x)(\mathcal{D} - \mathcal{J})(f^N(x) - f^{N-1}(x))dx + \mathcal{O}(\Delta t^2).$$

Proof: See Theorem 5.10 in Paper III.

Theorems 5.1 and 5.2 together with Theorem 4.5 are enough to determine how to choose the parameters Δx, Δt and ϵ, as illustrated by the following example.

Example 5.3 *Let X_t be a Lévy process with infinite jump activity. Suppose that the jump measure satisfies*

$$\nu(dy) \sim \frac{1}{y^{3-\beta}},$$

for some $0 < \beta < 1$. By Theorems 4.5, 5.1 and 5.2 there are constants C_1, C_2 and C_3, independent of Δx, Δt and ϵ, such that

$$|\mathcal{E}| \leq C_1 \left(1 + \int_{|y|\geq\epsilon} |e^y - 1|\, \nu(dy)\right) \Delta x + C_2 \Delta t + C_3 \epsilon^{\beta+1}. \tag{5.11}$$

The term $\int_{|y|\geq\epsilon} |e^y - 1|\, \nu(dy)$ can be estimated by

$$\int_{|y|\geq\epsilon} |e^y - 1|\, \nu(dy) \leq C \left(\epsilon^{\beta-1} + 1\right). \tag{5.12}$$

The computational complexity is proportional to $1/(\Delta x \Delta t)$ and independent of the choice of ϵ. Hence, ϵ should be chosen so that (5.11) is minimized. Furthermore, Δt is chosen so that the computational work is minimized for a given level of tolerance. These two conditions amount to letting ϵ and Δt decrease with Δx like

$$\begin{aligned} \epsilon &\sim \Delta x^{1/2}, \\ \Delta t &\sim \Delta x^{(1+\beta)/2}. \end{aligned} \tag{5.13}$$

The resulting computational work will increase as the tolerance, TOL, decreases as

$$\textit{Work} \sim \textit{TOL}^{-1-2/(\beta+1)}. \tag{5.14}$$

6 Optimal control, model calibration and Paper IV

Having spent some time on the forward problem, i.e. estimating the expected value (3.7), we now turn to the inverse problem of determining the pricing measure Q from observed option prices. Robust and efficient calibration is essential for the applicability of a model. The great success of the original Black & Scholes model is probably partly due to the fact that it is very

easy and fast to calibrate. The objective of the final paper of this thesis, Paper IV, is to describe how optimal control can be used to calibrate a jump-diffusion process.

Suppose that, under the pricing measure Q, the underlying S_t is an exponential jump-diffusion process,

$$\mathrm{d}S_t/S_{t-} = (r - \mu(t)m(t))\,\mathrm{d}t + \sigma(t, S_{t-})\mathrm{d}W_t + (J-1)\,\mathrm{d}\pi(t), \tag{6.1}$$

where the local volatility $\sigma = \sigma(t, S)$ is state and time dependent, the jump times are governed by a Poisson counting process $\pi(t)$ with time dependent intensity $\mu(t)$, and the relative jump size is a stochastic variable with density χ. The risk free interest rate is denoted by r and the drift term is determined by the fact that $e^{-rt}S_t$ is a martingale, forcing m to be

$$m = E[J-1].$$

In (6.1), S_{t-} denotes the left limit of S_t. Note that if σ and μ are constant then S_t is an exponential Lévy process with finite jump intensity.

A similar model, but with constant jump intensity, is calibrated by Andersen and Andreasen in [2]. Calibration of partial and stochastic differential equations is a huge field, with plenty of applications also outside finance. An accessible introduction to model calibration, with a focus on computations, is the book [46]. A more mathematical introduction to regularization techniques of inverse problems is the book [21].

Recall that the value of a call option with strike K and maturity T equals its future discounted expected payoff,

$$C(S_t, t; T, K) = e^{-r(T-t)}E^Q\left[\max(S_T - K, 0) \mid S_t\right]. \tag{6.2}$$

Let $\{C_m(T,K)\}_{T,K}$ be a collection of call option prices, for different strikes $K \geq 0$ and maturities $0 < T < \hat{T}$, observed in a market at time $t = 0$. Calibration of the pricing measure amounts to determining $\sigma(t, S)$, $\mu(t)$ and χ in such a way that the difference between the observed price C_m and the theoretical price (6.2) is minimized. There are different measures of the difference between observed and theoretical prices, in Paper IV we study the L^2-distance

$$\int_0^{\hat{T}} \int_{\mathbb{R}_+} \left(C(S_0, 0; T, K) - C_m(T, K)\right)^2 \mathrm{d}K\mathrm{d}T.$$

Another popular measurement is the L^2-distance of implied volatility, see for instance [2].

Similarly to the exponential Lévy setting described in the previous section, the theoretical option price satisfies, in the variables t and S, the partial integro-differential equation

$$rC = C'_t - \mu(t)\Big(C + mSC'_S + E\left[C(t,S)\right]\Big) + \frac{1}{2}\sigma^2(t,S)S^2C''_{SS} + rSC'_S,$$
$$C(T,S) = \max(S-K,0), \tag{6.3}$$

where

$$\mathrm{E}[C(t,JS)] = \int_{\mathbb{R}_+} C(t,Sx)\chi(x)\mathrm{d}x.$$

Here C'_t, C'_S, etc. denote the derivative of C with respect to t, S, etc.

In (6.3) the strike K and maturity T are fixed. Hence to price a family of options on the same underlying, one needs to solve (6.3) once for each different value of (T,K). Following Dupire, [20], one can derive a forward equation in T and K that can be used to price a whole family of call options on the same underlying. A similar result is obtained in [2].

Proposition 6.1 *Suppose that the risk neutral dynamics of a stock is given by* (6.1). *The value of European call options $C(T,K) = C(0,S;K,T)$, for $t = 0$ and $S_0 = S$ fixed, solve the forward equation*

$$C'_T = \mu(T)\Big(mKC'_K - (m+1)C + E\big[J(T)C(T,J(T)^{-1}K)\big]\Big) + \frac{1}{2}\sigma^2(T,K)K^2C''_{KK} - rKC'_K, \tag{6.4}$$

where

$$E\left[J(T)C(T,J(T)^{-1}K)\right] = \int_{\mathbb{R}_+} zC(T,K/z)\chi(z)dz$$
$$C(0,K) = max(S-K,0).$$

Proof: See Proposition 1 in Paper IV.

For later use we introduce the two operators ψ_1 and ψ_2:

$$\psi_1(C) := (m+1)C - mKC'_K + E[JC(T,J^{-1}K)], \tag{6.5}$$
$$\psi_2(C) := \frac{1}{2}K^2C''_{KK}.$$

6.1 Optimal control and model calibration

The idea in Paper IV is to use optimal control to determine the local volatility and the jump intensity $(\sigma^2(t,S),\mu(t))$. The jump size density χ needs to be treated separately, so assume for now that χ is known.

Suppose that the price of call options with maturity T, $0<T<\hat{T}$ and strike $K>0$, are priced in a market with prices given by $C_m = C_m(T,K)$. The corresponding theoretical price $C(T,K) = C(0,S_0;T,K)$ is for fixed (σ^2,μ) given by (6.2). For the sake of readability we assume that the interest rate $r=0$.

The optimal control formulation amounts to determining

$$\Big(\sigma^2(t,S),\mu(t)\Big) \in \arg\min_{\mathcal{B}} \int_0^{\hat{T}} \int_{\mathbb{R}_+} \left(C(T,K) - C_m(T,K)\right)^2 \mathrm{d}K\mathrm{d}T, \quad (6.6)$$

where

$$\begin{aligned} C_T &= \psi_1(C) + \psi_2(C), \\ C(0,K) &= \max(S_0 - K, 0). \end{aligned}$$

It is assumed that (σ^2,μ) takes values in the compact set

$$\mathcal{B} = \Big\{(\sigma^2,\mu) : (\sigma^2(t,S),\mu(t)) \in [\sigma_-^2,\sigma_+^2]\times[\mu_-,\mu_+]\Big\}.$$

There are several different methods for solving optimal control problems. The method suggested in Paper IV was introduced by Sandberg and Szepessy in [41] and is based on a symplectic approximation of the Hamiltonian system for the control problem. In [12] the method is applied to reconstruct data for the heat and wave equations.

The Hamiltonian system associated to (6.6) is

$$\begin{aligned} C_T' &= H_\lambda'(C,\lambda), \\ \lambda_T' &= -H_C'(C,\lambda). \end{aligned} \quad (6.7)$$

One major reason why the method of Sandberg and Szepessy is suitable in the setting of Paper IV is that the Hamiltonian can be determined explicitly. In Paper IV it is shown that the Hamiltonian equals

$$\begin{aligned} H(C,\lambda) = s_{[\mu_-,\mu_+]}\Big(\int_{\mathbb{R}_+} \lambda\psi_1(C)\mathrm{d}K\Big) \\ + \int_{\mathbb{R}_+} s_{[\sigma_-^2,\sigma_+^2]}\Big(\lambda\psi_2(C)\Big)\mathrm{d}K + \int_{\mathbb{R}_+} (C-C_m)^2\mathrm{d}K, \end{aligned} \quad (6.8)$$

where $s_{[a,b]}(x)$ is the function

$$s_{[a,b]}(x) = \begin{cases} ax & \text{if } x < 0 \\ bx & \text{if } x > 0. \end{cases}$$

Note that the Hamiltonian is not differentiable, however, it is straightforward to construct an explicit regularization of H, see Section 1.4.2 in Paper IV.

Discretizing the regularized Hamiltonian system using an implicit symplectic scheme results in a coupled system of equations, see Section 1.4.2 in Paper IV. Since the regularized Hamiltonian is known explicitly, the discretized system of equations can be solved efficiently with the Newton method.

Paper IV is concluded with some numerical experiments, both on artificial, as well as real, data from call options on the S&P-500 Index.

Bibliography

[1] Aurélien Alfonsi, Antje Fruth, and Alexander Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10(2):143–157, 2010.

[2] Leif Andersen and Jesper Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(4):231–262, 2000.

[3] Søren Asmussen and Jan Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.*, 38(2):482–493, 2001.

[4] Marco Avellaneda and Sasha Stoikov. High-frequency trading in a limit order book. *Quant. Finance*, 8(3):217–224, 2008.

[5] Davis Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9:69–107, 1996.

[6] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[7] N. H. Bingham and R. Kiesel. *Risk-neutral valuation*. Springer Finance. Springer-Verlag London Ltd., London, second edition, 2004. Pricing and hedging of financial derivatives.

[8] Tomas Björk. *Arbitrage theory in continuous time*. Oxford University Press, Oxford, 1999.

[9] Fischer Black and Piotr Karasinksi. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 47(4):52–59, Jul.-Aug. 1991.

[10] Fischer Black and Myron S Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, May-June 1973.

[11] Jean-Philippe Bouchaud, Marc Mézard, and Marc Potters. Statistical properties of stock order books: empirical results and models. *Quant. Finance*, 2(4):251–256, 2002.

[12] Jesper Carlsson. Symplectic reconstruction of data for heat and wave equations. http://arxiv.org/abs/0809.3621, 2008.

[13] Peter Carr and B. Dilip Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61–73, 1998.

[14] Rama Cont, Sasha Stoikov, and Rishi Talreja. A stochastic model for order book dynamics. *SSRN eLibrary*, 2008. http://ssrn.com/paper=1273160.

[15] Rama Cont and Peter Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[16] Rama Cont and Ekaterina Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, 43(4):1596–1626 (electronic), 2005.

[17] Rama Cont and Ekaterina Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance Stoch.*, 9(3):299–325, 2005.

[18] Marcus G. Daniels, J. Doyne Farmer, László Gillemot, Giulia Iori, and Eric Smith. Quantitative model of price diffusion and market friction based on trading as a mechanistic random process. *Phys. Rev. Lett.*, 90(10):108102, Mar 2003.

[19] Hans Michael Dietz and Volker Nollau. *Markov decision problems with countable state spaces*, volume 15 of *Mathematical Research*. Akademie-Verlag, Berlin, 1983.

[20] Bruno Dupire. Pricing and hedging with smiles. In *Mathematics of derivative securities (Cambridge, 1995)*, volume 15 of *Publ. Newton Inst.*, pages 103–111. Cambridge Univ. Press, Cambridge, 1997.

[21] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[22] J. Doyne Farmer, László Gillemot, Fabrizio Lillo, Szabolcs Mike, and Anindya Sen. What really causes large price changes? *Quant. Finance*, 4(4):383–397, 2004.

[23] J. Doyne Farmer, Paolo Patelli, and Ilija Zovko. The predictive power of zero intelligence models in Þnancial markets. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6):2254–2259, 2005.

[24] Eugene A. Feinberg and Adam Shwartz, editors. *Handbook of Markov decision processes*. International Series in Operations Research & Management Science, 40. Kluwer Academic Publishers, Boston, MA, 2002.

[25] Mark D. Flood. An introduction to complete markets. *Review*, pages 32–57, March 1991. http://ideas.repec.org/a/fip/fedlrv/y1991imarp32-57.html.

[26] Thierry Foucault. Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, 2(2):99 – 134, 1999.

[27] Dhananjay K. Gode and Shyam Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101(1):119, 1993.

[28] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[29] Ismo Koponen. Analytic approach to the problem of convergence of truncated Lévy flights towards the gaussian stochastic process. *Phys. Rev. E*, 52(1):1197–1199, Jul 1995.

[30] Steven Kou. Lévy processes in asset pricing. 2008. www.columbia.edu/ sk75/levy.pdf.

[31] Steven G. Kou. A Jump Diffusion Model For Option Pricing. *SSRN eLibrary*, 2001.

[32] Hayne E. Leland. Option pricing and replication with transactions costs. *The Journal of Finance*, 40(5):1283–1301, 1985.

[33] Dilip B. Madan, Peter P. Carr, and Eric C. Chang. The Variance Gamma Process and Option Pricing. *European Finance Review*, 2(1):79–105, 1998.

[34] Benoit Mandelbrot. The variation of some other speculative prices. *The Journal of Business*, 40(4):393–413, 1967.

[35] Robert Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[36] Robert Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[37] James R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

[38] Anna A. Obizhaeva and Jiang Wang. Optimal Trading Strategy and Supply/Demand Dynamics. *SSRN eLibrary*, 2005. http://ssrn.com/paper=666541.

[39] Christine A. Parlour. Price dynamics in limit order markets. *Rev. Financ. Stud.*, 11(4):789–816, 1998.

[40] Nils Reich, Christoph Schwab, and C Winter. On kolmogorov equations for anisotropic multivariate Lévy processes. *Finance and Stochastics*.

[41] Mattias Sandberg and Anders Szepessy. Convergence rates of symplectic Pontryagin approximations in optimal control theory. *M2AN Math. Model. Numer. Anal.*, 40(1):149–173, 2006.

[42] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[43] Eric Smith, J. Doyne Farmer, Lászloó Gillemot, and Supiriya Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative Finance*, 3(6), 2003.

[44] Nassim Nicholas Taleb. *The Black Swan: Second Edition: The Impact of the Highly Improbable: With a new section: "On Robustness and Fragility"*. Random House Trade Paperbacks, New York, 2010.

[45] Henk C. Tijms. *Stochastic models*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1994.

[46] Curtis R. Vogel. *Computational methods for inverse problems*, volume 23 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. With a foreword by H. T. Banks.