Stochastic Modelling of Order Books with Non-Stationary Dynamics

Dr. Péter Bozsoki
Christ Church
University of Oxford

A thesis submitted for the degree of
Master of Science in Mathematical Finance
Hilary 2013
Zsuzsinak, Dávidnak és Annának.

A türelmükért, a segítségükért, a szeretetükért.
Acknowledgements

I am thankful to my supervisors, Prof. Philip Bond and Dr. Lajos Gergely Gyurkó for introducing me to this field, for their guidance, and their help throughout my work on this thesis.

I am grateful to my former employer, d-fine Ltd for the financial support they provided towards this course.

I would also like to acknowledge the use of the Oxford Supercomputing Centre (OSC) in carrying out this work.

This work would not exist without the relentless support and infinite patience of Zsuzsi. The number of nights, weekends, and holidays she sacrificed in order to help me to find time for my studies is too high to count.

Gyöngyi and Zsolt provided invaluable help with creating an environment to allow me to study. Gyöngyi has always been there to step in for me around the house and be a stimulating, caring grandmother for Dávid and Anna.

The numerical simulation part of this work would not be the same without Zsolt’s insight, expert advice, and hands-on help with the computational infrastructure. He also helped in such ingenuous ways as frequently as only a true friend does. Thank you!

Last but not least I would like to express my gratitude to my parents. The example of their life and the way they brought me up set me on the path to be a student at the University of Oxford.
Abstract

We develop a model for a simple order book and for the price evolution of an instrument traded on an exchange. The model can handle both stationary and highly non-stationary dynamics in the price of the instrument, including sudden and fast changes (“flash crash”). It also incorporates a feedback mechanism linked to the imbalance of the order book. Yet the model remains simple enough to allow for calculating results analytically. In particular, we derive an expression for the probability of the increase of the price.

We also establish a versatile volume-based numerical simulation framework. The order processing rules can be changed flexibly to a great extent and different order generation and submission processes can also be simulated in a numerically fast way.

We show via simulations that the feedback leads to sudden and large changes in the price. The crashes can equally go either down or up, which is a phenomenon observed in practice on real exchanges. Moreover, our simulation yields a crash in the price which is very similar to the real dynamics of Dow Jones Industrial Average during the flash crash.

We also consider order books typical at exchanges using first-in-first-out order processing as well as characteristic books of exchanges where the orders are processed on a pro-rata basis. The price volatility is observed to be less sensitive to the feedback in the case of pro-rata order processing.
Contents

1 Introduction 1

2 The Model of a Simple Order Book 3
   2.1 A Stochastic Model with Markovian Dynamics ................. 4
   2.2 Model for Non-Stationary Time Evolution ..................... 7

3 Simulation Methodologies 9
   3.1 Various Simulation Approaches .......................... 9
      3.1.1 Modelling the Dynamics of the State Descriptor with a Stochastic Process ....................... 10
      3.1.2 One-to-One Simulation ................................ 12
   3.2 Simulation of the Non-Stationary Model ..................... 13
      3.2.1 Detailed Simulation Algorithm ....................... 15

4 Simulation Results 19
   4.1 Single Realizations ...................................... 19
      4.1.1 Initial Condition .................................. 19
      4.1.2 Choice of Parameter Values ......................... 20
      4.1.3 Evolution of the Order Book ......................... 21
      4.1.4 Comparison of Simulated Crashes with the Flash Crash .......... 27
   4.2 Statistical Observations .................................. 27
      4.2.1 Simulation and Statistical Noise .................... 32

5 Analytical Results 37
   5.1 Laplace Transform and Its Application to Conditional Probabilities .......................... 37
      5.1.1 Definition and Basic Properties ...................... 38
      5.1.2 Inversion of Laplace Transform ...................... 39
      5.1.3 The Rational Approximation Method .................. 40
      5.1.4 Continued Fraction and Padé Approximation ........... 42
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.5 Laplace Transform of First Passage Times</td>
<td>44</td>
</tr>
<tr>
<td>5.2 Stationary Time Evolution</td>
<td>45</td>
</tr>
<tr>
<td>5.2.1 Inverse Laplace Transform</td>
<td>50</td>
</tr>
<tr>
<td>5.3 Non-Stationary Time Evolution</td>
<td>51</td>
</tr>
<tr>
<td>5.3.1 Inverse Laplace Transform</td>
<td>55</td>
</tr>
<tr>
<td>6 Summary</td>
<td>57</td>
</tr>
<tr>
<td>A Simulation Parameters</td>
<td>59</td>
</tr>
<tr>
<td>Bibliography</td>
<td>59</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

At 14:32 EDT on 6 May 2010 the Dow Jones Industrial Average (DJI) went from 10867.5 points to 9869 points in a few minutes resulting in a loss of 998.5 points or an intra-day loss of 9.2%. This is the largest intra-day point decline to that date [1]. Shortly after the E-Mini S&P 500 futures contracts (the “E-Mini”), one of the most active stock index instrument traded in electronic markets suffered a liquidity crisis. Between 14:45:13 and 14:45:27, automated high frequency traders traded over 27,000 contracts, which accounted for about 49 percent of the total trading volume, but the net effect was only buying approximately 200 additional contracts net [1]. These events were later named flash crash and were subject of a high level investigation by the U.S. Securities and Exchange Commission (SEC) and the Commodity Futures Trading Commission (CFTC) [1]. Analysis of historic tick data time series by Nanex, a market data provider company, revealed that the phenomenon of the flash crash is not unique [2]. Such a sudden and very fast changes in the price of a security or even index happen much more frequently than previously thought, albeit usually on a significantly smaller scale. As a result the causes and potential ways of preventing such flash crashes become the focus of numerous academic, industrial, and governmental research [3, 4, 5, 6, 7]

However, the major part of this research effort focused either on the particular events on 6 May 2010 [2, 3] or approached the phenomenon from the experimental aspect by looking for signals in the market data to predict a looming crash [4, 5]. While these are very important for the market participants we believe that a more fundamental understanding of the general phenomenon is also indispensable in order to see the complete picture.

To advance this understanding we develop a model for the order book of exchanges in our present work. There are two properties of such a model which are crucial for our aim: (i) it must be able to reproduce non-stationary dynamics, including sudden
and large changes in the price; (ii) it has to be simple enough to allow the analysis via analytical calculations. We investigate the properties of the new model by numerical simulation for which we introduce a simulation framework which allows for a flexible yet relatively fast way of simulating and processing large number of orders. We also derive an analytical expression to calculate the probability of increase of the price of the security traded on the modelled exchange within a given time.

This thesis is organized as follows. We discuss the model in Chapter 2, where we first introduce a stochastic model for the order book with Markovian time evolution based on a published work. Second, we extend this model such that it fulfils the two central conditions described above. The simulation methodology is described in Chapter 3 and the results of the simulations are presented and discussed in Chapter 4. Latter is followed by Chapter 5 where the mathematical analyses of the introduced models are derived. We conclude our work in Chapter 6.
Chapter 2

The Model of a Simple Order Book

This Chapter contains the description of the model of a simple exchange of a single instrument. We start with a review of the order book model and dynamics published by Cont et al. in Ref. [8]. Then we extend this model by introducing a feedback mechanism into the order generation process. To do so the way the dynamics is simulated numerically has to be modified significantly, as we will see in the Chapter 3. This feedback allows us to investigate non-stationary dynamics, including sudden changes in the price, e.g. price crashes.

We consider a single financial asset, e.g. a stock, which is traded in an order-driven market. The asset can be sold as well as bought and there are three types of orders which drive the dynamics of the price of the asset. The first type is the limit order which is an order to sell or buy a given number of instruments at a pre-defined price, fixed at the time of the submission of the order (for example, 100 IBM stocks at $4.12). Here the trader who submits the order is willing to risk that the order is completed only at a future time (or maybe not at all) but he/she can be sure that the transaction will be done at the price specified. The second type is a market order. Market orders specify only quantity and the side of the trade (i.e., sell or buy) and they are fulfilled immediately at the price of the most favourable price available on the current market. That is, a buy market order is matched to the lowest outstanding sell limit order on the book, while a sell market order is completed at the highest buy limit order on the order book. The third type is the cancellation of a submitted but not yet fulfilled limit order.

The exchange has to keep track of the limit orders because they may not be executed immediately, unlike the two other order types. The tracking is needed until the limit order is either cancelled or matched by an opposite order and executed. The limit order book, or simply the order book, is used to do this tracking of the outstanding limit orders. It is defined as a summary of all outstanding limit orders
which states the total quantities of the orders posted and whether they are buy or sell at each price level. The lowest price for which there is an outstanding limit sell order is called the ask price and the highest buy price is called the bid price.

Throughout the present work we omit one of the properties of the real world order submission from the current model: the size of the orders is not considered. More precisely, all orders are implied to have an order size of one unit, e.g. a single stock. While this is clearly different from how actual markets work it makes the analysis of the resulting dynamics substantially more transparent. We also note that this is only a limited constraint as a large order can always be broken down to a series of unit sized orders. Indeed it is a common practice in the real markets that large orders are split into smaller quantities to avoid moving the market.

2.1 A Stochastic Model with Markovian Dynamics

Cont, Stoikov, and Talreja have modelled the dynamics of the limit order book of a single instrument with a continuous-time stochastic process [8]. In their work the market is modelled by a price grid \( p = \{1, ..., n\} \), where \( p \) is the price and each of the \( n \) price values correspond to multiples of a price tick. For any time-scale investigated the maximum price \( p = n \) can be chosen large enough such that it is very unlikely that orders are placed at prices higher than \( n \).

The state of the order book is modelled by a continuous-time process \( X(t) \equiv (X_1(t), \ldots, X_n(t)) , \ t \geq 0 \), where \( |X_p(t)| \) is the number of outstanding limit orders at price \( p \). Negative \( X_p(t) \) values mean there are \( |X_p(t)| \) buy orders at price \( p \) and positive \( X_p(t) \) indicate \( |X_p(t)| \) sell orders at price \( p \). The ask price \( p_A(t) \) and bid price \( p_B(t) \) at time \( t \) are defined by

\[
p_A(t) = \inf \{ p = 1, \ldots, n, X_p(t) > 0 \} \wedge (n + 1) \tag{2.1}
\]

\[
p_B(t) = \sup \{ p = 1, \ldots, n, X_p(t) < 0 \} \lor 0 \tag{2.2}
\]

The last terms in both equations above are included to ensure there is always a bid and an ask price. These terms set the bid (ask) price to 0 (\( n+1 \)) when there is no bid (ask) order on the order book, respectively.

Using the ask and bid prices we define the mid-price \( p_M(t) \) and the bid-ask spread \( S(t) \) as

\[
p_M(t) \equiv \frac{p_A(t) + p_B(t)}{2} \tag{2.3}
\]

\[
S(t) \equiv p_A(t) - p_B(t) \tag{2.4}
\]
2.1. A STOCHASTIC MODEL WITH MARKOVIAN DYNAMICS

The time evolution of the order book is determined by how new orders flow in. For a state $X \in \mathbb{Z}^n$ and $1 \leq p \leq n$, define

$$X^{p\pm 1} \equiv X \pm (0, \ldots, 0, 1, 0, \ldots, 0) \quad (2.5)$$

where the 1 in the vector on the right-hand side is the $p^{th}$ component. Because we assumed that all orders are of unit size we can list the effect of the three order types for both the sell and buy sides [8]:

- a limit buy order at price level $p < p_A(t)$ increases the quantity at level $p : X \to X^{p-1}$
- a limit sell order at price level $p > p_B(t)$ increases the quantity at level $p : X \to X^{p+1}$
- a market buy order decreases the quantity at the ask price: $X \to X^{p_A(t)-1}$
- a market sell order decreases the quantity at the bid price: $X \to X^{p_B(t)+1}$
- a cancellation of an outstanding limit buy order at price level $p \leq p_A(t)$ decreases the quantity at level $p : X \to X^{p+1}$
- a cancellation of an outstanding limit sell order at price level $p \geq p_B(t)$ decreases the quantity at level $p : X \to X^{p-1}$
- cancellations at each price level, each of which can be represented as a counting process.

Bouchaud et al. empirically observed that incoming orders arrive more frequently in the vicinity of the current bid/ask price and the rate of arrival of these orders depends on the distance to the bid/ask [9]. To capture these empirical features Cont at al. propose a stochastic model where the events outlined above are modelled using independent Poisson processes. Let $i \geq 1$ denote the distance of $i$ ticks from the opposite best quote as number of ticks. Then it is assumed that

- Limit buy (respectively, sell) orders arrive at rate depending on $i$ at independent, exponential times with rate $\lambda(i)$,
- Market buy (respectively, sell) orders arrive at independent, exponential times with rate $\mu_0$,.
• Cancellations of limit orders also depend on the distance from the opposite best quote and they also occur at a rate proportional to the number of outstanding orders: if the number of outstanding orders at price $p$ is $|x_p|$ then the cancellation rate is $\theta(i) |x_p|$. This assumption can be understood as follows: if we have a batch of $|x_p|$ outstanding orders, each of which can be cancelled at an exponential time with parameter $\theta(i)$, then the overall cancellation rate for the batch is $\theta(i) |x_p|$. Here $p$ and $i$ are connected through the ask or bid price, depending on whether the order is a sell or buy.

• All events above are mutually independent.

The arrival rates of the incoming orders depend on the distance to the bid/ask. To increase the probability of the execution, traders usually place most orders close to the current ask/bid prices. Therefore the arrival rate is modelled as a function 

$$
\lambda : \{1, ..., n\} \rightarrow [0, \infty) \text{ of the distance to the bid/ask.}
$$

Empirical studies suggest a power law 

$$
\lambda(i) = \frac{k}{i^\alpha}
$$

as a plausible mathematical model. [9, 10]. $k$ is a model parameter.

Ref. [8] also specifies an empirical cancellation rate. We use a close analytical approximation of the published empirical values in our work which is 

$$
\theta(i) \equiv \left( \theta_1 \cdot e^{-\log^2(i)} + \theta_0 \right) \cdot |x_p|.
$$

The parameters $\theta_0$, $\theta_1$ are chosen such that $\theta(i)$ matches the empirical rates in Ref. [8]. They are also shown for parameter values used in the simulations in Fig. 3.1.

The market order rates are assumed to be a constant value, denoted here by $\mu_0$: 

$$
\mu \equiv \mu_0 = \text{const}
$$

With these assumptions, $X$ is a continuous-time Markov chain with state space $\mathbb{Z}^n$ and transition rates are given by

- $X \rightarrow X^{p-1}$ with rate $\lambda (p_A(t) - p)$ for $p < p_A(t)$
- $X \rightarrow X^{p+1}$ with rate $\lambda (p - p_B(t))$ for $p > p_B(t)$
- $X \rightarrow X^{p_B(t)+1}$ with rate $\mu_0$ for $p > p_B(t)$
- $X \rightarrow X^{p_A(t)-1}$ with rate $\mu_0$ for $p < p_A(t)$
- $X \rightarrow X^{p+1}$ with rate $\theta (p_A(t) - p) |x_p|$ for $p < p_A(t)$
- $X \rightarrow X^{p-1}$ with rate $\theta (p - p_B(t)) |x_p|$ for $p > p_B(t)$
In reality the ask price is greater than the bid price. When a sell (buy) limit order is submitted below (above) the bid (ask) price, respectively, it is called order crossing. It depends on the particular market how this is handled in practice. In our work we avoid such events by design. The criterion of not allowing order crossing can be formulated by defining the admissible state of the order book. Following Cont et al. we say a state $X(t)$ is admissible if it fulfills the requirement

$$\mathcal{A} \equiv \{ x \in \mathbb{Z}^n | \exists k, l \in \mathbb{Z} \text{ s.t. } 1 \leq k \leq l \leq n, x_p \geq 0 \text{ for } p \geq l, x_p = 0 \text{ for } k \leq p \leq l, x_p \leq 0 \text{ for } p \leq k \}.$$  

(2.9)

Cont et al. also proves that the order book remains admissible with probability one if the initial state admissible and only the transitions above are allowed [8]. Formally, if $X(0) \in \mathcal{A}$, then $P[X(t) \in \mathcal{A}, \forall t \geq 0] = 1$. In other words, if the order book initially has all buy limit orders at prices below all ask orders than this property will not change as the system evolves with the given transitions.

Another important property of this model is that $X$ is an ergodic Markov process. Thus $X$ has a stationary distribution. For detailed proof we refer the reader to Ref. [8].

## 2.2 Model for Non-Stationary Time Evolution

In order to be able to investigate sudden and large changes in the price we have to change the model described in the previous Section. While the new model will clearly have non-stationary distribution our aim is to preserve the possibility of relying on the analytic framework developed in the Chapter 3 of Ref. [8].

Our choice for the model extension is motivated by the findings of Easley, Lopez de Prado, and O’Hara [4, 5]. Their work establishes a link between the properties of the order arrival process and the subsequent changes in the price of the instrument. In particular, they assume that if all active traders on the market are uninformed then the order flow contains the same amount of buy and sell orders on average. However, if there are traders with information not known by all market participants the informed traders will submit either sell or buy orders predominantly. This can be viewed as the process which forces the price of the instrument to reflect all information, including the latest developments. As this obviously can lead to high losses for the uninformed traders, especially market makers, Easley et al. call such flow of incoming orders toxic. Their results include introducing a measure, the Volume-Synchronized Probability of Informed Trading (VPIN) to quantify the amount of information in
the flow of incoming orders, and showing that high VPIN values were observed just before the flash crash on 6 May 2010 [4].

We modify the market order submission rate, $\mu$, and make it dependent on the state of the order book as well as allowing for different $\mu$ values for the buy and sell orders.

\[
\mu_{\text{sell}}(X(t)) \equiv \max \left( \mu_0, s \cdot \sum_{p=1}^{n} x_p \right) \quad (2.10)
\]
\[
\mu_{\text{buy}}(X(t)) \equiv \max \left( \mu_0, s \cdot \sum_{p=1}^{n} (-x_p) \right) = \max \left( \mu_0, -\mu_{\text{sell}}(X(t)) \right) \quad (2.11)
\]

This yields an arrival rate at least $\mu_0$, i.e., that of the original model, but $\mu_{\text{sell}}$ ($\mu_{\text{buy}}$) start to increase as soon as the total number of sell (buy) orders outweighs the total number of buy (sell) orders, respectively. The arrival rate is proportional to the difference of total number of sell orders minus the total number of buy orders on the book at time $t$. The parameter $s$ determines the speed of the increase, i.e., the strength of the feedback.

An intuitive interpretation of this feedback is that the market participants follow the general market sentiment: if a trader sees that there are, say, more outstanding sell orders than buy order he/she can follow a momentum strategy (or simply assume that the market has some information he/she is not aware of) and submit a sell order as well. However, his/her new order will be a market order, which is executed immediately to ensure the positions are closed without delay. This execution also implies that a buy order on the book was matched and taken out. Thus the imbalance will be even greater after the execution of this trade. This translates into even stronger signal for other market participant to follow the same course of action.

Note that the distribution of the arrival times remains exponential, the same distribution type for the arrival times of the cancellation and the arrival times of the limit orders. The parameters of the distributions are different, of course. Indeed, as it is visible from Eqs. (2.6, 2.7, 2.10, 2.11) each price has six different order rate value associated with it: limit order, market order, cancellation for both buy and sell side. Two of these, $\mu_{\text{sell}}$ and $\mu_{\text{buy}}$ depend on the state of the order book and hence introduce a memory into the dynamics. This will lead to non-stationary time evolution.
Chapter 3

Simulation Methodologies

In this Chapter we describe the simulation setup used to investigate the behaviour of the extended model detailed in Chapter 2. We start with briefly discussing three approaches which turned out to be too limited to successfully tackle all aspects of the problem at hand. These are the following: (i) modelling the dynamics of the state descriptor of the order book with a stochastic process; (ii) one-to-one simulation of all market players, including the exchange, market data distributor, and multiple traders using the Simulink product of MathWorks [11]; (iii) one-to-one simulation of all market players using MATLAB code. We use the term one-to-one simulation to describe a simulation setup where there is a one-to-one relationship between the real world entities and the simulated objects. These objects in the computer mimic the major aspects of the behaviour of different market participants by simulating their actions numerically as closely as possible.

The setup which proved fast yet versatile enough for present work turned out to be a mixture of the two approaches mentioned above: the exchange was simulated in a one-to-one way by replicating the main operations of an exchange in the code but the traders and order submission were modelled by a random process. In the second, major part of this Chapter, we relate the simulation setup used to study the extended model with non-stationary transition probabilities. The results from this simulation setup are in Chapter 4.

3.1 Various Simulation Approaches

This section describes the main features of those different simulation approaches which were tried out during the research phase of current project but were deemed not satisfactory. In particular, they proved either to be too inflexible from modelling
variability point of view or too slow to numerically simulate the required system within a reasonable time.

The motivation to include them in present document is twofold: (i) We believe that their brief review including the shortcomings may be informative for any potential future work on the subject. (ii) The simulation arrangement used for the calculations is a combination of the methods described here and hence understanding these methods facilitates the documentation of the former.

### 3.1.1 Modelling the Dynamics of the State Descriptor with a Stochastic Process

The first method we discuss here is the direct simulation of the time evolution of the state descriptor. In the case of the model outlined in Chapter 2 the state descriptor is the order book. The states of the system described by $n$ discrete possible price values, indexed by $i = 1..n$. The arrival and cancellation times of a single order for each price level $p_i$ are modelled by an exponential distribution with a well-defined rate. The order book is represented by a state descriptor vector $X$, with elements $X_i, i = 1..n$. Here the integer value of $X_i$ denotes the number of outstanding orders at price $p_i$.

We have implemented two types of dynamics in this case:

**Event-driven simulation:** In this case the state descriptor was evolved by a change in the value of a single element of the vector with difference of one event. This is also how the dynamics is described by Cont et al. in the Section 1.2 of Ref. [8]. In particular, the simulation of transition probabilities included the following steps:

1. Drawing a random arrival time both for new orders and for cancellation at each price level from an exponential distribution with the corresponding rates. This yields $2n$ time values: $n$ new orders and $n$ cancellations.

2. Determining the earliest event, i.e., the minimum of the $2n$ times and drop the remaining $2n - 1$ values.

3. Changing the value of the corresponding $X_i$ by one. The sign is determined by the sell or buy nature of the event and also whether it is a new order or cancellation.
As indicated in the introduction of this Chapter, this was the first implementation chronologically. As the work progressed it became clear that this algorithm could be sped up by replacing the first two steps with drawing two random numbers: the first random number would be from an exponential distribution and it would specify the time if the next arrival. The second draw would determine the type and price of the order coming in. This modified algorithm, however, would nevertheless be less flexible from the point of view of introducing additional features than the algorithm we introduce in the next Section.

This models a situation where every event on the market is immediately visible to all participants and they all modify their behaviour in no time. One can draw analogies with classical physics, where interaction needs zero time as the propagation speed of the interaction (or here: information) is infinite.

**Time-step driven simulation:** The other possibility is a time-step driven scenario, where the time evolution is sliced up to small $dt$ intervals. It uses the Poisson distribution corresponding to the exponential arrival times with the proper rate to calculate the number new orders and/or cancellations within $(t, t + dt)$. The simulation now consists the following steps:

1. Draw a random number from the Poisson distribution both for new orders and cancellation for each price level.
2. Net the new orders and cancellations. The result is the change in $X$, denoted by $dX$.
3. Add $dX$ incrementally to the original $X$.

The event-based approach has the advantage that it follows the evolution in its deepest level. Its disadvantage is that it is far more computationally intensive than the time-step based evolution.

There is also a more fundamental difference between the two approaches. The time-step driven simulation allows for the possibility of order crossing: within $dt$ a buy *and* a sell order arrives with the buy submitted at higher price than the sell. This reflects the fact that market participants need a finite time to learn the state of the market and act on it. Thus it is the more realistic version.

However, in simulation arrangements where order crossing can happen additional rules and corresponding actions needed to tackle these situations. This complicates
significantly both the numerical and the analytical analysis of such systems. Therefore we did not pursue this direction further in our present work.

Unfortunately the simplicity of the direct simulation of the state descriptor turned out to be a serious limitation. Our implementations of this approach exhibited Markovian dynamics, as expected. The limitation is that, in our experience, it is not possible to modify the transition probabilities such that the resulting dynamics produces sudden and large changes in the price. In particular, we modified the type and parameters of the distribution of the inter-arrival times using the event-driven simulation setup. We also kept the modelling assumptions used in Ref. [8] in place, which prohibited crossing and allowed only “orderly” order book, with distinct sell and buy sides separated by a spread. It seems that even changing the distribution of the inter-arrival times to a long-tailed, Lévy alpha-stable distribution does not lead to a different type of dynamics in this framework. We emphasize that this is an observation of the particular simulation setup we used and clearly not a mathematical statement.

3.1.2 One-to-One Simulation

An alternative way to the direct simulation of the dynamics of the state descriptor of the order book is a simulation where the market participants are “replicated” in the virtual world of computers. There is a one-to-one relationship between the real world entities and the simulated objects as well as between the actions of the real world entities and the behaviour of those numerical objects. In other words, in a one-to-one simulation there are virtual objects for the market participants, one for each: the traders, the queuing system, the exchange, the market data system, etc.

We created two prototypes of such a one-to-one simulation. The individual traders were modelled with their own submission algorithms together with an exchange system, composed from a queue and order book processing mechanism. The numerical simulation of the queue was based on the book of W. J. Stewart, in Ref. [12].

The two prototypes differed in the underlying technological platforms:

SimEvents: SimEvents is a toolbox of Simulink [11], which is a product of MathWorks. It is designed to simulate discrete events (e.g., arrival of an order or its submission) and as such it appeared to be a good candidate for the current problem.

While this software product seems to support well a wide range of queue systems, there is no direct support for the more complex processing logic needed to process the incoming orders and maintain the order book.
3.2. SIMULATION OF THE NON-STATIONARY MODEL

Events and listeners based MATLAB code: An existing framework is available for exchange simulation [13]. It was developed and published by a member of the MATLAB user community on the official MathWorks forum. It is written in MATLAB, and relies on two object oriented programming techniques, events and listeners, to handle the communication of the various (virtual) market participants.

As the MATLAB source code was also available, this was a promising framework which can be used to create a more articulated simulation.

Albeit the published code provides a robust design and an intuitive visualisation for a small system, it proved computationally demanding to process a relatively small amount of orders. The reason of this is that it relies on the built-in implementation of events and listeners of MATLAB. These built-in parts of MATLAB have a limitation of lowest possible time resolution of 1ms which limits the number of processed orders per unit time significantly.

3.2 Simulation of the Non-Stationary Model

Our aim was to create a simulation setup where non-stationary dynamics (e.g. flash crash) can be simulated and investigated. At the same time it had to be reasonably fast as well in order to be able to generate and process a statistically significant number of orders. Moreover, it had to be flexible so different models can be simulated with minimal changes in the code.

Handling the orders submission by traders on the market separately from the processing of the orders by the exchange makes possible to modify easily either part without changing the rest of the system. It, however, poses a numerical challenge. Namely, the different parts of the system can operate at substantially different speeds leading to events happening on different time-scales (e.g., one microsecond ($\mu$s) and 1 millisecond = 100 $\mu$s). This mirrors reality, where one trader can act much faster than the other or the exchange can process submitted orders much faster than they are submitted or, vice versa, the exchange can be overwhelmed by a submission of a many orders, if a trader is allowed to be faster. The presence of multiple time-scales can usually not be handled in a numerically efficient and fast way in time-based simulation setup with fixed timesteps. This is so because the timestep of the simulation implies one time-scale which is distinguished over the others in the system.

Simulating multiple order sources (e.g., multiple traders) separately in the system can lead to simultaneous submissions. This can lead to order crossings and other complications which can heavily hinder the analysis and understanding of the dynamics.
Therefore we decided to exclude the possibility of simultaneous submission from the system despite that this phenomenon is clearly a natural part of real systems.

Based upon the experience we gained from the implementation of the simulation setup described in Sec. 3.1 we converge on an volume-based setup relying on the combination of one-to-one simulation of selected parts and a statistically modelled processes.

The outline of the setup is the following:

- The exchange is modelled by a numerical vector representing the order book and a separated order-processing function. The order processing is invoked after each order submission. The order submission is allowed to submit only one order at each step (see below). This ensures that no crossing happens and an “orderly” order book can be maintained during the evolution without more complex algorithms. Note that this implies a real-world situation where the exchange works much faster than any trader which may not always be the case.

- All traders are jointly represented by a single process which generates the orders. Please note that this is not intended in its current form as an explicit model of the market participants, but rather as a simple way to generate and feed orders to the rest of our model. Because of the modular nature of the model it can be later modified or even replaced. However, modelling traders is a field on its own right and thus it is a serious undertaking [14, 15].

- This process observes the changes in the order book immediately. In other words, no time needed for the information to propagate from the exchange to the traders. However, a delay can easily be introduced in the process.

- Simulation time advances in an event-oriented (volume-based) way. One timestep corresponds to exactly one order submission and also the processing of that order. The overview of the corresponding algorithm is:

1. Generate the type, side, and, if needed, the quote of the order based on the information currently on the order book (ask/bid price, etc.) using a uniform random distribution. The probability of each type/side/quote order is determined by the order submission rates published on page 4-5 of Ref. [8] and are listed in Appendix A.

2. The order is processed immediately by either executing it or by taking it on the order book.
In this arrangement the generation and the submission of new orders is now separated from the processing of incoming orders. This is in contrast with the setup outlined in Sec. 3.1.1 where the order book dynamics is simulated directly as a Markov chain following the algorithm described in Ref. [8]. This separation of steps also allows us to choose freely the processes which generates the orders. The process can be simple or complicated, can depend on the last known state of order book or the last \( n \) states, etc.

Another important feature is that there is no inherent time-scale included in the order generation or the order processing. In other words, the core code does not know if 1 \( \mu s \) or 1 sec has elapsed since the last order. This is central to allow simulation of traffic with large fluctuations in the volume. If the code had an internal fixed timestep clock it would not be able to cope with the high load and would waste CPU time when there is no order.

Note that this does not mean that the order generation or order processing can not depend on time. Indeed it can, only time has to be passed in as input parameter.

### 3.2.1 Detailed Simulation Algorithm

After outlining the used simulation setup we describe here the individual steps in the algorithm together with how the various objects are represented in the code. We start with the details of the exchange and continue with the order generation.

#### Exchange

The simulation of the exchange is a rather straightforward implementation of the mechanism of a simple exchange for a single stock:

- There is an order book, represented by a vector \( p \). The index of the vector is the price and the integer values of the vector are the number of outstanding orders. Negative numbers indicate buy orders, positive numbers mean sell orders. E.g. if \( p(34) = 5 \) it means there are 5 sell orders at price 34. Similarly, \( p(28) = -12 \) is 12 outstanding buy orders at price 28.

- Ask (bid) price is defined as the lowest (highest) index \( a \) \((b)\) in \( p \) such that there are only non-positive (non-negative) numbers in \( p \) for indices \( i < a \) \((i > b)\).

Index \( a \) \((b)\) is the ask (bid) price.
• Given the way the submission rates/probabilities are calculated and if the system started from a state where there is no crossing then no crossing can happen, as per Ref. [8]

• Market buy (sell) orders are always executed at the ask (bid) prices. I.e., a market buy (sell) decreases the \( p(a) \) \( p(b) \) by one.

• Limit order with price \( i \) is executed immediately after submission if there is a matching order at that price. I.e., the absolute value of \( p(i) \) is decreased by one with the correct sign, as the new order takes out one on the order book: +1 for limit buy, -1 limit sell.

If there is no matching order already on the order book it is added to it: absolute value of \( p(i) \) increased by one with the correct sign (-1 for limit buy, +1 limit sell).

• Cancel buy (sell) order at price \( i \) has the same effect as a limit sell (buy) with a matching opposite order: it decreases the absolute value of \( p(i) \).

Order Generation

With its origins from the model in the Cont paper [8], the order generation is composed of the following steps:

1. Calculate the order rates: calculate the rate \( r_{i,o} \) for price \( i \) and order type \( o \). There are six order types: (1-2) market sell or buy (3-4) limit sell or buy and (5-6) cancellation of an existing sell or buy order.

\( r_{i,o} \) is a non-negative number for all \( i, o \) values. How it is calculated in the different cases is detailed on the next page.

2. Calculate the order probabilities: The previous step yields 6 values for each price \( i \) and \( i = 1..D \), where \( D \) is the order book depth. Putting all rates into matrix with dimension \( 6 \times D \) this matrix is normalised such that its overall sum is equal to one:

\[
R \equiv \sum_{i=1}^{D} \sum_{o=1}^{6} \tilde{r}_{i,o} = 1 \quad (3.1)
\]

where

\[
\tilde{r}_{i,o} = \frac{r_{i,o}}{\sum_{i=1}^{D} \sum_{o=1}^{6} r_{i,o}} \quad (3.2)
\]
3.2. SIMULATION OF THE NON-STATIONARY MODEL

Figure 3.1: Model parameters as a function from the distance of opposite quote. (a) shows the submission rate of the limit orders for model parameter $\alpha = 0.2, 0.5, 1$. These values were the usual values of used in the simulations. (b) the rate of cancellation coefficient for the model parameters: $\theta_0 = 0.34$, $\theta_1 = 0.47$, $\theta_3 = 1/3$. These values are the representative numbers in both cases and were chosen such that the resulting rates matches the functional form and values of the published values [8].

3. *Calculate the cumulative mass function:* Matrix $R$ is viewed as a two-dimensional probability mass function. As a next step the corresponding cumulative mass matrix $C$ is calculated. At this point we have monotonously increasing values spanning the interval $(0, 1)$.

4. *Choose a price and order type randomly:* A random value is drawn form a uniform distribution from the interval $(0, 1)$. This random number is mapped to an order type (including side) and price using the cumulative mass matrix $C$ in the previous step.

This step is the sole source of randomness in the simulation at the moment.

5. *Submit the order to the exchange:* As mentioned earlier, order size is always one in the current simulation.

The submission rates $r_{i,o}$ are calculated using the model as described on page 4-5 of Ref. [8]:

**Market order** The following three cases are distinguished

- *No feedback* $r_{i,o}$, for $o = 1$, $i = a$ (ask) and $o = 2$, $i = b$ (bid). $r_{i,o} = 0$ otherwise. $\mu$ is constant, from Eq. (5.46) $\mu = \mu_0$. Typical value is between $0.42 - 1$. 

![Limit order parameter](image1)

![Cancellation parameter](image2)
CHAPTER 3. SIMULATION METHODOLOGIES

- **Feedback** \( r_{i,o} = \max(\mu, \Delta) \) for \( o = 1, i = a \) (ask) and \( o = 2, i = b \) (bid). \( r_{i,o} = 0 \) otherwise. \( \Delta \) is the imbalance, defined as the difference of number of sell order and number of buy orders. As state descriptor \( x_p \) has different signs for buy and sell \( \Delta \) is \( \Delta = \sum_{p=1}^{n} x_p \).

Note that \( \Delta \) is usually a large integer number, far surpassing \( \mu \). Larger the depth of the order book \( D \) is more space is available for orders and they all add up. The difference \( \Delta \) clearly depends on their position and distribution between the two sides, but bigger \( D \) is, larger \( \Delta \) values may take.

**Limit orders** \( r_{i,o} = \frac{k}{d^\alpha} \), \( o = 3, 4 \) where \( k, \alpha \) are constant parameters and \( d \) is the distance from the bid or ask price, i.e., \( d(i) = i - b \) for sell orders and \( d(i) = a - i \) for buy orders.

**Cancel orders** As described in Eq. (2.7) the submission rate of cancellation is given by

\[
 r_{i,o} = \left( \theta_1 \cdot e^{-\log^2(d(i) - \theta_2)} + \theta_0 \right) \cdot |p(i)| \quad \text{for } o = 5, 6. \tag{3.3} 
\]

\( d(i) \) is defined the same way as for limit orders and \( p(i) \) is the number of orders at price \( i \) (i.e., the value of element \( i \) in \( p \)). As mentioned in Chapter 2, this form above is an analytical approximation of the real-world pattern [8].
Chapter 4

Simulation Results

We present our results of simulating the dynamics of the two models described in Chapter 2 using the simulation algorithms detailed in Chapter 3 below.

This Chapter has two major parts. We investigate typical single realizations of the evolution of the order book in the first part and discuss the results obtained by ensemble averaging in the second.

4.1 Single Realizations

4.1.1 Initial Condition

Once the simulation algorithm, detailed in Sec. 3.2, has been implemented in MATLAB, which our choice of programming language, the initial condition had to be determined.

In the stationary case the distribution of the orders is mainly determined by the $\alpha$ parameter introduced in Eq. (2.6). It determines how the limit order submission rate, $\lambda$, depends on the distance from the opposite best quote. We displayed characteristic curves of $\lambda$ in Fig. 3.1(a) and it is clear that the low and high $\alpha$ values lead to considerably different submission patterns. In particular, low $\alpha$ leads to order submission characteristic in markets where the exchange processes the orders in a first-in-first-out (FIFO) basis. High $\alpha$ models situations where processing of the orders sitting at the same price is done by pro-rata basis.

The model has many parameters which leads to a large parameter space. Thus we decided to use the same initial condition in all simulations. This made separating the effect of various factors easier (by fixing potential variable, i.e., the initial condition). One disadvantage of this is that we have to evolve the system longer to minimize the effects of the initial condition on the results.
Figure 4.1: Graphical representation of the order book at the start of the simulation. Negative number of outstanding orders indicates buy orders and positive ones correspond to the sell orders.

The particular initial condition used during the simulation is shown in Fig. 4.1. We think that this shape resembles reasonably closely the shape of the $\lambda(i)$ for a relatively large range of $\alpha$ parameters. As a result we expect that the time evolution has short transient behaviour because we start the system from a state which is relatively close to the particular shape of the order book for all simulation parameters used.

In order to see if this expectation is reasonable we also started the system from a different, albeit somewhat unrealistic initial condition: we set the number of outstanding orders on the book to the same constant number. That is, there were $m$ buy orders sitting at all prices at and below the bid price and also $m$ sell orders at all prices at and above the ask price. The results confirmed our expectations: the resulting time evolution did not differ qualitatively for the chosen set of simulation parameters.

### 4.1.2 Choice of Parameter Values

In reality crashes are rare, so long time passes between them. Due to the limitation in the computing power we intentionally focus on parameter values where we can expect that the pattern we wish to investigate in more details emerges. This makes the simulated scenarios less realistic, but it allows better analysis. There are two
patterns we are particularly interested in: the extreme case where crashes always occur and another where the price evolution is non-stationary yet it does not involve any crash, at least not within the simulated period.

Thus we initially scanned the parameter space to find the range which best fits our requirements. We probed $\alpha \in [0.2, 8]$, $s \in [10^{-3}, 1]$, $\theta_1 \in [0.2, 8]$, where $\alpha$, $\theta_1$, and $s$ are introduced in Eqs. (2.6, 2.7, and 2.10−2.11).

Non-stationary yet crash-free dynamics is observed when the feedback $s$ is very weak. Based on our experience, if $s \leq 0.02$ then crashes are unlikely. To make reading the values of $s$ easier we quote them in percentage from now on, i.e., $s \leq 2\%$. This value is somewhat blurred by the length of the simulation: longer it is, more chance there is for a crash to emerge. Indeed, the shorter simulations with $10^6$ orders processed, discussed in Sec. 4.2, display crash-free dynamics for $s$ up to $5\%$, while the next section below, Sec. 4.1.3, contains results where $s = 3\%$ and the price crashes after $6.5 \cdot 10^6$ events are processed. This is even more pronounced in Fig. 4.5, where we see that higher $s$ is, sooner the crash happens.

Fig. 4.5 also shows the parameter range where crashes reliably occur: long simulations (up to $10^7$ orders processed) and $s \geq 3\%$. Alternatively, if $s > 25\%$ a crash occurs during a shorter simulation as well.

There is also interplay between other parameters than $s$ and simulation length. For example, we have observed that if $\alpha \geq 2$ than price was crashed independently from the value of $s$. Because of the limited computation time these options could have been investigated only superficially and thus we decided not to include them in this thesis.

4.1.3 Evolution of the Order Book

First we simulate the stationary model of Sec. 2.1 as its evolution is known from Ref. [8] and as such our simulation must be able to reproduce the characteristic of this known case. We simulate the evolution of the order book for a set of parameters using long simulations during which $10^7$ orders are submitted and processed. Independently from the chosen set of parameters the graphical analysis of the evolution of the mid-price does not show any sign of long term memory or sudden changes. The same is true for the ask and bid prices as well as the spread. Fig. 4.2 illustrates on one chosen example the evolution of the mid-price and the spread of the order book. Note that these two values carry exactly the same information as the bid ask and bid prices with the transition between the two pairs given by Eqs. (2.3−2.4).
Figure 4.2: Stationary evolution. Parameters: \( \mu = 0.42, \ k = 1.92, \ \alpha = 0.5, \) distributionParamsCancellationDefaultVal: 0.4700 distributionParams CancellationScale: 0.3400, Price range: \( 1 - 10^4 \) ticks.

We also calculated to autocorrelation function of \( p_M \), defined as

\[
\rho(k) \equiv \frac{\sum_{i=1}^{N-k} (p_M(i) - \mathbb{E}[p_M])(p_M(i + k) - \mathbb{E}[p_M])}{\sum_{i=1}^{N} (p_M(i) - \mathbb{E}[p_M]^2)}
\]

in order to detect any long-term memory. However, \( \rho(k) \) decayed extremely fast and there was not significant autocorrelation for \( k > 10 \) values. This is also in line with our expectations and supports the conclusion that the implemented numerical simulation behaves as expected.

As our aim is not to reproduce the quantitative results of Ref. [8] but to investigate non-stationary evolution we do not pursue the investigation of this model further here.

We start the investigation of the extended model introduced in Sec. 2.2 in a similar fashion to that of the previous case and we focus on a single realization first.
We plotted the evolution of price on the top panel of Fig. 4.3 in the case of a strong feedback. The range within the price is allowed to vary is ranging from 1 to $10^4$ with $p_M(0) = 5000$, i.e., set to the middle of the allowed range. In the general case, $\mu_{\text{sell}}$ and $\mu_{\text{buy}}$ defined in Eqs. (2.10–2.11) both are allowed to reach large values. The initial random fluctuation determine which of them will be the dominant. In this particular realization of the stochastic process the sell side becomes dominant. Because of the feedback present in the model, i.e., larger the imbalance of the order book is more the probability of the further market sell orders increase, once one side becomes dominant it is very unlikely that fluctuations caused by the cancellation and limit orders affect the outcome.

This can be seen on the middle panel in Fig. 4.3. We plotted three quantities here: the order probability of the market sell (green) and order probability of market buy (black) as they are defined in Eq. (3.2). The third quantity is the the maximum of all order probabilities in the system, i.e., we calculated the max $\tilde{r}_{i,o}$ over all prices ($i$) and order types ($o$). We highlight that both the market buy and the market sell order probabilities are below the maximum value initially. In other words one of the limit order probability or the cancellation probability is dominant initially. However, as more and more orders are submitted and processed the market sell order probability start to increase. Moreover, the speed of the increase is increasing as well and when this happens the mid-price crashes to 0.

The crash also observable in the size of the spread, shown on the bottom panel in Fig. 4.3. After the initial transitionary period the spread is only a few ticks. But when the market sell order probability becomes the most likely order submission in the system (the market “panics”) the spread suddenly widens to 5-10x to its usual value.

The inset in the upper panel shows that there are fluctuations in the mid-price in the early stage of the evolution, they only not visible in the main plot because of the scale differences. We also highlight that the size of fluctuations and characteristics of the dynamics is very similar to that of seen in Fig. 4.2 for the stationary case.

As we mentioned above, it depends on the initial fluctuations whether the buy or sell side becomes dominant. Fig. 4.3 showed an example for the former. In Fig. 4.4 we can see an example of the latter: the buys side becomes dominant, simulating a bubble on the market or an “upward crash”. Latter is also observed on the market and as we can see here it is not fundamentally different from a change in the opposite direction. The plateau levels observed in the upper panel of Fig. 4.4 is where the price reaches the highest allowed price in the simulation. If this barrier originating from
Figure 4.3: Price crash caused by strong feedback. Parameters: $s = 3\%$, $\alpha = 0.3$, all other parameters have default values.
Figure 4.4: *Upward* price crash caused by strong feedback. Upper price cap is $10^4$ ticks. Parameters: $s = 3\%$, $\alpha = 0.3$, all other parameters have default values.
CHAPTER 4. SIMULATION RESULTS

Figure 4.5: Price crash caused by strong feedback for different $s$ values. Upper panel: stronger feedback leads to faster crash. Lower panel: “inverted” log scale to see the magnitude of fluctuations during the initial period. Parameters: $s = 3\%$, $3.5\%$, $4\%$. $\alpha = .5$, all other parameters have default values.

The numerical nature of the simulation was not there $p_M$ would increase further, ever faster. The two subsets, in the upper and middle panels are to show the the initial evolution is also volatile but they are not visible on the main graphs because of the scale of the vertical axis.

Given that the upwards and downwards crashes are the same from mathematical point of view in this model we concentrate on the downwards crashes from now, but all the results are applicable for the cases of upwards crashes.

We expect that a stronger feedback, expressed by the parameter $s$ leads to shorter time, measured as number of orders, needed to a price crash to materialize. We simulated the model for different $s$ values with the result shown in Fig. 4.5. The upper panel in figure shows the evolution of $p_M$ for three different $s$ values: $s = 3\%$ (blue curve), $s = 3.5\%$ (green curve), and $s = 4\%$ (black curve). These simulation results show the expected results. The lower panel shows the same three curves of $p_M$, but
4.2 Statistical Observations

on a transformed vertical axis. Instead of depicting $p_M$ on a linear scale as it is in the upper panel, we plot $\log(p_M(0) - p_M(t))$. This allows us to observe the fluctuations of $p_M$ before it crashes to 0. Interestingly, the strength of the feedback also seems to influence the volatility of $p_M$ during the phase of evolution when the dynamics is not yet dominated by the “panic” on the market and it is close to the stationary case.

4.1.4 Comparison of Simulated Crashes with the Flash Crash

The joint investigation of CFTC and SEC identified a combination of a very large sell order and an aggressive selling wave by high frequency traders (HFT) as one of the major causes of the flash crash observed on 6 May 2010 [1]. This conclusion is also supported by the analysis of Nanex and the work of O’Hara et al. [2, 4, 5]. Besides this cause, Nanex proves that there was a flood of limit orders far from the ask and bid prices which were then cancelled almost without exception [2]. This is a not-so-rare form of price discovery practice and called “quote stuffing”.

Comparing these explanations with the simulation results of our model we see that highly asymmetric market sentiment can lead to crashes on its own. The way the events unfolded in the case of the flash crash suggests that many HFT algorithms take into account either the imbalance on the market or some other related measure. In other words they seems to follow the (market) trend. While we can only speculate if this is indeed the case or not, we are sure in the following: Our results show that such a behaviour aggravates the problem further and, if enough traders (humans or algorithms) follow suit, the price crashes.

For an in-depth quantitative comparison with the actual event access to the intraday market data is needed. This is unfortunately out of our reach. Nevertheless, comparing the published time evolution of Dow Jones Industrial Average (DJI) from Ref. [2] with the simulated crashes in our model in Fig. 4.3 or in Fig. 4.5 shows remarkable similarities between our simulation results and the observed price crash in reality.

4.2 Statistical Observations

After evaluating the simulation results in detail for single realizations we proceed to the decrease simulation noise. This is achieved by calculating an ensemble average for a given set of parameters. The numbers of orders processed in each realizations has to be decreased significantly, to $10^6$ processed orders due to the finite computational resources. Even so, the high performance computers at the Oxford Supercomputing
CHAPTER 4. SIMULATION RESULTS

Figure 4.6: The flash crash of the Dow Jones Industrial Average price and the following recovery. The white curve is the price and the crash happens between 14:32 and 14:47. C.f. Fig. 4.3 and Fig. 4.5. From Ref. [2].

Centre were needed to obtain the results presented in this Section. The computational resources also determined the number of realizations: 240 different realizations were simulated for a selected set of parameter values.

Here we focus on the parameter range which yields a $p_M$ dynamics which is not stationary and it is not yet in the crash phase. The way the parameter values were chosen are described in Sec. 4.1.2. We also investigate the effect of different behaviour from the market participant regarding the submission of limit orders. This is achieved by different $\alpha$ parameters. $\alpha$ determines the limit order submission rate at various prices of the book and is defined in Eq. (2.6) as

$$\lambda(i) \equiv \frac{k}{i^\alpha}$$

(4.2)

where $k$ is a model parameter and $i$ is the distance from the best opposite quote. Large $\alpha$ implies that most traders are keen to submit close to the ask/bid prices and only a small fraction of the limit orders are submitted far from these prices. Small $\alpha$ simulates a scenario where traders are more willing to submit limit orders further away from the bid and ask prices (even in this case that prices close to ask/bid are the most likely candidates for limit order submission).

Fig. 4.7 shows the simulation results for the latter case: $\alpha = 0.2$. Moreover, the displayed results belong to values of $s$, introduced in Eqs. (2.10–2.11), which
Figure 4.7: Example time evolution of price and volatility when no crash happens. Limit orders are submitted far from bid/ask prices with relatively high probability ($\alpha = 0.2$). The strength of the feedback, $s$, was week. Average of 240 simulations.

are among the weakest feedback strengths investigated in current thesis. The five different $s$ values are $s \in \{0.1\%, 0.3\%, 0.5\%, 0.7\%, 0.9\%, 1.1\%\}$. The upper panel shows the price and the lower one the empirical volatility, measured at time $t$ as the unbiased estimator of the variance of the different realizations of $p_M(t)$.

We can see that the mid-price slowly decreases in all cases except the two with weakest feedback. The volatility first decreases for all $s$ values until the transitional period is over and the order book has “forgot” the initial condition. After that the volatility slowly increases. Remarkable that in this case stronger the feedback is, lower the volatility is: for $s = 1.1\%$ it is between 30-40 ticks after the initial period is over but remains above 60 ticks for the smallest investigated $s$ values. Our interpretation of this results is that any feedback drives the system towards a more deterministic model and hence leads to lower volatility.

This does not change when the $\alpha$ parameters is increased slightly to $\alpha = 0.3$. The pattern in Fig. 4.8 is very similar to that of Fig. 4.7. The main differences are that
the different $s$ cases are more separated from each other and the volatility seems to be lower for the $\alpha = 0.3$ simulations than for $\alpha = 0.2$ for the same $s$ parameter and time values.

This, however, changes if we keep $\alpha$ fixed at $\alpha = 0.3$ and allow $s$ to have somewhat higher values: $s \in \{1.4\%, 1.6\%, 1.8\%, 2.0\%\}$. As Fig. 4.9 shows, these level of feedback leads to almost indistinguishable price and volatility values for the initial period. The observed difference towards the end of the simulated time (measured as a number of orders processed) is originating from early emergence of a crash.

Fig. 4.10 shows a similar picture but for $\alpha = 0.5$ and a wider range of $s$ values. The largest $s$ ($s = 4.5\%$), implies strong enough feedback in this case that the mid-price suffers a “slow crash”.

Increasing $\alpha$ further alters the dynamics significantly. When $\alpha = 1$ most of the limit orders are submitted very close to the ask/ask prices and this may interfere with the market order submission driven mainly by the feedback process. This can be seen
4.2. STATISTICAL OBSERVATIONS

Figure 4.9: Example time evolution of price and volatility when no crash happens. Limit orders are submitted far from bid/ask prices with relatively high probability ($\alpha = 0.3$). The strength of the feedback, $s$, was in the mid-range. Average of 240 simulations.

in Fig. 4.11 for $s \in \{0.1\%, 0.7\%, 1.3\%, 1.9\%, 4.5\%\}$. The interplay of these two processes causes much higher volatility in the mid-price: it is between 100-300 ticks as it is shown in the lower panel of Fig. 4.11 compared with the range of 30-60 ticks of the previous cases. While such high volatility levels were visible only in the case of a crash (e.g. $s = 4.5\%$ in Fig. 4.10), the mid-price does not show the characteristics of a crash: the overall change in the mid-price is below 200 ticks or even around 100 ticks in the upper panel of Fig. 4.11. This is in contrast with the 500 tick change observed in $p_M$ for $s = 4.5\%$ in Fig. 4.10.

To consolidate all of these we calculate the time average of the volatility and plot as a function of $s$ values for four different $\alpha$ parameters. This gives Fig. 4.12. Here $s \in [0.1\%, 2.0\%]$ with 0.1% increments and $s \in [2\%, 5\%]$ with 0.5% increments. This totals in 26 points on the horizontal axis for each curve. Fig. 4.12 clearly shows that the $\alpha = 1$ case is significantly different from the others. For all $\alpha$ values the volatility
Most limit orders are submitted close to bid/ask prices with some probability of submission farther from the bid/ask prices ($\alpha = 0.5$). The strength of the feedback, $s$, varies from very small to strong. Note that in case of $s = 4.5\%$ a “slow crash” is observed. As the price decreases so is increasing the volatility. Average of 240 simulations.

increases for large $s$ as such a strong feedback already causes crashes. The decrease in the volatility for low $s$ values are indications of the same phenomenon we already observed in Fig. 4.7, namely that the feedback drives the system towards a more deterministic model and hence leads to lower volatility.

### 4.2.1 Simulation and Statistical Noise

In order to gain an insight about how precise these results are from the statistical noise point of view we calculate the time-averaged volatility of the $p_M$ series as above, but for two smaller subsets of the overall number of realizations. This is clearly not a complete in-depth stability analysis. Nevertheless, it does show that our results are stable and provide an estimate about the size of the error.
4.2. STATISTICAL OBSERVATIONS

Figure 4.11: Example time evolution of price and volatility when no crash happens. Most limit orders are submitted close to bid/ask prices with low probability of submission farther from the bid/ask prices ($\alpha = 1$). The strength of the feedback, $s$, varied from very small to strong. Average of 240 simulations.

The reason of not carrying out a more detailed analysis is the computationally intensive nature of this problem. The large parameter space, the non-stationary dynamics which prevents ergodic averaging (i.e., substitute ensemble average with averaging over time is not possible), and the need to simulate long evolution to overcome transient effects together strongly limit the number of simulations even on the supercomputing facilities of OSC.

To circumvent this problem we split the 240 realizations simulated for each $\alpha$ and $s$ parameter pair into two sets each with 120 realizations. Then we calculate the time-averaged volatility for each point in the parameter space and display the results in Fig. 4.13.

Fortunately the two sets indicate that the results are stable and the size of the error is small enough to allow us to make meaningful observations about the trend in the time-averaged volatility as a function of $s$ and $\alpha$. 
Figure 4.12: The time-average of volatility as a function of the strength of the feedback for various limit order submission parameters $\alpha$. Average of 240 simulations.
Figure 4.13: Statistical stability of time-averaged volatility for two different set of seed values, \( A \) and \( B \), used for the pseudo-random number generation in the submission process. Both sets contains 120 distinct realizations. We can see that results are statistically stable even if the statistical error is not yet negligible for this number of samples of the statistics.
Chapter 5

Analytical Results

The authors of Ref. [8] present an analytical solution for the conditional probability of an increase in mid-price in their work. Their formulation is based on a birth-death process and it yields the Laplace transformation of the conditional probability of the increase. In this Chapter we extend these results and derive an analytical expression for the same quantity in the case of the non-stationary time evolution of prices, driven by the model introduced in Sec. 2.2 and analysed with the help of numerical simulation in Chapter 4.

In the first part of this Chapter we review and consolidate the relevant results from the literature of discussing inverse Laplace transform and its application to transitional probabilities [16, 17, 18, 19, 20, 21]. After that we outline the results for Markovian case [8] which is followed by the derivation of the same quantity for our extended model where the time evolution is not stationary.

5.1 Laplace Transform and Its Application to Conditional Probabilities

This Section contains the basic definition of Laplace transform and its inverse as well as a few basic properties of them which we rely on the latter parts of this Chapter. We also review an approximation technique for the inverse Laplace transform together with its relation to continued fractions [16]. The last part of this section shows how Laplace transform can be used for to calculate transition probabilities and why continued fractions and the rational approximation method is particularly useful to tackle such a problems [17, 21, 18, 19, 20].
5.1.1 Definition and Basic Properties

The definition of the Laplace transform of the $\mathbb{R} \to \mathbb{R}$ function $f(t)$ is defined by

$$L_{\mathbb{R}^+} \{ f(t) \} = \int_0^\infty e^{-st} f(t) dt$$

(5.1)

It is assumed that the function $f(t)$ is defined for all positive $t$ in the range $(0, \infty)$, $s$ is real and, most importantly, that the integral is convergent. As the integral runs over the non-negative values this is also called the one-sided Laplace transform. We introduce the notation of $\tilde{f}(s) \equiv L_{\mathbb{R}^+} \{ f(t) \}$ and also of $L \{ f(t) \} \equiv L_{\mathbb{R}^+} \{ f(t) \}$.

We are going to rely on the two-sided Laplace transformation, which is defined as

$$L_{\mathbb{R}} \{ f(t) \} = \int_{-\infty}^{\infty} e^{-st} f(t) dt.$$  

(5.2)

The one-sided transformation of $f(t)$, $L_{\mathbb{R}^+} \{ f(t) \}$, and the two-sided, $L_{\mathbb{R}} \{ f(t) \}$, are closely linked:

$$L_{\mathbb{R}} \{ f(t) \} (s) = \int_{-\infty}^{\infty} e^{-st} f(t) dt = \int_{0}^{\infty} e^{st} f(-t) dt + \int_{0}^{\infty} e^{-st} f(t) dt = L_{\mathbb{R}^+} \{ f(-t) \} (-s) + L_{\mathbb{R}^+} \{ f(t) \} (s)$$

(5.3)

We briefly review those aspects of the Laplace transformation theory which are relevant in present work on. We do so for the one-sided transformation to keep things simpler and short. But all these aspects are valid in the case of the two-sided transformation as well with two important differences. The first is that the connection between the cumulative density function and the probability density function is slightly simpler in case of two-sided transformation. The two-sided Laplace transformation of the derivative of the original $f(t)$ function, $\frac{d}{dt} f(t)$, can be obtained from the two-sided Laplace transformation of $f(t)$ multiplied by $s$:

$$L_{\mathbb{R}} \left\{ \frac{d}{dt} f(t) \right\} = s L_{\mathbb{R}} \{ f(t) \}.$$  

(5.4)

Here we have to made the following assumptions: $f(t)$ is differentiable and continuous; $f(t) \to O(e^{\gamma t})$ as $t \to \infty$; $f'(t)$ is continuous except at a finite number of points.
5.1. LAPLACE TRANSFORM AND ITS APPLICATION TO CONDITIONAL PROBABILITIES

$t_1, t_2, \cdots, t_n$ in any finite interval $[0, T]$. In case of the one-sided transformation the second difference is that the two-sided transformation does not respect causality in contrast with the one-sided version. However, this is not a problem in our case as we are employing this transformation to probability density functions and not time series.

The conditional Laplace transform of the random variable $X$, conditional on the event $A$, is defined as the Laplace transform of the conditional probability density function of $X$ given $A$.

Using the linear property of expectation and Laplace transformation it is easy to see that if $X$ and $Y$ are independent random variables with well-defined Laplace transforms, then

$$
\mathcal{L}\{f_{X+Y}\} = \mathbb{E}[e^{-s(X+Y)}] = \mathbb{E}[e^{-sX}] \mathbb{E}[e^{-sY}] = \mathcal{L}\{f_X\} \cdot \mathcal{L}\{f_Y\} \quad (5.5)
$$

We highlight the following special case of the Laplace transformation as it will be useful in our analysis. The Laplace transform of $f(t) = e^{\alpha t}$:

$$
\mathcal{L}\{e^{\alpha t}\} = \int_{0}^{\infty} e^{-st} e^{\alpha t} dt = \int_{0}^{\infty} e^{-(s-\alpha)t} dt = \frac{1}{s - \alpha} \quad (5.6)
$$

Note that this integral converges provided that $\text{Re}\ s > \alpha$.

The second property establishes a connection between a function and its integrand:

$$
\mathcal{L}\left\{ \int_{0}^{t} f(u) du \right\} = \frac{1}{s} \int_{0}^{\infty} e^{-st} f(t) dt = \frac{\tilde{f}(s)}{s} \quad (5.7)
$$

5.1.2 Inversion of Laplace Transform

Let $\mathcal{L}^{-1}\{\tilde{f}(s)\}$ denote the function whose Laplace transform is $\tilde{f}(s)$. Thus

$$
f(t) = \mathcal{L}^{-1}\{\tilde{f}(s)\} \quad (5.8)
$$

and $f(t)$ is called the inverse transform of $\tilde{f}(s)$. Note that if a function $f_1(t)$ differs from $f(t)$ only at a finite set of values $t_1, t_2, \cdots, t_n$ then

$$
\mathcal{L}\{f_1\}(t) = \mathcal{L}\{f(t)\} \quad (5.9)
$$

so that the inverse transform is not unique. It can be shown that the Laplace transform is unique if $f(t)$ is continuous in the interval $[0, \infty)$. 

39
CHAPTER 5. ANALYTICAL RESULTS

The Bromwich Inversion Theorem  One way to calculate the inverse of a $\tilde{f}(s)$ function is to use the Bromwich inversion theorem \cite{16}. It states that

$$f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{st} \tilde{f}(s) ds. \quad (5.10)$$

While this is a common way to calculate the inverse Laplace transform, the above integral is not particularly practical in our case. The reason for this is that, as we will see, the function in Laplace space, $\tilde{f}(s)$ is not given in closed form, but as a continued fraction.

5.1.3 The Rational Approximation Method

If the Laplace transform $\tilde{f}(s)$ can be expressed as a rational fraction of two polynomial terms, i.e., in the form $P(s)/Q(s)$ where $P(s)$ and $Q(s)$ are polynomials of degree $p$ and $q$ respectively with $p \leq q$

$$\tilde{f}(s) = \frac{P(s)}{Q(s)} \quad (5.11)$$

where

$$P(s) = s^p + a_1 s^{p-1} + \cdots + a_p \quad (5.12)$$

$$Q(s) = s^q + b_1 s^{q-1} + \cdots + b_q \quad (5.13)$$

and $Q(s)$ has $n$ distinct roots $\alpha_k$, $k = 1..n$

$$Q(s) = (s - \alpha_1)(s - \alpha_2) \cdots (s - \alpha_n), \quad (5.14)$$

then $\tilde{f}(s)$ can be written as

$$\tilde{f}(s) = A_0 + \frac{A_1}{s - \alpha_1} + \frac{A_2}{s - \alpha_2} + \cdots + \frac{A_q}{s - \alpha_q}. \quad (5.15)$$

Thus

$$f(t) = A_0 \delta(t) + A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} + \cdots + A_q e^{\alpha_q t}. \quad (5.16)$$

Writing the last two equations in a more general form we get

$$\tilde{f}(s) = \sum_{k=1}^{n} \frac{P(\alpha_k)}{(s - \alpha_k) Q'(\alpha_k)} \quad (5.17)$$

for $\tilde{f}(s)$ where $Q'(\alpha_k)$ is the derivative of $Q(s)$ evaluated at $s = \alpha_k$, i.e., $Q'(s) \equiv \frac{d}{ds}Q(s)\big|_{s=\alpha_k}$. Hence

$$f(t) = \mathcal{L}^{-1}\{\tilde{f}(s)\} = \sum_{k=1}^{n} \frac{P(\alpha_k)}{Q'(\alpha_k)} e^{\alpha_k t}. \quad (5.18)$$
It has been shown in Ref. [22] that the partial fractional decomposition is not
necessary. Assuming for simplicity that \( P(s) = 1 \) yields
\[
\bar{f}(s) = \frac{1}{Q(s)} = \sum_{k=1}^{n} \frac{1}{s - \alpha_k} \frac{1}{Q'(\alpha_k)} \quad (5.19)
\]
and
\[
f(t) = \sum_{k=1}^{n} \frac{1}{Q'(\alpha_k)} e^{\alpha_k t}. \quad (5.20)
\]
Expanding the exponential term in \( f(t) \) as Taylor series leads to
\[
f(t) = \sum_{k=1}^{n} \frac{1}{Q'(\alpha_k)} \sum_{j=0}^{\infty} \frac{\alpha_k^j t^j}{j!} = \sum_{j=0}^{\infty} \frac{t^j}{j!} \left( \sum_{k=1}^{n} \frac{\alpha_k^j}{Q'(\alpha_k)} \right). \quad (5.21)
\]
It is known from the theory of residues that if \( R \) is sufficiently large so that the
circle \( C : |z| = R \) includes all poles of the integrand \( z^k/Q(z) \) then
\[
u_j \equiv \sum_{k=1}^{n} \frac{\alpha_k^j}{Q'(\alpha_k)} = \frac{1}{2\pi i} \oint_{|z|=R} \frac{z^j}{Q(z)} \, dz. \quad (5.22)
\]
Denoting the length of contour \( C \) by \( L \) and defining \( M \equiv \max_C |f(z)| \), the absolute
value of the integrand is limited by \( LM \). More precisely
\[
\left| \oint_C f(z) \, dz \right| \leq LM. \quad (5.23)
\]
Combining this cap for the length of the integral with letting \( R \to \infty \) yields
\[
u_j = 0, \quad j = 0, 1, \ldots, n - 2. \quad (5.24)
\]
For \( j = n - 1 \) we substitute \( z = Re^{i\theta} \) and obtain
\[
u_{n-1} = 1 \quad (5.25)
\]
if \( R \) is let \( R \to \infty \) in this case, too.

For any other values of \( j \) \( u_j \) is calculated by recursion. This can be seen by
observing that if \( \alpha_k \) is a root \( Q(s) = 0 \)
\[
u_j = \sum_{k=1}^{n} \frac{\alpha_k^j}{Q'(\alpha_k)} = -\sum_{k=1}^{n} \frac{b_1 \alpha_k^{j-1} + b_2 \alpha_k^{j-2} + \cdots + b_n \alpha_k^{j-n}}{Q'(\alpha_k)} \quad (5.26)
\]
To extend this result to the general case when \( P(s) \neq 1 \) we incorporate the nominator in the Taylor-series and obtain
\[
f(t) = \sum_{k=1}^{n} \frac{P(\alpha_k)}{Q'(\alpha_k)} e^{\alpha_k t} = \sum_{k=1}^{n} \frac{P(\alpha_k)}{Q'(\alpha_k)} \sum_{j=0}^{\infty} \frac{\alpha_k^j t^j}{j!} = \sum_{j=0}^{\infty} \frac{t^j}{j!} \left( \sum_{k=1}^{n} \frac{\alpha_k^j P(\alpha_k)}{Q'(\alpha_k)} \right) \equiv v_j \quad (5.27)
\]
CHAPTER 5. ANALYTICAL RESULTS

This yields

\[ f(t) = \sum_{j=0}^{\infty} \frac{v_j t^j}{j!} \]  

(5.28)

with \( v_j \) defined in the Eq. (5.27). Expanding the definition of \( v_j \) with that of \( u_j \) we find

\[ v_j = u_{j+l} + a_1 u_{j+l-1} + a_2 u_{j+l-2} + \cdots + a_l u_j \]  

(5.29)

### 5.1.4 Continued Fraction and Padé Approximation

Cohen [16] shows that if a Laplace transform \( \tilde{f}(s) \) is given but it is not a rational function then continued fraction approximation can be used to obtain an \( \hat{f}(s) \approx \tilde{f}(s) \). \( \hat{f}(s) \) is by construction a rational function and the results of the Sec. 5.1.3 can be used to calculate \( f(t) \).

As we will see later in Sec. 5.2 the relevant Laplace transform is already given in our case as a continued fraction. Thus it is a rational function and the results of Sec. 5.1.3 can be applied to it. However, the formalism described in the previous Section relies on a Laplace transform given in the form of \( \tilde{f}(s) = \frac{P(s)}{Q(s)} \) which is different from the usual form of continued fractions,

\[ \tilde{f}(s) = a_0 + \frac{a_1}{b_1 + s + \frac{a_2}{b_2 + s + \frac{a_3}{b_3 + s + \frac{a_4}{b_4 + s + \cdots}}}} \]  

(5.30)

where \( a_n, b_n, n \in \mathbb{N} \) are known parameters.

While the version of Eq. (5.30) with only infinite number of fractions can clearly be transformed to \( P(s)/Q(s) \) form by algebraic steps this way quickly becomes rather labour-intensive. In order to facilitate an easier transformation we introduce a recursive formula and the Padé approximation based on Ref [16].

Let \( g(z) \) be a general \( \mathbb{C} \to \mathbb{C} \) function which has the form

\[ g(z) = \frac{\sum_{k=0}^{\infty} \alpha_1 k z^k}{\sum_{k=0}^{\infty} \alpha_0 k z^k} \]  

(5.31)
5.1. LAPLACE TRANSFORM AND ITS APPLICATION TO CONDITIONAL PROBABILITIES

\[ g(z) = \frac{\alpha_{10}}{\alpha_{00} + z(\alpha_{10}\alpha_{01} - \alpha_{00}\alpha_{11}) + (\alpha_{10}\alpha_{02} - \alpha_{00}\alpha_{12}) z + \cdots} \]

\[ \frac{\alpha_{20}}{\alpha_{00} + z(\alpha_{20}\alpha_{11} - \alpha_{10}\alpha_{21}) + (\alpha_{20}\alpha_{12} - \alpha_{10}\alpha_{22}) z + \cdots} \]

\[ \frac{\alpha_{30}z}{\alpha_{00} + z(\alpha_{30}\alpha_{21} - \alpha_{20}\alpha_{31}) + (\alpha_{30}\alpha_{22} - \alpha_{20}\alpha_{32}) z + \cdots} \]

\[ (5.32) \]

As Eq. (5.32) indicates, \( \alpha_{ij}, i, j = 0, 1, 2, \ldots \) can be computed recursively by the formula

\[ \alpha_{ij} = \alpha_{i-1,0}\alpha_{i-2,j+1} - \alpha_{i-2,0}\alpha_{i-1,j+1}. \]  

(5.33)

Therefore the ratio of two power series is can be expressed with continued fractions and vice versa as it is given above. Of course, the continued fraction form is truncated in practice which implies an approximation for the function \( g(z) \).

The Padé-type approximant links a \( f(z) \) function given by the ratio of two Taylor series, like \( g(z) \) [16, 23]. Let function \( f(z) \) have a Taylor expansion

\[ f(z) = \sum_{k=0}^{\infty} a_k z^k, \]  

(5.34)

and define \( v(z) \), \( w(z) \) respectively by

\[ v(z) = b_0 z^k + b_1 z^{k-1} + \cdots + b_k \]  

(5.35)

\[ w(z) = c_0 z^{k-1} + c_1 z^{k-2} + \cdots + b_{k-1} \]  

(5.36)

with such \( b_k \) and \( c_k \) seriad that

\[ c_i = \sum_{j=0}^{k-i-1} a_j b_{i+j+1}, \quad i = 0, 1, \ldots, k - 1. \]  

(5.37)
Then \( \frac{w(z)}{v(z)} \) is called the \((k - 1/k)\) Padé-type approximation to \( f(z) \) and the approximation error is

\[
\frac{w(z)}{v(z)} - f(z) = O(z^k), \quad z \to 0. \tag{5.38}
\]

### 5.1.5 Laplace Transform of First Passage Times

Abate and Whit showed in their work that the Laplace transform of the distribution of first passage times of a birth-death process can be described with continuous fractions \([17, 21, 18, 19, 20]\). Here we review the method as this provides the underlying modelling of our system.

Let \( T_{i,j} \) be a random variable representing the first passage time from state \( i \) to state \( j \). Such first passage times can easily be expressed in terms of first passage times to neighbouring states; e.g., if \( i < j \), then

\[
T_{i,j} = T_{i,i+1} + T_{i+1,i+2} + \cdots + T_{j-1,j} \tag{5.39}
\]

where the random variables on the right are mutually independent. Let \( f_{i,j} \) be the probability density function of \( T_{i,j} \) and let \( \tilde{f}_{i,j} \) be its Laplace transform, i.e., \( \tilde{f}_{i,j} \equiv \mathcal{L}\{f_{i,j}\} \). Combining this with Eq. (5.39) we obtain

\[
\tilde{f}_{i,j}(s) = \prod_{k=i}^{j-1} \tilde{f}_{k,k+1}(s) \tag{5.40}
\]

for \( i < j \). In other words, in order to compute the \( \tilde{f}_{i,j} \), it suffices to be able to compute the Laplace transform of the first passage time to a neighbouring state.

As the next step we construct the continuous fractions representing the Laplace transform of first passage times down with an infinite state space. Let \( \lambda_i \) and \( \mu_i \) denote the birth and death rates in state \( i \), respectively. By considering the first transition from state \( i \) to state \( i - 1 \) we obtain the recursion

\[
\tilde{f}_{i,i-1}(s) = \frac{\mu_i}{\lambda_i + \mu_i} \cdot \frac{\lambda_i + \mu_i}{\lambda_i + \mu_i + s} + \frac{\lambda_i}{\lambda_i + \mu_i} \cdot \left( \frac{\lambda_i + \mu_i}{\lambda_i + \mu_i + s} \tilde{f}_{i+1,i}(s) \tilde{f}_{i,i-1}(s) \right)
\]

\[
= \frac{\mu_i}{\lambda_i + \mu_i + s} + \frac{\lambda_i \tilde{f}_{i+1,i}(s) \tilde{f}_{i,i-1}(s)}{\lambda_i + \mu_i + s} \tag{5.41}
\]

Solving Eq. (5.41) for \( \tilde{f}_{i,i-1}(s) \) yields

\[
\tilde{f}_{i,i-1}(s) = \frac{\mu_i}{\lambda_i + \mu_i + s - \lambda_i \tilde{f}_{i+1,i}(s)} \tag{5.42}
\]
By introducing second, third, etc. transitions the iteration of the last equation produces

\[ \bar{f}_{i,i-1}(s) = \frac{-1}{\lambda_{i-1}} \Phi^\infty_{k=i} \frac{\lambda_k \mu_k}{\lambda_k + \mu_k + s} \]  

(5.43)

where the last term is denotes the continued fraction and is defined as

\[ \Phi^\infty_{k=i} \frac{-\lambda_{k-1} \mu_k}{\lambda_k + \mu_k + s} \equiv \frac{-\lambda_{i-1} \mu_i}{\lambda_i + \mu_i + s + \frac{-\lambda_i \mu_{i+1}}{\lambda_{i+1} + \mu_{i+1} + s + \frac{-\lambda_{i+1} \mu_{i+2}}{\lambda_{i+2} + \mu_{i+2} + s + \cdots}}} \]  

(5.44)

Eq. (5.43) provides the Laplace transform of the distribution of the first passage times when the a queue of initial length \( i \) empties [17]. In this case the birth-death process implies “birth” as an arrival of a new entity in the queue (i.e., increasing \( i \) by one) and “death” is when the number of entities in the queue decreases because one of them is processed.

Note that the result and its derivation allows for birth and death rates which depends on the state of the system. The index \( k \) in Eq. (5.43) indicates that birth rate \( \lambda_k \) and death rate \( \mu_k \) can be different for different \( k \) states of the queue.

### 5.2 Stationary Time Evolution

In order to apply Eq. (5.43) for modelling of an order book we observe that an entry of the order book at a given price with \( j \) outstanding orders is an analogous to a queueing problem. The rate of order arrival corresponds to the birth rate and the rate they are removed from the order book is a death process (it is not important at this stage whether an order is cancelled or executed as long as it is not on the book anymore). Thus \( \bar{f}_{i,i-1}(s) \) gives the Laplace transform of the distribution of the first time when the number of orders decreased by one at the given price. If this given price is either the ask price, \( p_A \) or the bid price \( p_B \) and if the number of outstanding limit orders of either of these prices reaches zero then \( p_A \) or \( p_B \) will change and thus the mid-price, defined in Eq. (2.3), will change as well.

The general form of \( \bar{f}_i(s) \) in Eq. (5.43) contains rate which depends on the number of outstanding orders. Two rates out of the three introduced in Chapter 2 for the models, specified in Eqs. (2.6–2.7), do not depend on the number of orders outstanding. They, however, depend on how far the given price is from the opposite best quote. The third rate, the cancellation rate does depend on the number of orders at that price both for stationary dynamics, given in Eq. (2.8) and for the non-stationary case in Eqs. (2.10–2.11). This yields the following relationship
where \( S \) denotes the spread given in Eq. (2.4) as
\[ S(t) = p_A(t) - p_B(t). \] (5.47)
Thus Eq. (5.43) becomes
\[
\bar{f}_{i,i-1}^s(s) = \frac{-1}{\lambda(S)} \Phi_{k=i}^{\infty} \frac{\lambda(S) \mu_k}{\lambda(S) + \mu_k + s} 
\] (5.48)
and the continued fraction is given as
\[
\Phi_{k=i}^{\infty} \frac{-\lambda \mu_k}{\lambda + \mu_k + s} \equiv \frac{-\lambda \mu_i}{\lambda + \mu_i + s + \frac{-\lambda \mu_{i+1}}{\lambda + \mu_{i+1} + s + \frac{-\lambda \mu_{i+2}}{\lambda + \mu_{i+2} + s + \cdots}}} \] (5.49)
The dependence of parameters on the spread \( S \) is omitted in the last expression for the sake of brevity. We also do so below, except where it is necessary to indicate the dependence explicitly in order to avoid ambiguity.

Eq. (5.48) corresponds to the transition time from having \( i \) order to \( i - 1 \). In the general case we are interested in the transition time from having \( j \) orders to having no outstanding limit orders on the book at that price. Combining Eqs. (5.40, 5.48) yields
\[
\bar{f}_{j,S}(s) \equiv \bar{f}_{(j,0),S}(s) = \left( -\frac{1}{\lambda(S)} \right)^j \left( \prod_{i=1}^{j} \Phi_{k=i}^{\infty} \frac{-\lambda(S) \mu_k(S)}{\lambda(S) + \mu_k(S) + s} \right) 
\] (5.50)
Evaluating Eq. (5.43) requires the calculating the infinite continued fraction which is usually not feasible. Hence we introduce a finite approximation of the continued fraction by terminating it after \( n \) steps, i.e.,
\[
\Phi_{k=i}^{n} \frac{-\lambda \mu_k}{\lambda + \mu_k + s} \equiv \frac{-\lambda \mu_i}{\lambda + \mu_i + s + \frac{-\lambda \mu_{i+1}}{\lambda + \mu_{i+1} + s + \frac{-\lambda \mu_{i+2}}{\lambda + \mu_{i+2} + s + \cdots}}} \] (5.51)
5.2. STATIONARY TIME EVOLUTION

The $n \to \infty$ limit case is clearly the complete continued fraction:

$$\lim_{n \to \infty} \Phi^n_{k=1} \frac{\lambda \mu_k}{\lambda + \mu_k + s} = \Phi^\infty_{k=1} \frac{-\lambda \mu_k}{\lambda + \mu_k + s}. \quad (5.52)$$

Similarly, the finite approximation of $f_{j,S}(s)$ is given by using the first $n$ terms in the continued fractions and it is denoted by $f^{(n)}_j(s)$ for all $n > j$

$$f^{(n)}_j(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \Phi^n_{k=1} \frac{-\lambda \mu_k}{\lambda + \mu_k + s} \right). \quad (5.53)$$

It is also true that

$$\lim_{n \to \infty} f^{(n)}_j(s) = f_{j,S}(s). \quad (5.54)$$

The dependence of $\mu_k$ on spread $S$ in Eq. (5.46) can be written in a simpler form:

$$\mu_k(S) = \mu + k \cdot \theta(S). \quad (5.55)$$

Combining Eq. (5.51), Eq. (5.53), and Eq. (5.55):

$$f^{(n)}_{j,S}(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \frac{-\lambda (\mu + i \theta)}{\lambda + \mu + i \theta + s + \cdots + \mu + n \theta + s} \right). \quad (5.56)$$

$$f^{(n)}_{j,S}(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \frac{-\lambda (\mu + i \theta)}{\lambda + \mu + i \theta + s + \cdots + \mu + n \theta + s} \right). \quad (5.57)$$

$$f^{(n)}_{j,S}(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \frac{-\lambda (\mu + i \theta)}{\lambda + \mu + i \theta + s + \cdots + \mu + n \theta + s} \right). \quad (5.58)$$

$\tilde{f}_{j,S}$ and its approximation $\tilde{f}^{(n)}_{j,S}$ tell us the Laplace transform of the distribution of the transition time from having $j$ orders at a given price to having none. If the spread is one then the mid-price, $p_M$, changes when the number of orders at either of ask or bid price decreases to zero. Let $a$ ($b$) denote the number of outstanding order at the ask (bid) price, respectively. Thus $\tilde{f}_{a,S}$ ($\tilde{f}_{b,S}$) quantify the distribution of time needed to deplete the orders at ask (bid) price, respectively. Let $T$ be the time of the
first change in $p_M$ measured from any arbitrary starting time $t = 0$. Cont et al. prove in Ref. [8] that the Laplace transformation of the distribution of $p_M$ increasing at $T$ (rather than decreasing) is

$$F_{a,b}(s) = \frac{1}{s} f_a(s)f_b(-s)$$

(5.59)

In the case when the spread is larger than 1, $S > 1$, $p_M$ also changes when a new order submitted at price $i$ such that $p_B < i < p_A$. In this case $F_{a,b}$ is given by [8]

$$F_{a,b}^{(S)}(s) = \frac{1}{s} \left( f_{a,S} (\Lambda_S + s) + \frac{\Lambda}{\Lambda + s} (1 - f_{a,S} (\Lambda_S + s)) \right)$$

$$\times \left( f_{b,S} (\Lambda_S - s) + \frac{\Lambda}{\Lambda - s} (1 - f_{b,S} (\Lambda_S - s)) \right)$$

(5.60)

where $\Lambda_S \equiv \sum_{i=1}^{S-1} \lambda(i)$ and $\lambda(i)$ is the function determining the limit order submission rate at price $i$, defined in Eq. (2.6). The $(s)$ superscript in $F_{a,b}^{(S)}(s)$ indicates that the quantity also depends on the spread.

In order to compute $F_{a,b}(s)$ we approximate it by substituting $f_a(s)$ and $f_b(-s)$ with their respective $n$-step approximations

$$F_{a,b}^n(n,s) = \frac{1}{s} f_a^n(s)f_b^n(-s).$$

(5.61)
5.2. STATIONARY TIME EVOLUTION

Figure 5.2: Evaluating Eq. (5.62) with $n = 2$ (upper panel) and $n = 10$ (lower panel), where $n$ is the number of levels in the continued fractions. Other parameters: $\lambda + \mu + \theta = 2$. $\mu_A = \mu_B = 1.2$, $\lambda = 0.7$, $\theta = 0.1$, $a = b = 1$.

Using the explicit forms of $f^n_a(s)$ and $f^n_b(-s)$, given by Eq. (5.56), into this expression we obtain

$$F_{a,b}(n,s) = \frac{1}{s} \left( -\frac{1}{\lambda} \right)^{a+b} \frac{-\lambda(\mu + i\theta)}{\lambda + \mu + i\theta + s + \frac{-\lambda(\mu + (i+1)\theta)}{\lambda + \mu + (i+1)\theta + s} + \cdots + \frac{-\lambda(\mu + n\theta)}{\lambda + \mu + n\theta + s}} \right) \frac{1}{s} \left( -\frac{1}{\lambda} \right)^{a+b} \frac{-\lambda(\mu + j\theta)}{\lambda + \mu + j\theta - s + \frac{-\lambda(\mu + (j+1)\theta)}{\lambda + \mu + (j+1)\theta - s} + \cdots + \frac{-\lambda(\mu + n\theta)}{\lambda + \mu + n\theta - s}} \right) \right)$$

Equation (5.62)

$F_{a,b}(n,s)$ can be evaluated numerically for a given set of parameters $\lambda$, $\mu$, and $\theta$ and fixed values of $a$, $b$, and $n$. The results of such evaluation is shown in Fig. 5.1 and in Fig. 5.2. It is clearly visible from Fig. 5.2 that as $n$ increases the denominator will have more and more roots which leads to singularities in $F_{a,b}(n,s)$. As we have shown in Sec. 5.1.3 and Sec. 5.1.4 each singularity yields an exponential term in the inverse Laplace transform of $F_{a,b}(s)$. The other feature of the result in Eq. (5.62) is how the market order rate $\mu$ influences the position of these singularities. This is visualized...
in Fig. 5.1. The other two order arrival rates, $\lambda$ and $\theta$, obviously effect it too, but we focus on $\mu$ as this parameter is the one which is changed by the feedback.

### 5.2.1 Inverse Laplace Transform

While the $F_{a,b}(n,s)$ is evaluated numerically, it is also possible to calculate the the same quantity on a purely analytical way. In order to do this, we consider the simplest case: $n = 1$, $a = b = 1$. That is, only the first term in the continued fractions is included and there is only a single outstanding order at the ask and bid price. As above, the spread is also $S = 1$.

In this case Eq. (5.53) simplifies to

$$f_1^{(1)}(s) = \frac{1}{\lambda} \cdot \frac{-\lambda \mu_k}{\lambda + \mu_k + s} = \frac{\mu + \theta}{\lambda + \mu + \theta + s}. \quad (5.63)$$

Thus

$$F_1(s) = \frac{1}{s (\lambda + \mu + \theta + s)} \frac{(\mu + \theta)^2}{(\lambda + \mu + \theta - s)} \quad (5.64)$$

$$F_1(s) = \frac{1}{s (\lambda + \mu + \theta - s)} \frac{(\mu + \theta)^2}{(\lambda + \mu + \theta - s)^2} \quad (5.65)$$

Now we can utilize the results of Sec. 5.1.3 and Sec. 5.1.4. The roots of the denominator of $F_1(s)$ can be read from Eq. (5.64)

$$\alpha_1 = 0 \quad (5.66)$$
$$\alpha_2 = \lambda + \mu + \theta \quad (5.67)$$
$$\alpha_3 = -\alpha_2 = -(\lambda + \mu + \theta) \quad (5.68)$$

Applying Eq. (5.18) the inverse Laplace transform of $F_1(s)$ reads

$$F_1(t) \equiv L_s^{-1}\{F_1(s)\} = \sum_{k=1}^{n} \frac{P(\alpha_k)}{Q'(\alpha_k)} e^{\alpha_k t} = A_0 + e^{\alpha_0 t} + A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} \quad (5.69)$$

Because the denominator of Eq. (5.65) is a third order polynomial with roots at $\alpha_0$, $\alpha_1$, and $\alpha_2$. The root values are given by Eqs. (5.66–5.68).
5.3. NON-STATIONARY TIME EVOLUTION

Calculating the coefficients $A_0$, $A_1$, and $A_2$ yields

\[
A_0 = \frac{(\mu + \theta)^2}{(\lambda + \mu + \theta)^2} \tag{5.70}
\]

\[
A_1 = -\frac{(\mu + \theta)^2}{2(\lambda + \mu + \theta)^2} = \frac{1}{2}A_0 \tag{5.71}
\]

\[
A_2 = \frac{(\mu + \theta)^2}{2(\lambda + \mu + \theta)^2} = \frac{1}{2}A_0 \tag{5.72}
\]

Substituting these values into Eq. (5.69) we obtain the inverse Laplace transform of $F_1(s)$

\[
F_1(t) = A_0 + e^{\alpha_0 t} + A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t}
\]

\[
= \frac{(\mu + \theta)^2}{(\lambda + \mu + \theta)^2} - \frac{1}{2} \frac{(\mu + \theta)^2}{(\lambda + \mu + \theta)^2} e^{-(\lambda + \mu + \theta)t} \left[ 2 - e^{-(\lambda + \mu + \theta)t} + e^{(\lambda + \mu + \theta)t} \right] \tag{5.73}
\]

\[
= \frac{(\mu + \theta)^2}{(2\lambda + \mu + \theta)^2} \left[ \sinh((\lambda + \mu + \theta)t) + 1 \right]. \tag{5.74}
\]

5.3 Non-Stationary Time Evolution

After deriving an analytical result for the stationary time evolution we apply the same technique to the non-stationary case. As described in Sec. 2.2, the sole difference between the two cases from the point of view of the model is that market order submission rate depends on the state of the order book and it can have different values for ask and bid prices.

The feedback is introduced by changing the rate of market orders $\mu$. The market order rate for sell, $\mu_A$, and buy orders, $\mu_B$, are determined by the total number of outstanding orders as defined in Eqs. (2.10–2.11) and repeated here for convenience:

\[
\mu_A \equiv \max \left( \mu_0, s \cdot \sum_{p=1}^{n} x_p \right) \tag{5.75}
\]

\[
\mu_B \equiv \max \left( \mu_0, s \cdot \sum_{p=1}^{n} (-x_p) \right). \tag{5.76}
\]

Here $A$, $B$ denote ask and bid sides, respectively. The overall death rate of is given by the sum of market order rate and the cancellation rate. We need to distinguish
the two sides as these rates are different as well in this case

\[ \mu_{k,A} = \mu_A + k\theta(S) \]  
\[ \mu_{k,B} = \mu_B + k\theta(S) \]  

(5.77)  
(5.78)

\( \bar{f}_{j,S}(s) \), defined in Eq. (5.50), will also be different for the sell and buy sides. They read

\[ \bar{f}_{j,S,A}(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \Phi_{k=i}^{\infty} \frac{-\lambda \mu_{k,A}}{\lambda + \mu_{k,A} + s} \right) \]  
\[ \bar{f}_{j,S,B}(s) = \left( -\frac{1}{\lambda} \right)^j \left( \prod_{i=1}^j \Phi_{k=i}^{\infty} \frac{-\lambda \mu_{k,B}}{\lambda + \mu_{k,B} + s} \right) \]  

(5.79)  
(5.80)

Note that in order for Eqs. (5.79–5.80) to be valid we need to assume that \( \mu_{k,A} \) and \( \mu_{k,B} \) do not change while the number of outstanding orders at the ask or bid price decreases. As the market order rates depend on the imbalance this is not strictly true. However, so far the overall number of sell (buy) orders is much bigger than the number of orders at the ask (bid) price, respectively, it is safe to make this assumption. Hence the validity of the following results is based on this assumption being a reasonable one.

Repeating the steps of taking the first \( n \) level of the continued fraction, presented in the Sec. 5.2, we define analogously \( \bar{f}_{j,S,A}(s) \) and \( \bar{f}_{j,S,B}(s) \) to Eq. (5.53). The only difference from the previous forms is the different value of \( \mu \) for the buy and sell sides. \( F_{a,b}^{(S)}(s) \) and its \( n \)-step approximation \( F_{a,b}^{(S)}(n,s) \) are modified only by substituting the corresponding side-dependent \( \bar{f}_{j}(s) \) values. Accordingly, \( F_{a,b}^{(S)}(s) \) becomes

\[ F_{a,b}^{(S)}(s) = \frac{1}{s} \left( \bar{f}_{a,S,A}(\Lambda_S + s) + \frac{\Lambda}{\Lambda + s} (1 - \bar{f}_{a,S,A}(\Lambda_S + s)) \right) \left( \bar{f}_{b,S,B}(\Lambda_S - s) + \frac{\Lambda}{\Lambda - s} (1 - \bar{f}_{b,S,B}(\Lambda_S - s)) \right) \]  

(5.81)

In the case when the spread is only one tick

\[ F_{a,b}(s) = \frac{1}{s} \bar{f}_{a,S,A}(s) \bar{f}_{b,S,B}(-s). \]  

(5.82)

It is important to distinguish the different meaning of lower case \( a \) (\( b \)) and upper case \( A \) (\( B \)). Former denotes the number of outstanding orders at the ask (bid) prices, while upper case \( A \) (\( B \)) denotes the ask (bid) side and correspondingly the potentially different parameters applicable to the different sides. In particular, \( A \) indicates that
Figure 5.3: Evaluating Eq. (5.84) for different parameter values. Black curve: $\mu_A = \mu_B = 1.2$. Blue curve: $\mu_A \neq \mu_B$: $\mu_A = 1.2$ and $\mu_B = 2.4$. $\lambda + \mu + \theta = 2$. $\lambda = 0.7$, $\theta = 0.1$, $a = b = 1$.

Figure 5.4: Evaluating Eq. (5.83) with $n = 2$ (upper panel) and $n = 10$ (lower panel), where $n$ is the number of levels in the continued fractions. $\mu_A = 1.2$, $\mu_B = 2.4$, $\lambda = 0.7$, $\theta = 0.1$, $a = b = 1$. 
market order rate $\mu_A$ is used in calculating $f_{j,S,A}^{(n)}(s)$ and $B$ shows that $f_{j,S,B}^{(n)}(s)$ is dependent on $\mu_B$.

Note that the Eq. (5.60) and thus Eq. (5.82) assume that $\bar{f}_{a,S,A}(s)$ and $\bar{f}_{b,S,B}(s)$ are independent. This is true in the case of Eq. (5.60), but not for Eq. (5.82): $\bar{f}_{b,S,B}(s)$ depends on $\mu_{k,B}$ which includes all outstanding orders both sell and buy, by its definition Eq. (2.11). Analogously, $\bar{f}_{a,S,A}(s)$ depends on all order, too. Thus they are not independent. However, when there are much more orders not at the ask and bid prices it is a good approximation to treat them as independent. Because most of the orders are elsewhere the change in the number of orders at the bid or ask price barely effects $\mu_{k,B}$ and $\mu_{k,A}$. Thus $\bar{f}_{a,S,A}(s)$ and $\bar{f}_{b,S,B}(s)$ can be approximated to be independent.

Using the n-step approximation version of Eq. (5.82) and writing out the RHS we get

$$F_{a,b}(n,s) = \frac{1}{s} \left(-\frac{1}{\lambda}\right)^{a+b} \prod_{i=1}^{a} \frac{-\lambda (\mu_A + i\theta)}{\lambda + \mu_A + i\theta - s + \lambda (\mu_A + n\theta)} \cdot \prod_{j=1}^{b} \frac{-\lambda (\mu_B + j\theta)}{\lambda + \mu_B + j\theta - s + \lambda (\mu_B + n\theta)}$$

The calculation of $F_{a,b}(n,s)$ is possible numerically for fixed parameter values. To obtain an insight how different $F_{a,b}(n,s)$ is from its version in the case of stationary time evolution we numerically evaluated and compare the two cases in Fig. 5.3. The black curve corresponds to stationary time evolution and the blue one depict $F_{a,b}(n,s)$ for a fixed $n$ in the case of non-stationary dynamics. The main difference is that the black curve is clearly an odd function (i.e., symmetric to the origin), the blue curve does not maintain this symmetry. We remind that the market order submission rates are the same for both sell and buy side in the case of the black curve (stationary dynamics) but they are not symmetric for the non-stationary case. Our result shows that this difference translates directly into a significantly different dynamics. This is in line with what we have aimed for and what was shown by the numerical simulation of the model presented in Chapter 4.
In order to see the effect of a better approximation of \( F_{a,b}(s) \) we evaluate \( F_{a,b}(n, s) \) for two different cases: for smaller and larger \( n \). The results are displayed in Fig. 5.4 and they are in line what we have seen for the stationary dynamics in Fig. 5.2: larger \( n \) introduces more roots in the denominator of \( F_{a,b}(n, s) \) so we observe more singularities in the figure. According to the results in Sec. 5.1.3 this will yield more exponential terms in the series expansion of \( F_{a,b}(t) \) and will provide a more precise result.

Note that following the same logic outlined in the derivation above side dependent limit order submission rates \( \lambda_A, \lambda_B \) and cancellation rates \( \theta_A, \theta_B \) can in principle also be introduced and the model can be extended further. This, however, is out of the scope of present work.

### 5.3.1 Inverse Laplace Transform

Similarly to the symmetric case, \( F_1 \equiv F_{a=1,b=1}^{(1)}(n, s) \) simplifies significantly when \( n = 1, a = b = 1 \). In particular, it becomes a product of three fractions

\[
F_1(s) = \frac{1}{s} \cdot \left( -\frac{1}{\lambda} \right) \frac{-\lambda (\mu_A + \theta)}{\lambda + \mu_A + \theta + s} \cdot \left( -\frac{1}{\lambda} \right) \frac{-\lambda (\mu_B + \theta)}{\lambda + \mu_B + \theta - s} = \frac{1}{s (\lambda + \mu_A + \theta + s) (\lambda + \mu_B + \theta - s)}
\]  

(5.84)

It is straightforward to determine the roots of the denominator from the above form. The three roots are

\[
\alpha_1 = 0 \\
\alpha_2 = \lambda + \mu_B + \theta \\
\alpha_3 = -(\lambda + \mu_A + \theta)
\]  

(5.85–5.87)

Upon reflecting on this results it is clear that in the case \( \mu_A = \mu_B \) we get back the stationary case. Also the source of the non-symmetric nature of \( F_1(s) \) shown in Fig. 5.3 is explained by Eqs. (5.85–5.87).

Following the same steps as for the stationary case, the inverse Laplace transform of \( F_1(s) \) is calculated using Eq. (5.18). The general form of the result is a sum of three exponential terms in this case as well, but the roots and coefficients are different

\[
F_1(t) = A_0 + e^{\alpha_1 t} + A_1 e^{\alpha_2 t} + A_2 e^{\alpha_3 t}.
\]  

(5.88)
We get

\[ A_0 = -A_1 - A_2 \] (5.89)

\[ A_1 = \frac{-\left(\mu_A + \theta\right)\left(\mu_B + \theta\right)}{2(\lambda + \theta)^2 + (3\mu_A + \mu_B - 1)(\lambda + \theta) + \mu_A^2 - \mu_B + \mu_A\mu_B + \frac{(\lambda + \mu_A + \theta)^2}{(\lambda + \mu_B + \theta)}} \] (5.90)

\[ A_2 = \frac{-\left(\mu_A + \theta\right)\left(\mu_B + \theta\right)}{2(\lambda + \theta)^2 + (\mu_A + 3\mu_B + 1)(\lambda + \theta) + \mu_A + \mu_B^2 + \mu_A\mu_B - \frac{(\lambda + \mu_B + \theta)^2}{(\lambda + \mu_A + \theta)}} \] (5.91)

for the coefficients \( A_0 \), \( A_1 \), and \( A_2 \). To obtain a better overview of the structure of these coefficients we write them in the form

\[ A_0 = -A_1 - A_2 \] (5.92)

\[ A_1 = \frac{-\mu_1\mu_2}{\lambda_1\lambda_2 + \lambda_1^2 - \lambda_2 + \lambda_2^2/\lambda_2} \] (5.93)

\[ A_2 = \frac{-\mu_1\mu_2}{\lambda_1\lambda_2 + \lambda_2^2 + \lambda_1 - \lambda_2^2/\lambda_1} \] (5.94)

where we introduced the following shorthand notations

\[ \mu_1 = \mu_A + \theta \] (5.95)

\[ \mu_2 = \mu_B + \theta \] (5.96)

\[ \lambda_1 = \lambda + \mu_A + \theta \] (5.97)

\[ \lambda_2 = \lambda + \mu_B + \theta \] (5.98)

Thus the inverse Laplace transform of \( F_1(s) \) reads

\[ F_1(t) = A_0 + e^{\alpha_0 t} + A_1 e^{\alpha_1 t} + A_2 e^{\alpha_2 t} \]

\[ = -A_1 - A_2 \]

\[ + A_1 \exp\left[-(\lambda + \mu_A + \theta) t\right] \]

\[ + A_2 \exp\left[(\lambda + \mu_B + \theta) t\right] \] (5.99)

with the coefficients \( A_0 \), \( A_1 \), and \( A_2 \) given in Eqs. (5.89–5.91).
Chapter 6

Summary

In this thesis we have developed a model for a simple order book and for the price evolution of an instrument traded on the exchange. The model is capable of exhibiting both stationary and also highly non-stationary dynamics in the price of the instrument, including sudden and fast changes. At the same time the model is simple enough that analytical results for the probability of the increase of the price can be derived. This is achieved by extending the model of Cont et al. [8] via introducing a feedback into the model. The feedback is included through allowing the market order submission rate to be dependent on the imbalance of the order book. The imbalance is defined as the difference of the total number of sell orders and the total number of buy orders on the book.

We developed a versatile volume-based numerical simulation framework for this model. This versatility makes possible to change the order processing rules to a great extent and also to simulate different order generation and submission processes. Moreover, the implementation of this framework proved fast enough to process large number of orders within a few of hours on a modern workstation.

Using this simulation methodology we have shown that the feedback introduced does lead to sudden, fast, and large changes (crashes) in the price. The crashes can equally have downward or upward direction, which is a phenomenon observed in practice on real exchanges. Moreover, our simulation yields a crash in the price which is very similar to the real dynamics of Dow Jones Industrial Average during the flash crash.

We have also investigated the effect of imbalance dependent feedback on order books typical at exchanges using first-in-first-out (FIFO) order processing as well as at exchanges where the orders are processed on a pro-rata basis and observed that the volatility of the price is less sensitive to the feedback in the latter case than in the former case.
CHAPTER 6. SUMMARY

Besides the numerical simulation we derived an analytical expression to calculate the probability of the increase of the price. The result shows that the probability of a price increase depends on the order arrival rates, on the size of spread, and on the number of outstanding orders at the ask and bid prices.

While we strongly believe that present work is an important step towards creating a comprehensive model of sudden price changes on the markets we also realise that much more work is needed on the field. The simulation framework can be used to analyse a wider range of phenomena than the scope of this thesis allowed. To highlight a few, for example, the in-depth understanding of the effect of scaling parameter on the volatility is an obvious candidate for further analysis. Another possibility is to investigate the differences of markets with FIFO order processing and of pro-rata based execution. The analytical part can also be extended e.g. to explain the observed price crashes within this model.

We sincerely hope that there will be opportunities in the future to build further on the foundation laid down in this work and to learn more about this exciting field.
## Appendix A

### Simulation Parameters

<table>
<thead>
<tr>
<th>Description</th>
<th>Symbol</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limit order rate – coefficient</td>
<td>$k$</td>
<td>1.92</td>
</tr>
<tr>
<td>Limit order rate – power</td>
<td>$\alpha$</td>
<td>.5</td>
</tr>
<tr>
<td>Market order rate</td>
<td>$\mu$</td>
<td>.42</td>
</tr>
<tr>
<td>Cancellation order rate – baseline</td>
<td>$\theta_0$</td>
<td>.47</td>
</tr>
<tr>
<td>Cancellation order rate – vertical scale</td>
<td>$\theta_1$</td>
<td>.34</td>
</tr>
<tr>
<td>Cancellation order rate – horizontal scale</td>
<td>$\theta_2$</td>
<td>$\frac{1}{3}$</td>
</tr>
<tr>
<td>Price range</td>
<td>-</td>
<td>$10^4$ ticks</td>
</tr>
<tr>
<td>Initial spread</td>
<td>-</td>
<td>100 ticks</td>
</tr>
<tr>
<td>Initial price</td>
<td>-</td>
<td>@5000 ticks</td>
</tr>
<tr>
<td>Simulation length</td>
<td>-</td>
<td>$10^7$ (orders submitted and processed)</td>
</tr>
<tr>
<td>Scaling weight of imbalance</td>
<td>$s$</td>
<td>.6</td>
</tr>
</tbody>
</table>
Bibliography


