TECHNISCHE UNIVERSITÄT WIEN

DIPLOMARBEIT

Comparison of Limit Order Book Models

ausgeführt am

Institut für
Stochastik und Wirtschaftsmathematik
TU Wien

unter der Anleitung von

Associate Prof. Dipl.-Ing. Dr.techn. Friedrich Hubalek

durch

Julian Feurhuber
Matrikelnummer: 01320081

Wien, am Datum

Kurzfassung

In den letzten Jahren haben fast alle großen Börsen der Welt auf elektronische Orderbücher umgestellt. Diese speichern eingehende Limit Orders automatisch und führen diese gegen den bestmöglichen Preis aus. Durch diese Veränderung haben sich auch die Modelle der Orderbücher geändert. Diese Arbeit befasst sich mit der Analyse und Beschreibung von verschiedenen Modellen der Orderbücher. In letzter Zeit sind sehr viele Publikationen zu diesem Thema veröffentlicht worden. Ziel dieser Diplomarbeit ist es daher verschiedene Varianten der Modelle zu erklären und diese zu vergleichen. Aufgrund der großen Menge an verschiedenen Ansätzen wird in dieser Arbeit der Fokus auf drei verschiedene gelegt. Der erste beruht auf der Theorie der Markov-Prozesse und auf diesem Ansatz bauen die ersten drei Modelle der Arbeit auf. Danach wird ein Modell präsentiert was das Orderbuch als Warteschlange simuliert. Die letzten Modelle, die behandelt werden, verwenden Hawkes Prozesse, um die Orders zu beschreiben.

Abstract

In recent years, almost all major exchanges in the world have switched to electronic order books. These automatically store incoming limit orders and execute them against the best possible price. Due to this change, the models of order books have also changed. This thesis deals with the analysis and description of different order book models. Recently, many papers have been published on this topic. Therefore, the aim of this thesis is to explain and compare different variants of these models. Because of the large number of different approaches, this thesis will focus on three different concepts. The first one is based on the theory of Markov processes and on this approach the first three models of the thesis are built. After that, a model is presented which simulates the order book as a queueing system. The last models discussed use Hawke's processes to describe the orders.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am Datum

Julian Feurhuber

Contents

1 Introduction

In this first chapter some required definitions are going to be repeated. Most of the results are from the theory of stochastic processes which are necessary for further understanding of this work. However theorems and proofs are dispensed but can be looked up in the literature as Øksendal [14] and Bass [3]. In the second part of the introduction the limit order book in general will be explained.

1.1 Mathematical Background

For the whole thesis let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Furthermore let $T \subset \bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ be an arbitrary non-empty index set for the time parameter. The next few definitions are mostly known but are needed quite often therefore this short review should help to recall them.

Definition 1.1.1. *A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ of $\mathbb{F}$ is defined as an increasing collection of sub-σ-algebras of $\mathcal{F}$, such that $\mathcal{F}_t \subset \mathcal{F}$ for each t and $\mathcal{F}_s \subset \mathcal{F}_t$ if $s < t$.*

Definition 1.1.2. *Let $(U,\mathcal{U})$ be a measurable space. A process $X : T \times \Omega \to U$ is called $\mathcal{F}$-progressive if for every $t \in T$ its restriction to $\{s \in T|s \leq t\} \times \Omega$ is $\mathcal{B}(\{s \in T|s \leq t\}) \otimes \mathcal{F}_t$-measurable.*

Definition 1.1.3. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and T an index set with $T \in \{\mathbb{N},\ \mathbb{R}_+\}$. Then a real-valued stochastic process is a collection of random variables $\{X_t : t \in T\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with index set T. So, to every $t \in T$ corresponds a random variable:*

$$\begin{aligned} X_t &: \Omega \to \mathbb{R} \\ (t, w) &\mapsto X_t(\omega). \end{aligned}$$

Definition 1.1.4. *Let $(S,\mathcal{S})$ be a metric space, $T \subset \mathbb{R}$ and $X : T \times \Omega \to S$ a stochastic process. If X is right-continuous and for every $\omega \in \Omega$ the path $X(\cdot,\omega) \to S$ has left hand side limits for every $t \in T \setminus \{\inf T\}$*

The same is true if the process is left-continuous except the then right hand side limits have to be for every $t \in T \setminus \{\sup T\}$ The consecutive definitions are about stochastic processes which are later on used to model incoming orders or cancellations.

Definition 1.1.5. *Let $(U,\mathcal{U})$ be a measurable space. An $\mathcal{F}$-adapted U-valued process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbb{P})$ is called a Markov process with respect to the filtration $\mathcal{F}$ if for all $t \leq s$ in T the random variable X_s is conditionally independent of $\mathcal{F}_t$ given X_t.*

Therefore a Markovian process at a state depends only on the outcome of the previous stage and not on the ones it was before.

Definition 1.1.6. *Let $\mathbb{F}$ be a filtration. A Poisson process with parameter $\lambda > 0$ is a stochastic process X satisfying the properties:*

- $X_0 = 0$ *a.s.*
- *The paths of X_t are right continuous with left hand side limits*
- *If $s < t$ then $X_t - X_s$ is a Poisson random variable with parameter $\lambda(t-s)$*
- *If $s < t$ then $X_t - X_s$ is independent of $\mathcal{F}_s$*

Definition 1.1.7. *A Hawkes process X_t is a point process whose conditional intensity can be written as*

$$\lambda(t) = \mu(t) + \int_{-\infty}^{t} v(t-s)dX_s$$

where $v : \mathbb{R}^+ \to \mathbb{R}^+$ is a kernel function which describes the influence of past events on the intensity process $\lambda(t)$ and $\mu(t)$ is a deterministic function.

The upcoming definitions helps to specify Markov processes in terms of their behavior at each point. This is done with the infinitesimal generator which is a linear operator which is unbounded in general.

Definition 1.1.8. *For a Markov process $(X_t)_{t\geq 0}$ the generator $\mathcal{L}$ defined as*

$$\mathcal{L}f(x) := \lim_{t\to 0} \frac{\mathbb{E}^x(f(X_t)) - f(x)}{t}$$

whenever the limit exists in $(C_\infty, ||\cdot||_\infty)$. Here $\mathbb{E}^x$ denotes the expectation with respect to the semi-group of $(X_t)_{t\geq 0}$.

1.2 Limit Order Book

In stock exchanges nowadays orders are matched and executed by computer systems. The unmatched orders are stored in the so called limit order book. Each limit order book has two sides, the ask side and the bid side. The ask side represents sell limit order and the bid side buy limit orders. There exist two different kind of orders in the market. Those follow the upcoming terms

- Market order: An order to buy or sell a certain amount of a stock for the best available price
- Limit order: An order to buy or sell a certain amount of a stock for a specified price

Market orders are matched immediately at the best available price whereas limit orders are stored in the book if it is not executable at the arrival. A limit order is then either executed by a market order or is canceled after some time by the issuer. Otherwise it will remain in the limit order book. The best quote on the ask side is referred as best ask price whereas the best quote on the bid side is called best bid price. In these circumstances best means for each side something different. The best ask price is therefore the lowest price for which a stock can be paid. On the bid side the best quote means it is the highest prices someone is willing to pay for a stock. In further consequence the best ask price at time t will denoted by $P_a(t)$ and the best bid price at t by $P_b(t)$. The minimum price movement in a limit order book is called tick size and is denoted by δ_x. All orders have to arrive with a price specified to the tick size. The difference between best bid and ask price is called the spread. For further information about limit order books can be recommended the paper "Limit order books" [6].

2 Empirical Illustrations

In this section some illustrations of a real limit order book are shown. The data for those illustration is provided by LOBSTER which is an online limit order book data tool and provides high-quality limit order book data. In these example plots data of the Google stock from the NASDAQ stock exchange is used. The data contains 10 levels which refers to the first ten non empty entries standing in the order book. So that means that the distance between two levels in the data is the minimum tick size. Simply because if there is no volume standing at a certain price it does not count as level and it is therefore not in the data. For the illustration of the price dynamics during a day a whole day is used for a better understanding how prices change during a day. To show the standing volume in the limit order book a certain point during the same trading day as before is chosen randomly. The day which was chosen to illustrate the real data is the 21.05.2012. The first figure shows the best ask price and also the best bid price evolution during a trading day. The plot already suggests that it is pretty difficult model those dynamics.

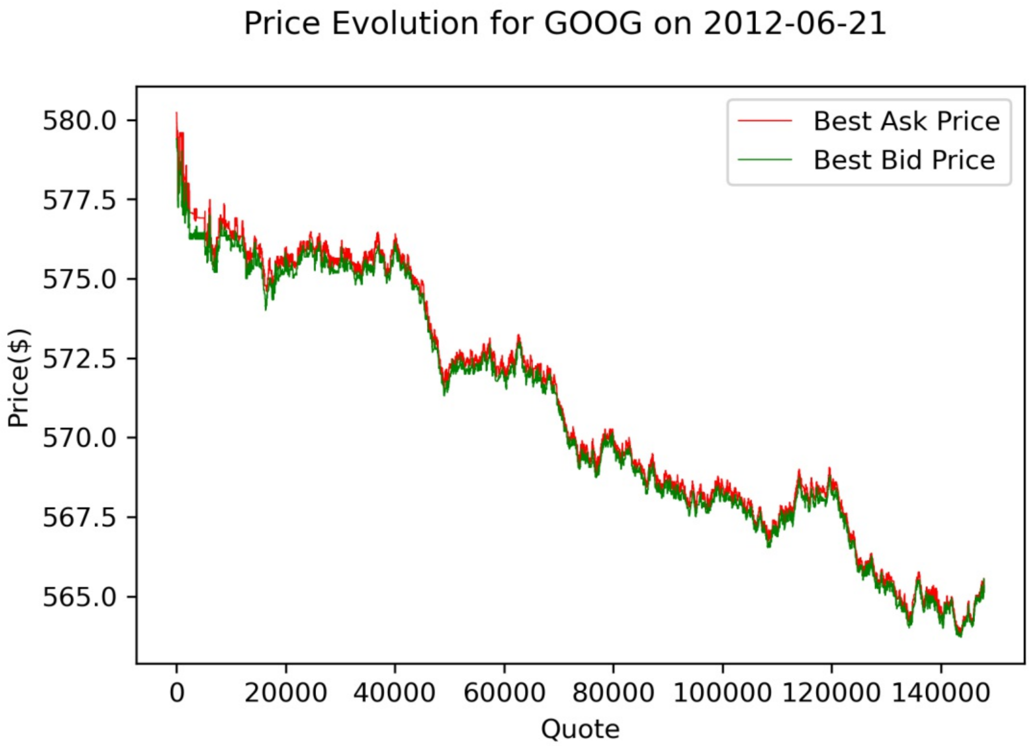

Figure 2.1: Price Dynamics

The x-axis needs further explanation because it shows the number of events taking place during the trading day and the trading day is limited from 09:30 to 16:00. To be register as an event in the data counts for example a submission of a new limit order, a market order, a cancellation of a standing order and executions of limit orders. The spread in this example varies during the day from 0.01$ to 2.30$.
The next figure shows the volume standing in the limit order book. The y-axis is the number of shares standing in the limit order book at each price. As already mentioned the first ten levels are shown if there non empty. Here it gets clearer that the levels are not separated by the tick size. For example here at the best ask price is 575.46$ and the volume at this price is 100.

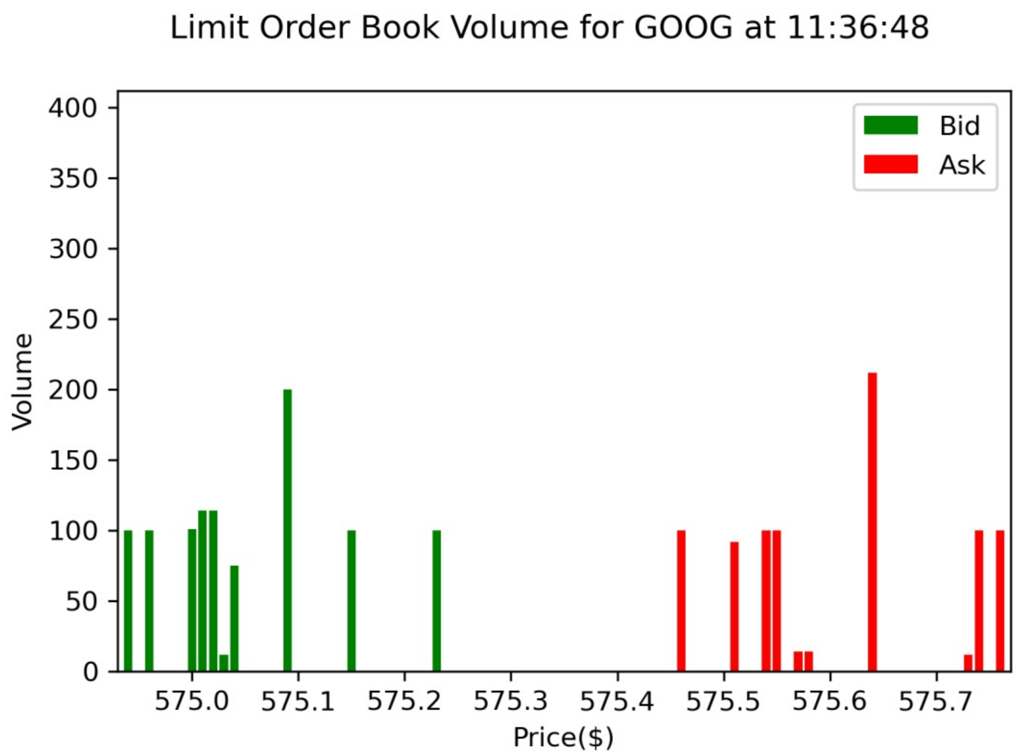

Figure 2.2: Standing volume

The last plot in this section shows the depth of the limit order book. The depth is of great importance for market participants because it is a measure for liquidity in the market. The plot describes on which level what percentage of the total volume in the order book is present. Therefore the y-axis shows the percentage of the total volume and the x-axis represents the ten levels.

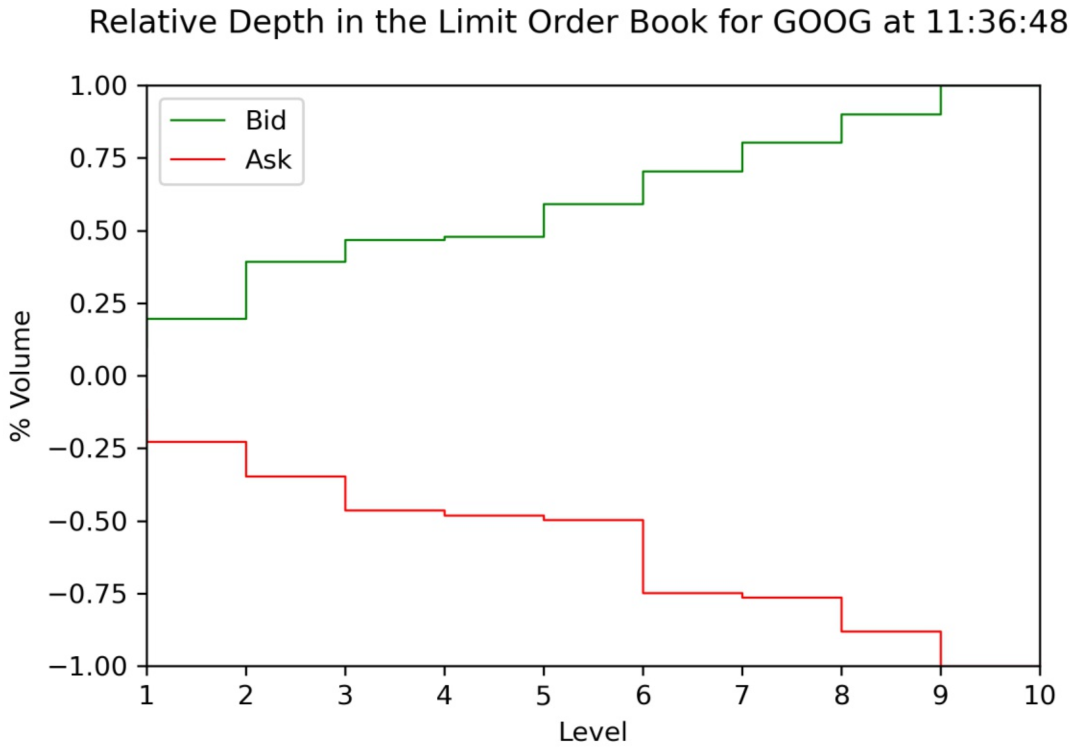

Figure 2.3: Depth in the limit order book

3 Markovian Limit Order Book Models

3.1 Zero-Intelligence Model

3.1.1 Introduction

The zero-intelligence model uses a general mathematical framework to study properties of limit order books in a Markovian context. For financial markets participants it is of great interest to understand the price dynamics. In this relatively simple model some conclusion about those dynamics can be made. The order flows are described by Markovian point processes. With the help of this assumption it is possible to derive several mathematical results considering independent Poisson arrival times. The model also shows that the cancellation rate cannot be neglected. The price process which is caused by the limit order book dynamics convergence to a diffusive process at visible time scales. The model largely follows the presentation in the book "Limit Order Books" [1, Section 6].

3.1.2 Order Book Dynamics

In the setup for this model the ask and bid side of the limit order book are described by a finite number of limits N, this means the opposite best quote can be 1 to N ticks away. The ticks away the difference from the opposite best price to the current price of the order. Whereby the best price for the ask side is described by P_a and the best price on the bid size is given by P_b. The amount of available orders at each tick away at time t is described by

$$(\boldsymbol{v}_a(t); \boldsymbol{v}_b(t)) := (v_{a_1}(t), \dots v_{a_N}(t); v_{b_1}(t), \dots, v_{b_N}(t)),$$

where $\boldsymbol{v}_a(t) := (v_{a_1}(t), \dots, v_{a_N}(t))$ belongs to the ask side of the limit order book and $\boldsymbol{v}_b(t) := (v_{b_1}(t), \dots, v_{b_N}(t))$ to the bid side. Therefore $v_{a_i}(t)$ represents the shares standing in the limit order book on the ask side i ticks away from the opposite best quote. The same applies to the bid side only that it has negative entries. This means in case of a bid limit order the sign of the volume is negative. The quantities in this framework take values in the discrete space $q\mathbb{Z}$, where $q \in \mathbb{N}$ is the minimum order size. The integrated quantities represents the shape of the limit order book and in this model the cumulative depth is presented by

$$V_a(t,i) := \sum_{k=1}^{i} v_{a_k}(t) \qquad \text{and} \qquad V_b(t,i) := \sum_{k=1}^{i} |v_{b_k}(t)|.$$

In the context of this mathematical framework the generalized inverse function of the cumulative market depth is helpful. The inverse function returns for a certain quantity q'

of shares the minimum number of ticks to get q' shares

$$V_a^{-1}(t,q') := \inf\{i : V_a(t,i) > q'\} \qquad \text{and} \qquad V_b^{-1}(t,q') := \inf\{i : V_b(t,i) > q'\}.$$

The functions for the ask and bid side are equal for the first non empty entry

$$V_a^{-1}(t,0) = V_b^{-1}(t,0).$$

If the moving frame of size $2N$ leaves a price the boundary condition sets in. The boundary condition ensures that the number of shares outside the frame is constant set to v_{a_∞} or v_{b_∞} for the ask and bid side respectively.
The agents operating on the market can participate with different actions. They can place market orders or limit orders and also have the possibility to cancel one of their already existing limit order. Those different events are mathematically described by independent Poisson processes.

- $M_{a/b}(t)$: Counting processes of market orders, with constant intensities λ^{M_a} and λ^{M_b}
- $L^i_{a/b}(t)$: Counting processes of limit orders at level i, with constant intensities $\lambda_i^{L_a}$ and $\lambda_i^{L_b}$
- $C^i_{a/b}(t)$: Counting process of cancellations of limit orders at level i, with stochastic intensities $\lambda_i^{C_a} v_{a_i}$ and $\lambda_i^{C_b} |v_{b_i}|$.

The subscript "a" denotes the ask side and "b" the bid side. Therefore for example a sell market order has the subscript "b" because it takes place on the bid side. All different kind of orders are by assumption of unit size q. For the cancellation process at level i the intensity is corresponding to the available amount of orders at that level. The time is indicated by the letter t. Additionally limit orders arrive always below the best quote of the opposite side. This is a realistic approach and does not make the model less valuable. In mathematical terms this means that sell limit orders $L^i_a(t)$ are submitted above $P_b(t)$ and vice versa buy limit orders $L^i_b(t)$ are submitted below $P_b(t)$.

3.1.3 Evolution Of The Order Book

To model the outstanding quantities process in the limit order book coupled stochastic differential equations are used. Depending on which side of the order book is considered different equations arise. The quantities for each side are described by the following equations:

$$
\begin{aligned}
dv_{a_i}(t) = &-\mathbb{1}_{\{v_{a_i}(t)\neq 0\}}\mathbb{1}_{\{V_a(t,i-1)=0\}} q dM_a(t) + qdL_a^i(t) - qdC_a^i(t) \\
&+ (J^{M_b}(\boldsymbol{v}_a, \boldsymbol{v}_b) - \boldsymbol{v}_a)_i dM_b(t) + \sum_{i=1}^{N} (J^{L_b^i}(\boldsymbol{v}_a) - \boldsymbol{v}_a)_i dL_b^i(t) \\
&+ \sum_{i=1}^{N} (J^{C_b^i}(\boldsymbol{v}_a, \boldsymbol{v}_b) - \boldsymbol{v}_a)_i dC_b^i(t),
\end{aligned}
$$

$$
\begin{aligned}
dv_{b_i}(t) = &\mathbb{1}_{\{v_{b_i}(t)\neq 0 \wedge V_b(t,i-1)=0\}} q dM_b(t) - qdL_b^i(t) + qdC_b^i(t) \\
&+ (J^{M_a}(\boldsymbol{v}_a, \boldsymbol{v}_b) - \boldsymbol{v}_b)_i dM_a(t) + \sum_{i=1}^{N} (J^{L_a^i}(\boldsymbol{v}_b) - \boldsymbol{v}_b)_i dL_a^i(t) \\
&+ \sum_{i=1}^{N} (J^{C_a^i}(\boldsymbol{v}_a, \boldsymbol{v}_b) - \boldsymbol{v}_b)_i dC_a^i(t).
\end{aligned}
$$

One should bear in mind that the bid size is modeled in negative terms, therefore the absolute value is needed. The first term of the equations refers to an incoming market order. The indicator function checks whether the entry at level i is empty and if the limit order reservoir is hit. It follows that a market order decreases the available orders by q if at level i is the first non-zero limit. The second term increases the amount of limits in the book through an arriving limit order, whereas the third term causes the opposite because it describes the influence from incoming cancellations. The other three terms describe the influence of orders arriving at the bid side on the ask side of the order book. The same is true for the other way around. The new introduced shift operators J's rearrange the level order by incoming events. These operators are defined in the following form using the inverse functions:

$$
\begin{aligned}
J^{M_b}(\boldsymbol{v}_a, \boldsymbol{v}_b) &= (0, \ldots, 0, v_{a_1}, \ldots, v_{a_{N-k}}) && \text{with } k = V_b^{-1}(t,q) - V_b^{-1}(t,0) \\
J^{M_a}(\boldsymbol{v}_a, \boldsymbol{v}_b) &= (0, \ldots, 0, v_{b_1}, \ldots, v_{b_{N-k}}) && \text{with } k = V_a^{-1}(t,q) - V_a^{-1}(t,0) \\
J^{L_b^i}(\boldsymbol{v}_a) &= (v_{a_{1+k}}, \ldots, v_{a_N}, v_{a_\infty}, \ldots, v_{a_\infty}) && \text{with } k = V_a^{-1}(t,0) - i \\
J^{L_a^i}(\boldsymbol{v}_b) &= (v_{b_{1+k}}, \ldots, v_{b_N}, v_{b_\infty}, \ldots, v_{b_\infty}) && \text{with } k = V_b^{-1}(t,0) - i \\
J^{C_b^i}(\boldsymbol{v}_a, \boldsymbol{v}_b) &= (0, \ldots, 0, v_{a_1}, \ldots, v_{a_{N-k}}) && \text{with } k = V_b^{-1}(t,q) - V_b^{-1}(t,0) \\
J^{C_a^i}(\boldsymbol{v}_a, \boldsymbol{v}_b) &= (0, \ldots, 0, v_{b_1}, \ldots, v_{b_{N-k}}) && \text{with } k = V_a^{-1}(t,q) - V_a^{-1}(t,0).
\end{aligned}
$$

It should be considered that the cancellation of a limit order at the best quote has the same impact on the dynamic as a market order. In the following the concept of the zero-intelligence model will be explained on the basis of an simple example.

3.1.4 Illustration of the Zero-Intelligence Model

The following figures and explanations make the model more understandable and imaginable. Let the boundary condition or the size of the moving frame be denoted by $2N$ and let $N = 10$. Therefore 10 ticks away of the opposite best quote the last volume is standing in the limit order book before the boundary condition sets in. The minimum order size $q = 1$ and the tick size $\delta_x = 1$. Furthermore v_{a_∞} or v_{b_∞} are 4 and -4 respectively. P_a and P_b are standing for the best ask price or rather best bid price. The spread between them is denoted by S. Knowing that the order book can be described. The initial shape of the order book is such that $\boldsymbol{v}_a = (0,0,0,0,0,3,5,2,5)$ and $\boldsymbol{v}_b = (0,0,0,0,0,-1,-4,-3,0,-6)$.

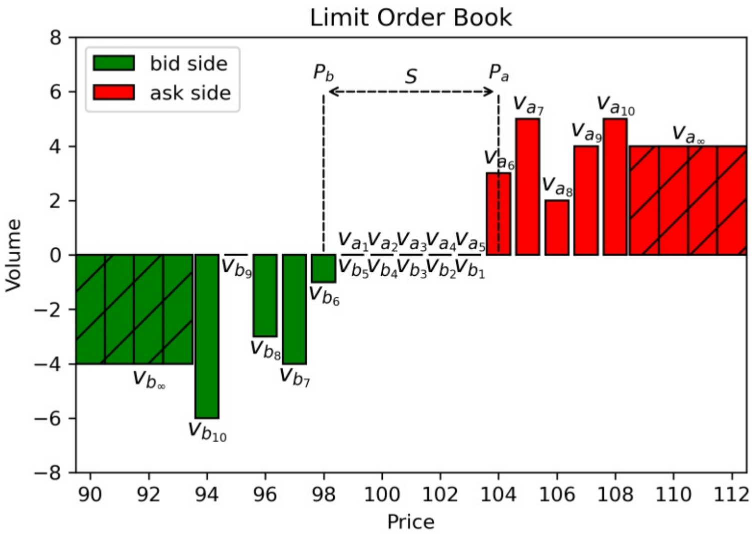

Figure 3.1: Limit Order Book

Assume that a ask limit order comes in. This order is placed at v_{a_3} therefore it is an order placement within the spread. This triggers the shift operator $J^{L_a^i}(\boldsymbol{v}_b)$. After that the new shape of the limit order book is $\boldsymbol{v}_a = (0,0,1,0,0,3,5,2,5)$ and $\boldsymbol{v}_b = (0,0,-1,-4,-3,0,-6,-4,-4,-4)$. Assume that the next order which is submitted is a sell market order. This would activate the shift operator operator $J^{M_b}(\boldsymbol{v}_a, \boldsymbol{v}_b)$. The shape of the order book would become afterwards $\boldsymbol{v}_a = (0,0,0,1,0,0,3,5,2,4)$ and $\boldsymbol{v}_b = (0,0,0,-4,-3,0,-6,-4,-4,-4)$. The last thing occurring is a cancellation at $v_{b_{10}}$. This would just change the bid side and does not have any influence on the ask side. The volume at the bid side would then be given by $\boldsymbol{v}_b = (0,0,0,-4,-3,0,-6,-4,-4,-3)$. The visual illustrations to those events can be seen on the upcoming page. The first plot shows the incoming ask order inside the spread. Afterwards the execution of the market order is shown in the figure. And the last plot shows the cancellation of limit order on the bid side.

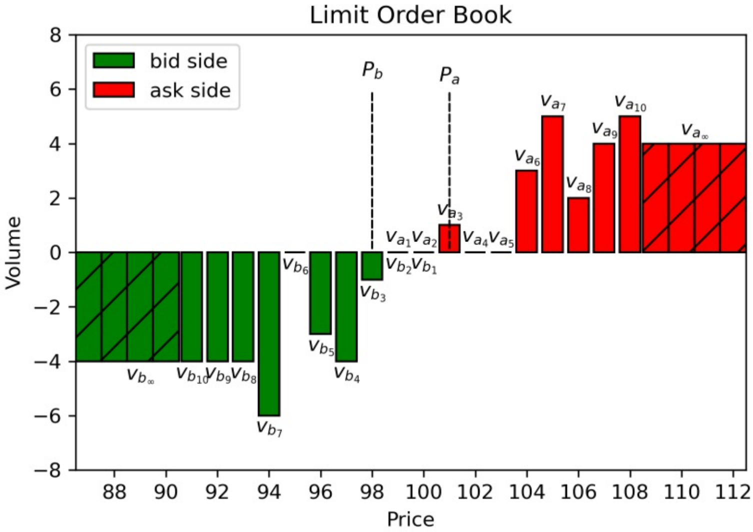

Figure 3.2: Incoming ask limit order

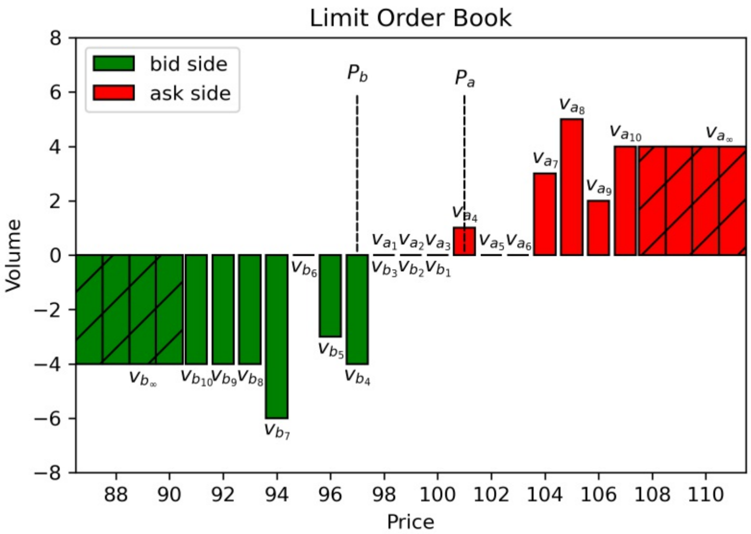

Figure 3.3: Execution of market order

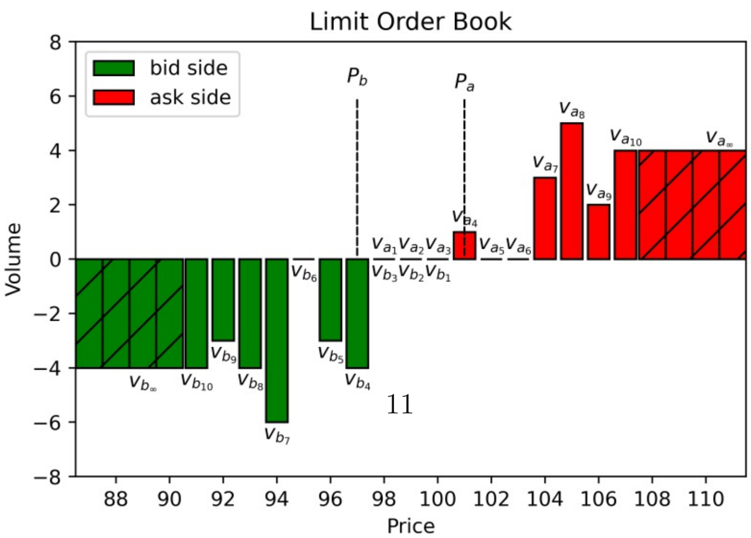

Figure 3.4: Cancellation on the bid side

3.1.5 Infinitesimal Generator and Price dynamics

The infinitesimal operator is derived using the theory of Markovian processes and Poisson processes. For the dynamics of the order book in this model the operator $\mathscr{L}$ denotes the infinitesimal generator. The infinitesimal operator is then given by:

$$\begin{aligned}
\mathscr{L}f(\boldsymbol{v}_a,\boldsymbol{v}_b) = {} & \lambda^{M_a}\big(f([v_{a_i}-(q-V_a(t,i-1))_+]_+;J^{M_a}(\boldsymbol{v}_a,\boldsymbol{v}_b))-f\big)\\
&+\sum_{i=1}^{N}\lambda_i^{L_a}\big(f(v_{a_i}+q;J^{L_a^i}(\boldsymbol{v}_b))-f\big)\\
&+\sum_{i=1}^{N}\lambda_i^{C_a}v_{a_i}\big(f(v_{a_i}-q;J^{C_a^i}(\boldsymbol{v}_a,\boldsymbol{v}_b))-f\big)\\
&+\lambda^{M_b}\big(f(J^{M_b}(\boldsymbol{v}_a,\boldsymbol{v}_b);[v_{b_i}+(q-V_b(t,i-q))_+]_-)-f\big)\\
&+\sum_{i=1}^{N}\lambda_i^{L_b}\big(f(J^{L_b^i}(\boldsymbol{v}_a);v_{b_i}-q)-f\big)\\
&+\sum_{i=1}^{N}\lambda_i^{C_b}|v_{b_i}|\big(f(J^{C_b^i}(\boldsymbol{v}_a,\boldsymbol{v}_b);v_{b_i}+q)-f\big).
\end{aligned}$$

The notation for the subscript + or - at the brackets denote in this formula the max and minimum function defined by $(x)_+ := \max(x,0)$ and $(x)_- := -\min(x,0)$. To describe the price dynamics stochastic differential equations are used. The best ask prices is denoted by $P_a(t)$ and the best bid price by $P_b(t)$, meaning the lowest ask respectively highest bid price and the dynamic of them is described by:

$$\begin{aligned}
dP_a(t) &= \Delta P\big[(V_a^{-1}(t,q)-V_a^{-1}(t,0))dM_a(t)\\
&\qquad -\sum_{i=1}^{N}(V_a^{-1}(t,0)-i)_+dL_a^i(t)+(V_a^{-1}(t,q)-V_a^{-1}(t,0))dC_a^{V_a^{-1}(t,0)}\big]\\
dP_b(t) &= -\Delta P\big[(V_b^{-1}(t,q)-V_b^{-1}(t,0))dM_b(t)\\
&\qquad -\sum_{i=1}^{N}(V_b^{-1}(t,0)-i)_+dL_b^i(t)+(V_b^{-1}(t,q)-V_b^{-1}(t,0))dC_b^{V_b^{-1}(t,0)}\big].
\end{aligned}$$

Here it becomes apparent that the price dynamics depend on three different events. The first thing to mention is that the price is influenced by market orders which consume liquidity at the top of the order book. By consuming liquidity the reduction of available orders at the top of the order book is meant. Another factor influencing the price is if a limit order is placed inside the spread. Through such orders the price moves the distance of the previous best quote to the new order which arrive within the spread. The last event which can influence the price is a cancellation of a limit order at the best price which has the same effect as a market order. Knowing these dynamics, it is now possible to derive the midprice and the spread. First let us look at the midprice:

$$
\begin{aligned}
dP(t) = \frac{\Delta P}{2}\Big[&(V_a^{-1}(t,q) - V_a^{-1}(t,0))dM_a(t) - (V_b^{-1}(t,q) - V_a^{-1}(t,0))dM_b(t) \\
&- \sum_{i=1}^{N}(V_b^{-1}(t,0) - i)_+ dL_a^i(t) + \sum_{i=1}^{N}(V_a^{-1}(t,0) - i)_+ dL_b^i(t) \\
&+ (V_a^{-1}(t,q) - V_a^{-1}(t,0))dC_a^{V_a^{-1}(t,0)}(t) - (V_b^{-1}(t,q) - V_b^{-1}(t,0))dC_b^{V_b^{-1}(t,0)}(t)\Big].
\end{aligned}
$$

The spread indicates the difference between best ask price and best bid price. The importance of the spread lies in the fact that it describes the cost of each transaction for the agents. In this model the spread has the following dynamic:

$$
\begin{aligned}
dS(t) = \Delta P\Big[&(V_a^{-1}(t,q) - V_a^{-1}(t,0))dM_a(t) + (V_b^{-1}(t,q) - V_a^{-1}(t,0))dM_b(t) \\
&- \sum_{i=1}^{N}(V_b^{-1}(t,0) - i)_+ dL_a^i(t) - \sum_{i=1}^{N}(V_a^{-1}(t,0) - i)_+ dL_b^i(t) \\
&+ (V_a^{-1}(t,q) - V_a^{-1}(t,0))dC_a^{V_a^{-1}(t,0)}(t) + (V_b^{-1}(t,q) - V_b^{-1}(t,0))dC_b^{V_b^{-1}(t,0)}(t)\Big].
\end{aligned}
$$

Those equations are interesting in the sense that they link the price dynamics to the order flow. The conditional infinitesimal drifts of the mid-price and the spread can explain what the expected shape of the limit order book at time t might be. Mathematically the expectations are denoted by:

$$
\begin{aligned}
\mathbb{E}[dP(t)|(\boldsymbol{v}_a, \boldsymbol{v}_b)] = \frac{\Delta P}{2}[&(V_a^{-1}(t,q) - V_a^{-1}(t,0))\lambda^{M_a} - (V_b^{-1}(t,q) - V_b^{-1}(t,0))\lambda^{M_b} \\
&- \sum_{i=1}^{N}(V_a^{-1}(t,0) - i)_+ \lambda^{L_a^i} + \sum_{i=1}^{N}(V_b^{-1}(t,0) - i)_+ \lambda^{L_b^i} \\
&+ (V_a^{-1}(t,q) - V_a^{-1}(t,0))\lambda^{C_a}_{V_a^{-1}(t,0)} v_{V_a^{-1}(t,0)} \\
&- (V_b^{-1}(t,q) - V_b^{-1}(t,0))\lambda^{C_b}_{V_b^{-1}(t,0)} |v_{V_b^{-1}(t,0)}|]dt,
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[dS(t)|(\boldsymbol{v}_a, \boldsymbol{v}_b)] = \Delta P[&(V_a^{-1}(t,q) - V_a^{-1}(t,0))\lambda^{M_a} + (V_b^{-1}(t,q) - V_b^{-1}(t,0))\lambda^{M_b} \\
&- \sum_{i=1}^{N}(V_b^{-1}(t,0) - i)_+ \lambda^{L_i^+} - \sum_{i=1}^{N}(V_a^{-1}(t,0) - i)_+ \lambda^{L_b^i} \\
&+ (V_a^{-1}(t,q) - V_a^{-1}(t,0))\lambda^{C_a}_{V_a^{-1}(t,0)} v_{V_a^{-1}(t,0)} \\
&+ (V_b^{-1}(t,q) - V_b^{-1}(t,0))\lambda^{C_b}_{V_b^{-1}(t,0)} |v_{V_b^{-1}(t,0)}|]dt.
\end{aligned}
$$

3.1.6 Ergodicity And Diffusive Limit

In this subsection two important questions will be answered. The first question is if the order book is stable and the second one is what the stochastic process limit of the price at large time scales is. To answer this questions the transition probability function at time t of the Markov process $\mathbf{S}_t$ is given by $\mathrm{Q}_t(\mathbf{S},\cdot)$ is used. Furthermore for the total variation norm of a probability measure μ denoted by the symbol $\|\mu\|$ is needed. Through theses assumptions and definitions the following statement holds.

Theorem 3.1.1. *If* $\underline{\lambda^C} := \min_{1\leq i\leq K}\{\lambda_i^{C_a/b}\} > 0$, *then* $(\boldsymbol{S}(t))_{t\geq 0} = (\boldsymbol{v}_a(t);\boldsymbol{v}_b(t))_{t\geq 0}$ *is an ergodic Markov process and has a unique stationary distribution* Π. *The rate of convergence to this stationary distribution of the order book is exponential. In case there exists* $r, 0 < r < 1$ *and* $R < \infty$ *such that*

$$\|\mathrm{Q}^t(\boldsymbol{S},\cdot) - \Pi(\cdot)\| \leq Rr^t V((\boldsymbol{S})$$

with $t \in R^+$, $(\boldsymbol{v}_a(t);\boldsymbol{v}_b(t)) \in \mathcal{S}$ *and* $V((\boldsymbol{v}_a;\boldsymbol{v}_b) := \sum_{i=1}^N a_i + \sum_{i=1}^N |b_i| + q$ *which are the total number of shares in the book plus* q *shares.*

Proof. See [1, Section 6]. □

In the assumptions for the Theorem it first can be observed how important a positive cancellation intensity is for the convergence. Additionally for the spread it can be stated that it has a well-defined stationary distribution.
To answer the second question about the asymptotic of the re-scaled centered price process some theory needs to be recalled. Mainly theory about ergodicity of Markovian processes and the martingale convergence theorems are needed where the books of Øksendal [14] and Bass [3] can be used to refresh ones memory. The arrival times are Poisson arrival times and this should be kept in mind. In general terms the prices process has the form

$$P_t = \int_0^t \sum_i F_i(\mathbf{S}(u))dN^i(u).$$

In this formula N^i's are the point processes driving the events affecting the limit order book as market orders, limit orders submitted inside the spread and cancellations at the top of the book. The $y^i \equiv y^i(\boldsymbol{v}_a,\boldsymbol{v}_b)$ is the intensity of N^i which can by definition depend on the state of the order book. The different F_i are the jumps in the price of interest always when also the process driving the order book jumps. Using this notation and the ergodic theorem in combination with the martingale theorem the following proposition holds.

Proposition 3.1.1. *The price process given by* $P(t) = \int_0^t \sum_i F_i(\boldsymbol{S}(u))dN^i(u)$ *and introducing the sequence of martingales* $\hat{P}^n$ *resulting from the centered, re-scaled price*

$$\hat{P}^n(t) \equiv \frac{P(nt) - \mathrm{Q}(nt)}{\sqrt{n}}$$

where Q *is the predictable compensator of* P

$$\mathrm{Q}_t = \sum_i \int_0^t y^i(\boldsymbol{S}(s))F_i(\boldsymbol{S}(s))ds$$

If those are fulfilled $\hat{P}^n$ *converges in distribution to a Brownian motion* $\hat{\sigma}B$ *where the volatility of this process is given by*

$$\hat{\sigma}^2 = \lim_{t\to+\infty} \frac{1}{t} \sum_i \int_0^t y^i(\boldsymbol{S}(s))(F_i(\boldsymbol{S}(s)))^2 ds$$
$$= \sum_i \int y^i(\boldsymbol{S})(F_i(\boldsymbol{S}))^2 \Pi(d(\boldsymbol{S}))$$

It remains still difficult to make precise forecasts about the re-scaled price dynamics. To make such it is important to understand the behavior of its compensator Q_{nt}. And although Q_{nt} satisfies itself an ergodic theorem and assuming that the asymptotic variance is not insignificant with respect to nt, it is not possible to state from the previous Proposition that the re-scaled price $P(nt)/\sqrt{n}$ behaves like a Brownian motion with deterministic drift. The following Theorem however provides a more detailed answer under more general assumptions regarding ergodicity.

Theorem 3.1.2. *Let the price process be defined by*

$$P(t) = \sum_i \int_0^t F_i(\boldsymbol{S}(s)) dN^i(s)$$

and the corresponding predictable compensator of the price process is given by

$$\mathrm{Q}(t) = \sum_i \int_0^t y^i(\boldsymbol{S}(s)) F_i(\boldsymbol{S}(s)) ds.$$

Furthermore let h *be the sum over all jumps of the process times the respective intensity*

$$h = \sum_i y^i(\boldsymbol{S}) F_i(\boldsymbol{S})$$

and let α *be defined as*

$$\alpha \overset{a.s.}{=} \lim_{t\to+\infty} \frac{1}{t} \sum_i \int_0^t y^i(\boldsymbol{S}(s)) F_i(\boldsymbol{S}(s)) ds$$
$$= \int h(\boldsymbol{S}) \Pi(d(\boldsymbol{S})).$$

For the solution to the Poisson equation g *is introduced and solves*

$$\mathscr{L}g = h - \alpha.$$

The associated martingale for the Poisson equation is given by

$$Z_t = g(\boldsymbol{S}(t)) - g(\boldsymbol{S}(0)) - \int_0^t \mathscr{L}g(\boldsymbol{S}(s)) ds$$
$$\equiv g(\boldsymbol{S}(t)) - g(\boldsymbol{S}(0)) - \mathrm{Q}(t) - \alpha t.$$

Thus the re-scaled centered price

$$\hat{P}^n(t) \equiv \frac{P(nt) - \alpha nt}{\sqrt{n}}$$

converges in distribution to a Brownian motion $\bar{\sigma}B$. *The asymptotic volatility* $\bar{\sigma}$ *satisfies the identity*

$$\begin{aligned} \bar{\sigma}^2 &= \lim_{t \to +\infty} \frac{1}{t} \sum_i \int_0^t y^i(\boldsymbol{S}(s))((F_i - \Delta^i(g))(\boldsymbol{S}(s)))^2 ds \\ &= \sum_i \int y^i(\boldsymbol{S})((F_i - \Delta^i(g))(\boldsymbol{S}))^2 \Pi(d(\boldsymbol{S})) \end{aligned}$$

where $\Delta^i(g)(\boldsymbol{v}_a; \boldsymbol{v}_b)$ *denotes the jump of the process* $g(\boldsymbol{v}_a; \boldsymbol{v}_b)$ *when the process* N^i *jumps and the limit order book is in the state* $(\boldsymbol{v}_a; \boldsymbol{v}_b)$

Proof. See [1, Section 6]. □

One difficulty in using the formula is to find the stationary distribution of the order book. However all the derived results hold also under less strong restrictions. In order to still arise the same outcome Lyapunov functions are used to model cancellations.

3.2 Weak Law Of Large Numbers For A Limit Order Book

3.2.1 Introduction

This model builds on the Markovian theory. The order flow dynamics are Markovian and this brings the advantage that the type of order, the order size and the price at which the order is placed can all depend on the current state of the order book. The state of the limit order book is the combination of the price and the standing volumes. The main result derived in this model is a limit for a fully state dependent Markovian order book dynamics. The obtained theorem shows that if the amount of orders submitted in a fixed time scale goes to infinity, while the number of active orders, the tick size as well as the number of active orders tend to zero the dynamics of the price dynamics and the volume density functions converge to the unique solution of a coupled ODE/PDE system which is non linear. The model described in this chapter is following the paper 'A Weak Law of Large Numbers for a Limit Order Book Model with Fully State Dependent Order Dynamics' [10].

3.2.2 Setup

In this model only the buy side of the limit order book will be analyzed. This is just to ease notation to obtain both sides of the limit order book one has just to define the sell side analogous. Another restriction which is made but can be relaxed due to the time changing theorem is that the order arrival times are deterministic. All random variables which will occur in this section are defined on a common complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The cádlág process $\mathbf{S}^{(n)} = (S^{(n)}(t))_{0\leq t\leq T}$ describes the dynamics of the limit order book. The values that process can take are in the Hilbert space

$$E := \mathbb{R} \times L^2(\mathbb{R}), \qquad \|\alpha\|_E := |\alpha_1| + \|\alpha_2\|_{L^2}.$$

Changes of the state of the order book can occur due to market or limit orders. The amount of such events in the n-th model are given by $\lfloor T/\Delta t^{(n)} \rfloor$. For the times those events take place

$$t_k^{(n)} := k\Delta t^{(n)}, \qquad k = 1, \dots, \left\lfloor \frac{T}{\Delta t^{(n)}} \right\rfloor$$

is defined where $\Delta t^{(n)}$ is a scaling parameter converging to zero as $n \to \infty$ and $t_0^{(n)} = 0$. For the state of the order book $S_k^{(n)}$ is defined. The k stands for the amount of events that happened in the n-th model. In detail the state of the book is given by

$$S^{(n)}(t) := S_k^{(n)} := (P_k^{(n)}, v_k^{(n)}) \qquad \text{for } t \in [t_k^{(n)}, t_{k+1}^{(n)}].$$

The best bid price after k events in the n-th model is denoted with $P_k^{(n)}$. The buy side volume density function relative to the best quote is denoted by $v_k^{(n)}$ where again k stands for the number of events that took place in the n-th model. The number of available stocks which are $j \in \mathbb{N}$ ticks away from the best price at time $t_k^{(n)}$ is described by

$$\int_{x_{-j-1}^{(n)}}^{x_{-j}^{(n)}} v_k^{(n)}(x)dx,$$

where $x_j^{(n)} := j\Delta x^{(n)}$ for $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. To be able to model also limit orders placed inside the spread the function $v_k^{(n)}$, $k \in \mathbb{N}$, is defined over the whole real line. The orders standing in the order book at a positive distance of the best price are referred as shadow book. This helps to model the distribution of sizes from orders that were placed inside the spread such that those orders enhance the current volume density function of the order book which is visible in a 'smooth' way to the right. The dynamics of the shadow book and the visible one are the same. Moreover the shadow book becomes a part of the visible book trough prices changes. At time $t = 0$ the state of the order book is known for all models with $n \in \mathbb{N}$ and is defined as

$$S_0^{(n)} = (P_0^{(n)}, v_0^{(n)}) \in \mathbb{R} \times L^2(\mathbb{R}).$$

The translation operators $T_+^{(n)}$ and $T_-^{(n)}$ work for every $n \in \mathbb{N}$ and are needed to claim the convergence condition on the sequence of initial states. Those operators act on functions $f : \mathbb{R} \to \mathbb{R}$ the following way

$$T_+^{(n)}(f)(\cdots) := f(\cdot + \Delta x^{(n)}), \quad T_-^{(n)}(f)(\cdot) := f(\cdot - \Delta x^{(n)}).$$

Furthermore it is to consider that the translation operator is isometric for all $f \in L^2$

$$\left\| T_+^{(n)}(f) \right\|_{L^2} = \|f\|_{L^2}.$$

Moreover, in this model $M > 0$ is a fixed constant.

Assumption 3.2.1. *The initial volume function $v_0^{(n)}$ is a non-negative step-function taking values on the grid $\{x_j^{(n)},\ j \in \mathbb{Z}\}$. It is uniformly bounded by the constant M and has compact support for all $n \in \mathbb{N}$ in the interval $[-M, M]$. Furthermore, let $v_0 \in L^2$ be a non-negative continuously differentiable function in such a way that*

$$\left\| v_0^{(n)} - v_0 \right\|_{L^2} = \mathcal{O}(\Delta x^{(n)}).$$

Besides that let the initial price $P_0 \in \mathbb{R}_+$ such that $P_0^{(n)} \to P_0$. The initial state is then denoted by $S_0 := (P_0, v_0) \in E$.

There are three events that affect the state of the limit order book. Those three events change the limit order book if one of the upcoming points hold.

- An arriving market sell order has the same size as the volume standing at the best price in the limit order book. In this situation the best price would decrease by one tick. Events of this type are referred as A type events. Such an event shifts the relative volume density function by one tick to the right.
- An arriving buy limit order which is submitted one tick above the best current price. Therefore the price would increase by one tick. This type of events belong to the B type events. And for the relative density function this means that it will be shifted one tick to the left.

- An arriving buy limit order of size $(\Delta v^{(n)}/\Delta x^{(n)})w_k^{(n)}$ at price $p_k^{(n)}$. In the case that $w_k^{(n)} < 0$ it corresponds to a cancellation of an existing order in the book. Those events are assigned in group C.

The scaling parameter $\Delta v^{(n)}$ specify the size of a particular submission or cancellation. Events of type A and B lead to price changes whereas events of type C do not change the price. Assuming that for category A submitted market orders equal the size of the standing volume at the best price is made for convenience and can be relaxed because the framework is flexible enough. In the case there stands more volume than a market order at the top of the book an incoming market order would be treated as a cancellation.
The different types which can occur are described by a field of random variables $(\phi_k^{(n)})_{k,n\in\mathbb{N}}$. These random variables are taking values in $\{A, B, C\}$. For the size and price at which an order submission or cancellation is made is modeled by the different field of random variables $(w_k^{(n)}, p_k^{(n)})_{k,n\in\mathbb{N}}$. However for the size and price the following assumption is needed.

Assumption 3.2.2. *There exist a field of random variables $(\pi_k^{(n)})_{k,n\in\mathbb{N}}$ which are living in the compact interval $[-M, M]$ with probability one and*

$$p_k^{(n)} := P_k^{(n)} + j\Delta x^{(n)} \quad \textit{for } \phi_k^{(n)} \in [x_{j-1}^{(n)}, x_j^{(n)}).$$

In addition there also exist a field of random variables $(w_k^{(n)})_{k,n\in\mathbb{N}}$ with the property $w_k^{(n)} \in [-M, M]$ for all $k, n \in \mathbb{N}$.

The random variables $\pi_k^{(n)}$ with $k,\ n \in \mathbb{N}_0$ describe the price levels relative to the best quote of order submissions respectively cancellations. In the case where $p_k^{(n)} = P_k^{(n)}$ the submission or cancellation is happening at the best price of the buy side. For the case $p_k^{(n)} < P_k^{(n)}$ the events appear deeper in the order book. The events which occur at a price $p_k^{(n)} > P_k^{(n)}$ are taking place in the shadow book. Due to price changes and the resulting shift of the relative volume density function $v^{(n)}$ the shadow book connects with the visible part of the book.
To avoid negative volumes because of cancellations one can assume that $w_k^{(n)}$ is dependent on $\pi_k^{(n)}$. The advantage of this model is that the conditional distribution of the random variables $(\phi_k^{(n)}), w_k^{(n)}, \pi_k^{(n)}$ can be dependent on the present price and volumes. In order to continue, further notations need to be presented. For instance the σ-field $\mathcal{F}_k^{(n)} := \sigma(S_j^{(n)},\ j \leq k)$ which is defined for every $n \in \mathbb{N}$ and $k = 0, 1, \ldots, \lfloor T/\Delta t^{(n)} \rfloor$. Furthermore, for each $n \in \mathbb{N}$ it is assumed that $S^{(n)}$ is a Markovian process with its own filtration. The space $E' := \{s = (P, v) \in E : v \in C^1\}$ is necessary for the upcoming assumption.

Assumption 3.2.3.

- *There exists Lipschitz continuous functions $p^A, p^B : E \to [0, 1]$ with Lipschitz constant L. In addition there exists a scaling parameter $\Delta p^{(n)}$ such that for all $n \in \mathbb{N}$ and $k \leq \lfloor T/\Delta t^{(n)} \rfloor$*

$$\mathbb{P}[\phi_k^{(n)} = I|\mathcal{F}_k^{(n)}] = \Delta p^{(n)} p^I[S_k^{(n)}] \quad \textit{a.s.} \quad \textit{for } I = A, B.$$

- *There exists a Lipschitz continuous function* $f^{(n)}: E \to L^2$ *for* $n \in \mathbb{N}$ *with Lipschitz constant* $L > 0$ *fro all* $n \in \mathbb{N}$ *such that for all* $k \leq \lfloor T/\Delta t^{(n)} \rfloor$

$$f^{(n)}[S_k^{(n)}](\cdot) = \frac{1}{\Delta x^{(n)}} \mathbb{E}\Big[w_k^{(n)} \sum_{j \in \mathbb{Z}} \mathbb{1}_{\{\pi_k^{(n)} \in [x_j^{(n)}, x_{j+1}^{(n)})\}}(\cdot) \mathbb{1}_C(\phi_k^{(n)}) | \mathcal{F}_k^{(n)} \Big] \quad a.s.$$

and

$$\sup_{s \in E} \left\| f^{(n)}[s](\cdot) \right\|_\infty \leq M.$$

Additionally a function $f: E \to L^2$ *exists with*

$$\sup_{s \in E} \left\| f^{(n)}[s] - f[s] \right\|_{L^2} = \mathcal{S}(\Delta x^{(n)})$$

where $f[s](\cdots): \mathbb{R} \to [-M, M]$ *is continuously differentiable in* x *for all* $s \in E'$*. The derivation of this function is uniformly bounded in absolute value by the constant* M*.*

To simplify notation some definitions will be introduced in order to then define the full dynamics of the limit order book for this model. For $I \in \{A, B, C\}$ and $k, n \in \mathbb{N}$ the event indicator variable is given by

$$\mathbb{1}_k^{(n),I} := \mathbb{1}_I(\phi_k^{(n)}).$$

Moreover, the shorthand notation $I = A, B$ is used

$$p^{(n),I}[\cdot] := \Delta p^{(n)} p^I[\cdot], \qquad p^{(n),B-A} := p^{(n),B} - p^{(n)A},$$
$$p^{B-A} := p^B - p^A, \qquad \mathbb{1}_k^{(n),B-A} := \mathbb{1}_k^{(n)B} - \mathbb{1}_k^{(n),A}.$$

Definition 3.2.1. *The dynamics of the state process of the limit order book* $S^{(n)} = (P^{(n)}, v^{(n)})$ *for each* $n \in \mathbb{N}$ *is denoted by* $S_0^{(n)} := s_0^{(n)}$ *and for* $k = 1, \dots, \lfloor T/\Delta t^{(n)} \rfloor$ *it is defined by*

$$\begin{aligned} P_k^{(n)} &= P_{k-1}^{(n)} + \Delta x^{(n)} \mathbb{1}_{k-1}^{(n),B-A} \\ v_k^{(n)} &= v_{k-1}^{(n)} + (T_-^{(n)} - I)(v_{k-1}^{(n)}) \mathbb{1}_{k-1}^{(n),A} + (T_+^{(n)} - I)(v_{k-1}^{(n)}) \mathbb{1}_{k-1}^{(n),B} + \Delta v^{(n)} M_{k-1}^{(n)}, \end{aligned}$$

where

$$M_k^{(n)}(\cdot) := \mathbb{1}_k^{(n),C} \frac{w_k^{(n)}}{\Delta x^{(n)}} \sum_{j \in \mathbb{Z}} \mathbb{1}_{\{\pi_k^{(n)} \in [x_j^{(n)}, x_{j+1}^{(n)})\}}(\cdot).$$

In order to be able to use the law of large numbers some assumptions regarding the scaling parameters needs to be chosen. For limit order arrivals and cancellations are fast time scales selected. In contrast a comparatively slow one is picked for arriving market orders and limit orders which are placed inside the spread.

Assumption 3.2.4. *There exist constants* $c_0, c_1, c_2 > 0$ *and* $\beta \in (0, 1)$ *such that*

$$\lim_{n \to \infty} \frac{\Delta x^{(n)} \Delta p^{(n)}}{\Delta t^{(n)}} = c_0, \quad \lim_{n \to \infty} \frac{\Delta v^{(n)}}{\Delta t^{(n)}} = c_1, \quad \lim_{n \to \infty} \frac{\Delta x^{(n)}}{(\Delta t^{(n)})^\beta} = c_2.$$

Without loss of generality it can be assumed that the constants c_0, c_1 and c_2 are all equal to one for the upcoming part.
Through the weak law of large numbers it can be stated that the state process of the order book converges in probability to a deterministic limit. This limit is the solution of a system of nonlinear differential equations depending on an initial boundary condition.

Theorem 3.2.1. *Assuming the assumptions 3.2.1, 3.2.2, 3.3.3 and 3.2.4 hold and there exists a deterministic process $S: \ [0,T] \to E$ such that for all $\epsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\left[\sup_{0\le t\le T} \left\| S^{(n)}(t) - S(t) \right\|_E > \epsilon \right] = 0.$$

The function $S = (P, v)$ is the unique solution of the following coupled ODE/PDE initial boundary value problem:

$$\begin{aligned} S(0) &= s_0 \\ dP(t) &= p^{B-A}[S(t)]dt, \qquad t \in [0,T], \\ v_t(t,x) &= p^{B-A}[S(t)]v_x(t,x) + f[S(t)](x), \qquad (t,x) \in [0,T] \times \mathbb{R}. \end{aligned}$$

Proof. See [10]. □

3.3 Second Order Approximations For Limit Order Books

3.3.1 Introduction

This model allows to work with an infinite dimensional limit order book. The dynamics of incoming orders are permitted to depend on the current price along with a volume indicator. Two different cases are studied which differ in the scaling regime. In the first case price changes hardly take place. This leads to a constant first order approximation of the price process and a stochastic differential equation for the volume process in the second order approximation. In the other case a slower rescaling rate is chosen. Which results in a different first as well as second order approximation. The first order approximation then becomes non-degenerate and for the second order approximation this leads to a partial differential equation with random coefficients. The model discussed in this chapter is following the paper of Horst and Kreher [11].

3.3.2 Setup

Because of the nearly symmetric sides of the limit order book and with the aim to keep it clear only the bid side of the book is discussed. Therefore the extension of the model to both sides is not that difficult. The random variables are as always defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The cádlág process $\mathbf{S}^{(n)} = (\mathbf{S}^{(n)}(t))_{0\leq t\leq T}$ describes the state of the bid side of the order book in the nth model. The process takes values in the Hilbert space

$$E := \mathbb{R} \times L^2(\mathbb{R}), \qquad \|\alpha\|_E := |\alpha_1| + \|\alpha_2\|_{L^2}.$$

The change of the limit order book happens due market and limit orders but also due to cancellations. The amount of such events taking place is denoted by $T_n := \lfloor T/\Delta t^{(n)} \rfloor$ where the n stands for the n-th limit order book model. The times in which those events occur is denoted by

$$t_k^{(n)} := k\Delta t^{(n)}, \quad k = 1, \dots, T_n$$

where $\Delta t^{(n)}$ describes a scaling parameter. This parameter converges to 0 as $n \to \infty$ and the starting point $t_0^{(n)} = 0$ is fixed. At time t the state of the order book becomes

$$\mathbf{S}_k^{(n)}(t) := (P_b^{(n)}(t), v_b^{(n)}(t)) \quad \text{for } t \in [t_k^{(n)}, t_{k+1}^{(n)}] \cap [0, T].$$

Whereas $P_b^{(n)}$ describes the best bid price process and is real-valued and $v_b^{(n)}$ is the volume density function on the buy side. It should be noted that the volume process is L^2-valued and that the density is relative to the best price process. While after k events the state of the order book is described by $\mathbf{S}_k^{(n)} := (P_k^{(n)}, v_k^{(n)})$. Due only one side is modeled and for clarity reasons the subscript b is omitted.
The tick size $\delta_x^{(n)}$ leads to the price grid denoted by $x_j^{(n)} := j\delta_x^{(n)}$ for $j \in \mathbb{Z},\ n \in \mathbb{N}$. For all $n \in \mathbb{N}$ and all $x \in \mathbb{R}$ an interval $I^{(n)}(x)$ is defined as

$$I^{(n)}(x) := (x_j^{(n)}, x_{j+1}^{(n)}] \quad \text{for } x_j^{(n)} < x < x_{j+1}^{(n)}.$$

For any $k = 0, \dots, T_n$ the random variable $v_k^{(n)}$ is L^2-valued and it is assumed that the variable is a cádlág step function on the set of prices $\{x_j^{(n)} : j \in \mathbb{Z}\}$. The volume standing in the limit order book at time $t_k^{(n)}$ and at price level $x_j^{(n)}$, $j \in \mathbb{Z}$ can be obtained by

$$\int_{x_{j-1}^{(n)}}^{x_j^{(n)}} v_k^{(n)}(x)dx = \delta_x^{(n)} v^{(n)}(x_j^{(n)}).$$

Starting at $t = 0$ the state of the limit order book is known for all $n \in \mathbb{N}$. The state at the beginning is defined as

$$\mathbf{S}_0^{(n)} = (P_0^{(n)}, v_0^{(n)}) \in \mathbb{R} \times L^2(\mathbb{R}).$$

The events driving the limit order book are described below. Those events can take place at each time $t_k^{(n)}, k = 1, \dots, T_n$ and they change the state of the order book if

- a market sell order arrives which has size of $\delta_x^{(n)} v_{k-1}^{(n)}(0)$. This equals by assumption the volume at the top of the book. Therefore if such an event occurs the best price declines by one tick. Thereby the volume density function is shifted one tick to the right. The set of theses events is defined as A.
- a buy limit order is submitted within the spread. It is assumed that the order is just one tick above the best quote and therefore the price increases just by one tick. For the volume density function this leads to a shift to the left by one tick. Events associated to this type are in the set B.
- a buy limit order is placed with size $\Delta v^{(n)} w_k^{(n)}$ at price level $\eta_k^{(n)}$. Should $w_k^{(n)}$ be smaller than 0 this results in a cancellation of volume. All kind of those events are in the set C.

The assumption that market orders have exactly the size of the volume standing at the best price has primarily mathematical reasons but there is evidence in the literature that supports that [5]. With this restriction the price changes just by one tick after an arriving market order. If a market order is submitted and has a lower size than the volume at the top of the book it is treated as a cancellation. Given that the size of the market order is larger than the liquidity available at the top of the book it is split up into smaller orders where each of this small orders change the price just by one tick. Retrieved from the assumption that market orders can change the price just by one tick the breakdown of a large market order has to be done.
In the definition of the last action driving the state of the order book $\Delta v^{(n)}$ describes a scaling parameter. This parameter specify the amount of a limit order submission or cancellation. A restriction on the tick $\delta_x^{(n)}$ is that if $n \to \infty$ then $\Delta v^{(n)} \to 0$. The relative price level of a limit order submission or placement is described by the random variable $\eta_k^{(n)}$. The random variable $\eta_k^{(n)}$ is defined on the price space $\{x_j^{(n)} : j \in \mathbb{Z}\}$. The exact size of orders is defined by the random variable $w_k^{(n)}$. Furthermore the sign of $w_j^{(n)}$ determines if it is a limit order placement or cancellation. Therefore such an event

affects the volume standing in the order book at price level $\eta_k^{(n)}$ which is the integral of $u_{k-1}^{(n)}$ over $I^{(n)}(\eta_k^{(n)})$ by the size $\Delta v^{(n)} w_k^{(n)}$. The step function $v^{(n)}(x)$ can change by $\Delta v^{(n)} w_k^{(n)}/\delta_x^{(n)}$ for all $x \in I^{(n)}(\eta_k^{(n)})$. For this model $\eta_k^{(n)} := \delta_x^{(n)} \lceil \phi_k^{(n)}/\delta_x^{(n)} \rceil$ with $\pi_k^{(n)}$ is a random variable which takes real values. Through that the interval $I^{(n)}(\eta_k^{(n)})$ can be written as $I^{(n)}(\pi_k^{(n)})$. Additionally the events introduced last can be written as $(w_k^{(n)}, \pi_k^{(n)})$ rather than $(w_k^{(n)}, \eta_k^{(n)})$.
A further assumption is that only one event can take place at a time. The events that can occur are described by the random variables $(\phi_k^{(n)})_{k,n\in\mathbb{N}}$ and take values in one of the three sets described before . The restriction that only one event at a time is taking place is not just for mathematical convenience it is also compatible with the real world. Because it is not possible to execute more than one order at a time. For the events of set C a placement operator is introduced and is defined by

$$M_k^{(n),C}(\cdot) := \mathbb{1}_C(\phi_k^{(n)}) \frac{w_k^{(n)}}{\delta_x^{(n)}} \mathbb{1}_{I^{(n)}(\pi_k^{(n)})}(\cdot).$$

That operator describes the change of the volume density function if a event of type C occurs. Due to the fact that the volume density function is defined on the real line and that it shifts one tick either to the left if a limit order is placed inside the spread or to the right if a market order is submitted, the influence on the limit order book due to a price change can be described by translation operators $T_-^{(n)}$ and $T_+^{(n)}$. Those act on functions $f:\ \mathbb{R} \to \mathbb{R}$ in the following way

$$\begin{aligned} T_-^{(n)}(f)(\cdot) &:= f(\cdot - \delta_x^{(n)}), \\ T_+^{(n)}(f)(\cdot) &:= f(\cdot + \delta_x^{(n)}). \end{aligned}$$

The dynamics of the limit order book can now be described by a stochastic differential equation.

Definition 3.3.1. *The dynamics of the process $\boldsymbol{S}^{(n)} = (P^{(n)}, v^{(n)})$ is given by $\boldsymbol{S}_0^{(n)} := \boldsymbol{s}_0^{(n)}$ for each $n \in \mathbb{N}$ and for every $k = 1, \dots, T_n$,*

$$\begin{aligned} P_k^{(n)} &= P_{k-1}^{(n)} + \delta_x^{(n)}(\mathbb{1}_B(\phi_k^{(n)}) - \mathbb{1}_a(\phi_k^{(n)})) \\ v_k^{(n)} &= v_{k-1}^{(n)} + (T_-^{(n)} - I)(v_{k-1}^{(n)})\mathbb{1}_A(\phi_k^{(n)}) + (T_+^{(n)} - I)(v_{k-1}^{(n)})\mathbb{1}_B(\phi_k^{(n)}) \\ &\quad + \Delta v^{(n)} M_k^{(n),C}. \end{aligned}$$

For the purpose to derive the first order approximation some other assumptions on the initial value of the stochastic differential equation and the random variables driving the limit order book are required.

Assumption 3.3.1. *The volume density function $v^{(n)}$ at time 0 is a step function on the price grid $\{x_j^{(n)}\ :\ j \in \mathbb{Z}\}$. Moreover the function is uniformly bounded through M and* *in $[-M,\ M]$ it has compact support for all $n \in \mathbb{N}$. In addition a differentiable function $v_0^{(n)} \in L^2(\mathbb{R})$ exists and $V_0^{(n)} \in \mathbb{R}$ such that*

$$\left\| v_0^{(n)} - v_0 \right\|_{L^2} = \mathcal{O}(\delta_x^{(n)})$$
$$|P_0^{(n)} - P_0^{(n)}| = o(\Delta t^{(n)})^{1/2}.$$

The function $v_0^{(n)}$ and the corresponding constant $P_0^{(n)}$ are jointly specified as $s_0 = (P_0, v_0) \in E$.

Assumption 3.3.2. *There exists a constant M greater than 0 such that for all $n \in \mathbb{N}$ and $k \leq T_n$*

$$\mathbb{P}[|\pi_k^{(n)}| > M] = \mathbb{P}[|w_k^{(n)}| > 0] = 0.$$

For every $n \in \mathbb{N}$ and for $k = 1, \ldots, T_n$ a σ-field $\mathcal{F}_k^{(n)} := \sigma(S_j^{(n)},\ j \leq k)$ is defined. It is assumed that for every $n \in \mathbb{N}$ the state process $\mathbf{S}^{(n)}$ first conditional moment relies on a volume indicator and the price at that time. The volume indicator is denoted with $Y^{(n)}$. To define the volume indicator a function $h \in L^2(\mathbb{R})$ is fixed which has support in $\mathbb{R}_-$. And with the help of h and the inner product the volume indicator is given by

$$Y_k^{(n)} := \langle h\ , v_k^{(n)} \rangle, \qquad k = 0, \ldots, T_n,\ n \in \mathbb{N}.$$

The upcoming assumption is the reason why this limit order book model has a Markovian structure. As a result of the conditional first moments of order placements or cancellations as well as price changes the Markovian structure is received.

Assumption 3.3.3.

- *For all $n \in \mathbb{N}$ there exist Lipschitz continuous functions*

  $$p^{(n),A}, p^{(n),B} :\ \mathbb{R} \times \mathbb{R} \to [0,1]$$

  with the Lipschitz constant L which is independent of n and scaling parameter $\Delta p^{(n)}$ such that for all $k = 1, \ldots, T_n$ the following equation holds

  $$\mathbb{P}[\phi_k^{(n)} = I|\mathcal{F}_{k-1}^{(n)}] = \Delta p^{(n)} p^{(n),I}(P_{k-1}^{(n)}, Y_{k-1}^{(n)}) \quad a.s.\ for\ I = A, B.$$

  Besides that there exist functions $p^A, p^B :\ \mathbb{R} \times \mathbb{R} \to [0,1]$ such that

  $$\sup_{(b,y)\in\mathbb{R}^2} |p^{(n),I}(b,y) - p^I(b,y)| = o(\Delta x^{(n)})^{1/2}$$

- *Let $f^{(n)} :\ \mathbb{R} \times \mathbb{R} \to L^2(\mathbb{R}), n \in \mathbb{N}$ be Lipschitz continuous functions with mutual Lipschitz constant $L > 0$ such that for all $k = 1, \ldots, T_n$ the equation*

  $$\delta_x^{(n)} f^{(n)}(P_{k-1}^{(n)}, Y_{k-1}^{(n)}; \cdot) := E[\mathbb{1}_C(\phi_k^{(n)}) w_k^{(n)} \mathbb{1}_{I^{(n)}(\pi^{(n)})}(\cdot)|\mathcal{F}_{k-1}^{(n)}] \quad a.s.$$

holds and in addition

$$\sup_{(b,y)\in\mathbb{R}^2} \left\| f^{(n)}(b,y;\cdot) \right\|_{\infty} \leq M$$

for a constant M. Further a function $f: \mathbb{R}\times\mathbb{R} \to L^2$ exists and it is such that

$$\sup_{(b,y)\in\mathbb{R}^2} \left\| f^{(n)}(b,y) - f(b,y) \right\|_{L^2} = \mathcal{O}(\Delta x^{(n)})$$

whereby the function $f(b,y;\cdot): \mathbb{R} \to [-M,M]$ is continuously differentiable in x for all $(b,y)\in\mathbb{R}\times\mathbb{R}$. The derivative of that is then uniformly bounded by the absolute value of the constant M.

3.3.3 First Order Approximation

Under an appropriate scaling assumption the state of the order book process $S^{(n)}$ can be approached by a deterministic process $S = (P, v)$. This process solves an ODE-PDE system. In order to derive the approximation two different time scales are necessary. One for the changes in volume which requires a fast time scale and a slower one for price changes. The intention behind deriving the first order approximation is that the expected impact on the volume of orders in the book is given by $\Delta t^{(n)}$, since the number of order book events is given by $T_n = \lfloor T/\Delta t^{(n)} \rfloor$. The probability of a price change is $\mathcal{O}(\Delta p^{(n)})$. If a price change occurs this will move the volume density function by $\delta_x^{(n)}$. Therefore the expected impact is of dimension $\Delta p^{(n)}\Delta x^{(n)}$ and it is also required that $\Delta p^{(n)}\delta_x^{(n)} = \mathcal{O}(\Delta t^{(n)})$ holds. An event of category C, which are cancellations and order placements affect the volume density function in size $\mathcal{O}(\Delta v^{(n)}/\delta_x^{(n)})$ on the interval relying on the price level at which the submission appears. The expected impact of such on the volume in the limit order book is $\Delta v^{(n)}$. The following scaling assumption refers to that.

Assumption 3.3.4. *There exist $\alpha \in (0,1)$ and $\beta \geq 1-\alpha$ in such a way that the following holds*

$$\Delta t^{(n)} = \Delta v^{(n)}, \qquad \delta_x^{(n)} = (\Delta t^{(n)})^{\alpha}, \qquad \Delta p^{(n)} = (\Delta t^{(n)})^{\beta}.$$

If $\beta = 1-\alpha$ this would refer to the critical case where a non-trivial prices process occurs for the first order approximation. While $\beta > 1-\alpha$ relates to the case where price movements hardly occur and therefore the price process is constant.

Theorem 3.3.1. *Under the assumptions 3.3.1-3.3.4 there exists a deterministic process $S: [0,T] \to E$ such that $\forall \epsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}[\sup_{0\leq t\leq T} \left\| S^{(n)}(t) - S(t) \right\|_{E} > \epsilon] = 0.$$

Additionally is the process $S = (P, v)$ the unique solution for the ODE-PDE system: $\forall (t, x) \in [0, T] \times \mathbb{R}$,

$$\begin{aligned} P_t &= P_0 + \mathbb{1}_{\{\alpha=1-\beta\}} \int_0^t p^{B-A}(P_s, Y_s) ds, \\ v(t, x) &= v_0(x) + \int_0^t f(P_s, Y_s; x) ds + \mathbb{1}_{\{1=1-\beta\}} \int_0^t p^{B-A}(P_s, Y_s) \delta_x v(s, x) ds, \\ Y_t &= \langle h, v(t; \cdot) \rangle \end{aligned}$$

Proof. See [11]. □

To write $p^A - p^B = p^{A-B}$ the weak law of large numbers is used. It should be noted that the volume function $v^{(n)}$ is not positive for sure in this setting except there are no cancellations. Two different approaches help to cope with this problem. The first is that the volume density function $v^{(n)}$ describes logarithmic volumes rather than actual volume. In this constellation negative entries would be legit. The other approach to guarantee a positive volume function but only for short time horizons is letting v_0 be positive and since the first order approximation is deterministic the volume function will be positive. That is why $v^{(n)}$ will be positive for large n and a short time scale with a high probability.

3.3.4 Second Order Approximation

The intention now is to obtain a second order approximation for $S^{(n)} = (P^{(n)}, v^{(n)})$ which denotes the state process. The complexity in doing so is that the price process and the volume process have different time scales. The amount of order placements and cancellations is expected around $T/\Delta t^{(n)}$, whereas the amount of prices changes is in the range of $T \Delta p^{(n)} / \Delta t^{(n)}$. In consequence of the dependency of the volume density process on the price process it is necessary that the same scaling parameters are used for the re-scaled deviation processes, $P^{(n)} - P$ and $v^{(n)} - v$. Subsequently δX_k describes the increment $X_k - X_{k-1}$ for a stochastic process $(X_k)_{k \in \mathbb{N}}$ in discrete time.
For simplicity another scaling parameter is introduced $\Delta^{(n)}$ which will get linked to the other scaling parameters later. The re-scaled discrete fluctuation best price process and the volume process are defined as

$$\begin{aligned} Z_k^{(n),P} &:= \frac{P_k^{(n)} - P(t_k^{(n)})}{(\Delta^{(n)})^{1/2}}, \\ Z_k^{(n),v} &:= \frac{v_k^{(n)}(\cdot) - v(t_k^{(n)}, \cdot)}{(\Delta^{(n)})^{1/2}} \end{aligned}$$

for $k = 0, 1, \ldots, T_n$. The fluctuations of the volume indicator is given by

$$Z_k^{(n),Y} := \frac{Y_k^{(n)} - Y(t_k^{(n)})}{(\Delta^{(n)})^{1/2}} = \frac{\langle h, v_k^{(n)} - v(t_k^{(n)}) \rangle}{(\Delta^{(n)})^{1/2}}$$

for $k = 0, 1, \dots, T_n$. Modifying the definitions for any time $t \in [0, T]$ they become

$$Z^{(n),P}(t) := \frac{P^{(n)}(t) - P(t_k^{(n)})}{(\Delta^{(n)})^{1/2}},$$
$$Z^{(n),v}(t) := \frac{v^{(n)}(t,\cdot) - v(t,\cdot)}{(\Delta^{(n)})^{1/2}}$$
$$Z^{(n),Y}(t) := \langle h, Z^{(n),v}(t)\rangle.$$

This definitions lead to the following useful lemma.

Lemma 3.3.2. *If $\Delta t^{(n)} = o(\Delta^{(n)})^{1/2}$ then*

$$\sup_{k \le T_n} \sup_{t \in [t_k^{(n)}, t_{k+1}^{(n)}]} |Z^{(n),P}(t) - Z_k^{(n),P}| \to 0$$
$$\sup_{k \le T_n} \sup_{t \in [t_k^{(n)}, t_{k+1}^{(n)}]} \left\| Z^{(n),v}(t) - Z_k^{(n),v} \right\|_{L^2} \to 0.$$

For the rest of this model H^m with $m \in \mathbb{N}$ describes the Sobolev space of order m. The norm for the Sobolev space is the usual Sobolev norm. Furthermore let H^{-m} stand for the dual of H^m and $H^0 := L^2$. Then

$$\mathcal{E}' := \bigcup_m H^{-m} \subseteq \dots \subseteq H^{-1} \subseteq L^2 \subseteq H^1 \subseteq \dots \subseteq \bigcap_m H^m =: \mathcal{E} \subseteq C_\infty(\mathbb{R}).$$

The second order approximation will later on converge weak in the Skorokhod space $D([0,T]; \mathbb{R} \times H^{-3})$ with the normally used Skorokhod metric. But to show the second order approximation of the volume part of the state process the next lemma is fundamental.

Lemma 3.3.3. *If $\phi \in H^3$ then*

$$\sup_{t \le T} \left| \left\langle \frac{1}{\Delta^{(n)}} (T_+^{(n)} - I)(v^{(n)}(t)), \phi \right\rangle - \langle \delta_x v(t), \phi \rangle \right| = o_{\mathbb{P}}(\|\phi\|_{H^2})$$

and

$$\sup_{t \le T} \left| \frac{1}{(\Delta x^{(n)})^2} \langle (T_+^{(n)} - I)(v^{(n)}(t)), (I - T_+^{(n)})(\phi) \rangle + \langle \delta_x v(t), \phi' \rangle \right| = o_{\mathbb{P}}(\|\phi\|_{H^3})$$

Thereby $\phi \in H^3$ has to be required in order to obtain the second order approximation. Furthermore it is then just possible to show convergence in the Skorokhod space $D([0,T]; \mathbb{R} \times H^{-3})$.

Assumption 3.3.5. *There exists $h \in H^3$ that satisfies $Y^{(n)} := \langle v^{(n)}, h \rangle$ for all $n \in \mathbb{N}$.*

The following assumptions are about the differentiability of p^A, p^B and f. Those assumptions are necessary because the fluctuations of the price and volume indicator have an impact on the dynamics of S due to those functions.

Assumption 3.3.6. *The functions p^A and p^B are twice continuously differentiable in both arguments and for $I = A, B$*

$$\sup_{b,y}(|p^I_b(b,y)| + |p^I_y(b,y)| + |p^I_{bb}(b,y)| + |p^I_{by}(b,y)| + |p^I_{yy}(b,y)|) < \infty$$

Assumption 3.3.7. *The function f is twice continuously differentiable in the first two arguments and*

$$\sup_{b,y}(\|f(b,y)\|_{L^2} + \|f_b(b,y)\|_{L^2} + \|f_y(b,y)\|_{L^2} + \|f_{bb}(b,y)\|_{L^2} + \|f_{by}(b,y)\|_{L^2} + \|f_{yy}(b,y)\|_{L^2}) < \infty$$

Another assumption which is required to derive the main result will follow after the introduction of the function

$$p^{(n)}(b,y) := (\Delta t^{(n)})^{1/2-\alpha}(p^{(n),B}(b,y) - p^{(n),A}(b,y))$$

which is defined for all $n \in \mathbb{N}$.

Assumption 3.3.8. *There exists a Lipschitz-continuous function $p: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that*

$$\sup_{(b,y)\in\mathbb{R}^2} |p^{(n)}(b,y) - p(b,y)| \to 0.$$

The next assumption is about second moments of order placements or rather cancellations. Those are required to guarantee convergence of the volume fluctuations.

Assumption 3.3.9. *There exist measurable functions $g^{(n)}: \mathbb{R} \times \mathbb{R} \to L^1(\mathbb{R})$, $n \in \mathbb{N}$, such that for all $k = 1, \dots, T_n$*

$$\Delta x^{(n)} g^{(n)}(V_{k-1}, Y^{(n)}_{k-1}; \cdot) = \mathbb{E}[\mathbb{1}_C(\phi^{(n)}_k)(w^{(n)}_k)^2 \mathbb{1}_{I^{(n)}(\pi^{(n)}_k)}(\cdot)|\mathcal{F}_{k-1}] \quad a.s.$$

Besides that there exists a constant $C > 0$ and a Lipschitz continuous function $g: \mathbb{R}\times\mathbb{R} \to L^1(\mathbb{R})$ with

$$\sup_{b,y} \left\| g^{(n)}(b,y) \right\|_\infty \leq C \quad \forall n \in \mathbb{N}$$

$$\sup_{b,y} \int_{\mathbb{R}} |g^{(n)}(b,y;x) - g(b,y;x)| dx \to 0.$$

With all the assumptions introduced the main result of this model can be presented and it is formulated in the following theorem.

Theorem 3.3.4. *Let the assumptions 3.3.1-3.3.4, 3.3.6 and 3.3.7 be fulfilled and the function σ_B is defined as $\sigma_B := (p^A + p^B)^{1/2}$.*

- *If $\Delta^{(n)} = \Delta t^{(n)}$, $\alpha > 1/2$ and $\beta = 2(1-\alpha)$ and let also the assumptions 3.3.8 and 3.3.9 hold then there exists a function μ such that the process $Z^{(n)} = (Z^{(n),P}, Z^{(n),v})$ converges weakly in $D([0,T];\mathbb{R}\times H^{-3})$ to (Z^P, Z^v). This is the unique solution, which is starting from $Z^P = 0$ and $Z^v(0,\cdot) = 0$ to the infinite-dimensional SDE*

$$\begin{aligned} dZ^B(t) &= \mu(Y_t)dt + \sigma_P(P_0, Y_t)dB_t^P \\ dZ^v(t) &= f_b(P_0, Y_t)Z^P(t)dt + f_y(V_0, Y_t)\langle Z^v(t), h\rangle dt \\ &\quad + \delta_x v(t)dZ^P(t) + dM_t \end{aligned}$$

  where B^B is a standard Brownian motion and M is an L^2 valued martingale which covariance depends on (P, Y).

- *If $\Delta^{(n)} = \delta_x^{(n)}, \alpha < 1/2$ and $\beta = 1-\alpha$ and let assumption 3.3.6 be additionally fulfilled then $Z^{(n)} = (Z^{(n),P}, Z^{(n)}, v)$ converges weakly in $D([0,T];\mathbb{R}\times H^{-3})$ to (Z^P, Z^v). This is the unique weak solution, which is starting from $Z_0^P = 0$ and $Z^v(0,\cdot) = 0$, to the system*

$$\begin{aligned} dZ^P(t) &= p_b^{B-A}(P_t, Y_t)Z^P(t)dt + p_y^{B-A}(P_t, Y_t)\langle Z^v(t), h\rangle dt \\ &\quad + \sigma_P(P_t, Y_t)dB_t \\ dZ^v(t) &= f_b(P_t, Y_t)Z^P(t)dt + p_y(P_t, Y_t)\langle Z^v(t), h\rangle dt \\ &\quad + \delta_x v(t)dZ^P(T) + \delta_x Z^v(t)dB_t \end{aligned}$$

Proof. See [11]. □

4 Queueing System Models

4.1 Introduction

This model develops further the idea of zero-intelligence models with the main difference that random sized orders are allowed. The advantage that this model has compared to the first is that the relationship between incoming order sizes and the shape of the order book can be studied. Where the shape of the order book refers to a function which gives for any price the amount of shares standing at this price in the limit order book. To keep the entry simple one-sided order book models will be introduced. These are then subsequently extended to random order sizes on both sides. This model follows mostly the representation in the book "Limit Order Books"[1, Section 7].

4.2 Link Between the Flows of Orders and the Shape of an Order Book

4.2.1 One-sided Queueing System

This order book model builds on the theory of queueing system. As already mentioned for the beginning just one side of the order book will be discussed. One-sided means for this model that all limit order are ask limit orders, all market orders are buy orders and all cancellation are on the ask side. Therefore the side which is presented is the ask side but all results would also hold for the bid side of the book. As a consequence that only the ask side will be modeled the bid side will be hold fixed at zero. An immediate result of this assumption is that the spread always equals the best ask prize. Throughout the one-sided model the subscript a, b standing for ask or buy side is dropped because all the activity takes place on the ask side. $P_a(t)$ denotes as always the best ask price at time t. The price process $\{P_a(t), t \in [0, \infty)\}$ is a continuous time stochastic process which takes values in the discrete set $\{1, \dots, N\}$. The price could be interpreted as number of ticks, whereby the tick size is described by δ_x. The upper bound is given by the highest price N this can be chosen very large, without affecting the shape of the order book at lower prices. All ask limit orders at price $i \in \{1, \dots, N\}$ submitted follow a Poisson process with parameter λ_i^L. The arriving orders are assumed to be mutually independent. That assumption allows orders submitted at prices $1, \dots, r$ to be summarized to one Poisson process with parameter $\Lambda_r^L := \sum_{i=1}^{r} \lambda_i^L$. The limits standing in the order book can be canceled and the intervals those limits are standing in the order book before being deleted form a set of mutually independent random variables. These are identically exponential distributed with parameter $\lambda^C > 0$. At random times buy market orders can be placed and that arises according to a Poisson process with parameter λ^M. A further limitation

for this model is that the size of orders which is of unit size.
The amount of limit orders standing in the order book at prices $1, \dots, k$ at time t is following a stochastic process which is described by $\{V_a(t,k)(t), t \in [0, \infty)\}$. The cumulative shape of the order book is the amount of available orders between the best limit and price k. The cumulative shape in this model will be denoted by the process $V_a(t,k)$. Under the use of queueing systems theory, $V_a(t,k)$ can be viewed as a $M/M/1+M$ queueing system. The arrival rate for this system is Λ_k^L, the service rate is given by λ^M and the reneging rate is λ^C. The queueing system will be referred as the "$1 \to k$" queueing system to illustrate the ticks. Knowing that $V_a(t,k)$ is ergodic it follows from theory that it has a stationary distribution $\pi_{V_a(t,k)}(\cdot)$ whenever the cancellation rate is strict positive $\lambda^C > 0$ which it is by assumption. In consequence it is possible to come up with the matrix form of the infinitesimal generator

$$\begin{pmatrix} -\Lambda_k^L & \Lambda_k^L & 0 & 0 & 0 & \cdots \\ \lambda^M + \lambda^C & -(\Lambda_k^L + \lambda^M + \lambda^C) & \Lambda_k^L & 0 & 0 & \cdots \\ 0 & \lambda^M + 2\lambda^C & -(\Lambda_k^L + \lambda^M + 2\lambda^C) & \Lambda_k^L & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The equations leading to this matrix are given by

$$\Lambda_k^L \pi_{i-1} - (\Lambda_k^L + \lambda^M - (i-1)\lambda^C)\pi_i + (\lambda^M + i\lambda^C)\pi_{i+1} = 0$$

for every i. The $\pi_{V_a(t,k)}$ represent the stationary probability and can be written for all $n \in \mathbb{N}$ as

$$\pi_{V_a(t,k)}(n) = \pi_{V_a(t,k)}(0) \prod_{i=1}^{n} \frac{\Lambda_k^L}{\lambda^M + i\lambda^C}.$$

Specify the summation over all n of the stationary probabilities to one, $\sum_{n=0}^{\infty} \pi_{V_a(t,k)}(n) = 1$ provides the following for $\pi_{V_a(t,k)}(0)$

$$\pi_{V_a(t,k)}(0) = \left(\sum_{n=0}^{\infty} \prod_{i=1}^{n} \frac{\Lambda_k^L}{\lambda^M + i\lambda^C} \right)^{-1}$$

With the normalization of the rates, which happens by dividing the parameters by the cancellation rate, the parameters become $\bar{\Lambda}_k^L = \Lambda_k^L/\lambda^C$ and $\bar{\lambda}^M = \lambda^M/\lambda^C$. Taking into account the newly introduced definitions and doing some rearrangement as well as simplifications the equation develops into

$$\pi_{V_a(t,k)}(n) = \frac{e^{-\bar{\Lambda}_k^L} (\bar{\Lambda}_k^L)^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_{\bar{\Lambda}_k^L}(\bar{\lambda}^M)} \prod_{i=1}^{n} \frac{\bar{\Lambda}_k^L}{i + \bar{\lambda}^M}$$

where $\Gamma_y(x)$ is the incomplete Euler-gamma function

$$\Gamma_y : \mathbb{R}_+ \to \mathbb{R}, \ x \mapsto \int_0^y t^{x-1} e^{-t} dt.$$

The price in this model equals k exactly then when the "$1 \to k-1$" queueing system is empty but the "$1 \to k$" system is not. Subsequently if the cumulative shape of the limit order book $V_a(t,k)$ is distributed according to the invariant distribution $\pi_{V_a(t,k)}$ the price P_a can be written for the first and last tick as

$$\pi_{P_a}(0) = 1 - \frac{e^{-\bar{\Lambda}_1^L}(\bar{\Lambda}_1^L)^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_{\bar{\Lambda}_1^L}(\bar{\lambda}^M)}$$

$$\pi_{P_a}(N) = \frac{e^{-\bar{\Lambda}_{K-1}^L}(\bar{\Lambda}_{K-1}^L)^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_{\bar{\Lambda}_{K-1}^L}(\bar{\lambda}^M)}$$

for all the k's in between, so $k \in \{2, \ldots, K-1\}$, the distribution of the price has the following form

$$\pi_{P_a}(k) = \frac{e^{-\bar{\Lambda}_{k-1}^L}(\bar{\Lambda}_{k-1}^L)^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_{\bar{\Lambda}_{k-1}^L}(\bar{\lambda}^M)} - \frac{e^{-\bar{\Lambda}_k^L}(\bar{\Lambda}_k^L)^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_{\bar{\Lambda}_k^L}(\bar{\lambda}^M)}.$$

If the expectation is applied to the shape of the order book $V_a(t,k)$ it is possible to obtain the average size $\mathbb{E}[V_a(t,k)]$ of the "$1 \to k$" queueing system. By further transformations the average is given by

$$\mathbb{E}[V_a(t,k)] = \bar{\Lambda}_k^L - \frac{\Gamma_{\bar{\Lambda}_k^L}(1+\bar{\lambda}^M)}{\Gamma_{\bar{\Lambda}_k^L}(\bar{\lambda}^M)}.$$

In order to get the amount of limit orders standing in the book at price $k \in \{1, \ldots, N\}$ one hast to take the difference between $V_a(t,k)$ and $V_a(t,k-1)$. The number of orders available is then defined by $v_{a_k} = V_a(t,k) - V_a(t,k-1)$. If the expected value is applied to v_{a_k} the average shape of the order book at price k becomes

$$\mathbb{E}[a_k] = \bar{\lambda}_k^L - \left(\frac{\Gamma_{\bar{\Lambda}_k^L}(1+\bar{\lambda}^M)}{\Gamma_{\bar{\Lambda}_k^L}(\bar{\lambda}^M)} - \frac{\Gamma_{\bar{\Lambda}_{k-1}^L}(1+\bar{\lambda}^M)}{\Gamma_{\bar{\Lambda}_{k-1}^L}(\bar{\lambda}^M)} \right)$$

where $\bar{\lambda}_k^L = \frac{\lambda_k^L}{\lambda^C}$.

4.2.2 Continuous Extension of The Basic Model

The extension concerns the price which can be any positive real number. The market order process is as before a Poisson process with parameter λ^M. No changes are made regarding the size of market orders, those are still of unit size. The elimination of existing limit orders standing in the order book take place after some time which is exponentially random distributed with parameter λ^C. Due to the continuity the submission of limit order changes. As of now the price follow a spatial Poisson process on the positive quadrant $\mathbb{R}_+^2$. The intensity for this process is given by $\lambda^L(p,t)$ which is a non negative function. With that intensity for the spatial Poisson process the arrival of limit orders will be simulated.

The first entry p refers to the price whereas the second entry t indicates the arrival time of orders. About the process is assumed that it is time homogeneous. Furthermore it is assumed that the process is price and time separable. The spatial intensity function of the limit orders which are random events is defined by $h_{\lambda^L} : \mathbb{R}_+ \to \mathbb{R}_+$. The intensity for arrival of limit orders is expressed by $\lambda^L(p,t) = \alpha h_{\lambda^L}(p)$. For any price $p \in [p_1, p_2]$ with $p_1 < p_2 \in [0, \infty)$ the limit order at that price is a homogeneous Poisson process with intensity $\int_{p_1}^{p_1} \lambda^L(p,t)dp$. The cumulative size of the order book is described by $A([0,p])$. This describes the shares standing in the book from zero up to price $p \in \mathbb{R}_+$. This can again be seen as a $M/M/M{+}1$ queueing system. The arrival rate for this system is $\alpha \int_0^p h_{\lambda^L}(u)du$, the service rate is λ^M and reneging is in this setting denoted by cancellations which have rate λ^C. Using the new introduced definitions and the results from the zero-intelligence model it is possible to get the average cumulative shape of the limit order book

$$\mathbb{E}[V_a([0,p])] = \int_0^p \bar{\lambda}^L(u)du - f\left(\bar{\lambda}^M, \int_o^p \bar{\lambda}^L(u)du\right).$$

Whereby the two following definitions were used to keep it clear

$$\bar{\lambda}^L = \frac{\alpha h_{\lambda^L}(u)}{\lambda^C}$$
$$f(x,y) = \frac{\Gamma_y(1+x)}{\Gamma_y(x)}.$$

In addition let $\bar{\Lambda}^L(p) = \int_0^p \bar{\lambda}^L(u)du$ be the normalized arrival rate of limit orders up to price p. For the cumulative shape of the order book $\mathbb{E}[V_a([o,p])]$ the expression $V(p)$ will be used. The average shape of the limit order book is the deviation of the cumulative shape according to the price, $a(p) = \partial A(p)/\partial p$. So that it is the average shape of the order book per price unit. Differentiating the formula for the average cumulative shape of the order book with respect to the price leads to the upcoming proposition.

Proposition 4.2.1. *In a continuous order book framework with homogeneous Poisson arrival of market orders with intensity λ^M, spatial Poisson arrival of limit orders with intensity $\alpha h_{\lambda^L}(p)$ and exponentially distributed time between submitting limit orders and the cancellation of those with parameter λ^C. The average shape in such an environment for the limit order book model is computed for all $p \in [0, \infty)$ by:*

$$a(p) = \bar{\lambda}^L(p)[1 - \bar{\lambda}^M(g_{\bar{\lambda}^M} \circ \bar{\Lambda}^L)(p)[1 - \bar{\lambda}^M[\bar{\Lambda}^L(p)]^{-1}[1 - (g_{\bar{\lambda}^M} \circ \bar{\Lambda}^L)(p)]]]$$

where

$$g_{\bar{\lambda}^M} = \frac{e^{-y}y^{\bar{\lambda}^M}}{\bar{\lambda}^M \Gamma_y(\bar{\lambda}^M)}.$$

The function $g_{\bar{\lambda}^M}(\bar{\Lambda}^L(p))$ can be seen as the probability that the limit order book is empty up to price p. Letting $\lambda^M \to 0$, which means no more market orders are submitted, leads to the average shape $a(p) \to \bar{\lambda}^L(p)$ in view of the proposition. This result is not really a surprise because without market orders the average shape of the order book would only depend on the normalized arrival rate of limit orders. Letting $p \to \infty$ the average shape has an approximation $a(p) \sim k\bar{\lambda}^L(p)$ with some constant k. This result will be investigated in the following proposition.

Proposition 4.2.2. *The shape of the limit order book can be written as*

$$v_a(p) = \bar{\lambda}^L(p)C(p)$$

where $C(p)$ denotes the probability that a limit order submitted at price p will be canceled before executed.

It can be concluded that the shape of the order book corresponds to the fraction of arriving limit orders that will be canceled. The difference between the flows of limit orders and the order book can be exactly described by the fraction of arriving limit orders that will be traded.
Using the equation for the average shape of the order book, carry out some computations and with some results of the theory for queueing systems, the fraction C_k of limit orders submitted at price k in the discrete framework which are canceled is

$$C_k = 1 - \frac{\bar{\lambda}^M}{\bar{\Lambda}^L_k}(g_{\bar{\lambda}^M}(\bar{\Lambda}^L(p)) - g_{\bar{\lambda}^M}(\bar{\Lambda}^L(p+\epsilon))).$$

The proportion of canceled limit orders at a price in the interval $[p, p+\epsilon]$ in the setting of the continuous model and with the average cumulative shape given by $A(p)$ is denoted by

$$1 - \frac{\bar{\lambda}^M}{\bar{\Lambda}^L(p+\epsilon) - \bar{\Lambda}^L(p)}(g_{\bar{\lambda}^M}(\bar{\Lambda}^L(p)) - g_{\bar{\lambda}^M}(\bar{\Lambda}^L(p+\epsilon))).$$

To gain the fraction $C(p)$ which describes the limit orders submitted at price p which are canceled, one has to let $\epsilon \to 0$. If this is done the following equation for $C(p)$ is obtained

$$\begin{aligned} C(p) &= 1 - \bar{\lambda}^M g'_{\bar{\lambda}^M}(\bar{\Lambda}^L(p)) \\ &= 1 - \bar{\lambda}^M (g_{\bar{\lambda}^M} \circ \bar{\Lambda}^L)(p)\left(1 - \frac{\bar{\lambda}^M}{\bar{\Lambda}^L(p)}(1 - (g_{\bar{\lambda}^M} \circ \bar{\Lambda}^L)(p))\right). \end{aligned}$$

Taking this formula into consideration the relationship between the shape of the order book and the flows of arrival limit orders can be examined. For high prices two main cases exist which can be analyzed. The first case assumes a positive finite constant α for the total arrival rate of limit orders. Then the proportional constant C_∞ for $p \to +\infty$ which describes the relation between the shape of the order book $a(p)$ and the normalized limit order flow $\bar{\lambda}^M(p)$ is defined as

$$C_\infty = \lim_{p\to\infty} C(p) = 1 - \bar{\lambda}^M g_{\bar{\lambda}^M}(\alpha)\left(1 - \frac{\bar{\lambda}^M}{\alpha}(a - g_{\bar{\lambda}^M}(\alpha))\right) < 1.$$

In this case the shape of the order book does not equal the normalized rate of arrival limit orders $\bar{\lambda}^L(p)$ it is just proportional to it. The proportion of canceled limit orders does not go to 1 as the price tends to infinity. One can deduce from that market orders play an important role for the shape of limit order books even at high prices. The other case is where $\lim_{p\to\infty} \bar{\Lambda}^L(p) = \infty$ and in this setting market orders do not reach the highest prices. This affects the tail of the limit order book because the order book behaves like there are no market orders at all, $a(p) \sim \bar{\lambda}^L(p)$ as $p \to +\infty$.

4.3 Varying Size Of Limit Orders

This model allows random sized limit orders in the framework of the one sided model. Just remark that all limit orders are ask orders, market orders are buy orders and cancellations just hit the ask side. Although a significant assumption changes particularly the order size most of the other assumptions remain. For the size of the limit orders it is assumed that these are independent random variables. A further assumption for sizes of limit orders is that they are identically distributed if they are submitted at a given price. To put it in more general terms the distribution depends on the price level and it can also depend on the price itself. To describe the probability that a limit order is submitted at price k is of size n the function g_n^k is introduced. This function is defined for any price $k \in \mathbb{N}$ and any order size $n \in \mathbb{N}$. The average over all submitted orders at price k is defined by $\bar{g}^k$ and is assumed to be finite. The arrival rate of limit orders of size n at price k is denoted by $\lambda_i^L g_n^i$. Using then the summation property of the independent Poisson process it is possible to write $\sum_{i=1}^k \lambda_i^L g_n^i$ for the arrival of limit orders at size n with a price up to k. To describe the probability of a limit order with a price up to k and size n is $G_n^k = \sum_{i=1}^k \frac{\lambda_i^L}{\Lambda_k^L} g_n^i$. Straightforward the mean size of a limit order is then $\bar{G}^K = \sum_{i=1}^k \frac{\lambda_i^L}{\Lambda_k^L} \bar{g}^i$.
Regarding the cancellation the mechanism the main properties remain. Limit orders can be canceled once standing in the limit order book although not all at once but unit by unit. The time intervals between submission of a limit order and cancellation of that order are assumed to be mutually independent random variables identically distributed according to an exponential distribution with parameter $\lambda^C > 0$. For market orders nothing changes in this model, they are still of unit size and are submitted at random times following a Poisson process with parameter λ^M.
The cumulative shape of the limit order book $V_a(t,k)$ is modeled by a stochastic process $\{V_a(t,k), t \in [0,\infty)\}$ describing the number of limit orders in the book at prices $1,\ldots,k$. In the context of a queueing system this would be a $M^x/M/1+M$ queueing system. The bulk arrival rate is the normalized limit order rate Λ_k^L, the bulk volume distribution is $(G_n^k)_{n\in\mathbb{N}}$, the service system for this system is λ^M and the cancellations are the reneging rate λ^C. The infinitesimal generator for the cumulative shape $V_a(t,k)$ is

$$\begin{pmatrix} -\Lambda_k^L & \Lambda_k^L G_1^k & \Lambda_k^L G_2^k & \Lambda_k^L G_3^k & \Lambda_k^L G_4^k & \cdots \\ \lambda^M + \lambda^C & -(\Lambda_k^L + \lambda^M + \lambda^C) & \Lambda_k^L G_1^k & \Lambda_k^L G_2^k & \Lambda_k^L G_3^k & \cdots \\ 0 & \lambda^M + 2\lambda^C & -(\Lambda_k^L + \lambda^M + 2\lambda^C) & \Lambda_k^L G_1^k & \Lambda_k^L G_2^k & \cdots \\ 0 & 0 & \lambda^M + 3\lambda^C & -(\Lambda_k^L + \lambda^M + 3\lambda^C) & \Lambda_k^L G_1^k & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The stationary distribution $\pi_{V_a(k)} = (\pi_{V_a(k)}(n))_{n\in\mathbb{N}}$ of $V_a(k)$ is the solution to the system of equations:

$$\begin{cases} 0 = -\Lambda_k^L \pi_{V_a(k)}(0) + (\lambda^M + \lambda^C)\pi_{V_a(k)}(1) & \text{n=0} \\ 0 = -(\Lambda_k^L + \lambda^M + n\lambda^C)\pi_{V_a(k)}(n) + (\lambda^M + (n+1)\lambda^C)\pi_{V_a(k)}(n+1) & \\ \quad + \sum_{i=1}^n \Lambda_k^L G_i^k \pi_{V_a(k)}(n-i) & \text{n}\geq 1. \end{cases}$$

This system of equations is solved with the help of probability generating functions. Before the theory of moment generating functions can be used those need to be defined. Therefore let $\Phi_{V_a(k)}(z) = \sum_{n\in\mathbb{N}} \pi_{V_a(k)}(nz^n)$ and let $\Phi_{G^k}(z) = \sum_{n\in\mathbb{N}} G_n^k z^n$. The normalized market parameter $\bar{\lambda}^M = \lambda^M/\lambda^C$ and $\bar{\Lambda}^L = \Lambda_k^L/\lambda^C$ which is the normalized parameter for limit orders, where the limit order rate gets divided by the cancellation rate, are going to be used. Multiplying the $n-$th line by z^n of the previous system of equations and summing over all n results in the differential equation

$$\frac{\partial}{\partial z}\Phi_{V_a(k)}(z) + \left(\frac{\bar{\lambda}^M}{z} - \bar{\Lambda}_k^L \phi_{G^k}(z)\right)\Phi_{V_a(k)}(z) = \frac{\bar{\lambda}^M}{z}\pi_{V_a(k)}(0)$$

with $\phi_{G^k} = (1-\Phi_{G^k}(z))/(1-z)$. Solving this equation for $\Phi_{V_a(k)}$ leads to the following

$$\Phi_{V_a(k)}(z) = z^{-\bar{\lambda}^M}\bar{\lambda}^M \pi_{V_a(k)}(0) e^{\bar{\Lambda}_k^L \int_0^z \phi_{G^k}(u)du} \int_0^z v^{\bar{\lambda}^M - 1} e^{-\bar{\Lambda}_k^L \int_0^v \phi_{G^k}(u)du} dv.$$

Using the property of moment generating functions that $\Phi_{V_a(k)}(1) = 1$ results in

$$\pi_{V_a(k)}(0) = \left(\bar{\lambda}^M \int_0^1 v^{\bar{\lambda}^M-1} e^{\bar{\Lambda}_k^L \int_v^1 \phi_{G^k}(u)du} dv\right)^{-1}.$$

Substituting this into the general solution formula yields

$$\Phi_{V_a(k)}(z) = z^{-\bar{\lambda}^M} \frac{\int_0^z v^{\bar{\lambda}^M-1} e^{\bar{\Lambda}_k^L \int_0^z \phi_{G^k}(u)du} dv}{\int_0^1 v^{\bar{\lambda}^M-1} e^{\bar{\Lambda}_k^L \int_v^1 \phi_{G^k}(u)du} dv}.$$

If this result in combination with the differential equation from above is considered and taking the limit when z tends increasingly to 1, the upcoming proposition is obtained. But before that can be stated some basic properties of moment generating functions are additionally required namely:

$$\lim_{\substack{z\to 1\\ t<1}} \Phi_{V_a(k)}(z) = 1$$

$$\lim_{\substack{z\to 1\\ t<1}} \frac{\partial}{\partial z}\Phi_{V_a(k)}(z) = \mathbb{E}[V_a(k)]$$

$$\lim_{\substack{z\to 1\\ t<1}} \phi_{G^k}(z) = \bar{G}^k.$$

Thus the following proposition can be derived.

Proposition 4.3.1. *In the discrete limit order book model were orders are only submitted at one side with Poisson arrival rate λ^M for market orders with unit size, Poisson arrival of limit order with rate λ_k^L and random size which is distributed according to $(g_n^k)_{n\in\mathbb{N}}$, the life span of limit orders which are not executed is exponential distributed with parameter λ^C, the average cumulative shape of the order book up to price k is then given by:*

$$\mathbb{E}[V_a(k)] = \bar{\Lambda}_k^L \bar{G}^k - \bar{\lambda}^M + \left(\int_0^1 v^{\bar{\lambda}^M - 1} e^{\bar{\Lambda}_k^L \int_v^1 \phi_{G^k}(u)du} dv \right)^{-1}.$$

Let now assume that the size of limit orders are geometrically distributed with parameter $q \in (0,1)$ and furthermore independent of the price. The price can be any $k \in \mathbb{N}$ and the distribution function is given by $g_n^k = (1-q)^{n-1}q$. The moment generating function is received using the volume distribution for geometrically distributed variables. In further consequence the moment generating function is given by $\phi_{G^k}(z) = 1/(1-(1-q)z)$. Thus the average cumulative shape of the order book is described by

$$\mathbb{E}[V_a(k)] = \frac{\bar{\Lambda}_k^L}{q} - \bar{\lambda}^M + \frac{\bar{\lambda}^M q^{\frac{\bar{\Lambda}_k^L}{1-q}}}{{}_2F_1(\bar{\lambda}^M, \frac{-\bar{\Lambda}_k^L}{1-q}, 1+\bar{\lambda}^M, 1-q)}$$

where ${}_2F_1$ is the hyper-geometric function

$${}_2F_1(a,b,c,z) = \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b+k)\Gamma(c)z^k}{\Gamma(a)\Gamma(b)\Gamma(c+k)k!}.$$

The price in the limit order book model is still continuous and the limit orders are submitted following a spatial Poisson process with intensity $\lambda^L(p,t) = \alpha h_{\lambda^L}(p)$. For h_{λ^L} it is assumed that it is a real non-negative function with positive support and it describes the spatial intensity of arrival rates. Therefore the amount of limit orders submitted in the limit order book in a price range $[p_1, p_2]$ is a homogeneous Poisson process with rate $\int_{p_1}^{p_"} \alpha h_{\lambda^L}(u)du$. If the definitions introduced in this special case are used the cumulative shape of the continuous order book at a price $p \in [0,\infty)$ is given by

$$V_a(p) = \frac{1}{q}\bar{\Lambda}^L(p) - \bar{\lambda}^M + \frac{\bar{\lambda}^M q^{\frac{\bar{\Lambda}^L(p)}{1-q}}}{{}_2F_1(\bar{\lambda}^M, \frac{-\bar{\Lambda}_k^L}{1-q}, 1+\bar{\lambda}^M, 1-q)}.$$

Taking the derivative with respect to p of the cumulative shape results in the average shape of the order book $v_a(p)$. The result of this derivation is contained in the subsequent proposition.

Proposition 4.3.2. *In a continuous one-sided limit order book model with homogeneous Poisson arrival market orders with intensity λ^M and unit order size, spatial Poisson rate for arriving limit orders with intensity $\alpha h_{\lambda^L}(p)$ at price p, geometric distribution of the limit order sizes with parameter q and life span of limit orders which are not executed is exponential distributed with parameter λ^C, the average shape of the limit order book v is then computed for $p \in [0,\infty)$ by:*

$$v_a(p) = \frac{\bar{\lambda}^L(p)}{q} + \frac{\partial}{\partial p}\left(\frac{\bar{\lambda}^M q^{\frac{\bar{\Lambda}^L(p)}{1-q}}}{{}_2F_1(\bar{\lambda}^M, \frac{-\bar{\Lambda}_k^L}{1-q}, 1+\bar{\lambda}^M, 1-q)} \right)$$

5 Hawkes Processes Models

5.1 Limit Order Book Driven by Hawkes Processes

5.1.1 Introduction

Hawkes processes are widely used nowadays in all different areas of finance [2, 9]. The advantage of Hawkes processes is that they allow for a great flexibility and versatility in models for processes that mutually excite each other. Hawkes processes were first introduced in the pioneering work of Hawkes and belong to the class of point processes [7, 8]. The first model which will be introduced using Hawkes processes is mainly from the book "Limit Order Book" [1, section 8].

5.1.2 Hawkes Processes and Model Set Up

For this limit order book model multivariate Markovian Hawkes processes are required. Therefore some definitions and results will be introduced at the beginning. Thus let $\mathbf{N} = (N^1, \ldots, N^D)$ be a D-dimensional point process with intensity vector $\lambda = (\lambda^1, \ldots, \lambda^D)$.

Definition 5.1.1. *The D-dimensional point process $\boldsymbol{N} = (N^1, \ldots, N^D)$ is called a multivariate Hawkes process with exponential kernel if there exists $(\lambda_0^i)_{1\leq i\leq D} \in (\mathbb{R}_+)^D$, $(\alpha_{ij})_{1,\leq i,j\leq D} \in (\mathbb{R}_+)^{D^2}$ and $(\beta_{ij})_{1\leq i,j\leq D} \in (\mathbb{R}_+)^{D^2}$ such that the intensities of the process satisfy the following set of relations:*

$$\lambda^M(t) = \lambda_0^M + \sum_{j=1}^{D} \alpha_{mj} \int_0^t e^{-\beta_{mj}(t-s)} dN^j(s)$$

for $1 \leq m \leq D$.

This particular kernels are chosen in order to satisfy the requirements for the following proposition.

Proposition 5.1.1. *If the process μ^{ij} is defined as*

$$\mu^{ij}(t) = \alpha_{ij} \int_0^t e^{-\beta_{ij}(t-s)} dN^j(s)$$

for $1 \leq i, j \leq D$ and $\boldsymbol{\mu} = \{\mu^{ij}\}_{1\leq i,j\leq D}$. It follows that the process $(\boldsymbol{N}, \boldsymbol{\mu})$ is Markovian.

Stability for multivariate Hawkes processes in the form they were just introduced is in general given [13]. Furthermore it can be shown that a Lyapunov function exists for such processes. Due to the existence of a Lyapunov function exponential convergence against the stationary distribution can be concluded.

Proposition 5.1.2. *Let the matrix* $\mathbf{A}$ *be defined as*

$$\mathbf{A}_{ij} = \frac{\alpha_{ij}}{\beta_{ij}}$$

for $1 \leq i, j \leq D$*. Further let* $\mathbf{A}$ *be a positive matrix and that is spectral radius* $\rho(\mathbf{A})$ *fulfills the condition*

$$\rho(\mathbf{A}) < 1$$

Then, there exits a (unique) multivariate point process $\mathbf{N} = (N^1, \ldots, N^D)$ *whose intensity is given by*

$$\lambda^M(t) = \lambda_0^M + \sum_{j=1}^{D} \alpha_{mj} \int_0^t e^{-\beta_{mj}(t-s)} dN^j(s)$$

In addition this process is stable and converges exponential in total variation norm against its unique stationary distribution.

Three different type of events can occur and influence the limit order book. First new limit orders can arrive, those are described by mutually exciting Hawkes processes. The second event that can appear are new arrivals of market orders, again described by mutually exciting Hawkes processes. The last event which can influence the limit order book are cancellations of limit orders already standing in the limit order book. The time those limit orders are standing in the order book is modeled by a Poisson process. Summarized and with mathematical notation processes are given by:

- $M_{a/b}(t)$: Hawkes processes of market orders which either can be a buy or sell order with constant intensities λ^{M_a} and λ^{M_b}
- $L^i_{a/b}(t)$: Hawkes processes of limit orders at level i, with constant intensities $\lambda_i^{L_a}$ and $\lambda_i^{L_b}$
- $C^i_{a/b}(t)$: Counting process for cancellations of limit orders at level i, with intensities $\lambda_i^{C_a} a_i$ and $\lambda_i^{C_b} |b_i|$.

Where the "a" stands for events happening on the ask side and "b" for the events affecting the bid side.

5.1.3 Infinitesimal Generator And Stability

The intensities of market and limit orders arrivals are described by a Markovian $(2N+2)$-dimensional Hawkes process. The limit order book as a whole is represented by the D-dimensional process $(\boldsymbol{v}_a;\boldsymbol{v}_b;\boldsymbol{\mu})$. In this process the available quantities standing in the order book and the intensities of the Hawkes process are included. The dimension of the state space is given by D and equals $D=(2N+2)^2+2N$. The evolution of the limit order book is characterized by the infinitesimal generator and has the subsequent form

$$\begin{aligned}
\mathscr{L}f(\boldsymbol{v}_a;\boldsymbol{v}_b;\boldsymbol{\mu}) &= \lambda^{M_a}\big(F([a_i-(q-V_a(i-1))_+]_+;J^{M_a}(\boldsymbol{v}_a,\boldsymbol{v}_b);\boldsymbol{\mu}+\Delta^{M_a}(\boldsymbol{\mu}))-F\big)\\
&+\sum_{i=1}^{N}\lambda_i^{L_a}\big(F(V_{a_i}+q;J^{L_a^i}(\boldsymbol{v}_b);\boldsymbol{\mu}+\Delta^{L_a^i}(\boldsymbol{\mu}))-F\big)\\
&+\sum_{i=1}^{N}\lambda_i^{C_a}v_{a_i}\big(F(v_{a_i}-q;J^{C_a^i}(\boldsymbol{v}_a,\boldsymbol{v}_b);\boldsymbol{\mu})-F\big)\\
&+\lambda^{M_b}\big(F(J^{M_b}(\boldsymbol{v}_a,\boldsymbol{v}_b);[v_{b_i}+(q-V_b(i-1))_+]_-;\boldsymbol{\mu}+\Delta^{M_b}(\boldsymbol{\mu}))-F\big)\\
&+\sum_{i=1}^{N}\lambda_i^{L_b}\big(F(J^{L_i^-}(\boldsymbol{v}_a);v_{b_i}-q;\boldsymbol{\mu}+\Delta^{L_b^i}(\boldsymbol{\mu}))-F\big)\\
&+\sum_{i=1}^{N}\lambda_i^{C_b}|v_{b_i}|\big(F(J^{C_b^i}(\boldsymbol{v}_a,\boldsymbol{v}_b);v_{b_i}+q;\boldsymbol{\mu})-F\big)\\
&=\sum_{i,j=1}^{N}\beta_{ij}\mu^{ij}\frac{\partial F}{\partial\mu^{ij}}.
\end{aligned}$$

The bid side volume is modeled by negative numbers therefore the absolute value is necessary. The J's are the shift operators introduced in the zero-intelligence model. For convenience $F(v_{a_i};\boldsymbol{v}_b;\boldsymbol{\mu})$ is written instead of $F(v_{a_1},\dots,v_{a_i};\dots,v_{a_n};\boldsymbol{v}_b;\boldsymbol{\mu})$. By the same reasoning the process and the corresponding state variable in the state space have the same symbol. The newly introduced $\Delta^{(\cdots)}(\boldsymbol{\mu})$ describes the jump of the intensity vector $\boldsymbol{\mu}$ whenever the process $N^{(\cdots)}$ jumps.

The infinitesimal operator $\mathscr{L}$ is a combination of standard difference operators and drift terms. The standard difference operators come from the arrival or cancellation of orders at every level and the shift operators J which are indicating the movements at the best limits. In contrary the drift term comes from the mean-reverting behavior property of intensities for Hawkes processes between jumps.

In the following part of this chapter the stability and long-term behavior of the order book will be discussed. To investigate the long-term behavior a Lyapunov function is constructed. Under the assumption that the spectral radius of the matrix $\mathbf{A}$ is smaller than one in combination with the existence of the Lyapunov function can be concluded that the limit order book is ergodic. The following proposition will describe this more in detail.

Proposition 5.1.3. *Under the assumptions regarding the Hawkes processes of arriving limit orders, market orders and the process for cancellations of limit orders and the as-*

sumption that the spectral radius for $\boldsymbol{A}$ is smaller than one the limit order book process $\boldsymbol{S}$ is ergodic. Moreover it converges exponential towards its unique stationary distribution Π.

The for this result required Lyapunov function is generated by the following Lemma.

Lemma 5.1.1. *For $\eta > 0$ small enough, the function V defined by*

$$V(\boldsymbol{v}_a; \boldsymbol{v}_b; \boldsymbol{\mu}) = \sum_{i=1}^{N} v_{a_i} + \sum_{i=1}^{N} |v_{b_i}| + \frac{1}{\eta} \sum_{i,j=1}^{(2N+2)} \delta_{ij} \mu^{ij}$$
$$\equiv V_1 + \frac{1}{\eta} V_2$$

where V_1 is a function just depending on $(\boldsymbol{v}_a; \boldsymbol{v}_b)$ and V_2 corresponds to $\boldsymbol{\mu}$, is a Lyapunov function which satisfies the geometric drift condition

$$\mathscr{L} V \leq -\zeta V + C$$

for some $\zeta > 0$ and $C \in \mathbb{R}$. The coefficients δ_{ij} are defined by the Perron-Frobenius theorem and it holds that: $\forall i, \epsilon_i > 0$

$$\delta_{ij} = \epsilon_i \frac{\mu^{ij}}{\beta_{ij}}.$$

5.1.4 Large Scale Limit of the Price Process

The long-term behavior of the price process is for researchers and market participants from great interest. Because based on that, you can infer the volatility in the market. However, to study the long-term behavior of the price process also the stochastic behavior of the intensities of the point processes which set off the order book events is included. The equation for the price dynamics is given by

$$P_t = \int_0^t \sum_i F_i(\mathbf{S}(u)) dN^i(u)$$

where $\mathbf{S} = (\boldsymbol{v}_a, \boldsymbol{v}_b, \boldsymbol{\mu})$ is the Hawkes process which also describes the state of the order book. The N^i, with have state-dependent intensities $y^i(\mathbf{S})$, are the Poisson and Hawkes processes describing the events which in turn determine the evolution of the order book. The F_i are bounded function due to the limitation of price changes. Since prices can only change by the amount of limits in the order book, as the non-zero boundary conditions $v_{a_\infty}, v_{b_\infty}$ holds.
The following theorem summarizes the main findings in this model driven by Hawkes processes about the price dynamics. It describes the convergence speed and the deterministic centered price.

Theorem 5.1.2. *Let the price process be defined as*

$$P(t) = \sum_i \int_0^t F_i(\boldsymbol{S}(s))dN^i(s)$$

and the corresponding predictable compensator of the price process is given by

$$\mathrm{Q}(t) = \int_0^t \sum_i y^i(\boldsymbol{S}(s))F_i(\boldsymbol{S}(s))ds.$$

Furthermore let h be the sum

$$h = \sum_i y^i F_i(\boldsymbol{S})$$

and α is determined by the following equation

$$\begin{aligned}\alpha &= \lim_{t\to+\infty}^{a.s.} \frac{1}{t}\int_0^t \sum_i y^i(\boldsymbol{S}(s))F_i(\boldsymbol{S}(s)ds \\ &= \int h(\boldsymbol{S})\Pi(d(\boldsymbol{S}).\end{aligned}$$

The solution to the Poisson equation is given by g and it solves

$$\mathscr{L}g = h - \alpha.$$

The associate resulting martingale is given by

$$\begin{aligned}Z_t &= g(\boldsymbol{S}(t)) - g(\boldsymbol{S}(0)) - \int_0^t \mathscr{L}g(\boldsymbol{S}(s))ds \\ &\equiv g(\boldsymbol{S}(t)) - g(\boldsymbol{S}(0)) - \mathrm{Q}(t) - \alpha t.\end{aligned}$$

Through that the deterministic centered, re-scaled price

$$\bar{P}^n(t) \equiv \frac{P(nt) - \alpha nt}{\sqrt{n}}$$

converges in distribution to a Brownian motion $\bar{\sigma}B$. The asymptotic volatility $\bar{\sigma}$ fulfills the identity condition

$$\begin{aligned}\bar{\sigma}^2 &= \lim_{t\to+\infty} \frac{1}{t}\int_0^t \sum_i y^i(\boldsymbol{S}(s))(F_i - \Delta^i(g)(\boldsymbol{S}(s))^2 ds \\ &\equiv \int \sum_i y^i(\boldsymbol{S})(F_i - \Delta^i(g)(\boldsymbol{S}))^2 \lambda\Pi(d(\boldsymbol{S})\end{aligned}$$

Proof. See [1, section 8]. □

5.2 Hawkes Random Measures Model

5.2.1 Introduction

This model is based on Hawkes processes and follows a paper published by Horst and Xu [12]. To be more precise the model uses Hawkes random measures to describe the order book. Hawkes random measures are an extension of the Hawkes processes previously presented and can be seen as infinite-dimensional Hawkes process. For the dynamics of incoming order flows they are not just depending on the present market price but also on the volume index. With this mathematical framework the incoming order arrival can depend not only on the past order placements but also on cancellations.

5.2.2 Hawkes Random Measures and Model Set Up

In order to introduce Hawkes random measures a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is required. The filtration has the common properties as such it is right-continuous. Further let $(U, \mathcal{U})$ be a measurable space endowed with a measure $m(du)$. In addition let p_t be an $(\mathcal{F}_t)$-point process on U. Furthermore let $N(dt, du)$ be a point measure on $[0, \infty) \times U$ defined as

$$N(I, A) = \#\{s \in I : p_s \in A\}, \qquad I \in \mathscr{B}(\mathbb{R}_+), A \in \mathcal{U}$$

where p_t is an $(\mathcal{F}_t)$-point process on U and $\mathscr{B}$ is the Borel σ-algebra. For further understanding and because it is necessary for some definitions it should be explained what is meant if a real-valued two-parameters process is $(\mathcal{F})$-progressive. Therefore let

$$\{h(t, x) : t \geq 0, x \in U\}$$

be a real-valued two-parameter process and this process is considered to be $(\mathcal{F})$-progressive if for every $t \geq 0$ the mapping $(w, s, x) \to h(w, s, x)$ limited to $\Omega \times [0, t] \times U$ is measurable relative to $\mathcal{F}_t \times \mathscr{B}([0, t]) \times \mathcal{U}$. Before Hawkes random measures can be introduced, Poisson random measures are going to be presented. The following definition is about the intensity processes of those.

Definition 5.2.1. *A non-negative $(\mathcal{F}_t)$-progressive process $\lambda(t, u)$ is called intensity process of $N(dt, du)$ with respect to the measure $m(du)$ if for any non-negative $(\mathcal{F}_t)$-predictable process $H(t, u)$ on U,*

$$\mathbb{E}\left[\int_0^t \int_U H(s, u) N(ds, du)\right] = \mathbb{E}\left[\int_0^t ds \int_U H(s, u) \lambda(s, u) m(du)\right].$$

With all the required definitions known it is possible to define Poisson random measures. Hence let $\lambda(t, u)$ be a non-negative $(\mathcal{F}_t)$-progressive process defined on U. For this type of processes it is achievable to construct a point measure $N(dt, du)$ on $[0, \infty) \times U$ with intensity process $\lambda(t, u)$ such that:

$$N([0, t], A) = \int_0^t \int_A \int_0^\infty \mathbf{1}_{\{z \leq \lambda(s,u)\}} N_0(ds, du, dz) \qquad t \geq 0, A \in \mathcal{U},$$

with $N_0(ds, du, dz)$ is a Poisson random measure on $[0, \infty) \times U \times [0, \infty)$ with intensity $dsm(du)dz$.
The upcoming definition forms the theoretical basis of this model.

Definition 5.2.2. *A random measure $N(dt, du)$ on $[0, \infty) \times U$ is called Hawkes random measure if its intensity process $\lambda(t, u)$ can be written as*

$$\lambda(t, u) = \mu(t, u) + \int_0^t \int_U \phi(s, u, v, t - s) N(ds, dv)$$

with $\mu(t, u) : [0, \infty) \times U \to [0, \infty)$ and $\phi(t, u, v, r) : [0, \infty) \times U^2 \times [0, \infty) \to [0, \infty)$ being $(\mathcal{F}_t)$-progressive.

The newly introduced processes of the definition of Hawkes random measures namely $\mu(t, u)$ and $\phi(t, u, v, r)$ designates the exogenous intensity respectively kernel of the Hawkes random measure $N(dt, du)$.
To ensure the existence of such Hawkes random measures the following Lemma is necessary. This is of importance to model the limit order book otherwise it could not be stated that the processes driving the limit order book are well defined.

Lemma 5.2.1. *Let $(U, \mathcal{U})$ be a measurable space then for any non-negative $(\mathcal{F}_t)$-progressive processes $\mu(t, u)$ and $\phi(t, u, v, r)$ satisfying*

$$\int_U \mu(t, u) m(du) + \sup_{u \in U} \int_U \phi(t, u, v, s) m(du) \leq C_0$$

with $C_0 > 0$ for any $t \in [0, T]$ there exists a Hawkes random measure with intensity process

$$\lambda(t, u) = \mu(t, u) + \int_0^t \int_U \phi(s, u, v, t - s) N(ds, dv).$$

Through this Lemma it is possible to model a limit order book with those measures. To be able to apply the Lemma further assumptions regarding the processes are needed. Since the requirements of the Lemma needs to be fulfilled by the processes modeling the events. All random processes in this chapter are defined on a filtered probability space $(\Omega, \mathbb{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, \mathbb{P})$. The dynamics of the $n-$th order book model will be denoted by a continuous-time stochastic process $(\mathbf{S}^{(n)}(t))_{0 \leq t \leq T}$ at any given time horizon $T > 0$. The stochastic process takes values in the Hilbert space

$$\mathcal{S} := \mathbb{R}^2 \times (L^2(\mathbb{R}; \mathbb{R}_+))^2, \qquad \|S\|^2_{\mathcal{S}^2} := |p_a|^2 + |p_b|^2 + \|v_a\|^2_{L^2} + \|v_b\|^2_{L^2}\, l.$$

The state of the limit order book at time $t \in [0, T]$ is described by

$$S^{(n)}(t) := \big(P_a^{(n)}(t), P_b^{(n)}(t), V_a^{(n)}(t), V_b^{(n)}(t)\big).$$

The best ask price is defined by the $\mathbb{R}$-valued process $P_a^{(n)}(t)$ and therefore the best bid price is the $\mathbb{R}$-valued process $P_b^{(n)}(t)$. The $V_a^{(n)}(t)$ denotes the volume density function at the ask side and $V_b^{(n)}(t)$ respectively on the bid side. The volume function is taking values in the Hilbert space L^2. The tick size in this continuous model is defined by $\delta_x^{(n)}$. Prices can become any value in $\{x_j^{(n)}, j \in \mathbb{Z}\}$, with $x_j^{(n)} := j\delta_x^{(n)}$ for $j \in \mathbb{Z}$ and $n \in \mathbb{N}$. The price interval which includes $x \in \mathbb{R}$ for all $n \in \mathbb{N}$ is defined by

$$\Delta^{(n)}(x) := [x_j^{(n)}, ^{(n)} x_{j+1}) \quad \text{for } x_j^{(n)} \leq x \leq x_{j+1}^{(n)}.$$

For any $t \in [0, T]$ the volume density function for the ask side $V_a^{(n)}(t, \cdot)$ is a càdlàg step function on the price grid, the same is true for the volume density function of the bid side $V_b^{(n)}(t, \cdot)$. To get the available volume for trading at price $x_j^{(n)}$ at time $t \in [0, T]$ the volume density function $V_{a/b}^{(n)}$ has to be integrated over $[x_j^{(n)}, x_{j+1}^{(n)})$. The state of the book at time $t = 0$ is deterministic for all $n \in \mathbb{N}$ and denoted by $S^{(n)}(0)$. There are eight different event types in this model which are able to change the state of the state of the limit order book:

- M_a: buy market orders
- M_b: sell market orders
- L_a: sell limit orders
- L_b: buy limit orders
- L_a^*: sell limit orders inside the spread
- L_b^*: buy limit orders inside the spread
- C_a: cancellations of sell limit orders
- C_b: cancellations of buy limit orders

Two minor assumptions are made regarding the market orders and about the limit orders placed inside the spread. The assumption about market orders is that these orders are not larger than the volume available at the top of the order book. This means that market orders do not lead to a price change. That assumption is just for mathematical convenience. Should however a market order arriving which is consuming liquidity it will be split up into smaller orders and be treated like small market orders. For limit orders placed inside the spread it is assumed that those orders change the price just by one tick.
To keep the model clearly arranged three sets will be defined. First a set which describes on which side of the order book an event takes place $\mathcal{I} = \{a, b\}$, where a stands for the ask side and b for the bid side. Then a set which helps to distinguish between market orders ("M") and limit orders placed inside the spread ("L"), $\mathcal{J} = \{M, L\}$. And the last set is $\mathcal{K} = \{L, C\}$ which is for limit orders placed outside the spread ("L") and cancellations ("C"). So if in the model appear subscripts as I, i, J, j and K, k it is assumed that $I, i \in \mathcal{I}$, $J, j \in \mathcal{J}$ and $K, k \in \mathcal{K}$. The processes driving the events affecting the order book are now presented for this model. Market orders on the sell side arrive according to an $(\mathcal{F}_t)$-random point measure $N_{aM}^{(n)}(dt)$ which takes values in $\mathbb{R}_+$ with intensity $\rho_{aM}^{(n)}(\mathbf{S}^{(n)}(t))\mu_{aM}^{(n)} dt$. The bid side follows the same random point measure assumptions but the subscript a is replaced by b and the same is true for the intensity where also the subscript is changed. The limit order which are placed inside the spread are submitted according to an $(\mathcal{F}_t)$-random point measure $N_{aL}^{(n)}(dt)$ and takes values in $\mathbb{R}_+$. The intensity for that measure is denoted by $\rho_{aL}^{(n)}(\mathbf{S}(t))\mu_{aL}^{(n)}$. Again this is for the ask side to describe the bid side the subscript a is replaced by b. The timestamp for arriving orders is described by $t \in [0, T]$.

The deterministic non-negative functions $\{\rho_{IJ}^{(n)}(S)\}_{I\in\mathcal{I},J\in\mathcal{J}}$ are chosen in such a way that they guarantee that bid and ask prices never cross and are defined on $\mathcal{S}$. Whereby the non-negative $(\mathcal{F}_t)$-progressive processes $\{\mu_{IJ}^{(n)}(t)\}_{I\in\mathcal{I},J\in\mathcal{J}}$ describe the dependence of price dynamics due to price changes in the past. For limit orders submitted outside the spread and cancellations of such it is assumed that those do not change prices. Furthermore in this model it is assumed that cancellations appear at random distances to the best price of the same side. Also the volume of the cancellations is random and can take different amounts of standing limit orders. For the limit orders outside the spread it is assumed that they also occur at random distance to the best price of the same side and with random volume. In mathematical detail they are modeled by an $(\mathcal{F}_t)$-random point measure $M_{a/bL}^{(n)}(dt,dx,dz)$ which takes values in $\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$. The z describes the order size and x denotes the difference between the best price and the price at which it is submitted. The intensity of the measure is $\lambda_{a/bL}^{(n)}(t,x)dtdxv_{a/bL}(dz)$. Cancellations of limit orders arrive x ticks away from the best quote on the same side of the book according to an $(\mathcal{F}_t)$-random point measure $M_{a/bC}^{(n)}(dt,dx,dz)$ on $\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$. The intensity of the measure for cancellations is given by $\lambda_{a/bC}^{(n)}(t,x)dtdxv_{a/bC}(dz)$ and a is the subscript for the ask side and b describes the bid side. The triple (t,x,z) stands for the time at which the event arrives, the number of ticks the limit order or cancellation is away from the best price at the same side and the size of an event. The processes $\{\lambda_{IK}^{(n)}(t,\cdot)\}_{I\in\mathcal{I},J\in\mathcal{J}}$ are mathematically speaking $(\mathcal{F}_t)$-progressive non-negative functions. Those describe the intensities of submitted limit orders and cancellations at different prices as a function depending on past orders and cancellations. The $(\mathcal{F}_t)$-progressive non-negative functions $\{v_{IK}(dz)\}_{I\in\mathcal{I},K\in\mathcal{J}}$ are probability measures taking values in $\mathbb{R}_+$. For these measures it is assumed that the fulfill $v_{IK}(|e^z-1|^4)<\infty$ for each $n\in\mathbb{N}$. It is deterministic and it denotes the size of arriving events. In order to make $M_{a/bL}^{(n)}(dt,dx,dz)$ a Hawkes random measure $v_{IK}(dz)$ has to be a Dirac measure. An overview over the different processes influencing the order book provides the table below.

Type	M_a	L_a^*	M_b	L_b^*
Notation	$N_{aM}^{(n)}(dt)$	$N_{aL}^{(n)}(dt)$	$N_{bM}^{(n)}(dt)$	$N_{bL}^{(n)}(dt)$
Space	$\mathbb{R}_+$	$\mathbb{R}_+$	$\mathbb{R}_+$	$\mathbb{R}_+$
Intensity	$\rho_{aM}^{(n)}(\mathbf{S})\mu_{aM}^{(n)}(t)$	$\rho_{aL}^{(n)}(\mathbf{S})\mu_{aL}^{(n)}(t)$	$\rho_{bM}^{(n)}(\mathbf{S})\mu_{bM}^{(n)}(t)$	$\rho_{bL}^{(n)}(\mathbf{S})\mu_{bL}^{(n)}(t)$

The processes which are not influencing the order book are summarized in that table.

Type	L_a	C_a	L_b	C_b
Notation	$M_{aL}^{(n)}(dt,dx,dz)$	$M_{aC}^{(n)}(dt,dx,dz)$	$M_{bL}^{(n)}(dt,dx,dz)$	$M_{bC}^{(n)}(dt,dx,dz)$
Space	$\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$	$\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_-$	$\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$	$\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_-$
Intensity	$\lambda_{aL}^{(n)}(t,x)dtdxv_{aL}(dz)$	$\lambda_{aC}^{(n)}(t,x)dtdxv_{aC}(dz)$	$\lambda_{bL}^{(n)}(t,x)dtdxv_{bL}(dz)$	$\lambda_{bC}^{(n)}(t,x)dtdxv_{bC}(dz)$

Subsequently, the results are used to derive the dynamics of the order book and this happens in the upcoming chapter.

5.2.3 Limit Order Book Dynamics and Scaling Limit

Keeping in mind that the prices can only change by one tick in this model the dynamics for the price are obtained by

$$P_a(t) = P_a(0) + \int_0^t \delta_x^{(n)} N_{aM}^{(n)}(ds) - \int_0^t \delta_x^{(n)} N_{aL}^{(n)}(ds),$$
$$P_b(t) = P_b(0) + \int_0^t \delta_x^{(n)} N_{bM}^{(n)}(ds) - \int_0^t \delta_x^{(n)} N_{bL}^{(n)}(ds).$$

That definition of price dynamics allows the spread to become negative. This means that the best ask price is smaller than the best bid price. Since this is not going to happen in reality the next condition assures that this also not gonna happen in this setting.

Assumption 5.2.1. *Let $S = (p_a, p_b, v_a, v_b)$ be any element in $\mathcal{S}$ with $p_a - p_b < \delta_x^{(n)}$ then it holds that*

$$\rho_{aL}^{(n)}(S) = \rho_{bL}^{(n)}(S) = 0.$$

For the purpose of describing the size of an order or cancellation in the n-th model $\delta_v^{(n)}$ is used. To obtain the dynamics of the volume density function one should consider an assumption made in advance. That is limit orders are submitted at a random price from which follows that the distance to the best quote of the same side is also random. For this reason and the fact that limit order placements are additive and cancellations are not but proportional to the volume standing in the order book the volume density functions are given by

$$\begin{aligned} V_a^{(n)}(t,x) = V_a^{(n)}(0,x) &+ \int_0^t \int_{\Delta^{(n)}(x-P_a^{(n)}(s-))} \int_{\mathbb{R}_+} \frac{\delta_v^{(n)}}{\delta_x^{(n)}} (e^z - 1) M_{aL}^{(n)}(ds,dy,dz) \\ &+ \int_0^t \int_{\Delta^{(n)}(x-P_a^{(n)}(s-))} \int_{\mathbb{R}_+} \frac{\delta_v^{(n)}}{\delta_x^{(n)}} V_a^{(n)}(s-, y + P^{(n)}(s-))(e^{-z} - 1) M_{aC}^{(n)}(ds,dy,dz), \\ V_b^{(n)}(t,x) = V_b^{(n)}(0,x) &+ \int_0^t \int_{\Delta^{(n)}(P_b^{(n)}(s-)-x)} \int_{\mathbb{R}_+} \frac{\delta_v^{(n)}}{\delta_x^{(n)}} (e^z - 1) M_{bL}^{(n)}(ds,dy,dz) \\ &+ \int_0^t \int_{\Delta^{(n)}(P_b^{(n)}(s-)-x)} \int_{\mathbb{R}_+} \frac{\delta_v^{(n)}}{\delta_x^{(n)}} V_b^{(n)}(s-, P_b^{(n)}(s-) - y)(e^{-z} - 1) M_{bC}^{(n)}(ds,dy,dz). \end{aligned}$$

Orders which influence the state of the limit order book arrive with a rate of $|\delta_n^{(n)}|^{-2}$. The events that do not have an impact on the limit order book arrive at rate $|\delta_v^{(n)}|^{-1}$. The assumptions of arrival rates are important in order to get a diffusive limiting dynamics for the price process and to receive a deterministic limiting dynamics for the volume density functions. To use the main advantage of Hawkes random measures clustering and cross-dependencies between orders have to be explained. For order arrivals it is allowed that the intensities depend not only on past price movements but also on past limit order placements and cancellations. This means that the model takes into account that limit orders are influenced by past events. That past events are influencing the order book is for example in [4]. Mathematically this dependence on past events is considered by the following arrival intensities

$$\mu_{IJ}^{(n)}(t) = \frac{1}{|\delta_x^{(n)}|^2}\hat{\mu}_{IJ}^{(n)}(t, \mathbf{S}^{(n)}(t-)) + \sum_{i\in\mathcal{I}, j\in\mathcal{J}} \int_0^t \phi_{IJ,ij}^{(n)}(t-s) N_{ij}^{(n)}(ds)$$
$$\sum_{i\in\mathcal{I}, k\in\mathcal{K}} \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\delta_x^{(n)}|^2}{\delta_v^{(n)}} \Phi_{IJ,ik}^{(n)}(y, t-s) M_{ik}^{(n)}(ds, dy, dz)$$

and

$$\lambda_{IK}^{(n)}(t,x) = \frac{1}{\delta_v^{(n)}}\hat{\lambda}_{IK}(t, \mathbf{S}^{(n)}(t-), x) + \sum_{i\in\mathcal{I}, j\in\mathcal{J}} \int_0^t \frac{|\delta_x^{(n)}|^2}{\delta_v^{(n)}} \psi_{IK,ij}(x, t-s) N_{ij}^{(n)}(ds)$$
$$+ \sum_{i\in\mathcal{I}, k\in\mathcal{K}} \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} \Psi_{IK,ik}(x, y, t-s) M_{ik}^{(n)}(ds, dy, dz).$$

For further understanding some terms will be presented. First the exogenous densities $\hat{\mu}_{IK}^{(n)}$ and $\hat{\lambda}_{IK}$ are going to be explained. Those densities depend just on the current state of the limit order book. Therefore these intensities take into account at what price level the order book is at the moment for the next submitted orders. The kernels $\phi_{IJ,ij}^{(n)}$ and $\Phi_{IJ,,ik}^{(n)}$ describe the influence of past events that have occurred on the price dynamics. Whereas the kernels $\psi_{IK,ij}^{(n)}$ and $\Psi_{IK,ik}^{(n)}$ define the effect of former events which did not influence the state of the book on placements and cancellations. The subscripts of the functions can be explained using $\phi_{bM,aL}^{(n)}$. Sticking with this example, it measures the influence of an ask limit order placed within the spread at time s on the intensity of a sell market order submitted at time t. The functions $\psi_{bL,bL}(x, t-s)$ and $\psi_{bC,bL}(x, t-s)$ describe the quantities. In this case it determine the impact of a bid side limit order x ticks away from the best bid quote at time s on the intensity of a bid side limit submitted order respectively cancellation at the same distance x from the at time t current price. The last function which needs to be explained is $\Psi_{bC,bL}(x, y, t-s)$ and this function provides information about the quantity. It measures the influence of an bid side limit order submission in the price interval $\delta^{(n)}(y)$ which contains y at time s at the event of a cancellation of a bid order at price level $\delta^{(n)}(y)$ at time t.

In order to derive the scaling limit for the order book some conditions and assumptions need to be made. The main assumptions concern the arrival intensities and the Hawkes kernels. Those guarantee in further consequence the convergence in law of the limit order book model consisting of stochastic differential equations and ordinary differential equations. The stochastic differential specifies the limiting price dynamics. The ordinary differential equation on the contrary describes the limiting volume dynamics. The first condition is about the convergence and moment on the initial states.

Assumption 5.2.2. *There exists a constant $C_0 > 0$ so that for any $n > 1$ and $I \in \mathcal{I}$,*

$$\mathbb{E}\Big[\Big\|\boldsymbol{S}^{(n)}(0)\Big\|^2_{\mathcal{S}^2}\Big] + \mathbb{E}\Big[\Big\|V_I^{(n)}(0,\cdot)\Big\|^4_{L^4}\Big] \leq C_0.$$

Further let $\boldsymbol{S}(0)$ be an $\mathcal{S}$-valued random variable which for $n \to \infty$ satisfies

$$\mathbb{E}\Big[\Big\|\boldsymbol{S}^{(n)}(0) - \boldsymbol{S}(0)\Big\|^2_{\mathcal{S}^2}\Big] \to 0.$$

Additionally assumptions are required regarding the order arrival intensities for market orders and limit orders placed inside the spread. These intensities have the product form $\rho_{IJ}^{(n)}\hat{\mu}_{IJ}^{(n)}$. For those assumptions are required to assure the convergence of the drift and volatility of the price process. Considering the difference between market order and submission inside the spread arrival intensities one get the expected increments of the ask and bid side

$$|\delta_x^{(n)}|^{-1}\big(\rho_{IM}^{(n)}(S)\hat{\mu}_{IM}^{(n)}(t,S) - \rho_{IL}^{(n)}(S)\hat{\mu}_{IL}^{(n)}(t,S)\big).$$

This difference can be rewritten and it becomes

$$\varrho_I^{(n)}(S)\hat{\mu}_{IM}^{(n)}(t,S) + \rho_{IL}^{(n)}(S)\hat{\beta}_I^{(n)}(t)$$

with

$$\varrho_I^{(n)}(S) := |\delta_x^{(n)}|^{-1}\big(\rho_{IM}^{(n)}(S) - \rho_{IL}^{(n)}(S)\big),$$
$$\hat{\beta}_I^{(n)}(t) := |\delta_x^{(n)}|^{-1}\big(\hat{\mu}_{IM}^{(n)}(t,S) - \hat{\mu}_{IL}^{(n)}(t,S)\big).$$

That condition assures subsequently the convergence of the factors to a continuous limit.

Assumption 5.2.3.

- *The functions $(\rho_{IJ}^{(n)}, \varrho_I^{(n)})$ are uniformly bounded.*
- *The functions $\{(\rho_{IJ}^{(n)}, \varrho_I^{(n)})\}_{n\geq 0}$ converge uniformly to Lipschitz continuous functions (ρ_{IJ}, ϱ_I).*

Due to the condition and assumption 5.2.1, that the spread never gets negative it can be concluded that

$$\rho_I := \rho_{IM} = \rho_{IL}.$$

In combination with the second condition follows that orders placed within the spread and market orders are on average equally probable

$$\hat{\mu}_I := \hat{\mu}_{IM} = \hat{\mu}_{IL}.$$

The following condition ensures that the limitation of intensities for the processes.

Assumption 5.2.4.

- *There exists a constant C_0 so that for every $p \in \{1,2,4\}$,*

$$\sup_{t\in[0,T],S\in\mathcal{S}} \left\{ |\hat{\mu}^{(n)}(t,S)| + |\hat{\beta}_I^{(n)}(t,S)| + \left\|\hat{\lambda}_{IK}(t,S,\cdot)\right\|_{L^p} \right\} \leq C_0$$

and for any $\epsilon > 0,\ t,t' \in [0,T], S, S' \in \mathcal{S}$,

$$\left\|\hat{\lambda}_{IK}(t',S',\cdot+\epsilon) - \hat{\lambda}_{IK}(t,S,\cdot)\right\|_{L^p} \leq C_0(\epsilon + |t-t'| + \left\|S-S'\right\|_{\mathcal{S}^2})$$

- *Let $\hat{\mu}_{IJ}(t,S)$ and $\hat{\beta}_I(t,S)$ be Lipschitz continuous functions such that*

$$\sup_{t\in[0,T],S\in\mathcal{S}} \left\{ |\hat{\mu}_{IJ}^{(n)}(t,S) - \hat{\mu}_{IJ}(t,S)| + |\hat{\beta}_I^{(n)}(t,S) - \hat{\beta}_I(t,S)| \right\} \to 0$$

Scaling conditions on Hawkes kernels are still missing.The excepted price increments contain in this model an extra term. The additional part originates from the impact of past events on orders influencing the order book and those new parts are given by

$$\theta_{I,ij}^{(n)}(t) := |\delta_x^{(n)}|^{-1}\Big(\phi^{(n)}i_{IM,ij}(t) - \phi_{IL,ij}^{(n)}(t)\Big),$$

$$\Theta_{I,ik}^{(n)}(y,t) := |\delta_x^{(n)}|^{-1}\Big(\Phi_{IM,ik}^{(n)}(y,t) - \Phi_{IL,ik}^{(n)}(y,t)\Big).$$

The last condition remaining provides regularity conditions on the Hawkes kernels. Especially the kernels describing the impact of past events on limit order submissions and cancellation arrivals. Moreover the condition ensures the convergence of Hawkes kernels characterizing the influence of past events on prices to regular functions.

Assumption 5.2.5.

- *For any $\epsilon > 0,\ p \in 1,2,4$ there exists a constant C_0 such that*

$$\sup_{t\in[0,T],y\in\mathbb{R}} \left\{ \|\psi_{IK,ij}(\cdot,t)\|_{L^p} + \|\Psi_{IK,ik}(\cdot,y,t)\|_{L^p} \right\} < C_0,$$

$$\sup_{t\in[0,T],y\in\mathbb{R}} \Big\{ \|\psi_{IK,ij}(\cdot+\epsilon,t) - \psi_{IK,ij}(\cdot,t)\|_{L^p} + \|\Psi_{IK,ik}(\cdot+\epsilon,y,t) - \Psi_{IK,ik}(\cdot,y,t)\|_{L^p} \Big\} \leq C_0\epsilon$$

- *The functions*

$$\kappa^{(n)}(y,t) := \big(\phi_{IJ,ij}^{(n)}(t), \Phi_{IJ,ij}^{(n)}(y,t), \theta_{I,ik}^{(n)}(t), \Theta_{I,ik}^{(n)}(y,t)\big)_{I,i\in\mathcal{I},J\in\mathcal{J},k\in\mathcal{K}}$$

are uniformly bounded and furthermore converge consistently to the following functions

$$\kappa(y,t) := \big(\phi_{IJ,ij}(t), \Phi_{IJ,ij}(y,t), \theta_{I,ik}(t), \Theta_{I,ik}(y,t)\big)_{I,i\in\mathcal{I},J\in\mathcal{J},k\in\mathcal{K}}$$

which are uniformly Lipschitz continuous in the variable t that specify the time

$$\sup_{t\in[0,T],y\in\mathbb{R}} |\kappa^{(n)}(y,t) - \kappa(y,t)| \to 0.$$

With the aid of this condition and the definitions of $\theta^{(n)}_{I,ij}$ and $\Theta^{(n)}_{I,ik}$ it can be claimed that

$$\phi_{I,ij} := \phi_{IM,ij} = \phi_{IL,ij},$$
$$\Phi_{I,ik} := \Phi_{IM,ik} = \Phi_{IL,,ik}.$$

Due to this mathematical deduction it follows that the impact of market orders from the same side and limit orders placed inside the spread are equal. In order to derive the main result of this model some more definitions need to be introduced. Therefore let

$$\alpha_{iL} = v_{IL}(e^z - 1),\ \alpha_{iC}(e^{-z} - 1)$$

and

$$\bar{\phi}_{Ii} = \phi_{I,iM} + \phi_{I,iL},\ \bar{\psi}_{IK,i} = \psi_{IK,iM} + \psi_{IK,iL},\ \bar{\theta}_{Ii} = \theta_{I,iM} + \theta_{I,iL}.$$

The effect of events which drive the price on themselves is defined by $\bar{\phi}_{Ii}$. Whereas the influence of those on events that do not change the price is measured by $\bar{\psi}_{IK,i}$. The impact of price determining events on the price dynamics is expressed by $\bar{\theta}_{Ii}$. In order to better present the main result the following functions are used

$$\beta^{(n)}_I(t) := \delta^{(n)}_x(\mu^{(n)}_{IM}(t) - \mu^{(n)}_{IL}(t)),$$
$$\mathbf{D}^{(n)}(t,S) := \left(|\delta^{(n)}_x|^2\mu^{(n)}_{ij}(t,S), \delta^{(n)}_v\lambda^{(n)}_{ik}(t,S,\cdot)\right)_{i\in\mathcal{I},j\in\mathcal{J},k\in\mathcal{K}}.$$

Whereby the latter needs some more explanation. It concerns the vector $\mathbf{D}^{(n)}(t,S)$ which exists in the space $\mathcal{D} := \mathbb{R}^4 \times (L^1(\mathbb{R};\mathbb{R}_+) \cap L^2(\mathbb{R};\mathbb{R}_+))^4$ for every $n \in \mathbb{N}$. Furthermore with the norm $\|\cdot\|_{\mathcal{D}^2_{1,2}} := \|\cdot\|_{\mathcal{D}^2_1} + \|\cdot\|_{\mathcal{D}^2_2}$, where $\|\cdot\|_{\mathcal{D}^p_q}$ is defined for any $p,q \in \mathbb{Z}_+$ and for every $D := (D_1,\dots,D_S) \in \mathcal{D}$ by

$$\|D\|^p_{\mathcal{D}^p_q} = \sum_{k=1}^{4} |D_k|^p + \sum_{k=5}^{8} \|D_k\|^p_{L^q}$$

the space $\mathcal{D}$ becomes a Banach space. Now it is possible to state the main result regarding the convergence of the limit order book.

Theorem 5.2.2. *Assuming that the conditions 5.2.2-5.2.5 hold then it can be stated that*

$$\left(\boldsymbol{S}^{(n)}, \boldsymbol{D}^{(n)}, \beta^{(n)}_a, \beta^{(n)}_b\right) \Rightarrow (\boldsymbol{S}, \boldsymbol{D}, \beta_a, \beta_b)$$

weakly in $\mathbb{D}(\mathbb{R}_+, \mathcal{S} \times \mathcal{D} \times \mathbb{R}^2)$, *with* $\boldsymbol{S} = (P_a, P_b, V_a, V_b)$ *and* $\boldsymbol{D} = (\mu_{ij}\lambda_{ik})_{i\in\mathcal{I},j\in\mathcal{J},k\in\mathcal{K}}$ *with* $\mu_i := \mu_{iM} = \mu_{IL}$ *for* $i \in \mathcal{I}$. *Additionally the limit solves the following stochastic dynamic* *system*

$$P_a(t) = P_a(0) + \int_0^t \Big[\rho_a(\boldsymbol{S}(s))\beta_a(s) + \varrho_a(\boldsymbol{S}(s))\mu_a\Big]ds + \int_0^t \sqrt{2\rho_a(\boldsymbol{S}(s))\mu_a(s)}dB_a(s),$$
$$P_b(t) = P_b(0) - \int_0^t \Big[\rho_b(\boldsymbol{S}(s))\beta_b(s) + \varrho_b(\boldsymbol{S}(s))\mu_b\Big]ds + \int_0^t \sqrt{2\rho_b(\boldsymbol{S}(s))\mu_b(s)}dB_b(s),$$
$$V_a(t,x) = V_a(0,x) + \int_0^t \Big[\alpha_{aL}\lambda_{aL}(s, x - P_a(s)) + \alpha_{aC}\lambda_{aC}(s, x - P_a(s))V_a(s,x)\Big]ds,$$
$$V_b(t,x) = V_b(0,x) + \int_0^t \Big[\alpha_{bL}\lambda_{bL}(s, P_b(s) - x) + \alpha_{bC}\lambda_{bC}(s, P_b(s) - x)V_b(s,x)\Big]ds,$$

where (B_a, B_b) *is a standard two-dimensional Brownian motion as explained in the theoretical part of this thesis, and*

$$\mu_I(t) = \hat{\mu}_I(t, \boldsymbol{S}(t)) + \sum_{i\in\mathcal{I}} \int_0^t \bar{\phi}_{Ii}(t-s)\rho_i(\boldsymbol{S}(s))\mu_i(s)ds$$
$$+ \sum_{i\in\mathcal{I},k\in\mathcal{K}} \int_0^t \int_{\mathbb{R}} \Phi_{I,ik}(y, t-s)\lambda_{ik}(s,y)dsdy,$$
$$\lambda_{IK}(t,x) = \hat{\lambda}_{IK}(t, \boldsymbol{S}(t), x) + \sum_{i\in\mathcal{I}} \int_0^t \bar{\psi}_{IK,i}(x, t-s)\rho_i(\boldsymbol{S}(s))\mu_i(s)ds$$
$$\sum_{i\in\mathcal{I},k\in\mathcal{K}} \int_0^t \int_{\mathbb{R}} \Psi_{IK,ik}(x, y, t-s)\lambda_{ik}(s,y)dsday,$$
$$\beta_I(t) = \hat{\beta}(t, \boldsymbol{S}(t)) + \sum_{i\in\mathcal{I}} \int_0^t \bar{\theta}_{Ii}(t-s)\rho_i(\boldsymbol{S}(s))\mu_i(s)ds$$
$$\sum_{i\in\mathcal{I},k\in\mathcal{K}} \int_0^t \int_{\mathbb{R}} \Theta_{I,ik}(y, t-s)\lambda_{ik}(s,y)dsdy.$$

Proof. See [12]. □

Bibliography

[1] F. Abergel, M. Anane, A. Chakraborti, A. Jedidi, and I. Muni Toke. *Limit Order Books*. Cambridge University Press, 2016.

[2] E. Bacry, I. Mastromatteo, and J. Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1):1550005, 2015.

[3] R. F. Bass. *Stochastic Processes*. Cambridge University Press, 2011.

[4] Z. Eisler, J.-P. Bouchaud, and J. Kockelkoren. The price impact of order book events: Market orders, limit orders and cancellations. *Quantitative Finance*, 12(9):1395–1419, 2012.

[5] J. Farmer, L. Gillemot, F. Lillo, S. Mike, and A. Sen. What really causes large price changes? *Quantitative Finance*, 4(4):383–397, 2004.

[6] M. D. Gould, M. A. Porter, S. Williams, M. McDonald, D. J. Fenn, and S. D. Howison. Limit order books. *Quantitative Finance*, 13(11):1709–1742, 2013.

[7] A. Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B*, 33(3):438–443, 1971.

[8] A. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[9] A. Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198, 2018.

[10] U. Horst and D. Kreher. A weak law of large numbers for a limit order book model with fully state dependent order dynamics. *SIAM Journal on Financial Mathematics*, 8(1):314–343, 2017.

[11] U. Horst and D. Kreher. Second order approximations for limit order books. *Finance and Stochastics*, 22(4):827–877, 2018.

[12] U. Horst and W. Xu. A scaling limit for limit order books driven by Hawkes processes. *SIAM Journal on Financial Mathematics*, 10(2):350–393, 2019.

[13] L. Massoulié. Stability results for a general class of interacting point processes dynamics, and applications. *Stochastic Processes and their Applications*, 75:1–30, 1998.

[14] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, Berlin, Heidelberg, 3 edition, 1992.