Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Quoting behaviour of a market-maker under different exchange fee structures

Author: **Bc. Rastislav Kiseľ**

Supervisor: **PhDr. Jozef Baruník, Ph.D.**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, December 21th, 2017

Signature

Acknowledgments

The author is especially grateful to his supervisor PhDr. Jozef Baruník, Ph.D. and to his parents.

Furthermore, this thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 681228.

Abstract

During the last few years, market micro-structure research has been active in analysing the dependence of market efficiency on different market characteristics. Make-take fees are one of those topics as they might modify the incentives for participating agents, e.g. broker-dealers or market-makers. In this thesis, we propose a Hawkes process-based model that captures statistical differences arising from different fee regimes and we estimate the differences on limit order book data. We then use these estimates in an attempt to measure the execution quality from the perspective of a market-maker. We appropriate existing theoretical market frameworks, however, for the purpose of finding optimal market-making policies we apply a novel method of deep reinforcement learning. Our results suggest, firstly, that maker-taker exchanges provide better liquidity to the markets, and secondly, that deep reinforcement learning methods may be successfully applied to the domain of optimal market-making.

JEL Classification	C32, C45, C61, C63
Keywords	make-take fees, Hawkes process, limit order book, market-making, deep reinforcement learning

Author's e-mail	kiselrastislav@gmail.com
Supervisor's e-mail	barunik@fsv.cuni.cz

Abstrakt

Posledných pár rokov sa výskum ohľadom mikro-štruktúry trhu snaží odhaliť závislosti tržnej efektivity na rozličných charakteristikách trhov. Poplatkový systém make-take je práve jedným z takýchto tém, keďže môže mať vplyv na správanie agentov, napr. broker-dealerov či market-makerov. V tejto diplomovej práci navrhneme model založený na Hawksovom procese, ktorý bude mať za cieľ zachytiť štatistické rozdiely vyplývajúce z odlišných poplatkových režimov a zároveň tieto rozdiely odhadnúť na dátach z limitnej knihy. Následne sa pokúsime použiť tieto odhady na zmeranie kvality obchodovania z pohľadu market-makera. Za týmto účelom použijeme existujúce tržné modely, avšak, optimálnu funkciu market-makera budeme hľadať pomocou metódy hlbokého spätnoväzbového učenia. Naše výsledky implikujú, že maker-taker burzy poskytujú kvalitnejšiu likviditu a zároveň, že hlboké spätnoväzbové učenie môže byť úspešne použité v oblasti hľadania optimálnych politík market-makera.

Klasifikace JEL	C32, C45, C61, C63
Klíčová slova	make-take poplatky, Hawkesov proces, kniha limitných objednávok, market-making, hlboké spätnoväzbové učenie

E-mail autora	kiselrastislav@gmail.com
E-mail vedoucího práce	barunik@fsv.cuni.cz

Contents

List of Tables

List of Figures

Acronyms

ADF Alternative Display Facility

AMEX American Stock Exchange, currently NYSE MKT LLC

ATS Automated Trading System

CTA Consolidated Tape Association

ECN Electronic Communication Network

FINRA Financial Industry Regulatory Authority

GMM Generalized Method of Moments

HFT High-Frequency Trading

HJB Hamilton-Jacobi-Bellman equations

ISO Inter-market Sweep Order

MLE Maximum Likelihood Estimation

NASDAQ National Association of Securities Dealers Automated Quotations

NBBO National Best Bid/Offer

Reg NMS Regulation National Market Service

RL Reinforcement Learning

SEC U.S. Securities and Exchange Commission

SIP Security Information Processor

SLSQP Sequential Least Squares Programming

TAQ Trade and Quote

TRF Trade Reporting Facility

TSX Toronto Stock Exchange

UTP Unlisted Trading Privileges

Master's Thesis Proposal

Author	Bc. Rastislav Kiseľ
Supervisor	PhDr. Jozef Baruník, Ph.D.
Proposed topic	Quoting behaviour of a market-maker under different exchange fee structures

Motivation The exchanges in the US have been undergoing some detrimental changes in the last decade. Stabilisation of algorithmic trading as a new normal practice among all market participants, proliferation of exchanges or migration of volume to non-lit venues present a challenging environment for operations of individual exchanges. The desire to remain profitable constantly forces them to implement innovations that try to cater to different market participants. Either renting co-location facilities to attract HFT, enabling proprietary dark liquidity trading options (e.g. hidden orders or exchange-owned dark pools) to attract institutional investors or implementing tailored fee structures.

This thesis will consider the effect of the last of the above-mentioned innovations on the quality of the equity markets. There are two fee structures present across US exchanges (although they are applied in several types of markets we will primarily have equity markets in mind). The first one, so called maker-taker arrangement, supplies rebates to parties quoting limit orders (i.e. order to buy/sell at the defined price) and bills parties that "take" this liquidity by posting market orders (i.e. order to buy/sell at the best existing price). The taker-maker setting follows the same but only inverted scheme, meaning paying rebates for market orders and extracting fees for limit orders. There is an empirical evidence that parameters describing market quality change with the fee structure and the level of rebate. Particularly, several authors have found that the fill rate is higher for markets with taker-maker setting or low taker fees (see Aldridge 2013; Battalio, R. *et al.* 2015; Cardella, L. *et al.* 2013; Yiping, L. *et al.* 2016).

We will try to incorporate this empirical evidence into modelling the behaviour of HFT market-maker on markets with different fee structures and with different fee levels. Therefore, the HFT market-maker will be deciding to quote under the trade-off between the fill rate and rebate, i.e. the higher the arrival of market orders the lower the rebate (or even payment of fees) for limit orders and vice versa. We will try to evaluate the market quality in regard to market-maker quoting implied by different fee structures.

Hypotheses

> Hypothesis #1: P&L of the market-maker will be higher in a market with maker-taker setting
>
> Hypothesis #2: Bid-ask spread will be lower in a market with fee structure preferring market orders
>
> Hypothesis #3: Bid-ask spread will be lower under stress conditions in a market with fee structure preferring market orders

Methodology The study will be based on a zero intelligence model of the limit order book. Zero intelligence models are a class of models assuming random trading of agents on the market in a continuous double action (i.e. buying and selling) setting. Two zero intelligence order books with different fee structures (and therefore different marketable order arrivals) will provide the background where the rational agent will be situated. Therefore, integrated with the model will be objective function of this market-maker, which will incorporate his aim to maximize expected terminal profit.

The study will simulate the model after establishing it. The simulation will be calibrated either on the existing empirical evidence mentioned earlier or on the estimation of fill rates on particular datasets (the path taken will depend on the author's access to proper datasets). The result of simulation is expected to be the relationship between the quoted optimal bid-ask spread of the market-maker and the fee structure of the market.

Expected Contribution We are persuaded of the soundness of our venture by the fact that the high-frequency companies are extensively overtaking market-making business from traditional players in equity markets and the rebates from liquidity provision are evidenced to constitute the bulk of the profit of several HFT traders (see Biais, B. *et al.* 2014). However, existing academic literature doesn't consider the nexus of market-making behaviour and exchange fee structure. Bearing these points in mind, to gauge the impact of market fee structures on the quoting behaviour of market-makers is a topic that warrants attention.

Therefore, we believe our study could enrich current debate on market microstructure by presenting the effects of institutional fee arrangement on the quality of liquidity provision.

Outline

1. Introduction
2. Literature Review
3. Model
4. Calibration on the data
5. Simulations
6. Conclusions

Core bibliography

Abergel, F., M. Anane, A. Chakraborti, A. Jedidi & I. M. Toke (2016): *Limit Order Books.* Cambridge University Press. ISBN ISBN 978-1-107-16398-0

Ait-Sahalia, Y. & M. Saglam (2016): "High Frequency Market Making." *SSRN Electronic Journal*

Avellaneda, M. & S. Stoikov (2008): "High-frequency trading in a limit order book." *Quantitative Finance*, **8(3)**: 217-224

Battalio, R., S.A. Corwin & R. Jennings (forthcoming 2016): "Can Brokers Have it All? On the Relation between Make?Take Fees And Limit Order Execution Quality." *The Journal of Finance*

Biais, B. & T. Foucault (2014): "HFT and market quality." *Bankers, Markets & Investors*, **128**, 5-19

Cardella, L., J. Hao & I. Kalcheva (2015): "Make and take fees in the US equity market." *SSRN Electronic Journal*

Carmona, R. & K. Webster (2012): "High frequency market making." *arXiv preprint* arXiv: 1210.5781

Cont, R., S. Stoikov & R. Talreja (2010): "A stochastic model for order book dynamics." *Operations research*, **58(3)**, pp. 549-563

Fodra, P. & M. Labadie (2012): "High-frequency market-making with inventory constraints and directional bets." *arXiv preprint* arXiv:1206.4810

Guéant, O., Ch.-A. Lehalle & J. Fernandez-Tapia (2012): "Dealing with the Inventory Risk: A solution to the market making problem" *Mathematics and Financial Economics* **7(4)**: pp. 477-507

Lin, Y. & P. L. Swan (2016): "Maker-Taker Fee, Liquidity Competition, and High Frequency Trading." *SSRN Electronic Journal*

Peržina, V. & J. M. Swart (2016): "How many market makers does a market need?." *arXiv preprint* arXiv:1612.00981

Šmíd, M. (2016): "Estimation of zero-intelligence models by L1 data." *Quantitative Finance*, **16(9)**, pp. 1423-1444.

Author Supervisor

Chapter 1

Introduction

Currently, a highly fractioned nature of US markets forces exchanges to compete with each other in the quest to attract high-quality liquidity. They may decide to provide different features, e.g. multiple options of data access or plethora of order types and fee structures, catering to individual market populations. Each modification of market micro-structure propels the overall dynamics of the markets. However, not all directions financial markets take may be normatively warranted. For example, if there would be an evidence that providing access to proprietary high-frequency trading worsens the execution quality for large buy-side institutions (e.g. insurance companies, pension funds, asset-management companies...), the sensibility of fixing this market structure might be questioned.

Providing quality analysis of market micro-structure is therefore of significant importance not only to the financial institutions themselves, but also to the policy-makers. It is therefore not surprising that even Permanent Subcommittee on Investigations of the US Senate requested a hearing attended by several executives of prominent US stock exchanges regarding possible conflicts of interests stemming from currently popular market liquidity-based fee models (Permanent Subcommittee on Investigations, 2014). The discussion of the Subcomittee also revolved around the study of Battalio *et al.* (2016) who have investigated the quality of market liquidity with regard to different fee structures. US stock exchanges currently employ either maker-taker or taker-maker fees (exchanges charge several other types of fees, however, we will focus exclusively on liquidity-based fees). In the former case, the exchanges pay rebates to liquidity providers and take fees from liquidity demanders, while in the later case, the liquidity providers pay fees that are further disseminated to liquidity demanders. Battalio *et al.* (2016) gathered evidence that maker-taker exchanges provide worse liquidity conditions (e.g. lower probability of fill and lower fill speed conditional on fill). Therefore, potential conflict of interest might arise when broker-dealers send their orders to maker-taker exchanges, where they obtain financial rebates, while at the same time decreasing execution quality for their clients.

Although solely financial incentive of broker-dealers is unequivocal under these market circumstances, the decision-making of market-makers is more complicated. If the claims of Battalio *et al.* (2016) would be true, market-makers have to find a balance between market quality and rebate provision. Higher probability of order fill on some exchanges would mean that the quantity of trading would be higher and the risk of informed trading would be lower there. However, by posting quotes on such exchanges the market-maker would incur taker-maker fees. The opposite situation holds for his engagement on maker-taker exchanges.

Although several authors have weighted in the debate (even before Battalio *et al.*, 2016), no comprehensive modelling approach for this feature of market micro-structure has been yet proposed. In this diploma thesis, we try to create a model that would be able to capture differences between probability of fill and fill rates in connection with maker-taker and taker-maker exchanges. This modelling will be based on a self-exciting point process system, called Hawkes process. Our initial idea was then to calibrate the model on market data and by tools of optimal control show what are the optimal quotes from the perspective of the market-maker. However, the structure and quality of available datasets challenged the first part of this direction. We have therefore decided to optimize the Hawkes model, obtain the corresponding market quality parameters and then use these parameters as an input to some existing model of optimal market-making. We have decided to implement a well-known model developed by Avellaneda & Stoikov (2008) that has been employed by several other research teams since its publication (see Fodra & Labadie, 2012, Guéant *et al.*, 2013, or Guéant, 2017). However, we have appropriated only the fundamental market framework of the original paper. We will not be using methods of stochastic optimal control via Hamilton-Jacobi-Bellman (HJB) equations, as the authors did. Instead, we propose solving the problem of finding the optimal quotes of the market-maker by reinforcement learning. Particularly, we will be implementing Double Deep Q-Network used initially by Mnih *et al.* (2015) to solve optimal control of game playing. An advantage stemming from using this method instead of analytical solution via HJB equations is the generality this optimal control method provides. In the original paper, in order to quote optimally, the market-maker needs to have a knowledge of the underlying market parameters (i.e. volatility and market order arrival intensity). Also, more complicated market structure models may be intractable. Both of these problems are surpassed by reinforcement learning. As of the date of writing this thesis, the author is not aware of any existing literature discussing deep reinforcement learning tools in the domain of market-making.

The thesis is structured as follows. Second chapter briefly summarizes the current state of research in the domains of make-take fees, Hawkes process limit order book modelling and finally optimal market-making and deep reinforcement learning. Next chapter introduces Hawkes processes and proposes a Hawkes process-based model for the limit order arrival from maker-taker and taker-maker exchanges and market order arrivals. The simulation of the process with successive optimization is provided as an evidence that the maximum likelihood optimization techniques work well even in multidimensional case. A chapter describing origin of the data and adjustments needed for the modelling follows. Chapter 5 presents the results of optimizations of the Hawkes process on the data. Next chapter introduces and applies deep reinforcement learning methods on the market framework of Avellaneda & Stoikov (2008). Conclusion follows.

Chapter 2

Literature review

2.1 Impact of Liquidity-based Fees on Market Quality

The spread of make-take fees [1] can be traced back to the dissemination of alternative trading systems (ATS) in the late 90ties. One of the first exchanges to offer maker-taker fee structure was Island Electronic Communication Network (ECN). This plan to attract order flow was met with success and other ATSs followed the suit. Traditional stock exchanges had to jump on the bandwagon in order to compete with the ATSs. By the mid-2000s, most US equity trading was done on platforms that charged some kind of fee for demanding liquidity and paid rebates to liquidity suppliers. The federal agency for securities supervision, U.S. Securities and Exchange Commission (SEC), quickly noticed the raging race for attracting order flow through liquidity rebates and approved a regulation regarding access fees as part of Reg NMS (Regulation National Market Service) in 2005 (implemented in 2007). Rule 610 established ceiling of 30 cents per 100 shares on the taker fee. This ceiling is still in place and besides that, all equity exchanges have to file fee changes to the SEC (see Securities and Exchange Commission Division of Trading and Markets, 2015). Currently, all of the 13 registered stock exchanges follow some kind of liquidity-based fee pricing system. While three of those (Bats BYX Exchange, Inc., Bats EDGA Exchange, Inc. and NASDAQ BX, Inc.) follow inverted, i.e. taker-maker, pricing system, majority adhere to the maker-taker fee schedule.

[1] In the following text, I will refer to the system when an exchange imposes fees or rebates based on the liquidity supply or demand as make-take fees, while by maker-taker (taker-taker) system I will denote the specific version of make-take fees when liquidity providers (demanders) receive rebates and liquidity demanders (providers) pay fees.

The benchmark argument on the impact of make-take fees on the quality of trading assumes frictionless market conditions. Under this assumption, any change in the fee distribution, i.e. any change in the maker fee and taker fee (or equivalently rebates) as a proportion of the total fee (holding this total fee constant) is reflected in the change of raw, i.e. quoted, bid-ask spread. Therefore, in case of substitution of maker-taker fees for undiscriminating fee payment, net spread, computed as a raw bid-ask spread plus two times the taker fee (i.e. actual costs for a roundabout market order), should remain the same. The reason is that higher taker fee would force some (less informed) agents to trade with limit orders, therefore increasing competition on the limit order book. This increased competition will tighten the spreads and the resulting net fees will be identical to the ones paid by the agents before the fee reform. Possible frictions (e.g. agency problems of brokers, obligatory minimum tick size, inter-exchange competition or competition of exchanges with limited display venues, so-called "dark pools") are not taken into account in this argument. We will firstly take a look at the formulation of the impact of make-take fees under perfect competition as postulated by Angel *et al.* (2011) and later formalized by Colliard & Foucault (2012) and then we will consider possible deviations in form of market imperfections and consequences on market quality implied by these deviations.

Colliard & Foucault (2012) model the market of investors trying to trade with a predetermined deadline for trade execution. They can either choose to enter a limit order market as a maker or taker or a dealer market. As a maker they run the risk of missing the deadline for the execution, however, they can obtain the desired price for their asset, and as a taker, they have to accept market maker's quotations, but there is no risk of missing the deadline. Similar trade-off is present in the dealer market. Investors are further differentiated according to their deadline intensity on patient and impatient ones - discount factors for both groups are different. The first corollary originating from an equilibrium states that the cum fee bid-ask spread is independent of the maker/taker fee breakdown as the total fee stays the same. Therefore, any change in the composition of the maker/taker fee payment structure impacts solely the raw bid-ask spread, while leaving all other market quality parameters (e.g. trading volume, fill rate, investors' welfare, etc.) untouched.

One of the first responses to the thesis stating that the composition of fee payment does not affect market quality came from Foucault *et al.* (2013). The authors model a market with maker and taker agents with specific monitoring costs. Their model predicts, besides other things, that when the tick size is zero, any change in the fee composition is neutralized by the adequate change in the raw bid-ask spread, leaving the cum fee spread and market quality parameters unaffected. However, when there is a positive tick size, the neutralization response is a step function. Under these conditions, any fee change not copying the price grid will change the market quality parameters. The logical conclusion is that influencing the market by changing the market microstructure is possible only by imposing "between tick size" fee changes. As the authors state: "... as long as the tick size is not zero, the make/take pricing model has true economic consequences: it affects the monitoring intensities, the trading rate, and market participants' welfare."

Another possible source of market imperfection hindering the smooth adjustment of raw spread to the changes in the fee composition was described by Brolley & Malinova (2013). The main contribution of the study lies in investigating the differences in the impact between the setting when brokers are passing the maker/taker fees to end investors and the setting when investors are paying flat fees to the broker who pays the liquidity access costs or receives the rebates from the exchange. Under former conditions, the authors arrive at the same conclusion as obtained by Colliard & Foucault (2012) that the market quality is independent of the fee distribution between makers and takers. However, under the latter setting, the investors are primarily affected by the changes in the raw bid-ask spread. Therefore, when the maker rebate increases, the raw bid-ask spread decreases and investors are more motivated to post market orders. The trading volume increases as a consequence. However, the authors also find a certain population of uninformed investors who decide to abstain from the market because the spread is smaller and decreased adverse selection costs (due to an increase in market order trading of other uninformed investors) are not enough to offset the first effect. Similar conclusions are drawn by O'Donoghue (2015) who strips the model of Colliard & Foucault (2012) of deadlines and a chance of non-execution for investors' orders. The author finds that in the case of keeping the net fee fixed, the increase in the taker fee and maker rebate lowers the market participation, while increasing the share of market orders. The change is driven by investors who would prefer to trade by posting limit orders before, but now decide to trade with market orders. The change in their behaviour is affected by the fact that investors pay flat commission to the broker and therefore the lower bid-ask spread incentivizes posting market orders as the surplus from them becomes relatively higher than the surplus from posting limit orders. Therefore, simultaneously with rising taker fees and maker rebates the fill rate increases.

Empirical analysis have consistently shown that the benchmark argument does not hold in real markets and the make-take fee structures do indeed affect market quality, thus giving credit to the imperfect market models. However, empirical analysis are not unanimous regarding the sign of the effect of the change in the maker/taker fee distribution (keeping the total fee fixed).

Malinova & Park (2015) find evidence in accordance with the model of Brolley & Malinova (2013) and O'Donoghue (2015). They studied the selected group of stocks that underwent a change in the fee structure on the Toronto Stock Exchange (TSX) in 2005. The change consisted of the transformation from volume-based taker fee (makers did not pay or receive any payments) to the maker-taker pricing system. The authors document a betterment of prices (therefore lowering of raw spreads) as a consequence of installing a maker rebate, thus corroborating the hypothesis of Colliard & Foucault (2012). However, unlike the perfect market models predict, market quality parameters changed. The authors present evidence of an increase in the trading volume, the fill rate for limit orders and a decrease of price impact (i.e. signed change of the midpoint following a trade) of marketable orders (that is equivalent to lower adverse selection costs).

However, some authors have questioned the generalizability of the findings of Malinova & Park (2015). Lin *et al.* (2016) argue that TSX was a monopoly exchange in Canada at that period (i.e. during 2005), and therefore the impact of market competition was almost non-existent in those circumstances. They hypothesize that there might have been no change in the informativness of the order flow as a consequence of starting the TSX experiment due to the monopolistic nature of the exchange. However, similar venture would result in a major shift in the character of the order flow in the US circumstances. Even though Malinova & Park (2015) mention that the stocks they study were cross-listed with NASDAQ and AMEX, the investors' desire to trade equities on foreign exchanges may have been restrained back then. Also, it could be argued that the presence of HFT companies have been lower back in 2005 and therefore the circumstances now differ in a significant way from those studied. Lin *et al.* (2016) together with Battalio *et al.* (2016) analyse more recent data from US stock exchanges and find contrary evidence to Malinova & Park (2015). The former study analyses data documenting temporary decrease in the access fees (from 30/29 basis points per share to 5/4 basis points per share for takers/makers) for a sample of 14 selected stocks on NASDAQ during 2015. The authors have found that the market share, routed volume and depth share at the NBBO (i.e. average percentage quote size when a NASDAQ quote is at the NBO/NBB) all declined as a result of the reform. They argue that this effect was caused by a drop in the position of NASDAQ in the routing tables of liquidity providers' (who were discouraged by small maker rebates). Similar chase for taker rebates by market orders is documented by Aldridge (2013). Lin *et al.* (2016) interestingly also find that market share decreased by a much smaller coefficient than the depth share and they deduce that liquidity demanders were more prone to trade on the NASDAQ compared to the passive liquidity side. This result was further strengthened by the findings that fill rate and speed of fill both improved at the selected sample of NASDAQ stocks. The authors also do not find strong support of the perfect market hypothesis - although they find a decrease in the effective spread (i.e. two times signed difference between trading price and midpoint), they also notice a decrease in the cum-fee effective spread. The explanation they provide is that the decline in the maker fee shifts the informed trading from NASDAQ to other, higher rebate-paying, exchanges, and as a result the adverse selection costs for liquidity providers also lower. Therefore, liquidity providers may be more willing to forgo some part of the original revenue by quoting prices slightly lower than were prevalent before the reform. Results of Battalio *et al.* (2016) are in line with the ones of Lin *et al.* (2016) in the domain of fill rates and fill speed. By analysing proprietary data together with NYSE TAQ data, authors consistently find that fill rates are much higher for exchanges that pay smaller maker rebates (and extract smaller take fees) and that conditional on fill, the speed of fill is much slower for those exchanges.

The baseline prediction of the neutral effect of the make-take fees on the market quality relied on the assumption of frictionless markets. More complex view was considered by some authors by adding either the mechanisms of fee passing (e.g. brokers) or inter-market competition. The theoretical models surpassing the perfect market circumstances predict that higher maker rebate and higher taker fee (keeping net fee constant) will decrease market participation but increase the share of market orders and therefore the fill rate. However, empirical studies of the US markets were not able to corroborate the latter hypothesis. Currently, besides a drop in the market share, the most evidenced effect of the higher maker rebate (and simultaneous higher taker fee) is a decrease in the fill rate and a slower speed of fill.

2.2 Modelling of the Limit Order Book

Our main goal consists in capturing potential statistical differences in patterns of dependence between market and limit orders within different exchange fee structures and then simulating such markets. Therefore, models that allow interdependence between arrival rates of individual order types would be the most befitting ones for our purposes. Also, we need to obtain models that allow estimation of its parameters on high-frequency market data. Several recent studies (Bowsher, 2007, Toke, 2011, Muni Toke & Pomponio, 2011, Gould *et al.*, 2013, Bacry *et al.*, 2013 or Lallouache & Challet, 2016) have evidenced that Hawkes process, a special type of self-exciting point process, can be used to model trade and limit order events and provides reasonable, and in some cases remarkable, fits to empirical datasets. It easily lends itself to multivariate extensions, it allows closed form maximum likelihood formula, also several non-parametric estimation approaches have been derived, goodness-of-fit tests are available and there are multiple methods for its simulation. In contrast to other zero-intelligence models, it does not suppose independent and identically distributed inter-arrival times of individual orders and it also might be easily adjusted to a real-time grid.

The first formulation of the model is due to Hawkes (1971), who besides proposing the general univariate and multivariate versions of the process, also derived its Fourier spectra for the special case of exponentially decaying kernel. Few years after the formulation saw contributions to the explicit formulation of maximum likelihood estimator (Ozaki, 1979), proof of consistency, asymptotic normality and efficiency of the MLE in a univariate stationary case (Ogata, 1978), and simulation techniques (e.g. Ogata, 1981). Since then, Hawkes process has been applied to a wide range of scientific areas, Lallouache & Challet (2016) mention earthquake occurrences, neuroscience, criminology and social networks modelling.

Bowsher (2007) was the first one to extensively appropriate the model for realities of financial econometrics. He started by introducing a generalized version of the Hawkes process (called by him g-HawkesE(k) in case of k intensity processes) that appended a dependence of the intensity function on the intensity from the end of the previous trading day (due to the closing of equity markets during the night) and derived the likelihood function for estimating parameters of such multivariate models. However, main contribution of the study consists in applying the random time change argument to the multivariate case in order to derive appropriate specification tests. He proposed two such tests, called "o-tests" and "m-tests", that test the correct specification jointly and separately for each intensity function, respectively. He then went on to analyse General Motors Corporation equity data from 2000 with respect to mid-quote changes and trades by the g-HawkesE(2) model. He found strong evidence of cross- and self-excitation in both directions. In the case of cross-excitation, the effects were of very high magnitude, but low half-lives and vice versa for the self-excitation.

Since then, several studies have focused on modelling limit order book by incorporating Hawkes process. Toke (2011) tried to model the dependencies in market making. After empirically observing shorter time spans between arrivals of market orders and subsequent arrivals of limit orders, he decided to capture this phenomena by modelling it with a bivariate Hawkes process with both self-excitation effects and one cross-excitation effect in the direction from market orders to limit orders. In comparison to simplified models, this bivariate Hawkes presented a reasonable fit and after simulation, exhibited distributions of time intervals between market and limit orders resembling the ones seen in data. Muni Toke & Pomponio (2011) concentrated on possible interdependence between trades-through (single market orders that trade at least the size available at the best ask/bid) at different market sides. They found that when comparing fits to BNP stock data of model with cross- and self-excitation terms and models with just self-excitation terms, the two were almost the same in their likelihoods. The authors concluded that cross-excitation is unimportant in modelling trades-through.

Bacry & Muzy (2014) proposed a comprehensive treatment of trade and price dynamics. They suggested a model with 4 underlying Hawkes processes. The first two were connected with trades at bid and ask and the second two were connected with prices changes (i.e. one for upward and one for downward price moves). The authors established auto-covariance function and asymptotic diffusive properties of the model. As an addition, they proposed another pair of Hawkes processes that were connected with market orders (bid and ask) of a single agent and derived the market impact profile of an order of this agent. Furthermore, they derived a new non-parametric estimator of the kernel function and used it to estimate the parameters of the model on the futures contracts of EuroStoxx and EuroBund. They found evidence for strong self-excitation in trades and cross-excitation (i.e. mean-reversion) in price dynamics.

Similar model was used by Bacry *et al.* (2013) in an analysis of two inconsistencies observed by current market microstructure research between theoretical diffusion models of prices and stylized facts of a behaviour of variance and covariance between assets at small time scales. The first one, belonging to the microstructure noise effects (graphically represented by a signature plot), states that with increasing granularity of data the empirical measures of daily variance increases without bound, while the second one, so-called Epps effect, describes the empirical observation that the higher granularity, the lower is the correlation of prices of two assets. Both observations are inconsistent with properties of diffusion models. The aim of authors was to capture these two effects while still keeping the Brownian characteristics of price dynamics on higher granularity scales. They started with modelling univariate and bivariate price dynamics as a difference of two counting processes (one for positive and one for negative price moves) and they established that, in a limit, the processes converge to a univariate or bivariate Brownian motions. The authors estimated the models on Euro-Bund and Euro-Boble data and showed that, indeed, both models capture the empirical realities of the signature plot and the Epps effect quite well.

One of the major threads in the literature on Hawkes process is investigating properties and appropriateness of application of different kernels. Rambaldi *et al.* (2014) try to model the activity in FX markets around the announcement of major macroeconomics news. They estimate Hawkes process models with double exponential and approximation of power-law kernel on currency quote data. They consistently find better fits with power-law kernel. As a next step, they add to the intensity equation another (exponential) kernel that is connected with a Poisson process of news. These second models are evidenced to reproduce changes of quotes around the announcement of major news well, with the power-law kernel being qualitatively more precise in capturing trading activity shortly after the announcement. Different conclusions are drawn by Lallouache & Challet (2016), who fit EUR/USD trade arrival data on different kernels of Hawkes process. They estimate the models with exponential, power-law, power-law with a cut-off, and a mixed kernel. In the case of decomposition of data into hourly cycles as well as in the case of daily decomposition, exponential kernel was the most appropriate kernel. However, a multiplicity of other studies (Hardiman *et al.*, 2013, Hardiman & Bouchaud, 2014, Bacry *et al.*, 2012, Bacry & Muzy, 2014 or Bacry *et al.*, 2015) have corroborated the notion that power-law kernels are more appropriate in modelling limit order book. It is surprising that this statement is evidenced to hold for different types of events (e.g. trade arrivals and quote changes).

A marked extension of Hawkes process was treated by Embrechts *et al.* (2011). They proposed two versions of marked Hawkes process - the first one in a multivariate setting and the second one in a vector-valued setting. The first one is defined by univariate marks associated with several counting processes and intensity functions, while in the second case, the marks are multivariate and there is only one underlying counting process and intensity function. The authors then test their models on empirical data - the first model is fit on extreme values of closing prices of Dow Jones Industrial Index (DJIA) for the period of 1994-2010 (i.e. values below 10% and above 90% quantiles) and the second model is fit on extreme hourly log returns (i.e. below 1% and above 99% quantile) of DJIA, NASDAQ-100 and SP500 Composite (time range from 1997 to 2010). Both estimations provided reasonable fits.

Flexibility inherent in the Hawkes process was appropriated for very diverse range of problems in the area of financial econometrics besides modelling limit order book and price dynamics. Large (2007) studied the resiliency of the market, i.e. the ability of an order book to replenish quickly after a large market order. He used characterization of the intensity function as a conditional expectation of a jump happening in the infinitesimally small neighbourhood of an occurrence of some other event. The author concluded that "in over 60 per cent of cases, the order book does not replenish reliably after a large trade. However, if it does replenish, it does so with a fairly fast half life of around 20 s." Filimonov & Sornette (2012) exploited the characterization of the Hawkes process as a composition of mother events and descendants. The mother events were considered a representation of the price movements due to the exogenous arrival of news (i.e. changes in the fundamental value of the underlying asset), while descendants were perceived as the price movements stemming from the market itself. The branching ratio, i.e. the ratio of descendant events to the ratio of all events stemming from one mother event, can be explicitly formulated in case of several kernel specifications of Hawkes process and the authors relate this number to a quantitative measure of the endogenous character of the price movements. By estimating the model on E-mini S&P 500 futures data, authors have found that the level of endogeneity has increased from 30% in 1998 to 70% in 2007. Hardiman *et al.* (2013) reacted to the Filimonov & Sornette (2012) by showing that when one changes the underlying kernel of the Hawkes process (from exponential to power-law) and the time window of the estimation, the results obtained can be very different. The authors have found that the ratio of non-fundamental trading in markets is steady over time and close to critical (i.e. spectral radius of the kernel matrix is close to 1), however, the reactions of markets get swifter with time. Filimonov & Sornette (2015) answered by making an extensive overview of possible biases (e.g. presence of outliers, edge effects in estimation of power-law kernel or non-stationary character of data) that could be a source of differences between the two studies. Hardiman & Bouchaud (2014) reacted by introducing a non-parametric approximation to the branching ratio that requires only knowledge of the first two moments. This asymptotically unbiased estimator was applied to the same dataset as used in Filimonov & Sornette (2012). Authors provided evidence supporting their previous claims about stability and criticality of the endogeneity parameter. Contagion of financial crises between different markets was considered in a recent study by Aït-Sahalia *et al.* (2015). The authors suggested substituting traditional Poission counting process in the jump term of the Levy diffusion by the Hawkes process as a means to allow the dependency structure in the jumps. The resulting Hawkes jump-diffusion, as the authors call it, does not have stationary and independent increments, but allows for rich mutual excitation between different processes. The authors provided GMM estimation technique and presented evidence of self-excitation and asymmetric cross-excitation between markets (direction is stemming from US markets). A remarkable study was published by Linderman & Adams (2014), who estimated the Hawkes process model enhanced with latent distance random graph on all individual components of SP100. They obtained a measure of distance between individual companies that exhibited clustering patterns in certain sectors of economy (e.g. energy and financials). They have also found that certain companies (e.g. Apple and Exxon) have strong cross-excitation effects on the performance of large number of companies.

A comprehensive treatment of stylized facts and modelling practices of limit order books can be found in Abergel *et al.* (2016). Modelling approach based on Hawkes process dynamics is also represented and several important theoretical results (e.g. generator, ergodicity, large scale limit) and simulation techniques are synthesized. A concise recent summary of existing research on the usage of Hawkes process in financial econometrics was completed by Bacry *et al.* (2015). Besides introducing several theoretical results, the authors gather the most promising studies in the areas of market activity, market endogeneity, price creation, market impact, optimal execution strategies, order book modelling and other miscellaneous models exploiting the framework of Hawkes process.

2.3 Market Making and Reinforcement Learning

The theory of optimal market-making develops optimal strategies that market-maker should follow in order to maximize his returns, while simultaneously minimizing inventory risk from repricing and minimizing the risk of informed trading. The domain has practical applications and has attracted considerable attention in recent years. The study that invigorated the research into optimal market-making was Avellaneda & Stoikov (2008). The authors considered a simple diffusion mid-price process with a zero drift and a constant volatility and Poisson process-driven market order arrival rates grounded in recent econo-physics results. By specifying a Hamilton-Jacobi-Bellman (HJB) framework of the model, they obtained closed form approximations to optimal quotes.

Several authors continued in developing either the benchmark model and/or the solution of it. Guéant *et al.* (2013) modified the Avellaneda-Stoikov model by adding inventory constraints. Then they showed that a solution of the HJB control equations can be found as a solution to a system of ordinary differential equations. Thus they obtained optimal quotes and their asymptotic behaviour (asymptotic in terminal time). Solution, optimal quotes and asymptotic limits were also obtained for the case of a diffusion with non-zero drift as the underlying mid-price and for the case of market impact. Fodra & Labadie (2012) solved the HJB equations and obtained approximations to optimal quotes for the general cases of either linear or exponential utility functions with possible inventory-risk aversion. Quite recent addition to the literature came from Guéant (2017), who extended the Guéant *et al.* (2013) solution by allowing for a more general market order intensity function than the exponential one used in the original Avellaneda & Stoikov (2008).

However, although these results are already very general, they may still be applied only in settings where we do assume certain distributional characteristics of the mid-price dynamics and market order arrivals. This limitation could be avoided by using some model-free technique of optimal control. Reinforcement learning is a family of methods, where an agent "interacts with environment by adaptively choosing actions in order to achieve some long-term objectives" (Chan & Shelton, 2001). Until recently, more massive expansion of reinforcement learning was hindered by either tedious manual crafting of features to represent the range of the target function or by the necessity to discretize its domain. These limitations have been overcome by recent advances in deep reinforcement learning, a method appropriating neural networks in approximating target function that reinforcement learning agent uses for searching for the optimal behaviour. Deep reinforcement learning was shown to exhibit extraordinary learning skills, most notably in playing a large spectrum of video games (e.g. Mnih *et al.*, 2015) or an ancient table game Go (e.g. Silver *et al.*, 2017).

These advances seem to be very slowly encroaching into the domain of quantitative finance. Reinforcement learning tools have found their applications mostly in time-series predictions, with minor excursions into the topic of optimal execution of trades. Corazza & Bertoluzzo (2014) or recently Jiang *et al.* (2017) and Deng *et al.* (2017) may be consulted for the application of regular and deep reinforcement learning systems to the task of enhancing time-series prediction or portfolio management. Much closer to our task, Nevmyvaka *et al.* (2006) and Hendricks & Wilcox (2014) have studied the application of regular Q-learning techniques for optimizing execution costs in markets. Recently, Fernandez-Tapia (2015) developed an iterative algorithm based on the stochastic gradient descent for a version of market-making problem and proved its convergence to global minimum under certain conditions.

Closest to our endeavour may be considered Chan & Shelton (2001) who studied the performance of Monte Carlo, SARSA and actor-critic algorithms given a task of maximizing profits for a market-maker. In the first part of the study, the authors developed a simple model of monopolistic market-maker that is unaware of the fundamental price of the asset. He sets only a single price and waits for trades from informed and uniformed traders. His state variable is inventory imbalance and he decides his actions based on different levels of this imbalance. Authors proved theoretically optimal policies for different proportions of informed and uninformed traders. In practical application, Monte Carlo and SARSA approximated the theoretical optima very well. In the second part of the study, the market-maker was able to set bid and ask prices and therefore had to control both the spread and the direction of his quotes. Both SARSA and actor-critic tracked reasonably well the true price process, while actor-critic finishing with a much lower variance of spread distribution (spread being directly incorporated into rewards).

Chan & Shelton (2001) thus provided evidence that stochastic policies resulting from applying actor-critic algorithms may be well suited for optimal control of market-makers. However, the state space representation for the algorithms consists of 3 states in the first and 12 in the second case, while action space has cardinality 3 in the first and cardinality 9 in the second case. The simplicity of the model and inability to scale it simultaneously with the algorithms therefore hinders the generalizability of its findings. The potential of deep reinforcement learning in searching for optimal policies in large products of state and action spaces can be, in my opinion, applied with the aim of providing solutions to optimal quoting problem in a much more realistic market settings.

Chapter 3

Limit Order Book Model

3.1 Introduction to the Hawkes process

As suggested in the previous chapter, we will be using a multivariate Hawkes process in modelling the arrival of market and limit orders. This chapter will, after reviewing some technical preliminaries, introduce the Hawkes process in a multivariate setting. We will also present log-likelihood specification and statistical tests. Then, we will propose a particular version of a multivariate Hawkes process, that will be used in an empirical estimation, and finally, we will provide an evidence for reliability of the MLE estimation.

Definition 3.1 (Point process). Let $(\Omega, \mathscr{F}, P)$ be a probability space and $(E, \mathscr{E})$ a measurable space. A sequence of non-decreasing $\mathscr{F}$-measurable $\mathscr{E}$-valued random variables $\{T_k\}_{k=1,2,...}$ with $T_k : \Omega \to E$ is called a point process.

Point processes on $\mathbb{R}^+$ are generally considered, therefore the measurable space is a pair of $\mathbb{R}^+$ and an associated Borel $\sigma-$algebra. Besides this, simple point processes are most often considered, which means that increasing property of the sequence $\{T_k\}_{k=1,2,...}$ is required. Also, we will be interested only in processes that are non-explosive, i.e. $\lim_{k\to\infty} T_k = \infty$.

In the following sections, we will always consider processes defined on a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$ with filtration $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$ and if not stated otherwise, the processes will be adapted to the filtration $\mathbb{F}$.

Definition 3.2 (Counting process). A càdlàg process

$$N(t) = \sum_{k=1,2...} \mathbb{1}_{\{T_k \leq t\}} \tag{3.1}$$

is called a counting process associated with the point process $\{T_k\}_{k=1,2,...}$.

Definition 3.3 (Conditional intensity function). A conditional intensity function of a counting process $N(t)$ is a left-continuous function:

$$\lambda(t|\mathscr{F}_t) = \lim_{h \to 0^+} \frac{\mathbb{E}(N(t+h) - N(t)|\mathscr{F}_t)}{h}, \tag{3.2}$$

Definition 3.4 (Univariate Hawkes process). A simple non-explosive counting process $N(t)$ with intensity $\lambda(t|\mathscr{F}_t)$ satisfying

1. $N(0) = 0$.
2. $\lambda(t|\mathscr{F}_t)$ is a left-continuous process given by

$$\lambda(t|\mathscr{F}_t) = \mu + \int_{-\infty}^{t} g(t-u)dN(u), \tag{3.3}$$

   where the integral is a stochastic integral w.r.t. semimartingale and $g(v) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$.

3. $P(N(t+h) - N(t) = 1|\mathscr{F}_t) = \lambda(t|\mathscr{F}_t)h + o(h)$
4. $P(N(t+h) - N(t) \geq 2|\mathscr{F}_t) = o(h)$

where by $o(h)$ we denote a quantity that satisfies:

$$lim_{h \to 0^+} \frac{o(h)}{h} = 0 \tag{3.4}$$

is called a univariate Hawkes process.

Although the integral presented by Hawkes (1971) has lower bound at $-\infty$, generally the processes starting at 0 are considered and therefore the integration is done only on some subset of $\mathbb{R}_0^+$. Hawkes (1971) also derived the condition on the function $g(v)$ in order for the process to be stationary, and that is:

$$\int_0^{\infty} g(v)dv < 1. \tag{3.5}$$

A multivariate Hawkes process is defined similarly as the univariate one - the difference is that a system of point processes and an inter-dependence between their intensities is introduced. Also, the underlying probability space needs to be adjusted. We now have several filtrations, i.e. $\mathbb{F}^m = \{\mathscr{F}^m(t)\}_{t \geq 0}$. Furthermore, the conditional intensity function needs to be adapted to the filtration defined by $\mathbb{F}' = \{\mathscr{F}_t'\}_{t\geq 0}$, where $\mathscr{F}_t' = \bigcup_{m=1}^{M} \mathscr{F}_t^m$. In the following, we assume exactly that different point processes are adapted to individual filtrations $\mathbb{F}^m$ and that the conditional intensity is adapted to $\mathbb{F}'$.

Definition 3.5 (Multivariate Hawkes process). A vector $\mathbf{N(t)} = (N^1(t), ..., N^M(t))$ of simple non-explosive counting processes satisfying $\forall m \in \{1, ..., M\}$:

1. $N^m(0) = 0$
2. $\lambda^m(t|\mathscr{F}_t')$ is a left-continuous process given by

$$\lambda^m(t|\mathscr{F}_t') = \mu_m + \sum_{n=1}^{M} \int_{-\infty}^{t} g^{m,n}(t-u)dN^n(u), \tag{3.6}$$

   where $\forall(m,n) \in B$, where $B = \{(i,j) : i \in \{1,...,M\}, j \in \{1,...,M\}\}, g^{m,n}(v) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$.

3. $P(N^m(t+h) - N^m(t) = 1|\mathscr{F}_t) = \lambda^m(t|\mathscr{F}_t')h + o(h)$
4. $P(N^m(t+h) - N^m(t) \geq 2|\mathscr{F}_t) = o(h)$

is called a multivariate Hawkes process.

Brémaud & Massoulié (1996) proved that there exists a condition under which the multivariate Hawkes process allows uniqueness and stability, the latter in the form of exponentially fast convergence towards stationary distribution. And the condition specifies that a matrix $A(MxM)$ with entries $a_{m,n} = \int_0^\infty g^{m,n}(v)dv$ needs to have spectral radius strictly less than 1.

There are different types of functions used as kernels in the literature. One can usually find strictly decreasing functions that try to imitate the nature of clustering of events, where the impact of an event is decreasing with the time from the occurrence of it. Also, most regularly, positivity of the base intensity and kernel intensity is applied, which causes an absence of inhibitory impacts in the process.

Exponential kernel of the form $\alpha e^{-\beta t}$ is the most commonly used one, as it allows for an $O(n)$ reduction of the usual $O(n^2)$ complexity of the computation of the log-likelihood, compensator or simulation of the process through the use of the recursive formula (following Ozaki, 1979). α parameter represents the measure of an instantaneous impact of the event, while β parameter represents the speed of decay of the impact. Larger β signifies that the event doesn't affect the intensity of the process after a short period of time.

The condition for the stationarity of the exponential kernel can be found by using the above-mentioned formula. We obtain that the matrix A is composed of elements $a_{m,n} = \frac{\alpha_{m,n}}{\beta_{m,n}}$. Therefore, the spectral radius of this matrix needs to be strictly lower than one in order for the Hawkes process to be stationary.

Another popular kernel is a power-law kernel parametrized by multiple slightly different versions in the literature. However, the basic representation takes the form $\frac{\alpha}{(t+\gamma)^{\beta}}$, where $\beta > 1$ (Bacry *et al.*, 2012). This kernel, although documented in the literature to capture the realities of the limit order book events well (see Section 2.2), is prohibitively slow due to its $O(n^2)$ complexity. Therefore, several exponential approximations have been tested in the literature in order to ease the computational burden. Lallouache & Challet (2016) used this parametrization composed of power-law factors in sum of exponentials:

$$g(v|n, \epsilon, \tau_0) = \frac{n}{Z}\Big(\sum_{i=0}^{J-1} \xi_i^{-(1+\epsilon)} e^{-\frac{v}{\xi_i}}\Big), \tag{3.7}$$

where $\xi_i = \tau_0 j^i$ for $0 \leq i < J$. Parameter Z is chosen such that $\int_0^{\infty} g(v)dv = n$, therefore $Z = \sum_{i=0}^{J-1}(\tau_0 m^i)^{-\epsilon}$. Parameter j controls the precision of the approximation and J specifies the range of it (the authors chose $j = 5$ and $J = 15$). $1 + \epsilon$ term approximates the power-law decay with an identical exponent.

Stationarity condition for this kernel is easily obtained from the condition for Z. Therefore, in a multivariate case, the A matrix will be composed of $a_{m,n} = n_{m,n}$ and if this matrix has a spectral radius lower than 1, the resulting process is stationary.

Other possible parametrizations of the kernel (mostly with power-law dynamics) can be found in Filimonov & Sornette (2015), Hardiman *et al.* (2013), or Lallouache & Challet (2016).

However, because most of the kernels besides exponential have been scarcely applied outside the univariate setting, we will be using the exponential kernel as the base kernel for our estimations. In Section 3.5, we will slightly adjust the base intensity of our conditional intensity function in order to approximate the dynamics specific to the intra-day fluctuations of financial markets.

3.2 Estimation of the Hawkes process

In our study, we will be using the MLE method in order to pinpoint the parameters of our model. However, reader should be aware that there are several other possible estimation techniques available for the Hawkes process. One can consult Bacry *et al.* (2012) for a non-parametric estimation of symmetric kernels through covariance matrix, Bacry & Muzy (2014) for another non-parametric estimation method, Olson & Carley (2013) for an estimation using branching property of the Hawkes process, Rasmussen (2013) for a method using Bayesian inference, Da Fonseca & Zaatour (2014) and Aït-Sahalia *et al.* (2015) for a GMM method of estimation or Kirchner (2017) for one recent formulation of another non-parametric method. A more comprehensive overview of possible methods and their description can be found in Bacry *et al.* (2015). Our choice of MLE was dictated by the fact that a lot of the presented estimation methods were tested mainly for univariate cases and also that the MLE was used by several influential papers in the area, see for example Bowsher (2007), Filimonov & Sornette (2012), Lallouache & Challet (2016). However, we shall accentuate that to the best of our knowledge, there is no proof of consistency and asymptotic normality of the MLE in the multivariate setting (only in the univariate one given by Ogata, 1978). Also, we are not aware that there exists a comprehensive study evaluating empirical qualities of different estimators, therefore we will provide a small simulation study presenting the quality of the MLE-based optimization on an artificial dataset.

Now, we will introduce the log-likelihood function for the general case of point processes and then for a specific case of the Hawkes process with an exponential kernel.

Theorem 3.1. *Let N be a counting process on [0, T] for some finite positive T, let $\{t_i\}_{1\leq i\leq N(T)}$ be a realization of N on this interval and let $\lambda(t|\mathscr{F}_t)$ denote its conditional intensity. Then the likelihood function of N satisfies:*

$$L = \Big[\prod_{i=1}^{N(T)} \lambda(t_i|\mathscr{F}'_{t_i})\Big] exp\Big(-\int_0^T \lambda(u|\mathscr{F}'_u)du\Big) \tag{3.8}$$

Proof. For a proof see e.g. Laub *et al.* (2015). □

The likelihood of the multivariate point process is just a product of likelihoods of individual components. The final log-likelihood of an M-variate point process takes therefore the following form:

$$logL = \sum_{m=1}^{M}\Big[\sum_{i=1}^{N(T)} log(\lambda^m(t_i^m|\mathscr{F}'_{t_i})) - \int_0^T \lambda^m(u|\mathscr{F}'_u)du\Big] \tag{3.9}$$

In the case of the Hawkes process with an exponentially decaying kernel, the formula for partial log-likelihood of single dimension m looks like this (see for example Abergel *et al.*, 2016):

$$\begin{aligned} logL^m = & - \mu_m T - \sum_{n=1}^{M} \frac{\alpha_{m,n}}{\beta_{m,n}} \Big\{ \sum_{k:t_k^n<T} [1 - e^{(-\beta_{m,n}(T-t_k^n))}] \\ & + \sum_{k:t_k^n<T} log\Big[\mu_m + \sum_{n=1}^{M} \alpha_{m,n} R_{m,n}(k)\Big]\Big\}, \end{aligned} \tag{3.10}$$

where

$$R_{m,n}(k) = e^{-\beta_{m,n}(t_k^m - t_{k-1}^m)} R_{m,n}(k-1) + \sum_{i:t_{k-1}^m \leq t_i^n < t_k^m} e^{-\beta_{m,n}(t_k^m - t_i^n)} \tag{3.11}$$

and $R_{m,n}(0) = 0$. For identical dimension also $R_{m,m}(1) = 0$ and $R_{m,m}(k)$ for $k \geq 2$ simplifies to:

$$R_{m,m}(k) = e^{-\beta_{m,n}(t_k^m - t_{k-1}^m)}(1 + R_{m,m}(k-1)) \tag{3.12}$$

Because individual dimensions are not interdependent in their parameters, we can simply maximize the log-likelihood of individual dimensions separately. Methods of optimization will be described in Section 3.5.

3.3 Goodness-of-fit statistics

Several qualitative and quantitative tools have been proposed in the literature in order to test the quality of the estimation and assess the hypothesis that given stochastic process is indeed a Hawkes process. All of the tools are connected with the random time change theorem. We will present the multivariate version here:

Theorem 3.2. *Let us have a vector* $\mathbf{N}(\mathbf{t}) = (N^1(t), ..., N^M(t))$ *of counting processes with conditional intensity satisfying for each* $m \in \{1, ..., M\}$*:*

$$\int_0^\infty \lambda^m(t|\mathscr{F}_t')dt = \infty \qquad a.s. \tag{3.13}$$

and define the $\mathscr{F}_t'$*-stopping time* $\tau_m(t)$ *as the unique solution to:*

$$\int_0^{\tau_m(t)} \lambda^m(s|\mathscr{F}_s^{'})ds = t. \tag{3.14}$$

Then a vector of counting processes $\tilde{\mathbf{N}}(\mathbf{t}) = (\tilde{N}^1(t), ..., \tilde{N}^M(t))$ *defined by*

$$\tilde{N}^m(t) = N^m(\tau_m(t)) \tag{3.15}$$

is composed of independent Poisson processes with unit intensity and the inter-arrival times of these processes are given by:

$$\tilde{t}_{i+1}^m - \tilde{t}_i^m = \int_{t_i^m}^{t_{i+1}^m} \lambda^m(t|\mathscr{F}_t^{'})dt \tag{3.16}$$

Proof. For a proof see Bowsher (2007). □

The integral of the intensity function on the specified time domain of the counting process is generally called the compensator and the inter-arrival times of the transformed Poisson processes are called residuals. As a consequence of the theorem, one can test whether the distribution of the residuals $\{\tilde{t}_{i+1}^m - \tilde{t}_i^m\}_{i=1,...,N^m(T)}$ follows an exponential distribution with parameter 1 (due to the fact that inter-arrival times of a unit Poisson process are distributed this way).

Several test statistics and qualitative measures of similarity are available for a researcher. The one used in several studies (Lallouache & Challet, 2016 or Embrechts *et al.*, 2011) as a qualitative measure is to compare the Q-Q plot of the exponential distribution with unit rate and the residuals. A rigorous way to test the similarity, applied for example by Lallouache & Challet (2016), is to use the Kolmogorov-Smirnov test statistic. Lallouache & Challet (2016) also exploited the excess dispersion test defined by Engle & Russell (1998) for ACD models, that measures whether the empirical second moment of residuals resembles the expected theoretical variance. Under the null hypothesis, the variable defined as

$$S = \frac{\hat{\sigma}^2 - 1}{\sqrt{8}}, \tag{3.17}$$

where $\hat{\sigma}^2$ is an empirical variance of residuals, has a limiting normal distribution.

Several studies (see e.g. Laub *et al.*, 2015) have used an equivalent version of previous tests by using transformation to uniform distribution. By defining the integral

$$\Lambda^m(t_i) = \int_0^{t_i^m} \lambda^m(t|\mathscr{F}_t^{'})dt \tag{3.18}$$

it can be easily deduced that a sequence $\{\Lambda^m(t_i)/\Lambda^m(T)\}_{i=1,\ldots,N(T)}$ is distributed according to the uniform distribution. Kolmogorov-Smirnov test may be used to measure the relationship quantitatively.

The independence property of residuals can be checked by some autocorrelation-based test statistic, e.g. Ljung-Box or Box-Pierce tests. The inability to reject either one or both of the last two tests, however, does not necessarily imply the independence property. Also, visual plots of pairs (U_i, U_{i+1}), where $U_i = F(\tilde{t}_i^m - \tilde{t}_{i-1}^m) = 1 - e^{-(\tilde{t}_i^m - \tilde{t}_{i-1}^m)}$ (F being the cumulative distribution function), can be studied for any regular patterns which may signify deviation from the underlying model specification. Even tests based on convergence properties of Hawkes processes to Brownian motion have been constructed and applied, see e.g. Laub *et al.* (2015) or Bacry *et al.* (2015).

In our study, we will be using Kolmogorov-Smirnov test to measure the similarity between the exponential distribution with unit rate and the empirical distribution of the residuals. We will also present Q-Q plots between these two distributions. The excess dispersion test will be used in order to measure specifically the similarity between the variances of these two distributions. Independence property will be tested by the Ljung-Box test statistic.

For an exponential kernel, the integral of the conditional intensity that is used to obtain the residual process can be found for example in Abergel *et al.* (2016) and takes the following form:

$$\begin{aligned}\int_{t_{k-1}^m}^{t_k^m} \lambda^m(s|\mathscr{F}_s^{'})ds =&\mu_m(t_k^m - t_{k-1}^m)\\ &+ \sum_{n=1}^{M} \frac{\alpha_{m,n}}{\beta_{m,n}} \Bigg\{ \left[1 - e^{-\beta_{m,n}(t_k^m - t_{k-1}^m)}\right] R_{m,n}(k-1)\\ &+ \sum_{i:t_{k-1}^m \le t_i^n < t_k^m} \left[1 - e^{-\beta_{m,n}(t_k^m - t_i^n)}\right]\Bigg\},\end{aligned} \tag{3.19}$$

where $R_{m,n}(k)$ is an identical recursion formula as in the log-likelihood case.

3.4 Model

As mentioned in the Chapter 2, our goal is to measure quantitatively the character of reaction of trades on quotes coming from exchanges with different fee structure. Therefore, in order to make our model parsimonious, but at the same time appropriate for the task, we shall assume that relevant events in the limit order book can be modelled by three dependent point processes. The first one shall be a point process of arrival of NBBO quotes from taker-maker exchanges, the second one shall be a point process of arrival of NBBO quotes from maker-taker exchanges and the third point process shall be a point process of arrival of trades. The NBBO condition is included due to the fact that Order Protection Rule 611 of Reg NMS requires the exchanges to allow all trades to happen at the best protected quotes. Protected quotes represent the best quotes from individual trading centres. Therefore, market participants that post trade orders non-exempt from Reg NMS (e.g. intermarket sweep orders or ISOs) trade always at the NBBO, defined as the best protected quote from all trading centres. Including in our model only NBBO quotes means we are assuming that market participants decide on trading because of noticing the arrival of a best quote. However, several other adjustments will be needed due to data specifics that are commented on in Chapter 4.

Mathematically, the composition of our model will take the form:

$$
\begin{aligned}
\lambda^1(t|\mathscr{F}_t^{'}) &= \mu_1 + \sum_{i=1}^{3} \int_0^t \alpha_{1,i} e^{-\beta_{1,i}(t-s)} dN^i(s) \\
\lambda^2(t|\mathscr{F}_t^{'}) &= \mu_2 + \sum_{i=1}^{3} \int_0^t \alpha_{2,i} e^{-\beta_{2,i}(t-s)} dN^i(s) \\
\lambda^3(t|\mathscr{F}_t^{'}) &= \mu_3 + \sum_{i=1}^{3} \int_0^t \alpha_{3,i} e^{-\beta_{3,i}(t-s)} dN^i(s),
\end{aligned}
\tag{3.20}
$$

where the first index refers to the process of arrival of quotes from taker-maker exchanges, the second index refers to the process of arrival of quotes from maker-taker exchanges and the third index refers to the process of occurrences of trades.

In order to fit the intraday fluctuation in the arrival of quotes and trades more closely, we have also tried an alternative of the above model that does not have a constant base intensity. Even though some authors have used linear interpolation functions (e.g. Bowsher, 2007 or Lallouache & Challet, 2016), due to our desire to minimize the number of parameters, we have specified quadratic function of the form $\mu_1 + \mu_2(t - 0.5)^2$, where t represents the current time on the interval $[0, 1]$, as a base intensity (we will represent constant case as $\mu_2 = 0$). We believe quadratic pattern could potentially approximate the larger traffic present in markets during opening and closing minutes/hours.

Parameters that will be of utmost interest for us are $\alpha_{3,1}$, $\beta_{3,1}$, $\alpha_{3,2}$ and $\beta_{3,2}$. The first one represents the instantaneous reaction of trade activity on the arrival of a single quote from a taker-maker exchange, while the second one represents the time evolution of this reaction. The term $\alpha_{3,1}e^{-\beta_{3,1}t}$ will correspond to the intensity of the non-homogeneous Poisson process that was initialized as a result of the event of type 1 and its integral on our empirical time interval (with lower bound at the time of the occurrence of the event) can be interpreted as how much a single event of this type contributed to the overall intensity process of dimension three. The same interpretations apply to parameters that correspond to the reaction of trades on the arrival of quotes from maker-taker exchanges. As we have mentioned in Section 2.1, current results of the literature on the make-take fees establish that inverted fee regimes will induce higher fill rates and conditional on fill, higher fill speeds. To test the first conclusion, we should derive the probability of fill connected with the corresponding non-homogeneous Poisson process introduced as a consequence of an occurrence of an event of type 1 or 2. The theorem and the proof are given below:

Theorem 3.3. *The probability of an occurrence of at least one event for a non-homogeneous Poisson process N(t) with intensity function* $\alpha e^{-\beta t}$ *on a finite interval [0, T] is given by:*

$$P[N(T) \geq 1] = 1 - e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})} \tag{3.21}$$

Proof. A result on the distribution function of a non-homogeneous Poisson process with right-continuous intensity function $\lambda(t)$ and a property of being bounded away from zero states that (see e.g. Gallager, 2013):

$$P(N(t) = n) = \frac{\Lambda(t)^n e^{-\Lambda(t)}}{n!}, \tag{3.22}$$

where $\Lambda(t) = \int_0^t \lambda(s)ds$.

However, we have a function decaying on $\mathbb{R}^+$ and therefore it is not bounded away from zero on its domain. In order to use this theorem, we need to divide our intensity function $\lambda(t)$ domain into two intervals, [0, T') and [T', ∞) with T' being an arbitrary finite number $\in \mathbb{R}^+$. On the first interval, our intensity function will take the form $\alpha e^{-\beta t}$ and on the second interval it will be a constant c (e.g. $\alpha e^{-\beta T'}$ for the purpose of continuity):

$$\lambda(t) = \mathbb{1}(t < T')\alpha e^{-\beta t} + \mathbb{1}(t \geq T')\alpha e^{-\beta T'} \tag{3.23}$$

And with this adjustment, we can compute the probability of at least one event happening in a finite interval [0, T]:

$$\begin{aligned} P(N(T) \geq 1) =& 1 - P(N(T) = 0) = 1 - e^{-\Lambda(T)} \\ &= 1 - e^{\mathbb{1}(T<T')(-\frac{\alpha}{\beta}(1-e^{-\beta T})) + \mathbb{1}(T \geq T')(T-T')\alpha e^{-\beta T'}}, \end{aligned} \tag{3.24}$$

which simplifies to:

$$P(N(T) \geq 1) = 1 - e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}, \tag{3.25}$$

for the case of $T < T'$.

□

Therefore, we can use these probabilities to measure fill probabilities of different triggers in the trade intensity function.

The second empirical conclusion present in the literature, i.e. that conditional on fill, the fill speed is higher on taker-maker exchanges, can be also analytically tested. We present the case for an exponential kernel in the theorem below:

Theorem 3.4. *The expected time of a first event t_1 for a non-homogeneous Poisson process N(t) with intensity function $\alpha e^{-\beta t}$ for the cases when at least one event happens until some finite T is given by:*

$$E[t_1 | N(T) \geq 1] = \int_0^T t \frac{\alpha e^{-\frac{\alpha}{\beta}(1-e^{-\beta t})} e^{-\beta t}}{1 - e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}} dt \tag{3.26}$$

Proof. In order to use the theorem related to the distribution function of number of arrivals of a non-homogeneous Poisson process, we need to apply the same idea as in the proof of Theorem 3.3 of dividing our time domain into two parts. Therefore, our intensity function will again take the form:

$$\lambda(t) = \mathbb{1}(t < T')\alpha e^{-\beta t} + \mathbb{1}(t \geq T')\alpha e^{-\beta T'}, \tag{3.27}$$

for some finite $T' > T$.

We continue by deriving the conditional distribution function of a first event given an event $N(T) \geq 1$ for some $T < T'$:

$$\begin{aligned}
F_{t_1}(t) &= P[t_1 \leq t | N(T) \geq 1] \\
&= 1 - P[t_1 > t | N(T) \geq 1] \\
&= 1 - \frac{P[t_1 > t, N(T) \geq 1]}{P[N(T) \geq 1]} \\
&= 1 - \frac{P[N(t) = 0, N(T) - N(t) \geq 1]}{P[N(T) \geq 1]} \\
&= 1 - \frac{P[N(t) = 0](1 - P[N(T) - N(t) = 0])}{1 - P[N(T) = 0]} \\
&= 1 - \frac{e^{-\Lambda(t)}(1 - e^{-(\Lambda(T) - \Lambda(t))})}{1 - e^{-\Lambda(T)}} \\
&= \frac{1 - e^{-\Lambda(t)}}{1 - e^{-\Lambda(T)}}
\end{aligned} \tag{3.28}$$

by the conditional probability law, independence of increments of a non-homogeneous Poisson process and the formula for the distribution of arrivals of a non-homogeneous Poisson process given in the proof of Theorem 3.3, respectively.

Although this function has domain on [0, T], it can be elongated to $-\infty$ to ∞ by specifying that:

$$F_{t_1}(t) = \begin{cases} 0, & -\infty < t < 0 \\ \frac{1-e^{-\Lambda(t)}}{1-e^{-\Lambda(T)}}, & 0 \leq t < T \\ 1, & T \leq t < \infty \end{cases} \tag{3.29}$$

We can see that our function now is a distribution function (i.e. non-decreasing and right-continuous with $lim_{t \to -\infty} = 0$ and $lim_{t \to \infty} = 1$). It is continuously differentiable on its domain, except for the points 0 and T, where one-sided derivatives exist, but are not equal. The derivative of this function at all sets except these two points is:

$$F'_{t_1}(t) = \begin{cases} f^1_{t_1}(t) = 0, & -\infty < t < 0 \\ f^2_{t_1}(t) = \frac{\alpha e^{-\frac{\alpha}{\beta}(1-e^{-\beta t})} e^{-\beta t}}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}}, & 0 < t < T \\ f^3_{t_1}(t) = 0, & T < t < \infty \end{cases} \tag{3.30}$$

In order to find an expectation of t_1 given $N(T) \geq 1$, we will need to calculate it by using improper integrals (with y being in the interval $(0, T)$):

$$\begin{aligned} E[t_1|N(T) \geq 1] &= \int_{-\infty}^{\infty} tF'_{t_1}(t)dt \\ &= \lim_{x \to 0-} \int_{-\infty}^{x} tf^1_{t_1}(t)dt + \lim_{x \to 0+} \int_{x}^{y} tf^2_{t_1}(t)dt \\ &+ \lim_{x \to T-} \int_{y}^{x} tf^2_{t_1}dt + \lim_{x \to T+} \int_{x}^{\infty} tf^3_{t_1}dt \end{aligned} \tag{3.31}$$

The first and the fourth integral converge to zero, so they can be omitted. Therefore, by just inserting the formulas for density functions:

$$E[t_1|N(T) \geq 1] = \lim_{x \to 0+} \int_{x}^{y} t\frac{\alpha e^{-\frac{\alpha}{\beta}(1-e^{-\beta t})}e^{-\beta t}}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}}dt + \lim_{x \to T-} \int_{y}^{x} t\frac{\alpha e^{-\frac{\alpha}{\beta}(1-e^{-\beta t})}e^{-\beta t}}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}}dt \tag{3.32}$$

Because the terms inside the integral are smaller or equal to $\frac{\alpha t}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}}$ for $\forall t \in [0, T]$, we can use a limit comparison test. The integrals:

$$\begin{aligned} \lim_{x \to 0+} \int_{x}^{y} \frac{\alpha t}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}} &= \lim_{x \to 0+} \left[\frac{\alpha t^2}{2(1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})})}\right]_x^y \\ &= \frac{\alpha y^2}{2(1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})})} \end{aligned} \tag{3.33}$$

and

$$\begin{aligned} \lim_{x \to T-} \int_{y}^{x} \frac{\alpha t}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}} &= \lim_{x \to T-} \left[\frac{\alpha t^2}{2(1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})})}\right]_y^x \\ &= \frac{\alpha(T^2 - y^2)}{2(1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})})} \end{aligned} \tag{3.34}$$

both converge and therefore our original integrals converge. The result follows:

$$E[t_1|N(T) \geq 1] = \int_{0}^{T} t\frac{\alpha e^{-\frac{\alpha}{\beta}(1-e^{-\beta t})}e^{-\beta t}}{1-e^{-\frac{\alpha}{\beta}(1-e^{-\beta T})}}dt \tag{3.35}$$

□

3.5 Simulations and estimations of the Hawkes process

In this chapter, we will present simulations and estimations of a three- dimensional Hawkes process. The reason for this exercise is to provide some evidence of a reliability of the estimation procedures as there is no analytical proof of the consistency of the MLE in the multivariate setting. We would also like to provide some evidence that non-stationary character of the series introduced by the quadratic base intensity does not hinder the estimation quality. The routines for the simulation, calculation of residuals and log-likelihood were written in C++ for efficiency purposes, while the optimization is done in Python (the binding was achieved through Cython). The algorithm for simulation was adapted from Abergel *et al.* (2016). The code underlying the simulation, statistical testing, estimation, and plotting may be seen on the github page of the author.[1] The reason why we have chosen to implement the routines by ourselves is due to the fact, that at the time of the preparation of the thesis, there were no libraries providing reliable functionalities befitting our purpose. R library hawkes `https://cran.r-project.org/web/packages/hawkes/hawkes.pdf` is not correctly implemented, as it omits certain parts of the simulation and log-likelihood algorithms, and it also does not provide compensator calculation routines. `https://github.com/dunan/MultiVariatePointProcess` is a C++ library that provides, according to the author, simulation, estimation and residual analysis routines, however, the binding to Python in order for an MLE optimization would be unnecessarily complicated. A recent addition to the computational tools for Hawkes processes is tick library (Bacry *et al.*, 2017) for Python available at `https://github.com/X-DataInitiative/tick/tree/master/doc`. Even though the library provides several simulation and inference routines, it lacks support for statistical analysis. At the end, we have therefore decided to implement the necessary functions by ourselves.

We have simulated a three-dimensional Hawkes process with an exponential kernel and $T = 22,800$ (equal to the number of seconds from 9.35am to 3.55pm that will be our empirical time range) and parameters:

[1] `https://github.com/ragoragino/py-hawkes`

Table 3.1: Optimizations - statistics

		1	2	3
C	KS test	0.765	0.530	0.439
	ED test	0.381	0.342	0.343
	LB test	0.925	0.478	0.454
	- LogL	-13,259.32		
Q	KS test	0.637	0.642	0.543
	ED test	0.357	0.250	0.370
	LB test	0.342	0.684	0.438
	- LogL	-45,607.261		

Note: The p-values for Kolmogorov-Smirnov, Excess-Dispersion and Ljung-Box test and negative log-likelihood for constant (C) and quadratic (Q) base intensity of a simulated three-dimensional Hawkes process.

$$\mu_{\mathbf{1}} = \begin{pmatrix} 0.2 & 1.7 & 0.6 \end{pmatrix}$$

$$\mu_{\mathbf{2}} = \begin{pmatrix} 0.8 & 0.4 & 0.5 \end{pmatrix}$$

$$\alpha = \begin{pmatrix} 0.4 & 0.1 & 0.1 \\ 0.3 & 0.5 & 0.4 \\ 0.18 & 0.2 & 0.31 \end{pmatrix}$$

$$\beta = \begin{pmatrix} 1.8 & 3 & 3.1 \\ 1.2 & 1 & 1.3 \\ 4.2 & 2.15 & 2.5 \end{pmatrix}$$

The choice of parameters was random and was adjusted in order to fit approximately the length characteristics of the empirical data samples, where the second component has around 100,000 events, while the first one only around 10,000, and the third one around 20,000 events. We have generated 11,195 events in the first dimension, 100,431 events in the second dimension, and 26,971 events in the third dimension (for the case of quadratic base intensity specification, the lengths of the series are 22,465, 119,147, and 35,245, respectively). We provide the p-values for Kolmogorov-Smirnov, Excess Dispersion and Ljung-Box tests for individual components in Table 3.1.

We have used optimization routine Sequential Least Squares Programming (SLSQP) provided by the scipy module developed by Jones *et al.* (2001–). The reason why this routine was chosen was due to its allowance of bounds and constraints which are both present in the Hawkes process.[2] Initial values of the parameters were set to resemble results found in the literature: μ_1 and μ_2 values were drawn from an exponential distribution with rate 1 and α and β parameters were drawn from a uniform distribution on $(0, 15)$. In the case of SLSQP optimization, the bounds were set in range $(0, 5)$ for μ_1 and μ_2 and in range $(0, 15)$ for other parameters. The constraints were the stationarity conditions for the exponential kernel. The lower bound for the parameters is due to the positivity of parameters in a general Hawkes process. We have tried different upper bounds together with ranges of initial parameters for α and β as μ_1 and μ_2 parameters are expected to be in the given range (in case μ_1 or μ_2 will be estimated on the edge of the parameter space, we will know that we might need to change the bounds for it too). The bounds that were chosen as final were the ones, where optimization routine was not behaving chaotically, i.e. the negative log-likelihood in individual optimizations was stable. Therefore, different initial values will be set for the optimization of the empirical dataset, as the parameter space might be different and therefore the optimizations might behave differently.

The results for the SLSQP optimization for the case of constant and quadratic base intensity are presented in Table 3.2. The optimization results paint an optimistic picture. The median values for μ_1, μ_2 and α parameters have been very precisely captured with little uncertainty. β values have also been optimized pretty well, however, there is much more uncertainty compared to the first two parameters. Also, we may notice much worse optimizations for parameters that are related to the first dimension, e.g. $\beta_{2,1}, \beta_{3,1}$ and $\beta_{1,3}$. This may be connected to the fact that the first dimension represents the shortest leg of the whole process. Therefore, in optimizations on real data we might have to be cautious about interpretations of results connected with the first dimension.

[2]We have also tried other routines provided in the scipy module and referenced in the literature, e.g. bounded L-BFGS-B or Nelder-Mead, however, they provided reliable, but slightly worse results compared to the SLSQP and therefore, we will be omitting these results from our study.

Table 3.2: Optimization of a simulated Hawkes process

		1			2			3		
		C	Q	OR	C	Q	OR	C	Q	OR
μ_1		0.182	0.183	0.2	1.671	1.682	1.7	0.631	0.629	0.6
		(4e-05)	(1e-04)		(3e-04)	(3e-04)		(1e-04)	(2e-04)	
μ_2			0.783	0.8		0.444	0.4		0.534	0.5
			(1e-04)			(5e-04)			(2e-04)	
α	**1**	0.372	0.380	0.4	0.105	0.108	0.1	0.067	0.068	0.1
		(3e-05)	(6e-05)		(3e-05)	(6e-05)		(4e-05)	(1e-04)	
	2	0.336	0.310	0.3	0.519	0.505	0.5	0.386	0.422	0.4
		(9e-04)	(6e-04)		(5e-05)	(6e-05)		(2e-04)	(3e-04)	
	3	0.127	0.167	0.18	0.209	0.206	0.2	0.299	0.315	0.31
		(6e-04)	(5e-04)		(4e-05)	(6e-05)		(1e-04)	(1e-04)	
β	**1**	1.665	1.688	1.8	2.974	3.012	3	1.814	1.996	3.1
		(2e-04)	(4e-04)		(0.001)	(0.003)		(0.002)	(0.007)	
	2	1.606	1.578	1.2	1.03	1	1	1.111	1.253	1.3
		(0.007)	(0.005)		(2e-04)	(3e-04)		(0.001)	(0.001)	
	3	4.955	4.226	4.2	2.339	2.363	2.15	2.407	2.557	2.5
		(0.035)	(0.026)		(0.001)	(0.001)		(0.001)	(0.002)	

Note: The result of 100 optimizations of a three-dimensional Hawkes process with an exponential kernel and a constant (C) and quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of the optimization results, while in brackets are median absolute deviations of optimized parameters (with normal adjustment). The maximum number of iterations was set to 20,000, all iterations ended successfully. OR abbreviation means original values of the parameters. The original negative log-likelihood for constant base intensity was -13,259.319, while for quadratic base intensity it was -45,607.261. The median for the two cases was -13,276.001 and -45,619.954 respectively, and median absolute deviation was 0.000 and 0.001 respectively. Median-based measures were chosen for robustness reasons as less than 5% of optimizations ended heavily dispersed from the rest of the optimizations.

Chapter 4

Data Description and Statistics

The primary data source for our study consists of the whole order book data for a trading during a single day, the October 24th of 2016. The TAQ (Trades and Quotes) data[1] cover market activity in all equity issues traded on NYSE, NASDAQ and regional exchanges that is reported to one of two central Security Information Processors (SIP) either by participants of Consolidated Tape Association (CTA) or Unlisted Trading Privileges (UTP) plans. All stock exchanges currently participate in CTA and UTP plans. These facilities provide collection, processing and distribution of the market data, the former for Tape A (i.e. listed on NYSE) and Tape B securities (i.e. listed on NYSE Arca, NYSE MKT, or BATS BZX Exchange), and the latter for Tape C securities (i.e. listed on NASDAQ).[2] Therefore, we will be able to discriminate between exchanges with taker-maker and maker-taker fees. At the date of the origination of our data sample, there were three exchanges, BATS BYX Exchange, BATS EDGA Exchange, and NASDAQ OMX BX, that applied taker-maker fee schemes.

We will consider the Daily TAQ Quotes file and the Daily TAQ Trades file. These files are a complete overview of all quotes and trades reported to the SIP by UTP or CTA participants. The Daily TAQ Quotes for our selected stocks contain nanosecond timestamp, exchange symbol, bid price, bid size, offer price, offer size, and other fields. The Daily TAQ Trades data contain nanosecond timestamp, exchange symbol, trade volume, trade price, and other fields. The complete enumeration of all fields and their short descriptions can be found in the Daily TAQ Client Specification (NYSE, 2016). However, we will be using only the ones mentioned.

The number of unique stock tickers being represented in the dataset is

[1] Obtained from NYSE: `ftp://ftp.nyxdata.com/Historical%20Data%20Samples/`.

[2] As of date of origination of our data sample.

around 8400. We will be focusing our study on four securities - Netflix (NFLX), Nvidia (NVDA), T-Mobile US (TMUS), and Shire PLC (SHPG). Our choice was dictated firstly by the focus on Tape C securities, that are processed by UTP and therefore have timestamps in nanosecond precision in contrast to microsecond precision for securities on Tape A or Tape B. The second criterion was given by the size of the data for NBBO quotes in general and for NBBO quotes originating in taker-maker exchanges in particular. Big stock names (e.g. Apple, Microsoft...) do not have such a large traffic of quotes originating in taker-maker markets, therefore we had to abandon selection of these otherwise most highly traded and quoted stocks. Hence, our data selection may have already created a bias as there may be some underlying reasons why some stocks are more heavily traded on taker-maker exchanges than other stocks. Unfortunately, we have not been able to gather some possible explanations regarding this difference and we are not able to pinpoint the sign or the magnitude of the bias.

The trading on U.S. exchanges happens during two time slots - one is regular trading hours (9.30am - 4.00pm) and one is pre- and after-hours activity (4.00am - 9.30am, 4.00pm - 8.00pm). However, due to the small amount of trades in the latter trading hours and higher volatility during opening and closing minutes, we will be focusing on the market activity during the time slots of 9.35am - 3.55pm. Besides this modification, we will also clean the data from the FINRA Alternative Display Facility (ADF) and FINRA Trade Reporting Facility (TRF) trade observations, because they originate from trades with non-displayed liquidity. As was already mentioned, we will be focusing only on NBBO quotes because we want to measure the time it takes trades to respond to quotes coming from markets with different fee structure. We will also exclude large amount of trades that are not regular or odd-lot. This step targets almost exclusively trade-throughs, e.g. ISO orders, because they are effectively trades not obeying the Order Protection Rule 611 and we have no reliable method how to classify whether the side of the quote matching the ISO (both on a single exchange) is the NBBO side or not. Due to the fact that certain quotes definitely traded, this operation may cause misclassifying quotes (instead of matching them with ISOs, which they could have traded in reality, we match them with regular trades), and therefore we have to acknowledge here another potential source of bias. However, we were not able to find resources suggesting some alleviation techniques for this issue and so again, we may not be able to say what may be the sign or the magnitude of the bias.

Table 4.1: Length of data samples for selected stocks

	BID			ASK		
	1	**2**	**3**	**1**	**2**	**3**
NFLX	6,139	40,784	11,959	4,611	43,329	14,303
NVDA	1,442	68,596	9,345	3,752	63,863	10,353
SHPG	2,312	14,387	3,079	3,688	14,063	3,710
TMUS	3,565	47,762	16,664	3,377	47,288	15,080

Note: Length of data samples for selected stocks. For quote and trade classification we have used quote and tick rules applied in Lee-Ready algorithm, see e.g. Lee & Ready (1991).

The classification of quotes and trades was done firstly with Lee-Ready algorithm that combines a quote and a tick rule for trades (see Lee & Ready, 1991 or Odders-White, 2000) and the quotes that remained non-classified were classified according to the tick rule for quotes. The quote rule says that trades should be classified as buys when their transaction price is higher than the midprice and as sells when their transaction price is lower than the midprice. When the transaction price is equal to the midprice, tick rule is used, meaning that when price increases relative to the previous trade price, the trade is classified as buy and as sell when the price decreases relative to the previous trade price. The quotes are subsequently classified on opposite sides (i.e. when trade is buy, the quote is ask, and when trade is sell, the quote is bid). As a final step, all the non-classified quotes are classified according to the tick rule, i.e. when the quote moves better or less worse on one side than on the other side relative to the previous quote, then that quote is classified as NBBO on that side. However, in contrast to Lee and Ready, we are not delaying the trades by a 5-second rule, as there is some evidence that the limit order book data are currently best classified without the application of delays, see e.g. Carrion & Kolay (2016).

The precise enumeration of observations in our data sample is presented in Table 4.1.

Chapter 5

Estimation results

In this chapter, we provide optimization results for individual stocks, NFLX, NVDA, SHPG, and TMUS, for both constant and quadratic intensity specifications, and for both sides, i.e. bid and ask, of the order book. We then continue with results concerning the execution quality measures for all these branches that will provide comparison between taker-maker and maker-taker exchanges relating to probabilities of fill and fill speed conditional on fill.

5.1 NFLX

In the case of NFLX, optimizations turn out very well for the base intensity, which was restricted to the range (0,5), and α parameters. For both cases, ASK and BID, it seems that the quadratic specification is helpful in maker-taker dimension as there seems to be a significant intra-day variation. As we have already seen in the case of simulations, β parameters are optimized with very high MAD. Therefore, there is quite a lot of uncertainty regarding the precision of the parameters. The optimizations themselves, however, ended all successfully and the MAD of negative log-likelihood is around 2-3%. This might suggest that the optimizations have found a local minimum, but a one where changes in some parameters' values have little impact on the final log-likelihood. From the values of parameters, it seems that immediate reactions to individual events are very high, most pronounced in the case of self-excitations and instantaneous impact of dimension 2 on dimension 3 (i.e. maker-taker events on trades). The decay of these impacts is proportionally quicker than for other events, except $\beta_{3,2}$ which seems to be relatively small and in connection with corresponding α parameter suggests prolonged intensity impact. The reasoning that might explain high parameter values is that, indeed, there is a high level of clustering of events on extremely small time-scales (i.e. order of microseconds). In the quest to capture this high-frequency clustering, the optimizer had to bloat the values of parameters in order the make the effect palpable. We have tried experimenting with different time scales, i.e. setting base units smaller than 1 second, but the resulting MADs of negative log-likelihoods were much more unstable than in the case of seconds, suggesting no such local minima existed for these different time specifications.

The statistics of the resulting optimization are also not entirely satisfactory. There is no test that could be passed on any reasonable significance level, suggesting the specified three-dimensional Hawkes model does not capture the dynamics of limit order book events. Plots in the Appendix present qualitative measures of the reliability of the estimation. In A.3, we can see the intensity function of taker-maker exchanges reacting to maker-taker events and trade arrivals. We can see an extremely high instantaneous impact simultaneously with extremely quick decays. A.2 compares event arrivals for a specific time range for maker-taker quotes and simulated events of this dimension. As we can see, the clustering of events and long no-event periods are captured authentically by the simulated series. A.1 presents the compensator connected with the trade arrivals, and as we can see, the quantiles of the exponential distribution and the compensator series are not following the same patterns, suggesting again model misspecification.

Table 5.1: NFLX BID optimizations - statistics

		1	**2**	**3**
C	KS test	4e-171	3e-565	9e-168
	ED test	0	0	8e-281
	LB test	0	0	0
	- LogL		-144,792.770	
Q	KS test	1e-170	2e-226	3e-181
	ED test	0	0	6e-303
	LB test	0	0	0
	- LogL		-146,987.287	

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for NFLX BID optimizations.

Table 5.2: NFLX BID optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.127	0.118	0.724	0.460	0.146	0.087
		(0.037)	(0.004)	(0.072)	(0.023)	(0.09)	(0.006)
μ_2			0.019		0.534		0.133
			(0.003)		(0.108)		(0.004)
α	**1**	751.341	796.684	51.517	54.878	125.538	127.054
		(594.828)	(604.953)	(21.256)	(29.417)	(57.787)	(47.912)
	2	424.152	332.888	946.622	1047.196	526.075	255.886
		(477.223)	(387.424)	(298.747)	(201.286)	(745.706)	(343.070)
	3	574.875	556.515	1740.169	1875.266	771.595	627.644
		(199.356)	(216.804)	(696.925)	(748.242)	(642.393)	(438.126)
β	**1**	3410.931	3575.577	1311.832	1407.870	7277.022	7409.851
		(2767.260)	(3090.794)	(644.982)	(865.625)	(3415.513)	(2500.461)
	2	3823.980	2831.656	2113.472	2176.112	2098.333	661.428
		(5077.977)	(3740.925)	(834.053)	(400.328)	(3028.375)	(893.663)
	3	8415.279	8154.182	14574.892	15184.948	3125.495	2232.257
		(2287.157)	(3130.288)	(5096.440)	(6203.271)	(2766.617)	(1533.554)

Note: The result of 100 optimizations of NFLX NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -144,792.770, while for quadratic case it was -146,987.287. Median absolute deviation of the negative log-likelihood was 3,952.230 for the constant case and 3,284.859 for the quadratic case. Median-based measures were chosen for robustness reasons.

Table 5.3: NFLX ASK optimizations - statistics

		1	2	3
C	KS test	8e-134	2e-304	9e-200
	ED test	0	0	0
	LB test	0	0	0
	- LogL		-168,476.069	
Q	KS test	4e-164	1e-268	1e-189
	ED test	0	0	8e-270
	LB test	0	0	0
	- LogL		-168,972.729	

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for NFLX ASK optimizations.

Table 5.4: NFLX ASK optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.09	0.084	0.77	0.528	0.164	0.094
		(0.005)	(0.009)	(0.065)	(0.044)	(0.006)	(0.003)
μ_2			0.021		0.627		0.161
			(0.005)		(0.114)		(0.008)
α	**1**	1094.297	959.693	51.255	48.931	81.584	86.712
		(576.949)	(615.207)	(32.682)	(32.305)	(40.733)	(49.948)
	2	598.504	583.474	1017.302	938.850	477.604	698.564
		(486.311)	(402.912)	(379.597)	(366.489)	(663.520)	(988.679)
	3	631.983	623.585	2093.276	2210.183	1549.229	1666.252
		(282.584)	(332.985)	(1238.588)	(1131.207)	(1242.843)	(1359.351)
β	**1**	5781.958	5562.447	1590.908	1656.858	6939.068	8007.255
		(2615.400)	(3447.485)	(1073.048)	(1212.134)	(2908.594)	(3507.981)
	2	4851.720	5389.313	2337.259	2162.727	2148.206	2703.831
		(4871.304)	(4226.257)	(1122.064)	(887.958)	(3069.009)	(3886.358)
	3	8249.662	7634.843	14638.850	15989.005	5118.335	5308.573
		(3284.117)	(2756.740)	(8465.995)	(8345.470)	(4010.160)	(4283.654)

Note: The result of 100 optimizations of NFLX NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -168,476.069, while for quadratic case it was -168,972.729. Median absolute deviation of the negative log-likelihood was 4,195.388 for the constant case and 4,665.950 for the quadratic case. Median-based measures were chosen for robustness reasons.

5.2 NVDA

For NVDA, the picture resulting from optimizations is not very different from the one we have seen in the NFLX case. μ parameters are estimated with high precision, α with less acuity and β with even larger imprecision. The quadratic case also captures the intra-day variation, most pronounced for the maker-taker quote arrivals. All optimizations ended successfully, with MADs of negative log-likelihood being always less than 2% (for ASK case even less than 1%). Large values of parameters again suggest a picture of high clustering of events. It is interesting to note that β values are always higher than 5000, except for the case of $\beta_{2,2}$. This parameter being relatively low (and with also relatively low MAD) signifies longer intensity impact in self-excitation on maker-taker exchanges.

The statistics of resulting optimizations are again unsatisfactory. There is no test that could be passed on any reasonable significance level, suggesting the specified three-dimensional Hawkes model does not capture the dynamics of limit order book events. Plots in the Appendix present qualitative measures of the quality of the estimation. In A.6, we can see the intensity function of maker-taker exchanges reacting to taker-maker events and trade arrivals. We can see the extremely high instantaneous impact simultaneously with extremely quick decays. A.5 compares event arrivals for a specific time range for taker-maker quotes and simulated events of this dimension. Although there are not so many events in the first dimension, the scarcity is similar in both cases. A.4 presents the compensator connected with trade arrivals, and even though the quantiles of the exponential distribution and the compensator series are not following the same patterns, the difference is not as pronounced as in the previous case.

Table 5.5: NVDA BID optimizations - statistics

		1	2	3
C	KS test	0	0	7e-174
	ED test	0	0	4e-116
	LB test	0	0	0
	- LogL	-356,058.321		
Q	KS test	0	0	5e-183
	ED test	0	0	2e-84
	LB test	0	0	0
	- LogL	-356,474.881		

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box test (LB), and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for NVDA BID optimizations.

Table 5.6: NVDA BID optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.032	0.022	0.879	0.700	0.090	0.051
		(0.003)	(0.006)	(0.017)	(0.090)	(0.004)	(0.010)
μ_2			0.023		0.403		0.087
			(0.008)		(0.126)		(0.008)
α	1	1201.547	1069.253	44.990	55.255	37.005	31.191
		(838.504)	(634.171)	(28.083)	(34.232)	(43.819)	(42.426)
	2	1191.870	1351.010	2245.125	2309.880	2679.086	2539.946
		(826.685)	(983.862)	(350.707)	(522.194)	(1787.072)	(1912.287)
	3	924.988	861.995	530.109	512.400	2975.826	3039.172
		(1050.460)	(975.679)	(234.089)	(235.879)	(1032.312)	(1197.070)
β	1	7195.038	7100.036	7613.063	8545.347	7916.362	8104.938
		(2613.762)	(1998.815)	(3239.853)	(3372.544)	(2732.685)	(3698.781)
	2	6510.376	7182.339	3586.000	3589.249	6948.695	6379.076
		(3016.080)	(2824.229)	(660.730)	(758.051)	(2451.156)	(3103.201)
	3	7735.160	8072.822	9190.783	9008.058	7710.445	8049.520
		(2969.707)	(2637.903)	(2935.730)	(3254.137)	(2146.834)	(2339.546)

Note: The result of 100 optimizations of NVDA NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -356,058.321, while for quadratic case it was -356,474.880. Median absolute deviation of the negative log-likelihood was 1,821.280 for the constant case and 6,802.818 for the quadratic case. Median-based measures were chosen for robustness reasons.

Table 5.7: NVDA ASK optimizations - statistics

		1	2	3
C	KS test	1e-169	0	2e-222
	ED test	0	0	1e-112
	LB test	0	0	0
	- LogL	-330,330.420		
Q	KS test	3e-150	0	4e-228
	ED test	0	0	6e-101
	LB test	0	0	0
	- LogL	-331,008.656		

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for NVDA ASK optimizations.

Table 5.8: NVDA ASK optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.087	0.086	0.821	0.616	0.110	0.065
		(0.006)	(0.006)	(0.008)	(0.007)	(0.004)	(0.003)
μ_2			0.000		0.476		0.104
			(0.000)		(0.007)		(0.005)
α	1	1450.344	1481.721	84.553	86.847	97.318	98.501
		(862.864)	(1092.081)	(52.214)	(58.296)	(63.809)	(61.942)
	2	496.666	508.718	1935.101	1915.462	2571.334	2779.551
		(274.639)	(393.397)	(196.821)	(181.760)	(1818.482)	(1579.184)
	3	482.317	481.115	504.401	482.490	3481.432	3386.527
		(315.826)	(336.177)	(198.180)	(218.212)	(1029.100)	(908.016)
β	1	5631.521	5312.752	6705.603	7081.026	7886.873	8577.666
		(2984.348)	(3045.707)	(3844.083)	(4301.729)	(2782.543)	(3845.109)
	2	7365.430	7385.836	3069.564	3033.176	7366.814	7185.738
		(3010.964)	(2417.843)	(301.505)	(269.948)	(2837.650)	(3066.504)
	3	7654.972	7634.843	9253.305	8553.228	8416.105	8495.226
		(2597.741)	(2978.633)	(3463.401)	(3664.692)	(2190.255)	(1968.855)

Note: The result of 100 optimizations of NVDA NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -330,330.420, while for quadratic case it was -331,008.655. Median absolute deviation of the negative log-likelihood was 1,594.195 for the constant case and 2,071.103 for the quadratic case. Median-based measures were chosen for robustness reasons.

5.3 SHPG

SHPG estimation does not in any way modify our existing views on the optimization properties. μ and α parameters are again estimated with a relatively high precision, while β with very large MAD. The quadratic base intensity specification helps in capturing the intra-day variation mainly for maker-taker quote arrivals. All optimizations ended successfully, with MADs of negative log-likelihood being always less than 3%. Large values of parameters again induce a picture of high clustering of events. The self-excitation effects present for maker-taker quotes dimension resemble the ones in NVDA case. Time-scale transformations again did not bring any success.

The statistics of optimizations are also unsatisfactory, even though the p-values for trade dimension are much better than in previous cases. Plots in the Appendix present qualitative measures of the quality of the estimation. A.8 compares event arrivals for a specific time range for processes of actual and simulated trades. A.7 presents the compensator connected with maker-taker quotes, where we can notice large discrepancy between the two compared distributions.

Table 5.9: SHPG BID optimizations - statistics

		1	**2**	**3**
C	KS test	0	3e-106	7e-26
	ED test	0	0	3e-38
	LB test	0	0	4e-5
	- LogL	-25,131.338		
Q	KS test	0	3e-101	1e-23
	ED test	0	0	3e-30
	LB test	0	0	5e-5
	- LogL	-25,646.790		

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for SHPG BID optimizations.

Table 5.10: SHPG BID optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.038	0.024	0.305	0.171	0.036	0.019
		(0.000)	(0.001)	(0.003)	(0.002)	(0.001)	(0.000)
μ_2			0.033		0.313		0.039
			(0.001)		(0.002)		(0.001)
α	**1**	157.967	163.054	55.427	56.185	66.748	67.105
		(31.552)	(45.959)	(1.570)	(2.038)	(35.997)	(35.748)
	2	387.334	397.367	223.631	226.651	1291.761	1200.900
		(318.616)	(270.406)	(7.749)	(7.159)	(451.555)	(440.399)
	3	455.995	487.160	551.121	531.270	1409.728	1563.749
		(213.264)	(178.944)	(168.829)	(141.323)	(563.761)	(535.173)
β	**1**	929.831	972.041	772.896	785.741	8055.185	7550.542
		(208.499)	(315.710)	(29.188)	(52.003)	(3520.239)	(2344.106)
	2	5009.776	4977.158	485.917	493.653	5959.027	5991.876
		(5672.634)	(4832.433)	(9.781)	(8.266)	(2513.795)	(2462.493)
	3	8121.400	8251.354	6811.777	6854.793	4452.040	4877.877
		(2444.031)	(2629.275)	(2202.120)	(1849.484)	(1747.918)	(1740.895)

Note: The result of 100 optimizations of SHPG NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -25,131.338, while for quadratic case it was -25,646.790. Median absolute deviation of the negative log-likelihood was 280.093 for the constant case and 339.211 for the quadratic case. Median-based measures were chosen for robustness reasons.

Table 5.11: SHPG ASK optimizations - statistics

		1	2	3
C	KS test	1e-63	1e-139	4e-48
	ED test	0	0	3e-47
	LB test	0	0	0.001
	- LogL		-27,268.010	
Q	KS test	2e-64	4e-145	5e-48
	ED test	0	0	2e-34
	LB test	0	0	0.013
	- LogL		-27,935.626	

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for SHPG ASK optimizations.

Table 5.12: SHPG ASK optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.070 (0.000)	0.053 (0.000)	0.288 (0.007)	0.140 (0.006)	0.049 (0.001)	0.023 (0.000)
μ_2			0.040 (0.000)		0.349 (0.006)		0.060 (0.001)
α	**1**	182.470 (7.489)	182.187 (6.964)	79.691 (2.980)	80.179 (2.544)	246.933 (127.813)	241.929 (126.822)
	2	201.403 (215.409)	181.516 (225.660)	232.199 (18.068)	233.267 (14.645)	970.079 (413.988)	1012.885 (362.466)
	3	473.431 (152.410)	456.264 (210.864)	686.416 (1209.802)	711.260 (177.768)	1055.840 (205.753)	1057.017 (141.794)
β	**1**	776.383 (37.558)	773.886 (29.901)	1028.008 (33.993)	1034.720 (32.829)	6549.578 (2918.176)	6372.586 (3377.675)
	2	2364.438 (3499.712)	4047.608 (5844.225)	510.215 (57.543)	516.229 (50.752)	5089.733 (2395.595)	5623.382 (2454.446)
	3	8152.654 (2222.397)	7853.278 (3050.286)	9204.777 (2612.497)	9512.860 (1801.659)	2911.401 (515.865)	2907.622 (356.855)

Note: The result of 100 optimizations of SHPG NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -27,268.010, while for quadratic case it was -27,935.626. Median absolute deviation of the negative log-likelihood was 483.921 for the constant case and 530.756 for the quadratic case. Median-based measures were chosen for robustness reasons.

5.4 TMUS

Results for TMUS closely follow the pattern for NVDA. μ parameters are again optimized very precisely, while α with less acuity and β with large imprecision. The quadratic parameter is high for maker-taker quote arrivals. All optimizations ended successfully, with MADs of negative log-likelihood being always around 1%. Strong clustering of events on small time-scales is captured by high values of parameters. $\beta_{2,2}$ is again significantly smaller than other β parameters, thus capturing strong self-excitation in maker-taker exchanges. We have also tried modifications of time-scales, however, without much success.

The statistics of the optimizations are again not providing strong reasons for a selection of our Hawkes model in capturing the dynamics of the empirical limit order data. Plots in the Appendix present qualitative measures of the quality of the estimation. A.10 compares event arrivals for a specific time range for actual and simulated trade process. A.9 presents compensator series connected with taker-maker arrivals that, as we can observe, follow a distribution with fatter tails than the exponential one.

Table 5.13: TMUS BID optimizations - statistics

		1	**2**	**3**
C	KS test	1e-157	0	4e-268
	ED test	0	0	6e-295
	LB test	0	0	5e-12
	- LogL	-263,973.723		
Q	KS test	1e-153	0	6e-243
	ED test	0	0	7e-313
	LB test	0	0	0
	- LogL	-264607.423		

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for TMUS BID optimizations.

Table 5.14: TMUS BID optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.06	0.003	0.630	0.440	0.197	0.102
		(0.011)	(0.005)	(0.008)	(0.007)	(0.006)	(0.002)
μ_2			0.053		0.442		0.221
			(0.013)		(0.011)		(0.007)
α	**1**	1158.486	1176.267	130.798	126.159	77.342	77.750
		(957.911)	(792.913)	(119.295)	(103.207)	(42.239)	(39.838)
	2	539.191	545.377	1521.633	1575.232	1333.262	1304.654
		(302.130)	(247.863)	(150.887)	(208.371)	(568.720)	(625.609)
	3	904.140	952.938	1165.082	1085.224	2279.795	2548.140
		(380.784)	(502.413)	(374.519)	(333.290)	(1120.274)	(939.963)
β	**1**	5208.971	5171.371	4622.947	4777.956	7853.645	7873.282
		(4220.648)	(4141.647)	(4671.873)	(4508.334)	(3161.090)	(2567.624)
	2	6658.078	7346.089	2439.232	2532.115	6763.670	7149.317
		(3831.235)	(2937.118)	(232.596)	(339.334)	(3056.668)	(3418.764)
	3	8134.119	8515.154	9396.709	8573.582	7054.835	7365.792
		(2483.843)	(2253.206)	(2967.611)	(2405.511)	(3220.594)	(2522.575)

Note: The result of 100 optimizations of TMUS NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -263,973.723, while for quadratic case it was -264,607.423. Median absolute deviation of the negative log-likelihood was 2,702.453 for the constant case and 1910.808 for the quadratic case. Median-based measures were chosen for robustness reasons.

Table 5.15: TMUS ASK optimizations - statistics

		1	2	3
C	KS test	2e-160	0	1e-168
	ED test	0	0	2e-178
	LB test	0	8e-09	0
	- LogL		-27,268.010	
Q	KS test	2e-167	0	1e-171
	ED test	0	0	3e-137
	LB test	0	4e-13	0
	- LogL		-27,935.626	

Note: The p-values for Kolmogorov-Smirnov (KS), Excess-Dispersion (ED), and Ljung-Box (LB) test, and negative log-likelihood for the constant (C) and the quadratic (Q) base intensity for TMUS ASK optimizations.

Table 5.16: TMUS ASK optimizations

		1		2		3	
		C	Q	C	Q	C	Q
μ_1		0.055	0.023	0.628	0.442	0.177	0.088
		(0.007)	(0.003)	(0.009)	(0.009)	(0.006)	(0.003)
μ_2			0.072		0.428		0.207
			(0.009)		(0.014)		(0.008)
α	**1**	1166.962	1135.326	141.161	139.679	82.408	79.367
		(636.833)	(695.371)	(97.288)	(118.219)	(54.193)	(45.914)
	2	576.833	598.445	1622.082	1639.169	1304.318	1231.081
		(438.352)	(323.283)	(182.743)	(240.675)	(833.125)	(815.574)
	3	924.015	880.488	1161.833	1096.971	2117.319	2093.742
		(372.082)	(424.814)	(323.222)	(379.964)	(904.941)	(929.258)
β	**1**	5868.378	6714.034	5254.403	4945.736	7745.235	6892.537
		(2963.878)	(3238.562)	(4095.375)	(4267.425)	(3202.260)	(3814.443)
	2	7087.393	7520.002	2586.105	2670.087	6630.284	5989.607
		(3081.338)	(2306.977)	(292.981)	(389.335)	(3858.576)	(4457.623)
	3	7835.784	7334.944	8825.339	8445.110	7144.622	6957.073
		(2526.519)	(2628.371)	(2737.982)	(2893.778)	(3133.367)	(2558.092)

Note: The result of 100 optimizations of TMUS NBBO series with an exponential kernel and the constant (C) and the quadratic (Q) base intensity with SLSQP optimization routine. The values are medians of optimization results, while in brackets are median absolute deviations of optimized parameters. The maximum number of iterations was set to 20,000, all iterations ended successfully. The median of negative log-likelihoods for the constant case was -253,323.147, while for quadratic case it was -254,308.422. Median absolute deviation of the negative log-likelihood was 2,654.213 for the constant case and 2,555.849 for the quadratic case. Median-based measures were chosen for robustness reasons.

5.5 Execution quality

As we could see, the statistical tests did not detect a presence of the Hawkes process specified by us in any stock at any reasonable significance level. Besides that, we could also observe overall statistical uncertainty connected with the estimation of β parameters. Therefore, we have to be very cautious about the relevance of our interpretations. Nonetheless, we believe, that the optimized Hawkes process might still be of some theoretical value. The main reason for this optimism of ours is that the simulated series obtained from the optimized parameters resemble the original series. Not only in several qualitative measures we have referenced in previous sections (mainly present for trade dimension), but also in their lengths. Table 5.17 presents a comparison of lengths of actual point process series and mean lengths of 100 simulated Hawkes processes with parameters set to medians of optimized parameters. For all presented stocks, the lengths of both series are almost identical. The resemblance is slightly better for the quadratic base intensity case, as this intensity might have captured some intra-day fluctuations (palpable mainly for NFLX and NVDA).

Finally, keeping these dilemmas in mind, we might be able to present the results concerning the execution quality statistics. Tables 5.18 and 5.19 show the probabilities of fill and fill speeds in case of fill connected to the taker-maker and maker-taker exchanges for all stocks and constant and quadratic base intensity. The picture that we have before provides evidence that maker-taker exchanges might be supplying higher quality execution services for their clients. Probabilities of fill are, except NVDA stock, always higher for maker-taker exchanges, sometimes even by a factor of 2 (NFLX). This holds regardless of the base intensity specification. The fill speed quality is even more unequivocal (shown in seconds with 6 decimal places, i.e. the last decimal place presents 1 microsecond). Besides BID side of SHPG, all stocks have a higher fill speed (meaning lower expected time to fill) in the case of maker-taker exchanges compared to taker-maker exchanges (the difference is in the order of tens of microseconds), for NFLX the difference is also by a factor of 2. The relationship holds for both intensity specifications.

We again stress that these probabilities of fill and fill speed are not representing the overall probabilities of fill and fill speed for trades' reactions, however, they represent an additional probability that is engendered when a new (taker-maker or maker-taker) quote arrives and fill speed connected with

Table 5.17: Comparison of length of actual and simulated processes

		NFLX	NVDA	SHPG	TMUS
BID	1	6,139	1,442	2,312	3,565
		5,744	1,388	2,330	3,705
		6,039	1,428	2,319	3,487
	2	40,784	68,596	14,387	47,762
		35,895	64,280	14,477	47,272
		40,844	67,527	14,391	46,689
	3	11,959	9,345	3,079	16,664
		10,625	9,632	3,098	15,891
		12,170	9,640	3,041	16,548
ASK	1	4,611	3,752	3,688	3,377
		4,327	3,876	3,695	3,357
		4,143	3,942	3,670	3,265
	2	43,329	63,863	14,063	47,288
		37,424	61,161	13,966	47,103
		39,376	62,540	13,656	45,676
	3	14,303	10,353	3,710	15,080
		13,492	10,420	3,737	15,128
		13,840	10,353	3,709	14,835

Note: Comparison of length of actual processes and 100 simulated processes with parameters set to medians of estimated parameters. Actual process length is always in the upper cell, while simulated process lengths for constant and quadratic base intensity are always in the bottom cells in this order.

Table 5.18: Execution quality statistics for the constant base intensity

			NFLX	NVDA	SHPG	TMUS
BID	PF	TM	0.066	0.113	0.055	0.105
		MT	0.113	0.056	0.078	0.117
	FS	TM	0.000117	0.000125	0.000121	0.000120
		MT	0.000067	0.000107	0.000144	0.000103
ASK	PF	TM	0.074	0.066	0.056	0.111
		MT	0.133	0.053	0.072	0.123
	FS	TM	0.000119	0.000139	0.000121	0.000124
		MT	0.000066	0.000107	0.000107	0.000110

Note: The probability of fill (PF) and fill speed in case of fill (FS) compared for bid and ask sides for all analysed stocks. TM represents taker-maker exchanges, while MT represents maker-taker exchanges.

this additional probability. Overall probabilities/fill speeds for trades will be higher/lower as they contain an additional base intensity (i.e. a simple Poisson process) and a sum of sensitivities to previous events, i.e. intensities of other non-homogeneous Poisson processes. However, these results are still of high informative values for our purposes, as additional probabilities and fill speeds provide a measure of execution quality gained from quoting on that particular type of exchange.

Table 5.19: Execution quality statistics for the quadratic base intensity

			NFLX	**NVDA**	**SHPG**	**TMUS**
BID	PF	TM	0.066	0.101	0.057	0.106
		MT	0.116	0.055	0.075	0.119
	FS	TM	0.000120	0.000121	0.000119	0.000114
		MT	0.000063	0.000109	0.000143	0.000113
ASK	PF	TM	0.078	0.061	0.056	0.113
		MT	0.129	0.055	0.075	0.122
	FS	TM	0.000128	0.000129	0.000125	0.000132
		MT	0.000060	0.000115	0.000103	0.000115

Note: The probability of fill (PF) and fill speed in case of fill (FS) compared for bid and ask sides for all analysed stocks. TM represents taker-maker exchanges, while MT represents maker-taker exchanges.

Our results are in direct contradiction with the results of Battalio *et al.* (2016) and Lin *et al.* (2016), who provide an evidence of qualitative difference in execution measures favouring taker-maker exchanges. However, documented higher fill rate for maker-taker exchanges is in accordance with the theoretical work of O'Donoghue (2015) and empirical results of Malinova & Park (2015) on TSX. However, even if our results would be correct, we would have to still understand, why there are stocks that are not in line with general results, e.g. NVDA for probability of fill and SPHG for fill speed. Besides that, in order for our results to provide a strong evidence, the Hawkes processes would need to be statistically significant. Therefore, replication on other datasets is warranted.

Chapter 6

Optimal Market-Making with Deep Reinforcement Learning

6.1 Market model

Current literature on the optimal market-making is almost uniform in its usage of the framework of Avellaneda & Stoikov (2008) for the characterization of the market mid-price evolution and the distribution of market order arrivals. Original model is appealing due to its simplicity and the fact that it appropriates results from econophysics of market micro-structure. It assumes a diffusive asset mid-price evolution $\{S_u\}_{u \geq t}$ with a zero drift and a constant volatility:

$$dS_u = \sigma dW_u, \tag{6.1}$$

where W_u is a standard Brownian motion and initial value $S_t = s$.

The probability of trading with market orders depends on the distance form the mid-price. The amount of shares bought/sold is formalized as a Poisson process N_t^b/N_t^a with an intensity λ^b/λ^a (b index denotes bid and a ask). The intensities are symmetric and are formalized in the following way:

$$\lambda(\delta) = Aexp(-k\delta) \tag{6.2}$$

where A and k are parameters and δ is the distance from the mid-price. The market-maker strives to maximize his objective function of the form:

$$u(s, x, q, t) = \max_{\delta^a, \delta^b} -E_t[exp(-\gamma(X_T + q_T S_T))], \tag{6.3}$$

where s denotes the asset price, x the money account of the market-maker,

q the inventory of the market-maker (all at initial time t of the trading), and γ is the discount factor. Avellaneda & Stoikov (2008) solved the problem by using Hamilton-Jacobi-Bellman equations and an asymptotic expansion in the inventory variable. They obtained a result that optimal quotes of the market-maker revolve around a reservation price

$$r(s,t) = s - q\gamma\sigma^2(T-t). \tag{6.4}$$

Market-maker computes this reservation price and then quotes around this price with a symmetric spread given by:

$$\delta^a + \delta^b = \gamma\sigma^2(T-t) + \frac{2}{\gamma}ln(1+\frac{\gamma}{k}) \tag{6.5}$$

The reservation price has a clearly stabilizing effect on market-makers' inventory, as in the case of positive inventory, it moves his ask quotes closer to the mid-price and his bid quotes far away from it, and vice versa. It can be also observed that several factors - negligible inventory (meaning minimal repricing risk), small discount factor (representing short-sighted agent), low volatility, or closeness to the terminal state - may make the quoting behaviour very similar to the symmetric quoting w.r.t. to the mid-price (i.e. always quoting a time average of Equation 6.5).

Avellaneda & Stoikov (2008) run simulations comparing this optimal, inventory, strategy with symmetric strategy. They found that generally, the optimal strategy has a lower mean, but it also has a lower variance.

We will appropriate this market framework. The market mid-price will evolve according to a diffusion with a constant volatility and a zero drift, market arrivals will follow Poisson processes based on the distance from the mid-price. The market-makers' final utility will be specified as an expectation of his terminal money account balance and the value of (stock) assets held at terminal time. We will only substitute for the exponential expectation function its linear counterpart (see also Fodra & Labadie, 2012):

$$u(s,x,q,t) = \max_{\delta^a,\delta^b} E_t[\gamma(X_T + q_T S_T))] \tag{6.6}$$

The motivation is that then the specification of rewards in the reinforcement learning setting (see the next section) will be more straightforward. However, in contrast to Avellaneda & Stoikov (2008) and other optimal market-making literature, we will use deep reinforcement learning techniques instead of analytic solutions in finding optimal quotes of the market-maker. Precisely, we will follow the framework of Double Deep Q-network (DDQN) used by Mnih *et al.* (2015) that trained reinforcement learning agents that were able to emulate or even surpass human performance in 29 of 46 different Atari games.

6.2 Reinforcement Learning

Reinforcement learning[1] is a machine learning domain where agents interact in an environment specified by a finite state space $\mathcal{S}$, finite action space $\mathcal{A}$, reward space $\mathcal{R} \subset \mathrm{R}$, and in case of Markov Decision Processes (MDP) also a probability transition mapping $P(s', r|s \in \mathcal{S}, a \in \mathcal{A}), \forall s \in \mathcal{S}, a \in \mathcal{A}, r$ in $\mathcal{R}$ and $s' \in S^+$, where $\mathcal{S}^+$ is $\mathcal{S}$ including the terminal state, and reward probability distribution $P(r|s \in \mathcal{S}, a \in \mathcal{A}), \forall s \in \mathcal{S}, a \in \mathcal{A}$ and r in $\mathcal{R}$. The agent performs actions by following certain policy $\pi(a|s)$, obtains rewards r and moves to new states s'. The problem may be either episodic with finite terminal states or infinite.

The return from a state (in a discrete and finite case) is defined as $R_t = \sum_{k=0}^{T} \gamma^k r_{t+k}$, where γ is a discount factor. The goal of the agent is to either evaluate states or state/action pairs given a policy or find a policy that maximizes the expected returns from each possible state or state/action pair. The value function $v_\pi(s) = E[R_t|s_t = s]$ measures exactly what is the value of the state given certain policy π, while the action-value function $q_\pi(s, a) = E[R_t|s_t = s, a_t = a]$ measures what is the value of the state and performing an action in that state given certain policy π. Both functions can be written in a form called Bellman equations:

$$v_\pi(s) = E_\pi^{a,s',r}[r + \gamma v_\pi(s')|s] \tag{6.7}$$

$$q_\pi(s, a) = E_\pi^{s',r}[r + \gamma E_\pi^{a'}[q_\pi(s', a')|s']|s, a] \tag{6.8}$$

Bellman optimality equations specify values of states or state/action pairs when the agent tries to find a policy that maximizes his/her expected returns from a state or a state/action pair:

[1] This brief introduction is based on Sutton & Barto (1998) and Li (2017).

$$v_*(s) = \max_{\pi} v_\pi(s) = \max_{\pi} \max_{a} E_{\pi}^{a,s',r}[r + \gamma v_*(s')|s] \tag{6.9}$$

$$q_*(s,a) = \max_{\pi} q_\pi(s,a) = \max_{\pi} E_{\pi}^{s',r}[r + \gamma \max_{a'} q_*(s',a')|s,a] \tag{6.10}$$

Plethora of algorithms may be used in order to iteratively evaluate some states or state/action pairs given a policy, or to iteratively find a maximizing policy. Dynamic programming may be used for problems where probability transition matrix is available. However, for most practical application this is not the case. Then, most often Monte Carlo methods or Temporal Difference learning is used. Monte Carlo methods, although proven to minimize the MSE on the training set, are not efficient, as they update the target function at the end of each episode, and therefore they are not even useful for non-episodic problems. Temporal Difference (TD) learning overcomes this hurdle and updates target functions by certain amount that is proportional to the error between the current target function and the estimate of it each single step (or with delays of few steps). SARSA, an on-policy method, or Q-Learning, an off-policy method, are the most common ones used for the purpose of policy evaluation or the search for maximizing policy.

The most commonly applied algorithms use some form of Bellman optimality equation. We will present only the relevant case of Q-learning that is a model-free, off-policy (i.e. it does not use the same function for updating the action-value as for the decision-making), and online algorithm that searches for optimal values of the action-value function. The pseudo-code for the Q-Learning is shown in Algorithm 1.

Algorithm 1 Q-Learning (off-policy TD control) for estimating $\pi \approx \pi_*$

```
Initialize Q(s,a), ∀s ∈ 𝒮, a ∈ 𝒜, arbitrarily, and Q(S_terminal, .) = 0
for episode = 1 to END do
    Initialize S_t = S_1
    while True do
        Choose A_t from S_t using the policy derived from Q(S_t, A_t) (e.g., ϵ-greedy)
        Take action A_t, observe R_{t+1}, S_{t+1}
        Q(S_t, A_t) ← Q(S_t, A_t) + α[R_{t+1} + γ max_a Q(S_{t+1}, a) − Q(S_t, A_t)]
        If S_t is terminal : break
```

However, one can notice that in order to find the optimal policy, one needs to find the values of states or state/action pairs for each combination of eligible states or state/action pairs. This can be done with either tabular methods, where the value of each combination is separately estimated. For large state and/or action spaces, this can pose an insurmountable hurdle. One then has to to turn to function approximation methods, where the state values or state/action values are formalized by a certain function with a parameter vector $\mathbf{w}$: $v(s; \mathbf{w})$ or $q(s, a; \mathbf{w})$. Semi-gradient learning methods, such as semi-gradient TD(0), can then be used for optimal policy control. For these methods, convergence guarantees are provided for a linear function approximation. Their disadvantage lies in the need of the feature specification.

This is the place where artificial neural networks come into play. They can act as a universal continuous function approximator and therefore may serve as an automatic engine for a feature selection. This property was successfully exploited by several research teams, most notably by the DeepMind researchers in several studies (e.g. Mnih *et al.*, 2013, Mnih *et al.*, 2015, Lillicrap *et al.*, 2015, Van Hasselt *et al.*, 2016, Silver *et al.*, 2016, Silver *et al.*, 2017). As was already noted, we will appropriate the Double Deep Q-Learning algorithm with experience replay that was concisely presented in Mnih *et al.* (2015).

In this study, the authors used deep convolutional neural networks for approximating the action-value Q-function. The loss measure of the learning process at step t was defined as:

$$L_t(\mathbf{w_t}) = E_{\pi}^{s,a}[(y_t - Q(s, a; \mathbf{w_t}))^2], \tag{6.11}$$

where $y_t = E_{\pi}^{s'}[r + \gamma \max_{a'} Q(s', a'; \mathbf{w_{t-1}})|s, a]$ is the target value at time t. As we can observe, one obtains the maximum with respect to the current Q-function and uses this value to update the parameters of this Q-function (similar as in the case of the classical Q-learning). Gradient of the loss function takes the form:

$$\nabla_{\mathbf{w_t}} L_t(\mathbf{w_t}) = E_{\pi}^{s,a,s'}\left[\left(r + \gamma \max_{a'} Q(s', a'; \mathbf{w_{t-1}}) - Q(s, a; \mathbf{w_t})\right) \nabla_{\mathbf{w_t}} Q(s, a; \mathbf{w_t}))\right]. \tag{6.12}$$

This loss function may be optimized by stochastic gradient descent methods (e.g. vanilla SGD, Adam, RMSProp, etc.). The algorithm becomes classical Q-learning when one replaces expectation by sample averages and update of weights occurs at each time step. Therefore the iteration update looks likes this:

$$\mathbf{w_{t+1}} = \mathbf{w_t} + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, a; \mathbf{w_t}) - Q(S_t, A_t, \mathbf{w_t})]\nabla \mathbf{w_t} Q(S_t, A_t, \mathbf{w_t}) \tag{6.13}$$

For a long time it was observed that non-linear function approximators very often diverge in reinforcement learning applications. In order to stabilize the learning, the authors appropriated an experience replay and a different target and update network - in the former case, the problem of correlated samples and in the latter case, the problem of biased learning and possible large swings in successive policies, was mitigated. Experience replay stores a buffer of samples $(S_t, A_t, R_{t+1}, S_{t+1})$ from previous steps and draws randomly from this buffer a selection of samples as an input to the deep Q-network. The second novelty, i.e. separating the target and update network, consists of creating a second convolutional deep neural network that is used for maximizing the target value. The update equation therefore takes the following form:

$$\mathbf{w_{t+1}} = \mathbf{w_t} + \alpha[R_{t+1} + \gamma \max_a \hat{Q}^t(S_{t+1}, a, \mathbf{w_t}) - \hat{Q}^u(S_t, A_t, \mathbf{w_t})]\nabla \mathbf{w_t} \hat{Q}^u(S_t, A_t, \mathbf{w_t}), \tag{6.14}$$

where $\hat{Q}^t$ is the estimated target network and $\hat{Q}^u$ is the estimated update network. In order for the target network to track the development of the update network, each C steps the target network copies the weights from the update network. Between these updates, the target network stays unchanged. The full pseudocode of the Double Deep Q-learning is presented in Algorithm 2 (we present the version without pre-processing, as our data input does not have a visual origin).

6.3 Optimal Quoting with Deep Reinforcement Learning

Our goal will be therefore to situate an agent that will try to optimize his expected terminal gains from the continuous quotation on the market in the market-making framework of Avellaneda & Stoikov (2008). By continuing the thread of the Hawkes order book model estimations, we will differentiate between two markets - maker-taker and taker-maker. These markets will have different probabilities of trading, specifically Equation 6.2 will incorporate an

Algorithm 2 Double Deep Q-learning with experience replay (without pre-processing)

Initialize replay memory D to capacity N, initialize weights for target and update network, Q^t and Q^u, respectively
for episode = 1 to END **do**
  Initialize sequence $S_t = S_1$
  while True **do**
    With probability ϵ select a random action a_t, otherwise select $a_t = \max_a Q^u(S_t, a; \theta)$
    Execute action A_t and observe reward R_{t+1} and new state S_{t+1}
    Store transition $(S_t, A_t, R_{t+1}, S_{t+1})$ in D
    Sample random minibatch of transitions $(S_j, A_j, R_{j+1}, S_{j+1})$ from D
    Set

$$y_j = \begin{cases} R_{j+1}, & \text{for terminal } S_{j+1} \\ R_{j+1} + \gamma \max_{a'} Q^t(S_{j+1}, a'; \theta), & \text{for non-terminal } S_{j+1} \end{cases} \tag{6.15}$$

    Perform a gradient descent step on $(y_j - Q^u(S_t, A_t; \theta))^2$ according to the equation 6.12
    If S_t is terminal : **break**

additional probability for trading on different exchange types. Although results from previous sections suggest that quoting on maker-taker exchanges is generally more profitable for market-makers, as they obtain liquidity rebates and simultaneously have higher probabilities of fill and speeds of fill, this section proposes a way to measure these trading gains. The optimization procedure itself will be done with a Double Deep Q-Network. One may ask, whether the results of the market-making framework of Avellaneda & Stoikov (2008) are not enough for computing the profitability gains/losses. However, the Avellaneda-Stoikov results are correct only under a specific assumption on the dynamics of the underlying stochastic process and are relevant only for a market-maker with a knowledge of the market parameters (σ and k). Reinforcement learning is a model-free tool for optimal control, hence no knowledge of the underlying process is required. Therefore, the deep Q-network applied in this section might converge to optimum even under very different market conditions. The generality of the optimum search is the advantage of this novel approach.

Our input data consist of the state of the market - current asset (stock) price, current inventory position of the market-maker, time distance from the last bid and ask trade and time distance to the terminal time. The choice of this state representation was guided by our experimentation. We specify possible actions on a discrete grid from -0.3 to 2.1 (by a step of 0.3), i.e. the action set has cardinality 81. Other parameters are overtaken from Avellaneda & Stoikov (2008), i.e. $S_t = 100$, $T = 1$, $\sigma = 2$, $dt = 0.005$, $k = 1.5$, and $A = 140$. Initial money account and inventory are both set to zero (short-selling is allowed).

As our input data are not of visual character, we will use only a feedforward neural network instead of a convolutional one. The neural networks will consist of 2 hidden layers (128 and 256 neurons, respectively) and an output layer comprising activations for all actions (similar to Mnih *et al.*, 2013 and Mnih *et al.*, 2015). In setting the parameters of the neural network, we have been guided by current research practices and our own experimentation. We will be using normal initialization for weights adjusted by the number of inputs to that particular layer and zero initialisation of biases. Leaky RELU will be considered as an activation function. We set the learning rate and regularizer to 0.001. Size of the experience replay buffer was set to 2,056, while the batch size is 64. Target network is updated every 100 steps. Exploration is linearly interpolated from 0 to 5,000 episodes, ending at the value of 0.001. Simulation length is 10,000 episodes (each episode being of length 200), and all outputs are taken for the last 2,500 episodes. Discount parameter was set to 0.95.

The setting of a fee value is quite subtle. Even though the price of the simulated asset is realistic, spread of the optimal quote from the solution of Avellaneda & Stoikov (2008) with linear utility function is around 0.66 (see next paragraph), which would imply an extremely illiquid market. In order to adjust the fee, we have set it to the same value as when comparing the fee rebates/charges to the usual spreads on the US markets. Maximum fee charge in the US markets specified by the Rule 610 of RegNMS is 0.003$ for one share transaction (Securities and Exchange Commission Division of Trading and Markets, 2015). When we take the median spread of the US stocks as 0.05$ (see Angel *et al.*, 2015), then applying the ratio to our case results in the fee rebate/charge of 0.04. The probability of fill differential between taker-maker and maker-taker exchanges will be set to an average value of our empirical estimation, i.e. 0.03.

Firstly, we will show that deep reinforcement learning agent is able to approximate the optimum quotation in the base scenario, i.e. without the fees/rebates and probability of fill adjustments. From Fodra & Labadie (2012) we know, that for a linear utility function and a martingale process, the optimal controls (δ^b_*, δ^a_*) around the mid-price are given by:

Table 6.1: Statistics of the performance of agents - base scenario

Agent	Mean of P&L	St. dev. of P&L	Mean of inventory	St. dev. of inventory
RL	68.27	13.52	-0.09	9.18
AS 0.1	65.05	6.50	0.03	3.06
AS 0.01	68.75	8.95	0.14	5.31
SQ	69.02	13.64	-0.12	8.60

Note: Selected statistics for the reinforcement learning (RL), our replication of Avellaneda-Stoikov (AS) with the discount parameter either 0.1 or 0.01 and symmetric quoting (SQ) of optimal spread ($1/k$) agents.

$$\delta_*^{a,b} = \frac{1}{k} \tag{6.16}$$

This means that for our case of $k = 1.5$, the optimal control of our agent should be either 0.6 or 0.9. Indeed, it turns out that the higher one is 0.6 and the agent learns this value (see Figure 6.3). Figure 6.2 presents final P&L for the RL agent and symmetric strategies around the mid-price with different spreads. We can see here that the RL agent is closely imitating the profit of the strategy with a constant 0.6 bid/ask spread. Even though we have different baseline utility specification than Avellaneda & Stoikov (2008), we can try to compare the two as decreasing the discount parameter in the original study moves its solution closer to the strategy of symmetric quoting (with the bid and ask spread given by the time average of Equation 6.5). Figure 6.1 and Table 6.1 show the resemblance between the final P&L of the strategy from the Avellaneda & Stoikov (2008) and between our RL agent. Our agent learns to maximize the mean, however, s/he is sacrificing higher variance of the P&L and final inventory. The cause of this discrepancy is due to the utility function specification as the exponential implicitly penalizes larger negative P&L results more than larger positive ones (for similar results see also Fodra & Labadie, 2012). Therefore, volatility of the P&L and inventory will be higher under the linear utility specification. We can also observe that the performance of our agent closely resembles the strategy of symmetric quoting of optimal quotation corroborating the quality of learning of the agent.

Next, we shall add an access market fee and a rebate of 0.04 for taker-maker and maker-taker exchanges, respectively, and a probability of fill decrease for taker-maker exchanges of 0.03. Table 6.2 summarizes the results for both exchanges for this case.

As was expected, the final P&L is higher when quoting on maker-taker exchanges. The standard deviations of both P&L and inventory are lower for

Figure 6.1: Final P&L histogram of the Reinforcement Learning agent (RL) and a solution of Avellaneda-Stoikov (AS) 0.1, i.e. with a discount parameter $\gamma = 0.1$

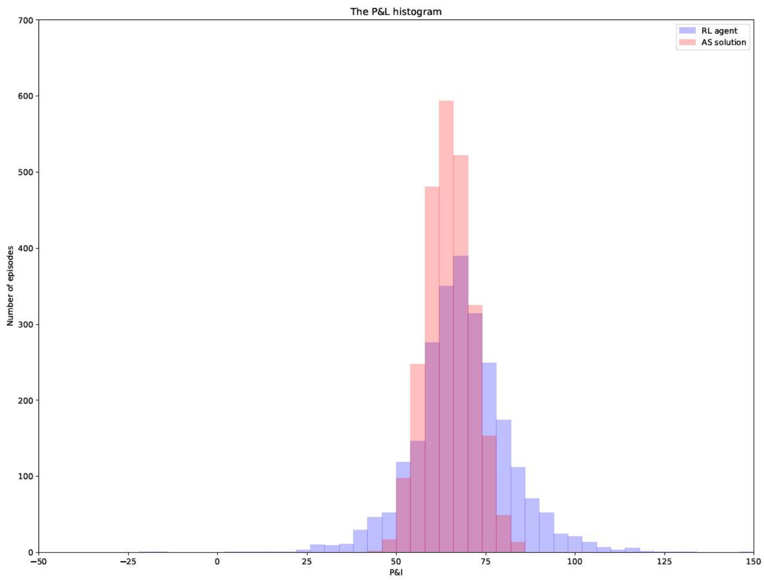


Figure 6.2: Final P&L histogram of the Reinforcement Learning agent (RL) and strategies of symmetric quoting around the mid-price.

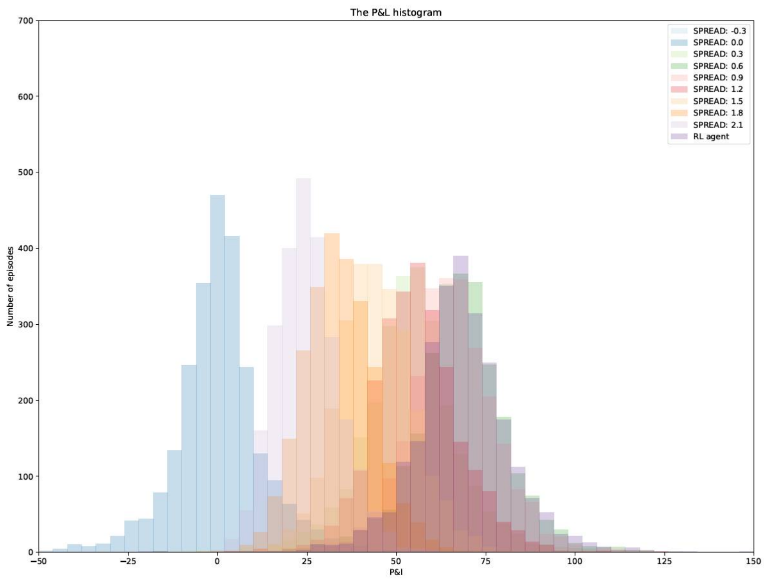


Table 6.2: Statistics of the performance of agents w.r.t. different exchange fees

Exchange type	Mean of P&L	St. dev. of P&L	Mean of inventory	St. dev. of inventory
Maker-taker	72.82	13.66	-0.09	9.18
Taker-maker	65.18	13.35	-0.09	8.87

Note: Selected statistics for reinforcement learning agents for maker-taker and taker-maker exchanges. Access fee was set to 0.04 and probability differential was chosen as 0.03.

Figure 6.3: A sample of the mid-price evolution with optimal bid and ask quotes from the RL agent

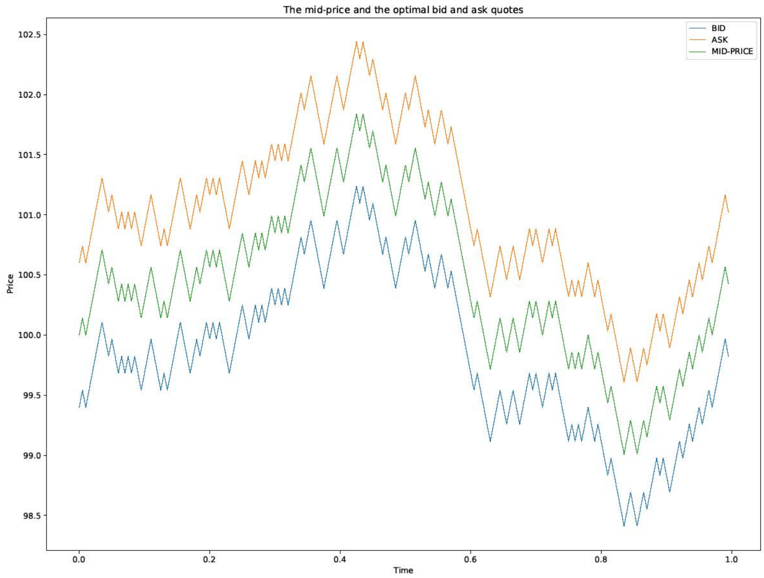


Table 6.3: Statistics of the performance of agents - base scenario with $\sigma = 4$

Agent	Mean of P&L	St. dev. of P&L	Mean of inventory	St. dev. of inventory
RL	68.34	25.49	-0.09	9.18
AS 0.1	52.75	6.72	-0.0056	2.20
AS 0.01	67.58	10.07	0.05	3.78
SQ	69.06	25.25	-0.12	8.60

Note: Selected statistics for the reinforcement learning (RL), our replication of Avellaneda-Stoikov (AS) with discount parameter either 0.1 or 0.01 and symmetric quoting (SQ) of optimal spread ($1/k$) agents.

taker-maker exchanges which is caused by a lower amount of trades due to the smaller probability of fill.

The final learning will occur under a heightened market stress, i.e. we have set $\sigma = 4$. Firstly, we will show the results for the base scenario without any probability and fee adjustments in order to see the behaviour of the RL agent in comparison with the analytical results.

The strategy of Avellaneda & Stoikov (2008) results in a significant decrease in the P&L, but the variation of it is almost identical to the case of the benchmark market volatility. On the other hand, our agent finishes with an almost identical P&L, but with a doubled standard deviation of it. Inventory measures remain the same for our agent, as the optimal quoting did not change and therefore also the number and composition of trades was stable during the learning.

Figure 6.4: Final P&L histogram of the Reinforcement Learning agent (RL) and a solution of Avellaneda-Stoikov (AS) 0.1, i.e. with a discount parameter $\gamma = 0.1$, under the market stress, meaning $\sigma = 4$.

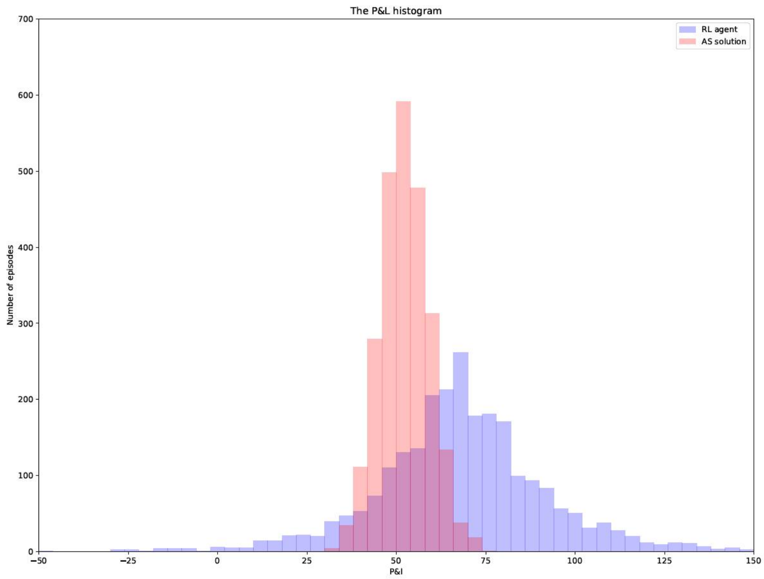


Figure 6.5: Final P&L histogram of the Reinforcement Learning agent (RL) and strategies of symmetric quoting around the mid-price with $\sigma = 4$.

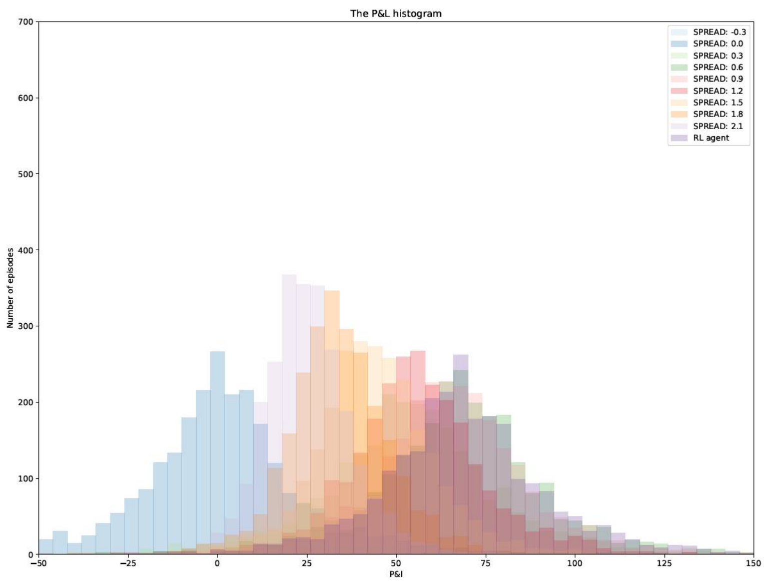


Figure 6.6: A sample of the mid-price evolution with optimal bid and ask quotes from the RL agent under $\sigma = 4$.

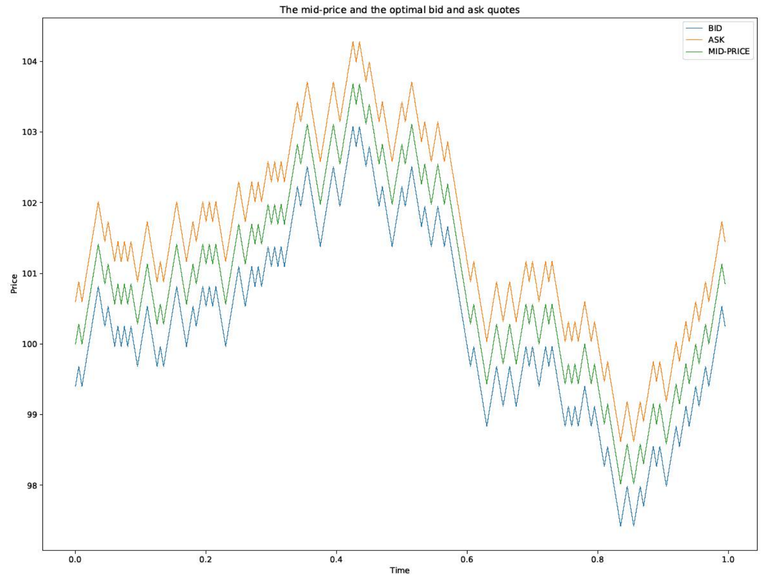


Table 6.4: Statistics of the performance of agents w.r.t. different exchange fees under the market stress

Exchange type	Mean of P&L	St. dev. of P&L	Mean of inventory	St. dev. of inventory
Maker-taker	72.89	25.56	-0.09	9.18
Taker-maker	65.32	24.92	-0.09	8.87

Note: Selected statistics for reinforcement learning agents for maker-taker and taker-maker exchanges under the conditions of doubled market volatility, i.e. $\sigma = 4$. Access fee was set to 0.04 and probability differential was chosen as 0.03.

Now we shall add the probability add-on and the fee rebate/charge to the market environment. The comparison between optimal values for both exchanges is presented in Table 6.4.

The situation is similar as in the previous case. Quoting on maker-taker exchanges results in a higher terminal profit, while the statistics for inventory remain the same as for the benchmark volatility case, because the optimal quoting does not change.

Finally, we can answer the hypotheses we have posed in the Proposal of our thesis. To the first question, the P&L of the market-maker is higher when quoting on maker-taker exchanges. To the second and third question, the spreads are the same for both volatility specifications for both fee regimes. The reason lies in the linear utility specification which does not penalize large inventory positions and its optimal quotes depend only on the parameters of the market order arrivals. Possible extensions with an inventory penalization or an exponential utility specification might improve on the latter results.

Chapter 7

Conclusion

The role of the (designated) market-makers is to provide continuous liquidity to the markets. Under normal circumstances, their quotes compete with those of a wide range of other actors (e.g. non-designated liquidity providers) and their absence on the top of the limit order book may not affect adversely the quality of the order flow. However, in more volatile times, their quotes may constitute only liquidity available. Therefore, from the perspective of the market-makers and other market actors, the knowledge of an impact of market micro-structure changes on the P&L of the market-maker is non-trivial.

In this thesis, we have tried to model the impact of one of the market micro-structure features, i.e. different fee regimes, on the P&L of the market-maker. We began with a specification of the model that captures probability of fill and fill speed for different types of fee structures. However, applying the model on selected order book data proved to be disappointing as no statistical importance was detected. Nonetheless, we believe that data inconsistency may have been the root cause of this discrepancy, and further replication may turn out to be more meaningful. We then cautiously interpreted the results of the optimizations of the model relating to the execution quality. It turns out that maker-taker exchanges may provide better liquidity to the markets.

These preliminary estimations were later on applied in the framework of optimal market-making. We have appropriated the model of Avellaneda & Stoikov (2008) for the purpose of setting up a general market framework. However, we have applied a novel deep reinforcement learning technique, Double Deep Q-Network, and obtained results that are comparable to the results of the original study. The advantage of deep reinforcement learning is that it is model-free. In contrast to the existing literature, we do not need to assume any knowledge of market parameters by market-making agents. The only requirement is a large dataset of historical or simulated time series that could be used for the training of the agent.

By using the Double Deep Q-Network in the context of make-take fee regimes, we have obtained an approximate enumeration of the optimal P&L and spreads. The fact that maker-taker exchanges provide better profit maximization opportunities for market-makers was already expected from the first part, as on these markets the liquidity providers obtain besides official fee rebates also better execution quality conditions. We have also seen that under the linear utility specification, the optimal spreads under both exchanges remain the same.

Although the first part of the thesis has not provided reliable answers, we believe that the results of the second part are of some importance. As it was already mentioned, this thesis, as far as the knowledge of the author goes, is the first attempt to apply deep reinforcement learning techniques to the domain of market-making. And the quality of the approximation of analytical results by our agent leaves us very optimistic. Future extensions are quite diverse - from specifying much more complicated underlying processes (e.g. Levy processes), to appending inventory constraints or using deep reinforcement learning techniques applicable to continuous action spaces (e.g. deep deterministic policy gradients).

Bibliography

Abergel, F., M. Anane, A. Chakraborti, A. Jedidi, & I. M. Toke (2016): *Limit order books.* Cambridge University Press.

Aït-Sahalia, Y., J. Cacho-Diaz, & R. J. Laeven (2015): "Modeling financial contagion using mutually exciting jump processes." *Journal of Financial Economics* **117(3)**: pp. 585–606.

Aldridge, I. (2013): *High-frequency trading: a practical guide to algorithmic strategies and trading systems*, volume 459. John Wiley and Sons, 2. edition.

Angel, J. J., L. E. Harris, & C. S. Spatt (2011): "Equity trading in the 21st century." *The Quarterly Journal of Finance* **1(01)**: pp. 1–53.

Angel, J. J., L. E. Harris, & C. S. Spatt (2015): "Equity trading in the 21st century: An update." *Quarterly Journal of Finance* **05(01)**: p. 1550002.

Avellaneda, M. & S. Stoikov (2008): "High-frequency trading in a limit order book." *Quantitative Finance* **8(3)**: pp. 217–224.

Bacry, E., M. Bompaire, S. Gaïffas, & S. Poulsen (2017): "tick: a Python library for statistical learning, with a particular emphasis on time-dependent modeling." *ArXiv e-prints* .

Bacry, E., K. Dayri, & J.-F. Muzy (2012): "Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data." *The European Physical Journal B-Condensed Matter and Complex Systems* **85(5)**: pp. 1–12.

Bacry, E., S. Delattre, M. Hoffmann, & J.-F. Muzy (2013): "Modelling microstructure noise with mutually exciting point processes." *Quantitative Finance* **13(1)**: pp. 65–77.

Bacry, E., I. Mastromatteo, & J.-F. Muzy (2015): "Hawkes processes in finance." *Market Microstructure and Liquidity* **1(01)**: p. 1550005.

Bacry, E. & J.-F. Muzy (2014): "Hawkes model for price and trades high-frequency dynamics." *Quantitative Finance* **14(7)**: pp. 1147–1166.

Battalio, R., S. A. Corwin, & R. Jennings (2016): "Can brokers have it all? on the relation between make-take fees and limit order execution quality." *The Journal of Finance* **71(5)**: pp. 2193–2238.

Bowsher, C. G. (2007): "Modelling security market events in continuous time: Intensity based, multivariate point process models." *Journal of Econometrics* **141(2)**: pp. 876–912.

Brémaud, P. & L. Massoulié (1996): "Stability of nonlinear hawkes processes." *The Annals of Probability* pp. 1563–1588.

Brolley, M. & K. Malinova (2013): "Informed trading and maker-taker fees in a low-latency limit order market." *Social Science Research Network* `https://ssrn.com/abstract=2178102`. [Online; accessed 08-May-2017].

Carrion, A. & M. Kolay (2016): "Trade signing in fast markets." *Social Science Research Network* `https://ssrn.com/abstract=2489868`. [Online; accessed 11-September-2017].

Chan, N. T. & C. Shelton (2001): "An electronic market-maker." .

Colliard, J.-E. & T. Foucault (2012): "Trading fees and efficiency in limit order market." *Review of Financial Studies* **25(11)**: pp. 3389–3421.

Corazza, M. & F. Bertoluzzo (2014): "Q-learning-based financial trading systems with applications." .

Da Fonseca, J. & R. Zaatour (2014): "Hawkes process: Fast calibration, application to trade clustering, and diffusive limit." *Journal of Futures Markets* **34(6)**: pp. 548–579.

Deng, Y., F. Bao, Y. Kong, Z. Ren, & Q. Dai (2017): "Deep direct reinforcement learning for financial signal representation and trading." *IEEE transactions on neural networks and learning systems* **28(3)**: pp. 653–664.

Embrechts, P., T. Liniger, & L. Lin (2011): "Multivariate hawkes processes: an application to financial data." *Journal of Applied Probability* **48(A)**: pp. 367–378.

Engle, R. F. & J. R. Russell (1998): "Autoregressive conditional duration: a new model for irregularly spaced transaction data." *Econometrica* pp. 1127–1162.

Fernandez-Tapia, J. (2015): "High-frequency trading meets reinforcement learning: Exploiting the iterative nature of trading algorithms." .

Filimonov, V. & D. Sornette (2012): "Quantifying reflexivity in financial markets: Toward a prediction of flash crashes." *Physical Review E* **85(5)**: p. 056108.

Filimonov, V. & D. Sornette (2015): "Apparent criticality and calibration issues in the hawkes self-excited point process model: application to high-frequency financial data." *Quantitative Finance* **15(8)**: pp. 1293–1314.

Fodra, P. & M. Labadie (2012): "High-frequency market-making with inventory constraints and directional bets." *arXiv preprint arXiv:1206.4810* .

Foucault, T., O. Kadan, & E. Kandel (2013): "Liquidity cycles and make/take fees in electronic markets." *The Journal of Finance* **68(1)**: pp. 299–341.

Gallager, R. G. (2013): *Stochastic processes: theory for applications.* Cambridge University Press.

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

Guéant, O. (2017): "Optimal market making." *Applied Mathematical Finance* **24(2)**: pp. 112–154.

Guéant, O., C.-A. Lehalle, & J. Fernandez-Tapia (2013): "Dealing with the inventory risk: a solution to the market making problem." *Mathematics and financial economics* pp. 1–31.

HARDIMAN, S., N. BERCOT, & J.-P. BOUCHAUD (2013): "Critical reflexivity in financial markets: a hawkes process analysis." *The European Physical Journal B* **86(10)**: pp. 1–9.

HARDIMAN, S. J. & J.-P. BOUCHAUD (2014): "Branching-ratio approximation for the self-exciting hawkes process." *Physical Review E* **90(6)**: p. 062807.

HAWKES, A. G. (1971): "Spectra of some self-exciting and mutually exciting point processes." *Biometrika* pp. 83–90.

HENDRICKS, D. & D. WILCOX (2014): "A reinforcement learning extension to the almgren-chriss framework for optimal trade execution." In "Computational Intelligence for Financial Engineering & Economics (CIFEr), 2104 IEEE Conference on," pp. 457–464. IEEE.

JIANG, Z., D. XU, & J. LIANG (2017): "A deep reinforcement learning framework for the financial portfolio management problem." *arXiv preprint arXiv:1706.10059* .

JONES, E., T. OLIPHANT, P. PETERSON *et al.* (2001–): "SciPy: Open source scientific tools for Python."

KIRCHNER, M. (2017): "An estimation procedure for the hawkes process." *Quantitative Finance* **17(4)**: pp. 571–595.

LALLOUACHE, M. & D. CHALLET (2016): "The limits of statistical significance of hawkes processes fitted to financial data." *Quantitative Finance* **16(1)**: pp. 1–11.

LARGE, J. (2007): "Measuring the resiliency of an electronic limit order book." *Journal of Financial Markets* **10(1)**: pp. 1–25.

LAUB, P. J., T. TAIMRE, & P. K. POLLETT (2015): "Hawkes processes." *arXiv preprint arXiv:1507.02822* .

LEE, C. & M. J. READY (1991): "Inferring trade direction from intraday data." *The Journal of Finance* **46(2)**: pp. 733–746.

LI, Y. (2017): "Deep reinforcement learning: An overview." *arXiv preprint arXiv:1701.07274* .

Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, & D. Wierstra (2015): "Continuous control with deep reinforcement learning." *arXiv preprint arXiv:1509.02971* .

Lin, Y., P. L. Swan *et al.* (2016): "Maker-taker fee, liquidity competition, and high frequency trading." *Social Science Research Network* `https://ssrn.com/abstract=2826376`. [Online; accessed 08-May-2017].

Linderman, S. & R. Adams (2014): "Discovering latent network structure in point process data." In "International Conference on Machine Learning," pp. 1413–1421.

Malinova, K. & A. Park (2015): "Subsidizing liquidity: The impact of make/take fees on market quality." *The Journal of Finance* **70(2)**: pp. 509–536.

Mnih, V., K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, & M. Riedmiller (2013): "Playing atari with deep reinforcement learning." *arXiv preprint arXiv:1312.5602* .

Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.* (2015): "Human-level control through deep reinforcement learning." *Nature* **518(7540)**: pp. 529–533.

Muni Toke, I. & F. Pomponio (2011): "Modelling trades-through in a limited order book using hawkes processes." *Social Science Research Network* `https://ssrn.com/abstract=1973856`. [Online; accessed 12-June-2017].

Nevmyvaka, Y., Y. Feng, & M. Kearns (2006): "Reinforcement learning for optimized trade execution." In "Proceedings of the 23rd international conference on Machine learning," pp. 673–680. ACM.

NYSE (2016): "Daily taq client specification - version 2.2." `http://www.nyxdata.com/Data-Products/Daily-TAQ`. [Online; accessed 22-June-2017].

Odders-White, E. R. (2000): "On the occurrence and consequences of inaccurate trade classification." *Journal of Financial Markets* **3(3)**: pp. 259–286.

O'Donoghue, S. M. (2015): "The effect of maker-taker fees on investor order choice and execution quality in us stock markets." *Social Science Research*

Network https://ssrn.com/abstract=2607302. [Online; accessed 08-May-2017].

Ogata, Y. (1978): "The asymptotic behaviour of maximum likelihood estimators for stationary point processes." *Annals of the Institute of Statistical Mathematics* **30(1)**: pp. 243–261.

Ogata, Y. (1981): "On lewis' simulation method for point processes." *IEEE Transactions on Information Theory* **27(1)**: pp. 23–31.

Olson, J. F. & K. M. Carley (2013): "Exact and approximate em estimation of mutually exciting hawkes processes." *Statistical Inference for Stochastic Processes* **16(1)**: pp. 63–80.

Ozaki, T. (1979): "Maximum likelihood estimation of hawkes' self-exciting point processes." *Annals of the Institute of Statistical Mathematics* **31(1)**: pp. 145–155.

Permanent Subcommittee on Investigations (2014): "Conflicts of interest, investor loss of confidence, and high speed trading in u.s. stock markets." https://www.hsgac.senate.gov/subcommittees/investigations/hearings/conflicts-of-interest-investor-loss-of-confidence-and-high-speed-trading-in-us-stock-markets. [Online; accessed 29-November-2017].

Rambaldi, M., P. Pennesi, & F. Lillo (2014): "Modeling fx market activity around macroeconomic news: a hawkes process approach. arxiv preprint." *arXiv preprint arXiv:1405.6047* **2(4)**.

Rasmussen, J. G. (2013): "Bayesian inference for hawkes processes." *Methodology and Computing in Applied Probability* **15(3)**: pp. 623–642.

Securities and Exchange Commission Division of Trading and Markets (2015): "Maker-taker fees on equities exchanges." "https://www.sec.gov/spotlight/emsac/memo-maker-taker-fees-on-equities-exchanges.pdf". [Online; accessed 08-May-2017].

Silver, D., A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.* (2016): "Mastering the game of go with deep neural networks and tree search." *Nature* **529(7587)**: pp. 484–489.

Silver, D., J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton *et al.* (2017): "Mastering the game of go without human knowledge." *Nature* **550(7676)**: pp. 354–359.

Sutton, R. S. & A. G. Barto (1998): *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge.

Toke, I. M. (2011): "Market making in an order book model and its impact on the spread." In "Econophysics of Order-driven Markets," pp. 49–64. Springer.

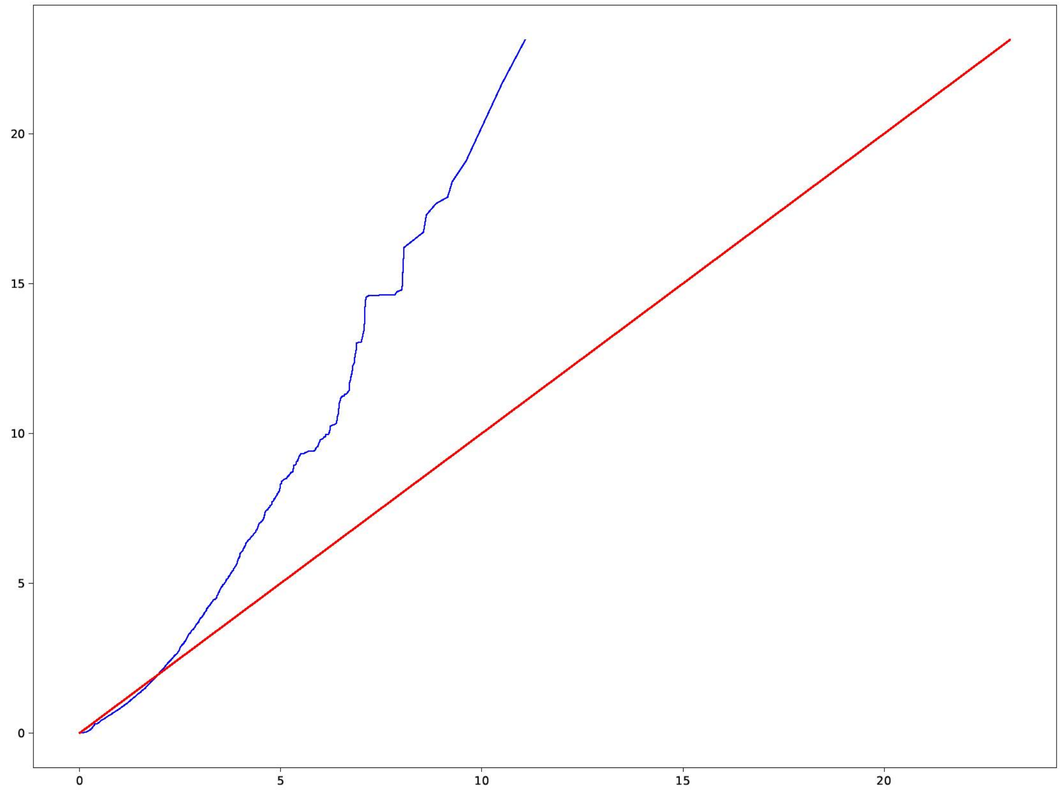
Van Hasselt, H., A. Guez, & D. Silver (2016): "Deep reinforcement learning with double q-learning." In "AAAI," pp. 2094–2100.

Appendix A

Plots of the Hawkes process-based limit order book model

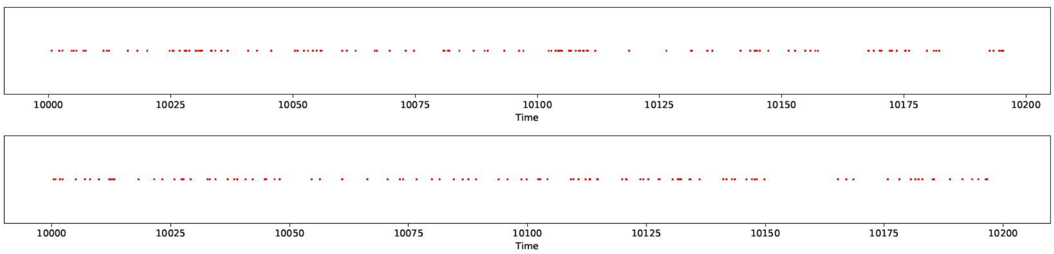
As a side note, we shall mention that all the following plots concerning the Hawkes process are done with a constant base intensity specification. No principal differences were observed between the constant and the quadratic intensity, and therefore for the sake of consistency we present only the plots for the base intensity one. Also, because we wanted to stay concise, only a selection of plots is provided. In case of interest, feel free to request the rest from the author.

Figure A.1: Q-Q plot of the compensator for the trade process for NFLX ASK



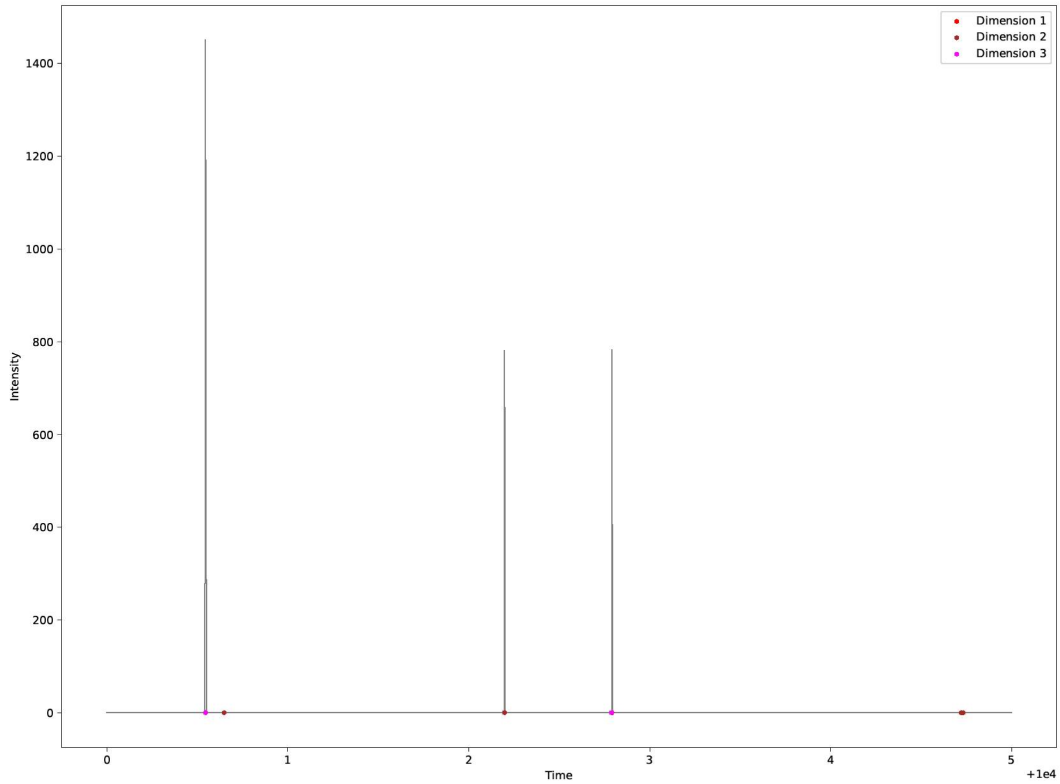
Note: An alignment of red and blue line would suggest the process specified is indeed a Hawkes process.

Figure A.2: Comparison of actual and simulated events for the maker-taker quote process for NFLX BID for range (10,000, 10,200)



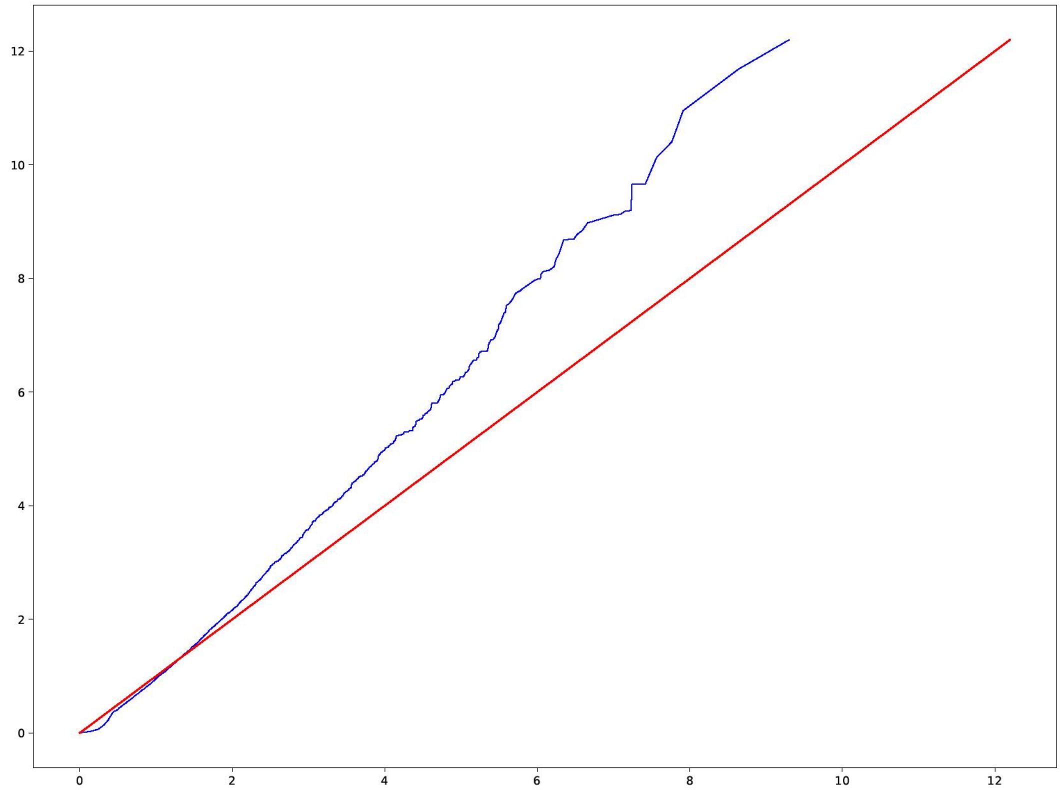
Note: Process in the upper cell is an extract from the dataset, while the lower one is a simulated Hawkes process with parameters set to median of estimated values.

Figure A.3: Plot of the conditional intensity function for the taker-maker quote process for NFLX BID for range (10,000, 10,005)



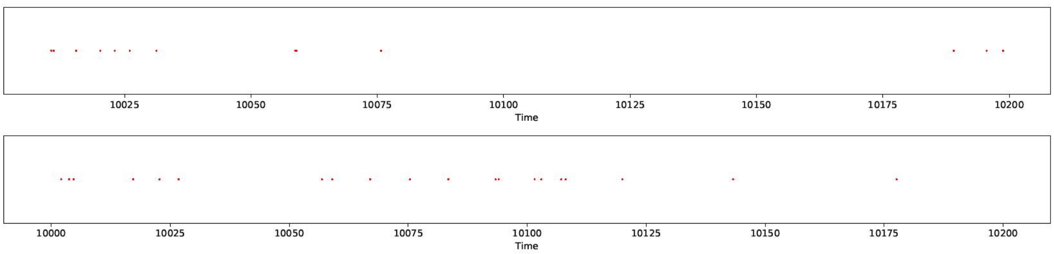
Note: Not each reaction to an event from a different dimension might be captured due to grid sampling of the intensity function for different dimensions.

Figure A.4: Q-Q plot of the compensator for the trade process for NVDA BID



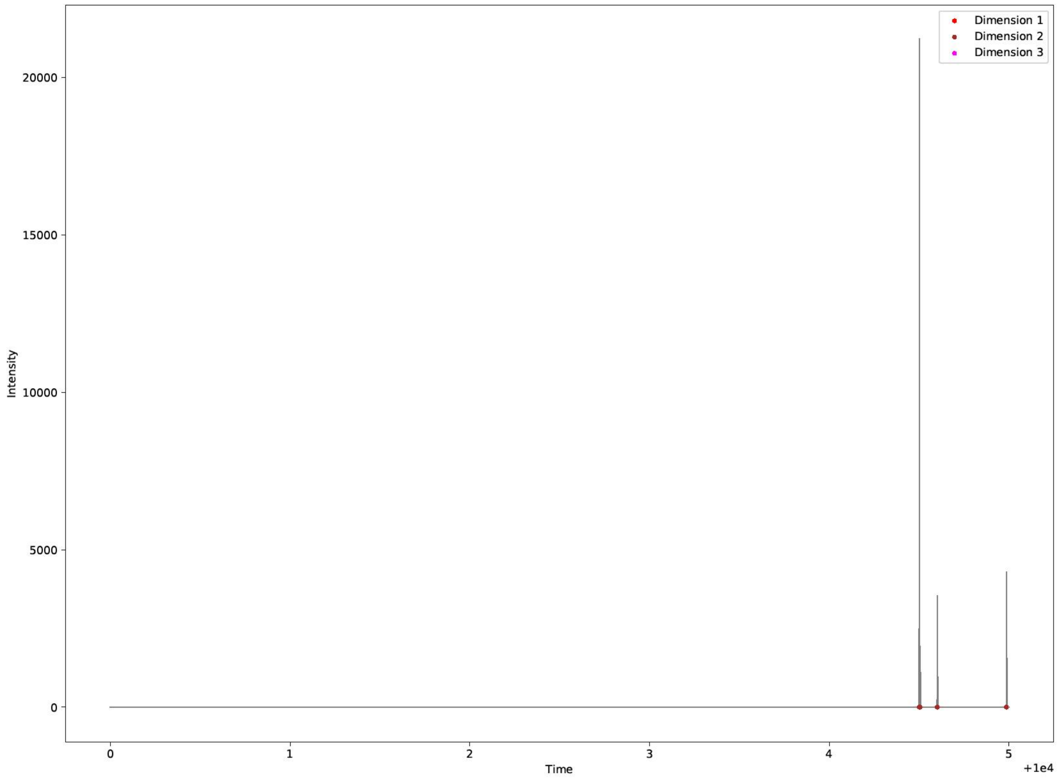
Note: An alignment of red and blue line would suggest the process specified is indeed a Hawkes process

Figure A.5: Comparison of actual and simulated events for the maker-taker quote process for NVDA ASK for range (10,000, 10,200)



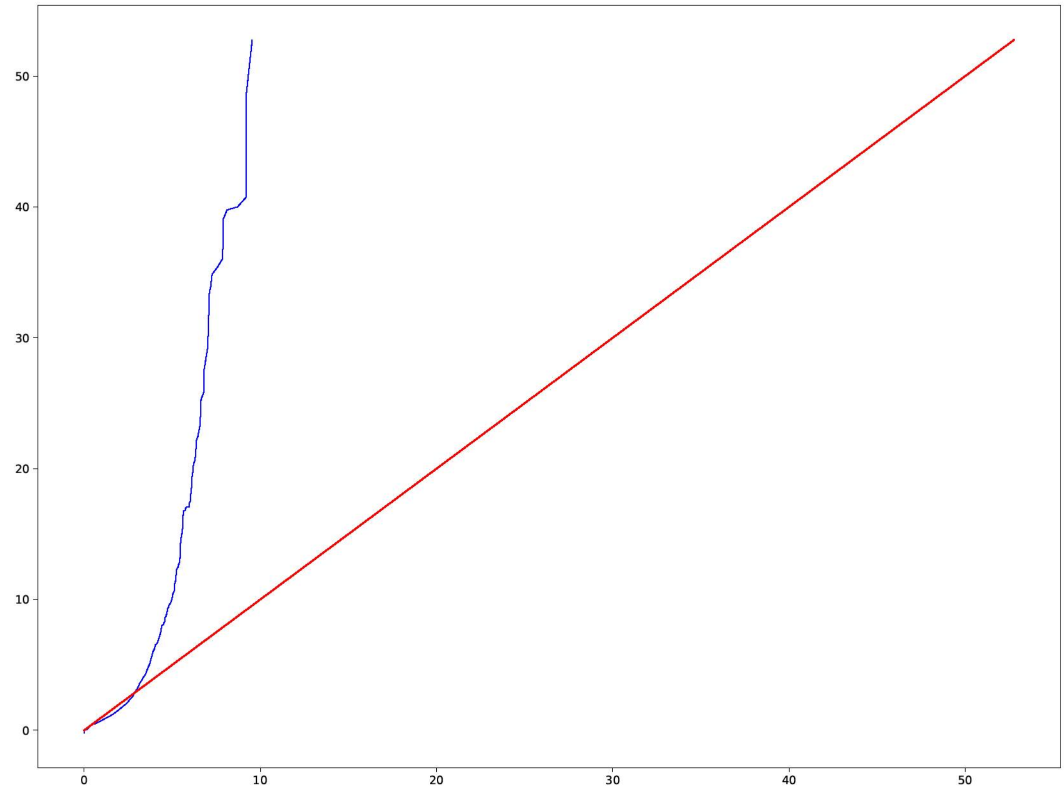
Note: Process in the upper cell is an extract from the dataset, while the lower one is a simulated Hawkes process with parameters set to median of estimated values.

Figure A.6: Plot of the conditional intensity function for the taker-maker quote process for NVDA ASK for range (10,000, 10,005)



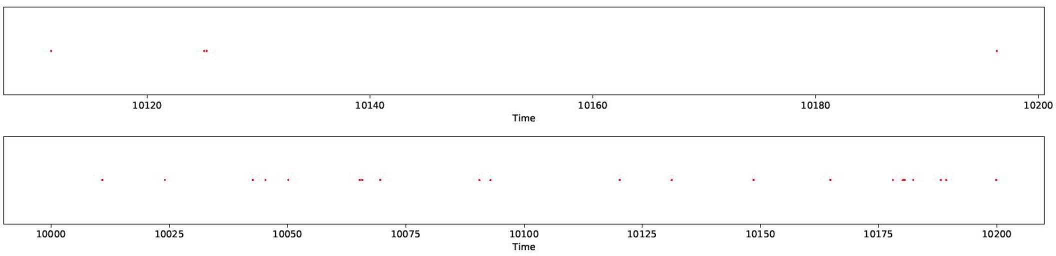
Note: Not each reaction to an event from a different dimension might be captured due to grid sampling of the intensity function for different dimensions.

Figure A.7: Q-Q plot of the compensator for the maker-taker quote process for SHPG ASK



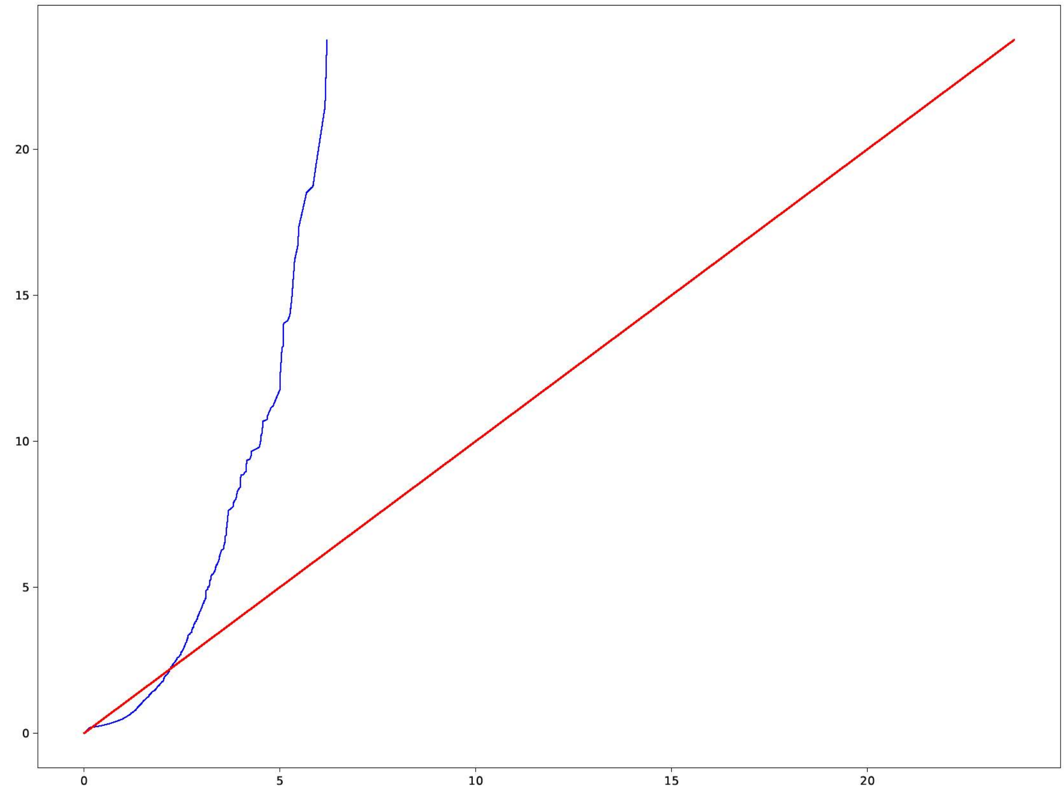
Note: An alignment of red and blue line would suggest the process specified is indeed a Hawkes process

Figure A.8: Comparison of actual and simulated events for the trade process for SHPG BID for range (10,000, 10,200)



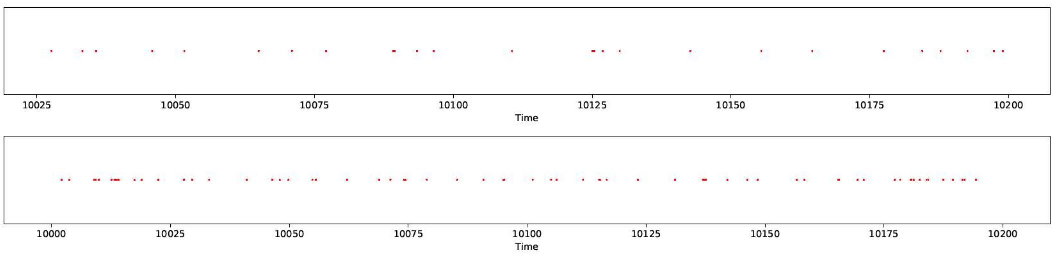
Note: Process in the upper cell is an extract from the dataset, while the lower one is a simulated Hawkes process with parameters set to median of estimated values.

Figure A.9: Q-Q plot of the compensator for the taker-maker quote process for TMUS ASK



Note: An alignment of red and blue line would suggest the process specified is indeed a Hawkes process

Figure A.10: Comparison of actual and simulated events for the trade process for TMUS BID for range (10,000, 10,200)



Note: Process in the upper cell is an extract from the dataset, while the lower one is a simulated Hawkes process with parameters set to median of estimated values.