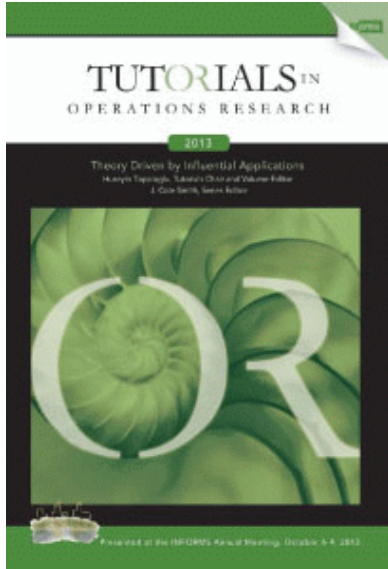




INFORMS TutORials in Operations Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Optimal Placement in a Limit Order Book

Xin Guo

To cite this entry: Xin Guo. Optimal Placement in a Limit Order Book. *In* INFORMS TutORials in Operations Research. Published online: 14 Oct 2014; 191-200.

<https://doi.org/10.1287/educ.2013.0113>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2013, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Optimal Placement in a Limit Order Book

Xin Guo

Department of Industrial Engineering and Operations Research, University of California,
Berkeley, Berkeley, California 94720, xinguo@ieor.berkeley.edu

Abstract Algorithmic trading refers to the automatic and rapid trading of large quantities with orders specified and implemented by an algorithm. Roughly speaking, algorithmic trading is based on two different time scales: the daily or weekly scale and a smaller (10- to 100-second) time scale. These two time scales essentially reflect the two steps by which the traders slice and place orders. The first step is to optimally slice big orders into smaller ones on a daily basis with the goal to minimize the price impact and/or to maximize the expected utility; the second step is to optimally place the orders within seconds. The former is the well-known optimal execution problem, and the latter is the much less studied optimal placement problem. This paper reviews several simple models and approaches for the optimal placement problem. Several most relevant statistical issues are presented, together with a brief discussion on the key differences between the system of limit order book and the multiclass queues with reneging.

Keywords limit order book; high-frequency trading; optimal placement; correlated random walk; diffusion limit; queues with reneging

1. Introduction

Technological innovation has completely transformed the fundamentals of the financial market. As a result, automatic and electronic order-driven trading platforms have largely replaced the traditional floor-based trading. In an electronic order-driven market, orders arrive at the exchange and wait in the *limit order book (LOB)* to be executed. In the United States, high-frequency trading (HFT) firms represent 2% of the approximately 20,000 firms operating today, but account for 73% of all equity orders volume. In most exchanges, order flow is heavy with thousands of orders in seconds and tens of thousands of price changes in a day for a liquid stock. Meanwhile, the time for the execution of a market order has dropped below one millisecond. This new era of trading is commonly referred to as the *high-frequency trading* or *algorithmic trading*.

In an order-driven market, there are two types of buy/sell orders for market participants to post: market orders and limit orders. A *limit order* is an order to trade a certain amount of security (stocks, futures, etc.) at a given specified price. The lowest price for which there is an outstanding limit sell order is called the *ask price*, and the highest limit buy price is called the *bid price*. Limit orders are collected and posted in the LOB, which contains the quantities and the price at each price level for all limit buy and sell orders. A *market order* is an order to buy/sell a certain amount of the equity at the best available price in the LOB; it is then matched with the best available price, and a trade occurs immediately and the LOB is updated accordingly. A limit order stays in the LOB until it is executed against a market order or until it is canceled; cancellation is allowed at any time. The closer a limit order is to the bid/ask, the faster it may be executed. Most exchanges are based on the first-in, first-out (FIFO) policy for orders on the same price level, although some derivatives

FIGURE 1. Orders happened in the bid/ask queues.

	Price	Size		Price	Size		Price	Size		Price	Size
Ask	9.12	1,525	⇒	9.12	1,525	⇒	9.12	1,525	⇒	9.12	1,525
	9.11	3,624		9.11	3,624		9.11	3,624		9.11	3,624
	9.10	4,123		9.10	4,123		Ask 9.10	4,123		Ask 9.10	4,100
	9.09	1,235		9.09	1,235		9.09	1,235		9.09	1,235
	9.08	3,287		9.08	3,287		9.08	3,687		9.08	3,687
Bid	9.07	4,895	⇒	9.07	3,695	⇒	9.07	3,695	⇒	9.07	3,695
	9.06	3,645		9.06	3,645		9.06	3,645		9.06	3,645
	9.05	2,004		9.05	2,004		Bid 9.05	2,004		Bid 9.05	2,004
	9.04	3,230		9.04	3,230		9.04	3,230		9.04	3,230
	9.03	7,246		9.03	7,246		9.03	7,246		9.03	7,246

on some exchanges have the prorate microstructure; that is, an incoming market order is dispatched on all active limit orders at the best price, with each limit order contributing to execution in proportion to its volume. In this paper, unless specified otherwise, we always focus the discussions on the FIFO market, with extensions to the prorate microstructure whenever appropriate.

Order books are available with different levels of details. For example, the so called “level 1 order book” contains the best price level of the order book, whereas the “level 2 order book” provides the prices and quantities of the best five levels on both the ask and the bid sides. Figure 1 is a typical example showing the dynamics of the LOB of the top five levels: a market sell order of size 1,200, followed by a limit ask order of size 400 at price 9.08, and then a cancellation of size 23 for a limit ask order at the price 9.10.

Cancellation is a major and distinct feature of algorithm trading. It is a critical strategic tool used by high-frequency traders to test the market. On average, more than 80% of the orders are canceled within a second after they are submitted. An exemplary case was reported in October 2012 by CNBC news (Melloy [48]):

A single mysterious computer program that placed orders—and then subsequently canceled them—made up 4 percent of all quote traffic in the U.S. stock market last week [the week of October 5th, 2012] The program placed orders in 25-millisecond bursts involving about 500 stocks The algorithm never executed a single trade.

2. Optimal Placement with LOB

2.1. Optimal Placement vs. Optimal Execution

Roughly speaking, algorithmic trading is based on two different time scales: the daily or weekly scale and a smaller (10- to 100-second) time scale. These two time scales reflect two different types of traders: the slow scale for the ordinary investors and the fast scale for the others. They are referred to as fundamental traders and high-frequency traders, respectively, in a recent empirical work of Kirilenko et al. [43]. They show that these two different types of traders displayed statistically different behavior and played different roles during the flash crash of May 2010. In general, the former try to minimize the price impact of their trades and the latter tend to predict and act faster to price changes. The two time scales essentially

correspond to the two steps by which the traders slice and place orders. The first step is to optimally slice big orders into smaller ones on a daily basis and the second step is to optimally place the orders within seconds. The former is the well-known *optimal execution problem* and the latter is much less studied, and we call it the *optimal placement problem*.

2.1.1. Optimal Execution. The key issue in the optimal execution problem is the *price impact*. The major modeling hypothesis behind the optimal execution problem is that any trading strategy—especially that involves large amount of buying and selling within a short period of time—will have an impact on the stock price: too large of an order may depress the price and reduce the potential profit, and too many small transactions may be costly and may take too long to complete. These price impact models, unlike those in classical hedging and pricing, focus on controlling the speed/quantity of trading to minimize the price impact and/or to maximize expected utility functions. The study of the optimal execution problem was pioneered by Bertsimas and Lo [9], who analyze a discrete random walk model, and by Almgren and Chriss [4, 5] and Almgren [3], who consider continuous-time Bachelier-type Brownian motion models. The trading impact is assumed to be additive, and their analysis yields a volume-average-type static/deterministic trading strategy. Based on different model approaches, the price impact is then evolved into three different types: permanent, temporary, and transient. Later on, the notion of price manipulation is developed as analog of no arbitrage for derivative pricing. Standard references for these models with additive and possible nonlinear price impact include Huberman and Stanzl [40], Obizhaeva and Wang [50], Schied and Schöneborn [57], Schied et al. [58], Almgren and Lorenz [6], Alfonsi et al. [1, 2], Gatheral [27], Predoiu et al. [55], Schied et al. [58], Weiss [62], and works involving the geometric Brownian motion and/or multiplicative price impact include Gatheral and Schied [28], Forsyth et al. [23], and Guo and Zervos [33].

2.1.2. Optimal Placement. Optimal placement problem studies how to optimally place the small-sized orders in the LOB. Specifically, when using limit orders, traders do not need to pay the spread and most of the time even get a rebate. This rebate structure varies from exchange to exchange and leads to different optimization problems. For instance, in the Hong Kong Stock Exchange, successful executions of limit orders get a discount, i.e., a fixed percentage of the execution price, whereas in other places such as the London Stock Exchange, the discount may be a fixed amount. This rebate, however, comes with a risk because there is no guarantee of execution for limit orders. On the other hand, when using market orders, one has to pay both the spread between the limit and the market orders and the fee in exchange for a guaranteed immediate execution. Thus, the traders have to decide: Given a number of shares to buy or sell, should one use market orders, or limit orders, or both? How many orders should be placed at different price levels? What is the optimal sequence of order placement in a give time frame with multitrades? In essence, traders have to balance the trade-off between paying the spread and fees and execution/inventory risk when placing market orders and limit orders.

There is also the price impact issue in the optimal placement problem. For instance, the price impact of a market order is generally believed to be larger than that of a limit order, and the empirical study by Cont et al. [19] suggests that the price impact is more or less linear at the trade level.

Though optimal execution and optimal placement are two different problems, the former is sometimes studied with the incorporation of certain aspects of the latter, especially when the LOB is taken into consideration (see Alfonsi et al. [1], Predoiu et al. [55]). The latter, on the other hand, could be viewed as the former when the execution risk is removed, as shown by Guo et al. [34].

2.2. Optimal Placement

The optimal placement problem can be summarized as follows:

- One needs to buy N orders before time $T > 0$ ($T \approx 1/5$ minutes).

- One can split the N orders into $(N_{0,t}, N_{1,t}, \dots)$, where $N_{0,t}$ is the number of orders at the market price, $N_{1,t}$ is the number of orders at the best bid, $N_{2,t}$ is the number of orders at the second best bid, $\dots$, for $t = 0, 1, \dots, T$.
- No intermediate selling is allowed before time T .
- If the limit orders are not executed by time T , then one has to buy the nonexecuted orders at the market price at time T .
- When one share of limit order is executed, the market gives a rebate of $r > 0$.
- When the trader submits a share of market order, there is a fee of $f > 0$.

Given N , the objective is to find the optimal strategy $(N_{0,t}, N_{1,t}, \dots, N_{k,t})_t$ to minimize the overall total expected cost.

2.3. Discrete Models

2.3.1. Markov Chain Model for the Full LOB with $N = 1$. Hult and Kiessling [41] build a high-dimensional Markov chain model for the state and evolution of the entire LOB. Specifically, they assume that there are d price levels in the order book, say, $\pi^1 < \dots < \pi^d$, and that the Markov chain $X_t = (X_t^1, \dots, X_t^d)$ represents the volume at time t of buy orders with negative values and of sell orders with positive values at each price level. The generator matrix of X is $Q = (Q_{xy})$, where Q_{xy} denotes the transition matrix from state $x = (x^1, \dots, x^d)$ to $y = (y^1, \dots, y^d)$. For example, a limit sell order of size k at level j corresponds to a transition from state x to state $x + ke^j$. For the convenience of analysis, they also assume that the highest bid level is always lower than the lowest ask level and that there is always someone to sell at the highest possible price and buy at the lowest possible price. There is no specific constraint on the spread between the best bid and ask levels. Within this framework, they consider a special version of the optimal placement problem with $N = 1$; that is, an agent has to buy one unit and has to decide between place a market order at the best ask level or place the limit buy at a lower level. The key is then to compute the probability that a limit buy order is executed before the price moves up as well as the expected buy price resulting from a limit buy order. Using the potential theory for Markov chains, they derive conditions for the existence of an optimal strategy and provide a value-iteration algorithm to numerically find the optimal strategy. They also provide a calibration method for the Markov chain model using real data from a foreign exchange market.

2.3.2. Correlated Random Walk Model. Guo et al. [34] propose a correlated random walk model for the optimal placement problem. Their starting point is the characteristics of the LOB rather than the whole LOB and they focus on modeling the ask price. Fixing a finite time horizon $[0, T]$, they assume that the spread between the bid price and the ask price is constant, the number of price changes is the same as the time steps, and the ask price follows a correlated random walk with increment $+1$ or -1 ; that is, $A_n = \sum_{i=1}^n X_i$, with X_i taking two values ± 1 with transition probabilities

$$\begin{array}{c|cc} & +1 & -1 \\ \hline +1 & p_1 & q_1 \\ -1 & q_2 & p_2 \end{array},$$

where $p_1 + q_1 = 1 = p_2 + q_2$. The initial condition is $P_r(X_1 = 1) = p = 1 - P_r(X = -1)$. Note that in this model, $p_1 - p_2$ and p measure the “drift” of the correlated walk. For simplicity, they further assume that $p_1 = p_2 = p_c$. In these particular parameter choices, if $p_c = \frac{1}{2}$, then the price is a time-changed random walk, and if $p_c < \frac{1}{2}$, the price “mean reverts.”

They first consider the case without any price impact so that the optimal placement problem for N is reduced to finding the order level with the lowest expected cost to buy one share. They show for a one-period model that the optimal strategy is to use either the market order or the best bid. The threshold strategy is explicitly given in terms of the model

parameters, and is shown to be monotonic with respect to time and the fees; that is, with more time remaining, and/or with higher payoff from using limit orders, one is more likely to use limit orders.

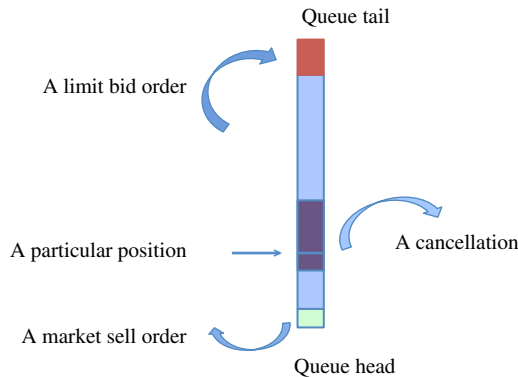
They then continue the analysis to a multiperiod model and also add a linear price impact. In these extended models they focus on the best bid and ask prices and add a (constant) probability q of execution of each limit order. In the multiperiod model without the price impact, they again derive a threshold-type optimal strategy among the choices of market order, limit order, and neither. In the price impact model, they draw analogy to the work on the optimal execution problem by Bertsimas and Lo [9] by replacing the correlated random walk model with a continuous-time, piecewise linear martingale process and consider the following simplified problem: for a given pair of N, T , the number of shares and the time frame, how many should be put in the market order and how many in the limit order in each time step between 0 and T ? They prove that the optimal strategy at each step is to use either the market order or the limit order, but not both, and they present a numerical algorithm for finding the optimal sequence for the number of market order and limit order at each step. In the particular case when there is no execution risk, i.e., $q = 0$, their solution reduces to that of Bertsimas and Lo [9].

2.4. Continuous-Time Model with Reduced Form

In general, the probability of a particular limit order being executed depends on the queue length, its position in the LOB, the frequency of price changes, the arrival rate of market orders, and the cancellation of orders, as Figure 2 shows. To capture the essence of the optimal placement problem, it is important to understand $Z(t)$, the dynamics of an order position at time t after it is placed on the LOB, because $Z(t)$ contains crucial information of the balance and relation between the execution risk and the microstructure of the LOB.

To study $Z(t)$, Guo and Ruan [32] focus on the queues of the best bid and the best ask. This “reduced form” approach was first proposed by Cont and de Larrard [17], where they draw connections between the LOB and multiclass priority queues. One of their key ideas to simplify the problem is the random initialization of the queues once a queue is depleted. With this assumption and by viewing the arrival of market, limit, and cancellation orders as point processes, they establish a functional central limit theorem after some technical assumptions including the stationary uniformly mixing condition on the sequence of order sizes (market, limit, and cancellation). They show that the limit dynamics of the LOB may be approximated by a Markov process $Q = (Q_t^a, Q_t^b)$ in the positive quadrant R_+^2 , which behaves like a planar Brownian motion in the interior of the quadrant but with a boundary

FIGURE 2. Orders happened in the best bid queue.



condition of an integral type, known also as the Wentzel condition. More precisely, Q is a Markov process with an infinitesimal generator of the form

$$\mathcal{A}h(x, y) = A \frac{\partial h}{\partial y} + B \frac{\partial h}{\partial x} + C \frac{\partial^2 h}{\partial y^2} + D \frac{\partial^2 h}{\partial x^2} + E \frac{\partial^2 h}{\partial y \partial x},$$

defined on a smooth function h with the boundary conditions

$$\begin{aligned} \int_{R_+^2} [h(u, v) - h(0, y)] F(du, dv) &= 0, \\ \int_{R_+^2} [h(u, v) - h(x, 0)] F(du, dv) &= 0, \end{aligned}$$

where A, B, C, D, E are constants associated with the means, variances, and correlations of various quantities from the LOB, and F is a bivariate distribution function. (See Cont [16], Cvitanic and Kirilenko [21], and Rosu [56] for other models of LOB dynamics).

Built on the results of Cont and de Larrard [17] for (Q_t^a, Q_t^b) , Guo and Ruan [32] provide several fluid and diffusion limits for $Z(t)$ under various cancellation assumptions. The fluid limit is relatively straightforward once the convergence of the limit established. The key to establish the diffusion limits comes from a nice result of Kurtz and Protter [44], which is a generalization of Wong and Zakai's [64] earlier work concerning passing the convergence relation between the càdlàg processes (X_n, Y_n) to (X, Y) in the Skorohod topology to the convergence relation between $\int X_n dY_n$ to $\int X dY$. With the fluid and diffusion limit for $Z(t)$, it is possible to calculate various quantities of interest for the LOB, including (i) the probability that an order is filled before the price increases, (ii) the expected price to pay if the order is not executed, (iii) the expected cost to place N_0 orders at the bid and $N - N_0$ at the ask, and (iv) the optimal strategy for placing orders.

2.4.1. Connection Between LOB and Classical Queues. To get a better sense of $Z(t)$, it helps to understand the fundamental differences between the classical queuing theory and the aforementioned reduced-form stochastic model of LOB. First, unlike the regulated Brownian motion in the classical heavy-traffic-limit queuing models (see, for example, Harrison [35], Whitt [63]), the underlying multidimensional diffusion process for the LOB involves a nonlocal Wentzel-type condition and is much less understood. Itô and McKean [42] explicitly prescribe how to construct a Brownian motion with such a Wentzel-type condition. However, characterization of such diffusions remains open, and it is analogous to the characterization of one-dimensional diffusion problem; the latter is now completely solved and characterized by Feller semigroups. There are some recent studies using the partial differential equation (PDE) approach to analyze the associated differential operator (Galakhov and Skubachevskii [25, 26]). Second, $Z(t)$ is not the same as the “workload process” in the classical queuing theory. In fact, to the best of our knowledge, this quantity has never been studied in the classical queuing, which seems to be more interested in the stability and analysis of the whole system rather than the dynamics of each individual in the system. Finally, despite some similarity between the limit order with order cancellation and the reneging queues (Ward and Glynn [60, 61]), the motivation and characteristics of cancellations in the limit order are very complex, especially when there are nondisplayable orders. This is an important issue because different cancellation assumptions lead to different fluid and diffusion limits, and the corresponding optimal trading strategies could be qualitatively distinct.

3. Beyond Optimal Placement

3.1. Related Statistical Issues

There have been many studies on the statistical properties of LOB, including autocorrelation of trade sizes (buy or sell), the average shape of LOB, the volume of orders, and the

duration between orders. (See, for instance, Bouchaud et al. [12] and He and Mamaysky [38].) There are also studies linking the characteristics of LOB with the trading behavior. For instance, Biais et al. [11] find that traders are more likely to submit limit orders when the LOB contains relatively few orders, or when the bid–ask spread widens. Hollifield et al. [39] use empirical analysis to test whether the optimal order submission strategy is a monotonic function of a trader’s valuation for the asset. There is a rich literature on market microstructure (O’Hara [52], Foucault et al. [24], Hasbrouck [36]), but most of it is for specialist-based markets and not for order-driven markets. Interested readers are referred to Cont [16] for more discussion on the statistical aspects of the LOB.

For the purpose of analyzing the optimal placement problem, several most relevant characteristic of LOBs are order cancellation, price mean reversion, and the order flow imbalance.

- *Order cancellation.* Because cancellation is frequently and strategically exploited by high-frequency traders to test the market, and because different assumptions on order cancellations may lead to different trading strategies, it is important to analyze the distribution of canceled orders versus total orders, and the distribution of canceled orders in terms of their positions in a LOB. When the data of LOB contain higher-level information such as the trader’s IDs, one can get a better picture for the motivation of order cancellation, by tracking, for example, the appearance or disappearance of certain IDs and the volume changes in the corresponding positions of the LOB.

- *Price mean reversion.* Mean reversion in prices has been observed by many practitioners in HFT, and some may even build trading strategies on betting the mean reversion. If one focuses on the six types of events, i.e., market buy and sell, limit buy and sell, and cancellation of buy and sell, then it is natural to model the arrival intensities of events by the multivariate Hawkes [37] process because it captures some of the clustering effect found in the LOB. The standard method for parameter estimation of Hawkes processes seems to be the maximum likelihood estimator (MLE) method (Ogata [51], Ozaki [53], Muni Toke and Pomponio [49]). The MLE method has a definite advantage of achieving the smallest possible asymptotic variance, at least in the univariate case. Studies on the consistency and asymptotic behavior of the MLE can be found in Chen and Hall [15] for the multivariate case and in Ogata [51] for the univariate case.

- *Order flow imbalance.* Order flow imbalance could be a strong indication of the market momentum. Recently, Cont et al. [19] examined the roles of order flow imbalance, trade imbalance, and the depth of LOB in affecting the price and found that order flow imbalance is a more general metric than the other two in limit order executions. Volume synchronized probability of informed trading (VPIN) is one of the well-known indicators developed by Easley et al. [22] for modeling order flow imbalance. (See also a recent report for testing its effectiveness by Bethel et al. [10].) Zheng et al. [65] use the lasso method to predict price jump and show that in the French stock market, trade sign and market order size as well as the liquidity on the best bid (best ask) are consistently informative for predicting the incoming price jump. Also, Lai and Xing [45] describe exponential autoregressive conditional duration (EACD) and Weibull autoregressive conditional duration (WACD) models to estimate limit order duration, and they also provide two statistical models (the ordered probit model and Rydberg–Shephard model) to predict price changes.

3.2. Market Making and Others

There are many optimization problems in the LOB beyond the particular one presented in this paper. For example, closely related to the optimal placement problem is the market-making problem, where trading strategies involve simultaneously placing limit orders to buy and sell. The idea is to maximize the profit by playing with the spread between the bid and ask prices while controlling the inventory risk and the execution risk. The profit is a reward for the market maker for his or her service to provide the liquidity to the market. For instance, Avellaneda and Stoikov [7], Bayraktar and Ludkovski [8], Cartea and

Jaimungal [13], Cartea et al. [14], Veraarta [59], Guilbaud and Pham [31], and Guéant et al. [30] study minimizing the inventory risk and balancing the execution risk with consideration of microstructure of LOB. There are also order scheduling problems, especially when orders can be placed in different exchanges with different fee structures. Cont and Kukanov [18] use the stochastic programming approach for the optimal placement problems across different LOB markets. See also Laruelle et al. [46] and Maglaras et al. [47].

4. Summary

This paper presents the optimal placement problem and several models and approaches. Related optimal execution problem and statistical issues are discussed.

Acknowledgments

The author would like to thank R. Cont, S. Jaimungal, A. Kercheval, A. de Larrard, D. Leinweber, J. Ma, C. Moallemi, J. Wu, S. Wong, K. Wong, and the participants in WCMF 2012, SIAM Conference on Financial Mathematics and Engineering 2012, and IMS-FPS2013 for comments and discussions. Generous financial and data support from NASDAQ OMX Educational Foundation, New York and CASH Dynamic Opportunities Investment Ltd., Hong Kong is acknowledged.

References

- [1] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance* 10(2):143–157, 2010.
- [2] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM Journal on Financial Mathematics* 3(1):511–533, 2012.
- [3] R. Almgren. Optimal execution with nonlinear impact function and trading enhanced risk. *Applied Mathematical Finance* 10:1–18, 2003.
- [4] R. Almgren and N. Chriss. Value under liquidation. *Risk* 12:61–63, 1999.
- [5] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk* 3:5–39, 2000.
- [6] R. Almgren and J. Lorenz. Adaptive arrival price. *Institutional Investor Journals* 1:59–66, 2007.
- [7] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance* 8(3):217–224, 2008.
- [8] E. Bayraktar and M. Ludkovski. Liquidation in limit order books with controlled intensity. *Mathematical Finance*. Forthcoming, 2013.
- [9] D. Bertsimas and A. W. Lo. Optimal control of execution codes. *Journal of Financial Markets* 1:1–50, 1998.
- [10] E. W. Bethel, D. Leinweber, O. Rübel, and K. Wu. Federal market information technology in the post flash crash era: Roles for supercomputing. Preprint, Lawrence Berkeley National Laboratory, Washington, DC, 2011.
- [11] B. Biais, P. Hillion, and C. Spatt. An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance* 50(5):1655–1689, 1995.
- [12] J.-P. Bouchaud, M. Mezard, and M. Potters. Statistical properties of stock order books: Empirical results and models. *Quantitative Finance* 2:251–256, 2002.
- [13] A. Cartea and S. Jaimungal. Modeling asset prices for algorithmic and high frequency trading. *Applied Mathematical Finance*. Forthcoming, 2013.
- [14] A. Cartea, S. Jaimungal, and J. Ricci. Buy low sell high: A high frequency trading perspective. Preprint, University College London, London, 2011.
- [15] F. Chen and P. Hall. Inference for a non-stationary self-exciting point process with an application in ultra-high frequency financial data modeling. *Journal of Applied Probability*. Forthcoming, 2013.
- [16] R. Cont. Statistical modeling of high frequency financial data: Facts, models and challenges. *IEEE Signal Processing* 28(5):16–25, 2011.

- [17] R. Cont and A. de Larrard. Price dynamics in a Markovian limit order book market. *SIAM Journal for Financial Mathematics* 4(1):1–25, 2013.
- [18] R. Cont and A. Kukanov. Optimal order placement in limit order markets. Preprint, Imperial College London, London, 2012.
- [19] R. Cont, A. Kukanov, and S. Stoikov. The price impact of order book events. *Journal of Financial Econometrics*. Forthcoming, 2013.
- [20] R. Cont, R. Talreja, and S. Stoikov. A stochastic model for order book dynamics. *Operations Research* 58(3):549–563, 2010.
- [21] J. Cvitanic and A. Kirilenko. High-frequency traders and asset prices. Working paper, California Institute of Technology, Pasadena, 2010.
- [22] D. Easley, L. M. de Prado, and M. O’Hara. Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies* 25(5):1457–1493, 2012.
- [23] P. A. Forsyth, J. S. Kennedy, S. T. Tse, and H. Windcliff. Optimal trade execution: A mean-quadratic-variation approach. *Journal of Economic Dynamics and Control* 36(2):1971–1991, 2012.
- [24] T. Foucault, O. Kadan, and E. Kandal. Limit order book as a market for liquidity. *Review of Financial Studies* 18(4):1171–1217, 2005.
- [25] E. I. Galakhov and A. L. Skubachevskii. On non-negative contractive semigroups with non-local conditions. *Sbornik: Mathematics* 189(1):43–74, 1998.
- [26] E. I. Galakhov and A. L. Skubachevskii. On Feller semigroups generated by elliptic operators with integrodifferential boundary conditions. *Journal of Differential Equations* 176(2):315–355, 2001.
- [27] J. Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance* 10(7):749–759, 2010.
- [28] J. Gatheral and A. Schied. Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework. *International Journal on Theoretical and Applied Finance* 14(3):353–368, 2011.
- [29] M. D. Gould, M. A. Porter, S. Williams, M. MacDonald, D. Fenn, and S. D. Howison. Limit order books. *Quantitative Finance*. Forthcoming, 2013.
- [30] O. Guéant, C.-A. Lehalle, and F. J. Tapia. Optimal portfolio liquidation with limit orders. *SIAM Journal of Financial Mathematics* 3(1):740–764, 2012.
- [31] F. Guillaud and H. Pham. Optimal high frequency trading with limit and market orders. *Quantitative Finance* 13(1):79–94, 2013.
- [32] X. Guo and Z. Ruan. Dynamics of the order position in a limit order book. Working paper, University of California, Berkeley, Berkeley, 2013.
- [33] X. Guo and M. Zervos. Optimal execution with nonlinear and multiplicative price impact. Preprint, University of California, Berkeley, Berkeley, 2012.
- [34] X. Guo, A. de Larrard, and Z. Ruan. Optimal order placement in a limit order book. Preprint, University of California, Berkeley, Berkeley, 2012.
- [35] M. Harrison. *Brownian Motion and Stochastic Flow Systems*. Krieger, Malabar, FL, 1990.
- [36] J. Hasbrouck. *Empirical Market Microstructure*. Oxford University Press, Oxford, UK, 2007.
- [37] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58(1):83–90, 1971.
- [38] H. He and H. Mamasky. Dynamic trading with price impact. *Journal of Economic Dynamics and Control* 29:891–930, 2005.
- [39] B. Hollifield, R. A. Miller, and P. Sandas. Empirical analysis of limit order markets. *Review of Economic Studies* 71:1024–1063, 2004.
- [40] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica* 72(4):1247–1275, 2004.
- [41] H. Hult and J. Kiessling. Algorithmic trading with Markov chains. Doctoral thesis, Stockholm University, Stockholm, 2010.
- [42] K. Itô and H. P. McKean Jr. *Diffusion Processes and Their Sample Paths*. Springer, New York, 2008.

- [43] A. Kirilenko, A. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. Preprint, Commodity Futures Trading Commission, Washington, DC, 2011.
- [44] T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 19(3):1035–1070, 1991.
- [45] T. L. Lai and H. Xing. *Statistical Models and Methods for Financial Markets*. Springer, New York, 2007.
- [46] S. Laruelle, C.-A. Lehalle, and G. Pagès. Optimal split of orders across liquidity pools: A stochastic algorithm approach. *SIAM Journal on Financial Mathematics* 2(1):1042–1076, 2011.
- [47] C. Maglaras, C. Moallemi, and H. Zheng. Optimal Routing in a fragmented market. Working paper, Columbia University, New York, 2012.
- [48] J. Melloy. Mysterious algorithm was 4% of trading activity last week. *CNBC* (October 8), <http://www.cnbc.com/id/49333454>, 2012.
- [49] I. Muni Toke and F. Pomponio. Modelling trades-through in a limit order book using Hawkes processes. *Economics: The Open-Access, Open-Assessment E-Journal* 6:2012–22, 2012.
- [50] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets* 16(1):1–32, 2013.
- [51] Y. Ogata. The asymptotic behavior of maximum likelihood estimators for stationary point processes. *Annals of the Institution of Statistical Mathematics* 30(1):243–261, 1978.
- [52] M. P. O'Hara. *Market Microstructure Theory*. Blackwell, Cambridge, MA, 1995.
- [53] T. Ozaki. Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institution of Statistical Mathematics* 31(1):145–155, 1979.
- [54] M. Potters and J.-P. Bouchaud. More statistical properties of order books and price impact. *Physica A* 3(24):133–140, 2003.
- [55] S. Predoiu, G. Shaikhet, and S. Shreve. Optimal execution in a general one-sided limit-order book. *SIAM Journal Financial Mathematics* 2:183–212, 2011.
- [56] I. Rosu. A Dynamic model of the limit order book. *Review of Financial Studies* 22:4601–4641, 2009.
- [57] I. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastics* 13:181–204, 2009.
- [58] I. Schied, T. Schöneborn, and Tehranchi. Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance* 17(6):471–489, 2010.
- [59] L. A. M. Veraarta. Optimal market making in the foreign exchange market. *Applied Mathematical Finance* 17(4):359–372, 2010.
- [60] A. Ward and P. Glynn. A diffusion approximation for a Markovian queue with reneging. *Queueing Systems* 43(1–2):103–128, 2003.
- [61] A. Ward and P. Glynn. A diffusion approximation for a GI/GI/1 queue with balking or reneging. *Queueing Systems* 50(4):371–400, 2005.
- [62] A. Weiss. Executing large orders in a microscopic market model. E-print, arXiv, <http://arxiv.org/abs/0904.4131v1>, 2009.
- [63] W. Whitt. *Stochastic-Process Limits*. Springer-Verlag, New York, 2002.
- [64] E. Wong and M. Zakai. Riemann-Stieltjes approximations of stochastic integrals. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 12:87–97, 1969.
- [65] B. Zheng, E. Moulines, and F. Abergel. Price jump prediction in limit order book. Working paper, Natixis Equity Markets, Paris, France, 2012.