Limit Order Books

A limit order book is essentially a file in a computer that contains all orders sent to the market, with their characteristics such as the sign of the order, price, quantity and a timestamp. The majority of organized electronic markets rely on limit order books to store lists of the interests of market participants in their central computer. A limit order book contains all information available on a specific market and it reflects the way the market moves under the influence of its participants.

This book discusses several models of limit order books. It begins by assessing the empirical properties of data, and then moves on to mathematical models in order to reproduce the observed properties. It finally presents a framework for numerical simulations. It also covers important modelling techniques including agent-based modelling, and advanced modelling of limit order books based on Hawkes processes. The book also provides in-depth coverage of simulation techniques and introduces general, flexible, open source library concepts useful to readers in studying trading strategies in order-driven markets.

The book will be useful to graduate students in the field of econophysics, financial mathematics and quantitative finance. The contents of this book are taught by the authors at CentraleSupélec (France) for a course on “Physics of Markets”. A short course based on the content of this book has been taught at the Graduate School of Mathematical Sciences, University of Tokyo (Japan), and it will be used at the Université Paris Saclay (France) for a course in quantitative finance.
Frédéric Abergel is a Professor and Director of the Chair of Quantitative Finance, CentraleSupélec, France. Beginning as a CNRS scientist at Université Paris Sud Orsay, he acquired several years of industrial experience in investment banking at BNP Paribas, CAI Cheuvreux, Barclays Capital and Natixis CIB. His research interests include financial markets, pricing and hedging of derivatives, quantitative finance and empirical properties of financial data.

Marouane Anane is a Quantitative Analyst at the BNP Paribas, Paris. His research interests include market making strategies, price dynamics and automated technical analysis.

Anirban Chakraborti is a Professor and Dean of the School of Computational and Integrative Sciences, Jawaharlal Nehru University, India. He has held academic/research positions at the Saha Institute of Nuclear Physics, Helsinki University of Technology, Brookhaven National Laboratory, Banaras Hindu University and the Ecole Centrale Paris. He is a recipient of the Young Scientist Medal of the Indian National Science Academy in 2009. His research areas include econophysics, statistical physics and quantum physics.

Aymen Jedidi is a Quantitative Analyst at HSBC Bank, Paris area, France. His research interests are quantitative risk management and stochastic order book modelling.

Ioane Muni Toke is an Associate Professor and Dean of studies at the Université de la Nouvelle-Calédonie, New Caledonia. He has held academic/research positions at the Ecole Centrale Paris and University of Texas at Dallas. He has research interests in financial markets modelling and microstructure, quantitative finance, statistical finance, applied mathematics and applied probability.
Physics of Society: Econophysics and Sociophysics

This book series is aimed at introducing readers to the recent developments in physics inspired modelling of economic and social systems. Socio-economic systems are increasingly being identified as ‘interacting many-body dynamical systems’ very much similar to the physical systems, studied over several centuries now. Econophysics and sociophysics as interdisciplinary subjects view the dynamics of markets and society in general as those of physical systems. This will be a series of books written by eminent academicians, researchers and subject experts in the field of physics, mathematics, finance, sociology, management and economics.

This new series brings out research monographs and course books useful for the students and researchers across disciplines, both from physical and social science disciplines, including economics.

Series Editors:

Bikas K. Chakrabarti
Professor, Saha Institute of Nuclear Physics, Kolkata, India

Mauro Gallegati
Professor of Economics, Polytechnic University of Marche, Italy

Alan Kirman
Professor emeritus of Economics, University of Aix-Marseille III, Marseille, France

H. Eugene Stanley
William Fairfield Warren Distinguished Professor Boston University, Boston, USA

Editorial Board Members:

Frédéric Abergel
Professor of Mathematics
CentraleSupélec, Chatenay-Malabry, France

Hideaki Aoyama
Professor, Department of Physics, Kyoto University, Kyoto, Japan

Anirban Chakraborti
Professor of Physics
Dean, School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, India

Satya Ranjan Chakravarty
Professor of Economics
Indian Statistical Institute, Kolkata, India

Arnab Chatterjee
Visiting Scientist of Physics
Saha Institute of Nuclear Physics, Kolkata, India

Shu-Heng Chen
Professor of Economics and Computer Science Director, AI-ECON Research Center, National Chengchi University, Taipei, Taiwan

Cars Hommes
Professor of Economics
Amsterdam School of Economics, University of Amsterdam Director, Center for Nonlinear Dynamics in Economics and Finance (CeNDEF), Amsterdam, Netherlands

Domenico DelliGatti
Professor of Economics
Catholic University, Milan, Italy
Kausik Gangopadhyay  
Professor of Economics  
Indian Institute of Management, Kozhikode, India

Giulia Iori  
Professor of Economics  
School of Social Science, City University, London, United Kingdom

Taisei Kaizoji  
Professor of Economics  
Department of Economics and Business, International Christian University, Tokyo, Japan

Kimmo Kaski  
Professor of Physics  
Dean, School of Science, Aalto University, Espoo, Finland

János Kertész  
Professor of Physics  
Center for Network Science, Central European University, Budapest, Hungary

Akira Namatame  
Professor of Computer Science and Economics Department of Computer Science, National Defense Academy, Yokosuka, Japan

Parongama Sen  
Professor of Physics  
University of Calcutta, Kolkata, India

Sitabhra Sinha  
Professor of Physics  
Institute of Mathematical Science, Chennai, India

Victor Yakovenko  
Professor of Physics  
University of Maryland, College Park, USA

Physics of Society: Forthcoming Titles

- **Macro-Econophysics: New Studies on Economic Networks and Synchronization** by Yoshi Fujiwara, Hideaki Aoyama, Yuichi Ikeda, Hiroshi Iyetomi, Wataru Souma, Hiroshi Yoshikawa

- **Interactive Macroeconomics: Stochastic Aggregate Dynamics with Heterogeneous and Interacting Agents** by Mauro Gallegati, Corrado Di Guilmi and Simone Landini

- **A Statistical Physics Perspective on Socio Economic Inequalities** by Victor Yakovenko and Arnab Chatterjee
Physics of Society: Econophysics and Sociophysics

Limit Order Books

Frédéric Abergel
Marouane Anane
Anirban Chakraborti
Aymen Jedidi
Ioane Muni Toke
# Contents

*Figures* xi  
*Tables* xv  
*Foreword* xvii  
*Preface* xix  
*Acknowledgments* xxi  

## 1 Introduction

### PART ONE  
**EMPIRICAL PROPERTIES OF ORDER-DRIVEN MARKETS**

### 2 Statistical Properties of Limit Order Books: A Survey  
2.1 Introduction 9  
2.2 Time of Arrivals of Orders 9  
2.3 Volume of Orders 12  
2.4 Placement of Orders 13  
2.5 Cancellation of Orders 15  
2.6 Average Shape of the Order Book 16  
2.7 Intraday Seasonality 18  
2.8 Conclusion 19  

### 3 The Order Book Shape as a Function of the Order Size  
3.1 Introduction 20  
3.2 Methodology 20  
3.3 The Regression Model 22  
3.4 Conclusion 28  

### 4 Empirical Evidence of Market Making and Taking  
4.1 Introduction 29  
4.2 Re-introducing Physical Time 29  
4.3 Dependency Properties of Inter-arrival Times 31
6.3.3 Interpreting the asymptotic volatility 74
6.4 The Role of Cancellations 75
6.5 Conclusion 76

7 The Order Book as a Queueing System 77
7.1 Introduction 77
7.2 A Link Between the Flows of Orders and the Shape of an Order Book 78
  7.2.1 The basic one-sided queueing system 78
  7.2.2 A continuous extension of the basic model 80
7.3 Comparison to Existing Results on the Shape of the Order Book 83
  7.3.1 Numerically simulated shape in Smith et al. (2003) 83
  7.3.2 Empirical and analytical shape in Bouchaud et al. (2002) 84
7.4 A Model with Varying Sizes of Limit Orders 88
7.5 Influence of the Size of Limit Orders on the Shape of the Order Book 92
7.6 Conclusion 96

8 Advanced Modelling of Limit Order Books 97
8.1 Introduction 97
8.2 Towards Non-trivial Behaviours: Modelling Market Interactions 97
  8.2.1 Herding behaviour 98
  8.2.2 Fundamentalists and trend followers 99
  8.2.3 Threshold behaviour 101
  8.2.4 Enhancing zero-intelligence models 101
8.3 Limit Order Book Driven by Hawkes Processes 102
  8.3.1 Hawkes processes 103
  8.3.2 Model setup 104
  8.3.3 The infinitesimal generator 105
  8.3.4 Stability of the order book 106
  8.3.5 Large scale limit of the price process 108
8.4 Conclusion 110

PART THREE
SIMULATION OF LIMIT ORDER BOOKS

9 Numerical Simulation of Limit Order Books 113
9.1 Introduction 113
9.2 Zero-intelligence Limit Order Book Simulator 113
  9.2.1 An algorithm for Poissonian order flows 113
  9.2.2 Parameter estimation 115
PART FOUR
IMPERFECTION AND PREDICTABILITY IN ORDER-DRIVEN MARKETS

10 Market Imperfection and Predictability
10.1 Introduction 145
10.2 Objectives, Methodology and Performances Measures 146
  10.2.1 Objectives 146
  10.2.2 Methodology 147
  10.2.3 Performance measures 148
10.3 Conditional Probability Matrices 148
  10.3.1 Binary case 150
  10.3.2 Four-class case 154
10.4 Linear Regression 156
  10.4.1 Ordinary least squares (OLS) 156
  10.4.2 Ridge regression 158
  10.4.3 Least Absolute Shrinkage and Selection Operator (LASSO) 163
  10.4.4 Elastic net (EN) 165
10.5 Conclusion 167

Appendix A A Catalogue of Order Types 169
Appendix B Limit Order Book Data 171
Appendix C Some Useful Mathematical Notions 176
Appendix D Comparison of Various Prediction Methods 187

Bibliography 209
Figures

1.1 A schematic illustration of the order book. 3
2.1 Distribution of interarrival times for stock BNPP.PA in log-scale. 10
2.2 Distribution of interarrival times for stock BNPP.PA. 10
2.3 Distribution of the number of trades in a given time period \( \tau \) for stock BNPP.PA. 11
2.4 Distribution of volumes of market orders. 12
2.5 Distribution of normalized volumes of limit orders. 13
2.6 Placement of limit orders. 14
2.7 Placement of limit orders. 14
2.8 Distribution of estimated lifetime of cancelled limit orders. 15
2.9 Distribution of estimated lifetime of executed limit orders. 16
2.10 Average quantity offered in the limit order book. 17
2.11 Average limit order book: price and depth. 17
2.12 Normalized average number of market orders in a 5-minute interval. 18
2.13 Normalized average number of limit orders in a 5-minute interval. 19
3.1 Mean-scaled shapes of the cumulative order book. 22
3.2 Mean-scaled number of market orders. 24
3.3 Mean-scaled shapes of the cumulative order book. 25
4.1 Empirical distribution function of the bid-ask spread in event time and in physical time. 30
4.2 Empirical distributions of the time intervals between two consecutive orders and of the time intervals between a market order and the immediately following limit order. 32
4.3 Empirical distributions of the time intervals between a market order and the immediately following limit order. 33
xii  Figures

4.4 Empirical distributions of the time intervals between two consecutive orders and of the time intervals between a limit order and an immediately following market order. 34
4.5 Impact functions on $M^1_{\text{buy}}$ arrival intensity. 39
4.6 Impact functions on the six events $O^1$. 40
5.1 Illustration of the Bak, Paczuski and Shubik model. 50
5.2 Snapshot of the limit order book in the Bak, Paczuski and Shubik model. 51
5.3 Empirical probability density functions of the price increments in the Maslov model. 52
5.4 Average return $\langle r_{\Delta t} \rangle$ as a function of $\Delta t$ for different sets of parameters and simultaneous depositions allowed in the Challet and Stinchcombe model. 54
5.5 Lifetime of orders for simulated data in the Mike and Farmer model. 56
5.6 Cumulative distribution of returns in the Mike and Farmer model. 56
6.1 Order book dynamics. 63
7.1 Shape (top panel) and cumulative shape (bottom panel) of the order book. 85
7.2 Comparison of the shapes of the order book in our model (black curves) and using the formula proposed by Bouchaud et al. (2002) (gray curves). 86
7.3 Price density function $\pi_{pa}$ as a function of the price. 88
7.4 Shape of the order book as computed in Eq. (7.34) (top) and cumulative shape of the order book as computed in Eq. (7.33) (bottom). 93
7.5 Shape of the order book as computed in Eq. (7.34). 94
7.6 Shape of the order book as computed in Eq. (7.34). 95
9.1 Model parameters: arrival rates and average depth profile (parameters as in Table 9.2). 118
9.2 Model parameters: volume distribution. Panels (a), (b) and (c) correspond respectively to market, limit and cancellation orders volumes. 119
9.3 Average depth profile. 120
9.4 Probability distribution of the spread. 121
9.5 Autocorrelation of price increments. 122
9.6 Price sample path. 122
9.7 Probability distribution of price increments. 123
9.8 Q-Q plot of mid-price increments. 123
9.9 Signature plot: $\sigma^2_{\hat{h}} := \mathbb{V} [P(t + h) - P(t)]/h$. 125
9.10 A cross-sectional comparison of liquidity and price diffusion characteristics between the model and data for CAC 40 stocks (March 2011). 126
9.11 Simulation of a two-dimensional Hawkes process with parameters given in Eq. (9.8) 132
9.12 Simulation of a two-dimensional Hawkes process parameters given in Eq. (9.8). (Zoom of Fig. 9.11) 132
9.13 Empirical density function of the distribution of the durations of market orders (left) and limit orders (right) for three simulations. 135
9.14 Empirical density function of the distribution of the time intervals between a market order and the following limit order for three simulations. 136
9.15 Empirical density function of the distribution of the bid-ask spread for three simulations. 136
9.16 Empirical density function of the distribution of the bid-ask spread for three simulations. 137
9.17 Empirical density function of the distribution of the 30-second variations of the mid-price for five simulations. 138
9.18 Hawkes parameters for aggressive limit orders for various CAC40 stocks. 139
9.19 Hawkes parameters for aggressive market orders for various CAC40 stocks. 140
10.1 The quality of the binary prediction: The AUC and the Accuracy are higher than 50%. 151
10.2 The quality of the binary prediction. 151
10.3 The quality of the binary prediction. 152
10.4 The quality of the binary prediction. 152
10.5 The quality of the binary prediction. 153
10.6 The quality of the 4-class prediction. 155
10.7 The quality of the OLS prediction. 157
10.8 The quality of the OLS prediction. 158
10.9 The quality of the OLS prediction. 159
10.10 The quality of the Ridge HKB prediction. 161
10.11 The quality of the Ridge LW prediction. 162
10.12 The quality of the Ridge prediction. 162
10.13 The quality of the LASSO prediction. 163
10.14 The quality of the LASSO prediction. 164
10.15 The quality of the LASSO prediction. 164
10.16 The quality of the EN prediction. 165
10.17 The quality of the EN prediction. 166
Tables

3.1 Basic statistics on the number of orders and the average volumes of orders per 30-minute time interval for each stock. 21
3.2 Panel regression results for the models defined in Eq. (3.1), using raw data, for $B_{k,t} = B_{k,t}^5$ (top panel), $B_{k,t} = B_{k,t}^{10}$ (middle panel) and $B_{k,t} = B_{k,t}^{10} - B_{k,t}^5$ (lower panel). 26
3.3 Panel regression results for the models defined in Eq. (3.1), using deseasonalized data, for $B_{k,t} = B_{k,t}^5$ (top panel), $B_{k,t} = B_{k,t}^{10}$ (middle panel) and $B_{k,t} = B_{k,t}^{10} - B_{k,t}^5$ (lower panel). 27
4.1 Event types definitions. 35
4.2 Event occurrences statistics summary. 36
4.3 Conditional probabilities of occurrences per event type. 36
4.4 Median conditional waiting time. 38
4.5 Mean conditional waiting time. 38
5.1 Analogy between the $A + B \rightarrow \emptyset$ reaction model and the order book in Bak et al. (1997) 50
5.2 Analogy between the deposition-evaporation process and the order book in Challet and Stinchcombe (2001). 53
5.3 Results of Smith et al. 58
9.1 Model parameters for the stock SCHN.PA (Schneider Electric) in March 2011 (23 trading days). 116
9.2 Model parameters for the stock SCHN.PA (Schneider Electric) in March 2011 (23 trading days). 116
9.3 CAC 40 stocks regression results. 129
9.4 Estimated values of parameters used for simulations. 133
10.1 Historical occurrences matrix for Deutsche Telekom over 2013. 148
10.2 Monthly historical conditional probabilities. 149
B.1 Tick by tick data file sample. 172
B.2 Trades data file sample. 173
B.3 Number of limit and markets orders. 174
Tables

D.1 The quality of the binary prediction: 1-minute prediction AUC and accuracy per stock. 187
D.2 The quality of the binary prediction: The daily gain average and standard deviation for the 1-minute prediction (without trading costs). 189
D.3 The quality of the binary prediction: The daily gain average and standard deviation for the 1-minute prediction (with trading costs). 190
D.4 The quality of the 4-class prediction: 1-minute prediction AUC and accuracy per stock. 192
D.5 The quality of the 4-class prediction: The daily gain average and standard deviation for the 1-minute prediction (without trading costs). 194
D.6 The quality of the 4-class prediction: The daily gain average and standard deviation for the 1-minute prediction (with trading costs). 195
D.7 The quality of the 4-class prediction: The daily gain average and standard deviation for the 30-minute prediction (without trading costs). 197
D.8 The quality of the binary prediction: The daily gain average and standard deviation for the 30-minute prediction (with 0.5 bp trading costs). 199
D.9 The quality of the OLS prediction: The AUC and the accuracy per stock for the different horizons. 200
D.10 The quality of the Ridge HKB prediction: The AUC and the accuracy per stock for the different horizons. 202
D.11 The quality of the Ridge LW prediction: The AUC and the accuracy per stock for the different horizons. 204
D.12 The quality of the LASSO prediction: The AUC and the accuracy per stock for the different horizons. 205
Foreword

When physicists became convinced that matter was not continuous but made from atoms, new ideas on old subjects started flourishing. Not only well-known macroscopic laws (thermodynamics, hydrodynamics) became better understood and bolstered by a more fundamental underlying reality, but a host of spectacular and often unexpected effects were rationalized, in particular collective emergent phenomena phase transitions, superconductivity, avalanches, etc. Similarly, after decades of mathematical finance devoted to the study of effective low frequency models of markets (chiefly based on variations on the Brownian motion), the increasing availability of high frequency data now allows a comprehensive study of price formation and of the microstructure of supply and demand. A new era of financial modelling is opening up, with the hope of addressing a hitherto neglected yet crucial aspect of price dynamics: feedback effects that can lead to market anomalies, instabilities and crashes. Instead of considering the market as an inert, reliable measurement apparatus that merely reveals the fundamental value of assets without influencing it, the empirical study of the order book reveals that markets do generate their own dynamics. New intuitions about market dynamics are necessary. New fascinating statistical regularities are collected and modelled, in particular using numerical simulations of agent based models. New analytical tools are being built to account for these observations. The final goal is, much as in physics, to understand the emergent phenomena and replace ad-hoc models of prices by micro-founded ones where jumps, fat-tails and clustered volatility would have a clear origin. This is important on many counts: while the intellectual endeavour is of course exciting in itself, its offshoots will deeply influence the way we think about market regulation in the wake of high-frequency trading, and the models we use for financial engineering (from derivative pricing to algorithmic trading and optimal execution).

Limit order books offers a much needed, broad review of a field that has literally exploded in the last 20 years, where researchers from economics, financial mathematics, physics, computer science, etc. compete and confront. This diversity is well illustrated by the content of the present book that covers a very wide ground, from empirical facts to advanced mathematical techniques and numerical simulation tools. It will be a very useful
Foreword

and inspiring entry point for all scientists, engineers, regulators and traders interested in understanding how financial markets really work at the basic level.

Jean-Philippe Bouchaud
Capital Fund Management & École Polytechnique
Preface

The Chair of Quantitative Finance was created at École Centrale Paris, now CentraleSupélec, in 2007. Since its inception, most of its research activities were devoted to the study of high frequency financial data. The interdisciplinary nature of the team, composed of mathematicians, financial engineers, computer scientists and physicists, gave it a special dimension. A sizeable portion of its research efforts has been focused on the characterization and mathematical modelling of limit order books.

Literally at the core of every modern, electronic financial market, the limit order book has triggered a huge amount of research in the past twenty years, marked by the seminal work of Biais et al. (1995) on the empirical analysis of the Paris exchange and revitalized a few years later, in a fascinating manner, by the work of Smith et al. (2003). However, much as this topic is interesting, important and challenging, we realized that there was still no reference book on the subject! We therefore decided to assemble in a single document a survey of the existing literature and our own contributions on limit order books, whether they were pertaining to their statistical properties, mathematical modelling or numerical simulation.

We have tried to follow the intellectual approach of an experimental physicist: empirical data should come first, and only empirical analyses may be considered as a reliable ground for building up any kind of theory. The mathematical modelling follows. Models address the different phenomena that are observed and highlighted, and provide a framework to explain and reproduce these phenomena, and they are studied from theoretical, analytical and numerical perspectives.

The book is thus organized as follows: The first part is devoted to the empirical properties of limit order books; the second part, to their mathematical modelling and the third, to their numerical analysis. The fourth part deals with some advanced topics such as imperfection and predictability. Each part presents a survey of the existing scientific literature, as well as our own contributions.

Significant parts of the material covered in this book have already been presented in bits and pieces in different research and survey articles, in particular Chakraborti et al. (2011a,b); Abergel and Jedidi (2013, 2015); Anane and Abergel (2015); Muni Toke (2015, 2011). However, what was lacking was a consistent and systematic compilation of these,
Preface

found in a single place where the emphasis was set on a single object of interest. We hope that this book will fulfil this need and complement the already existing abundant literature on market microstructure. The interdisciplinary approaches, with the stress on both empirical data analyses and theoretical studies, will hopefully render it useful to the reader – researcher, graduate student or practitioner, while facilitating him/her in finding most of the contemporary knowledge on this essential component of financial markets.
Acknowledgments

We are grateful to all our collaborators: Nicolas Huth, Anton Kolotaev, Mehdi Lallouache, Nicolas Millot, Marco Patriarca, Fabrizio Pomponio, Mauro Politi, Rémi Tachet, Riadh Zaatour and Ban Zheng, for their contributions to these developments. We also acknowledge Damien Challet, Rémy Chicheportiche, Charles-Albert Lehalle, Grégoire Loeper, Eric Moulines, Françoise Roueff, Mathieu Rosenbaum, Stéphane Tyc, Nakahiro Yoshida for fruitful discussions and inputs.

We are thankful to BNP Paribas for their generous funding to the Chair of Quantitative Finance.

Some sections of this book were written while Ioane Muni Toke was a Senior Fellow in the “Broad Perspectives and New Directions in Financial Mathematics” program of the Institute for Pure and Applied Mathematics, University of California at Los Angeles.

Ioane Muni Toke also acknowledges the support of the CREST project of the Japan Science and Technology Agency.

Several sections of this book were completed while Frédéric Abergel was visiting the Graduate School of Mathematical Sciences at the University of Tokyo, and the Laboratoire de Probabilités et Modèles Aléatoires at CNRS, Université Pierre et Marie Curie and Université Denis Diderot. He is grateful for the support of these institutions.

Some of the material presented in this book has been previously published and is used here with kind permission from Springer Science+Business Media, Taylor and Francis and World Scientific Publishing:


Acknowledgments

