Neural Stochastic Agent-Based Limit Order Book Simulation: A Hybrid Methodology

Extended Abstract

Zijian Shi
University of Bristol
Bristol, United Kingdom
zijian.shi@bristol.ac.uk

John Cartlidge
University of Bristol
Bristol, United Kingdom
john.cartlidge@bristol.ac.uk

ABSTRACT

Realistic limit order book (LOB) simulations are essential in understanding market dynamics. Mainstream simulation models include agent-based models (ABMs) and stochastic models (SMs). However, ABMs tend not to be grounded on real historical data, while SMs tend not to enable dynamic LOB interaction. Here, we propose a hybrid LOB simulation paradigm characterised by: (1) representing the aggregation of market events' logic by a neural stochastic background trader (BT) that is pre-trained on historical LOB data through a neural point process model; and (2) embedding the BT into an ABM to enable responsive interaction. Empirical results demonstrate that system behaviours exhibit multiple stylised facts, and the results of interaction between the BT and various trading strategies are in accordance with observations of real markets.

CCS CONCEPTS

• **Computing methodologies** → **Agent / discrete models**; **Neural networks**; • **Applied computing** → **Economics**.

KEYWORDS

Market simulation; Neural point process; Agent-based model

ACM Reference Format:
Zijian Shi and John Cartlidge. 2023. Neural Stochastic Agent-Based Limit Order Book Simulation: A Hybrid Methodology: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

In financial markets, a LOB is a continuously updating queueing system for limit orders: orders to buy (sell) some quantity q of a security for at most (least) price p. It depicts the most fine-grained level of demand and supply information concerning a specific financial asset, and has been used as a primary data source in financial microstructure studies [4, 17, 18]. Although the LOB is of critical importance in research, using historical LOB data suffers from two major problems: it is not possible to conduct 'what if' counterfactual analysis that requires perturbing historical prices, and the public availability of LOB data is limited. Simulations and synthetically generated data can help to overcome these problems.

Two mainstream models for LOB simulation are agent-based models (ABMs) and stochastic models (SMs). However, ABMs have subjectivity issues and are not guaranteed to behave in the same manner as the real world [11], while SMs lack adequate infrastructure to allow interaction between agents as in a real exchange. In recent years, deep learning has been used for realistic LOB simulation [13, 19, 21]. However, as far as the authors are aware, only [8] have proposed a method that enables responsive interaction, by embedding a CGAN-based generator inside an ABM framework.

Here, we present a hybrid neural stochastic agent-based model (NS-ABM) to simulate the LOB. We introduce a neural stochastic 'background trader' (BT) – pre-trained on real-world level-2 LOB data using a state-dependent parallel neural Hawkes process (sd-PNHP) model [20] – which mimics aggregated order events of the whole market. We then incorporate the BT into the open-source ABIDES [6] platform and show that the BT is able to reproduce a comprehensive list of ten stylised facts about real world LOBs. Finally, a series of interaction experiments demonstrate that the BT reacts realistically to endogenous agent-driven events.[1]

2 MODEL FORMULATION

Here, we deploy the model previously developed in [20] as an autonomous BT agent within ABIDES (see below) so that it can interact and respond to the actions of other trading agents in the market. We propose that this 'hybrid' ABM approach can generate realistic and responsive market dynamics.

2.1 Neural Stochastic Background Trader

Pure ABMs cannot learn an individual agent's behaviour from data, as LOB events are anonymous. However, it *is* possible to learn the overall order stream pattern of the whole market [8, 20]. Using a sd-PNHP model [20], coupled with several empirical distributions regarding order statistics and a stochastic sampling method, we are able to sample LOB order streams. The model can also be configured ('order flow impact' set on/off) to respond to interactions with other trading strategies by incorporating trading agents' orders into the BT's memory. When the BT is configured to include order flow impact, agents' order events cause the BT to respond, resulting in order flow price impact in the system.

2.2 ABIDES

ABIDES [6] – Agent-Based Interactive Discrete Event Simulation – is an open-source agent-based simulation platform for conducting LOB microstructure experiments. We adopt ABIDES as the agent simulation platform to embed our neural-stochastic BT, and we consider interactions with ABIDES' trend agents and value agents.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

[1] A longer preprint version of this paper is available: http://arxiv.org/abs/2303.00080

Trend agents assume that a trend in price will persist or reverse in the near future, and they are named momentum (MM) agents and mean reversion (MR) agents, respectively. Trend agents calculate the current trend in price using the difference between moving average price over a long and a short time window, and trade accordingly. Value agents trade according to observations on a fundamental value oracle that is exogeneous to the market [5]. The observed fundamental value represents the 'fair price' of the stock, and value agents trade whenever the price deviates from this value. Both zero intelligence (ZI) agents and heuristic belief (HBL) agents trade based on this logic, while HBL agents further exploit historical LOB transactions to maximise the probability of order execution.

3 EXPERIMENTS AND RESULTS

3.1 Sensitivity Analysis

Following the method of [16], we first conduct sensitivity analysis on a system in which all orders are generated by the BT. We consider key input parameters that relate to the empirical distributions of order price and volume, the percentage of market orders in all orders, and some control parameters that enforce the LOB will not run out of liquidity. Results show that the two most influential parameters are: (1) the exponent of distribution for market order volume; and (2) market order percentage imbalance. These findings concur with [16], even though the methods adopted differ.

3.2 Reproducing Stylised Facts

'Stylised facts' in economics are empirical findings that are so consistently observed that they are generally accepted as truth. Such facts can therefore be used to verify the fidelity of a simulation. The simulated LOB is able to reproduce a list of more than ten stylised facts, including: Hurst exponent for volatility [9]; auto correlation in order-sign series [14]; order flow imbalance impact [10]; price impact concavity [15]; and six facts previously verified in [20].

3.3 Agent Interactions

By allowing various trading strategies (MM, MA, ZI, HBL) to interact with the BT, we can gain insights into how the system reacts to endogenous events, and how those reactions compare to real markets. Several criteria are considered: (i) profitability; (ii) market volatility; (iii) BT's ask-bid order imbalance; and (iv) the correlation between fundamental stock value and realised stock price. We vary the existence of order flow impact, number of agents $n \in \{1, 15, 50\}$, and use a mean-reverting fundamental value for value agents.

3.3.1 Herding Effect. We study the influence of financial herding behaviours, taking reference from the change in market statistics caused by changing the number of agents holding the same strategy.

In terms of (i), we see that mean profitability increases with number of agents for all strategies. This conforms with empirical studies indicating that the intensity of herding is positively related to trading profitability, especially for trend strategies [2, 7]. In terms of (ii), we see that increasing the number of MM agents causes increasing volatility, while the opposite effect occurs for other agent types. This is to be expected for MMs as price momentum drive further movements in the same direction, which can even lead to tail events such as a market crash [3]. In contrast, MR agents and value agents follow a mean-reverting fundamental value oracle. Therefore, an increase in the number of agents tends to reduce market volatility. In terms of (iv), studies on the futures market indicated that HFTs help prices converge to the fundamental [12]. We find value agents facilitate this process, as they are constantly comparing between fundamental value and the realised stock price. Increasing the number of agents accelerates the rate of convergence.

3.3.2 Order Flow Impact. We compare markets with and without order flow impact, and see significant variations in (iii). For trend agents, the order flow impact manifests as causing further trend following events (i.e., pushing price farther away, or pulling price back towards, the mean). When the BT is set to have no order impact, the underestimation of volatility in a MM market is caused by overlooking the momentum ignition, which indicates that investors tend to follow the price trend inititated by HFTs [1]. As for value agents, we find that correlation is smaller when order flow impact is excluded. This result indicates that the inclusion of order flow impact is able to model the price discovery role of value agents.

3.3.3 Competition between strategies. To understand strategy interaction, we performed heterogeneous experiments, with the existence of order flow impact: (1) 'trend' markets containing 15 MM *vs* 15 MA; and (2) 'value' markets containing 15 ZI *vs* 15 HBL. In trend markets, profits of both strategies are inferior to the profits generated in homogeneous markets; and the resulting market volatility falls ≈50% compared with a homogeneous MM market, indicating MR agents' mean-reverting impact. In value markets, HBL achieve slightly higher profits than ZI, but differences are not significant.

3.3.4 Responsiveness. We investigate price impact by introducing a percent-of-volume (POV) agent to interact with the BT. A λ-POV agent submits a series of market orders as a λ fraction of total market volume. We separate price impact into two components: the plain price impact caused by the POV agent when the BT is not responding, and the additional order flow impact caused by the BT responding to the actions of the POV agent.

We find that the plain price impact increases approximately linearly with λ in range $[0.01, 0.5]$. In contrast, order flow impact increases superlinearly. For low values of $\lambda = 0.01$, trading volume is too low for the BT to respond; while for high values of λ, there is a non-trivial response from the BT which produces a convex influence on price. This portion of price impact is generally missing when backtesting trading algorithms, where the market is unrealistically assumed to be unresponsive. This again highlights the utility of dynamic simulation platforms, such as we have presented here.

4 CONCLUSIONS

NS-ABM combines the benefits of ABM with data-driven methods, by including a neural stochastic BT trader that is pre-trained on real data. Empirical results demonstrate that the BT can realistically simulate real world LOB dynamics, and react realistically to endogenous market events. NS-ABM, as a whole, provides a promising alternative perspective in realistic and interactive LOB simulation.

ACKNOWLEDGMENTS

ZS's PhD is funded by a China Scholarship Council / University of Bristol joint-scholarship.

REFERENCES

[1] Bruno Biais, Thierry Foucault, et al. 2014. HFT and market quality. *Bankers, Markets & Investors* 128, 1 (2014), 5–19.

[2] Sushil Bikhchandani and Sunil Sharma. 2000. Herd behavior in financial markets. *IMF Staff papers* 47, 3 (2000), 279–310.

[3] Ekkehart Boehmer, Dan Li, and Gideon Saar. 2018. The competitive landscape of high-frequency trading firms. *The Review of Financial Studies* 31, 6 (2018), 2227–2276.

[4] Jean-Philippe Bouchaud, Marc Mézard, and Marc Potters. 2002. Statistical properties of stock order books: empirical results and models. *Quantitative finance* 2, 4 (2002), 251.

[5] David Byrd. 2019. Explaining agent-based financial market simulation. arXiv:1909.11650v1. https://doi.org/10.48550/arXiv.1909.11650

[6] David Byrd, Maria Hybinette, and Tucker Hybinette Balch. 2020. ABIDES: Towards High-Fidelity Multi-Agent Market Simulation. In *Proceedings of the 2020 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation* (Miami, FL, Spain) *(SIGSIM-PADS '20)*. Association for Computing Machinery, New York, NY, USA, 11–22.

[7] Qiwei Chen, Xiuping Hua, and Ying Jiang. 2018. Contrarian strategy and herding behaviour in the Chinese stock market. *The European Journal of Finance* 24, 16 (2018), 1552–1568.

[8] Andrea Coletta, Aymeric Moulin, Svitlana Vyetrenko, and Tucker Balch. 2022. Learning to simulate realistic limit order book markets from data as a World Agent. In *Proceedings of the Third ACM International Conference on AI in Finance*. Association for Computing Machinery, New York, NY, USA, 428–436.

[9] Rama Cont. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance* 1, 2 (2001), 223.

[10] Rama Cont, Arseniy Kukanov, and Sasha Stoikov. 2014. The price impact of order book events. *Journal of financial econometrics* 12, 1 (2014), 47–88.

[11] Martin D Gould, Mason A Porter, Stacy Williams, Mark McDonald, Daniel J Fenn, and Sam D Howison. 2013. Limit order books. *Quantitative Finance* 13, 11 (2013), 1709–1742.

[12] Eun Jung Lee. 2015. High frequency trading in the Korean index futures market. *Journal of Futures Markets* 35, 1 (2015), 31–51.

[13] Junyi Li, Xintong Wang, Yaoyang Lin, Arunesh Sinha, and Michael Wellman. 2020. Generating realistic stock market order streams. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34(01). AAAI Press, Palo Alto, CA, 727–734.

[14] Fabrizio Lillo and J Doyne Farmer. 2004. The long memory of the efficient market. *Studies in nonlinear dynamics & econometrics* 8, 3 (2004), 1–33.

[15] Fabrizio Lillo, J Doyne Farmer, and Rosario N Mantegna. 2003. Master curve for price-impact function. *Nature* 421, 6919 (2003), 129–130.

[16] Frank McGroarty, Ash Booth, Enrico Gerding, and VL Chinthalapati. 2019. High frequency trading strategies, market fragility and price spikes: an agent based model perspective. *Annals of Operations Research* 282, 1 (2019), 217–244.

[17] Randi Næs and Johannes A Skjeltorp. 2006. Order book characteristics and the volume–volatility relation: Empirical evidence from a limit order market. *Journal of Financial Markets* 9, 4 (2006), 408–432.

[18] Ingmar Nolte, Mark Salmon, and Chris Adcock. 2016. *High frequency trading and limit order book dynamics*. Routledge, Abingdon, England.

[19] Zijian Shi and John Cartlidge. 2021. The Limit Order Book Recreation Model (LOBRM): An Extended Analysis. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, Cham, Switzerland, 204–220.

[20] Zijian Shi and John Cartlidge. 2022. State Dependent Parallel Neural Hawkes Process for Limit Order Book Event Stream Prediction and Simulation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, New York, NY, USA, 1607–1615.

[21] Zijian Shi, Yu Chen, and John Cartlidge. 2021. The LOB Recreation Model: Predicting the Limit Order Book from TAQ History Using an Ordinary Differential Equation Recurrent Neural Network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35(1). AAAI Press, Palo Alto, CA, 548–556.