Optimal Market Making Strategies Under Synchronised Order Arrivals with Deep Neural Network

So Eun Choi[*], Hyun Jin Jang[†‡], Kyungsub Lee[§], Harry Zheng[¶]

March 11, 2020

Abstract

This study investigates the optimal execution strategy of market making for the market and limit order arrival dynamics under a novel framework that includes a synchronised factor between buy and sell order arrivals. Through statistical tests we empirically confirm that the synchrony propensity appears in the market, wherein a buy order arrival tends to follow the sell order long-term mean level and vice versa. This is presumably closely related to a drastic increase in the influence of high frequency trading activities in the markets. For solving a high dimensional Hamilton–Jacobi–Bellman equation, we propose a neural network approximation with theoretical verification of the existence of a network structure that guarantees a small enough the loss function. Finally, we implement the terminal profit and loss profile of market making with the estimated optimal strategy and compare its performance distribution with that of other ordinary strategies. We find that our estimation of the optimal market making placement facilitates significantly stable and steady profit accumulation over time through strict inventory management.

keywords: Market making; Optimal strategy; Order arrival models; Synchrony; High dimensional Hamilton–Jacobi–Bellman; Deep neural network

JEL: G13, C63, C22

[*]Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea

[†]School of Business Administration, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea, E-mail: janghj@unist.ac.kr

[‡]Corresponding author

[§]Department of Statistics, Yeungnam University, Gyeongsan, Republic of Korea

[¶]Department of Mathematics, Imperial College, London SW7 2AZ, United Kingdom

1 Introduction

The emergence of innovative technologies has accelerated the paradigm shift in trading activities in financial markets. In particular, automated trading based on the ultra-low latency electronic system enables facilitation of order generation, routing, and execution, all within a fraction of a second. It uses computerized algorithmic trading called high frequency trading (HFT) that has grown dramatically over the past decade and now accounts for nearly half of all trades in the US and European equity markets – the estimates are 55% and 40%, respectively – while rapidly growing in the Asian market for a variety of asset classes (Miller and Shorter, 2016).

In general, very little is publicly known about the strategies employed in HFT. According to the US Securities and Exchange Commission (SEC, 2010, 2014) concept releases, these strategies can be categorised into two groups depending on the risk appetite. The first is a passive strategy and includes market making and arbitrage trading, which usually do not depend on the direction of price movements. Market making trades mainly provides liquidity to the marketplace, exploiting both bid and ask orders to generate profit from the bid and ask spread. Arbitrage trading is generally pursued to produce profit from price disparity among related securities such as exchange-traded fund and a basket of underlying stocks. Meanwhile, the second is an aggressive strategy that involves momentum ignition and order anticipation strategies, utilizing price direction along with either long or short position. A momentum ignition strategy attempts to trigger sharp price movements – either up or down – by initiating a series of orders. An order anticipation strategy seeks to identify large institutional orders and then trade ahead of those orders in anticipation that they will move market prices. Although these strategies are not new, the current advanced technology may empower traders to better identify profit opportunities and execute their strategies more effectively than in the past.

It is certain that the aforementioned strategies can lead financial markets to become more synchronised and substantially to increase correlation in the price structure. This is because HFT is more likely to occur by tracking price movement patterns rather than changes in market fundamentals. In addition, orders tend to be submitted as a pair of long and short sides (i.e. round-trip trades) and are also executed subsequently and repeatedly on one side – either buy or sell. Gerig (2015) proposes a single-period model of synchrony in financial markets caused by HFT by finding similarities with behaviours of animal groups, such as schooling fish or herding birds. Under market circumstances, it is worthwhile to discuss how the model can capture the highly dependent structure of order flow arrivals attributed to HFT activities. In addition, in terms of market makers that may be designated by a firm or may pursue a market making HFT strategy, the problem of optimal placement of bid and ask spread is a crucial issue.

Despite the market structure changing toward hyper-correlation regime, most studies

modelling high frequency dynamics still employ the exciting factor-based Hawkes models. However, the phenomenon of synchronisation in order flow arrivals might not be temporary, which means that it only can disappear if no other subsequent orders are placed. This is because the likelihood of orders to be synchronised still exists due to HFT activities. In this context, we propose a model for market order arrivals and limit order book dynamics based on the Hawkes process with adding a novel factor, referred to as the *synchronising factor*. This factor makes contemporaneous interactions between buy and sell orders feasible and also enables the interacting effect to remain permanent in the long-term mean of order flows. We then consider a decision making problem for a market maker under a situation where the synchronised market and limit orders enter a trading book. We derive the market maker's optimal trading strategy by maximising his/her profit and liquidating the inventory over a finite time period.

Our order dynamic model is established on the Hawkes processes (Hawkes, 1971a; Hawkes, 1971b; Hawkes and Oakes, 1974) which have been employed as a tool for modelling price movements in high frequency dynamics due to their great flexibility and versatility. As a pioneering work, Bowsher (2007) introduces a bivariate Hawkes process to model the joint dynamics of trades and mid-price changes on the New York Stock Exchange (NYSE) stocks. Large (2007) formulises the resiliency of limit order books based on a ten-variate Hawkes process by testing stocks on the London Stock Exchange. In addition, many studies have employed Hawkes processes in high-frequency finance (Bacry et al., 2012; Bacry et al., 2013; Da Fonseca and Zaatour, 2014; Aït-Sahalia et al., 2015; Bacry et al., 2015; Lee and Seo, 2017). The key feature of Hawkes-based models is the inclusion of exciting factors in their intensity processes. It means that a counting process is more likely to increase when the counting event arrives because its intensity can instantly jump up depending on the movement of the original process. Such a design enables the capture of a clustering phenomenon on arrivals in the counting process throughout a feedback kernel of its intensity which communicates with the counting arrivals.

In the proposed Hawkes' intensity process, the exciting and synchronising components appear to play similar roles, which lead to a strong correlation over time. Nevertheless, from the perspective of the mechanism causing the abnormal impacts, synchronisation is differentiated from excitation. Exciting events are activated by certain external stimulations such as a status change in the original and other relevant processes, and hence, they accelerate arrivals of subsequent orders instantaneously. By contrast, the synchronising factor contributes to enhancing the likelihood of integrating two processes irrespective of any exogenous events. In other words, in terms of the exciting factor, intensity processes are assumed to be independent unless the underlying process changes because order arrivals lead to a temporary increase in intensity. Meanwhile, the synchronising factor forces one process to lead the other and vice versa, implying that the two processes tend to be associated endogenously over a long-term beyond market fundamentals.

This study makes three primary contributions. First, we examine the evidence of the existence of the synchronising factor from market order data. For the six representative stocks in the pool of the US large caps – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft, we discover a remarkable increase in the amount of synchronising tendency between market buy and sell order arrivals in 2018 than in 2008. The increase over the past decade differs across stocks, with the highest change evident in JP Morgan stock. It implies that the frequencies of market buy and sell orders in a unit time tend to interact more dynamically with each other than in the past.

Second, we exploit the deep neural network (DNN) technique to solve the optimisation problem derived as a high dimensional partial differential equation (PDE) with discontinuous components, which is generally known to be difficult to solve analytically. In order to solve the PDE numerically, we propose the DNN-based approximation inspired by the deep Galerkin method (Sirignano and Spiliopoulos, 2018) which is mesh-free simulation suitable for application to the high dimensional PDE problems. We verify the existence of a DNN structure that guarantees the loss function – defined with the boundary and terminal conditions of a given PDE – to be small enough on a compact domain. The DNN algorithm enables us to construct an approximate solution with a very low numerical error; which is trained to satisfy that the loss function is minimised for the generated random samples over the domain of PDE.

Third, for the wealth management in high frequency market making, we show the importance of reflecting the synchrony effect in market and limit order dynamics through computational simulations. To do so, under the scenarios that order dynamics are synchronised, we compare the performances between posting optimal strategies with the synchronising factor and without it. The result shows that the strategy considering the syncronising effect produces significantly higher gains than the case not considering. It implies that if the synchronising tendency appears in the real market between buy and sell order arrivals, market makers need to include the synchronising factor in the order book dynamic model in order to obtain the 'fully' optimal trading quotes and to achieve a best practice.

The remainder of the paper is organised as follows. Section 2 reviews the relevant literature, while Section 3 develops the model of market and limit orders dynamics including the synchronising factor and builds an optimal execution problem for market makers with derivation of the associated PDE. Section 4 presents empirical evidence of the existence of synchrony in market order data for the representative US equities from 2008 to 2018, along with the estimate results under the proposed model. Section 5 presents a DNN to estimate the PDE, and Section 6 presents the simulation results from the efficiency and accuracy of the chosen architecture of the DNN to determine optimal strategies for market makers. Finally, Section 7 concludes. Technical proofs and additional figures/tables are presented in the Appendices A and B, respectively.

2 Literature Review

This study is related to prior research which applies DNNs to solve PDEs and optimisation problems for market makers under highly correlated price dynamics. This section reviews the literature on the application of machine learning to finance, optimal placement strategies for market making, and models of market price structure with hyper dependency.

The application of machine learning to quantitative finance areas has rapidly expanded in recent years and has been examined across a wide range of research. Studies have examined the problems of pricing and hedging derivatives (Hutchinson et al., 1994; Gramacy and Ludkovski, 2015;De Spiegeleer et al., 2018), financial market prediction (Sirignano, 2019l; Dixon et al., 2019) and credit risk management (Baesens et al., 2003; Fitzpatrick and Mues, 2016; Loterman et al., 2012; Khandani et al., 2010). Among a variety of machine learning schemes, DNNs employ a multilayer structure in the neural network framework to learn more about complex nonlinear relationships. In particular, DNNs are well-suited for modelling high dimensional nonlinear problems, which facilitate estimation of arbitrarily continuous functions on compact sets (Hornik et al., 1989; Hornik, 1991). As its seminal example, DNNs are utilised as a numerical scheme for high dimensional PDEs to overcome the curse of dimensionality problem. They enable the consideration of stochastic optimal control and exotic option pricing beyond the Black-Scholes setting which can be transformed to a problem of solving nonlinear PDEs (E et al., 2017; Han et al., 2018; Sirignano and Spiliopoulos, 2018; Fujii et al., 2019).

The optimal decision of a market maker under order dynamics has been widely studied in the market microstructure field. Ho and Stoll (1981) initiate a discussion on an optimal market making policy by specifying a true price for assets on the supply and demand curves of a public market. They derive optimal bid-ask quotes around the true price accounting for the inventory effect. In this spirit, Avellaneda and Stoikov (2008) propose a market making model in an order book by employing a diffusion process for a mid-price and Poisson process for executed limit orders. For the exponential utility function, it provides an asymptotic solution for quoting spread and reservation price. Rosu (2009) derives an equilibrium transaction price between market and limit orders in terms of utility considering a trade-off between execution prices and waiting costs at the bounded discrete price levels. Gueant et al. (2013) study the same problem as in Avellaneda and Stoikov (2008)'s work by adding the inventory volume constraints, and then approximate the optimal control with asymptotic limits in the infinite time horizon. Cartea et al. (2014) model the dynamics of the arrival of market orders and the resulting changes in the limit order book's shape with self-exciting and mutually exciting Hawkes processes. Guilbaud and Pham (2013) investigate an optimal market making policy when an agent uses both limit and market orders by providing a numerical method that estimates the optimal problem. Cartea and Jaimungal (2013) study modelling price revision and duration for HTF activities using the hidden

Markov model with regime switching. Veraart (2010) models two types of market making actions in foreign exchange markets, removing, and adding liquidity with two-dimensional Brownian motions. Guo et al. (2017) employ a correlated random walk for the best bid/ask prices and solve the optimal placement problem with a reflection principle.

Hyper dependence in price dynamics has been modelled in the context of cointegration in econometrics or flocking behaviour. Cointegration occurs when two or more non-stationary time series are driven by one or more common non-stationary time series, as proposed in seminal works by Granger (1981) and Engle and Granger (1987). Many financial data series exhibit cointegration, for example, international stock markets (Cerchi and Havenner, 1988; Taylor and Tonks, 1989; Duan and Pliska, 2004), foreign exchange rates (Baillie and Bollerslev, 1989; Kellard et al., 2010), and futures and spot prices (Ng and Pirrong, 1994, 1996; Maslyuka and Smyth, 2009). Similarly, flocking is referred to as a collective motion of a large number of self-propelled entities. Reynolds (1987) proposes the break-through algorithm that makes generating realistic computer simulations of flocking agents feasible. Flocking behaviour is observed in many contexts in physics, biology, engineering, and human systems, including financial markets (e.g. Rauch et al., 1995; Huepe and Aldana, 2008; Ha et al., 2015; Fang et al., 2017). In a similar manner, Gerig (2015) employs the term 'synchrony' in financial markets, which implies that market prices are monitored by several traders who quickly disseminate information to each other, thus possibly resulting in price comovement. Such a phenomenon resembles animal group behaviours, wherein individuals communicate with each other to share information through contemporaneously scanning the environment using many 'eyes'.

3 Market Models and Optimization Problems

In Section 3.1, we introduce in detail the model of arrivals for buy and sell orders that interact with each other in market and limit orders. We then pose an optimal execution problem for a market maker who seeks to maximise his/her wealth from his/her round-trip trades till the end of the day by penalising inventories, and derive the Hamilton–Jacobi–Bellman (HJB) equation in Section 3.2.

3.1 Model setup

We first define a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions. It is assumed that all stochastic processes introduced in this paper are defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Let S_t be the mid-price of the asset at time t with the dynamics

$$\mathrm{d}S_t = \sigma S_t \mathrm{d}W_t,$$

where σ is a positive constant and W is a standard Brownian motion.

Consider a market maker who continuously posts a limit buy order and sell order of the asset with depth $\delta_t^-, \delta_t^+ \geq 0$, respectively. In other words, the market maker posts a buy limit order at a price of $S_t - \delta_t^-$, and a sell limit order at a price of $S_t + \delta_t^+$. The market maker provides liquidity to the market and earns profits from the bid-ask spread.

We assume that transactions only occur when market orders arrive and match with pending limit orders posted by the market maker. Let the counting processes M_t^+ and M_t^- with intensities λ_t^+ and λ_t^- denote the arrival of other participants' buy and sell market orders, respectively. We denote the market maker's filled buy and sell limit orders by the counting processes N_t^- and N_t^+, respectively. As a measure of the chance with which the market maker's limit buy and sell orders are executed, we consider the fill probabilities $h(\delta_t^\pm, c_t^\pm)$ at time t for limit orders placed $\delta_t^\pm$ away from S_t. Since the market buy orders will lift the market maker's sell limit orders and market sell orders will hit the market maker's buy limit orders, the probability that limit orders are filled increases when the distance from the mid-price δ decreases, and vice versa. The process c_t can be interpreted as parameters directly determining the shape of the limit order book.

From this setup, the processes $N_t^\pm$ can be regarded as the pathwise stochastic integral with respect to $M_t^\pm$:

$$N_t^- = \int_0^t I_s^- \mathrm{d}M_s^- \text{ and } N_t^+ = \int_0^t I_s^+ \mathrm{d}M_s^+. \tag{1}$$

Here, $I_t^\pm$ are defined by

$$I_t^- = \begin{cases} 0 & \text{if } M_t^- - M_{t-}^- = 0 \\ \epsilon_t^- & \text{otherwise} \end{cases} \text{ and } I_t^+ = \begin{cases} 0 & \text{if } M_t^+ - M_{t-}^+ = 0 \\ \epsilon_t^+ & \text{otherwise} \end{cases}, \tag{2}$$

where $\epsilon_t^\pm$ are Bernoulli random variables with probability $h(\delta_{t-}^\pm, c_{t-}^\pm)$, respectively. If the expectation of $\lambda_t^\pm$ is bounded for $0 \leq t \leq T$, $M_t^\pm$ have only a finite number of jumps along the time interval $[0, T]$ almost surely. It follows that $I_t^\pm$ are progressively measurable, and therefore $N_t^\pm$ are well defined as the Lebesgue-Stieltjes integrals and also progressively measurable. We address the condition for bounded intensities later in Lemma 1.

In this study, we assume that market order volumes are independent and identically distributed and exponentially distributed, and the shape of the limit order book is flat, as similar to the setup used by Avellaneda and Stoikov (2008) and Cartea et al. (2014). Given that market orders arrive, the probability that a limit order at price level $S_t \pm \delta_t^\pm$ is executed is equal to $\mathrm{e}^{-\delta_t^\pm c_t^\pm}$.

From the flatness assumption of the limit order book, c_t can be regarded as the market depth of the order book. More precisely, c_t^- and c_t^+ represent the depth of the limit buy (-) and sell (+) order books at the price levels. For a small c, it asserts that the volume of accumulating limit orders prior to where the agent's posted is almost zero, which implies

that the posted limit orders are highly likely to be executed by market orders even for a long distance from the mid-price δ. Conversely, in the case of a large c, the posted limit orders should wait until the stacking limit orders are executed even for a short distance δ.

The market maker's execution processes have two driving factors, how often market orders are walking into the limit order book (i.e. λ_t) and how deep is the volume depth for the current limit order book (i.e. c_t). We model λ_t and c_t including comovement features and feedback effects in limit and market order arrival dynamics.

First, we design the intensity processes for market sell/buy orders $\boldsymbol{\lambda}_t = (\lambda_t^-, \lambda_t^+)_{t\geq 0}$ as a process that has interacting, self-exciting, and mutually exciting features, that is,

$$\begin{aligned} \mathrm{d}\lambda_t^- &= \beta(\theta^- - \lambda_t^- + \kappa\lambda_t^+)\mathrm{d}t + \eta\mathrm{d}M_t^- + \nu\mathrm{d}M_t^+, \\ \mathrm{d}\lambda_t^+ &= \beta(\theta^+ - \lambda_t^+ + \kappa\lambda_t^-)\mathrm{d}t + \eta\mathrm{d}M_t^+ + \nu\mathrm{d}M_t^- \end{aligned} \tag{3}$$

where $\beta, \theta^{\pm}$ are strictly positive and η, ν, κ are non-negative coefficients. The market order arrival intensity jumps up immediately after any market order arrival, where the parameters η and ν govern how responsive the self-exciting and mutually-exciting components of the intensity are due to additional market orders, respectively. However, their states revert to the original mean level with speed β since the exciting impacts from market order arrivals are temporary.

Unlike other existing models, we take a stochastic mean-reversion level of $\theta^- + \kappa\lambda_t^+$ for the sell intensity and $\theta^+ + \kappa\lambda_t^-$ for the buy intensity, where $\theta^{\pm}$ are constant mean levels. It indicates that two processes may have a higher endogenous correlation besides exogenous exciting responses due to the increase in strategy executions related to round-trip trading by HFT traders. It means that even when no buy or sell market order arrives, the two processes can retain their interaction tendency in dynamics. From a traditional economics perspective, the effect could be insignificant if there exist transactions only for typical supply and demand purposes. However, it is necessary to consider this effect for modelling the current market microstructure in electronic exchange-traded assets with high liquidity. Thus, the parameter κ represents how the strength of such endogenous interactive tendency is retained in buy and sell orders; hereafter, we call it a *synchronising factor*.

Second, we assume that the depth dynamics of the limit buy/sell order book $\boldsymbol{c}_t = (c_t^-, c_t^+)$ are the processes that are excited by market order arrivals and are coupled throughout the synchronising factor, that is,

$$\begin{aligned} \mathrm{d}c_t^- &= \xi(\alpha - c_t^- + \kappa_c c_t^+)\mathrm{d}t + \eta_c\mathrm{d}M_t^- + \nu_c\mathrm{d}M_t^+, \\ \mathrm{d}c_t^+ &= \xi(\alpha - c_t^+ + \kappa_c c_t^-)\mathrm{d}t + \eta_c\mathrm{d}M_t^+ + \nu_c\mathrm{d}M_t^- \end{aligned} \tag{4}$$

where α, ξ are strictly positive and η_c, ν_c, κ_c are nonnegative constants. It indicates that the depth of the limit order book jumps up when both market buy/sell orders arrive. This is a one-way effect because market orders cause jumps in following limit orders, but jumps

in limit orders do not induce jumps in market order arrivals (Large, 2007). In addition, the depth processes of both limit buy/sell orders interact each other similar to λ_t. More precisely, α is the long-run mean level, κ_c is the synchronising factor, ξ is the mean-reverting speed, and η_c, ν_c are exciting terms by market orders for the depth processes in buy/sell sides of the limit order book.

To examine the conditions that guarantee the intensity processes for market and limit orders be stable, we define the mean future rate, $m_t^{\pm}(u) = \mathbb{E}[\lambda_u^{\pm}|\mathcal{F}_t]$ and $n_t^{\pm}(u) = \mathbb{E}[c_u^{\pm}|\mathcal{F}_t]$, respectively, for $u \geq t$. For the intensity processes $\lambda_t^{\pm}$ ($c_t^{\pm}$) to be stable, $m_t^{\pm}$ ($n_t^{\pm}$) must remain bounded as a function of u for each t. As a stability condition of $\lambda_t^{\pm}$, we obtain $(1-\kappa)\beta > \eta + \nu$. More detailed results and proofs appear in Appendix A.1.

3.2 Optimisation problems for market makers

This section builds up the market maker's optimisation problem based on the proposed order book dynamic model. We consider a market maker's cash process X_t that satisfies

$$\mathrm{d}X_t = (S_t + \delta_{t-}^{+})\mathrm{d}N_t^{+} - (S_t - \delta_{t-}^{-})\mathrm{d}N_t^{-},$$

which accounts for the cash increase when a sell limit order is lifted by a buy market order, and the cash decrease when a buy limit order is hit by a sell market order. Accordingly, a market maker's inventory process q_t is given as

$$\mathrm{d}q_t = \mathrm{d}N_t^{-} - \mathrm{d}N_t^{+}.$$

A market maker seeks the strategy $(\delta_t^{-}, \delta_t^{+})_{0 \leq t \leq T}$ that maximises the cash value at the terminal date T. At time T, the market maker liquidates the terminal inventory q_T using market orders at a price lower than the mid-price to account for liquidity costs as well as the market orders walking the limit order book. The performance of the market maker achieved during $[t, T]$ is given by

$$\Phi_T = X_T + q_T(S_T - \phi q_T) - \psi \int_t^T q_u^2 du, \tag{5}$$

where $\phi \geq 0$ is a cost attributed to liquidity as well as the impact of the market order walking the limit order book, and $\psi \geq 0$ is the running inventory penalty parameter.

The value function of the market maker is given by

$$V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) = \max_{(\delta_u^{+}, \delta_u^{-})_{t \leq u \leq T}} \mathbb{E}\left(\Phi_T \middle| X_t = x, S_t = s, q_t = q, \boldsymbol{\lambda}_t = \boldsymbol{\lambda}, \boldsymbol{c}_t = \boldsymbol{c}\right), \tag{6}$$

where ϕ, $\psi \geq 0$, and the initial states of the cash amount x, the stock price s, the inventory amount q, and the intensity levels $\boldsymbol{\lambda} = (\lambda^{-}, \lambda^{+})$, $\boldsymbol{c} = (c^{-}, c^{+})$ are given.

The control problem in Eq.(6) can be employed to use the dynamic programming principle to show that the function V solves the following HJB equation

$$\begin{aligned}
&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + \beta(\theta^- - \lambda^- + \kappa\lambda^+)\frac{\partial V}{\partial \lambda^-} + \beta(\theta^+ - \lambda^+ + \kappa\lambda^-)\frac{\partial V}{\partial \lambda^+} \\
&+ \xi(\alpha - c^- + \kappa_c c^+)\frac{\partial V}{\partial c^-} + \xi(\alpha - c^+ + \kappa_c c^-)\frac{\partial V}{\partial c^+} \\
&+ \lambda^- \max_{\delta^- \geq 0}\{\mathrm{e}^{-\delta^- c^-}(\Delta^-_{x,q,\lambda,c}V - V) + (1 - \mathrm{e}^{-\delta^- c^-})(\Delta^-_{\lambda,c}V - V)\} \\
&+ \lambda^+ \max_{\delta^+ \geq 0}\{\mathrm{e}^{-\delta^+ c^+}(\Delta^-_{x,q,\lambda,c}V - V) + (1 - \mathrm{e}^{-\delta^+ c^+})(\Delta^+_{\lambda,c}V - V)\} - \psi q^2 = 0
\end{aligned} \tag{7}$$

with the terminal condition

$$V(T, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) = x + q(s - \phi q), \tag{8}$$

where the shift operators $\Delta^{\pm}_{x,q,\lambda,c}$ and $\Delta^{\pm}_{\lambda,c}$ are defined as

$$\begin{aligned}
\Delta^-_{x,q,\lambda,c}V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= V(t, x - s + \delta^-, s, q + 1, \boldsymbol{\lambda} + (\eta, \nu), \boldsymbol{c} + (\eta_c, \nu_c)), \\
\Delta^+_{x,q,\lambda,c}V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= V(t, x + s + \delta^+, s, q - 1, \boldsymbol{\lambda} + (\nu, \eta), \boldsymbol{c} + (\nu_c, \eta_c)), \\
\Delta^-_{\lambda,c}V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= V(t, x, s, q, \boldsymbol{\lambda} + (\eta, \nu), \boldsymbol{c} + (\eta_c, \nu_c)), \\
\Delta^+_{\lambda,c}V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= V(t, x, s, q, \boldsymbol{\lambda} + (\nu, \eta), \boldsymbol{c} + (\nu_c, \eta_c)),
\end{aligned}$$

with the given controls δ^- and δ^+.

To obtain the simpler form of the HJB, we consider the following ansatz solution. Since the initial cash and initial stock price do not affect the market maker's strategy, we propose

$$V(t, x, s, q, \boldsymbol{\lambda}, \boldsymbol{c}) = x + qs + g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}), \tag{9}$$

which is inspired by observing the terminal condition Eq.(8) for V.

From the substitutions, Eq.(7) is simplified to

$$\begin{aligned}
&\frac{\partial g}{\partial t} + \beta(\theta^- - \lambda^- + \kappa\lambda^+)\frac{\partial g}{\partial \lambda^-} + \beta(\theta^+ - \lambda^+ + \kappa\lambda^-)\frac{\partial g}{\partial \lambda^+} \\
&\quad + \xi(\alpha - c^- + \kappa_c c^+)\frac{\partial g}{\partial c^-} + \xi(\alpha - c^+ + \kappa_c c^-)\frac{\partial g}{\partial c^+} \\
&\quad + \lambda^- \max_{\delta^- \geq 0}\{\mathrm{e}^{-\delta^- c^-}(\delta^- + \Delta^-_{q,\lambda,c}g - g) + (1 - \mathrm{e}^{-\delta^- c^-})(\Delta^-_{\lambda,c}g - g)\} \\
&\quad + \lambda^+ \max_{\delta^+ \geq 0}\{\mathrm{e}^{-\delta^+ c^+}(\delta^+ + \Delta^+_{q,\lambda,c}g - g) + (1 - \mathrm{e}^{-\delta^+ c^+})(\Delta^+_{\lambda,c}g - g)\} - \psi q^2 = 0,
\end{aligned} \tag{10}$$

where the shift operators $\Delta^{\pm}_{q,\lambda,c}$ and $\Delta^{\pm}_{\lambda,c}$ are defined as follows:

$$\begin{aligned}
\Delta^-_{q,\lambda,c}g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= g(t, q + 1, \boldsymbol{\lambda} + (\eta, \nu), \boldsymbol{c} + (\eta_c, \nu_c)), \\
\Delta^+_{q,\lambda,c}g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= g(t, q - 1, \boldsymbol{\lambda} + (\nu, \eta), \boldsymbol{c} + (\nu_c, \eta_c)), \\
\Delta^-_{\lambda,c}g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= g(t, q, \boldsymbol{\lambda} + (\eta, \nu), \boldsymbol{c} + (\eta_c, \nu_c)), \\
\Delta^+_{\lambda,c}g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) &= g(t, q, \boldsymbol{\lambda} + (\nu, \eta), \boldsymbol{c} + (\nu_c, \eta_c)).
\end{aligned}$$

Thus, we explicitly obtain the optimal controls

$$
\begin{aligned}
(\delta_t^-)^* &= \frac{1 - c^-(\Delta^-_{q,\lambda,c} g - \Delta^-_{\lambda,c} g)}{c^-} \mathbb{1}_{\{c^-(\Delta^-_{q,\lambda,c} g - \Delta^-_{\lambda,c} g) < 1\}}, \\
(\delta_t^+)^* &= \frac{1 - c^+(\Delta^+_{q,\lambda,c} g - \Delta^+_{\lambda,c} g)}{c^+} \mathbb{1}_{\{c^+(\Delta^+_{q,\lambda,c} g - \Delta^+_{\lambda,c} g) < 1\}}.
\end{aligned} \tag{11}
$$

By substituting the derived optimal controls to the original Eq.(10), we finally obtain

$$
\begin{aligned}
&\frac{\partial g}{\partial t} + \beta(\theta^- - \lambda^- + \kappa\lambda^+)\frac{\partial g}{\partial \lambda^-} + \beta(\theta^+ - \lambda^+ + \kappa\lambda^-)\frac{\partial g}{\partial \lambda^+} \\
&\quad + \xi(\alpha - c^- + \kappa_c c^+)\frac{\partial g}{\partial c^-} + \xi(\alpha - c^+ + \kappa_c c^-)\frac{\partial g}{\partial c^+} \\
&\quad + \lambda^-(\Delta^-_{q,\lambda,c} g - g)\mathbb{1}_{\{c^-(\Delta^-_{q,\lambda,c} g - \Delta^-_{\lambda,c} g) \geq 1\}} + \lambda^+(\Delta^+_{q,\lambda,c} g - g)\mathbb{1}_{\{c^+(\Delta^+_{q,\lambda,c} g - \Delta^+_{\lambda,c} g) \geq 1\}} \\
&\quad + \lambda^- \left(\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c} g - \Delta^-_{\lambda,c} g)}}{\mathrm{e} c^-} + \Delta^-_{\lambda,c} g - g \right) \mathbb{1}_{\{c^-(\Delta^-_{q,\lambda,c} g - \Delta^-_{\lambda,c} g) < 1\}} \\
&\quad + \lambda^+ \left(\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c} g - \Delta^+_{\lambda,c} g)}}{\mathrm{e} c^+} + \Delta^+_{\lambda,c} g - g \right) \mathbb{1}_{\{c^+(\Delta^+_{q,\lambda,c} g - \Delta^+_{\lambda,c} g) < 1\}} - \psi q^2 = 0
\end{aligned} \tag{12}
$$

with the terminal condition

$$
g(T, q, \boldsymbol{\lambda}, \boldsymbol{c}) = -\phi q^2.
$$

4 Evidences of Synchrony from the Market

Empirical observations can detect the properties of the stochastic process of the order arrival intensities. Figure 1 presents the empirical intensity processes for the Apple stock's market buy and sell orders, which are inferred from the estimated model defined by Eq. (3) on a particular date (3 January, 2018) for a five-minute time elapse (from 14:21 to 14:26). A unit time for the intensities is set to one second. Similar patterns will be observed in cases where the stock, date, and time are chosen randomly. An important feature that can be seen in the paths is that each market order arrival is highly clustered, and the spikes of arrivals tend to disappear very quickly. In addition, both buy- and sell-side order arrivals are not apparently independent with no one-sided order arrivals, and the bursts of activities tend to arrive almost contemporaneously.

To justify the proposed model, this section attempts to capture the existence of synchrony tendency through real market data. Due to the difficulties in accessing limit order book data, the test is conducted based on transacted prices, which may be assumed to be the market order executed by incoming limit orders. To investigate order arrival patterns, we calibrate all the model parameters using a maximum likelihood estimation (MLE)

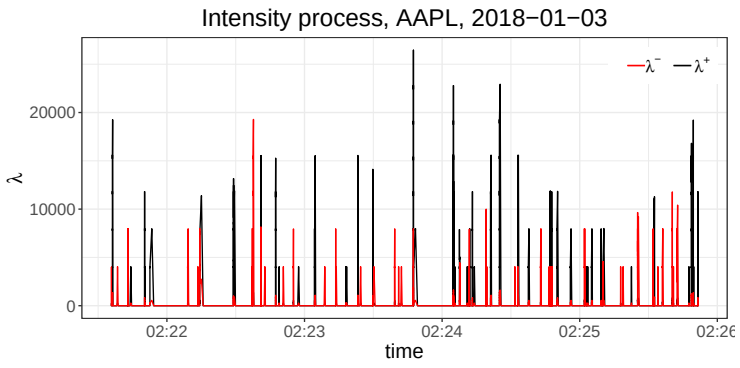

Figure 1: Empirical intensity processes λ^+, λ^- of the Apple stock (NYSE ticker: AAPL) for buy and sell market orders traded from 14:21 to 14:26 on 3 January, 2018

method, and then examine featured evidences especially whether there have been considerable changes in synchronising and exciting factors over time.

To perform MLE for the six parameters in the proposed model $(\beta, \theta^-, \theta^+, \kappa, \nu, \eta)$, we employ the Broyden-Fletcher-Goldfarb-Shanno algorithm (Fletcher, 1987), a kind of quasi-Newton method for numerical optimisation which facilitates adequate estimation of the Hawkes-type process. The numerical procedure for the MLE is implemented using a R optimization function and its wrapper package maxLik[1]. The methodology involves maximising the log-likelihood function

$$L(s) = \int_0^T \log \lambda_t^+ \mathrm{d}M_t^+ + \int_0^T \log \lambda_t^- \mathrm{d}M_t^- - \int_0^T (\lambda_t^+ + \lambda_t^-)\mathrm{d}t, \tag{13}$$

where s is a set of parameters to maximize (Ogata, 1978). The market order arrivals $\lambda_t^\pm$ are given as closed forms in Eqs.(24) and (26) presented in A.2, which are used to implement Eq.(13). The MLE algorithm is processed until a tolerance of 10^{-6} is achieved.

We select six representative stocks among the US large cap equities, namely, Microsoft, Amazon, Apple, JP Morgan, Chevron, and IBM. Our data consist of trades and quotes for NYSE-listed (JP Morgan, Chevron, and Apple) and NASDAQ-listed (Microsoft, Amazon, and IBM) equities[2]. The trade and quote data are recorded in micro or nanosecond decimal precision. The estimations are conducted using the trade prices and their arrival times reported on the exchanges where they are listed. To investigate how the dynamic patterns in market order arrivals have substantially changed in the recent decade, we collect data

[1] `https://cran.r-project.org/web/packages/maxLik/index.html`

[2] The trade and quote information of NYSE stocks and NASDAQ stocks are obtained from Consolidated Tape Association (CTA; `https://www.ctaplan.com/index`) and Unlisted Trading Privileges (UTP; `http://www.utpplan.com/`), respectively.

Table 1: Summary of data statistics for the number of market buy and sell order arrivals in a day.

		2008				2018			
		Mean	Std	Max	Min	Mean	Std	Max	Min
Buy	Microsoft	33,445	19,173	136,025	4,536	20,035	9,767	61,549	6,531
	Amazon	9,563	4,192	31,329	1,500	13,788	7,173	38,780	4,071
	Apple	36,111	15,078	10,3042	7,576	22,587	9,685	69,497	8,210
	JP Morgan	10,501	4,658	35,598	2,199	5,859	2,734	16,520	2,205
	Chevron	5,809	2,105	15,062	1,894	2,542	1,128	11,207	1,027
	IBM	4,466	1,618	12,367	1,289	2,295	1,294	8,729	762
Sell	Microsoft	33,702	18,763	137,227	5,346	21,159	10,690	68,887	6,558
	Amazon	9,440	4,005	28,434	1,317	14,777	8,082	50,443	4,478
	Apple	35,635	14,744	10,0928	8.082	24,382	10,200	72,341	9,778
	JP Morgan	10,773	4,581	35,329	2,202	6,523	2,994	17,922	2,599
	Chevron	6,173	2,179	11,975	1,849	2,659	1,073	9,074	1,114
	IBM	4,708	1,622	10,976	1,137	2,429	1,424	11,654	957

from 2008 to 2018. On each day, trade data between 10:30 and 15:30 are analysed, excluding the trades in the beginning and ending times to avoid seasonal effects.

For data preprocessing to classify transacted prices for buy-initiated and sell-initiated orders, we employ the Lee-Ready algorithm (Lee and Ready, 1991). Using the best bid and ask price flows from the quote data, the mid-price processes are derived and the trade type – whether the transaction is seller-initiated or buyer-initiated – is determined by the relative position of the trade price compared with the mid-price. More precisely, when the trade price is less (greater) than the mid-price, which means that the transaction occurs at the bid (ask) side, then the market order is considered a sell (buy) market order. When the mid and trade prices are equal, the tick test is applied: if the traded price reduces below the most recent traded price that is different from the current trade price, it is considered a sell market order. If it increases, the transaction is regarded as a buy market order.

For the two end-years 2008 and 2018, Table 1 presents descriptive statistics for the number of arrivals of market buy/sell orders in a day after following the preprocessing procedure. For all the stocks, the buy/sell order arrival frequency was 200,326 in 2008; and 139,039 in 2018; that is, the absolute size of orders in 2008 was larger than that in 2018 increase by 140%. One feature of the order flows is that the difference between buy and sell orders[3] widened in 2018 compared with 2008: for all the stocks, the difference was 536 in 2008, whereas it was 4,823 in 2018.

[3]The difference = the number of sell orders - the number of buy orders

Table 2: Results for mean of the estimated κ, the number of significant estimation for κ, and checking the stability condition $(1-\kappa)\beta/(\eta+\nu)$

	2008			2018		
	κ (%)	# significance	stability	κ (%)	# significance	stability
IBM	0.1343	8 (3.2%)	3.2342	7.0297	179 (71.3%)	2.9728
Chevron	0.3359	10 (3.9%)	3.3635	5.9676	163 (64.9%)	1.8656
Apple	0.8712	58 (23.1%)	1.8783	1.7943	146 (58.1%)	1.7780
Amazon	1.6632	76 (30.4%)	1.9250	0.2066	57 (22.8%)	1.6500
JP Morgan	0.0000	0 (0%)	2.6078	8.4821	208 (83.2%)	3.4385
Microsoft	0.0000	0 (0%)	1.1912	0.5315	63 (25.2%)	1.6503
Overall	0.5008	152 (10.1%)	2.3667	4.0020	816 (54.4%)	2.2259

We perform MLE based on data collected on a daily-basis 2 January to 31 December: for example, the estimation is conducted 252 times for the data in 2008 and 250 times in 2018. First, we investigate the result of the synchronising factor κ. Table 2 presents the average of the estimated κ, the number of significant estimates at the 5% significance level among the whole sample in each year, and the average of stability condition $(1-\kappa)\beta/(\eta+\nu)$ derived in Lemma 1 with the estimated samples. The column 'Overall' contains the average of all the stocks for κ, and the stability and sum of all the significant estimated samples. We see that the synchronising level in 2018 jumped up about 8 times more than that in 2008 on average and the number of significant estimations increased 5.4 times during the same period – the total number of significant samples was 152 in 2008, but 816 in 2018 for all the stocks, and the number of order arrivals every single day, on average, is specified in Table 1. Whereas the overall market stability has decreased marginally, the stability of the process can be highly guaranteed because all are greater than one[4].

Next, we illustrate the calibration results of $\theta^+, \theta^-, \eta, \nu, \beta$ in 2008 and 2018 for the basic statistics – mean, standard deviation, skewness, and kurtosis – in Table 5. The column 'Average' represents the average value of all the stock's estimations. All the estimations of the five parameters were statistically significant. We observe that the mean-reversion levels θ^+, θ^- of both buy and sell arrivals dropped by about 38% in 2018 than in 2008. The exciting levels η, ν decreased by 74% and 66%, respectively, and the reverting speed β also declined by 90% on average over the same time period. It is worthy to note that the self-exciting factor η has a remarkably greater value than the mutually-exciting factor ν across all the stock samples. We provide more detailed calibration results of different

[4]Using this condition without the synchronising parameter, the literature has discussed the measurement of the degree of market instability and resiliency (e.g.,Danielsson et al., 2012; Filimonov and Sornette, 2012; Hardiman et al., 2013).

aspects in Appendix B.

Such a consistent tendency observed in the market in terms of hyper dependency is presumably closely related to the dramatic increase in HFT activities over the past decade. Even though the total number of arrivals in 2018 decreased by 31% than in 2008, the synchrony tendency in buy and sell orders increased by 8 times over the same period. The empirical results above may support the argument that the arrival intensities for buy and sell orders are dynamically incorporating the synchrony impact as well as the exciting tendency. Thus, we need a better explainable model to develop more profitable trading strategies.

5 Neural Network Approximation for the HJB Equation

In this section, we build up an approximation method to find a solution to the high-dimensional HJB equation using DNN. To do so, we adopt a mesh-free scheme called the deep Galerkin method proposed by Sirignano and Spiliopoulos (2018). Even though mesh-based methods, such as a finite difference method, are known to the most common way of solving PDEs, they are computationally infeasible for high dimensional cases. However, the deep Galerkin method trains batches of randomly sampled time and space points through neural networks instead of forming a mesh, making it flexible and versatile, as required for high dimensional PDE problems. The deep Galerkin method takes input values from the domain of the PDE and produces a candidate of the PDE solution by composing a smooth activation function repeatedly. The extent to composing represents the depth of the network and the composed function can solve the PDE numerically through tuning the DNN parameters. The training is a procedure of finding the best DNN parameters that minimise a loss function, which indicates how close the DNN architecture is to satisfying the PDE's given conditions.

We first consider a domain of the PDE such that $\mathcal{D} = (-\infty, \infty) \times [0, \infty)^4 \subset \mathbb{R}^5$ and $\mathcal{D}_T = [0, T) \times \mathcal{D}$. We assume that there exists a unique solution $u(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) \in C^1(\mathcal{D}_T)$ to the following PDE:

$$\begin{aligned} &\frac{\partial u}{\partial t}(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) + \mathcal{L}u(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) = 0, \text{ for } (t, q, \boldsymbol{\lambda}, \boldsymbol{c}) \in \mathcal{D}_T \\ &u(T, q, \boldsymbol{\lambda}, \boldsymbol{c}) = -\phi q^2, \text{ for } (q, \boldsymbol{\lambda}, \boldsymbol{c}) \in \mathcal{D}, \end{aligned} \tag{14}$$

where $\mathcal{L}$ is a nonlinear PDE operator with one time variable and five state variables defined in Eq.(27). If the PDE (12) has a unique solution $g(t, q, \boldsymbol{\lambda}, \boldsymbol{c})$, then

$$u(t, q, \boldsymbol{\lambda}, \boldsymbol{c}) = g(t, q, \boldsymbol{\lambda}, \boldsymbol{c}), \quad \text{for } (t, q, \boldsymbol{\lambda}, \boldsymbol{c}) \in \mathcal{D}_T.$$

According to deep Galerkin method, u can be approximated with a DNN structure defined by $f(\cdot, \Theta)$, where Θ is a set of the DNN parameters. Among the various architectures of f, we employ a fully-connected feedforward network, which is the most basic type of neural network but has an outstanding performance in finding our PDE solution.

Now we consider a reduced compact domain so that differential operators with respect to all the variables can be well defined to use the universal approximation theorem by Hornik (1991). Let a compact set $\tilde{\mathcal{D}} = [-N_1, N_1] \times [n_2, N_2]^2 \times [n_3, N_3]^2 \subset \mathcal{D}$ for any large positive numbers N_1, N_2, N_3 and small positive numbers n_2, n_3, and let $\tilde{\mathcal{D}}_T = [0, T) \times \tilde{\mathcal{D}}$. On the reduced domain, we construct the loss function

$$L(f) = \left\| \frac{\partial f}{\partial t}(t, q, \boldsymbol{\lambda}, \boldsymbol{c}; \Theta) + \mathcal{L} f(t, q, \boldsymbol{\lambda}, \boldsymbol{c}; \Theta) \right\|^2_{\tilde{\mathcal{D}}_T, \mu_1} + \left\| f(T, q, \boldsymbol{\lambda}, \boldsymbol{c}; \Theta) + \phi q^2 \right\|^2_{\tilde{\mathcal{D}}, \mu_2}, \quad (15)$$

where μ_1 and μ_2 are probability measures on $\tilde{\mathcal{D}}_T$ and $\tilde{\mathcal{D}}$, respectively, which are absolutely continuous with respect to Lebesgue measure. The loss function represents the error of how far f deviates from the original PDE's operator and terminal conditions. We verify that there exists a feedforward neural network f that makes the loss function Eq.(15) small enough on the compact set $\tilde{\mathcal{D}}_T$ (see Theorem 3, Appendix A.3).

The goal is to find the DNN parameters Θ such that the error in Eq.(15) is minimized. To do so, we develop Algorithm 1 which states the DNN algorithm using the stochastic gradient decent method on a sequence of time and state points that are randomly sampled on $\tilde{\mathcal{D}}_T$ and $\tilde{\mathcal{D}}$ with respect to μ_1 and μ_2, respectively. In Step 2 the mean squared error $\tilde{L}(f; \Theta)$ is an unbiased estimator of the loss function $L(f)$. The algorithm proceeds in a descending direction, which means that the loss function decreases after an iteration and the next Θ can be a better parameter estimate than the previous one.

For the implementation of Algorithm 1, we employ the hyperbolic tangent function as an activation function so as to enable f to become smooth and to apply the differential operator to the neural network. We find the following DNN architecture to be effective:

$$\begin{aligned} h_1 &= \tanh(W_0 \boldsymbol{x} + b_0) \\ h_{l+1} &= \tanh(W_l h_l + b_l), \qquad l = 1, \cdots, L \\ f(\boldsymbol{x}; \Theta) &= W_{L+1} h_{L+1} + b_{L+1}, \end{aligned} \quad (16)$$

with $h_l \in \mathbb{R}^n$ and L hidden layers having n hidden units in each hidden layer. The function $\tanh(\cdot)$ is taken to each component of h_l. This network takes the input $\boldsymbol{x} = [t\ q\ \boldsymbol{\lambda}\ \boldsymbol{c}]^T$ and maps this input to the value of the function $f(\boldsymbol{x}; \Theta)$ a candidate for the solution of the PDE. The parameter Θ consists of weight matrices $W_0 \in \mathbb{R}^{n \times 6}$, $W_l \in \mathbb{R}^{n \times n}, W_{L+1} \in \mathbb{R}^{1 \times n}$ and bias vectors $b_0 \in \mathbb{R}^n$, $b_l \in \mathbb{R}^n$, $b_{L+1} \in \mathbb{R}^1$.

For the key hyperparameters L and n in the DNN, we find that three-hidden-layer (i.e. $L = 3$) and 900 hidden nodes (i.e. $n = 900$) are effective and provide a good performance

Algorithm 1 The DNN algorithm for solving PDE (14).

1: Generate random samples $\{(t_i, q_i, \boldsymbol{\lambda}_i, \boldsymbol{c}_i)\}_{i=1}^m$ from probability measure μ_1 in $\tilde{\mathcal{D}}_T$ and $\{(\tilde{q}_i, \tilde{\boldsymbol{\lambda}}_i, \tilde{\boldsymbol{c}}_i)\}_{i=1}^m$ from probability measure μ_2 in $\tilde{\mathcal{D}}$ with the batch size m.

2: Calculate the mean squared error $\tilde{L}(f;\Theta)$ as

$$\tilde{L}(f;\Theta) = \frac{1}{m}\sum_{i=1}^{m}\Bigg\{ \left(\frac{\partial f}{\partial t}(t_i, q_i, \boldsymbol{\lambda}_i, \boldsymbol{c}_i;\Theta) + \mathcal{L}f(t_i, q_i, \boldsymbol{\lambda}_i, \boldsymbol{c}_i;\Theta)\right)^2 + \left(f(T, \tilde{q}_i, \tilde{\boldsymbol{\lambda}}_i, \tilde{\boldsymbol{c}}_i;\Theta) + \phi\tilde{q}_i^2\right)^2\Bigg\}$$

3: Update the parameters Θ in the opposite direction of the gradient of $\tilde{L}(f;\Theta)$ with regard to the parameters:

$$\Theta \leftarrow \Theta - \ell\,\nabla_\Theta \tilde{L}(f;\Theta)$$

with learning rate ℓ.

4: Repeat until the gradient approach to zero.

on the given market model parameters. We use a batch size of 25,000 (i.e. $m = 25,000$) drawn from the uniform probability measures μ_1 and μ_2 on $\tilde{\mathcal{D}}$. The initial learning rate ℓ is chosen as 10^{-4} and reduced by a factor of 10 when the loss function stops decreasing. The DNN testing is conducted using TensorFlow in Python, and the DNN parameters are updated using the Adam optimisation algorithm, which is a momentum-based stochastic optimization method (Kingma and Ba, 2014). The DNN parameters are initialised using the Xavier initialisation.

In the parameter domains, the reduced domain $\tilde{\mathcal{D}}_T$ is defined with $T = 10,800$ and $N_1 = 100$, $N_2 = 250$, $N_3 = 150$, $n_2 = 0.5$, $n_3 = 0.3$ to generate random samples as inputs[5]. To ensure the five state variables stay in the reduced domain $\tilde{\mathcal{D}}_T$, we confirm that millions of simulated paths $q_t, \lambda_t^{\pm}$, and $c_t^{\pm}$ are contained in $\tilde{\mathcal{D}}_T$. Furthermore, we choose the parameters of the market model as follows: $\beta = 100$, $\theta^- = 0.5$, $\theta^+ = 0.4$, $\kappa = 0.2$, $\eta = 20$, $\nu = 15$ (for the intensity processes $\boldsymbol{\lambda}$ in Eq.(3)), $\xi = 40$, $\alpha = 0.3$, $\kappa_c = 0.1$, $\eta_c = 8$, $\nu_c = 5$ (for the depth processes $\boldsymbol{c}$ in Eq.(4)), $\phi = 0.1$, $\psi = 0.01$ (penalty on the inventory in Eq.(5)).

Figure 2 presents a contour plot of the optimal quote distances δ^- and δ^+ with regard to the inventory q and time t. The intensities for the market order and limit order depth processes are taken as $\lambda^- = 3.5$, $\lambda^+ = 3$ and $c^- = 1.5$, $c^+ = 2$, respectively. The level of optimal distances from the mid-price is outstandingly prominent when the market maker retains a large amount of inventory that needs to be liquidated or acquired, and little time

[5]The time T represents 10,800 seconds because all the parameters in our model are measured at the unit level of seconds, which means that the trading lasts three hours.

is left to trade the stocks. As the inventory q gets positively larger, the market maker's strategy is to become more reluctant to buy stocks over time, and the strategy is likely to place limit buy order δ^-, which is far away from the mid-price, to reduce the chance of executions.

As q gets negatively larger (the short position increases), the strategy tends to increase the stock by placing the limit buy order almost at the mid-price to enhance the chance of executions. In a similar manner, as q gets negatively larger, the strategy is to become more hesitant to sell stocks by placing limit sell order δ^+, which is far from the mid-price. As q gets positively larger, the strategy tends to reduce the stock position by placing the limit sell order almost at the mid-price.

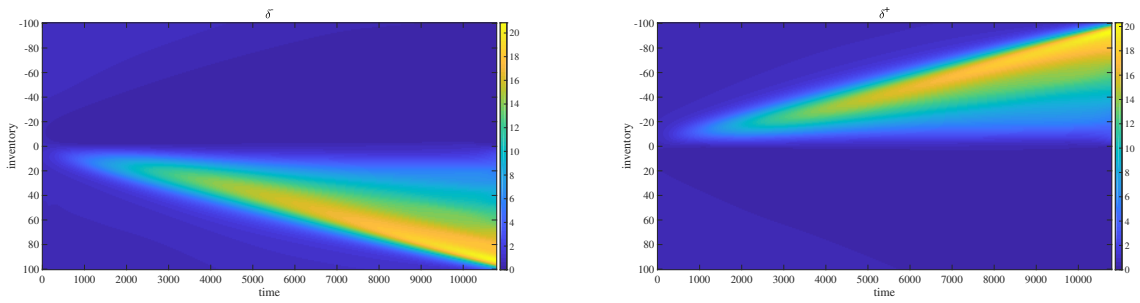

Figure 2: Contour plots of the optimal quotes δ^- for limit buy order (left) and δ^+ for limit sell order (right) depending on the change in inventory q when $\lambda^- = 3.5$, $\lambda^+ = 3$, $c^- = 1.5$, $c^+ = 2$.

Next, we show the estimated solution of the PDE (14) and its numerical error of the DNN algorithm. We compute the DNN-approximated solution f for $0 \leq t \leq T$ with respect to the inventory q with a given ϕ. The other variables $\lambda^\pm$ and $c^\pm$ are randomly sampled with 100,000 points for each and are taken on average for this visualisation. In Figure 3, the left panel illustrates contour lines of the estimated solution of g over the whole period, and the right panel shows only the part of the estiated value of g under the differential operator $\frac{\partial}{\partial t} + \mathcal{L}$, which has a value of zero in the true solution. In Figure 4, the left panel indicates the true solution g (red line) and the DNN-estimated solution (blue line) at the end time T over the q domain, and the right panel shows its error computed as the absolute difference between the two curves that appear in the right panel. As the inventory size either increases or decreases, the estimation error increases, but the absolute error level is not that significant as considering the solution values.

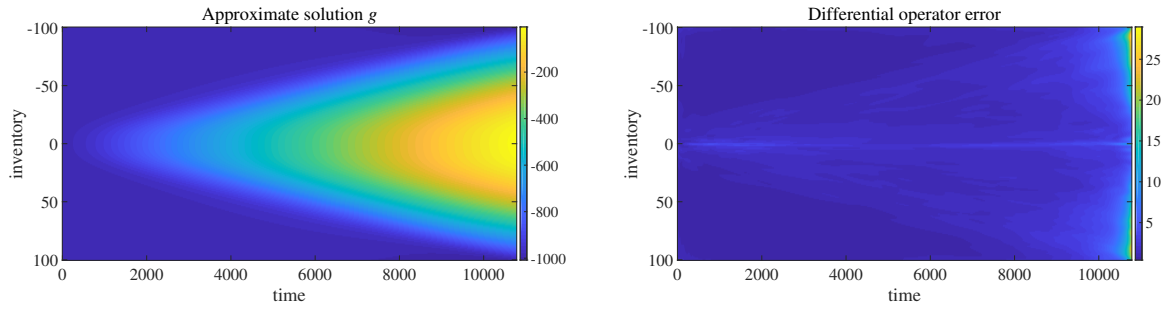

Figure 3: The DNN-approximate solution of g (left) and the DNN-approximate value of differential operator $\frac{\partial}{\partial t}g + \mathcal{L}g$ (right). Each value is obtained by taking average of all the values (g and $\frac{\partial}{\partial t}g + \mathcal{L}g$) computed with 100,000 random samples of $\lambda^{\pm}$ and $c^{\pm}$.

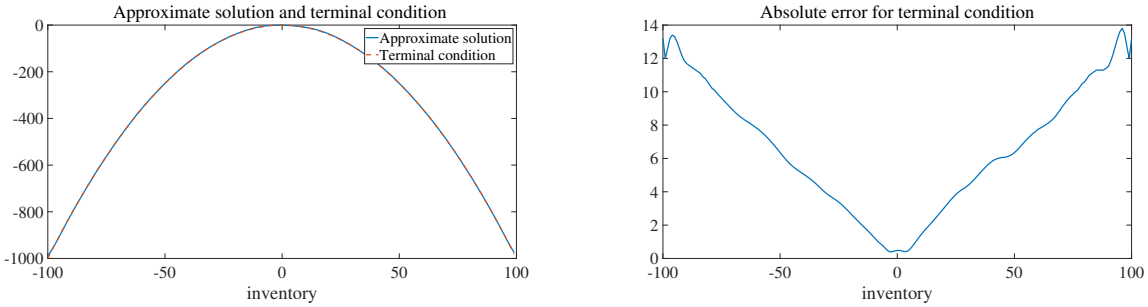

Figure 4: The terminal condition of g (blue line) and its DNN-approximate solution (red-dot) by taking average on the values with respect to $\lambda^{\pm}$ and $c^{\pm}$ with 100,000 random sampling for each (left) and its absolute difference (right).

6 Simulation

This section conducts various numerical simulations of the market maker's optimal execution strategy exploiting the DNN estimation discussed in Section 5. To do so, we compute the distribution of the terminal profit and loss (PnL) of a market maker, that is,

$$\text{PnL} = X_T + q_T(S_T - \phi q_T)$$

with the liquidation cost ϕ under the posting controls $\delta^{\pm}$, by generating required sample paths.

A sample path of the terminal PnL can be generated with the following procedure. First, based on the thinning algorithm (Ogata, 1978), sample paths of the intensity $\lambda_t^{\pm}$ and the market order processes $M_t^{\pm}$ are obtained. Since our model contains additional synchrony factors compared to a typical exciting Hawkes model, we modify the original sampling algorithm to the one specified in Algorithm 2 (presented in Appendix A.4). Accordingly, sample paths of depth of limit order book $c_t^{\pm}$ with $M_t^{\pm}$ are generated. Next, a sample

of $N_t^{\pm}$ is filtered from $M_t^{\pm}$ using a Bernoulli variable generator under the fill probability $h(\delta_t^{\pm}, c_t^{\pm})$ with the estimated (or given) $\delta_t^{\pm}$. Lastly, with a generated stock price path S_t, X_t and q_t are computed. Above procedure is running until the accumulated interarrival time of either N_t^+ or N_t^- is reaching T. The initial wealth starts at zero and we obtain K samples of terminal PnL by repeating the procedure.

Optimal strategy versus simple strategy.
We compute the performance of the optimal strategy and compare it with a rough constant strategy and a more elaborate strategy. The detailed strategies are given as follows: (i) the controls are given as constant, that is, $\delta_t^- = \delta_t^+ = d_0$; (ii) the controls are chosen asymmetrically and are linearly dependent on inventory q_t, that is,

$$\delta_t^-(q_t) = \begin{cases} 2q_t + \frac{1}{2} & \text{if } q_t \geq 0 \\ -\frac{1}{2}q_t + \frac{1}{2} & \text{if } q_t < 0 \end{cases}, \quad \delta_t^+(q_t) = \begin{cases} \frac{1}{2}q_t + \frac{1}{2} & \text{if } q_t \geq 0 \\ -2q_t + \frac{1}{2} & \text{if } q_t < 0 \end{cases}; \tag{17}$$

(iii) the optimal controls δ^* in Eq.(11) are taken.

The strategy (ii) is inspired by the estimation results of the optimal δ^* which has a very asymmetry pattern depending on the current inventory size. Although the appeared pattern is nonlinear, we consider the strategy with a linear but asymmetry relation with respect to the amount of inventory, because it is not required to have computational costs. The reason to choose the factor '2' in this strategy is that it achieved the best performance among other integer values.

For this simulation, we adopt the market parameters based on the calibration results from the market data shown in Section 4, which are also applied to the numerical tests in Section 5. The parameters are chosen as follows: the trading time $T = 10,800$; liquidation cost $\phi = 0.1$; market order intensities are determined under $\beta = 100, \theta^- = 0.5, \theta^+ = 0.4, \kappa = 0.2, \eta = 20, \nu = 15$; and limit order intensities under $\xi = 40, \alpha = 0.3, \kappa_c = 0.1, \eta_c = 8, \nu_c = 5$. With this selection, we generate 50,000 samples of the terminal PnL.

We illustrate the histograms of the terminal wealth PnL achieved by the three strategies of posting $\delta^{\pm}$ in Figures 5 and 6. Accordingly, Table 4 presents the descriptive statistics of the terminal PnLs (i.e. mean, standard deviation, minimum, and maximum) for the constant, asymmetric, and optimal posting strategies. In terms of the expected terminal PnL, posting at the optimal distance performs the best compared with the other cases. In addition, the optimal posting case significantly outperforms the others in all the samples, considering the minimum PnL of the optimal strategy and maximum PnLs of all other constant cases. The PnL's diversity is considerably smaller than that of the others, which implies that the optimal strategy facilitates achieving the highest expected profit with a comparatively lower risk than others.

We next show a whole simulated path of the cash process X_t and the inventory process q_t over time. Figure 7 illustrates the sample paths of X_t and q_t when using the optimal δ_t^*

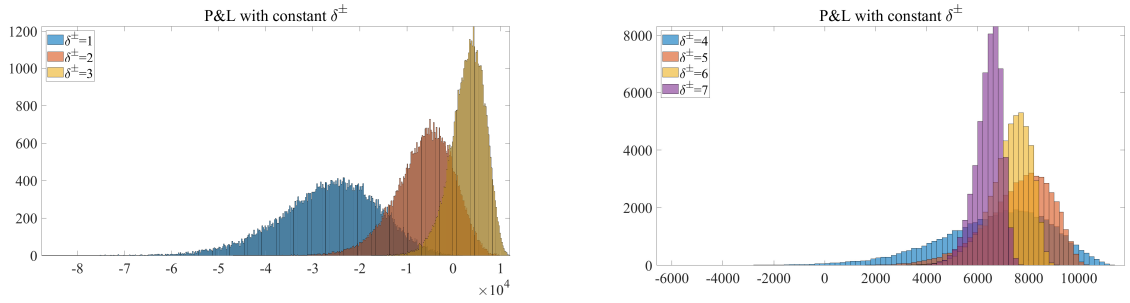

Figure 5: Distribution of the terminal PnL for a market maker with constant $\delta_0 = 1, 2, 3$ (left) and $\delta_0 = 4, 5, 6, 7$ (right) when trading lasts 10,800 seconds.

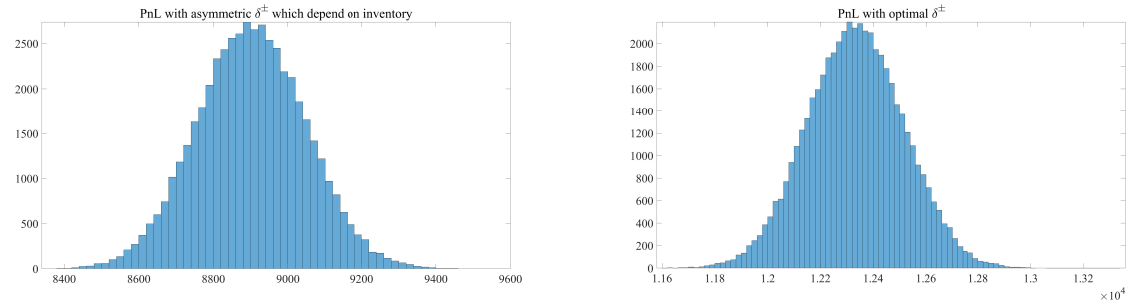

Figure 6: Distribution of the terminal PnL for a market maker with asymmetric δ (left) and with optimal δ^* (right) which depend on inventory when trading lasts 10,800 seconds.

Table 3: Descriptive statistics of the terminal PnLs using three strategies (constant, asymmetric, and optimal) when trading lasts 10,800 seconds.

Market making strategy		Mean	SD	Min	Max
Constant	$d_0 = 1$	-26,701	10,472	-87,524	1,937
	$d_0 = 2$	-6,242	6,174	-46,834	9,746
	$d_0 = 3$	2,969	3,667	-17,546	11,858
	$d_0 = 4$	6,629	2,226	-6,563	11,648
	$d_0 = 5$	7,578	1,346	-861	10,509
	$d_0 = 6$	7,253	832	1,452	9,230
	$d_0 = 7$	6,411	532	2,824	7,786
Asymmetric $\delta^{\pm}$ in Eq.(17)		8,902	**146**	8,351	9,582
Optimal δ^*		**12,333**	183	**11,583**	**13,349**

strategy (left) and a constant strategy with $\delta_0 = 5$ (right). Posting an optimal distance tends to lead a gradual increase in profit over time, while posting a naive constant can result in a cumulative loss. In terms of inventory management, the inventory process under

an optimal execution tends to move within a small range between ± 7 points at the most, whereas the inventory process under the constant execution increases over time and reaches over 150 points on this path.

Such inventory management within tight bounds of optimal posting can be understood as a consequence of the objective function that aims to maximise the final wealth by penalising large exposure to inventories. In practical terms, the capital amount available to market makers is strictly constrained by their internal rules, because firms or regulators penalise activities that take large exposures. Another reason of the necessity of tight inventory management is that the piling (long or short) positions during market making should be liquidated at the end of the trading day by paying transaction costs. Given strict control of exposure to inventories is required for market making, market makers might prefer to adopt a high penalising rate ψ when their risk appetite reduces.

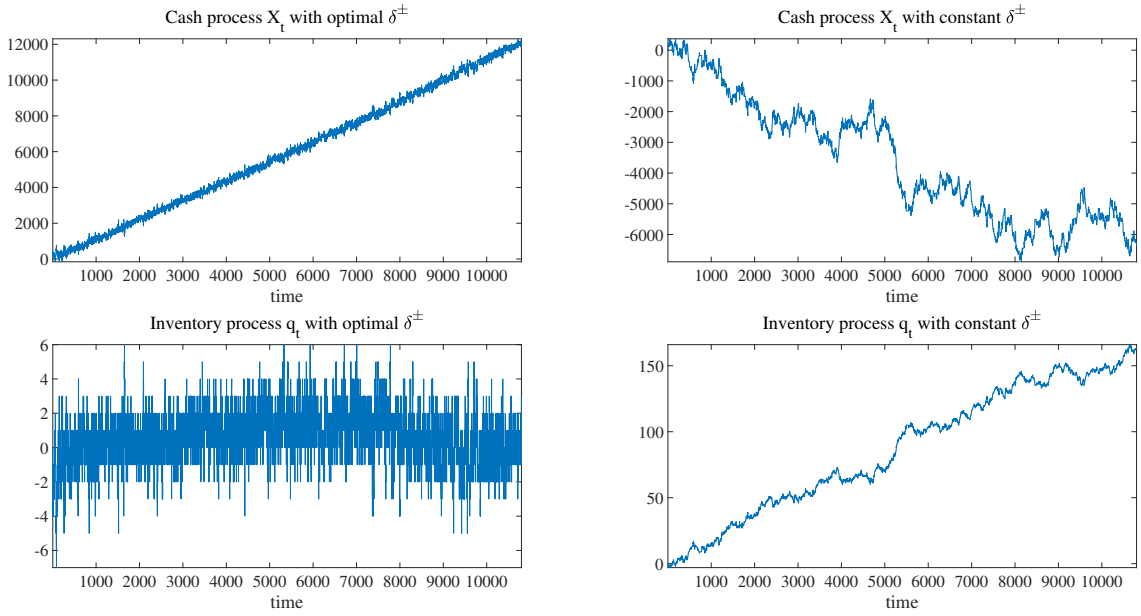

Figure 7: Sample paths of the cash process X_t (upper) and inventory process q_t (lower) for the optimal posting (left) and a constant posting (right) for 10,800 seconds trading. The values of terminal wealth PnL are 12,085 (optimal $\delta^{\pm *}$; upper left) and 7,477 (constant $\delta^{\pm}$; lower right).

Optimal strategies with and without the synchronising factor.
We conduct a test of comparing the terminal PnLs when posting the optimal strategy and posting a strategy without a synchronising factor under the circumstances of existence of synchrony tendency between buy and sell orders. The primary purpose of this test is to

investigate how the suboptimal posting strategy which does not consider the synchronizing factor could be aggravating the performance of market maker's PnLs comparing to the fully optimal strategy considering the synchronising factor. The optimal strategies are estimated with the same architecture of DNN used in Eq.(16). This simulation chooses the parameters[6]: $T = 1,800$; for market orders $\beta = 50, \theta^- = 0.5, \theta^+ = 0.4, \eta = 4, \nu = 0.1$; and for limit orders $\xi = 40, \alpha = 0.3, \eta_c = 3, \nu_c = 0.1$. For the synchronising parameter, we consider two cases that (i) Case 1: synchrony is taken with $\kappa = \kappa_c = 0.9$; and (ii) Case 2: no synchrony, that is, $\kappa = \kappa_c = 0$. Then, we generate 30,000 synchronising samples and apply the strategies estimated under the each case.

Figure 8 presents the terminal PnL distributions under optimal strategies with and without a synchronizing factor when there exists a synchronizing tendency in the market. The expected value of the terminal PnL when the market makers ignore κ and κ_c becomes significantly smaller than that when they does not. The corresponding statistics of PnL are given in Table 4. It is noticed that Case 1 outperforms the other by 40.2% on average, and the worst scenario of Case 1 is better than the best of Case 2. If the synchrony effect decreases (or the mutually-exciting effect increases), the performance differences between two cases can be reducing.

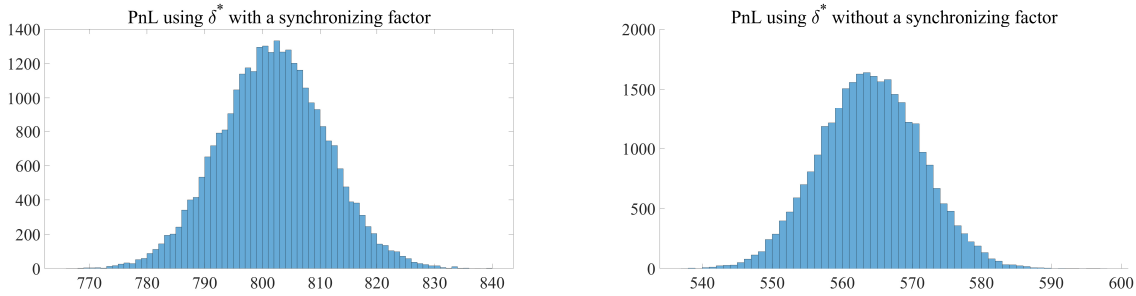

Figure 8: Under the circumstance of which market and limit orders are synchronised with $\kappa = \kappa_c = 0.9$, the distributions of the market maker's terminal PnL using δ^* with and without a synchronizing factor (i.e., $\kappa = \kappa_c = 0.9$ and $\kappa = \kappa_c = 0$, respectively) when trading lasts 1,800 seconds.

This simulation implies that it can be crucial to reflecting the synchronising effect in market and limit order dynamics in terms of wealth management for high-frequency market making. As shown in Section 4 that the synchronising tendency has clearly appeared in the real market between buy and sell order arrivals, market makers have to consider the

[6]To examine the impact of synchrony in simulation clearly, we decrease the mutually-exciting factor and increase the synchronising factors compared to the previous simulations.

Table 4: Descriptive statistics of the terminal PnLs using two optimal strategies of synchrony and no synchrony under the synchronising scenarios for market and limit orders with 1,800 seconds trading time.

Market making strategy δ^*	Mean	SD	Min	Max
Synchrony $\kappa = \kappa_c = 0.9$	801.9	9.2	766.7	839.6
No synchrony $\kappa = \kappa_c = 0$	564.0	7.2	537.6	597.1

synchronising factor in the order book model in order to obtain the 'fully' optimal trading quotes and to achieve the best performance.

7 Conclusion

This study investigates the optimal execution strategy of market making for the market and limit order arrival dynamics under a novel model with a synchronising factor between buy and sell orders. To examine the synchrony tendency in the market, we estimate the relevant model parameters through the MLE algorithm modified for the proposed model. The empirical result confirms that the arrival frequencies of market buy orders have a propensity to follow the sell orders' long-term mean level and vice versa. This is presumably caused by the huge increase in the influence of HFT activities in the markets, and such a phenomenon may continue to appear or become stronger with further innovation in technologies.

For solving the high dimensional HJB equation derived as the optimal decision problem, we propose DNN approximation instead of mesh-based methods. We verify the existence of a DNN structure that guarantees the loss function be small enough on the restricted compact domain. We then estimate the solution of the HJB and the optimal placement strategy within a tolerant numerical error bound. Finally, we simulate the performance of the end profit from market making under a variety of strategies (i.e. naive constant, asymmetric, and optimal placement) and compare their statistical distributions. We find that the estimated optimal strategy outperforms the other cases significantly and facilitates in providing a profit profile with very stable and steady accumulation over time with tight inventory management. In addition, we show the importance of reflecting the synchronising effect in market and limit order dynamics through simulations to achieve the best practice of high-frequency market making.

Acknowledgements

The authors are grateful to anonymous reviewers for providing their constructive comments and suggestions. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (no.2017R1C1B5074398) and Institute for Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2017-0-01779, A machine learning and statistical inference framework for explainable artificial intelligence).

References

Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R.J., 2015. Modeling financial contagion using mutually exciting jump processes. Journal of Financial Economics 117, 585–606. doi:10.1016/j.jfineco.2015.03.002.

Avellaneda, M., Stoikov, S., 2008. High-frequency trading in a limit order book. Quantitative Finance 8, 217–224.

Bacry, E., Dayri, K., Muzy, J.F., 2012. Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. The European Physical Journal B 85, 157.

Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.F., 2013. Modelling microstructure noise with mutually exciting point processes. Quantitative Finance 13, 65–77.

Bacry, E., Mastromatteo, I., Muzy, J.F., 2015. Hawkes processes in finance. Market Microstructure and Liquidity 1, 1550005.

Baesens, B., Setiono, R., Mues, C., Vanthienen, J., 2003. Solving high-dimensional partial differential equations using deep learning. Management Science 46, 312–329.

Baillie, R., Bollerslev, T., 1989. The message in daily exchange rates: A conditional-variance tale. Journal of Business & Economic Statistics 7, 197–305.

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

Cartea, A., Jaimungal, S., 2013. Modelling asset prices for algorithmic and high-frequency trading. Appl Math Fin 20, 512–547.

Cartea, A., Jaimungal, S., Ricci, J., 2014. Buy low, sell high: A high frequency trading perspective. SIAM Journal on Financial Mathematics 60, 415–444.

Cerchi, M., Havenner, A., 1988. Cointegration and stock prices: The random walk and wall street revisited. Journal of Economic Dynamics and Control 12, 333–346.

Da Fonseca, J., Zaatour, R., 2014. Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. Journal of Futures Market 34, 548–579.

Danielsson, J., Shin, H.S., Zigrand, J.P., 2012. Endogenous and systemic risk, in: Quantifying systemic risk. University of Chicago Press, pp. 73–94.

De Spiegeleer, J., Madan, D., Reyners, S., Schoutens, W., 2018. Machine learning for quantitative finance: fast derivatives pricing, hedging and fitting. Quantitative Finance 18, 1635–1643.

Dixon, M., Klabjan, D., Bang, J., 2019. Classification-based financial markets prediction using deep neural networks. Working paper -, –.

Duan, J.C., Pliska, S., 2004. Option valuation with co-integrated asset prices. Journal of Economic Dynamics & Control 28, 727–754.

E, W., Han, J., Jentzen, A., 2017. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. Commun. Math. Stat. 5, 349–380.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society , 251–276.

Fang, F., Sun, Y., Spiliopoulos, K., 2017. On the effect of heterogeneity on flocking behavior and systemic risk. Statistics & Risk Modeling 34, —-.

Filimonov, V., Sornette, D., 2012. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. Physical Review E 85, 56108.

Fitzpatrick, T., Mues, C., 2016. An empirical comparison of classification algorithms for mortgage default prediction: evidence from a distressed mortgage market. European Journal of Operational Research 249, 427–439.

Fletcher, R., 1987. Practical methods of optimization . Wiley 2.

Fujii, M., Takahashi, A., Takahashi, M., 2019. Asymptotic expansion as prior knowledge in deep learning method for high dimensional bsdes. Asia-Pacific Financial Markets 26, 391–408.

Gerig, A., 2015. High-frequency trading synchronizes prices in financial markets. US Securities and Exchange Commission January, –. URL: `https://www.sec.gov/files/dera-wp-hft-synchronizes.pdf`.

Gramacy, R., Ludkovski, M., 2015. Sequential design for optimal stopping problems. SIAM Journal of Financial Mathematics 6, 748–745.

Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specification. Journal of econometrics 16, 121–130.

Gueant, O., Lehalle, C.A., Fernandez-Tapia, J., 2013. Dealing with the inventory risk: a solution to the market making problem. Math Finan Econ 8, 477–507.

Guilbaud, F., Pham, H., 2013. Optimal high-frequency trading with limit and market orders. Quantitative Finance 13, 79–94.

Guo, X., Larrard, A., Ruan, Z., 2017. Optimal placement in a limit order book: an analytical approach. Math Finan Econ 11, 189–213.

Ha, S.Y., Kim, K.K., Lee, K., 2015. A mathematical model for multi-name credit based on community flocking. Quantitative Finance 15, 841–851.

Han, J., Jentzen, A., E, W., 2018. Solving high-dimensional partial differential equations using deep learning. PNAS 115, 8505–8510.

Hardiman, S.J., Bercot, N., Bouchaud, J.P., 2013. Critical reflexivity in financial markets: a Hawkes process analysis. The European Physical Journal B 86, 442.

Hawkes, A.G., 1971a. Point spectra of some mutually exciting point processes. Journal of the Royal Statistical Society. Series B (Methodological) 33, 438–443.

Hawkes, A.G., 1971b. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 83–90.

Hawkes, A.G., Oakes, D., 1974. A cluster process representation of a self-exciting process. Journal of Applied Probability 11, 493–503.

Ho, T., Stoll, H.R., 1981. Optimal dealer pricing under transactions and return uncertainty. Journal of Financial Economics 9, 47–73.

Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural Networks 4, 251–257. doi:10.1016/0893-6080(91)90009-T.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366. doi:10.1016/0893-6080(89)90020-8.

Huepe, C., Aldana, M., 2008. New tools for characterizing swarming systems: A comparison of minimal models. Physic A: Statistical Mechanics and its Applications 387, 2809–2822.

Hutchinson, J., Lo, A., Possio, T., 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. Journal of Finance 49, 851–889.

Kellard, N., Dunis, C., Sarantis, N., 2010. Foreign exchange, fractional cointegration and the implied–realized volatility relation. Journal of Banking & Finance 34, 882–891.

Khandani, A., Kim, A., Lo, A., 2010. Consumer credit-risk models via machine-learning algorithms. International Journal of Forecasting 34, 2767–2787.

Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization arXiv:1412.6980.

Large, J., 2007. Measuring the resiliency of an electronic limit order book. Journal of Financial Market 10, 1–25.

Lee, C.M.C., Ready, M.J., 1991. Inferring trade direction from intraday data. The Journal of Finance 46, 733–746.

Lee, K., Seo, B.K., 2017. Modeling microstructure price dynamics with symmetric Hawkes and diffusion model using ultra-high-frequency stock data. Journal of Economic Dynamics and Control 79, 154–183.

Loterman, G., Brown, I., Martens, D., Mues, C., Baesens, B., 2012. Benchmarking regression algorithms for loss given default modeling. International Journal of Forecasting 28, 161–170.

Maslyuka, S., Smyth, R., 2009. Cointegration between oil spot and future prices of the same and different grades in the presence of structural change. Energy Policy 37, 1687–1963.

Miller, R.S., Shorter, G., 2016. High frequency trading: Overview of recent developments. Congressional Research Service .

Ng, V.K., Pirrong, S.C., 1994. Fundamentals and volatility: Storage, spreads, and the dynamics of metals prices. The Journal of Business 67, 203–230.

Ng, V.K., Pirrong, S.C., 1996. Price dynamics in refined petroleum spot and futures markets. Journal of Empirical Finance 2, 359–388.

Ogata, Y., 1978. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. Annals of the Institute of Statistical Mathematics 30, 243–261. doi:10.1007/BF02480216.

Rauch, E.M., Millonas, M.M., Chialvo, D.R., 1995. Pattern formation and functionality in swarm models. Physics Letters A 207, 185–193.

Reynolds, C.W., 1987. Flocks, herds and schools: A distributed behavioral model. ACM SIGGRAPH Computer Graphics 21, 25–34.

Rosu, I., 2009. A dynamic model of the limit order book. The Review of Financial Studies 22, 4601–4641.

SEC, 2010. Concept Release on Equity Market Structure. US Securities and Exchange Commission 17 CFR Part 242, 34–61358. URL: https://www.sec.gov/rules/concept/2010/34-61358.pdf.

SEC, 2014. Equity market structure literature review part ii: High frequency trading. US Securities and Exchange Commission March, –. URL: https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf.

Sirignano, J., 2019. Deep learning for limit order books. Quantitative Finance 19, 549–570.

Sirignano, J., Spiliopoulos, K., 2018. DGM: A deep learning algorithm for solving partial differential equations. Journal of Computational Physics 375, 1339–1364. URL: https://linkinghub.elsevier.com/retrieve/pii/S0021999118305527, doi:10.1016/j.jcp.2018.08.029.

Taylor, M., Tonks, I., 1989. The internationalisation of stock markets and the abolition of uk exchange control. The Review of Economics and Statistics 71, 332–336.

Veraart, L.A.M., 2010. Optimal market making in the foreign exchange market. Appl Math Fin 17, 359–372.

A Proofs

A.1 Lemma 1

We derive the conditions that guarantee the intensity processes for market orders be stable in Lemma 1 with the mean future rate $m_t^{\pm}(u) = \mathbb{E}[\lambda_u^{\pm}|\mathcal{F}_t]$ for $u \geq t$.

Lemma 1. *The mean future rates $m_t^{\pm}(u)$ are bounded for all $u \geq t$ if and only if $(1-\kappa)\beta > \eta + \nu$. Furthermore,*

$$\lim_{u\to\infty} m_t^{\pm}(u) = D^{-1}\pi, \quad \textit{where} \quad D = \begin{bmatrix} \beta-\eta & -\kappa\beta-\nu \\ -\kappa\beta-\nu & \beta-\eta \end{bmatrix} \quad \textit{and} \quad \pi = \begin{bmatrix} \beta\theta^- \\ \beta\theta^+ \end{bmatrix}.$$

Proof. Similar to the scheme of Cartea et al. (2014), taking the integral of both sides of Eq.(3), the conditional expectation $\mathbb{E}[\cdot|\mathcal{F}_t]$ by applying Fubini's theorem, and the derivative gives the following system of ordinary differential equations (ODEs) for $m_t^{\pm}(u)$, then we have

$$\frac{d}{du}\begin{bmatrix} m_t^-(u) \\ m_t^+(u) \end{bmatrix} = \begin{bmatrix} \beta\theta^- \\ \beta\theta^+ \end{bmatrix} + \begin{bmatrix} -\beta+\eta & \kappa\beta+\nu \\ \kappa\beta+\nu & -\beta+\eta \end{bmatrix}\begin{bmatrix} m_t^-(u) \\ m_t^+(u) \end{bmatrix} \tag{18}$$

with initial values $m_t^-(t) = \lambda_t^-, m_t^+(t) = \lambda_t^+$. If D has no zero eigenvalues, it takes a unique solution such that

$$\begin{bmatrix} m_t^-(u) \\ m_t^+(u) \end{bmatrix} = e^{-D(u-t)}\left(\begin{bmatrix} \lambda_t^- \\ \lambda_t^+ \end{bmatrix} - D^{-1}\pi\right) + D^{-1}\pi. \tag{19}$$

Since the eigenvalues of D are $(1\pm\kappa)\beta - (\eta\mp\nu)$, $\lim_{u\to\infty} m_t(u)$ converges to $D^{-1}\pi$ if and only if $(1\pm\kappa)\beta > \eta\mp\nu$. It implies $(1-\kappa)\beta > \eta+\nu$ since $\beta,\eta,\nu,\kappa > 0$. □

We also obtain the conditional expectation $n_t^{\pm}(u) = \mathbb{E}[c_u^{\pm}|\mathcal{F}_t]$ by solving a system of ODEs derived by taking the integral, conditional expectation, and derivative in Eq.(4).

Remark 2. *The mean future rate $n_t^{\pm}(u)$ is given by, for all $u \geq t$,*

$$\begin{bmatrix} n_t^-(u) \\ n_t^+(u) \end{bmatrix} = e^{-\tilde{D}(u-t)}\int_t^u e^{\tilde{D}(s-t)}G(s)ds + e^{-\tilde{D}(u-t)}\begin{bmatrix} n_t^- \\ n_t^+ \end{bmatrix}, \tag{20}$$

where

$$\tilde{D} = \begin{bmatrix} \xi & -\kappa_c\xi \\ -\kappa_c\xi & \xi \end{bmatrix} \quad \textit{and} \quad G(s) = \begin{bmatrix} \xi\alpha \\ \xi\alpha \end{bmatrix} + \eta_k\begin{bmatrix} m_t^-(s) \\ m_t^+(s) \end{bmatrix} + \nu_k\begin{bmatrix} m_t^+(s) \\ m_t^-(s) \end{bmatrix} \tag{21}$$

with the initial condition $n_t^{\pm} = n_t^{\pm}(t)$ and $m_t^{\pm}(s)$ in Eq.(19).

A.2 Solution for the Equation (3)

In Section 3, we consider the intensity processes $(\lambda_t^-, \lambda_t^+)_{t\geq 0}$ for market sell and buy orders $(M_t^-, M_t^+)_{t\geq 0}$ such that

$$\begin{aligned} \mathrm{d}\lambda_t^- &= \beta(\theta^- - \lambda_t^- + \kappa\lambda_t^+)\mathrm{d}t + \eta\mathrm{d}M_t^- + \nu\mathrm{d}M_t^+ \\ \mathrm{d}\lambda_t^+ &= \beta(\theta^+ - \lambda_t^+ + \kappa\lambda_t^-)\mathrm{d}t + \eta\mathrm{d}M_t^+ + \nu\mathrm{d}M_t^- \end{aligned} \tag{22}$$

with all non-negative coefficients.

For the goodness-of-fit test to the proposed intensity model for the ML estimator, the explicit form for $(\lambda_t^-, \lambda_t^+)_{t\geq 0}$ is needed to have a likelihood function. Since the synchronising terms in $\mathrm{d}\lambda_t^+$ and $\mathrm{d}\lambda_t^-$, the system of the stochastic differential equations (SDEs) is not solved straightforward. By Ha et al. (2015), we split each process into an average process and each fluctuation process, which are the average process $\bar{\lambda}_t = \frac{1}{2}(\lambda_t^+ + \lambda_t^-)$, and the

fluctuation processes $\tilde{\lambda_t^-} = \lambda_t^- - \bar{\lambda}_t$ and $\tilde{\lambda_t^+} = \lambda_t^+ - \bar{\lambda}_t$. From this, $\tilde{\lambda_t^-} + \tilde{\lambda_t^+} = 0$ which is applied to derive the solution $\tilde{\lambda}_t$.

First, we have the average process such as

$$\mathrm{d}\bar{\lambda_t} = \frac{\beta}{2}\left(\theta + 2(\kappa-1)\bar{\lambda_t}\right)dt + \zeta \mathrm{d}M_t^- + \zeta \mathrm{d}M_t^+ \tag{23}$$

where $\theta = \theta^- + \theta^+$ and $\zeta = \frac{1}{2}(\eta + \nu)$. By letting $A = \beta(\kappa - 1)$, a (strong) solution of the linear SDE in Eq.(23) is given by

$$\bar{\lambda_t} = e^{At}\bar{\lambda_0} + \frac{\theta}{2(\kappa-1)}(e^{At} - 1) + \zeta e^{At}\left(\int_0^t e^{-As}dM_s^- + e^{-As}dM_s^+\right). \tag{24}$$

Second, we have the SDE for each fluctuation process $\tilde{\lambda_t^+}, \tilde{\lambda_t^-}$ such as

$$\begin{aligned}
\mathrm{d}\tilde{\lambda_t^-} &= \frac{\beta}{2}\left(\tilde{\theta} - 2(\kappa+1)\tilde{\lambda_t^-}\right)dt + \tilde{\zeta}\mathrm{d}M_t^- - \tilde{\zeta}\mathrm{d}M_t^+ \\
\mathrm{d}\tilde{\lambda_t^+} &= -\frac{\beta}{2}\left(\tilde{\theta} + 2(\kappa+1)\tilde{\lambda_t^+}\right)dt - \tilde{\zeta}\mathrm{d}M_t^- + \tilde{\zeta}\mathrm{d}M_t^+
\end{aligned} \tag{25}$$

where $\tilde{\theta} = \theta^- - \theta^+$ and $\tilde{\zeta} = \frac{1}{2}(\eta - \nu)$. By letting $\tilde{A} = \beta(\kappa + 1)$, the (strong) solutions of Eq.(25) are given by

$$\begin{aligned}
\tilde{\lambda_t^-} &= e^{-\tilde{A}t}\tilde{\lambda_0^-} + \frac{\tilde{\theta}}{2(\kappa+1)}(1 - e^{-\tilde{A}t}) + \tilde{\zeta}e^{-\tilde{A}t}\left(\int_0^t e^{-\tilde{A}s}dM_s^- - e^{-\tilde{A}s}dM_s^+\right) \\
\tilde{\lambda_t^+} &= e^{-\tilde{A}t}\tilde{\lambda_0^+} - \frac{\tilde{\theta}}{2(\kappa+1)}(1 - e^{-\tilde{A}t}) - \tilde{\zeta}e^{-\tilde{A}t}\left(\int_0^t e^{-\tilde{A}s}dM_s^- - e^{-\tilde{A}s}dM_s^+\right),
\end{aligned} \tag{26}$$

respectively. From Eqs.(24) and (26), we have $\lambda_t^- = \bar{\lambda}_t + \tilde{\lambda_t^-}$ and $\lambda_t^+ = \tilde{\lambda_t^+} + \bar{\lambda}_t$.

A.3 Convergence theorem for loss function $L(f)$

Theorem 3. *Denote the set of all functions implemented by one hidden layer with n hidden units and one output unit as*

$$\mathfrak{N}_k^n(\varphi) = \left\{ g : \mathbb{R}^k \to \mathbb{R} \middle| g(x) = \sum_{j=1}^n b_j \varphi\left(a_j' x + c_j\right) \right\}$$

where φ is the activation function of the hidden units. When a function is implemented by a network with an arbitrarily large number of hidden units, the function is contained in

$$\mathfrak{N}_k(\varphi) = \bigcup_{n=1}^{\infty} \mathfrak{N}_k^n(\varphi).$$

For every $\epsilon > 0$, there exists a function $f \in \mathfrak{N}_k(\varphi)$ such that

$$L(f) \leq \epsilon$$

for the nonlinear operator $\mathcal{L}$ defined by

$$\begin{aligned}
\mathcal{L}u(t,q,\boldsymbol{\lambda},\boldsymbol{c}) = {} & \beta(\theta^- - \lambda^- + \kappa\lambda^+)\frac{\partial u}{\partial \lambda^-} + \beta(\theta^+ - \lambda^+ + \kappa\lambda^-)\frac{\partial u}{\partial \lambda^+} \\
& + \xi(\alpha - c^- + \kappa_c c^+)\frac{\partial u}{\partial c^-} + \xi(\alpha - c^+ + \kappa_c c^-)\frac{\partial u}{\partial c^+} \\
& + \lambda^-(\Delta^-_{q,\lambda,c}u - u)\mathbb{1}_{\{c^-(\Delta^-_{q,\lambda,c}u - \Delta^-_{\lambda,c}u)\geq 1\}} + \lambda^+(\Delta^+_{q,\lambda,c}u - u)\mathbb{1}_{\{c^+(\Delta^+_{q,\lambda,c}u - \Delta^+_{\lambda,c}u)\geq 1\}} \\
& + \lambda^-\left(\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c}u - \Delta^-_{\lambda,c}u)}}{\mathrm{e}c^-} + \Delta^-_{\lambda,c}u - u\right)\mathbb{1}_{\{c^-(\Delta^-_{q,\lambda,c}u - \Delta^-_{\lambda,c}u)<1\}} \\
& + \lambda^+\left(\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c}u - \Delta^+_{\lambda,c}u)}}{\mathrm{e}c^+} + \Delta^+_{\lambda,c}u - u\right)\mathbb{1}_{\{c^+(\Delta^+_{q,\lambda,c}u - \Delta^+_{\lambda,c}u)<1\}} - \psi q^2.
\end{aligned} \tag{27}$$

Proof. By the universal approximation theorem proposed by Hornik (1991), for any $\epsilon > 0$, there exists a function $f \in \mathfrak{N}_6(\varphi)$ such that

$$\max_{|\alpha|\leq 1} \sup_{\tilde{\mathcal{D}}_T} |D^\alpha(f(t,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) - u(t,q,\boldsymbol{\lambda},\boldsymbol{c}))| < \epsilon, \tag{28}$$

where D^α is the αth order derivative operator.

The second term of loss function $L(f)$ in Eq.(15) is

$$\left\| f(T,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) - \phi q^2 \right\|^2_{\tilde{\mathcal{D}},\mu_2} = \int_{\tilde{\mathcal{D}}} |f-u|^2 \, d\mu_2. \tag{29}$$

For the first term of Eq.(15), we can write that

$$\begin{aligned}
& \left\| \frac{\partial f}{\partial t}(t,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) + \mathcal{L}f(t,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) \right\|_{\tilde{\mathcal{D}}_T,\mu_1} \\
& = \left\| \frac{\partial f}{\partial t}(t,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) - \frac{\partial u}{\partial t}(t,q,\boldsymbol{\lambda},\boldsymbol{c}) + \mathcal{L}f(t,q,\boldsymbol{\lambda},\boldsymbol{c};\Theta) - \mathcal{L}u(t,q,\boldsymbol{\lambda},\boldsymbol{c}) \right\|_{\tilde{\mathcal{D}}_T,\mu_1}.
\end{aligned} \tag{30}$$

Let $A_h^\pm = \{(t,q,\boldsymbol{\lambda},\boldsymbol{c}) \in \tilde{\mathcal{D}}_T \,\big|\, c^\pm(\Delta^\pm_{q,\lambda,c}h - \Delta^\pm_{\lambda,c}h) < 1\}$. By the Minkowski inequality,

Eq.(30) is less than the following:

$$
\begin{aligned}
&\left(\int_{\tilde{\mathcal{D}}_T}\left|\frac{\partial f}{\partial t}-\frac{\partial u}{\partial t}\right|^2 d\mu_1\right)^{1/2}+\left(\int_{\tilde{\mathcal{D}}_T}\left|\beta(\Theta^- -\lambda^- +\kappa\lambda^+)\left(\frac{\partial f}{\partial\lambda^-}-\frac{\partial u}{\partial\lambda^-}\right)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{\tilde{\mathcal{D}}_T}\left|\beta(\Theta^+ -\lambda^+ +\kappa\lambda^-)\left(\frac{\partial f}{\partial\lambda^+}-\frac{\partial u}{\partial\lambda^+}\right)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{\tilde{\mathcal{D}}_T}\left|\xi(\alpha-c^- +\kappa_c c^+)\left(\frac{\partial f}{\partial c^-}-\frac{\partial u}{\partial c^-}\right)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{\tilde{\mathcal{D}}_T}\left|\xi(\alpha-c^+ +\kappa_c c^-)\left(\frac{\partial f}{\partial c^+}-\frac{\partial u}{\partial c^+}\right)\right|^2 d\mu_1\right)^{1/2}+\left(\int_{\tilde{\mathcal{D}}_T}\left|(\lambda^- +\lambda^+)(f-u)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{\tilde{\mathcal{D}}_T\setminus(A_f^-\cup A_u^-)}\left|\lambda^-(\Delta^-_{q,\lambda,c}f-\Delta^-_{q,\lambda,c}u)\right|^2 d\mu_1\right)^{1/2}+\left(\int_{\tilde{\mathcal{D}}_T\setminus(A_f^+\cup A_u^+)}\left|\lambda^+(\Delta^+_{q,\lambda,c}f-\Delta^+_{q,\lambda,c}u)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{A_f^-\cap\overline{A_u^-}}\left|\lambda^-\left(\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c}f-\Delta^-_{\lambda,c}f)}}{\mathrm{e}c^-}-\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c}u-\Delta^-_{\lambda,c}u)}}{\mathrm{e}c^-}\right)\right|^2 d\mu_1\right)^{1/2} \qquad (*)\\
&+\left(\int_{A_f^+\cap\overline{A_u^+}}\left|\lambda^+\left(\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c}f-\Delta^+_{\lambda,c}f)}}{\mathrm{e}c^+}-\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c}u-\Delta^+_{\lambda,c}u)}}{\mathrm{e}c^+}\right)\right|^2 d\mu_1\right)^{1/2} \qquad (*)\\
&+\left(\int_{A_f^-\cap A_u^-}\left|\lambda^-(\Delta^-_{\lambda,c}f-\Delta^-_{\lambda,c}u)\right|^2 d\mu_1\right)^{1/2}+\left(\int_{A_f^+\cap A_u^+}\left|\lambda^+(\Delta^+_{\lambda,c}f-\Delta^+_{\lambda,c}u)\right|^2 d\mu_1\right)^{1/2}\\
&+\left(\int_{A_u^-\setminus A_f^-}\left|\lambda^-\left(\Delta^-_{q,\lambda,c}f-\Delta^-_{\lambda,c}u-\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c}u-\Delta^-_{\lambda,c}u)}}{\mathrm{e}c^-}\right)\right|^2 d\mu_1\right)^{1/2} \qquad (**)\\
&+\left(\int_{A_f^-\setminus\overline{A_u^-}}\left|\lambda^-\left(\Delta^-_{q,\lambda,c}u-\Delta^-_{\lambda,c}f-\frac{\mathrm{e}^{c^-(\Delta^-_{q,\lambda,c}f-\Delta^-_{\lambda,c}f)}}{\mathrm{e}c^-}\right)\right|^2 d\mu_1\right)^{1/2} \qquad (\dagger)\\
&+\left(\int_{A_u^+\setminus A_f^+}\left|\lambda^+\left(\Delta^+_{q,\lambda,c}f-\Delta^+_{\lambda,c}u-\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c}u-\Delta^+_{\lambda,c}u)}}{\mathrm{e}c^+}\right)\right|^2 d\mu_1\right)^{1/2} \qquad (**)\\
&+\left(\int_{A_f^+\setminus\overline{A_u^+}}\left|\lambda^+\left(\Delta^+_{q,\lambda,c}u-\Delta^+_{\lambda,c}f-\frac{\mathrm{e}^{c^+(\Delta^+_{q,\lambda,c}f-\Delta^+_{\lambda,c}f)}}{\mathrm{e}c^+}\right)\right|^2 d\mu_1\right)^{1/2}. \qquad (\dagger)
\end{aligned}
$$

For the integrals $(*)$, the exponential functions are uniformly continuous because

$$c^{\pm}\left(\Delta^{\pm}_{q,\lambda,c}f-\Delta^{\pm}_{\lambda,c}f\right)<1 \ \text{ and } \ c^{\pm}\left(\Delta^{\pm}_{q,\lambda,c}u-\Delta^{\pm}_{\lambda,c}u\right)\le 1.$$

We can make the integrand small enough by choosing $f \in \mathfrak{N}_k(\varphi)$ sufficiently close to u. For the integrals $(**)$,

$$\left(\Delta^{\pm}_{q,\lambda,c}u - \Delta^{\pm}_{\lambda,c}u\right) < \frac{1}{c^{\pm}} \leq \left(\Delta^{\pm}_{q,\lambda,c}f - \Delta^{\pm}_{\lambda,c}f\right)$$

and with Eq.(28), $A^{\pm}_u \setminus A^{\pm}_f \subset \{(t,q,\boldsymbol{\lambda},\boldsymbol{c}) \in \tilde{\mathcal{D}}_T \,\big|\, \Delta^{\pm}_{q,\lambda,c}u - \Delta^{\pm}_{\lambda,c}u \in \left(\frac{1}{c^{\pm}} - 2\epsilon, \frac{1}{c^{\pm}}\right)\}$. Thus, $\mu_1(A^{\pm}_u \setminus A^{\pm}_f)$ decreases to zero if ϵ becomes arbitrarily small. The integrals $(\dagger)$ can be treated similarly. Other integrals and Eq.(29) are bounded by ϵ times some constant related to a size of domain $\mathcal{D}_T$, therefore we can make the loss $L(f)$ arbitrarily be small.

□

A.4 Thinning algorithm for the proposed model

Let $t_1, t_2, \ldots$ with $0 < t_1 < t_2 < \ldots$ denote the random arrival times at which the counting processes M^-_t or M^+_t jump. Algorithm 2 describes a thinning algorithm to generate the random arrival times.

B Figures and tables related to the MLE results

Table 5 illustrates the MLE calibration results in 2008 and 2018 for the constant mean-reversion levels θ^+, θ^-, self- and mutually-exciting factors η, ν, and mean-reverting speed β with mean, standard deviation, skewness, and kurtosis for the 6 selected stocks. The column 'Average' represents the value from taking average on all the stock's estimations. All the estimations were statistically significant.

Figures 9-16 illustrate the daily changes in our model parameters MLE-estimated from market order data for each stock in 2008 and 2018. The 2008-year results supper-impose three episode dates related to defaults[7]. The specific date of each episode is given as follows: (a) March 14, 2008, Bear Stearns' default; (b) September 15, 2008, Lehman Brothers' default; and (c) November 23, 2008, Citigroup's insolvency. In particular, after the Bear Stearns' default, the IBM and Chevron's synchronising effect κ has almost disappeared, and their self-exciting pattern η has jumped down remarkably, as shown in Figures 9 and 13.

Figures 17-19 present the yearly changes in the model parameters from 2008 to 2018 by taking average on all estimations from daily data for each stock. Figure 20 shows the same results by taking average on all the stocks per year. It is noticed that, as a whole, synchrony tendency between buy and sell orders has been getting firm and evident over time, even though its level has pretty fluctuated.

[7] `https://lauder.wharton.upenn.edu/wp-content/uploads/2015/06/Chronology_Economic_Financial_Crisis.pdf`

Algorithm 2 A thinning algorithm to simulate the processes $M_t^{\pm}$.

Initialize $t_0 = 0$ and $n = 0$.
while $t_n < T$ **do**
  Set $s^- = 0$ and $s^+ = 0$.
  while true do
    Draw $E^- \sim \exp(\lambda^-_{t_n})$ and $U^- \sim \mathrm{Unif}(0, \lambda^-_{t_n})$.
    Update $s^- = s^- + E^-$.
    Set $\lambda^-_{new} = \left(\frac{\lambda^-_{t_n}+\lambda^+_{t_n}}{2} - \frac{\theta^-+\theta^+}{2(1-\kappa)}\right) \mathrm{e}^{-\beta(1-\kappa)s^-} + \left(\frac{\lambda^-_{t_n}-\lambda^+_{t_n}}{2} - \frac{\theta^--\theta^+}{2(1+\kappa)}\right) \mathrm{e}^{-\beta(1+\kappa)s^-} + \frac{\theta^-+\theta^+}{2(1-\kappa)} + \frac{\theta^--\theta^+}{2(1+\kappa)}$.
    if $U^- \leq \lambda^-_{new}$ **then**
      break
    end if
  end while
  while true do
    Draw $E^+ \sim \exp(\lambda^+_{t_n})$ and $U^+ \sim \mathrm{Unif}(0, \lambda^+_{t_n})$.
    Update $s^+ = s^+ + E^+$.
    Set $\lambda^+_{new} = \left(\frac{\lambda^-_{t_n}+\lambda^+_{t_n}}{2} - \frac{\theta^-+\theta^+}{2(1-\kappa)}\right) \mathrm{e}^{-\beta(1-\kappa)s^+} - \left(\frac{\lambda^-_{t_n}-\lambda^+_{t_n}}{2} - \frac{\theta^--\theta^+}{2(1+\kappa)}\right) \mathrm{e}^{-\beta(1+\kappa)s^+} + \frac{\theta^-+\theta^+}{2(1-\kappa)} - \frac{\theta^--\theta^+}{2(1+\kappa)}$.
    if $U^+ \leq \lambda^+_{new}$ **then**
      break
    end if
  end while
  Set $t_{n+1} = t_n + \min(s^-, s^+)$.
  if $s^- \leq s^+$ **then**
    Set $\lambda^-_{t_{n+1}} = \lambda^-_{new} + \eta$ and $\lambda^+_{t_{n+1}} = \frac{\theta^-+\theta^+}{2(1-\kappa)} - \frac{\theta^--\theta^+}{2(1+\kappa)} + \left(\frac{\lambda^-_{t_n}+\lambda^+_{t_n}}{2} - \frac{\theta^-+\theta^+}{2(1-\kappa)}\right) \mathrm{e}^{-\beta(1-\kappa)s^-} - \left(\frac{\lambda^-_{t_n}-\lambda^+_{t_n}}{2} - \frac{\theta^--\theta^+}{2(1+\kappa)}\right) \mathrm{e}^{-\beta(1+\kappa)s^-} + \nu$.
  else
    Set $\lambda^+_{t_{n+1}} = \lambda^+_{new} + \eta$ and $\lambda^-_{t_{n+1}} = \frac{\theta^-+\theta^+}{2(1-\kappa)} + \frac{\theta^--\theta^+}{2(1+\kappa)} + \left(\frac{\lambda^-_{t_n}+\lambda^+_{t_n}}{2} - \frac{\theta^-+\theta^+}{2(1-\kappa)}\right) \mathrm{e}^{-\beta(1-\kappa)s^+} + \left(\frac{\lambda^-_{t_n}-\lambda^+_{t_n}}{2} - \frac{\theta^--\theta^+}{2(1+\kappa)}\right) \mathrm{e}^{-\beta(1+\kappa)s^+} + \nu$.
  end if
  Update $n = n + 1$.
end while
return $\{t_k\}_{1 \leq k \leq n-1}$.

Table 5: Estimation results for $\theta^+, \theta^-, \eta, \nu, \beta$: mean(1), standard deviation(2), skewness(3), kurtosis(4)

		2008					2018				
		θ^+	θ^-	η	ν	β	θ^+	θ^-	η	ν	β
IBM	1	0.586	0.601	304.0	119.3	1287	0.141	0.150	224.4	27.39	749.0
	2	0.202	0.189	106.5	48.15	345.7	0.068	0.075	115.4	10.53	277.8
	3	0.978	1.176	0.344	0.506	1.528	2.027	2.479	1.990	0.905	1.175
	4	4.876	5.388	2.381	2.333	4.572	8.076	12.52	8.451	4.689	4.935
Chevron	1	0.770	0.781	306.9	117.4	1322	0.161	0.168	214.6	29.69	767.4
	2	0.255	0.247	143.1	48.15	394.5	0.069	0.060	99.95	12.05	270.6
	3	0.608	0.680	1.096	0.374	1.553	2.790	1.243	0.893	1.088	0.644
	4	2.699	2.804	4.410	3.092	4.690	19.59	5.908	3.594	5.401	3.072
Apple	1	2.451	2.423	859.9	179.9	1984	0.965	1.005	3447	434.5	7030
	2	0.787	0.767	335.1	127.1	794.4	0.417	0.389	1796	354.6	4261
	3	0.146	-0.060	-0.420	0.218	-0.453	1.005	0.661	4.262	5.244	7.452
	4	3.945	3.452	3.060	1.439	2.583	3.921	2.857	33.31	44.38	85.26
JP Morgan	1	1.437	1.436	2941	674.0	8025	0.380	0.379	156.0	30.13	684.4
	2	0.572	0.570	4176	2003	11040	0.159	0.170	53.14	11.23	193.9
	3	1.316	1.330	4.659	8.135	5.106	1.465	1.564	1.454	2.505	1.152
	4	5.791	5.831	31.18	84.65	40.18	5.176	6.166	7.062	17.26	5.192
Amazon	1	0.751	0.749	510.8	154.4	1267	0.519	0.499	3089	181.1	5349
	2	0.434	0.428	318.4	126.9	815.6	0.274	0.252	1311	104.3	2077
	3	0.772	0.785	-0.341	-0.084	-0.223	1.429	1.818	0.933	1.495	0.687
	4	3.419	3.410	1.970	1.247	1.987	5.224	7.994	3.835	5.908	3.266
Microsoft	1	1.792	1.793	11191	878.4	12239	0.762	0.779	4849	722.8	9146
	2	1.081	1.080	14280	3533	13938	0.465	0.485	2895	377.4	4884
	3	1.791	1.797	3.610	11.36	3.198	1.569	1.5798	2.964	1.029	2.612
	4	7.138	7.193	19.34	154.1	14.86	5.532	5.750	16.04	4.500	14.22
Average		1.3028	1.3023	2703	356.8	4417	0.4862	0.4947	1998	237.3	3957

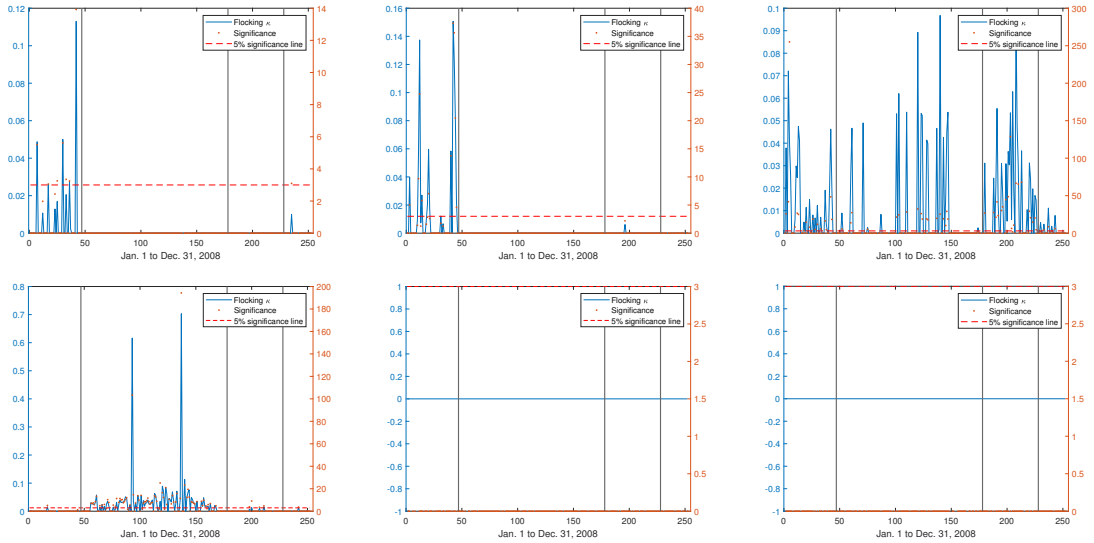

Figure 9: The estimates of a synchronising parameter κ for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2008

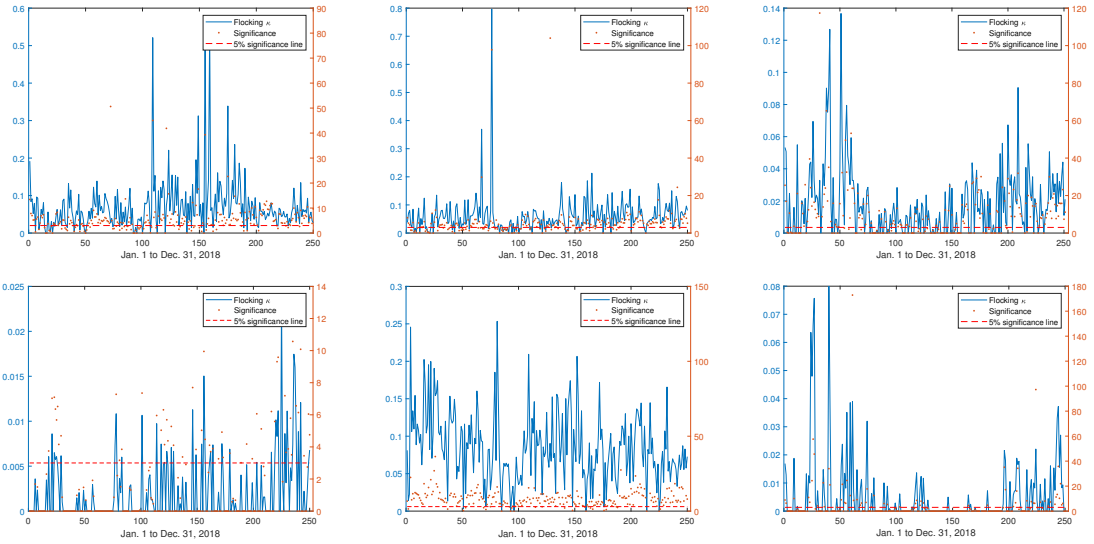

Figure 10: The estimates of a synchronising parameter κ for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2018

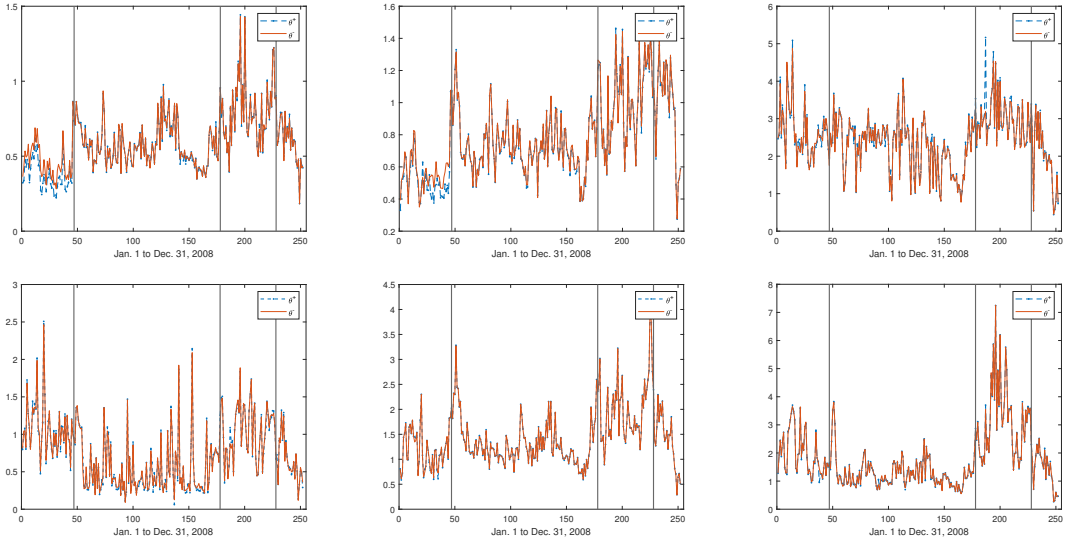

Figure 11: The estimates of the mean-reversion levels θ^-, θ^+ for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – for arrivals of market buy and sell orders from January 2 to December 31, 2008

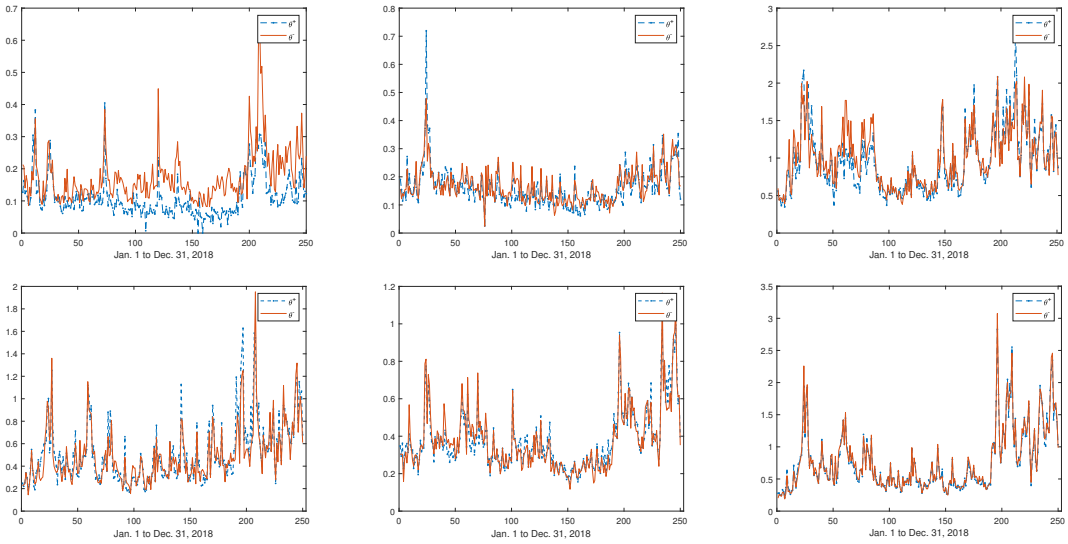

Figure 12: The estimates of the mean-reversion levels θ^-, θ^+ for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2018

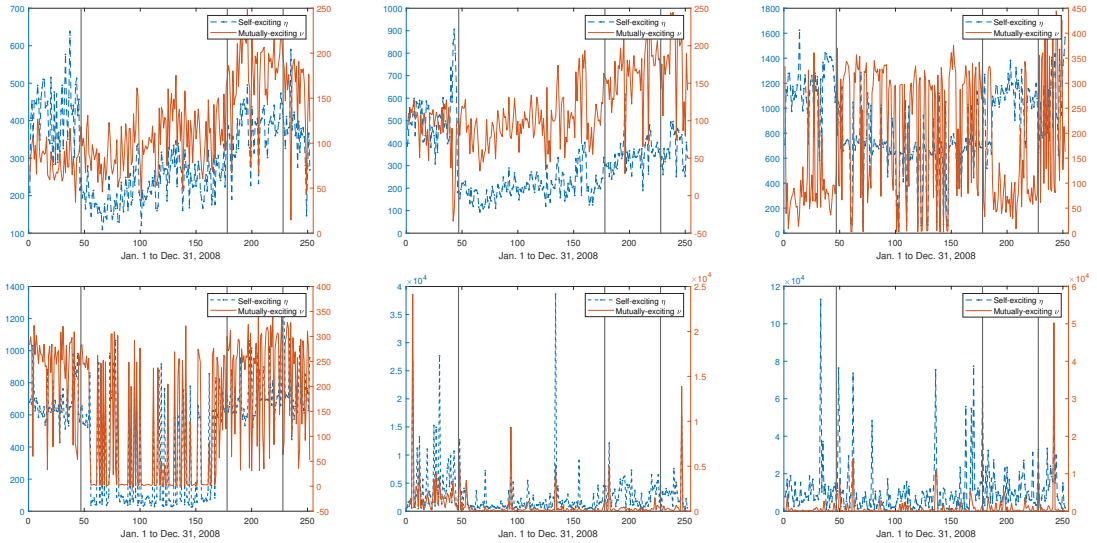

Figure 13: The estimates of a self- and mutually-exciting parameters η, ν for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2008

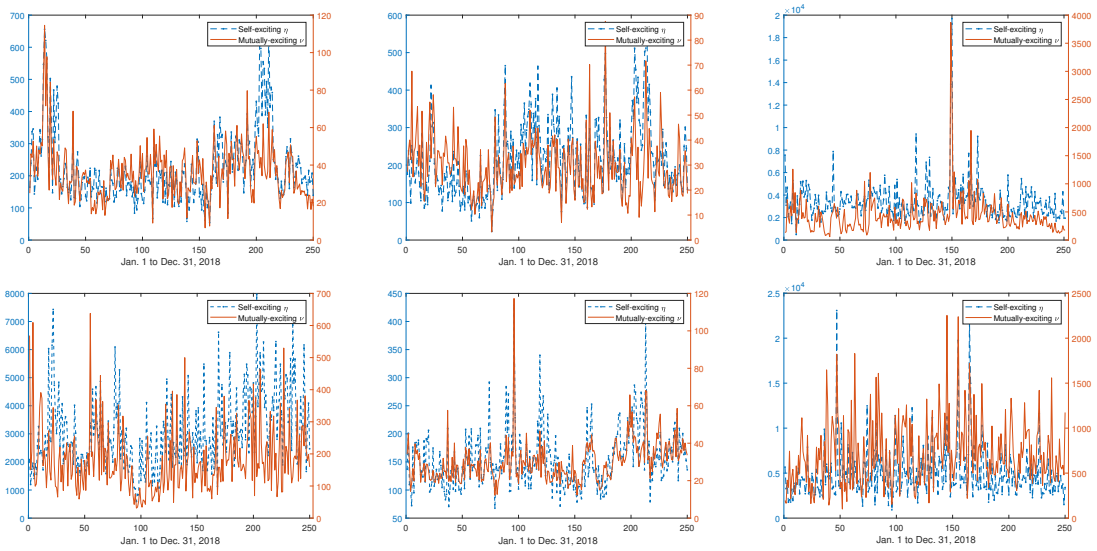

Figure 14: The estimates of a self- and mutually-exciting parameters η, ν for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2018

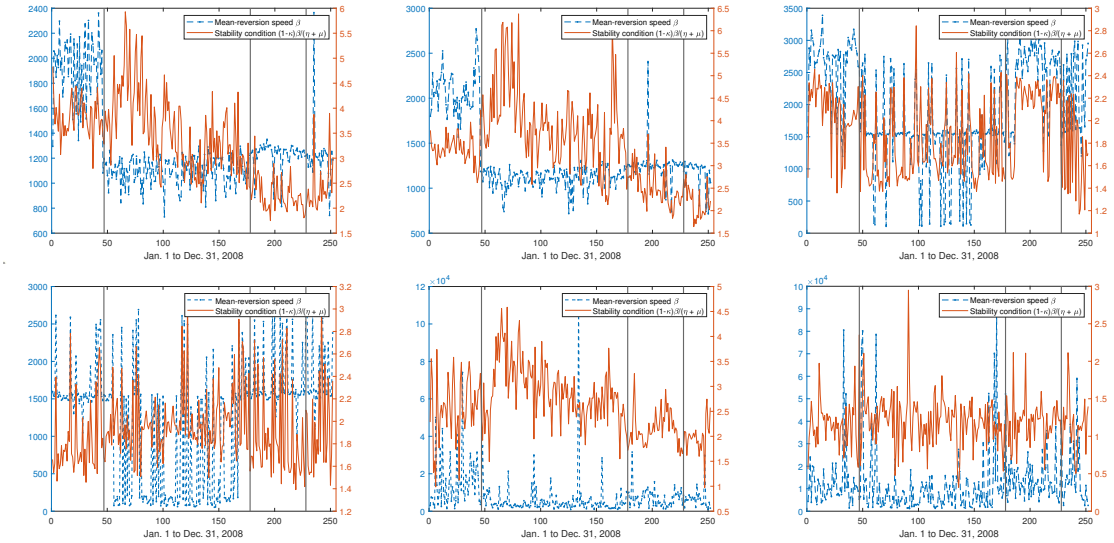

Figure 15: The estimates of a mean-reversion speed β and stability condition for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2008

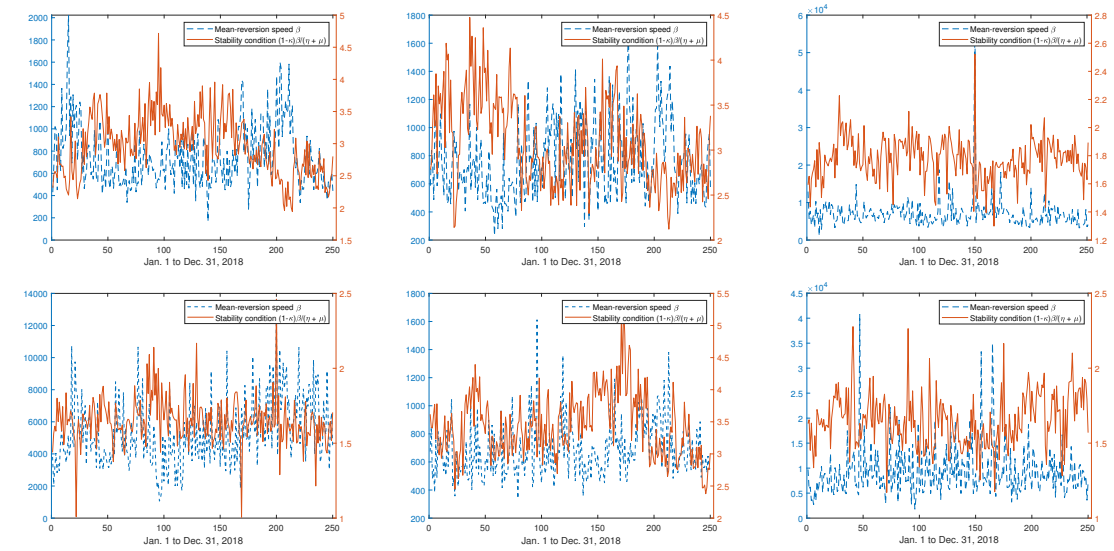

Figure 16: The estimates of a mean-reversion speed β and stability condition for six stocks – IBM, Chevron, Apple, Amazon, JP Morgan, and Microsoft in order from top and left – from January 2 to December 31, 2018

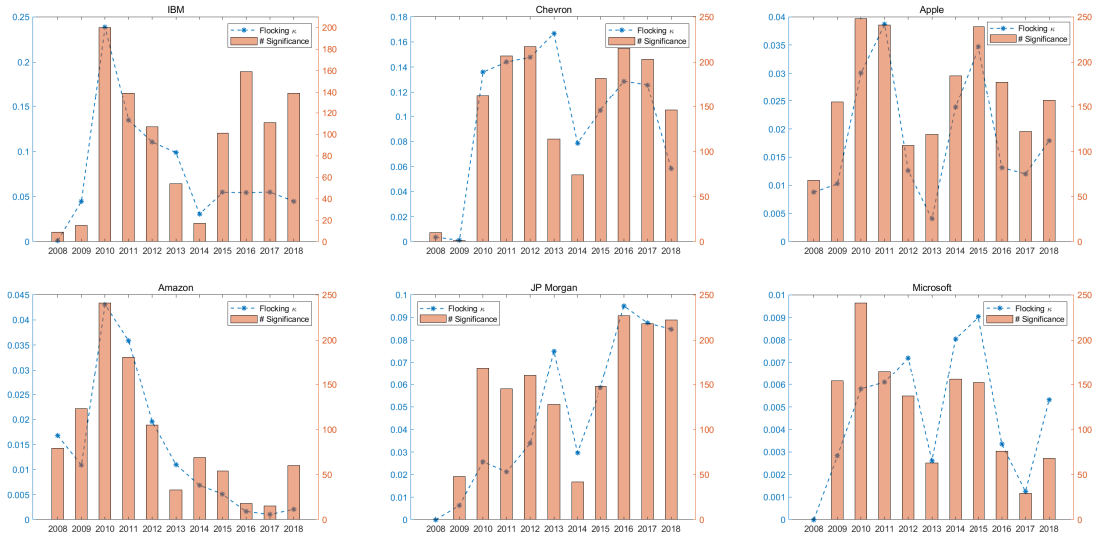

Figure 17: The average of the estimated synchronising parameter κ and the number of significant samples for six stocks in year from 2008 to 2018

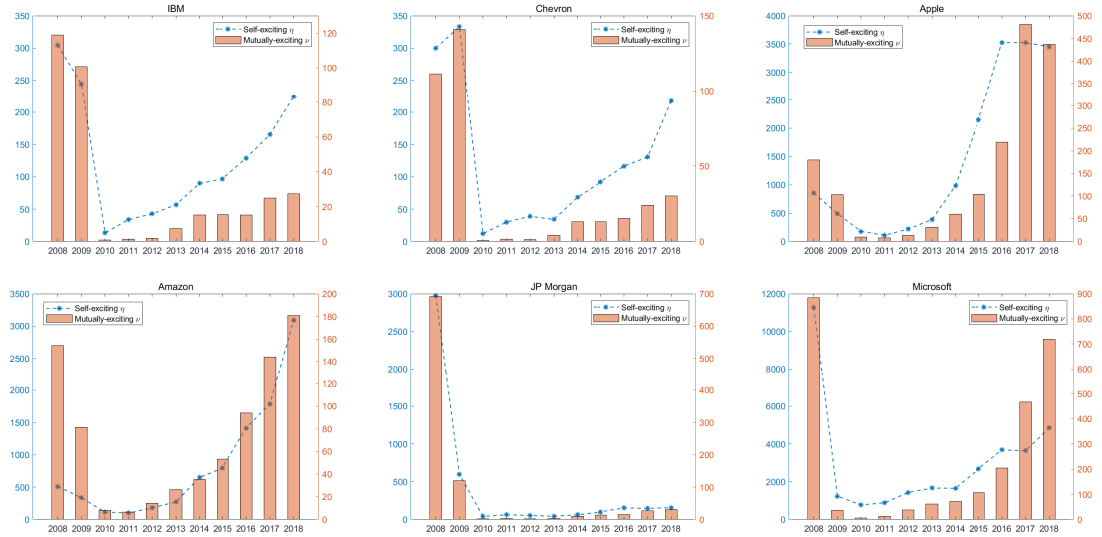

Figure 18: The average of the estimated self and mutually exciting parameters η and ν for six stocks in year from 2008 to 2018

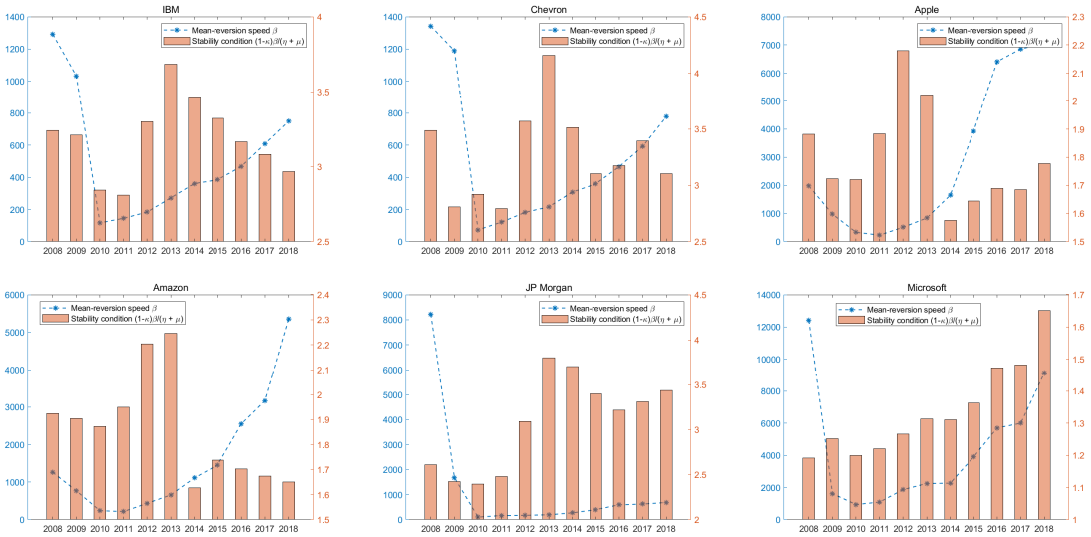

Figure 19: The average of mean-reversion speed β and the stability condition computed with the estimated parameters for six stocks in year from 2008 to 2018

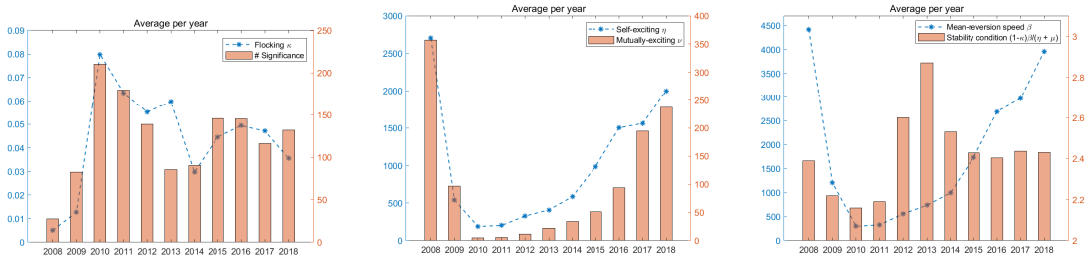

Figure 20: The average of the synchronising factor κ, the exciting factors η, ν, and mean-reversion speed β with stability condition taking average on six stocks in year from 2008 to 2018