

Price Impact Without Averaging

Claudio Bellani, Damiano Brigo, Mikko S. Pakkanen & Leandro Sánchez-Betancourt

To cite this article: Claudio Bellani, Damiano Brigo, Mikko S. Pakkanen & Leandro Sánchez-Betancourt (2023) Price Impact Without Averaging, Applied Mathematical Finance, 30:4, 175-206, DOI: [10.1080/1350486X.2024.2303078](https://doi.org/10.1080/1350486X.2024.2303078)

To link to this article: <https://doi.org/10.1080/1350486X.2024.2303078>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 25 Jan 2024.



[Submit your article to this journal](#)



Article views: 273



[View related articles](#)



[View Crossmark data](#)

Price Impact Without Averaging

Claudio Bellani^a, Damiano Brigo^a, Mikko S. Pakkanen^a and Leandro Sánchez-Betancourt^b

^aDepartment of Mathematics, Imperial College London, London, UK; ^bMathematical Institute and Oxford-Man Institute of Quantitative Finance, University of Oxford, Oxford, UK

ABSTRACT

We present a method to estimate price impact in order-driven markets that does not require averaging over executions or scenarios. Given order book data associated with one single execution of a sell metaorder, we estimate its contribution to price decrease during the trade. We do so by modelling the limit order book using a state-dependent Hawkes process, and by defining the price impact profile of the execution as a function of the compensator of the state-dependent Hawkes process. We apply our method to a dataset from NASDAQ, and we conclude that the scheduling of sell child orders has a bigger impact on price than their sizes.

ARTICLE HISTORY

Received 3 January 2024
Accepted 4 January 2024

KEYWORDS

Price impact; Hawkes processes; metaorder; order-driven markets

1. Introduction

Price impact is the phenomenon whereby trade executions affect the price of the asset being traded. The price is affected in a way that is unfavourable to the trade, i.e., it decreases as a consequence of sell orders and it increases as a consequence of buy orders. Hence, price impact is often regarded as a hidden transaction cost, and price impact detection is considered a branch of transaction cost analysis.

References in the literature are numerous and among them are Torre (1997), Almgren et al. (2005), Moro et al. (2009), Tóth et al. (2011), Engle, Ferstenberg, and Russell (2012), Bacry et al. (2015), Brokmann et al. (2015), Zarinelli et al. (2015), Tóth, Eisler, and Bouchaud (2016) and Patzelt and Bouchaud (2018). An overview of the research done on price impact is provided by Bouchaud (2017).

Recently, Capponi and Cont (2019) have however suggested that the empirical findings on price impact laws may be indistinguishable from the ubiquitous scaling of price variation, i.e., volatility, as a square root of time. Thus, they venture to question the existence of price impact altogether.

However, as observed in Bouchaud (2017), price impact can be regarded as an embodiment of the basic economic law of supply and demand, and it must play a role in price formation. This is to say that, despite the acknowledged entanglement between execution-induced price moves and volatility, the phenomenon of price impact must exist in a market with price mechanism. The consensus among these authors is that, in order to detect price

CONTACT Leandro Sánchez-Betancourt  sanchezbetan@maths.ox.ac.uk  Andrew Wiles Building, Woodstock Rd, Oxford OX2 6GG, UK

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

impact, one has to average across several executions and scenarios to attenuate the impact of volatility – see Bucci et al. (2019).

In this paper, we present a measurement of price impact in order-driven markets that does not require averaging over executions and scenarios. Given the order book data associated with one single execution of a sell metaorder (subdivided into smaller child orders distributed in time), we measure its contribution to price decrease during the trade.

Our motivation is twofold. On the one hand, the economic argument by which price impact must exist applies scenario-by-scenario and to every execution, other things being equal. Therefore, it makes sense to ask whether price impact is detectable without averaging. In other words, is averaging the only practical way to disentangle price impact from volatility? On the other hand, and from a more practical point of view, investors and traders might in fact be more interested in a trade-by-trade assessment of price impact rather than on statistics obtained by extensive averaging. This was pointed out in the aforementioned article by Capponi and Cont (2019), who argue that for institutional investors who execute infrequent but large trades, ‘the observed impact may deviate significantly from such an average’.

Our method to measure price impact is based on a granular model of order book dynamics. Our model extends that of Bacry and Muzy (2014), where a four-dimensional Hawkes process was used to describe arrivals of market orders into a limit order book and price changes. We increase the granularity of the model by describing price changes via a state variable that evolves in time, and by tracking not only market orders but also limit orders. The Hawkes process and state variable are modelled using a class of hybrid marked point processes introduced in Morariu-Patrichi and Pakkanen (2018), more specifically, *state-dependent Hawkes processes*.

The so-increased granularity allows us to: (i) have a snapshot-by-snapshot proxy for outstanding volume on every level of the order book, and (ii) assess the impact that a particular trader has on the market without assuming that their orders walk the book with the same frequency as the orders from other market participants – as assumed in Bacry and Muzy (2014).

For an averaged measure of price impact, one fixes a trade direction for order executions and assesses price movements across several scenarios. Since in these scenarios the signs of all other market variables on average cancel out, an average price movement emerging during these executions represents the impact sought. If we instead eschew averaged measures of price impact, we cannot count on this average cancellation of all market fluctuation other than the trade direction of the considered execution. This poses the counterfactual of what would have happened if the trade under consideration had not been executed. In other words, the execution-induced price movement during the single observed scenario has to be disentangled from the price movement induced by all the surrounding market noise. We do so by formalizing our notion of price impact without averaging around the concept of compensator of a counting process. This allows us to focus on a single realization of the process and yet factor out its volatility. The idea of using compensators of Hawkes processes is reminiscent of existing approaches in the literature, for example in Jusselin and Rosenbaum (2020) and Rosenbaum and Tomas (2021).

We demonstrate our approach by assessing price impact on NASDAQ data for the ticker INTC. Since the identities of market participants behind the observed market events are not disclosed, we cannot directly measure the impact of a chosen execution among those

in the raw data. Instead, we calibrate our market model on the provided data, and then we simulate a liquidation in the calibrated model. By so doing, we illustrate the usefulness of the class of market simulators to which our model belongs.

More precisely, we fix the size of the metaorder to be executed, and we examine how the price impact of such an execution varies with (i) the size of the child orders (whether they are likely to consume all available liquidity on the first price level of the limit order book or not), and (ii) the scheduling of the child orders (whether they cluster around particular times or are evenly distributed during the execution). Empirically, we find that the clustering of child orders has a bigger impact on the price than their sizes.

The paper is organized as follows. Section 2 gives some background about order-driven markets and limit order books. Section 3 recalls basic notions from the theory of counting processes and establishes the notation used in the paper. Section 4 describes our model for order-driven markets based on state-dependent Hawkes processes. Section 5 defines price impact in our model, and describes how we quantify it. Section 6 presents applications of our modelling framework to a dataset from NASDAQ, and provides some insights on the main drivers of price impact in this dataset. Section 7 concludes.

2. Background on Order-Driven Markets and Limit Order Books

Order-driven markets are trading venues organized around *limit orders*. A limit order is the fundamental action that market participants in order-driven markets can perform. It is represented by a 4-tuple (t, q, p, d) , where t denotes time, q denotes size, p denotes price and d denotes direction. We say that a market participant *posts* a limit order (t, q, p, d) if at time t they submit to the exchange their commitment to buy ($d = 1$) or sell ($d = -1$) the amount q at the price p . The price p is interpreted as the highest price at which they are committed to buy if $d = 1$, or the lowest price at which they are committed to sell if $d = -1$. By regulation, the price p must be an integer multiple of some fixed $\psi > 0$ known as the tick size. Market participants have also the possibility to withdraw their commitments to trade, by *cancelling* a previously submitted limit order.

A trading epoch in an order-driven market is the collection of all the limit orders submitted by market participants within some time interval. Limit orders cannot be submitted simultaneously, so that a limit order is identified by its time stamp t . We express this mathematically by saying that a trading epoch is the graph of a function from a subset of the positive half-line to the space $\{(q, p, d) : q \geq 0, p \geq 0, d = -1, 1\}$.

When a market participant submits a limit order (t, q, p, d) , a matching algorithm run by the trading platform looks for possible counterparts to their order: if other participants' commitments to trade exist to (partially) fulfil the order (at a price not worse than p from the submitter's point of view), then their order is (partially) cleared. The fraction of their order that is cleared disappears from the market, and it is said to be *executed*; the fraction that could not be fulfilled is recorded in the *limit order book*, waiting for counterparts to trade with.

A limit order (t, q, p, d) in a trading epoch is said to be *active* (or outstanding) at time u if $t \leq u$ and by time u the order has neither been executed nor been cancelled.

The search for possible trading counterparts for the matching of the incoming order (t, q, p, d) happens among the active orders with time stamp $s < t$ and opposite direction $-d$. We say that the active order $(s, \rho, \pi, -d)$ matches the incoming order (t, q, p, d) if $\pi d \leq$

pd ; in this case, the two orders are executed one against the other, and $\rho \wedge q$ units of shares are transacted at the price π per unit of share. After the match, the matched orders are replaced with the orders $(s, \tilde{\rho}, \pi, -d)$ and $(t, \tilde{q}, p, d)$, where $\tilde{\rho} = \rho - \rho \wedge q$ and $\tilde{q} = q - \rho \wedge q$. At least one of them has null size. If $\tilde{\rho} = 0$, i.e., $q \geq \rho$, the order with time stamp s ceases to be active and it is deleted from its queue. If $\tilde{q} = 0$, i.e., $q \leq \rho$, then the order with time stamp t (i.e., the incoming order) is fully executed, and the search for possible matching counterparts stops. If instead $\tilde{q} > 0$, i.e., $q > \rho$, then the search continues among active orders with opposite direction $-d$.

Definition 2.1: Let E be a trading epoch in an order-driven market. The ask order queue A_t at time t is defined as

$$A_t := \{(s, q, p, -1) \in E : (s, q, p, -1) \text{ is active at time } t\}.$$

Similarly, the bid order queue B_t at time t is defined as

$$B_t := \{(s, q, p, +1) \in E : (s, q, p, +1) \text{ is active at time } t\}.$$

A limit order book is a grid of equally spaced prices at which active limit orders sit. The space between consecutive grid nodes is called the tick size of the LOB, which we denote by ψ . All orders submitted to the exchange must have prices that are integer multiples of the tick size. Prices are increasing from left to right. At every node of the grid, outstanding limit orders to buy or sell at the corresponding price are collected; these limit orders are also said to be queuing there. Spontaneously – i.e., by market forces – such orders organize in a way that buy offers will be displaced on the left (the so-called *bid side*) and sell offers will be on the right (the so-called *ask side*). Indeed, if there were a buy (resp., sell) limit order to the right (resp., left) of some sell (resp., buy) limit orders, or at the same node, then the matching algorithm would have matched them, clearing them out from the order book. Therefore, the whole configuration of a limit order book at time t is given when the following variables are specified:

- (i) the best ask price P_t^a , i.e., the lowest price at which one can find sell offers active at time t ;
- (ii) the best bid price P_t^b , i.e., the highest price at which one can find buy offers active at time t ;
- (iii) the volume $V_t^{a,i}$ of ask offers at price $P_t^{a,i} = P_t^a + (i - 1)\psi$, for $i = 1, 2, \dots$, i.e., the quantity

$$V_t^{a,i} = \sum_{\{(s,q,P_t^{a,i},-1) \in A_t : s \leq t\}} q,$$

where the sum is over all the active sell limit orders submitted by time t for a price of $P_t^a + (i - 1)\psi$;

- (iv) the volume $V_t^{b,i}$ of bid offers at price $P_t^{b,i} = P_t^b - (i - 1)\psi$, for $i = 1, 2, \dots$, i.e., the quantity

$$V_t^{b,i} = \sum_{\{(s,q,P_t^{b,i},1) \in B_t: s \leq t\}} q,$$

where the sum is over all the active buy limit orders submitted by time t for a price of $P_t^b - (i - 1)\psi$.

The whole configuration of a limit order book at time t is described by the variables $(P_t^a, P_t^b, \{(V_t^{a,i}, V_t^{b,i}) : i = 1, 2, \dots\})$. Apart from these variables, other derived quantities are useful to assess the properties of an order book. These are the spread, the mid-price and the volume imbalance (or queue imbalance).

The spread at time t is the distance $|P_t^a - P_t^b|$ between best ask price and best bid price, which we denote by ϕ_t . The mid-price at time t is the mid point in between P_t^b and P_t^a , which we denote by P_t^m , and is given by $P_t^m = (P_t^a + P_t^b)/2$. Finally, the n -levels volume imbalance (or queue imbalance) at time t , denoted I_t^n , is the normalized excess of limit orders on the first n levels of the bid side compared to the limit orders on the first n levels of the ask side, namely

$$I_t^n = \frac{\sum_{i \leq n} V_t^{b,i} - \sum_{i \leq n} V_t^{a,i}}{\sum_{i \leq n} V_t^{b,i} + \sum_{i \leq n} V_t^{a,i}}, \quad (1)$$

where $\sum_{i \leq n} V_t^{b,i}$ (resp., $\sum_{i \leq n} V_t^{a,i}$) is the cumulative volumes on the first n bid (resp., ask) levels. The queue imbalance I_t^n is widely accepted as a reliable signal for the next mid-price move (see Cartea, Donnelly, and Jaimungal 2018): when it is close to -1 the mid-price will likely decrease, and when it is close to $+1$ it will likely increase.

The arrival of a limit order to the market triggers two events: one is the consumption of the liquidity capable to (partially) service the limit order, the other is the addition of the non-executed part of the limit order to the appropriate queue in the order book. Proposition 2.2 states the terms of this decomposition.

Proposition 2.2: *Given the configuration $(P_{t-}^a, P_{t-}^b, \{(V_{t-}^{a,i}, V_{t-}^{b,i}) : i = 1, 2, \dots\})$ of the limit order book immediately before time t , processing the sell limit order $(t, q, p, -1)$ is equivalent to processing the pair of orders $[(t, q_M^-, 0, -1), (t, q - q_M^-, p, -1)]$, with the former having priority over the latter, and where*

$$q_M^- := \min \left(q, \sum_{i \geq 1} V_{t-}^{b,i} \mathbb{1}_{\{P_{t-}^{b,i} \geq p\}} \right).$$

Similarly, processing the buy limit order $(t, q, p, +1)$ is equivalent to processing the pair of orders $[(t, q_M^+, \infty, +1), (t, q - q_M^+, p, +1)]$, where

$$q_M^+ := \min \left(q, \sum_{i \geq 1} V_{t-}^{a,i} \mathbb{1}_{\{P_{t-}^{a,i} \leq p\}} \right).$$

From this point forward, given a limit order (t, q, p, d) , we let $q_M = q_M^+$ if $d = +1$, and $q_M = q_M^-$ if $d = -1$. The first component $(t, q_M, 0, -1)$ in the decomposition $[(t, q_M, 0, -1), (t, q - q_M, p, -1)]$ of the sell limit order $(t, q, p, -1)$ is called *sell market order*. Notice that in the 4-tuple $(t, q_M, 0, -1)$, the price p is set to zero, and this guarantees immediate execution: the sale of the amount q_M is instantaneously matched with outstanding buy limit orders on the bid side, and no fraction of q_M is put into the queue. On the contrary, the second component $(t, q - q_M, p, -1)$ specifies exactly that part of $(t, q, p, -1)$ which will be queued. Similarly, $(t, q_M, \infty, +1)$ represents the market order component of the buy limit order $(t, q, p, +1)$, and it is referred to as *buy market order*. The price $p = \infty$ is a way to express the fact that a counterpart for the purchase of the amount q_M will instantaneously be found on the ask side of the order book. In the following, the term ‘market order’ (either buy or sell) will be referring to the first component in the decomposition of a limit order with non-zero market order size $q_M > 0$.¹

A sell market order $(t, q_M, 0, -1)$ is said to ‘walk the book’ if $q_M > V_t^{b,1}$. Similarly, a buy market order walks the book if its size is larger than the volume of offers on the first ask level.

Given a time window $[0, T]$, the evolution in time of the limit order book $\{(P_t^a, P_t^b, \{(V_t^{a,i}, V_t^{b,i}) : i = 1, 2, \dots\}) : 0 \leq t \leq T\}$ results from the history of all limit order submissions, cancellations and executions that happened in $[0, T]$. If every limit order in $\{(t, q, p, d) : 0 \leq t \leq T\}$ is decomposed as per in Proposition 2.2, then the seller-initiated trades that happened in $[0, T]$ are $\{(t, q_M, 0, -1) : q_M > 0, 0 \leq t \leq T\}$, and the buyer-initiated trades that happened in $[0, T]$ are $\{(t, q_M, \infty, +1) : q_M > 0, 0 \leq t \leq T\}$. Hence, in the following we will identify trades with market orders of non-zero size $q_M > 0$.

3. Background on Counting Processes and Hawkes Processes

3.1. Counting Processes

In this section we introduce our notation for counting processes, and we review basic concepts from the theory of such processes. Our main reference is Daley and Vere-Jones (2008, Chapter 14).

Let d_E be a positive integer. For each e ranging from 1 to d_E , let $T_j^e, j = 1, 2, \dots$, be a strictly increasing sequence of positive random times, and assume that $T_j^e \neq T_{j'}^{e'}$ if $(e, j) \neq (e', j')$. Then,

$$N_e(t) := \sum_j \mathbb{1} \{T_j^e \leq t\}, \quad t \geq 0,$$

is a non-decreasing right-continuous process; we call N_e the counting process associated with the sequence $(T_j^e)_j$. Notice that $(T_j^e)_j$ can be retrieved from N_e by

$$T_j^e = \inf \{t > 0 : N_e(t) \geq j\};$$

hence, there is a one-to-one correspondence between N_e and $(T_j^e)_j$.

For $t > 0$ we define

$$\delta N_e(t) := \lim_{h \downarrow 0} (N_e(t) - N_e(t - h)),$$

and we notice that $\delta N_e(t) = 1$ if and only if $t = T_j^e$ for some j , otherwise $\delta N_e(t) = 0$.

The d_E -dimensional vector $N(t) = (N_1(t), \dots, N_{d_E}(t))$ is referred to as multivariate counting process associated with the d_E sequences $(T_j^e)_j$, $e = 1, \dots, d_E$. Let $N_g(t) := N_1(t) + \dots + N_{d_E}(t)$ be the ground process of N , and let

$$T_n := \inf \{t > 0 : N_g(t) \geq n\}, \quad n = 1, 2, \dots,$$

be the ordered sequence of random times stemming from the union $\{T_j^e : j = 1, 2, \dots; e = 1, \dots, d_E\}$. By defining for $n = 1, 2, \dots$,

$$E_n := \sum_{e=1}^{d_E} e \mathbb{1} \{ \delta N_g(T_n) = \delta N_e(T_n) \},$$

we have that the pair (T_n, E_n) equivalently characterizes the multivariate counting process, because

$$N_e(t) = \sum_n \mathbb{1} \{ T_n \leq t, E_n = e \}, \quad (2)$$

for all $t > 0$ and all $e = 1, \dots, d_E$.

We can interpret this construction by saying that the index e labels d_E types of events that occur in time, and $N_e(t)$ counts the number of events of type e that have occurred by time t .

Example 3.1 (Poisson process): Let τ_e^j , $j = 1, 2, \dots$, $e = 1, \dots, d_E$ be independent random variables such that τ_e^j is exponentially distributed with parameter $\lambda_e > 0$, $e = 1, \dots, d_E$. Let $T_j^e := \sum_{k \leq j} \tau_e^k$, and notice that T_j^e has probability density function

$$f_{e,j}(t) = \frac{\lambda_e^j}{(j-1)!} t^{j-1} e^{-\lambda_e t} \mathbb{1} \{t > 0\}.$$

Then, the multivariate counting process N associated with the arrival times T_j^e is called d_E -dimensional Poisson process of rates $\lambda_1, \dots, \lambda_{d_E}$. This name is justified as follows. Since $\{N_e(t) \geq j\} = \{T_j^e \leq t\}$, we have that $\frac{d}{dt} P(N_e(t) \geq j) = f_{e,j}(t)$. On the other hand, if we define

$$S_{e,j}(t) := \sum_{k \geq j} \frac{(\lambda_e t)^k}{k!} e^{-\lambda_e t},$$

we also have that $\frac{d}{dt} S_{e,j}(t) = f_{e,j}(t)$, by telescopic sum. Since $S_{e,j}(0) = P(N_e(t) \geq 0)$, we deduce that $P(N_e(t) \geq j) = S_{e,j}(t)$, and that

$$P(N_e(t) = j) = \frac{(\lambda_e t)^j}{j!} \exp(-\lambda_e t).$$

Therefore for every t , $N_e(t)$ is a Poisson random variable of parameter $\lambda_e t$, and the ground process N_g of N is such that for every t , $N_g(t) \sim \text{Pois}(\lambda_1 t + \dots + \lambda_{d_E} t)$.

The minimal filtration to which a multivariate counting process N is adapted – and such that it satisfies the usual conditions of completeness and right-continuity – is called

the internal history of N . Any other filtration to which N is adapted is called a history of N , and it must be a superset of the internal history.

Definition 3.2: Let $(\Omega, \mathfrak{F} = (\mathfrak{F}_t)_t, P)$ be a filtered probability space where the multivariate counting process N is defined, and assume that $\mathfrak{F}$ is a history of N . We say that the d_E -dimensional stochastic process $\Lambda = (\Lambda_1, \dots, \Lambda_{d_E})$ is an $\mathfrak{F}$ -compensator for N if: (i) $\Lambda(0) = 0$ and Λ is of finite variation; (ii) Λ is $\mathfrak{F}$ -predictable; (iii) Λ is right-continuous; (iv) $N - \Lambda$ is a local martingale.

Given the counting process N and a history $\mathfrak{F}$, the $\mathfrak{F}$ -compensator is unique up to an evanescent set, and it is equivalently characterized as the $\mathfrak{F}$ -predictable projection of N , namely as the $\mathfrak{F}$ -predictable non-decreasing process Λ such that

$$E \left[\int_{\mathbb{R}_+} Y dN \right] = E \left[\int_{\mathbb{R}_+} Y d\Lambda \right] \quad (3)$$

for all non-negative $\mathfrak{F}$ -predictable processes Y (see Daley and Vere-Jones 2008, Proposition 14.2.II).

If Λ is absolutely continuous, we write

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

for some $\mathfrak{F}$ -predictable process $\lambda = (\lambda_1, \dots, \lambda_{d_E})$, which is called intensity of the counting process N . Combining this with Equation (3), one obtains the formula

$$E[N_e(t) - N_e(s) | \mathfrak{F}_s] = E \left[\int_s^t \lambda_e(u) du | \mathfrak{F}_s \right], \quad s \leq t,$$

which allows to interpret $\lambda_e(t)$ as a measure of the ‘instantaneous risk’ of a jump at time t in the e th component of the counting process N . Notice that this ‘risk’ evolves in time and it varies depending on the information available up to time s .

Compensators are crucial in the following time-change result, which will be used to perform goodness-of-fit diagnostics (see Section 6.2).

Theorem 3.3 (Meyer 1971): Let N be a d_E -dimensional counting process with arrival times T^e . Assume that N has continuous compensator Λ such that $\Lambda_e(t) \rightarrow \infty$ as $t \rightarrow \infty$ for all $e = 1, \dots, d_E$. Then, the random sequences $\{\Lambda(T_j^e) : j = 1, 2, \dots\}$, $e = 1, \dots, d_E$ are the arrival times of a d_E -dimensional unit-rate Poisson process, namely the time-changed inter-arrival times

$$\tau_e^j := \Lambda(T_j^e) - \Lambda(T_{j-1}^e) \quad (4)$$

are all independent exponentially distributed random variables for $j = 1, 2, \dots$ and $e = 1, \dots, d_E$.

For a proof of Theorem 3.3, see Brown and Nair (1988).

3.2. Multidimensional Hawkes Processes

In this section, we recall the basics of the theory of state-dependent Hawkes processes from Morariu-Patrichi and Pakkanen (2018, 2022).

Definition 3.4: A d_E -dimensional counting process N is called a Hawkes process if it admits an absolutely continuous compensator Λ with intensities

$$\lambda_e(t) = \nu_e + \sum_{e'=1}^{d_E} \int_0^t \kappa_{e',e}(t-s) dN_{e'}(s), \quad e = 1, \dots, d_E, \quad (5)$$

for some non-negative base rates $\nu_e \geq 0$, and some non-negative locally integrable functions $\kappa_{e',e} \geq 0$ that are supported on the non-negative half line.

The matrix-valued function $t \mapsto [\kappa_{e',e}(t)]_{e,e'=1,\dots,d_E}$ is referred to as the kernel of the Hawkes process N . If all the kernel functions are integrable, the spectral radius ρ of the $d_E \times d_E$ -matrix of L^1 norms $\|\kappa_{e',e}\|_1$ is called radius of the Hawkes kernel; if some of the kernel functions are not integrable, the spectral radius is set to $+\infty$.

A d_E -dimensional Hawkes process is asymptotically stationary if the radius of its kernel is smaller than 1; in this case the intensity process λ is asymptotically stationary.

Let $\mathcal{S}$ be a finite state space. We can label its elements as $x = 1, \dots, d_S$, where d_S is the number of possible states of the system. A state-dependent counting process is a pair (N, X) , where for all t , $N(t)$ records the number of events occurred by time t as per formula (2), and $X(t)$ records the state of the system at time t . More specifically, we have:

Definition 3.5 (Morariu-Patrichi and Pakkanen 2022, Definition 2.1): Let N be a d_E -dimensional counting process. Let X be a continuous-time piecewise-constant process in the finite state space $\mathcal{S}$ of cardinality d_S . Let $\mathfrak{F}$ be the minimal complete right-continuous filtration generated by the pair (N, X) . Then, we say that (N, X) is a state-dependent Hawkes process if

- (i) N admits an absolutely continuous $\mathfrak{F}$ -compensator with intensities

$$\lambda_e(t) = \nu_e + \sum_{e'=1}^{d_E} \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s), \quad e = 1, \dots, d_E, \quad (6)$$

for some d_E non-negative base rates $\nu_e \geq 0$, $e = 1, \dots, d_E$, and some d_E^2 measurable functions $\kappa_{e',e} : \mathbb{R}_+ \times \mathcal{S} \rightarrow \mathbb{R}_+$, $e, e' = 1, \dots, d_E$, such that $\kappa_{e',e}(\cdot, x)$ is locally integrable for all x in $\mathcal{S}$;

- (ii) X jumps only at arrival times T_n of N , and there exist d_E transition matrices $\phi_e(\cdot, \cdot)$, $e = 1, \dots, d_E$, defined on $\mathcal{S}$ such that for all n

$$P(X(T_n) = x | E_n, \mathfrak{F}_{T_n-}) = \phi_{E_n}(X(T_n-), x), \quad x = 1, \dots, d_S, \quad (7)$$

where $X(T_n-) = \lim_{t \uparrow T_n} X(t)$ is the state of the system immediately before the n th event E_n , and $\mathfrak{F}_{T_n-} = \bigvee_{\epsilon > 0} \mathfrak{F}_{T_n-\epsilon}$ represents the information available immediately before this event.

Given a state-dependent Hawkes process (N, X) , let T_n and E_n be the sequences of arrival times and events that equivalently describe the counting process component N of the pair (N, X) , as per Equation (2). Let X_n be the sequence of states $X(T_n)$, for $n = 1, 2, \dots$. Then, the $d_E d_S$ -dimensional counting process

$$\tilde{N}_{e,x}(t) := \sum_n \mathbb{1} \{T_n \leq t, E_n = e, X_n = x\} \quad (8)$$

is called the hybrid-MPP counterpart of (N, X) . We have that the j th jump time T_j^e of the e th component of N is the j th order statistic of $\{T_k^{e,x} : k = 1, 2, \dots; x = 1, \dots, d_S\}$, where $(T_k^{e,x})_k$ are the jump times of the (e, x) th component of $\tilde{N}$. Similarly, T_n is the n th order statistics of $\{T_k^{e,x} : k = 1, 2, \dots; e = 1, \dots, d_E; x = 1, \dots, d_S\}$. The (e, x) th component $\tilde{N}_{e,x}$ of the hybrid-MPP counterpart of (N, X) admits a continuous compensator with density given by

$$\tilde{\lambda}_{e,x}(t) = \phi_e(X(t), x) \left(\nu_e + \sum_{e', x'} \int_{[0,t)} \kappa_{e',e}(t-s, x') d\tilde{N}_{e',x'}(s) \right), \quad (9)$$

where ϕ_e is the transition matrix associated with event type e , and $\kappa_{e',e}$, for $e, e' = 1, \dots, d_E$, are the Hawkes kernels of N .

Let $\bar{\lambda} = \lambda_1 + \dots + \lambda_{d_E}$ be the sum of the intensities. If $\bar{\lambda}$ is decreasing in time, then a state-depended Hawkes process (T_n, E_n, X_n) can be simulated as detailed in Algorithm 3.1.

Algorithm 3.1 [Algorithm 2.4]MP18sta

Require: $(T_i, E_i, X_i)_{i=1, \dots, n-1}$

set $t := T_{n-1}$

set $\xi := 0$

while $\xi = 0$ **do**

draw $U \sim \text{Exp}(\bar{\lambda}(t))$

5: set $\xi := 1$ with probability $\bar{\lambda}(t+U)/\bar{\lambda}(t)$

update $t \leftarrow t + U$

end while

set $T_n := t$

draw E_n in $\{1, \dots, d_E\}$ with probabilities proportional to $\{\lambda_1(T_n), \dots, \lambda_{d_E}(T_n)\}$

10: draw X_n in $\{1, \dots, d_S\}$ with probabilities $\{\phi_{E_n}(X_{n-1}, 1), \dots, \phi_{E_n}(X_{n-1}, d_S)\}$

return (T_n, E_n, X_n)

4. State-Dependent Hawkes Model

We consider four streams of random times: the stream $(T_j^1)_j$ of times when limit orders are executed on the bid side (equivalently identified with the arrival times of sell market orders); the stream $(T_j^2)_j$ of times when limit orders are executed on the ask side (equivalently identified with arrival times of buy market orders); the stream $(T_j^3)_j$ of times when

either an ask limit order is inserted inside the spread, or the cancellation of a bid limit order depletes the liquidity available at the first bid level; the stream $(T_j^4)_j$ of times when either a bid limit order is inserted inside the spread, or the cancellation of an ask limit order depletes the liquidity available at the first ask level.

The four sequences of random times give rise to a four-dimensional counting process $N = (N_1, \dots, N_4)$ with the following interpretation of its components:

- (i) $N_1(t)$ denotes the number of seller-initiated trades that happened before or at time t (identified with the number of market orders arrived on the bid side of the order book by time t);
- (ii) $N_2(t)$ denotes the number of buyer-initiated trades that happened before or at time t (identified with the number of market orders arrived on the ask side of the order book by time t);
- (iii) $N_3(t)$ denotes the number of decreases in the mid-price caused by a limit order insertion or cancellation that happened before or at time t ;
- (iv) $N_4(t)$ denotes the number of increases in the mid-price caused by a limit order insertion or cancellation that happened before or at time t .

The counting process N is paired with the state variable X . At time t , the state variable $X(t)$ summarizes the configuration $(P_t^b, P_t^a, \{(V_t^{b,i}, V_t^{a,i}) : i = 1, 2, \dots\})$ of the limit order book at time t , by recording a proxy for the n -levels volume imbalance, and the variation of the mid-price compared to time $t-$. More precisely,

$$X(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} := \begin{pmatrix} \mathbb{1}\{\delta P^m(\hat{T}(t)) > 0\} - \mathbb{1}\{\delta P^m(\hat{T}(t)) < 0\} \\ \frac{1}{2} \sum_{k=0}^{K-1} (2k - K + 1) \mathbb{1}\left\{\frac{k-K}{K} \leq I_t^n < \frac{2(k+1)-K}{K}\right\} \end{pmatrix}, \quad (10)$$

where $\delta P^m(t) = \lim_{\epsilon \downarrow 0} (P^m(t) - P^m(t - \epsilon))$, $\hat{T}(t) = \sup\{T_j^e \leq t : e = 1, \dots, 4; j = 1, \dots\}$, and I_t^n is defined in (1).

The first component X_1 of the state variable X can take the values $-1, 0, +1$, respectively denoting downward jump in the mid-price, unchanged mid-price and upward jump in the mid-price. The second component X_2 of the state variable X is a discretisation of the n -levels queue imbalance I_t^n , and – assuming that K is odd – it takes integer values from $-(K-1)/2$ to $(K-1)/2$, spanning the full range of possible values of I^n from -1 to $+1$.

It follows from the definition of X_2 that if at time t we have that $X_2(t) = x_2$, then the n -levels queue imbalance I_t^n at time t must be in the half-open interval $[(2x_2 - K - 1)/2K, (2x_2 + 1)/K)$. Notice that X_2 depends on the two additional parameters n and K : the former is the number n of levels of the limit order books taken into account in the computation of the queue imbalance I^n ; the latter is the number K of points in the partition of the interval $[-1, 1]$ used for the discretisation of I^n .

The pair (N, X) is modelled as a state-dependent Hawkes process, hence we assume that there are base rates v_e , Hawkes kernels $\kappa_{e',e} = \kappa_{e',e}(t, x')$ and transition matrices ϕ_e such that Definition 3.5 is satisfied. The number of event types is $d_E = 4$ and the number of states is $d_S = 3K$.

When a new event occurs, i.e., when one of the components N_e of N jumps, the state variable X is updated as per in Equation (7). The update models the mechanism whereby trades on either side of the limit order book can trigger changes in the

mid-price and in the queue imbalance. Indeed, assume that a sell (resp., buy) market order arrives at time T_j^1 (resp., T_j^2), and that $X(T_j^1-) = (x_1, x_2)$ (resp., $X(T_j^2-) = (x_1, x_2)$) for some x_1 in $\{-1, 0, +1\}$ and some x_2 in $\{(1-K)/2, (3-K)/2, \dots, (K-1)/2\}$. Then, the mid-price jumps downward (resp., upward) with probability $p_- := \sum_{y_2=(1-K)/2}^{(K-1)/2} \phi_1((x_1, x_2), (-1, y_2))$ (resp., $p_+ := \sum_{y_2=(1-K)/2}^{(K-1)/2} \phi_2((x_1, x_2), (+1, y_2))$), and it remains unchanged with probability $p_0 := 1 - p_- = \sum_{y_2=(1-K)/2}^{(K-1)/2} \phi_1((x_1, x_2), (0, y_2))$ (resp., $p_0 := 1 - p_+ = \sum_{y_2=(1-K)/2}^{(K-1)/2} \phi_2((x_1, x_2), (0, y_2))$).² This jump of the state variable happens exactly at the arrival time T_j^1 (resp., T_j^2) of the sell (resp., buy) market order, and it naturally captures the mechanism responsible for the market-order-induced price change: p_- (resp., p_+) represents the probability that a sell (resp., buy) market order walks the book given its submission, and p_0 represents the probability that it does not. Notice that p_- (resp., p_+) and p_0 depend on the state of the limit order book immediately before the arrival of the sell (resp., buy) market order, and in particular they depend on x_2 . This is a granular description of the order book mechanism, and it accounts for the fact that it is less likely that a sell (resp., buy) market order walks the book when the volumes on the bid (resp., ask) side are high, namely $p_-(x_1, x_2) \leq p_-(x_1, \tilde{x}_2)$ if $x_2 \geq \tilde{x}_2$ (resp., $p_+(x_1, x_2) \leq p_+(x_1, \tilde{x}_2)$ if $x_2 \leq \tilde{x}_2$).

The first component X_1 of the state variable X enables to write the following proxy for the mid-price:

$$P_0^m + \frac{\psi}{2} \int_0^t X_1(s) dN_g(s), \quad (11)$$

where ψ is the tick size of the limit order book and $N_g = N_1 + \dots + N_4$ is the ground process of N .

Remark 4.1: Our model can be compared to that of Bacry and Muzy (2014). In their model, four streams of random times are considered: the stream $(T_j^1)_j$ of times when limit orders are executed on the bid side (equivalently identified with the arrival times of sell market orders); the stream $(T_j^2)_j$ of times when limit orders are executed on the ask side (equivalently identified with arrival times of buy market orders); the stream $(T_j^3)_j$ of times when the mid-price decreases; the stream $(T_j^4)_j$ of times when the mid-price increases. T^1 and T^2 are as in our model, whereas T^3 and T^4 represent what in our model we represent through the state variable X_1 . In Bacry and Muzy (2014), the four-dimensional counting process $N = N(t)$ associated with T^1 , T^2 , T^3 and T^4 is assumed to be a four-dimensional ordinary Hawkes process. In their model, a buy (resp., sell) market order coming into the exchange and walking the book at time t would be represented by the equation $\delta N_2(t) = \delta N_4(t) = 1$ (resp., $\delta N_1(t) = \delta N_3(t) = 1$). However, the components of a multidimensional Hawkes process jump simultaneously with probability zero. In other words, if $\{(T_j^e)_j : e = 1, \dots, 4\}$ are the arrival times associated with a 4-dimensional Hawkes process, it holds $P(T_j^e = T_{j'}^{e'} \text{, for some } j, j' \geq 1, \text{ and } e \neq e') = 0$. Since the direction of the causality is unambiguous (a market order originates first and as a result of its execution the mid-price jumps), Bacry and Muzy (2014) propose to add to the Hawkes kernels $\kappa_{1,3}$ and $\kappa_{2,4}$ an atomic component. This is the defining feature of the ‘impulsive impact

kernel' – see Bacry and Muzy (2014, Section 2.1.3). In our paper, the usage of the state variable X_1 circumvents the need of these atomic components and naturally accommodates mid-price changes triggered by market orders walking the book.

The second component X_2 of the state variable X reproduces the state variable of the queue-imbalance model in Morariu-Patrichi and Pakkanen (2022). It is conceived as the main indicator of the regime in which limit and market orders will arrive to the exchange: in high-frequency markets trading algorithms send their orders in response to observable quantities of the limit order book configuration, and a prominent one is indeed the queue imbalance. It is therefore expected that when X_2 is positive (resp., negative), the intensities of events of types $e = 2$ (resp., $e = 1$) will be higher, because market participants following the queue imbalance signal will expect the price to increase (resp., decrease). After the price change, the volumes of deeper queues on the ask (resp., bid) side enter the computation of the queue imbalance, and this will likely reset the signal. As noted in Morariu-Patrichi and Pakkanen (2022), this interaction can be deemed responsible for the mean-reverting behaviour of price dynamics in high-frequency markets.

Moreover, we use X_2 to reproduce the update of the limit order book configuration that happens when a labelled agent submits their market orders. Indeed, we consider normalized volumes up to level n , namely we assume that $\sum_{i=1}^n (V_t^{b,i} + V_t^{a,i}) \equiv 1$, and we assume that the $2n$ -tuple $(V_t^{a,1}, V_t^{b,1}, \dots, V_t^{a,n}, V_t^{b,n})$ is distributed as a Dirichlet random variable with $2n$ -dimensional parameter $\gamma = \gamma(X(t)) \in \mathbb{R}_+^{2n}$ that depends on the state variable at time t .

Given the time evolution of the limit order book in the time window $[0, T]$, an estimator for $\gamma(x)$, with x ranging from 1 to $3K$, can be obtained by maximum likelihood estimation. Once γ is known, the order book mechanics can be reproduced by drawing from the conditional distribution $\text{Dir}_{\gamma(X(t))}(\cdot | X_2(t))$. This is the Dirichlet distribution of the $2n$ -tuple $(V_t^{a,1}, V_t^{b,1}, \dots, V_t^{a,n}, V_t^{b,n})$ with parameter $\gamma(X(t))$ conditioned on

$$\frac{2X_2(t) - K - 1}{2K} \leq \underbrace{\sum_{i=1}^n (V_t^{b,i} - V_t^{a,i})}_{=I_t^n} < \frac{2X_2(t) + 1}{K}.$$

Algorithm 4.1 describes how to reproduce the order book update in the case of the arrival of a sell market order $(t, q_M, 0, -1)$. The case of buy market orders is analogous.

Line 4 in Algorithm 4.1 says that the bid price (and consequently the mid-price) decreases if the size of the sell market order is larger than the available liquidity sitting on the first bid level. Lines 6:10 cancel (from the bid queues) the orders whose execution has been triggered by the arrival of $(t, q_M, 0, -1)$. On line 7 we used the notation $\min_+(a, b) = \max(0, \min(a, b))$ for a and b real numbers.

5. Price Impact Profiles

Measuring price impact requires two things. The first is to modify the model (N, X) of Section 4 in a way to account for a labelled agent, whose impact we wish to measure. The second is to extrapolate to which extent the labelled agent is responsible for the evolution

Algorithm 4.1 State update via order book mechanics (sell market order)

Require: $X(t-), (t, q_M, 0, -1)$

set $\ell := (2X_2(t-) - K - 1)/(2K)$

set $u := (2X_2(t-) + 1)/K$

sample $V_{t-} = (V_{t-}^{a,1}, V_{t-}^{b,1}, \dots, V_{t-}^{a,n}, V_{t-}^{b,n}) \sim \text{Dir}_{\gamma(X(t-))}(\cdot | X_2(t-))$

set $X_1(t) := -1$ if $q_M \geq V_{t-}^{b,1}$; 0 otherwise

5: initialize $q := 0$

for i in $1, \dots, n$ **do**

 set $v := \min_+(V_{t-}^{b,i}, q_M - q)$

 set $V_t^{b,i} := V_{t-}^{b,i} - v$

 update $q \leftarrow q + v$

10: **end for**

set $B := \sum_{i=1}^n V_t^{b,i}$

set $A := \sum_{i=1}^n V_{t-}^{a,i}$

set $I := (B - A)/(B + A)$

set $X_2(t) := (2k - K + 1)/2$ if $(k - K)/K \leq I < (2k + 2 - K)/K$, where

$k = 0, \dots, K - 1$

15: **return** $(X_1(t), X_2(t))$

of the price dynamics that emerge from the state process $(X(t))_t$. Section 5.1 describes the former; Section 5.2 describes the latter.

5.1. Labelled Agent

We account for a labelled agent in the market, and we aim to measure their impact on the dynamics of the order book. We take the perspective of a liquidation, namely we consider our agent (also referred to as liquidator) to be selling the amount Q_0 of asset. The case of acquisition is mutatis mutandis the same.

We let $[0, T]$ represent the time window of the liquidation. The quantity Q_0 is referred to as the size of the liquidator's metaorder, or their initial inventory, and we normalize it with respect to the overall volume $\sum_{i=1}^n (V_0^{a,i} + V_0^{b,i})$ of offers sitting on the first n levels of the order book at the start of the liquidation window.

We assume that the liquidator intervenes in the market only by sending sell market orders; they will never place a limit order to queue on the ask side, but they will initiate trades with existing offers on the bid side.

Hence, the liquidation is described by the sequence $\{(T_j^0, q_{M,j}, 0, -1) : j = 1, 2, \dots\}$ of sell market orders sent by the liquidator. For every j , T_j^0 is the time stamp of the liquidator's j th child market order, and $q_{M,j}$ is its size.

We suppose that the stream of random times $T_1^0 < T_2^0 < \dots$ is restricted to $[0, T]$. We assume non-explosiveness, so that the number of liquidator's market orders is finite if the time horizon T of the execution window is not $+\infty$. Moreover, we let $t_0 = T_1^0$ represent the time at which the liquidator begins their intervention in the market, and we let $\tau := \sup\{T_j^0 \leq T : j = 1, 2, \dots\}$ be the time at which the liquidation stops.

Assumption 5.1: Let Q_0 be the size of the liquidator's metaorder, and for j in $\mathbb{N}$, let $z_j = q_{M,1} + \dots + q_{M,j}$ be the sum of all liquidity-normalized sizes of the first j child market orders sent by the liquidator. Then, the termination time τ of the liquidation is assumed to coincide with the smallest time stamp T_j^0 among the liquidator's market orders such that $z_j \geq Q_0$, namely

$$\tau = \inf \left\{ T_j^0 \geq t_0 : \sum_{k=1}^j q_{M,k} \geq Q_0 \right\}.$$

We introduce the liquidator's presence in the model described in Section 4 by expanding the dimension of the counting process N : we let the zero-th component $N_0(t)$ count the liquidator's market orders sent to the exchange by time t . In other words, from the overall sequence $(T_j^1)_j$ of arrival times of market orders described in Section 4, we extract those sent by the liquidator and we label them as $(T_j^0)_j$; we then let

$$N_0(t) := \sum_{j \geq 1} \mathbb{1} \left\{ T_j^0 \leq t \right\}$$

count the number of trades initiated by the liquidator that happened before or at time t . Notice that the map $t \mapsto N_0(t)$ represents how the liquidator is splitting in time the execution of their metaorder. In other words, this is the liquidator's execution schedule.

The pair (N, X) is a state-dependent Hawkes process where the counting process component N is five-dimensional, and the state process X is as in Equation (10). The event types will be labelled $e = 0, 1, \dots, 4$ and the states will be labelled $x = 1, \dots, 3K$ or $x = (x_1, x_2)$ with $x_1 = -1, 0, +1$ and $x_2 = -(K-1)/2, \dots, +(K+1)/2$. The following assumption is in place on the intensities.

Assumption 5.2: For all $e = 1, \dots, 4$, the Hawkes kernel $\kappa_{0,e}$ coincides with $\kappa_{1,e}$.

Assumption 5.2 guarantees consistency in the effect that trades have on the order book dynamics. It says that the rates of arrival of market orders to the exchange are modified by the liquidator's interventions in the same way as they are by other participants' sell market orders. More precisely, for $e = 1, \dots, 4$ it holds

$$\begin{aligned} \lambda_e(t) &= v_e + \sum_{e'=0}^4 \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s) \\ &= v_e + \sum_{e'=1}^4 \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s) + \int_{[0,t)} \kappa_{1,e}(t-s, X(s)) dN_0(s). \end{aligned} \quad (12)$$

The liquidator's execution schedule admits an absolutely continuous compensator Λ_0 with density

$$\lambda_0(t) = v_0 \mathbb{1}_{[0,\tau)}(t) + \sum_{e',x'} \mathbb{1}_{[0,\tau)}(t) \int_{[0,t)} \kappa_{e',0}(t-s, x') d\tilde{N}_{e',x'}(s). \quad (13)$$

For $j = 1, 2, \dots$ let $(T_j^0, q_{M,j}, 0, -1)$ be the liquidator's child market orders, as denoted above. The liquidator's order scheduling depends on the Hawkes parameters v_0 , and $\kappa_{e',0}$,

which modulate the sequence of arrival times T_j^0 . Additionally, the liquidation depends on the size $q_{M,j}$ of the j th child market order, for $j = 1, 2, \dots$ ³ The evolution of the limit order book is simulated by combining Algorithms 3.1 and 4.1, as detailed in Algorithm 5.1.

Algorithm 5.1 Simulation of order book in the presence of liquidator

Require: $(T_i, E_i, X_i)_{i=1, \dots, n-1}, q_M$
do Lines 1:8 of Algorithm 3.1
draw E_n with probabilities proportional to $\{\lambda_0(T_n), \lambda_1(T_n), \dots, \lambda_4(T_n)\}$
if $E_n = 0$ **then**
set $X_n \leftarrow \text{Algorithm 4.1}(X_{n-1}, (T_n, q_M, 0, -1))$
5: **else**
draw X_n in $\{1, \dots, 3K\}$ with probabilities $\{\phi_{E_n}(X_{n-1}, 1), \dots, \phi_{E_n}(X_{n-1}, 3K)\}$
end if
return (T_n, E_n, X_n)

Remark 5.3: In Bacry and Muzy (2014) (see Remark 4.1), a labelled agent is accounted for by considering the following intensities of the four-dimensional counting process N . For $e = 1, \dots, 4$ and $t > 0$ they set

$$\lambda_e(t) = \nu_e + \sum_{e'=1}^4 \int_{[0,t)} \kappa_{e',e}(t-s) dN_{e'}(s) + \int_{[0,t)} \theta_e(s) dA(s), \quad (14)$$

where $t \mapsto A(t)$ is the liquidator's execution schedule, $t \mapsto \theta_1(t)$ (resp., $t \mapsto \theta_2(t)$) represents the impact of the liquidator's market orders on the arrival of other participants' sell (resp., buy) market orders, and $t \mapsto \theta_3(t)$ (resp., $t \mapsto \theta_4(t)$) represents the impact of the liquidator's market orders on downward (resp., upward) jumps of the mid-price. In their model, to have consistency between the liquidator and other market participants one needs to impose $\theta_3(t) = \kappa_{1,3}(t)$ and $\theta_4(t) = \kappa_{1,4}(t)$ for $t \geq 0$. Practically, this implies that the atomic components in the Hawkes kernel are passed to the integrands θ_3 and θ_4 , which means that the liquidator walks the book at an average rate equal to the overall proportion of markets orders walking the book. The consequence that the liquidator walks the book in this way can be a potentially undesirable feature.⁴ In our model, there is no need for this to be assumed. We are able to test executions without this assumption, and still maintain consistency between the liquidator and other market participants.

Remark 5.4: The liquidator's interventions in the market have been modelled by expanding one component of the Hawkes process introduced in Section 4. The justification for this modelling choice is twofold. First, this guarantees consistency between executions of sell market orders sent by the liquidator and executions of sell market orders sent by other market participants. Given our interest in understanding the liquidator's impact, any other stochastic model would raise questions of granting the liquidator with a privileged order scheduling. Second, expanding the dimensions of N allows us to give a natural justification to Bacry and Muzy (2014)'s formula for the intensity (14) – $\kappa_{0,e}$ takes the role of θ_e and dN_0

takes the role of dA . Hence, our modelling choice resonates with existing models in the literature, and it is grounded in our phenomenological point of view. A future work could adopt the point of view of optimal execution, and optimize the liquidator's scheduling in a set of admissible liquidation strategies aimed at minimizing their price impact.

5.2. Definition of Price Impact

We partition the state space $\mathcal{S}$ according to the values of the first component X_1 of the state variable $X = (X_1, X_2)$. We define

$$\mathcal{S}^{x_1} := \{y = (y_1, y_2) \in \mathcal{S} : y_1 = x_1\}. \quad (15)$$

We refer to states x in $\mathcal{S}^+$ (resp., in $\mathcal{S}^-$) as inflationary (resp., deflationary) states.

The jump times for the mid-price consequently give rise to the counting processes

$$N^{x_1}(t) := \sum_n \mathbb{1} \{T_n \leq t, X_1(T_n) = x_1\}, \quad x_1 \in \{-1, 0, +1\}, \quad (16)$$

where T_n is the n th jump time of the ground process. The difference $N^+(t) - N^-(t)$ is a proxy for the mid-price in the order book. Indeed, we can rewrite the integral quantity in Equation (11) as

$$\int_0^t X_1(s) dN_{\mathbf{g}}(s) = N^+(t) - N^-(t).$$

Definition 5.5: The state-dependent Hawkes model (N, X) is said price-symmetric if for all $t \geq 0$

$$\left(\sum_{x \in \mathcal{S}^+} - \sum_{x \in \mathcal{S}^-} \right) \sum_{e=1}^4 \phi_e(X(t), x) \ell_e(t) = 0,$$

where

$$\ell_e(t) = v_e + \sum_{e'=1}^4 \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s).$$

Proposition 5.6: Assume that there exist a permutation σ_E of $\{1, \dots, 4\}$ and a bijective map $\sigma_S : \mathcal{S}^+ \rightarrow \mathcal{S}^-$ such that

- (i) $\phi_e(y, x) = \phi_{\sigma_E(e)}(y, \sigma_S(x))$ for all y in $\mathcal{S}$, all x in $\mathcal{S}^+$ and all $e = 1, \dots, 4$;
- (ii) $v_e = v_{\sigma_E(e)}$ for all $e = 1, \dots, 4$;
- (iii) $\kappa_{e',e}(t, x') = \kappa_{e',\sigma_E(e)}(t, x')$ for all x' in $\mathcal{S}$ all $e, e' = 1, \dots, 4$ and all $t \geq 0$.

Then, (N, X) is price-symmetric.

Remark 5.7: The condition in Proposition 5.6(i) captures the idea that, given the current state y , transitions to inflationary states and transitions to deflationary states are equally likely. The conditions in Proposition 5.6(iii) capture the idea that, given the current state y ,

every event-state pair (e', x') excites an event-state pair (e, x) with inflationary state $x \in \mathcal{S}^+$ the same way as it excites an event-state pair $(\sigma_E(e), \sigma_S(x))$ with deflationary state $\sigma_S(x) \in \mathcal{S}^-$; in other words, the offspring from every event-state pair (e', x') are equally likely to be associated with inflationary states or with deflationary states.

Definition 5.8: Let t_0 be the time when the liquidator becomes active in the market. Then, the price impact profile of the execution schedule N_0 is the primitive of $t \mapsto \text{Dir}(t) + \text{Indir}(t)$ pinned at 0 in t_0 , where

$$\text{Dir}(t) = \sum_{x \in \mathcal{S}^-} \phi_0(X(t), x) \left(\nu_0 + \sum_{e'=1}^4 \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{e',0}(t-s, x') d\tilde{N}_{e',x'} \right) \mathbb{1}_{[0,\tau)}(t),$$

where τ is the termination time of the liquidation,

$$\phi_0(x', x) = \frac{\sum_j \mathbb{1}\{X(T_j^0-) = x', X(T_j^0) = x\}}{\sum_j \mathbb{1}\{X(T_j^0-) = x'\}}, \quad (17)$$

and

$$\text{Indir}(t) = \sum_{e=1}^4 \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{1,e}(t-s, x') d\tilde{N}_{0,x'} \left(\sum_{x \in \mathcal{S}^-} - \sum_{x \in \mathcal{S}^+} \right) \phi_e(X(t), x).$$

The map $t \mapsto \text{Dir}(t) + \text{Indir}(t)$ is referred to as intensity of the price impact profile.

Remark 5.9: In Definition 5.8 we defined price impact as the time integral of intensities of counting processes. These counting processes count the increases and the decreases of the mid-price. The tick size is set by the trading venue to 0.01 USD, and so the minimum change in the mid-price is 0.005 USD. Therefore, in the examples below, the physical dimension of price impact is 0.005 USD.

Remark 5.10: The intensity of the price impact profile is decomposed in two components, namely $\text{Dir}(t)$ and $\text{Indir}(t)$. Both are null if $N_0(t) \equiv 0$. The former is referred to as ‘direct’ impact and stems from those summands of the execution schedule’s intensity $\lambda_0(t) = \sum_{x=1}^{3K} \tilde{\lambda}_{0,x}(t)$ that are associated with deflationary states, namely $\text{Dir}(t) = \sum_{x \in \mathcal{S}^-} \tilde{\lambda}_{0,x}(t)$. Notice that $\text{Dir}(t) \geq 0$ for all $t > 0$ and $\text{Dir}(t) = 0$ for all $t > \tau$. On the contrary, the second term $\text{Indir}(t)$ stems from events originated by participants other than the liquidator but in response to the liquidator’s interventions, hence the name of ‘indirect’ impact. It can have either sign and it is in general non-zero even beyond the termination time; for this reason it is linked to the transient impact. More precisely, for $t > \tau$ it holds $\text{Indir}(t) = \sum_{e=1}^4 \sum_j \kappa_{1,e}(t - T_j^0, X(T_j^0)) (\sum_{x \in \mathcal{S}^-} - \sum_{x \in \mathcal{S}^+}) \phi_e(X(t), x)$, and the transient price impact profile is the map $t \mapsto \int_0^t \text{Indir}(s) ds$, restricted to the interval $t \geq \tau$.

In a price-symmetric state-dependent Hawkes model, if $N_0 \equiv 0$, then $N^- - N^+$ is a martingale, and its compensator is identically null. Instead, when the liquidator is active in the market, the symmetry is disrupted, and we map this disruption to our measure of the price impact.

Hence, Definition 5.8 is vindicated by the following proposition.

Proposition 5.11: *If (N, X) is price-symmetric, then the price impact profile of N_0 is the $\mathfrak{F}$ -compensator of $N^- - N^+$, where $\mathfrak{F}$ is the minimal complete right-continuous filtration to which (N, X) is adapted.*

The direct impact component $\text{Dir}(t)$ of the intensity $(\lambda^- - \lambda^+)(t)$ encompasses the transition matrix ϕ_0 associated with the state update that occurs when liquidator's orders are executed. For x' and x in $\mathcal{S}$, $\phi_0(x', x)$ is estimated according to Equation (17); hence it summarizes the state transitions that stem from Algorithm 4.1 during the simulation of the execution. This disentangles the effects of liquidator's orders (whose sizes $q_{M,j}$ are set by the liquidator) from the effects of other market orders, i.e., $\phi_0 \neq \phi_1$ in general, allowing to investigate the impact of different execution strategies.

In particular, the liquidator might choose to send market orders with sizes that never exceed the available liquidity on the first bid level; this would cause $\phi_0(x', x) = 0$ for all x' in $\mathcal{S}$ and all deflationary x in $\mathcal{S}^-$, and thus $\text{Dir}(t) \equiv 0$. Nonetheless, the overall impact would not be null, because of the indirect term $\text{Indir}(t)$. Indeed, even without ever walking the book, the liquidator's orders would modify (i) the arrival of orders submitted by other market participants who react to the liquidator's executions; (ii) the volumes in the order book.

As far as (i) is concerned, if the dynamics of order submission is such that deflationary events trigger other events with deflationary effects on the price, then the price may plunge as an indirect consequence of the liquidation.

As far as (ii) is concerned, despite the fact that they do not walk the book, liquidator's executions consume liquidity on the bid side, pushing the state trajectory $t \mapsto X(t)$ to dwell in states $\{y = (y_1, y_2) \in \mathcal{S} : y_2 < 0\}$ for longer. The probability of transitioning from these states to deflationary states is higher than the probability of transitioning to inflationary states, hence making the term $(\sum_{x \in \mathcal{S}^-} - \sum_{x \in \mathcal{S}^+})\phi_e(X(t), x)$ positive, and contributing to the impact via the indirect term $\text{Indir}(t)$. Notice that this form of impact would not be captured by a less granular model where the update of the volumes in the limit order book is not reproduced as we do in Algorithm 4.1, and where the liquidator's child orders are assumed to walk the book at an average rate equal to the overall proportion of market orders that do so – see Remark 5.3.

In Section 6, we will see that, after calibrating our model on empirical data from NASDAQ, the indirect component of the price impact is actually the main driver of price impact during liquidation.

6. Applications

6.1. Description of the Dataset and Model Specifications

We study order book data provided by LOBSTER.⁵ LOBSTER is a provider of high-quality limit order book data that is reconstructed from NASDAQ's Historical TotalView-ITCH⁶ files with detailed event information. The reconstruction methodology is described in Huang and Polak (2011).

For every NASDAQ ticker and every active trading day, LOBSTER provides two files in .csv format: a 'message file' and an 'orderbook' file. The former is an event-by-event

Table 1. LOBSTER labels of order book events.

1:	Submission of a new limit order
2:	Cancellation (partial deletion of a limit order)
3:	Deletion (total deletion of a limit order)
4:	Execution of a visible limit order
5:	Execution of a hidden limit order
6:	Indicates a cross trade, e.g., auction trade
7:	Trading halt indicator

Table 2. Mapping of LOBSTER labels to event types.

LOBSTER event label	Bid/Ask	Event type
1	ask	3
1	bid	4
2	ask	4
2	bid	3
3	ask	4
3	bid	3
4	ask	2
4	bid	1
5	ask	2
5	bid	1

history of messages sent to the exchange that provoked an update in the configuration of the order book. The latter is an event-by-event snapshot of the order book, where the n th row corresponds to the configuration resulting from the n th message reported in the message file.

Prices are reported in 10^{-4} USD; hence the tick size, imposed by regulation⁷ and equal for all shares with price above 1USD, is set to 100. Time stamps are reported in seconds after midnight with resolution at the nanosecond scale. In the plots that follow prices are always reported in 10^{-4} USD and times are reported in seconds. Events happening in the trading venue are labelled according to Table 1.⁸

Table 2 shows how we map LOBSTER order book labels to the sequences of arrival times described in Section 4.

In the analysis that follows, we study order book data for the ticker INTC trading on 25 January 2019. First, we calibrate our state-dependent Hawkes model on the dataset of 25 January 2019; then, we simulate liquidations of a large number of shares using Algorithm 5.1; and finally we assess the price impact of such simulated liquidations as per Definition 5.8. At the end of the section we make remarks about the sensitivity of our results with respect to calibrated parameters, and we provide insights on the implementation for other dates and tickers.

6.2. Calibration

After filtering for the arrival times $(T_j^e)_j$, $e = 1, \dots, 4$, and defining the state variable $X = (X_1, X_2)$ as per Equation (10) with $n = 2$ and $K = 3$, the data sets of message file and order book for INTC on 25 January 2019 are as in Table 3.

Starting from these data sets we perform maximum likelihood estimation of our state-dependent Hawkes model.

Table 3. Ten time stamps from the filtered message file and order book file.

Time	Bid/Ask	LOBSTER event label	Event type	price	size
35400.092452	ask	4	2	460700	300
35400.092533	ask	1	3	460700	200
35400.092768	ask	4	2	460700	100
35400.113748	ask	4	2	460700	400
35400.113776	bid	1	4	460700	100
35400.121175	bid	4	1	460700	253
35400.121258	ask	1	3	460700	3300
35400.123294	ask	4	2	460700	100
35400.123334	ask	4	2	460700	6
35400.125010	ask	4	2	460700	14

Time	State	Ask price 1	Ask volume 1	Bid price 1	Bid volume 1
35400.092452	(1, 1)	460700	200	460600	900
35400.092533	(−1, 1)	460700	200	460600	1578
35400.092768	(0, 1)	460700	200	460600	1553
35400.113748	(1, 1)	460700	300	460600	1029
35400.113776	(1, −1)	460800	1600	460700	100
35400.121175	(−1, −1)	460800	1800	460700	153
35400.121258	(−1, 0)	460700	3300	460600	1029
35400.123294	(0, 1)	460700	352	460600	900
35400.123334	(0, 1)	460700	346	460600	900
35400.125010	(0, 1)	460700	332	460600	900

We show the first level of the limit order book for ease of exposition.

Transition probabilities are straightforwardly estimated from empirical frequencies. For every event $e = 1, \dots, 4$, we estimate a 9×9 -transition matrix ϕ_e that describes the law of the state-update in Equation (7). In Table 4, we show the result of this estimation focussing on events of type either 1 or 2, i.e., execution on either the bid or the ask side.

Assumption 6.1: Hawkes kernels are assumed in the parametric form

$$\kappa_{e',e}(t, x') = \alpha_{e',x',e} (t+1)^{-\beta_{e',x',e}}, \quad (18)$$

for some non-negative coefficients $\alpha_{e',x',e} \geq 0$ and $\beta_{e',x',e} > 1$.

Remark 6.2: Assumption 6.1 is grounded in the stylized fact that power-law kernels better fit real world data than exponential kernels do, albeit being more computational expensive. This assumption also builds on Bacry and Muzy (2014)'s findings. In the aforementioned paper, the authors devise a non-parametric estimation for Hawkes kernels. Once this non-parametric estimation has converged, they compare the estimated kernels

Table 4. Transition probabilities ϕ_e calibrated on INTC as of 25 January 2019.

Probabilities of mid-price movements $X_1(T) \in \{-1, 0, +1\}$ when event of type either 1 or 2 occurs at time T and queue imbalance $X_2(T-)$ is negative (−1), neutral (0) or positive (+1):

event	X_1 X_2	−1	0	1
1	−1	0.476657	0.523343	0.000000
	0	0.410729	0.589271	0.000000
	1	0.346157	0.653843	0.000000
2	−1	0.000000	0.702554	0.297446
	0	0.000000	0.617759	0.382241
	1	0.000000	0.533823	0.466177

Table 4. Continued.

Full transition probability from $X(T-) = (X_1(T-), X_2(T-))$ to $X(T) = (X_1(T), X_2(T))$ when an execution happens on the bid side at time T .											
event	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1
		X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1
1	-1	-1	0.176622	0.175752	0.000000	0.359614	0.283382	0.004630	0.0	0.0	0.0
		0	0.059468	0.279633	0.005352	0.077036	0.550193	0.028318	0.0	0.0	0.0
		1	0.009081	0.217948	0.071566	0.019519	0.433721	0.248165	0.0	0.0	0.0
	0	-1	0.287448	0.245331	0.003610	0.414895	0.048716	0.000000	0.0	0.0	0.0
		0	0.065844	0.350749	0.021704	0.083769	0.473489	0.004445	0.0	0.0	0.0
		1	0.005626	0.162936	0.130160	0.008709	0.242360	0.450209	0.0	0.0	0.0
	1	-1	0.257898	0.279307	0.004003	0.327795	0.130997	0.000000	0.0	0.0	0.0
		0	0.066807	0.357431	0.025199	0.091674	0.447274	0.011614	0.0	0.0	0.0
		1	0.013147	0.224193	0.203817	0.033174	0.353798	0.171872	0.0	0.0	0.0

Table 5. Hawkes parameters ν_e , $\alpha_{e',x',e}$, and $\beta_{e',x',e}$ calibrated on INTC as of 25 January 2019.

	event 1	event 2	event 3	event 4
Base rates ν_e , $e = 1, \dots, 4$.				
base rate	0.040201	0.050182	0.000735	0.000608
$\alpha_{e',x',e}$ for $e = 1, \dots, 4$ and $e' = 1, \dots, 4$ when queue imbalance $X_2 = 0$ and $X_1 = 0$.				
event 1	9.25	0.08	2.74	0.75
event 2	0.03	11.15	0.25	2.85
event 3	0.04	0.21	0.21	0.25
event 4	0.45	0.01	0.22	0.16
$\beta_{e',x',e}$ for $e = 1, \dots, 4$ and $e' = 1, \dots, 4$ when queue imbalance $X_2 = 0$ and $X_1 = 0$.				
event 1	14.71	5.32	7.22	5.64
event 2	4.61	16.92	4.92	7.92
event 3	1.57	1.92	3.41	2.84
event 4	3.53	2.70	2.84	3.56

with parametric ones, and confirm that indeed the decay of the kernels is of power-law type.

We estimate the parameters ν_e , $\alpha_{e',x',e}$, and $\beta_{e',x',e}$ using a gradient-descent algorithm. In Table 5, we present the result of this estimation by reporting the four dimensional vector ν , and both, $\alpha_{e',x',e}$ and $\beta_{e',x',e}$ for when $X_2 = 0$ and $X_1 = 0$ – the other cases are not presented in the interest of space.

Figure 1 shows QQ plots for goodness-of-fit diagnostics; we note that the fit is adequate for our purposes although there is some deviation in the tail part of the plots – perfecting the fit tends to be difficult with the amount of data we employ (e.g., for INTC on 25 January 2019 we employ 1,563,582 datapoints).

Finally, Figure 2 compares the trajectory of the mid-price as reported in LOBSTER, as reconstructed from Equation (11), and as simulated in the calibrated model (one sample); we plot two different time-scales.

6.3. Price Impact Assessment

We simulate liquidations in our state-dependent Hawkes model for an order book calibrated on LOBSTER data for the ticker INTC on 25 January 2019, and we assess the price impact of such liquidations using Definition 5.8.

QQ plot residuals

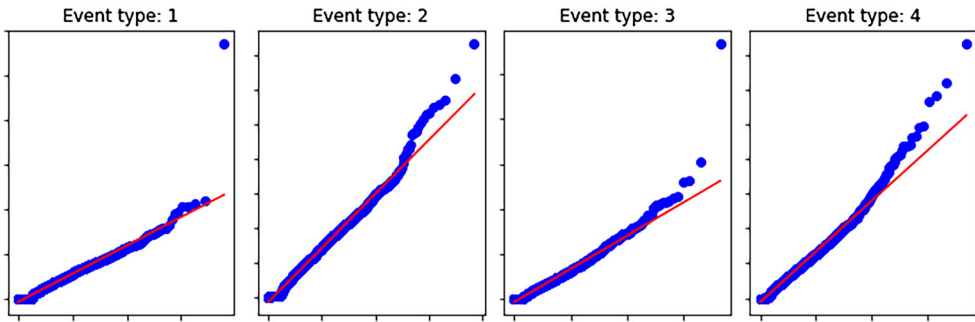


Figure 1. Goodness-of-fit diagnostics for the model calibrated on INTC as of 25 January 2019. We test that the time-changed inter-arrival times of Equation (4) are i.i.d. samples from a unit rate exponential distribution. Empirical quantiles of the time-changed inter-arrival times on the y-axis and theoretical quantiles of unit-rate exponential distribution on the x-axis.

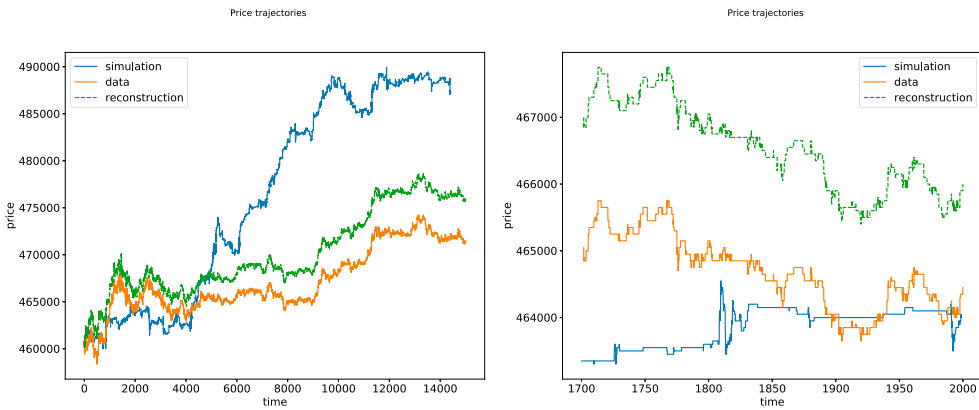


Figure 2. Mid-price trajectories on two time scales. Origin of time is set at 9.55am 25 January 2019. Time is measured in seconds. Prices are in 10^{-4} USD.

Table 6. Parameters of the liquidation schedule.

Name	Symbol	Formula
base rate	ν_0	Zero-th component of $\nu = (\nu_0, \nu_1, \dots, \nu_4)$
clustering rate	a	$\alpha_{e',x,0} = a \cdot \alpha_{e',x,1}$
order size	c	$q_{Mj} = c \cdot \sum_{i=1}^n V_{T_j^0-}^{b_j}$

We investigate two aspects of the liquidation schedule: The rate with which liquidator's orders walk the book (captured by the transition matrix ϕ_0), and the clustering of the liquidator's orders in response to events happening in the limit order book. We modulate these two aspects through the parameters reported in Table 6.

We run simulations for different values of these parameters and we find out that the clustering of liquidator's orders has a bigger price impact than the rate with which they

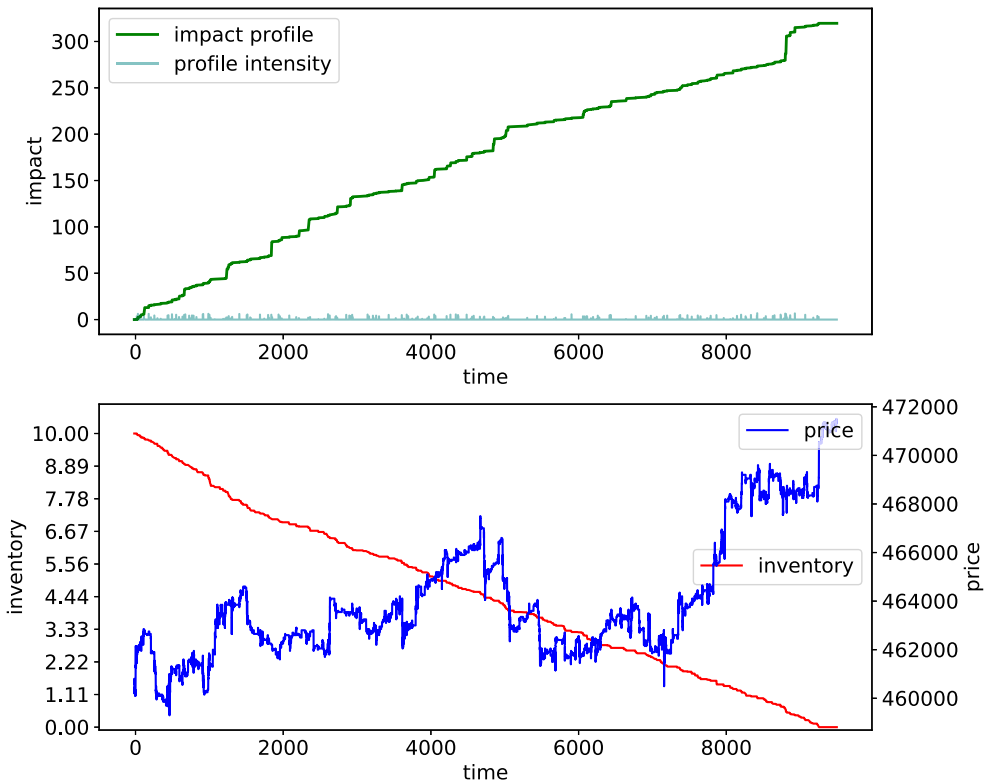


Figure 3. Price impact with low rate of walking the book and no clustering. Initial inventory $Q_0 = 10$, base rate $\nu_0 = 0.03$, clustering rate $a = 0$, order size $c = 0.075$, start time $t_0 = 0$, termination time $\tau = 9250.2$, and price impact score is 0.0346.

walk the book. This suggests that the dynamic evolution of the order book plays a bigger role in price formation than the instantaneous states of the queues.

Figures 3, 6 and 7 present three simulations representative of our findings.

More precisely, Figure 3 shows a liquidation in which there is no clustering of market orders (i.e., $\alpha_{e',x,0} = 0$), and the execution schedule follows a Poisson process with rate $\nu_0 = 0.03$, namely approximately 75% of the base rate of all other sell market orders.

All liquidator's orders have size equal to 7.5% of the volumes of bid offers queuing on the first n levels of the bid side at the moment of the order submission. This entails that liquidator's orders will walk the book only when the aggregate size of level 1 is less than 7.5% of the overall bid volume. As a result, the estimated transition matrix $\phi_0 = \phi_0(x', x)$ concentrates the mass on those states $x = (x_1, x_2)$ such that $x_2 = 0$. The more positive the queue imbalance at the moment of the order submission is, the more this concentration holds.

The liquidation takes approximately 9250 s to complete. The price impact score, defined as the maximum of the price impact profile divided by the duration, is 3.46%.

The line charts in Figure 3 (top left plot) present visualizations of the liquidation schedule (the red dots represent executions on the bid side triggered by one of liquidator's sell market orders), and its intensity (see Equation (13)), which in this case is simply

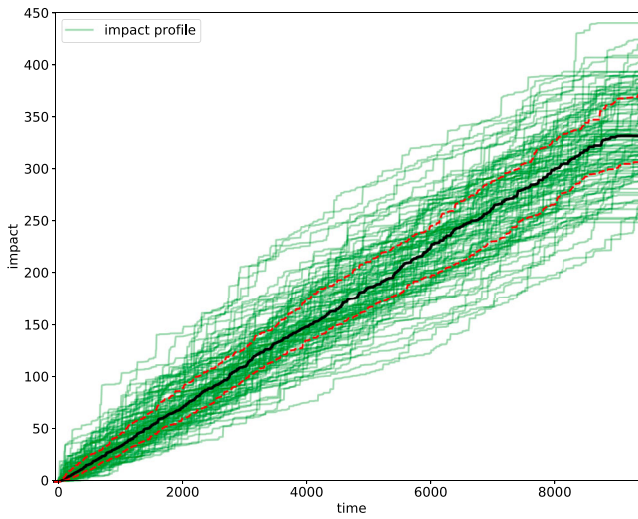


Figure 4. Simulated trajectories of the intensity of the price impact profile, i.e., the map $t \mapsto \text{Dir}(t) + \text{Indir}(t)$ (see Definition 5.8) for the scenario considered in Figure 3. The black solid line is the median trajectory of the impact profile across simulations and the red dotted lines are the 25% and 75% quantile trajectories.

$\lambda_0(t) = \nu_0 \mathbb{I}_{[0,\tau)}(t)$. This panel also depicts the impact profile trajectory (green line in top left plot) that we used to compute the price impact score. Moreover, the trajectories of liquidator's inventory and of the mid-price during execution are plotted (see bottom left plot). The latter can be compared to the mid-price simulated when the liquidator was not present in the market (see Figure 2), and it provides a graphical representation of price impact.

We remark that the stylized features one observes in the impact profile trajectory are consistent across simulations. In Figure 4 we show the impact profile trajectories for one hundred simulations.

Remark 6.3: Taking the mean across simulations provides a bridge from our scenario-dependent impact profiles to the average impact profiles studied in the literature. Since our model is rich enough to comprise all market variables appearing in the stylized facts of price impact, one can hold all model parameters fixed except for one variable of interest, and study average price impact scores as a function of the variable of interest. This opens the door to investigating whether a specific model adheres to the stylized facts in the literature, as those reviewed in Zarinelli et al. (2015). In Figure 5 below, we study two experiments where we stress the order size parameter c and the base rate parameter ν_0 .

The left panel of Figure 5 shows the average increase in the price impact profile at time τ as a function of the order size parameter c ; the right panel shows the price impact score as a function of the base rate parameter ν_0 – see Table 6 for the definition of c and ν_0 . The remaining model parameters are those in Figure 3.

Figure 6 shows a liquidation in which the execution schedule is as in Figure 3, namely no clustering and same base rate. However, in this case liquidator's orders have a much bigger size. They have size equal to 50% of the volumes of bid offers queuing on the first n levels of the bid side at the moment of the order submission. This entails that liquidator's orders

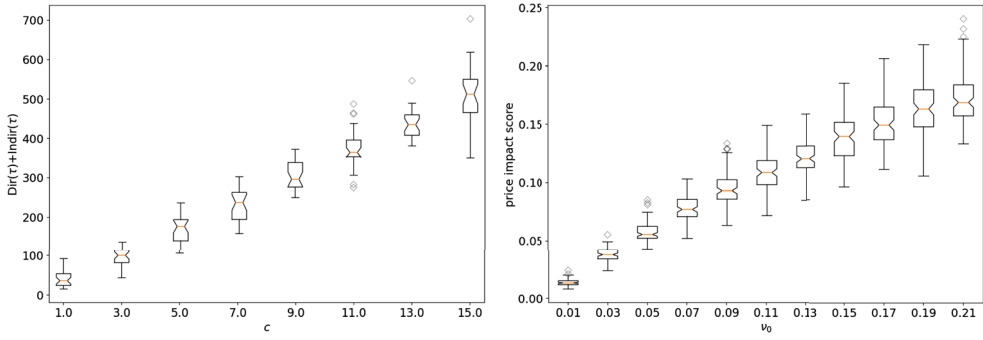


Figure 5. Left panel: box plots of the distribution of $\text{Dir}(\tau) + \text{Indir}(\tau)$ when stressing the order size parameter c . Right panel: box plots of the distribution of the price impact score as a function of the base rate ν_0 .

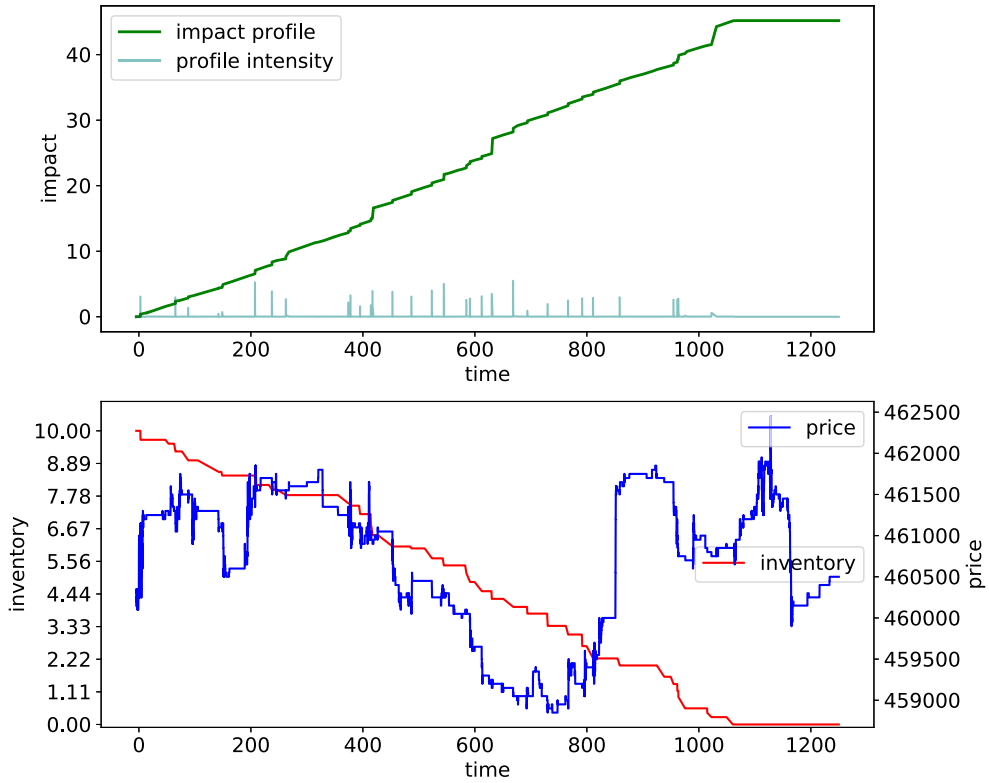


Figure 6. Price impact with high rate of walking the book and no clustering. Initial inventory $Q_0 = 10.0$, base rate $\nu_0 = 0.03$, clustering rate $a = 0$, order size $c = 0.5$, start time $t_0 = 0$, termination time $\tau = 1060.7$, and price impact score is 0.0426.

will likely walk the book. Indeed, they walk the book whenever the aggregate size of level 1 is less than 50% of the overall bid volume. As a result, the estimated transition matrix $\phi_0 = \phi_0(x', x)$ concentrates the mass on those states $x = (x_1, x_2)$ such that $x_2 = -1$. The

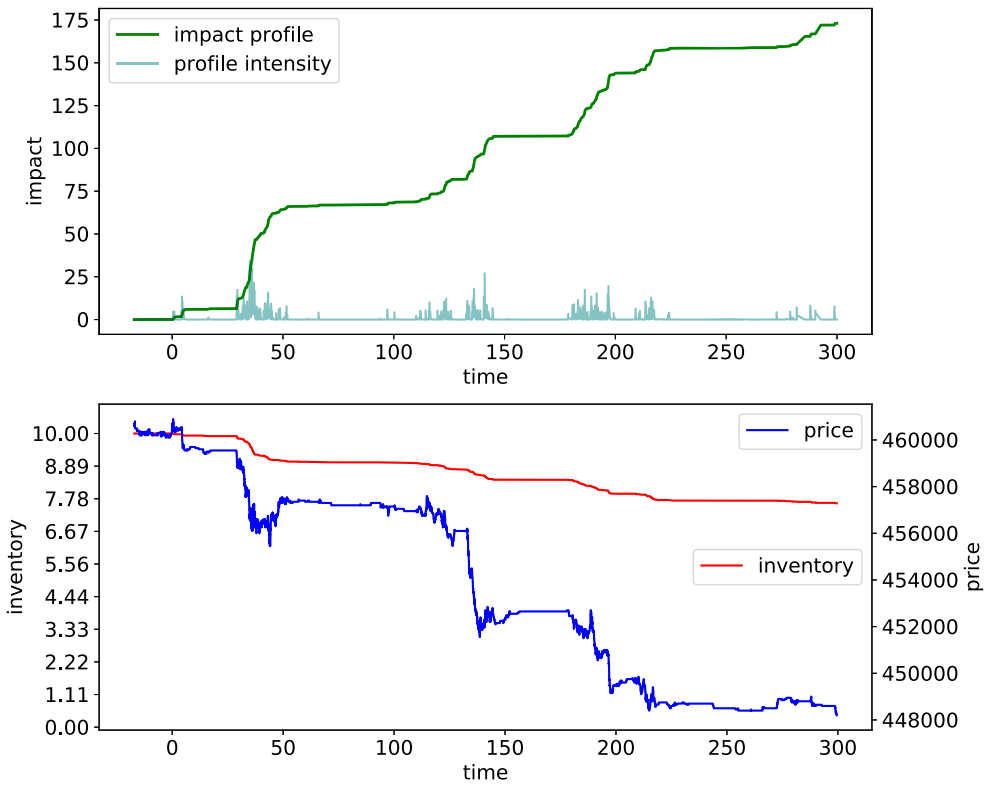


Figure 7. Price impact with clustering of liquidator's orders and low rate of walking the book. We show the first 300 s. Initial inventory $Q_0 = 10.0$, base rate $v_0 = 0$, clustering rate $a = 0.25$, order size $c = 0.015$, start time $t_0 = 0$, termination time $\tau = 722.3$, and price impact score is 0.999.

more negative the queue imbalance at the moment of the order submission is, the more this concentration holds.

The liquidation takes approximately 1061 s to complete. This is considerably shorter than in the previous simulation because every order executes more of the liquidator's inventory. The price impact score, is 4.26%. The plotted mid-price trajectory shows a sharper plunge during execution.

Hence, Figures 3 and 6 show how our model captures the consequence that order sizing has on price impact. Order sizing however is not the main driver of price impact. We demonstrate this in Figure 7.

Figure 7 shows a liquidation in which the intensity of the execution schedule has zero base rate, namely there is no exogenous cause for the liquidator to submit their orders. Instead, they submit their orders in response to orders submitted by other market participants. We set the liquidator's response to be proportional to the response that other market participants make when scheduling their sell market orders. That is, we set $\alpha_{e',x,0} = a \cdot \alpha_{e',x,1}$, for some coefficient $a \geq 0$. In the simulation of Figure 7 this coefficient is $a = 0.25$.

Recall that $\alpha_{e',x,1}$ was estimated by maximum likelihood estimation from the LOBSTER data set – see Table 5 where the L^1 -norms $\alpha_{e',x,e}/(\beta_{e',x,e} - 1)$ are reported. The

estimation revealed that seller-initiated executions are more likely to excite other events with deflationary pressure on the mid-price, i.e., events of type $e = 1$ or $e = 3$, whereas buyer-initiated executions are more likely to excite other events with inflationary pressure on the mid-price, i.e., events of type $e = 2$ or $e = 4$. As a consequence, when $\alpha_{e',x,0}$ is proportional to $\alpha_{e',x,1}$ deflationary events will tend to cluster, potentially triggering abrupt price changes. The line charts in Figure 7 give visual representations of this. We deliberately focussed on a short time horizon of 300 s, so that the phenomenon described is more apparent to the eye.

We set the order sizes to be very small, at 1.5% of the volumes available on the first n levels of the bid side. This has the purpose to isolate the phenomenon of indirectly induced price changes versus price changes induced by walking the book. Indeed, the estimated transition probabilities ϕ_0 on this simulation assign negligible probabilities to transitioning to a state $X(T)$ with $X_1(T) = -1$ when an event of type 0 occurs at time T .

Nonetheless, because of the clustering effect of deflationary events, the price impact score is the highest, at 99.9%.

The results presented in this section are robust to misspecification of model parameters. When we stress the base rates $\{\nu_e\}_e$, together with the coefficients $\{\alpha_{e',x',e}\}_{e',x',e}$ and $\{\beta_{e',x',e}\}_{e',x',e}$ with shocks between -5% and $+5\%$, we observe that the average price impact scores across one hundred simulations get affected by less than 10% of the average unstressed values. For example, if we consider the experiment in Figure 3, and we repeat the profiling with the modified parameters $\tilde{\nu}_e = 1.05 \nu_e$, $\tilde{\alpha}_{e',x',e} = 1.05 \alpha_{e',x',e}$ and $\tilde{\beta}_{e',x',e} = 1.05 \beta_{e',x',e}$ for all $\{e', x, e\}$, we find that the average price impact scores changes from 0.037 (with 0.006 standard deviation) to 0.041 (with 0.006 standard deviation), which represents an average increase of 8.2%. Similarly, in that same scenario of Figure 3, when we consider the modified parameters $\tilde{\nu}_e = 0.95 \nu_e$, $\tilde{\alpha}_{e',x',e} = 0.95 \alpha_{e',x',e}$ and $\tilde{\beta}_{e',x',e} = 0.95 \beta_{e',x',e}$ for all $\{e', x, e\}$, we find that the average price impact scores decrease to 0.036 (with 0.005 standard deviation), which represents an average decrease of 3.7%. The other two scenarios we consider in Figures 6 and 7 show a similar behaviour when stressing the calibrated parameters. This means that the results shown are indeed robust to calibration errors.

Finally, we repeat the above analysis employing the tickers AAPL and INTC for a range of dates during January 2019. We report that no further insights are derived from considering other dates or other tickers. For example, it was always the case that the highest price impact scores were achieved under the scenario we consider in Figure 7, i.e., where the clustering effect plays a key role in the price impact profile. Thus, we conclude that the intuition derived from the analysis in Figures 3, 6, and 7 remains valid for other dates and similar tickers.

In limit order books of small tick-size stocks (e.g., GOOG or TSLA), orders are posted and cancelled in adjacent price queues more frequently than for medium-size or large-size stocks. In other words, the mid-price of a small tick-size stock has far smaller constant traits. This is because, relative to the stock price, a one-tick change in the best bid or best ask is not as significant as for larger tick-size stocks. When studying small tick-size stocks with our model, such a reduced importance of single-tick changes can be accounted for by letting N_3 and N_4 jump only when the mid-price changes by two or more ticks, de-facto renormalizing the parameters of the limit order book (doubling or tripling the tick size and merging adjacent queues of orders). In our experiments, this led to a more robust calibration and prevented overfitting.

7. Discussion

Modelling price impact, direct and indirect, is an important challenge in market microstructure. In this paper we proposed a methodology for modelling both components (direct and indirect) separately on a path-by-path basis. Our definition of price impact assumes that price movements are driven by state-dependent Hawkes processes. Specifically, the (i) seller/buyer-initiated trades, and the (ii) decreases/increases in mid-price by limit order insertions or cancellations, follow a state-dependent Hawkes process. We conducted a goodness-of-fit analysis and concluded that the model fits the data adequately. We studied the average behaviour of the price impact score as a function the base rate of the execution and found evidence of a concave relationship between these variables; future work will explore alternative calibration or modelling assumptions that exhibit a better fit on the data together with further studies of the average behaviour of the price impact profiles we define. Additional avenues for future work include exploring the behaviour of the price impact profile after the executions are completed and use our model to solve control problems of optimal execution of large orders and liquidity provision.

Notes

1. In some trading venues, traders can actually submit market orders, i.e., buy or sell orders that are executed without price constraints, at least as long as offers with the opposite direction exist. Even in such cases, we keep our convention of referring to the fraction of a limit order that is executed upon submission as market order.
2. There is no chance that a sell (buy) market order can cause an increase (decrease, respectively) in the mid-price.
3. Every $q_{M,j}$ satisfies the measurability constraint $q_{M,j} \hat{\in} \mathfrak{F}_{T_j^0-} = \bigvee_{\epsilon>0} \mathfrak{F}_{T_j^0-\epsilon}$, where $\mathfrak{F}$ is a history of (N, X) .
4. To see why this can be a potentially undesirable feature, consider a liquidator that trades with child orders whose sizes are significantly different from the average market orders arriving in the market.
5. See <https://lobsterdata.com/>.
6. See <http://nasdaqtrader.com/Trader.aspx?id=ITCH>.
7. See rule 4701(k) at <https://listingcenter.nasdaq.com/rulebook/nasdaq/rules/nasdaq-4000>.
8. See <https://lobsterdata.com/info/DataStructure.php>.

Disclosure statement

This work reflects the analysis and views of the authors Claudio Bellani, Damiano Brigo, Leandro Sánchez-Betancourt, and Mikko Pakkanen. No reader should interpret this work to present the views of any third party. Assumptions, opinions, views and estimates constitute the authors' judgment as of the date given and are subject to change without notice and without duty to update.

References

- Almgren R., C. Thum, E. Hauptmann, and H. Li. 2005. "Direct Estimation of Equity Market Impact." *Risk* 18 (7): 58–62.
- Bacry E., A. Iuga, M. Lasnier, and C.-A. Lehalle. 2015. "Market Impacts and the Life Cycle of Investors Orders." *Market Microstructure and Liquidity* 1 (2): 1550009. <https://doi.org/10.1142/S2382626615500094>.
- Bacry E., and J.-F. Muzy. 2014. "Hawkes Model for Price and Trades High-frequency Dynamics." *Quantitative Finance* 14 (7): 1147–1166. <https://doi.org/10.1080/14697688.2014.897000>.

- Bouchaud J.-P. 2017. *Market Impact: A Review. Imperial-CFM Seminar*. <https://www.imperial.ac.uk/events/99125/jean-philippe-bouchaud-cfm-market-impact-a-review/>.
- Brokmann X., E. Serie, J. Kockelkoren, and J.-P. Bouchaud. 2015. “Slow Decay of Impact in Equity Markets.” *Market Microstructure and Liquidity* 1 (2): 1550007. <https://doi.org/10.1142/S2382626615500070>.
- Brown T. C., and M. G. Nair. 1988. “A Simple Proof of the Multivariate Random Time Change Theorem for Point Processes.” *Journal of Applied Probability* 25 (1): 210–214. <https://doi.org/10.2307/3214247>.
- Bucci F., I. Mastromatteo, M. Benzaquen, and J.-P. Bouchaud. 2019. “Impact Is Not Just Volatility.” *Quantitative Finance* 19 (11): 1763–1766. <https://doi.org/10.1080/14697688.2019.1622768>.
- Capponi F., and R. Cont. 2019. “Trade Duration, Volatility and Market Impact.” SSRN. <https://ssrn.com/abstract=3351736>.
- Cartea A., R. Donnelly, and S. Jaimungal. 2018. “Enhancing Trading Strategies with Order Book Signals.” *Applied Mathematical Finance* 25 (1): 1–35. <https://doi.org/10.1080/1350486X.2018.1434009>.
- Daley D., and D. Vere-Jones. 2008. *An Introduction to the Theory of Point Processes Volume II: General Theory and Structure*. Probability and Its Applications. 2nd ed. New York: Springer.
- Engle R., R. Ferstenberg, and J. Russell. 2012. “Measuring and Modeling Execution Cost and Risk.” *The Journal of Portfolio Management* 38 (2): 14–28. <https://doi.org/10.3905/jpm.2012.38.2.014>.
- Huang R., and T. Polak. 2011. “Lobster: Limit Order Book Reconstruction System.” *Lobsterdata Docs*.
- Jusselin P., and M. Rosenbaum. 2020. “No-arbitrage Implies Power-law Market Impact and Rough Volatility.” *Mathematical Finance* 30 (4): 1309–1336. <https://doi.org/10.1111/mafi.v30.4>.
- Meyer P.-A.. 1971. “Démonstration Simplifiée D’un Théorème De Knight.” *Séminaire de Probabilités de Strasbourg* 5:191–195.
- Morariu-Patrichi M., and M. S. Pakkanen. 2018. “Hybrid Marked Point Processes: Characterization, Existence and Uniqueness.” *Market Microstructure and Liquidity* 4 (3n04): 1950007. <https://doi.org/10.1142/S2382626619500072>.
- Morariu-Patrichi M., and M. S. Pakkanen. 2022. “State-Dependent Hawkes Processes and Their Application to Limit Order Book Modelling.” *Quantitative Finance* 22 (3): 563–583. <https://doi.org/10.1080/14697688.2021.1983199>.
- Moro E., J. Vicente, L. G. Moyano, A. Gerig, J. D. Farmer, G. Vaglica, F. Lillo, and R. N. Mantegna. 2009. “Market Impact and Trading Profile of Hidden Orders in Stock Markets.” *Physical Review E* 80 (6 Pt 2): 066102. <https://doi.org/10.1103/PhysRevE.80.066102>.
- Patzelt F., and J.-P. Bouchaud. 2018. “Universal Scaling and Nonlinearity of Aggregate Price Impact in Financial Markets.” *Physical Review E* 97:012304. <https://doi.org/10.1103/PhysRevE.97.012304>.
- Rosenbaum M., and M. Tomas. 2021. “A Characterisation of Cross-Impact Kernels.” Preprint, arXiv:2107.08684.
- Torre N. 1997. *Barra Market Impact Model Handbook*. Vol. 208. Berkeley: BARRA Inc.
- Tóth B., Z. Eisler, and J.-P. Bouchaud. 2016. “The Square-root Impace Law Also Holds for Option Markets.” *Wilmott* 2016 (85): 70–73. <https://doi.org/10.1002/wilm.2016.2016.issue-85>.
- Tóth B., Y. Lempérière, C. Deremble, J. de Lataillade, J. Kockelkoren, and J.-P. Bouchaud. 2011. “Anomalous Price Impact and the Critical Nature of Liquidity in Financial Markets.” *Physical Review X* 1 (2): 021006. <https://doi.org/10.1103/PhysRevX.1.021006>.
- Zarinelli E., M. Treccani, J. D. Farmer, and F. Lillo. 2015. “Beyond the Square Root: Evidence for Logarithmic Dependence of Market Impact on Size and Participation Rate.” *Market Microstructure and Liquidity* 1 (2): 1550004. <https://doi.org/10.1142/S2382626615500045>.

Appendix. Proofs

Proof of Proposition 2.2: Let $(t, q, p, -1)$ be a sell limit order. Let $N_v := \inf\{n \geq 0 : q < \sum_{i=1}^n V_{t-}^{b,i}\}$. Let $N_p := \inf\{n \geq 1 : P_{t-}^{b,n} < p\}$. Let

$$\begin{aligned} q^i &:= \max \left(0, q - \sum_{k=1}^{i \wedge N_v \wedge (N_p - 1)} V_{t-}^{b,k} \right) \\ &= \max \left(0, q - \sum_{1 \leq k \leq i} V_{t-}^{b,k} \mathbb{1}_{\{P_{t-}^{b,k} \geq p\}} \right) \end{aligned}$$

Let N_v^M , N_p^M and q_M^i be the corresponding quantities for the order $(t, q_M, 0, -1)$. Notice that $N_v^M = N_v^M \wedge N_p^M$. Processing $(t, q_M, 0, -1)$ does not affect the ask side because $q_M^\infty = 0$; moreover the effects on the bid side of processing $(t, q_M, 0, -1)$ are the same as those of $(t, q, p, -1)$, because $\inf\{n \geq 0 : q_M < \sum_{i=1}^n V_{t-}^{b,i}\} = N_v \wedge N_p$ and $q^{k+N-1} - q^{k+N} = q_M^{k+N_v^M-1} - q_M^{k+N_v^M}$. Therefore, after $(t, q_M, 0, -1)$ has been processed, the bid side is the same as the bid side after the processing of $(t, q, p, -1)$.

Processing $(t, q - q_M, p, -1)$ after $(t, q_M, 0, -1)$ does not alter the bid side because either $P_{t-1}^b < p$ or $q - q_M = 0$. Moreover, $(t, q - q_M, p, -1)$ has the same effects on the ask side as those of $(t, q, p, -1)$ because $q - q_M = q^\infty$.

The proof of the decomposition of a buy limit order is mutatis mutandis the same. ■

Proof of Proposition 5.6: Let (T_n, E_n, X_n) be the sequence of arrival times, events and states. Then, we can compute

$$\begin{aligned} & \sum_{x \in S^+} \sum_{e=1}^4 \phi_e(X(t), x) \sum_{e'=1}^4 \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s) \\ &= \sum_{n: T_n < t} \sum_{x \in S^+} \sum_{e=1}^4 \phi_e(X(t), x) \kappa_{E_n, e}(t - T_n, X_n) \\ &= \sum_{n: T_n < t} \sum_{x \in S^+} \sum_{e=1}^4 \phi_{\sigma_E(e)}(X(t), \sigma_S(x)) \kappa_{E_n, \sigma_E(e)}(t - T_n, X_n) \\ &= \sum_{n: T_n < t} \sum_{x \in S^-} \sum_{e=1}^4 \phi_e(X(t), x) \kappa_{E_n, e}(t - T_n, X_n) \\ &= \sum_{x \in S^-} \sum_{e=1}^4 \phi_e(X(t), x) \sum_{e'=1}^4 \int_{[0,t)} \kappa_{e',e}(t-s, X(s)) dN_{e'}(s). \end{aligned}$$
■

Proof of Proposition 5.11: We need to show that

$$(\lambda^- - \lambda^+)(t) = \text{Dir}(t) + \text{Indir}(t). \quad (\text{A1})$$

To this purpose, we compute λ^- (resp., of λ^+) as the sum of the intensities of $\tilde{N}_{e,x}$ for $e = 0, \dots, 4$ and x in $\mathcal{S}^-$ (resp., in $\mathcal{S}^+$). From Equations (9) and (12) it follows that

$$\begin{aligned} \lambda^-(t) = & \sum_{x \in \mathcal{S}^-} \left\{ \phi_0(X(t), x) \lambda_0(t) \right. \\ & + \sum_{e=1}^4 \phi_e(X(t), x) \left(\underbrace{v_e + \sum_{e'=1}^4 \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{e',e}(t-s, x') d\tilde{N}_{e',x'}(s)}_{=\ell_e(t)} \right. \\ & \left. \left. + \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{1,e}(t-s, x') d\tilde{N}_{0,x'}(s) \right) \right\}, \end{aligned}$$

where $\lambda_0(t)$ is as in (13), and

$$\begin{aligned} \lambda^+(t) = & \sum_{x \in \mathcal{S}^+} \sum_{e=1}^4 \phi_e(X(t), x) \left(\underbrace{v_e + \sum_{e'=1}^4 \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{e',e}(t-s, x') d\tilde{N}_{e',x'}(s)}_{=\ell_e(t)} \right. \\ & \left. + \sum_{x'=1}^{3K} \int_{[0,t)} \kappa_{1,e}(t-s, x') d\tilde{N}_{0,x'}(s) \right). \end{aligned}$$

By price-symmetry, the terms $\sum_{x \in \mathcal{S}^-} \sum_{e=1}^4 \phi_e(X(t), x) \ell_e(t)$ and $\sum_{x \in \mathcal{S}^+} \sum_{e=1}^4 \phi_e(X(t), x) \ell_e(t)$ will cancel out from the difference $\lambda^-(t) - \lambda^+(t)$. ■