

# Models for the impact of all order book events

Zoltán Eisler,<sup>1</sup> Jean-Philippe Bouchaud,<sup>1</sup> and Julien Kockelkoren<sup>1</sup>

<sup>1</sup>*Capital Fund Management, Paris, France*

(Dated: July 19, 2011)

We propose a general framework to describe the impact of different events in the order book, that generalizes previous work on the impact of market orders. Two different modeling routes can be considered, which are equivalent when only market orders are taken into account. One model posits that each event type has a temporary impact (TIM). The “history dependent impact” model (HDIM), on the other hand, assumes that only price-changing events have a direct impact, itself modulated by the past history of all events through an “influence matrix” that measures how much, on average, an event of a given type affects the immediate impact of a price-changing event of the same sign in the future. We find in particular that aggressive market orders tend to reduce the impact of further aggressive market orders of the same sign (and increase the impact of aggressive market orders of opposite sign). We discuss the relative merits of TIM and HDIM, in particular concerning their ability to reproduce accurately the price diffusion pattern. We find that in spite of theoretical inconsistencies, TIM appears to fare better than HDIM when compared to empirical data. We ascribe this paradox to an uncontrolled approximation used to calibrate HDIMs, calling for further work on this issue.

Keywords: price impact, market orders, limit orders, cancellations, market microstructure, order flow

## Contents

<b>1. Introduction</b>	1
<b>2. A short summary of market order impact models</b>	2
<b>3. Many-event impact models</b>	4
3.1. Notations and definitions	4
3.2. The transient impact model (TIM)	5
3.3. The history-dependent impact model (HDIM)	5
<b>4. Model calibration and empirical tests</b>	7
4.1. Data	7
4.2. The case of large ticks	7
4.3. The case of small ticks	8
4.3.1. TIM	8
4.3.2. HDIM	8
<b>5. Conclusion</b>	10
<b>Acknowledgments</b>	11
<b>A. Expression of the price diffusion for the TIM and HDIM</b>	11
<b>References</b>	12

## 1. INTRODUCTION

The relation between order flow and price changes has attracted considerable attention in the recent years [1–6]. Most empirical studies to date have focused on the impact of (buy/sell) market orders. Many interesting results have been obtained, such as the very weak dependence of impact on the volume of the market order, the long-range nature of the *sign* of the trades, and the resulting non-permanent, power-law decay of market order impact with time (see Ref. [6]). However, this representation of impact is incomplete in at least two ways. First, the impact of all market orders is usually treated on the same footing, or with a weak dependence on volume, whereas some market orders are

“aggressive” and immediately change the price, while other are more passive and only have a delayed impact on the price. Second, other types of order book events (limit orders, cancellations) must also directly impact prices: adding a buy limit order induces extra upwards pressure, and cancelling a buy limit order decreases this pressure. Within a description based on market orders only, the impact of limit orders and cancellations is included in an indirect way, in fact as an effectively decaying impact of market orders. This decay reflects the “liquidity refill” mechanism explained in detail in Refs. [4, 6–9], whereby market order trigger a counterbalancing flow of limit orders.

A framework allowing one to analyze the impact of all order book events, and to understand in detail the statistical properties of price time series, is clearly desirable. Surprisingly, however, there are only very few quantitative studies of the impact of limit orders and cancellations [10, 11] – partly due to the fact that more detailed data, beyond trades and quotes, is often needed to conduct such studies. The aim of the present paper is to provide a theoretical framework for the impact of all order book events, which allows one to build statistical models with intuitive and transparent interpretations [11].

## 2. A SHORT SUMMARY OF MARKET ORDER IMPACT MODELS

A simple model relating prices to trades posits that the mid-point price  $p_t$  just before trade  $t$  can be written as a linear superposition of the of the impact of all past trades [3, 4]:

$$p_t = \sum_{t' < t} [G(t - t')\xi_{t'} + \eta_{t'}] + p_{-\infty}, \quad \xi_t \equiv \epsilon_t v_t^\theta \quad (1)$$

where  $v_{t'}$  is the volume of the trade at time  $t'$ ,  $\epsilon_{t'}$  the sign of that trade (+ for a buy, – for a sell), and  $\eta_t$  is an independent noise term that models any price change not induced by trades (e.g. jumps due to news). The exponent  $\theta$  is found to be small. The most important object in the above equation is the function  $G(\ell)$  describing the temporal evolution of the impact of a single trade, which can be called a ‘propagator’: how does the impact of the trade at time  $t' = t - \ell$  propagate, on average, up to time  $t$ ? Because the signs of trades are strongly auto-correlated,  $G(\ell)$  must decay with time in a very specific way, in order maintain the (statistical) efficiency of prices. Clearly, if  $G(\ell)$  did not decay at all, the returns<sup>1</sup>  $r_t = p_{t+1} - p_t$  would simply be proportional to the sign of the trades, and therefore would themselves be strongly autocorrelated in time. The resulting price dynamics would then be highly predictable, which is not realistic. The result of Ref. [3] is that if the correlation of signs  $C(\ell) = \langle \epsilon_t \epsilon_{t+\ell} \rangle$  decays at large  $\ell$  as  $\ell^{-\gamma}$  with  $\gamma < 1$  (as found empirically), then  $G(\ell)$  must decay as  $\ell^{-\beta}$  with  $\beta = (1 - \gamma)/2$  for the price to be exactly diffusive at long times. The impact of single trades is therefore predicted to decay as a power-law (at least up to a certain time scale).

The above model can be rewritten in a completely equivalent way in terms of returns, with a slightly different interpretation [8, 9]:

$$r_t = G(0)\xi_t + \sum_{t' < t} \kappa(t - t')\xi_{t'} + \eta_t; \quad \kappa(\ell) \equiv G(\ell + 1) - G(\ell). \quad (2)$$

This can be read as saying that the  $t^{\text{th}}$  trade has a permanent impact on the price, but this impact is history dependent and depends on the sequence of past trades. The fact that  $G$  decays with  $\ell$  implies that the kernel  $\kappa$  is negative, and therefore that a past sequence of buy trades ( $\xi_{t' < t} > 0$ ) tends to *reduce* the impact of a further buy trade, but increase the impact of a sell trade. This is again a consequence of the dynamical nature of liquidity: when trades persist in a given direction, opposing limit orders tend to pile up and reduce the average impact of the next trade in the same direction. For the price to be an exact martingale, the quantity  $\hat{\xi}_t = -\sum_{t' < t} \kappa(t - t')\xi_{t'}$  must be equal to the conditional expectation of  $\xi_t$  at time  $t^-$ , such that  $\xi_t - \hat{\xi}_t$  is the surprise part of  $\xi_t$ . This condition allows one to recover the above mentioned decay of  $G(\ell)$  at large  $\ell$ .

In order to calibrate the model, one can use the empirically observable impact function  $\mathcal{R}(\ell)$ , defined as:

$$\mathcal{R}(\ell) = \langle (p_{t+\ell} - p_t) \cdot \xi_t \rangle, \quad (3)$$

and the time correlation function  $C(\ell)$  of the variable  $\xi_t = \epsilon_t v_t^\theta$  to map out, numerically, the complete shape of  $G(\ell)$ . This was done in Ref. [4], using the exact relation:

$$\mathcal{R}(\ell) = \sum_{0 < n \leq \ell} G(n)C(\ell - n) + \sum_{n > 0} [G(n + \ell) - G(n)]C(n). \quad (4)$$

---

<sup>1</sup> In the following, we only focus on price changes over small periods of time, so that an additive model is adequate.

Alternatively, one can use the ‘return’ version of the model, Eq. (2), which gives:

$$\mathcal{S}(\ell) = \langle r_{t+\ell} \cdot \xi_t \rangle, \quad (5)$$

that in turn leads to

$$\mathcal{S}(\ell) = G(0)C(\ell) + \sum_{n>-\ell} \kappa(\ell+n)C(n). \quad (6)$$

As noted in Ref. [12], this second implementation is in fact much less sensitive to finite size effects and therefore more adapted to data analysis.<sup>2</sup>

The above model, regardless of the type of fitting, is approximate and incomplete in two interrelated ways. First, Eqs. (1) and (2) neglect the fluctuations of the impact: one expects in general that  $G$  and  $\kappa$  should depend both on  $t$  and  $t'$  and not only on  $\ell = t - t'$ . Impact can indeed be quite different depending on the state of the order book and the market conditions at  $t'$ . As a consequence, if one blindly uses Eqs. (1) and (2) to compute the second moment of the price difference,  $D(\ell) = \langle (p_{t+\ell} - p_t)^2 \rangle$ , with a non-fluctuating  $G(\ell)$  calibrated to reproduce the impact function  $\mathcal{R}(\ell)$ , the result clearly underestimates the empirical price variance: see Fig. 1.

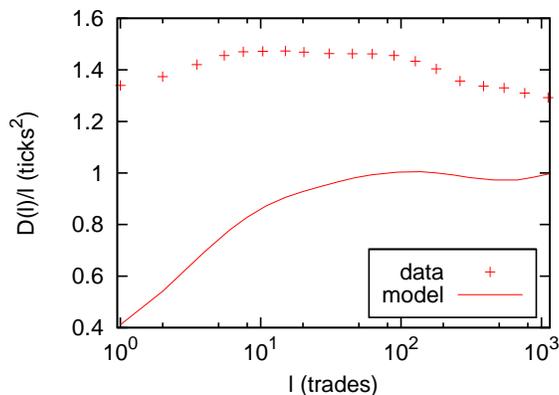


Figure 1:  $D(\ell)/\ell$  and its approximation with the transient impact model (TIM) with only trades as events, with  $\eta_t = 0$  and for small tick stocks. Results are shown when assuming that all trades have the same, non fluctuating impact  $G(\ell)$ , calibrated to reproduce  $\mathcal{R}(\ell)$ . This simple model accounts for  $\sim 2/3$  of the long term volatility. Other events and/or the fluctuations of  $G(\ell)$  must therefore contribute to the market volatility as well.

Adding a diffusive noise  $\eta_t \neq 0$  would only shift  $D(\ell)/\ell$  upwards, but this is insufficient to reproduce the empirical data. Second, as noted in the introduction, other events of the order book can also change the mid-price, such as limit orders placed inside the bid-ask spread, or cancellations of all the volume at the bid or the ask. These events do indeed contribute to the price volatility and should be explicitly included in the description. A simplified description of price changes in terms of market orders only attempts to describe other events of the order book in an effective way, through the non-trivial time dependence of  $G(\ell)$ .

In the following, we will generalize the above model to account for the impact of other types of events, beyond market orders. In this case, however, it will become apparent that the two versions of the above model *are no longer equivalent*, and lead to different quantitative results. Our main objective will be to come up with a simple, intuitive model that (i) can be easily calibrated on data, (ii) reproduces as closely as possible the second moment of the price difference  $D(\ell)$ , (iii) can be generalized to richer and richer data sets, where more and more events are observable and (iv) can in principle be systematically improved.

<sup>2</sup> The key difference is that a numerical solution necessarily truncates  $G(\ell)$  in Eq. (4) and  $\kappa(\ell)$  in Eq. (6) at some arbitrary  $\ell_{\max}$ . This truncation in the former case corresponds to the boundary condition of  $\kappa(\ell > \ell_{\max}) \equiv 0$ , hence a fully temporary impact at long times, while in the latter case to  $\kappa(\ell > \ell_{\max}) \equiv G(\ell_{\max})$ , hence a partially permanent impact. The latter solution is more smooth and consequently it is better behaved numerically.

$\pi$	event definition	event sign definition	gap definition ( $\Delta_{\pi,\epsilon}$ )
$\pi = \text{MO}^0$	market order, volume < outstanding volume at the best	$\epsilon = \pm 1$ for buy/sell market orders	0
$\pi = \text{MO}'$	market order, volume $\geq$ outstanding volume at the best	$\epsilon = \pm 1$ for buy/sell market orders	half of first gap behind the ask ( $\epsilon = 1$ ) or bid ( $\epsilon = -1$ )
$\pi = \text{CA}^0$	partial cancellation of the bid/ask queue	$\epsilon = \mp 1$ for buy/sell side cancellation	0
$\pi = \text{LO}^0$	limit order at the current best bid/ask	$\epsilon = \pm 1$ for buy/sell limit orders	0
$\pi = \text{CA}'$	complete cancellation of the best bid/ask	$\epsilon = \mp 1$ for buy/sell side cancellation	half of first gap behind the ask ( $\epsilon = 1$ ) or bid ( $\epsilon = -1$ )
$\pi = \text{LO}'$	limit order inside the spread	$\epsilon = \pm 1$ for buy/sell limit order	half distance of limit order from the earlier best quote on the same side

Table I: Summary of the 6 possible event types, the corresponding definitions of the event signs and gaps.

### 3. MANY-EVENT IMPACT MODELS

#### 3.1. Notations and definitions

The dynamics of the order book is rich and complex, and involves the intertwined arrival of many types of events. These events can be categorized in different, more or less natural types. In the empirical analysis presented below, we have considered the following six types of events:

- market orders that do not change the best price (noted  $\text{MO}^0$ ) or that do change the price (noted  $\text{MO}'$ ),
- limit orders at the current bid or ask ( $\text{LO}^0$ ) or inside the bid-ask spread so that they change the price ( $\text{LO}'$ ),
- and cancellations at the bid or ask that do not remove all the volume quoted there ( $\text{CA}^0$ ) or that do ( $\text{CA}'$ ).

Of course, events deeper in the order book could also be added, as well as any extra division of the above types into subtypes, if more information is available. For example, if the identity code of each order is available, one can classify each event according to its origin, as was done in Ref. [13]. The generic notation for an event type occurring at event time  $t$  will be  $\pi_t$ . The upper index ' ("prime") will denote that the event changed any of the best prices, and the upper index 0 that it did not. Abbreviations without the upper index ( $\text{MO}$ ,  $\text{CA}$ ,  $\text{LO}$ ) refer to both the price changing and the non-price changing event type. Every event is given a sign  $\epsilon_t$  according to its expected long-term effect on the price – the precise definitions are summarized in Table I. Note that the table also defines the gaps  $\Delta_{\pi,\epsilon}$ , which will be used later. We will rely on indicator variables denoted as  $I(\pi_t = \pi)$ . This expression is equal to 1 if the event at  $t$  is of type  $\pi$  and zero otherwise. We also use the notation  $\langle \cdot \rangle$  to denote the time average of the quantity between the brackets.

Let us now define the response of the price to different types of orders. The average behavior of price after events of a particular type  $\pi$  defines the corresponding *response function* (or average impact function):

$$\mathcal{R}_\pi(\ell) = \langle (p_{t+\ell} - p_t) \cdot \epsilon_t | \pi_t = \pi \rangle. \quad (7)$$

This is a correlation function between  $\epsilon_t I(\pi_t = \pi)$  at time  $t$  and the price change from  $t$  to  $t + \ell$ , normalized by the stationary probability of the event  $\pi$ , denoted as  $P(\pi) = \langle I(\pi_t = \pi) \rangle$ . This normalized response function gives the expected directional price change after an event  $\pi$ . Its behavior for all  $\pi$ 's is shown in Fig. 2(left). Tautologically,  $\mathcal{R}_\pi(\ell = 1) > 0$  for price changing events and  $\mathcal{R}_\pi(\ell = 1) = 0$  for the others. Empirically, all types of events lead, on average, to a price change in the expected direction, i.e.  $\mathcal{R}_\pi(\ell) > 0$ .

We will also need the "return" response function, upgrading the quantity  $\mathcal{S}(\ell)$  defined above to a matrix:

$$\mathcal{S}_{\pi_1, \pi_2}(\ell) = \langle I(\pi_{t+\ell} = \pi_2) \cdot r_{t+\ell} \cdot \epsilon_t | \pi_t = \pi_1 \rangle. \quad (8)$$

Clearly, as an exact identity:

$$\mathcal{R}_\pi(\ell) = \sum_{n=0}^{\ell-1} \sum_{\pi'} \mathcal{S}_{\pi, \pi'}(n). \quad (9)$$

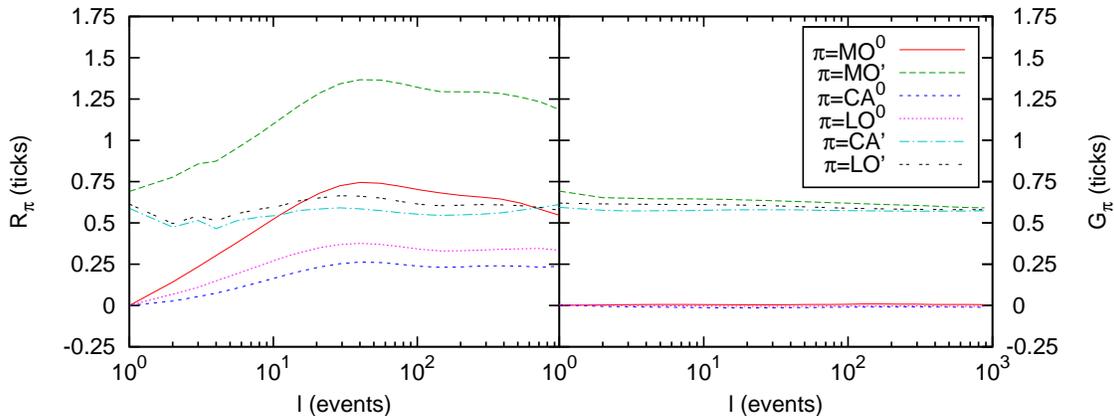


Figure 2: (*left*): The response function  $\mathcal{R}_\pi(\ell)$  and (*right*): the bare impact function  $G_\pi(\ell)$  for the TIM, for the data described in Sec. 4.1. The curves are labeled according to  $\pi$  in the legend. Note that  $G_\pi(\ell)$  is to a first approximation independent of  $\ell$  for all  $\pi$ 's. However, the small variations are real and important to reproduce the correct price diffusion curve  $D(\ell)$ .

Similarly, the signed-event correlation function is defined as:

$$C_{\pi_1, \pi_2}(\ell) = \frac{\langle I(\pi_t = \pi_1)\epsilon_t I(\pi_{t+\ell} = \pi_2)\epsilon_{t+\ell} \rangle}{P(\pi_1)P(\pi_2)}. \quad (10)$$

Our convention is that the first index corresponds to the first event in chronological order. Note that in general, there are no reasons to expect time reversal symmetry, which would impose  $C_{\pi_1, \pi_2}(\ell) = C_{\pi_2, \pi_1}(\ell)$ . If one has  $N$  event types, altogether there are  $N^2$  of these event-event correlation and return response functions.

### 3.2. The transient impact model (TIM)

Let us write down the natural generalization of the above transient impact model (TIM), embodied by Eq. (1), to the many-event case. We now envisage that each event type  $\pi$  has a “bare” time dependent impact given by  $G_\pi(\ell)$ , such that the price reads:

$$p_t = \sum_{t' < t} [G_{\pi_{t'}}(t - t')\epsilon_{t'} + \eta_{t'}] + p_{-\infty}, \quad (11)$$

where one selects for each  $t'$ , the propagator  $G_{\pi_{t'}}$  corresponding to the particular event type at that time. After straightforward calculations, the response function (7) can be expressed as

$$\mathcal{R}_{\pi_1}(\ell + 1) - \mathcal{R}_{\pi_1}(\ell) = \sum_{\pi_2} P(\pi_2) \left[ G_{\pi_2}(0)C_{\pi_1, \pi_2}(\ell) + \sum_{n > -\ell} [G_{\pi_2}(\ell + n + 1) - G_{\pi_2}(\ell + n)] C_{\pi_1, \pi_2}(n) \right]. \quad (12)$$

This is a direct extension of Eq. (4). One can invert the system of equations (12), to evaluate the unobservable  $G_\pi$ 's in terms of the observable  $\mathcal{R}_\pi$ 's and  $C_{\pi_1, \pi_2}$ 's – see Fig. 2(right). Note that we formulate the problem here in terms of the “derivatives” of the  $G_\pi$ 's since, as mentioned above, it is numerically much more stable to introduce new variables for the increments of  $\mathcal{R}$  and  $G$  and solve (12) in terms of those.

Once this is known, one can explicitly compute the lag dependent diffusion constant  $D(\ell) = \langle (p_{t+\ell} - p_t)^2 \rangle$  in terms of the  $G$ 's and the  $C$ 's, generalizing the corresponding result obtained in Ref. [3]: see Appendix A.

### 3.3. The history-dependent impact model (HDIM)

Now, if one wants to generalize the history dependent impact model (HDIM), Eq. (2), to the many-event case, one immediately realizes that the impact at time  $t$  depends on the type of event one is considering. For one thing, the

instantaneous impact of price non-changing events is trivially zero. For price changing events,  $\pi' = \text{MO}', \text{LO}', \text{CA}'$ , the instantaneous impact is given by:

$$r_t = \epsilon_t \Delta_t. \quad (13)$$

Here  $\Delta$  depends on the type of price changing event  $\pi'$  that happens then, and possibly also on the sign of that event. For example, if  $\pi_t = \text{MO}'$  and  $\epsilon_t = -1$  this means that at a sell market order executed the total volume at the bid. The midquote price change is  $-\Delta_t$ , which usually means that the second best level was at  $b_t - 2\Delta_t$ , where  $b_t$  is the bid price before the event. The factor 2 is necessary, because the ask did not change, and the impact is defined by the change of the midquote. Hence  $\Delta$ 's for MO's (and similarly CA's) correspond to half of the gap between the first and the second best quote *just before* the level was removed (see also Ref. [14]). Another example when  $\pi = \text{LO}'$  and  $\epsilon_t = -1$ . This means that at  $t$  a sell limit order was placed inside the spread. The midquote price change is  $-\Delta$ , which means that the limit order was placed at  $a_t - 2\Delta_t$ , where  $a_t$  is the ask price. Thus  $\Delta$  for LO's correspond to half of the gap between the first and the second best quote *right after* the limit order was placed. In the following we will call the  $\Delta$ 's *gaps*. For large tick stocks, these non-zero  $\Delta$ 's are most of the time equal to half a tick, and only very weakly fluctuating. For small tick stocks, substantial fluctuations of the spread, and of the gaps behind the best quotes, can take place. The generalization of Eq. (2) precisely attempts to capture these gap fluctuations, which are affected by the flow of past events. If we assume that the whole dynamical process is statistically invariant under exchanging buy orders with sell orders ( $\epsilon \Rightarrow -\epsilon$ ) and bids with asks, the dependence on the current non zero gaps on the past order flow can only include an even number of  $\epsilon$ 's. Therefore the lowest order model for non-zero gaps (including a constant term and a term quadratic in  $\epsilon$ 's) is:

$$\Delta_t |_{\pi_t = \pi'; \epsilon_t = \epsilon} = \Delta_{\pi'}^R + \sum_{t_1 < t} \kappa_{\pi_1, \pi'}(t - t_1) [\epsilon_{t_1} \epsilon - C_{\pi_1, \pi'}(t - t_1)] + \eta'_t, \quad (14)$$

where  $\kappa_{\pi_1, \pi'}$  are kernels that model the dependence of the gaps on the past order flow (note that  $\kappa_{\pi_1, \pi'}$  is a  $6 \times 3$  matrix) and  $\Delta_{\pi'}^R$  are the average realized gaps, defined as  $\langle \Delta_t |_{\pi_t = \pi'} \rangle$ , since the average of the second term in the right hand side is identically zero. Note that the last term, equal to  $\sum_{n>0} \sum_{\pi} P(\pi) \kappa_{\pi, \pi'}(n) C_{\pi, \pi'}(n)$ , was not explicitly included in our previous analysis, Ref. [11]. However, typical values are less than 1% of the average realized gap, and therefore negligible in practice. We set it to zero in the following.

Eq. (14), combined with the definition of the return at time  $t$ , Eq. (13), leads to our generalization of the HDIM, Eq. (2):

$$r_t = \Delta_{\pi_t}^R \epsilon_t + \sum_{t_1 < t} \kappa_{\pi_{t_1}, \pi_t}(t - t_1) \epsilon_{t_1} + \eta_t. \quad (15)$$

It is interesting to compare the above equation with its analogue for the transient impact model, which reads (after Eq. (11)):

$$r_t = G_{\pi_t}(1) \epsilon_t + \sum_{t_1 < t} [G_{\pi_{t_1}}(t - t_1 + 1) - G_{\pi_{t_1}}(t - t_1)] \epsilon_{t_1} + \eta_t. \quad (16)$$

The two models can only be equivalent if:

$$\kappa_{\pi_{t_1}, \pi}(\ell) \equiv G_{\pi_{t_1}}(\ell + 1) - G_{\pi_{t_1}}(\ell), \quad \forall \pi, \quad (17)$$

which means that the ‘‘influence matrix’’  $\kappa_{\pi_1, \pi}$  has a much constrained structure, which has no reason to be optimal. It is also *a priori* inconsistent since the TIM leads to a non zero price move even if  $\pi$  is a non price-changing event, since Eq. (17) is valid for all event types  $\pi$ . This is a major conceptual drawback of the TIM framework (although, as we will see below, the model fares quite well at reproducing the price diffusion curve).

The matrix  $\kappa_{\pi_1, \pi_2}$  can in principle be determined from the empirical knowledge of the response matrices  $\mathcal{S}_{\pi_1, \pi_2}$ , since:

$$\begin{aligned} \frac{1}{P(\pi_2)} \mathcal{S}_{\pi_1, \pi_2}(\ell) &= \frac{1}{P(\pi_2)} \langle I(\pi_{t+\ell} = \pi_2) \cdot r_{t+\ell} \cdot \epsilon_t |_{\pi_t = \pi_1} \rangle = \Delta_{\pi_2}^R C_{\pi_1, \pi_2}(\ell) \\ &+ \sum_{t' < t+\ell} \sum_{\pi} \kappa_{\pi, \pi_2}(t + \ell - t') \langle I(\pi_{t'} = \pi) \epsilon_{t'} I(\pi_t = \pi_1) \epsilon_t |_{\pi_{t+\ell} = \pi_2} \rangle. \end{aligned} \quad (18)$$

Note however that the last term includes a three-body correlation function which is not very convenient to estimate. At this stage and below, we need to make some approximation to estimate higher order correlations. We assume that

all three- and four-body correlation functions can be factorized in terms of two-body correlation functions, as if the variables were Gaussian. This allows to extract  $\kappa_{\pi_1, \pi_2}$  from a numerically convenient expression, used in [11]:

$$\frac{1}{P(\pi_2)} \mathcal{S}_{\pi_1, \pi_2}(\ell) = \Delta_{\pi_2}^R C_{\pi_1, \pi_2}(\ell) + \sum_{n > -\ell} \sum_{\pi} \kappa_{\pi, \pi_2}(n + \ell) C_{\pi_1, \pi}(n). \quad (19)$$

Knowing the  $\kappa_{\pi_1, \pi_2}$ 's and using the same factorization approximation, one can finally estimate the price diffusion constant, given in Appendix A. Although the factorization approximation used to obtain the diffusion constant looks somewhat arbitrary, we find that it is extremely precise when applied to the diffusion curve.

For large tick stocks, the gaps hardly vary with time and are all equal to 1 tick. In other words  $\kappa_{\pi_1, \pi} \approx 0$  and the model simplifies enormously, since now  $G_{\pi}(\ell) \equiv \Delta_{\pi}^R$ . In this limit, one therefore finds:

$$R_{\pi}(\ell) = \langle (p_{t+\ell} - p_t) \cdot \epsilon_t | \pi_t = \pi \rangle \approx \Delta_{\pi}^R + \sum_{0 < t' < \ell} \sum_{\pi_1} \Delta_{\pi_1}^R P(\pi_1) C'_{\pi, \pi_1}(t'), \quad (20)$$

which means that the total price response to some event can be understood as its own impact (lag zero), plus the sum of the biases in the course of future events, conditional to this initial event. These biases are multiplied by the average price change  $\Delta^R$  that these induced future events cause. Within the same model, the volatility reads:

$$D(\ell) = \langle (p_{t+\ell} - p_t)^2 \rangle \approx \sum_{0 \leq t', t'' < \ell} \sum_{\pi_1} \sum_{\pi_2} P(\pi_1) P(\pi_2) C_{\pi_1, \pi_2}(t' - t'') \Delta_{\pi_1}^R \Delta_{\pi_2}^R. \quad (21)$$

For small ticks, on the other hand, gaps do fluctuate and react to the past order flow; the influence matrix  $\kappa_{\pi_1, \pi'}$  describes how the past order flow affects the current gaps. If  $\kappa_{\pi_1, \pi'}(\ell)$  is positive, it means that an event of type  $\pi_1$  (price changing or not) tends to *increase* the gaps (i.e. reduce the liquidity) for a later price changing event  $\pi'$  in the same direction, and decrease the gap if the sign of the event  $\pi'$  is opposite to that of  $\pi_1$ .

## 4. MODEL CALIBRATION AND EMPIRICAL TESTS

### 4.1. Data

We have tested the above ideas on a set of data made of 14 randomly selected liquid stocks traded on the NASDAQ during the period 03/03/2008 – 19/05/2008, a total of 53 trading days (see Ref. [11] for a detailed presentation of these stocks and summary statistics). In order to reduce the effects of the intraday spread and liquidity variations we exclude the first 30 and the last 40 minutes of the trading days. The particular choice of market is not very important, many of our results were also verified on other markets, such as CME Futures, US Treasury Bonds and stocks traded at the London Stock Exchange.

Our sample of stocks can be divided into two groups: large tick and small tick stocks. Large tick stocks are such that the bid-ask spread is almost always equal to one tick, whereas small tick stocks have spreads that are typically a few ticks. The behavior of the two groups is quite different, for example, the events which change the best price have a relatively low probability for large tick stocks (about 3% altogether), but not for small tick stocks (up to 40%). Note that there is a number of stocks with intermediate tick sizes, which to some extent possess the characteristics of both groups. Technically, they can be treated in exactly the same way as small tick stocks, and all our results remain valid. However, for the clarity of presentation, we will not consider them explicitly.

As explained above, we restrict ourselves to events that modify the bid or ask price, or the volume quoted at these prices. Events deeper in the order book are unobserved and will not be described: although they do not have an immediate effect on the best quotes, our description is still incomplete. Furthermore, we note that the stocks we are dealing with are traded on multiple platforms. This may account for some of the residual discrepancies reported below.

Since we consider 6 types of events, there are 6 response functions  $\mathcal{R}_{\pi}$  and propagators  $G_{\pi}$ , 36 correlation functions  $C_{\pi_1, \pi_2}$ . However, since the return response functions  $\mathcal{S}_{\pi_1, \pi_2}$  and the influence kernels  $\kappa_{\pi_2, \pi_1}$  are non zero only when the second event  $\pi_2$  is a price changing event, there are only  $3 \times 6 = 18$  of them.

### 4.2. The case of large ticks

As explained in the previous section, the case of large ticks is quite simple since the gap fluctuation term of HDIM can be neglected altogether. As shown in Ref. [11], the predictions given by Eqs. (20) and (21) are in very good

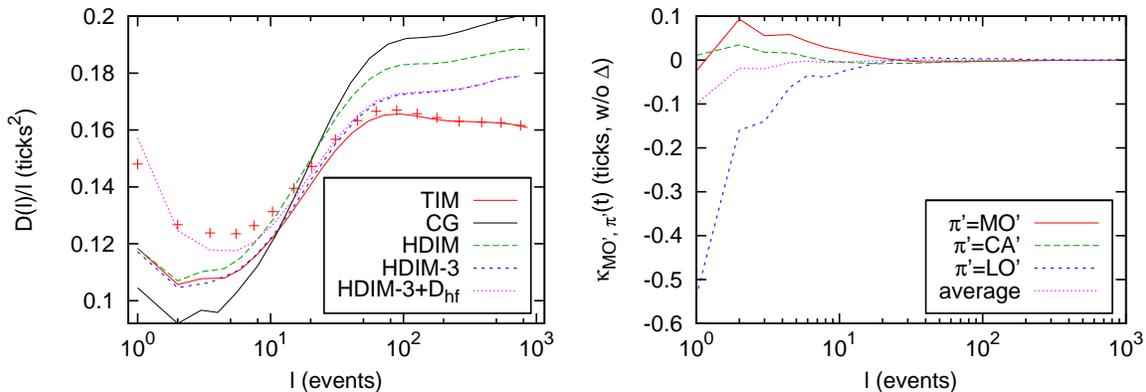


Figure 3: (left)  $D(\ell)/\ell$  and its approximations. Crosses correspond to the data. The constant gap (CG) model corresponds to  $G_\pi(\ell) \equiv G_\pi(1)$ . TIM corresponds to the temporary impact model calibrated on returns. The curve for HDIM uses the approximate calibration of  $\kappa$ 's, HDIM-3 is taking 3 times the  $\kappa$ 's as from the calibration. We also indicate HDIM-3 with adding the constant  $D_{hf} = 0.04$ . Note that the vertical scale is different from Fig. 1, since in the time clock is different in the two cases (all events vs. trades in Fig. 1). (right) Comparison of the three non zero  $\kappa_{MO', \pi'}(\ell)$  with their average over  $\pi'$ . Note that  $\kappa_{MO', MO'} < 0$ : after an  $MO'$  event, gaps on the same side are on average smaller.

agreement with the empirical determination of the  $\mathcal{R}_\pi(\ell)$  and the price diffusion  $D(\ell)$ . Small remaining discrepancies can indeed be accounted for by adding the gap fluctuation contribution, of the order of a few percents.

The temporary impact model, on the other hand, is not well adapted to describe large tick stocks, for the following reason: when Eq. (12) is used to extract  $G_\pi(\ell)$  from the data, small numerical errors may lead to some spurious time dependence. But as far as  $D(\ell)$  is concerned, any small variation of  $G_\pi$  is amplified through the second term of Eq. (A1) which is an infinite sum of positive terms. As noted in Ref. [11], this leads to large discrepancies between the predicted  $D(\ell)$  and its empirical determination. At any rate, one should clearly favor the calibration of  $G_\pi(\ell)$  using Eq. (12) rather than the analogue of Eq. (4).

### 4.3. The case of small ticks

The case of small ticks is much more interesting, since in this case the role of gap fluctuations is crucial, and is a priori a stringent test for the two models on stage.

#### 4.3.1. TIM

Within the temporary impact model, the response functions  $\mathcal{R}_\pi(\ell)$  are tautologically accounted for, since they are used to calibrate the propagators  $G_\pi(\ell)$  using Eq. (12). Once the  $G_\pi(\ell)$ 's are known (see Fig. 2, where  $\mathcal{R}_\pi(\ell)$  and  $G_\pi(\ell)$  are shown for small tick stocks), one can compute the time dependent diffusion coefficient  $D(\ell)/\ell$ , and compare with empirical data. This is shown in Fig. 3(left). Note that we calibrate the  $G_\pi(\ell)$  for each stock separately, compute  $D(\ell)/\ell$  in each case, and then average the results over all stocks. The agreement is surprisingly good for long times, while for shorter times the model underestimates price fluctuations, which is expected since the model does not allow for high frequency fluctuations. We also show the prediction based on the constant gap approximation,  $\kappa_{\pi_1, \pi_2} \equiv 0$ . Although Fig. 2 suggests that this is an acceptable assumption, we see that  $D(\ell)$  is overestimated. As will be argued below, gaps do adapt to past order flow, and the net effect of the gap dynamics is to reduce the price volatility. We finally note that calibrating the  $G_\pi(\ell)$  on the response functions directly (and not on their derivatives), as was done in [11], leads to much poorer results for the diffusion coefficient  $D(\ell)/\ell$ .

#### 4.3.2. HDIM

We now turn to the history-dependent impact model. As explained above, we determine the influence kernels  $\kappa_{\pi_1, \pi_2}(\ell)$  using Eq. (19). We plot in Fig. 4 the resulting “integrated impact” on the future gaps of all 6  $\pi_1$  events,

which we define as<sup>3</sup>:

$$\delta G_{\pi}^*(\ell) = \sum_{n=1}^{\ell-1} \sum_{\pi'} P(\pi') \kappa_{\pi, \pi'}(n), \quad (22)$$

As explained in Ref. [11],  $\delta G_{\pi}^*(\ell)$  captures the contribution of the gap “compressibility” to the impact of an event of type  $\pi$  up to a time lag  $\ell$ , leaving the sequence of events unchanged. If  $\kappa_{\pi, \pi'}(n)$  were independent of  $\pi'$ , as postulated in the TIM, one would have  $\delta G_{\pi}^*(\ell) = G_{\pi}(\ell) - G_{\pi}(1)$  as an identity. The agreement turns out to be excellent (see Fig. 4), which was not guaranteed a priori since the HDIM is calibrated on a much larger set of correlation functions.

However, this does not mean that  $\kappa_{\pi, \pi'}(n)$  are necessarily independent of  $\pi'$ . To illustrate the point that Eq. (17) is too restrictive, Fig. 3(right) compares the three  $\kappa_{MO', \pi'}$ s, which are clearly different from one another. Note that the average over  $\pi'$  is negative, meaning that  $MO'$  events tend to “harden” the book (i.e. after an  $MO'$  event, gaps on the same side are on average smaller). This is true for all price changing events, while (perhaps surprisingly) small market orders  $MO^0$  “soften” the book:  $\delta G_{MO^0}^*$  is positive and gaps tend to grow. Queue fluctuations ( $CA^0$  and  $LO^0$ ) seem less important, but for small ticks these types of events also harden the book. Note finally that for large ticks  $\delta G^*$ 's are found to be about two orders of magnitude smaller, which confirms that gap fluctuations can be neglected in that case.

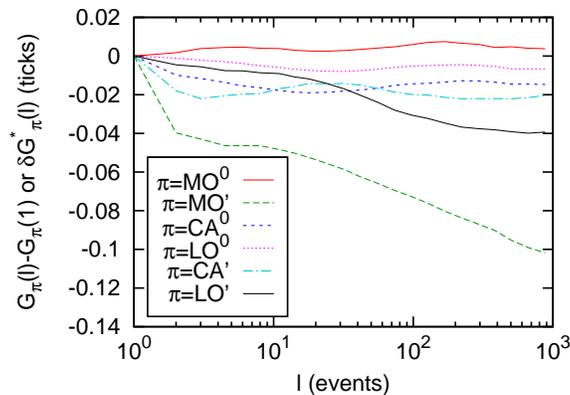


Figure 4: The integrated impact on the future gaps  $\delta G_{\pi}^*(\ell)$  in the HDIM estimated via Eq. (17). The results are indistinguishable from  $G_{\pi}(\ell) - G_{\pi}(1)$  calculated for the TIM. The curves are labeled according to  $\pi$  in the legend.

Now, Eq. (19) relies on the factorization of a three-point correlation function and is not exact, so there is no guarantee that the response functions  $\mathcal{R}_{\pi}(\ell)$  are exactly reproduced using this calibration method. In order to check this approximation, we have simulated an artificial market dynamics where the price evolves according to Eq. (15), with the true (historical) sequence of signs and events and  $\eta_t = 0$ . The kernels  $\kappa_{\pi_1, \pi_2}$  are calibrated using Eq. (19). This leads to the predictions shown as dashed lines in Fig. 5. The agreement can be much improved by simply multiplying all  $\kappa$ 's by a factor 3, see Fig. 5. Of course, some discrepancies remain and one should use the historical simulation systematically to determine the optimal  $\kappa$ 's. This is, however, numerically much heavier and an improved analytical approximation of the three-point correlation function, that would allow a more accurate workable calibration, would be welcome.

Finally, we computed  $D(\ell)$  for the HDIM using (A2) in the Appendix. Here again, we have tested the quality of the factorization approximation using the same historical simulation. In this case, the  $D(\ell)$  curve is indistinguishable from its approximation, so any discrepancy between the data and formula (A2) cannot be blamed on its approximate nature, but rather on an inadequate calibration of the  $\kappa$ 's.

The result is given in Fig. 3(left) together with the previous theoretical predictions and the empirical data.<sup>4</sup> With the naive calibration the HDIM turns out to be worse than the TIM for large lags: it overestimates  $D(\ell)$  by 15% or so. Increasing the  $\kappa$ 's by a factor 3 again greatly improves the fit but part of the discrepancy remains. For small lags, one

<sup>3</sup> Note that this definition is compatible with the one given in Ref. [11], because of a slight change in the interpretation in the  $\kappa$  kernels here.

<sup>4</sup> The  $D(\ell)$ -HDIM shown here is indistinguishable from the one appearing in Fig. 16 of Ref. [11].

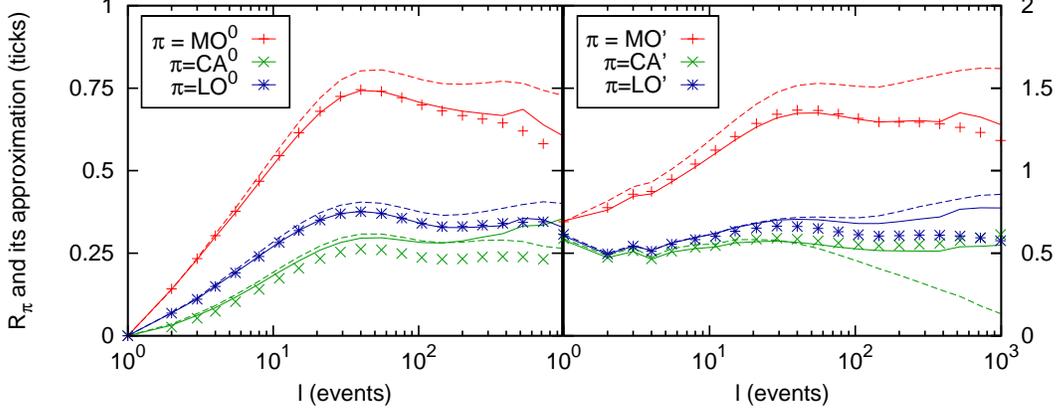


Figure 5:  $\mathcal{R}_\pi(\ell)$  and their approximation with the HDIM. Symbols correspond to data, they are perfectly in line with the model prediction under the assumption that approximation (19) is correct. The dashed lines correspond to the response function of an actual simulation of the model with the  $\kappa$ 's calibrated via Eq. (19). The solid lines correspond to the simulation if we increase all calibrated  $\kappa$ 's by a factor 3 (HDIM-3). The colors vary according to  $\pi$  as shown in the legend.

needs to add a constant contribution  $D_{\text{hf}} \approx 0.04$  ticks squared to match the data.<sup>5</sup> The HDIM produces a significant improvement over the constant gap model, because it explicitly includes the effect of gap fluctuations. However, since the calibration procedure relies on an approximation, we do not reproduce the response functions exactly. Hence the better founded model (HDIM) fares worse in practice than a model with theoretical inconsistencies (TIM). As noted above, a better calibration procedure for the  $\kappa$ 's could improve the situation.

At any rate, numerical discrepancies should be expected regardless of the fitting procedure, since we have neglected several effects, which must be present. These include (i) all volume dependence, (ii) unobserved events deeper in the book and on other platforms and (iii) higher order, non-linear contributions to model history dependence. On the last point, we note that based on symmetry arguments, the gap fluctuation term may include higher order terms of the form:

$$\sum_{t_1, t_2, t_3 < t'} \kappa_{\pi_{t_1}, \pi_{t_2}, \pi_{t_3}; \pi_{t'}} (t' - t_1, t' - t_2, t' - t_3) \epsilon_{t_1} \epsilon_{t_2} \epsilon_{t_3} \epsilon_{t'}, \quad (23)$$

or with a larger (even) number of  $\epsilon$ 's. The presence of a four  $\epsilon$  term is in fact suggested by the data shown in Fig. 13 of Ref. [11], and also by more recent analysis [15]. It would be interesting to study these effects in detail, and understand their impact on price diffusion.

## 5. CONCLUSION

Let us summarize what we have tried to accomplish in the present paper. Our aim was to provide a general framework to describe the impact of different events in the order book, in a way that is flexible enough to deal with any classification of these events (provided this classification makes sense).<sup>6</sup> We have specifically considered market orders, limit orders and cancellation at the best quotes, further subdividing each category into price-changing and non price-changing events, giving a total of 6 types. In trying to generalize previous work, which focused on the impact of market orders only, we have discovered that two different models can be envisaged. These are equivalent when only a single event type, market orders regardless of their aggressivity, are taken into account. One model posits that each event type has a temporary impact (TIM), whereas the other assumes that only price-changing events have a direct impact, which is itself modulated by the past history of all events, a model we called “history dependent impact” (HDIM).

<sup>5</sup> This contribution accounts for high frequency “noise” in the data that the model is not able to reproduce, as, for example, sequences of placement and cancellation of the same limit order inside the gap.

<sup>6</sup> See Ref. [13] for an application of this method to orders with brokerage codes.

The TIM is a natural extension of Hasbrouck’s VAR model to a multi-event setting: one writes a Vector Autoregression model for the return at time  $t$  in terms of *all* signed past events, but neglects the direct influence of past returns themselves (although these would be easy to include if needed). We have discussed the fact that TIMs are, strictly speaking, inconsistent since they assign a non-zero immediate impact to non price-changing events. Still, provided the model is correctly calibrated using returns (see Eq. (12)), we find that the TIM framework allows one to reproduce the price diffusion pattern surprisingly accurately.

The HDIM family can also be thought of as a VAR model, although one now distinguishes between different types of event-induced returns before regressing them on past events. The HDIM is interesting because it gives a very appealing interpretation of the price changing process in terms of history dependent “gaps” that determine the amplitude of the price jump if a certain type of price-changing event takes place. We have in particular defined a lag-dependent,  $6 \times 3$  “influence matrix” (called  $\kappa_{\pi, \pi'}$  in the text), which tells us how much, on average, an event of type  $\pi$  affects the immediate impact of a  $\pi'$  price-changing event of the same sign in the future.

The HDIM therefore envisages the dynamics of prices as consisting of three processes: instantaneous jumps due to events, events inducing further events and thereby affecting the future jump *probabilities* (described by the correlation between events), and events exerting pressure on the gaps behind the best price and thereby affecting the future jump *sizes* (described by the  $\kappa$ ’s). By describing this third effect with a linear regression process, we came up with the explicit model (15), that can be calibrated on empirical data provided some factorization approximation is made (which unfortunately turns out not to be very accurate, calling for further work on this matter). This allows one to measure the influence matrix  $\kappa$  and its lag dependence. We find in particular that price-changing events, such as aggressive market orders  $MO'$ , tend to reduce the impact of later events of the same sign (i.e. a buy  $MO'$  following a buy  $MO'$ ) but increase the impact of later events of the opposite sign. As stressed in Refs. [6, 8, 9], this history dependent asymmetric liquidity is the dominant effect that mitigates persistent trends in prices which would otherwise be induced by the long-ranged correlation in the sign of market orders.

In spite of these enticing features, we have found that the HDIM leads to a worse determination of the price diffusion properties than the TIM. The almost perfect agreement between the TIM prediction and empirical data is perhaps accidental, but it may also be that TIMs (that have less parameters) are numerically more robust than HDIMs. For HDIMs, a more accurate calibration procedure is needed. This could be achieved either by finding a better, workable approximation for the three-point correlation function, or by using a purely numerical approach based on a historical simulation of the HDIM. On the other hand, some effects have been explicitly neglected, such as the role of unobserved events deeper in the book and on other platforms, or possible non-linearities in the history dependence of gaps. It would be very interesting to investigate the relevance of these effects, and to come up either with a fully consistent version of HDIM, or with a convincing argument for why the TIM appears to be particularly successful.

In any case, we hope that the intuitive and versatile framework that we proposed above, together with operational calibration procedures, will help making sense of the highly complex and intertwined sequences of events that take place in the order books, and allows one to build a comprehensive theory of price formation in electronic markets.

### Acknowledgments

The authors are grateful to Emmanuel Sérié for his ideas on fitting impact kernels. They also thank Bence Tóth for his critical reading and comments.

### Appendix A: Expression of the price diffusion for the TIM and HDIM

We give here the rather ugly looking explicit expressions for the diffusion curve  $D(\ell)$  in both models. For the TIM, one gets as an exact expression:

$$\begin{aligned}
D(\ell) = & D_0 \ell + \sum_{0 \leq n < \ell} \sum_{\pi_1} G_{\pi_1}(\ell - n)^2 P(\pi_1) + \sum_{n > 0} \sum_{\pi_1} [G_{\pi_1}(\ell + n) - G_{\pi_1}(n)]^2 P(\pi_1) \\
& + 2 \sum_{0 \leq n < n' < \ell} \sum_{\pi_1, \pi_2} G_{\pi_1}(\ell - n) G_{\pi_2}(\ell - n') C_{\pi_1, \pi_2}(n' - n) \\
& + 2 \sum_{0 < n < n' < \ell} \sum_{\pi_1, \pi_2} [G_{\pi_1}(\ell + n) - G_{\pi_1}(n)] [G_{\pi_2}(\ell + n') - G_{\pi_2}(n')] C_{\pi_1, \pi_2}(n - n') \\
& + 2 \sum_{0 \leq n < \ell} \sum_{n' > 0} \sum_{\pi_1, \pi_2} G_{\pi_1}(\ell - n) [G_{\pi_2}(\ell + n') - G_{\pi_2}(n')] C_{\pi_2, \pi_1}(n' + n). \tag{A1}
\end{aligned}$$

where  $D_0$  is the variance of the noise term  $\eta_t$ .

For the HDIM, on the other hand, one has to use a factorization approximation to compute 3- and 4-point correlation functions in terms of 2-point correlations. One can finally estimate the price diffusion constant, which is given by the following approximate equation [11]:

$$\begin{aligned}
D(\ell) = \langle (p_{t+\ell} - p_t)^2 \rangle \approx D_0 \ell + & \sum_{0 \leq t', t'' < \ell} \sum_{\pi_1} \sum_{\pi_2} P(\pi_1) P(\pi_2) C_{\pi_1, \pi_2}(t' - t'') \Delta_{\pi_1}^R \Delta_{\pi_2}^R + \\
& 2 \sum_{-\ell < t < \ell} \sum_{\pi_2, \pi_3} \sum_{\tau > 0} (\ell - |t|) \Delta_{\pi_3}^R \kappa_{\pi_2, \pi_3}^+(\tau, t) C_{\pi_2, \pi_3}(t + \tau) P(\pi_2) P(\pi_3) + \\
& \sum_{-\ell < t < \ell} \sum_{\pi_2, \pi_4} \sum_{\tau, \tau' > 0} (\ell - |t|) \kappa_{\pi_2, \pi_4}^{++}(\tau, \tau', t) C_{\pi_2, \pi_4}(\tau - \tau' + t) P(\pi_2) P(\pi_4), \tag{A2}
\end{aligned}$$

where  $D_0$  is again the variance of the noise  $\eta_t$ ,

$$\kappa_{\pi_2, \pi_3}^+(\tau, t) = \sum_{\pi_1} \kappa_{\pi_2, \pi_1}(\tau) [I(t=0)I(\pi_1 = \pi_3) + I(t \neq 0)P(\pi_1) + I(t = -\tau)P(\pi_1)\Pi_{\pi_2 \pi_1}(\tau)], \tag{A3}$$

and, for  $t \geq 0$ ,

$$\begin{aligned}
\kappa_{\pi_2, \pi_4}^{++}(\tau, \tau', t) = \sum_{\pi_1, \pi_3} \kappa_{\pi_2, \pi_1}(\tau) \kappa_{\pi_4, \pi_3}(\tau') \{ & I(t = \tau') I(\pi_1 = \pi_4) P(\pi_3) + \\
& I(t \neq \tau') P(\pi_1) P(\pi_3) [\Pi_{\pi_1, \pi_3}(t) + 1] \}, \tag{A4}
\end{aligned}$$

whereas for  $t < 0$ , we use  $\kappa_{\pi_2, \pi_4}^{++}(\tau, \tau', -t) = \kappa_{\pi_4, \pi_2}^{++}(\tau', \tau, t)$ . We also introduced a correlation function between event types as [11]:

$$\Pi_{\pi_1, \pi_2}(\ell) = \frac{P(\pi_{t+\ell} = \pi_2 | \pi_t = \pi_1)}{P(\pi_2)} - 1 \equiv \frac{\langle I(\pi_t = \pi_1) I(\pi_{t+\ell} = \pi_2) \rangle}{P(\pi_1) P(\pi_2)} - 1. \tag{A5}$$

- 
- [1] J. Hasbrouck, *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading* (Oxford University Press, 2007).
- [2] S. Mike and J. D. Farmer, *Journal of Economic Dynamics and Control* **32**, 200 (2008).
- [3] J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart, *Quantitative Finance* **4**, 176 (2004).
- [4] J.-P. Bouchaud, J. Kockelkoren, and M. Potters, *Quantitative Finance* **6**, 115 (2006).
- [5] R. K. Lyons, *The Microstructure Approach to Exchange Rates* (MIT Press, 2006).
- [6] J.-P. Bouchaud, J. D. Farmer, and F. Lillo, *How markets slowly digest changes in supply and demand*, in *Handbook of Financial Markets: Dynamics and Evolution*, North-Holland, Elsevier (2009).
- [7] P. Weber and B. Rosenow, *Quantitative Finance* **5**, 357 (2005).
- [8] A. Gerig, *A theory for market impact: How order flow affects stock price*, *PhD thesis* (2008), arXiv:0804.3818.
- [9] J. D. Farmer, A. Gerig, F. Lillo, and S. Mike, *Quantitative Finance* **6**, 107 (2006).
- [10] N. Hautsch and R. Huang, *The market impact of a limit order* (2009), working Paper 2009/23, Center for Financial Studies, Frankfurt am Main.
- [11] Z. Eisler, J.-P. Bouchaud, and J. Kockelkoren (2011), arXiv:0904.0900, to appear in *Quantitative Finance*.
- [12] E. Sérié (2010), unpublished report, Capital Fund Management, Paris, France.
- [13] B. Tóth, Z. Eisler, F. Lillo, J.-P. Bouchaud, J. Kockelkoren, and J. Farmer (2011), arXiv:1104.0587.
- [14] J. D. Farmer, L. Gillemot, F. Lillo, S. Mike, and A. Sen, *Quantitative Finance* **4**, 383 (2004).
- [15] B. Tóth (2011), in preparation.