POLITECNICO DI TORINO

Master degree in
PHYSICS OF COMPLEX SYSTEMS

MASTER DEGREE THESIS

A mechanism for the latent liquidity revealing into the limit order book

Supervisor:
Dr. Luca DALL'ASTA

Candidate:
Lorenzo DALL'AMICO

Co - supervisor:
Dr. Michael BENZAQUEN

A.A. 2017/2018

Declaration of Authorship

I, Lorenzo DALL'AMICO, declare that this thesis titled, "A mechanism for the latent liquidity revealing into the limit order book" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: July 14, 2018

POLITECNICO DI TORINO

Abstract

A project funded by

LadHyX- École Polytechnique

and directed by

Dr. Michael BENZAQUEN

MASTER DEGREE THESIS

A mechanism for the latent liquidity revealing into the limit order book

by Lorenzo DALL'AMICO

Latent order book models have allowed for significant progress in the understanding of price formation in financial markets. In particular they were able to reproduce a number of stylized facts such as the square root impact of metaorders. An important question that is raised – if one wants to bring such models closer to real market data – is that of the connection between the latent order book and the revealed order book, observable and quantifiable. It is here suggested a self-consistent mechanism for the revelation of latent liquidity that allows for quantitative estimation of the latent order book from real market data. In particular, our setup allows to track the revealed liquidity as function of revelation rates and incentive to give away information by revealing one's intentions. We confront our results to real market data and discuss market stability. Finally we run a numerical simulation to compute the price impact and discuss different regimes.

Acknowledgements

A tree is as high as its roots grow deeper into the ground.

I feel very lucky today when I look back and see that I could get as high as my abilities allowed me to, with nothing that kept me behind.
I want to thank my parents for the strength and support they gave me and for having being a safe harbor where I was never asked to dock longer then I wished.
A hug goes to my sister who, after all, was able to live together with my presence that, I'm aware, might have been somewhat cumbersome from time to time. But she loved me anyway the way I was.
A thank goes to all the comrades of this small two years long adventure in which we had to share, willing or not, houses, professors, bureaucracy, worries and satisfactions. Things changed a lot for me and I'm glad I had someone to share these little but important steps with.
I thank then my closest friends, those with whom I would have liked to share something more then a quick beer during a weekend. It's true, is not about how much time you spend together, but its quality and, even if sometimes its hard to hear those words, I know that among them I have a place where to go and call home.

These are my roots, what kept me standing when things appeared harder and what allowed me to move forward.

Then there is a special thanksgiving that should be kept apart: to my wonderful girlfriend, Giulia. We took together this road that at some point looked harder then our expectations, but that gave us more then what the distance tried to take from us. She was for me a reason, a spur to go on. She was to me the sun that makes the tree to blossom.

Finally I'd like to thank my advisor, Dr. Michael Benzaquen, for the great support and the many stimuli he gave me, making me feel like a part of his team from day one. Thanks also to Antoine Fosset with whom I shared the desk in the past months and who was crucial to approach everyday problems.

Ringraziamenti

Un albero cresce quanto le sue radici sono profonde nella terra.

Mi sento molto fortunato oggi quando guardandomi indietro vedo di essere arrivato là dove le mie abilità mi hanno concesso, senza nulla che mi tenesse indietro.
Vorrei anzitutto rigraziare i miei genitori per la forza ed il supporto che mi hanno sempre dato e per essere stati per me un porto sicuro dove non mi è mai stato richiesto di attraccare più di quanto desiderassi.
Un abbraccio va a mia sorella che, dopo tutto, è stata capace di convivere con una presenza che, lo so, è stata un po' ingombrante di tanto in tanto. Ma mi ha sempre voluto bene, così com'ero.
Un grazie va ai compagni di questa piccola avventura di due anni in cui abbiamo condiviso, volenti o nolenti, case, professori, burocrazia, preoccupazioni e soddisfazioni. Molto è cambiato per me e sono contento di aver avuto qualcuno con cui condividere questi piccoli ma importanti passi.
Ringrazio poi i miei amici più intimi, quelli per i quali vorrei aver potuto condividere di più che una birretta veloce il sabato sera. E' vero, non è la quantità di tempo passato insieme, ma la qualità e, anche se alle volte è difficile sentire queste parole, io so che tra loro ho un posto dove tornare e che posso chiamare casa.

Queste sono le mie radici, quelle che mi hanno tenuto in piedi quando le cose sembravano più difficili e mi hanno concesso di andare avanti.

C'è poi un ringraziamento speciale, da fare a parte: alla mia bellissima fidanzata Giulia. Abbiamo intrapreso insieme questa strada che, a tratti, si è rivelata più dura di quanto atteso, ma che ci ha dato più di quanto la distanza ci abbia tolto. Lei è stata per me una ragione, uno sprone ad andare avanti. Lei è stata per me il sole che fa fiorire l'albero.

Infine vorrei ringraziare il mio relatore, Dr. Michael Benzaquen, per il supporto e gli stimoli che mi ha sempre dato, facendomi sentire parte del suo team dal primo giorno. Grazie anche ad Antoine Fosset con cui ho condiviso la scrivania in questi ultimi mesi e che è stato di cruciale importanza nell'approcciare i problemi di tutti i giorni.

Contents

List of Figures

Chapter 1

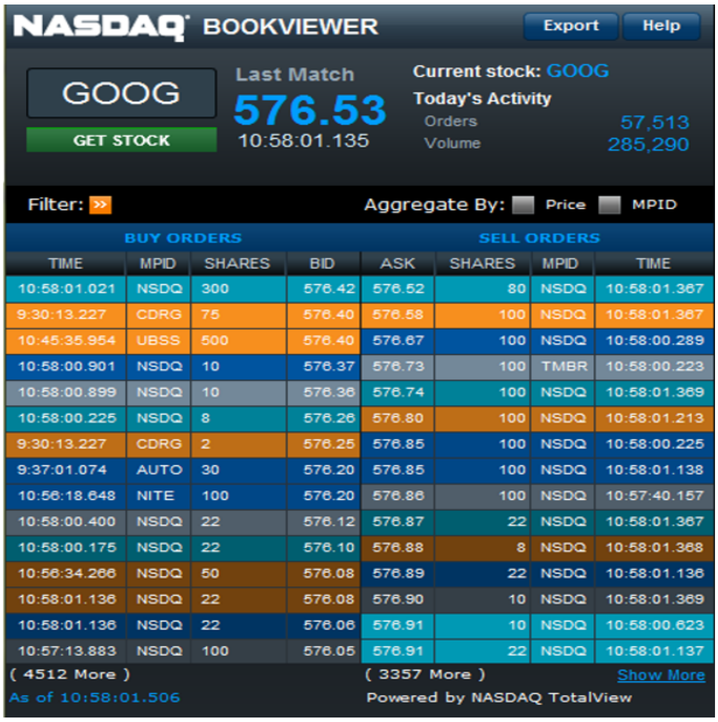
Introduction and motivation

The market is a complex systems where many agents act with the purpose of maximizing their utility. The sum of the individuals' actions will not correspond to the whole system [2] that presents a very fascinating behavior of self-organization that attracted the study of many theoretical physicists. Even if we are only at the beginning of such rigorous studies in this subject and only few stylized facts are known, the approach of physics proved to be able to give important elements in the microscopical modeling of financial market [15].
In this chapter we aim to give a description of the structure of the order book with some basic definitions and vocabulary. Few considerations will then be made about the strategies adopted by the traders in the market and, related to that, some relevant statistical observations. Finally we'll give a short review of how the problem of modeling the limit order book has been tackled so far concerning the contest in which this work is located.

1.1 The limit order book: structure and definitions

The most basic element on financial instruments is the *equity* or *share*. This corresponds to the ownership of a piece of a company and the profit that each trader is willing to make from it comes from how the stock value changes with time. It then becomes a gambling game where each traders aims at buying at a low price and selling back at higher one [53].
The limit order book is a device where the to-be-executed orders of the market participants are stored. It also takes the name of *continuous double auction* . *Continuous* is referred to time, meaning that at any moment – up to the market servers time discretization – people can make an action on the book; *double auction* is because it is divided into two sides: the buyers – that are on the *bid* side – and the sellers – that are on the *ask* side –. Each trader can then post an order that will be characterized by three key quantities: the sign (buy or sell), the volume (the

quantity he wants to trade) and the price at which he wants to trade it [15, 17, 38]. In Fig.(1.1) a snapshot of a limit order book.



TIME	MPID	SHARES	BID	ASK	SHARES	MPID	TIME
10:58:01.021	NSDQ	300	576.42	576.52	80	NSDQ	10:58:01.367
9:30:13.227	CDRG	75	576.40	576.58	100	NSDQ	10:58:01.367
10:45:35.954	UBSS	500	576.40	576.67	100	NSDQ	10:58:00.289
10:58:00.901	NSDQ	10	576.37	576.73	100	TMBR	10:58:00.223
10:58:00.899	NSDQ	10	576.36	576.74	100	NSDQ	10:58:01.369
10:58:00.225	NSDQ	8	576.26	576.80	100	NSDQ	10:58:01.213
9:30:13.227	CDRG	2	576.25	576.85	100	NSDQ	10:58:00.225
9:37:01.074	AUTO	30	576.20	576.85	100	NSDQ	10:58:01.138
10:56:18.648	NITE	100	576.20	576.86	100	NSDQ	10:57:40.157
10:58:00.400	NSDQ	22	576.12	576.87	22	NSDQ	10:58:01.367
10:58:00.175	NSDQ	22	576.10	576.88	8	NSDQ	10:58:01.368
10:56:34.266	NSDQ	50	576.08	576.89	22	NSDQ	10:58:01.136
10:58:01.136	NSDQ	22	576.08	576.90	10	NSDQ	10:58:01.369
10:58:01.136	NSDQ	22	576.06	576.91	10	NSDQ	10:58:00.623
10:57:13.883	NSDQ	100	576.05	576.91	22	NSDQ	10:58:01.137

FIGURE 1.1: *A snapshot of a limit order book* : on the left column we see buy orders starting from the best bid and on the right one we see the sell orders starting from the best ask, with $a_t > b_t$. Each column is characterized by the the number of shares at that price and the price itself.
Source: https://data.nasdaq.com/BookViewer.aspx

The lowest sell order in the limit order book is called *best ask* – a_t – and equivalently the highest buy order is the *best bid* – b_t –. Since the transaction happens as soon as two orders of the opposite kind meet at the same price, then $b_t < a_t \quad \forall t$. The difference between the best ask and the best bid is called bid-ask spread $s_t = a_t - b_t$. The price at which the last transaction was made is called *trade price* and all transactions are conveyed by the market order book. The trade price is the most important quantity for the market participant, however it is not well defined from the static point of view: what we know is that $p_t = b_{t-1}$ or $p_t = a_{t-1}$. This means that the trade price is a dynamical concept and we are not able to define it from a stationary snapshot of the order book as the one in Fig.(1.1). We therefore need to introduce another important quantity: the *mid price*. We define it as the mid-point between the best bid and the best ask, which is the price that is usually displayed as reference of a given stock and is always well defined.
The actions a trader can post are of three kinds. By denoting with x the price of the considered

order we can distinguish:

- *Limit orders*: in this case $x > b_t$ for sell orders and $x < a_t$ for buy orders, so we are in the circumstance in which one is not hitting the best offer on the opposite side. The trader secures the price at which its order will be executed, but not the time, entering in a waiting list. Limit orders are the ones that show up in the limit order book.
- *Market orders*: in this case $x \leq b_t$ for sell orders and $x \geq a_t$ for buy orders. Here is what happens when the bid and ask orders meet and the trade is executed.
- *Cancellation orders*: in this case the order that used to be in the limit order book is removed. The only way to change one's position inside the order book is to remove the order and then place it at a different price, so these orders are very important to describe the dynamics of the book.

As just mentioned, the limit orders are put in a queue. The sequence in which they are performed goes first according to the price – the execution of a market order will always happen at the best quote – then to the arrival time in a *first in first out* kind of way. By placing a large enough market order it is possible to change the best ask/bid and therefore the mid price. The average change of price due to the arrivals of market orders is called *impact* and it will be of central importance in the following chapters.

Another important concept is that of *liquidity*. We don't have a rigorous definition of liquidity, however we can say that a liquid market should have these three characteristics [21, 28, 42, 54]:

- *Depth*: the orders placed in the limit order book cover a large range of prices
- *Breath*: the market has a large number of participants so to guarantee a small price variation when placing an order (large liquidity $\rightarrow$ small impact)
- *Resilience*: the market reacts quickly when brought out of equilibrium

An alternative definition we can give is that market orders consume liquidity while limit orders provide it. The decision whether to post a market or a limit order is mainly dictated by the impatience of the trader and the need he has to have the deal done. However the equilibrium between liquidity taking and provision is a well established phenomenon and we'll give a more detailed description of how this happens in the next paragraph.
A fundamental aspect in the strategy making is the fact that people want to have their orders executed as soon as possible. To jump on top of the queue one simply has to propose the best offer. However, in the meantime, no one wants to trade at an inconvenient price, so the

position in the limit order book is the compromise between these two behaviors. As a consequence, the average shape of the limit order book will be an initially increasing function of the modulus of the distance from the mid price – the closer we get to the best quote the more we are likely to propose a disadvantageous price and so fewer people will be posting orders there – that reaches a maximum and then decreases – orders too far from the best quote have to wait too long to be executed and are posted less often – [17].

Of crucial importance when modeling the limit order book is the *tick size* – π –. It represents the discretization of the price axis where orders can be put. The tick size changes from market to market and there is not a general rule to define it. In some cases it is a fixed fraction of the average trade price, in others – like the US market – it is a fixed quantity regardless of the trade price. The key fact is that in some markets the discretization can be negligible with respect to the trade price – *small ticks* – and in other it isn't – *large ticks* –. For small ticks, since the price axis is nearly continuous the number of orders at each price x is vanishing and it is not uncommon to find holes inside the limit order book. For this reason in small ticks markets priority queues at fixed price are typically short and therefore they are of no particular importance. On the other hand, the rough discretization in the large ticks markets has the opposite effect, making crucial the role played by priority queues. This, combined with the fact that for large ticks moving one's position means to change sensibly the price, has as a consequence that in small ticks markets one expects to be much more likely to observe the cancellation and repost of an order at a different price then it is in large ticks markets.

A final remark should be done about the time scales. We shouldn't think of the traders as only physical people posting orders, but also (or mostly) as trained algorithms that study the time series of the data of the order book and take decisions in fractions of a second. In 2015 Bonart *et. al.* [11] measured the fast response of high frequency traders (HFT) to happen in approximately $30\mu s$ and this quantity is definitely a decreasing function of time. They also measured the time discretization of NASDAQ servers to be of order $1\mu s$ making it comparable with the reaction time of HFT. This lets us understand the problem of *lag*. We should be aware that the frequency at which decisions are taken and the price changes are so high that is very much likely that the orders are executed on a slightly different market then the one the decision was made on. This effect, combined with the fact that the trajectory of the price is not necessarily a smooth function of time is another fundamental point we will address in Chapter 3.
In Fig.(1.2) a little sketch to summarize some of the quantities defined and to graphically picture the problem.

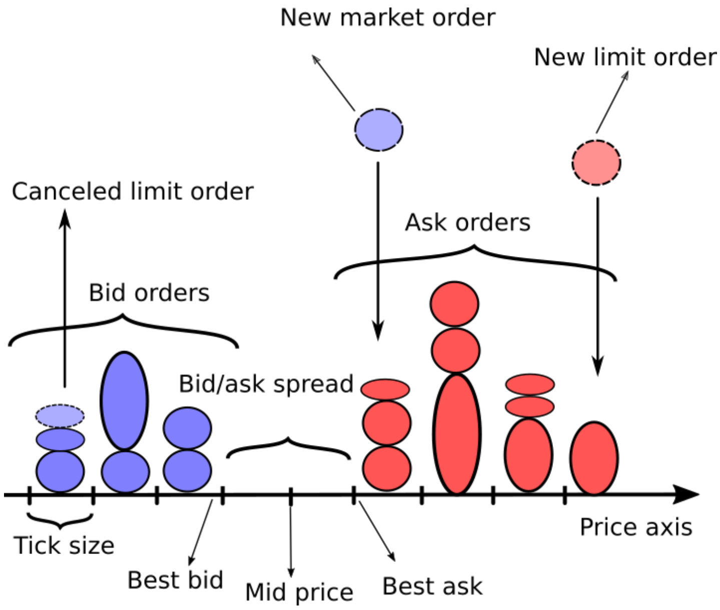


FIGURE 1.2: *A sketch of the limit order book* : in red the ask, in blue the bid. The arrows indicate the main quantities defined and the we observe the deposition of a new ask order inside the ask generates a limit order, while the deposition of a bid order inside the ask will generate a market order that consumes liquidity.

1.2 The behavior of the traders: strategies and observations

In this section we intend to give a stylized description of the behavior of the market participants and present some relevant experimental evidences that should be kept into account when trying to model the limit order book dynamics.

1.2.1 Liquidity taking and liquidity provision

The dynamics of the limit order book is the result of a fine tuning between the behavior of liquidity providers and liquidity takers. Liquidity providers must wait an initially unknown amount of time until their order is executed. A trade will always happen between a market order meeting a limit order, so the advantage of being a liquidity provider relies in being the one of the two traders to have the most advantageous position, allowing him to earn the spread. Having a large spread is therefore convenient for liquidity providers but it is risky in the mean time. We should remember that it is the interest of all the liquidity providers to have the shortest possible time to execution and hence to be at the top of the priority queue. To gain a better position it is then enough to place a new order inside the spread, becoming the best quote, so having a large bid/ask difference is precarious because the best position can be easily beaten by other traders. On the other hand, liquidity takers act in the opposite direction: by placing market orders they open the bid/ask spread, but is their convenience to keep it as small as possible. This very loose argument gives us an intuition of how liquidity takers and providers are in a fine tuned equilibrium that is the compromise between their competing strategies.

A very peculiar aspect of market orders is their potential information content. If a trader had some relevant knowledge about how the mid price was going to evolve in the immediate future, then he would like to take advantage from it by anticipating all the others. The presence of a "pressure" in market orders (*i.e.* an imbalance between buy and sell orders) allows the liquidity providers to understand their position is no longer convenient and makes them want to leave the best quote [25, 44].

We should now focus on a very particular kind of liquidity providers: the market makers (MM). In nowadays electronic trading system, anyone can be a liquidity provider, with its own strategy. Still, for long, liquidity provision was an exclusive task of market makers that where institutional organizations. What they do is to sell at the best ask and in the meantime buy at best bid, typically a large volume of shares. Their strategy allows them to earn the spread at each transaction. The drawback of such strategy is that it is perfectly working if the price doesn't move but it is counterproductive if it does because they find themselves in the condition of having bought at an unrealistic high price falling at the bottom of the priority queue and selling at an inconvenient price if the trade price raised and vice versa if it went down. So the goal of the market makers is too keep the spread as large as possible and to reduce price fluctuations, in order to be able to still make profit. It is observed that the activity of market makers is directly correlated with the market order pressure that naturally pushes the price in one direction [11, 13, 27]. The other fundamental aspect of MM is that they are HFT. It is particularly important to observe that they react in a much faster way with respect to the low frequency traders (LFT) creating a two time scale liquidity that will be further commented in the following.

We now bring light to another fundamental observation that characterizes the financial market: liquidity is vanishing. It has been measured [52] that the instantaneous volume displayed on the limit order book is approximately 0.1% of the total daily traded volume. The reason of this is because people want to give away as little information as possible about their intentions, at least until they have a fair confidence that their orders will be executed in a reasonable amount of time. Displaying one's orders at once would mean to give away a lot of information to the other traders that would evidently make the price move in the direction favorable to them.

It is quite common that some big traders may want to buy a very large amount of shares of a single company. A question then rises: how could the order be executed knowing that there won't be enough liquidity and it can't neither be posted in the limit order book at once if one wants to avoid a large information giveaway? To overcome this problem market attendants are then accustomed to split their orders into smaller pieces and to execute them over a long lapse of time that can go from hours to weeks. We will refer to this kind of orders in the following as *metaorders* and they represent a central aspect of the present study. Further discussions will follow.

1.2.2 Information: two diverging perspectives

A very well established concept in the classical financial studies is that of *fair price* or *fundamental value* [12, 14, 15]. The claim is that there exist a "true" value for each good and the trade price should be mainly determined by it. In particular, it is expected that there is a subset of traders that are informed and who know such value, being able to make profits when the trade price moves away from it and in the meantime their action has the effect to keep these two quantities together. Another subset of traders are the uninformed ones that introduce noise in the process. To the concept of fair price we add the customary *efficient market hypothesis* (EMH). In an informal definition, it consists of saying that the mid price mirrors all the available information. Another closely related notion is arbitrage efficiency, which states that no trader can make profit without taking any risk. Summing these ingredients, the pictures of classical economy sees all the traders as fully rational, some of which are perfectly informed and some aren't and that all the perceptions, news, strategies are perfectly contained in a single number: the mid price. As a consequence, all fluctuations of such quantity correspond to the arrival of news or to some relevant change in the price perception.
This interpretation of the market has been questioned by many. The assumption of full rationality of the traders in such a complex contest is of course very strong, but a part from it, there is an intimate contradiction in the EMH. Without the presence of informed traders the price can't be pushed to its true value, that contains all the information in some sense. However, since no trader could make a systematic profit without taking any risk according to the EMH, informed traders should gain as much as all the others, making this an evident contradiction with their own existence.
Perhaps the role played by information on the actual trade price is the most critical aspect of this interpretation. It has been observed that news play a minor role in the determination of the price [46] and in fact most of the volatility is aroused by the trading activity. A very interesting study was carried on by Jouilin *et.al* [29] about the correlation between the news arrivals and the so-called *4σ jumps*, so those variations of price that exceed 4 times the variance of the mid price. It was observed that such correlation is in fact very small and most

of the price changes aren't due to the arrival of new pieces of information. Also most news don't induce a jump at all, meaning that they were somehow expected. In this picture we should also keep in mind that price fluctuations happen at a much higher frequency then that of news arrivals. The conclusion of such paper is therefore that the price is only minimally affected by the news feed and that the origin of high price jumps has to be found in some liquidity dry outs that are possible – and somewhat likely – because the total liquidity of the order book is indeed vanishing.
In this perspective, we hence consider the price fluctuations as a completely *endogenous process*. To this end a very particular phenomenon is often brought as an example in the econophysics literature. During May 6^{th} 2010 between 14:42 and 15:07 the Dow Jones index registered a sudden collapse. The events where afterwards analyzed and the origin of such crash identified: that day the volatility was particularly high and a big trader placed a large order with an automated algorithm; the unbalance in the pressure made the other traders feel particularly uncomfortable and their reaction led to the sudden collapse of the price [30, 35]. Without getting further specific with the details the caused the crash, the message we here want to convey is that that sudden and large price jump wasn't caused at all by any news feed, but simply by the "mechanical" activity of traders. That was a very particular example and it is still considered as a reference, but many other smaller flash crashes happened ever since, giving full support to the idea that the price is mainly determined by the market activity and not by the news feed. In Fig.(1.3) a picture of the flash crash taken from [30].

1.2.3 The subtle nature of diffusion

While the HFM is indeed a hypothesis, the arbitrage efficiency is an important evidence, both from the experimental and intuitive point of view. The price must be at any moment unpredictable, because if it wasn't so it would mean that it would be possible to make a systematic gain over the market. This is absurd and it is in fact observed that if we do not consider too short time scales – $t \lesssim 1\mu s$ – and too long times scales – $t \gtrsim 1y$ – then the price follows effectively a random walk, *i.e.* [14, 18]:

$$\mathcal{D}(t) := \langle (p_t - p_0)^2 \rangle = \sigma^2 t \tag{1.1}$$

where p_t indicates the mid price at time t. Even if this relation looks very simple, we must underline that there is a complex process ongoing that is worth commenting.
We just claimed that the price changes are mainly due to the market activity. However by dividing their orders into smaller pieces, the liquidity takers create a correlated market order flow. This term alone would lead to super diffusion, so $\mathcal{D}(t) = \mathcal{D}t^{2H}$, with $H > \frac{1}{2}$ the

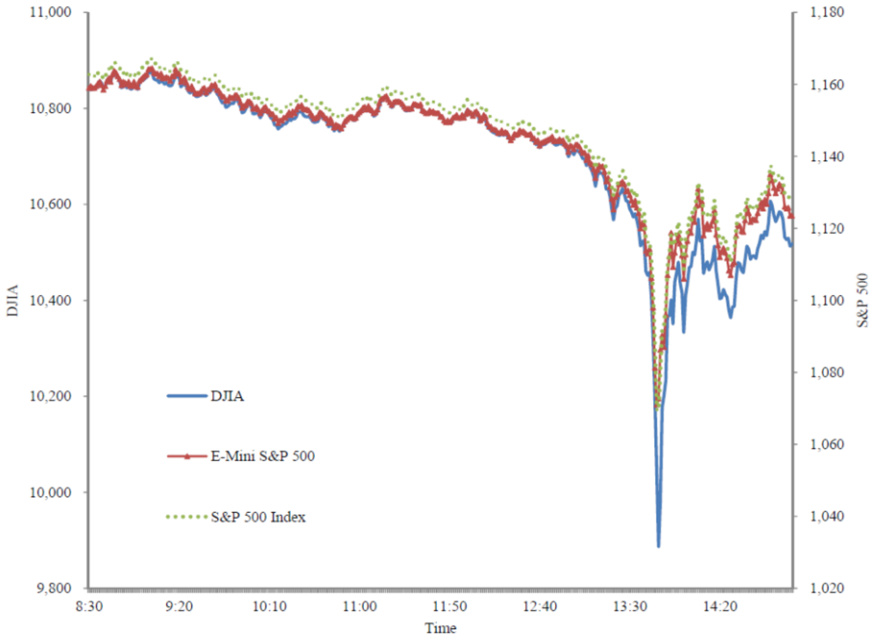


FIGURE 1.3: *The flash crash of May 6, 2010*: the end-of-minute transaction prices of the Dow Jones Industrial Average (DJIA), S&P 500 Index, and the June 2010 E-Mini S&P 500 futures contract on May 6, 2010 between 8:30 and 15:15 CT. *Source*: Kirilenko [30]

Hurst coefficient [31]. So the question that is raised is the so called *diffusivity puzzle*: how is it possible that correlated market orders generate uncorrelated prices? We can write the price at a given moment in time in a simplified form as [12, 16, 18]:

$$p_t = p_0 + \int_0^t ds\, \mathcal{G}(t-s) m_s \tag{1.2}$$

where time translational invariance was assumed, m_s indicates the market order – with sign – placed at time s and $\mathcal{G}$ represents a general propagator. We also moved to a continuous time limit. We now want to plug this expression into Eq.(1.1)

$$\mathcal{D}(t) = \langle (p_t - p_0)^2 \rangle = \int_0^t ds \int_0^t d\ell\, \mathcal{G}(t-s)\mathcal{G}(t-\ell)\langle m_s m_\ell \rangle \tag{1.3}$$

What we observe, as we said, is that market order are correlated, giving us an expression for $\langle m_s m_\ell \rangle$, Eq(1.4a), [37, 47, 51]. We also make the assumption that the kernel decays as a power law:

$$\mathcal{G}(t) \sim |t|^{-\beta} \tag{1.4a}$$

$$\langle m_s m_\ell \rangle := \mathcal{C}(s-\ell) \sim |s-\ell|^{-\gamma}, \quad \gamma \approx \frac{1}{2} \tag{1.4b}$$

By plugging these expressions in Eq.(1.3) we obtain the following scaling:

$$\mathcal{D}(t) \sim \int_0^t ds \int_0^t d\ell \, [(t-s)(t-\ell)]^{-\beta} |\ell - s|^{-\gamma} = \int_0^t ds' \int_0^t d\ell' \, (\ell' s')^{-\beta} |\ell' - s'|^{-\gamma} \sim$$
$$\sim \int_0^t ds' \, s'^{-2\beta-\gamma+1} \sim t^{-2\beta-\gamma+2} \tag{1.5}$$

So, to recover the result of Eq.(1.1) we must impose the following relation between the coefficients:

$$\beta = \frac{1-\gamma}{2} \approx \frac{1}{4} \tag{1.6}$$

We will talk about this relation later again. The message we here want to convey is that, in order to get the absence of arbitrage – that is a necessary condition – we need to finely tune the kernel to go against the persistent order flow created by the liquidity takers. The actuators of such mean reversion are the liquidity suppliers that impose a sub diffusive motion of the price.

In the end, we shouldn't look at the diffusion of the prices as simple random walk, but rather the result of these two competing terms that go one against the other. Under this perspective we can say the market is at a critical point [14, 16, 51].

1.3 Modeling the financial market

The classical economic perspective starts from a quite strong set of assumptions: there exists a fair price and it conveys all the information, traders are fully rational and the price fluctuations are determined by the news feed. We argued already about how questionable are these hypothesis, but let's mention the Kyle model, one of the central models in the financial literature [32]. Here three kinds of traders are distinguished: *noise traders* who make random trades[1], *market makers* who set the prices in order to keep market efficiency and the *insider*

[1] Note: rational doesn't mean informed

that knows the liquidation price at a given moment in the future and is provided with an infinite bank that he uses to make profit over the noise traders. The model then assumes the following evolution of the prices:

$$p_{t+1} = p_t + \lambda[\beta(p_\infty - p_t) + \xi_t] + \eta_t \tag{1.7}$$

p_∞ represents the liquidation price and $\beta(p_\infty - p_t)$ is the contribution given by the informed insider demand, while ξ_t is the effect of uninformed traders and η_t that of news feed.
Even if it is a very simplistic model with big assumptions, it has been a reference point in the modeling of the financial markets ever since its publication. We here bring it to attention for a fundamental aspect: price variations are linear. In particular, it is a simple concept that our action on the market has an effect that goes against it, so if we buy the price will tend to increase and vice versa to decrease if we sell. But how does it do that? We here rigorously introduce the concept of *price impact*:

$$\mathcal{I}(\mathcal{Q}_t) = \langle \epsilon(p_t - p_0) | \mathcal{Q}_t = \epsilon \int_0^t ds\, m_s \rangle \tag{1.8}$$

where $\mathcal{Q}_t$ represents the total traded volume until time t, ϵ its sign ($\epsilon = 1$ for buy orders and $\epsilon = -1$ for sell orders such that the impact is always positive) and m_t[2] the size of the order that was put at time t. What Kyle model predicts is that the impact is linear in the total traded volume, *i.e.* $\mathcal{I}(\mathcal{Q}_t) \propto \mathcal{Q}_t$. However this is not what we observe. It is in fact a well established evidence that the impact follows the so called *square-root law* and is often written in the following way [1, 3, 9, 19, 20, 23, 48]:

$$\mathcal{I}(\mathcal{Q}_t) = Y\sigma_{\mathrm{D}}\sqrt{\frac{\mathcal{Q}_t}{V_{\mathrm{D}}}} \tag{1.9}$$

where Y is a numerical prefactor of order one, σ_{D} is the daily volatility and V_{D} the total traded volume in one day. In Fig.(1.4) a picture taken from [50] of some experimental verifications of the square root law, compared to the linear impact predicted by Kyle model. Note that in the figure $\Delta \equiv \mathcal{I}$.

An important question we should address if we want to describe the market micro structure is what should be the building blocks of our model. In Lachapelle *et.al.* [34] the order book

[2] Note that we are assuming that the sign of the order doesn't depend on time: what we are looking at is the motion of the price due to a metaorder (since we are also assuming that the order $\mathcal{Q}$ is placed over an interval of time) placed by a single trader that will therefore act always in the same direction. We can then say that the impact is the average price variation due to a metaorder.

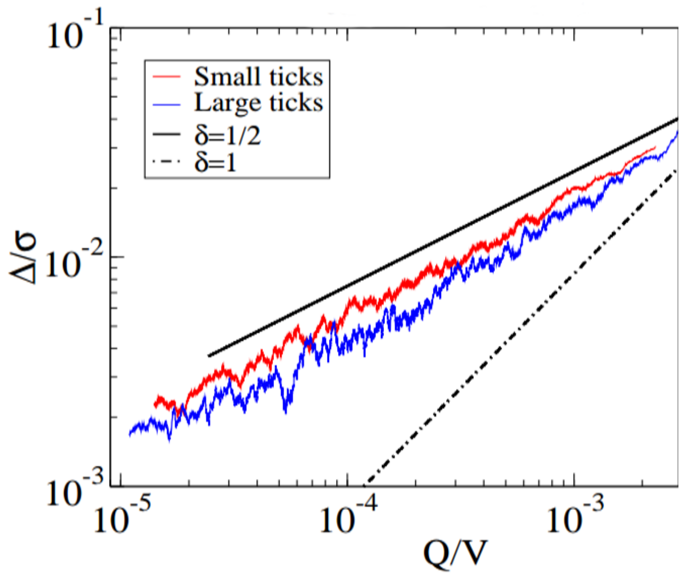


FIGURE 1.4: *Square root law for impact*: the impact of metaorders for CFM proprietary trades on futures markets, in the period June 2007 December 2010. Impact is measured as the average execution shortfall of a metaorder of size Q.
Source: Toth [50]

is modeled in terms of mean field game theory where every trader is rational and intends to maximize his utility. It is quite straightforward to think that every market participant has his own strategy, however the nature of the market is so intimately complex that each choice and information are very hard to interpret, making a rational behavior ultimately hard to assume. Any observation, is affected by a given error, of course, and the number of factors that are to be kept into account in order to determine a proper strategy is so high that one can't neglect the uncertainty associated to it. Then this leads us to think that we are allowed, without much loss of generality, to move from a model in which agents are fully rational to one in which they act completely at random: a so called *zero intelligence model*. This choice will relief us from the need of making a lot of assumptions on the behavioral aspects of traders, considering instead only a very small set of reasonable hypothesis. On the other hand though, we know traders do have an intelligence and a strategy and we are aware that some particular behaviors can't be encoded in a zero intelligence model. Let's make an example: referring to the flash crash of 2010, what was observed was that, even if the price was falling down, not only the sellers but also the buyers left the market. In complete absence of intelligence we would expect the buyers to profit of the price crashing down. However, the act of leaving the market during a crash is mainly due to fear. So, by talking of a zero intelligence model we expect to be able to capture some relevant statistical properties, but not other dynamical peculiarities [47, 49]. Also, it has been observed by many [7, 15, 33] how some regularities are effectively present and are invariant from one market to the other, making plausible the idea of a very simple, coarse grained model, bare of unjustified assumptions, able to depict such regularities.

Finally we here want to present one of the protagonists of the present dissertation. The LOB is a very important tool to describe the behavior of the market since it contains all the most important parameter: mid price, volume imbalance, liquidity, etc. In Focault [24] it is offered a model of the limit order book that starts from the rich enough phenomenology that determines the spread size, given by the competition between liquidity takers and liquidity providers. It is almost natural to think to start from a direct description of the LOB in order to model it. However, as we mentioned earlier, the LOB is just the minimal part of people's intentions. Most of the trading activity is not declared until the very end and therefore it doesn't appear inside the LOB. In Toth *et.al.* [50] it was proposed the concept of *latent order book*. Essentially the claim is that in order to properly model the limit order book one should keep into account all the desires: the latent and the declared ones. Here it was proposed a model in which traders could deposit new orders, remove the old ones and change their mind with a diffusive-like term. They also noted that the shape of the order book in the vicinity of the trade price had to be linear in order to keep the square root impact. In Mastromatteo *et.al.* [39] the model was further inspected and it was observed how the HFT were effectively unimportant in the determination of the statistical long term properties. Finally Donier *et.al* [22] built on this framework a consistent model able to theoretically describe the square root impact.
This last model represents the starting point of our research and in the next chapter we give it a detailed description.

Chapter 2

The latent order book model

In this chapter we provide a description of the model of Donier *et.al.* [22] in which it is proposed a description of the latent order book, *i.e.* of an ideal, not measurable book where all the intentions of the traders are kept. It is a zero intelligence model in which the quantities under analysis are the *averages* of the volumes of the densities at each price. We can look at this as a *mean field* formulation in which each agent acts at random and independently of all the others. Also, it is very important to stress the fact that we are talking of the average dynamics in the sense that the instantaneous shape of the LOB is of course a random multidimensional variable and to model it we should keep into account some stochastic term, that in fact does not appear in the equations we will meet. By considering the time average of the book we are necessarily getting interested in the long memory behavior of the low frequency traders, while the HFT are essentially not kept into account. We will come back on this important point in Chapters 3, 4 where the same question will be raised.

2.1 Description of the model

The starting point is a simple reaction diffusion system of equations. We here make the important – as unrealistic – assumption that the reaction happens inside such latent order book. We call it unrealistic for the order book somehow represents the thoughts of the traders: even if two of them agree at the same price they must go through the real limit order book in order to have the deal done. One could think that in this model all traders at the trade price reveal their intentions and therefore react, while all the others don't. This is however an interpretation that is not contained in the equations per se and in the spirit of the model that should in the end be considered as a latent order book plus a reaction term. By denoting with $\mathcal{A}$ the ask and with $\mathcal{B}$ the bid, the reaction amounts to say

$$\mathcal{A} + \mathcal{B} \to \emptyset \tag{2.1}$$

We denote with $\rho_{\mathrm{A}}(x,t)$, $\rho_{\mathrm{B}}(x,t)$ the average densities of orders at price x and time t. A very important remark we should do is that this model is continuous in space and time. This approximation shouldn't affect the relevant statistical properties, but of course won't be adapt to describe the spread dynamics. The terms that enter into the equations are the following:

a) *Drift diffusion* : this term models the fact that people can change their mind along time and they do so by moving on the price axis. The drift term represents a collective motion that is caused by a common source and can be, for example, a new piece of information. In general this is an exogenous time dependent term. The diffusion instead represents the independent random motion of each trader that is thought of as a zero intelligence agent[1].

b) *Cancellation* : this term allows the trader to remove orders from the book, partially or completely. It is modeled by the parameter ν.

c) *Deposition*: this term represents new orders incoming. It is modeled by the function $\lambda_{\mathrm{A,B}}(x)$, for ask and bid respectively. Deposition is assumed not to depend on the density of the book, unlike the cancellation and to satisfy the following properties

$$\lambda_{\mathrm{A}}(x) = \lambda_{\mathrm{B}}(-x) \tag{2.2a}$$

$$\lambda_{\mathrm{A}}(x) \geq \lambda_{\mathrm{A}}(x'), \forall\, x > x' \tag{2.2b}$$

In the model it is further assumed that $\lambda_{\mathrm{A}}(x) = \lambda\Theta(x)$ where Θ is the Heaviside function and λ a parameter.

d) *Reaction* : when two orders meet they react according to Eq.(2.1).

Note however that another implicit and very important assumption is made: this is a 1D model in which no interaction between different assets is considered. More explicitly, we are here claiming that each asset is decoupled from the others (*e.g.* Microsoft from Apple) that is however an assumption the authors made in order to keep the simplicity the model. We can write the reaction term as

$$R_{A,B} = \kappa\rho_{\mathrm{A}}(x,t)\rho_{\mathrm{B}}(x,t) \tag{2.3}$$

[1] It looks quite straightforward to model a random reassessment of intentions as a diffusion process. It should however be noticed that two limit orders can never effectively give place to a reaction that only happens via the market order book. The diffusive term hides this process allowing the best quote to further drift towards the opposite side and consuming liquidity and then becoming – by definition – a market order.

and, by indicating with p_t the trade price, the equations become:

$$\partial_t \rho_{\mathrm{A}}(x,t) = -V_t \partial_x \rho_{\mathrm{A}}(x,t) + D\partial_{xx}\rho_{\mathrm{A}}(x,t) - \nu\rho_{\mathrm{A}}(x,t) + \lambda\Theta(x-p_t) - R_{A,B} \tag{2.4a}$$
$$\partial_t \rho_{\mathrm{B}}(x,t) = -V_t \partial_x \rho_{\mathrm{B}}(x,t) + D\partial_{xx}\rho_{\mathrm{B}}(x,t) - \nu\rho_{\mathrm{B}}(x,t) + \lambda\Theta(p_t-x) - R_{A,B} \tag{2.4b}$$

By taking the limit for $\kappa \to \infty$, then the books don't overlap and the trade price is well defined. The deposition therefore is assumed to happen always on the "right side" with uniform probability *i.e.* no ask orders will be put below the fair price and vice versa for the bid orders. The function $\phi(x,t) = \rho_{\mathrm{B}}(x,t) - \rho_{\mathrm{A}}(x,t)$ is introduced and, since the there is no overlap in the books, there is no loss of information concerning the shape of both books; in particular

$$\rho_{\mathrm{A}}(x,t) = -\phi(x,t)\Theta(x-p_t) \tag{2.5a}$$
$$\rho_{\mathrm{B}}(x,t) = \phi(x,t)\Theta(p_t-x) \tag{2.5b}$$

We now perform the following change of variable:

$$y = x - \hat{p}_t \tag{2.6a}$$
$$\tau = t \tag{2.6b}$$

where $\hat{p}_t$ is a to-be-defined time dependent parameter. The new differential operators then become

$$\partial_x f = \partial_y f \tag{2.7a}$$
$$\partial_t f = -\partial_\tau \hat{p}_\tau \partial_y f + \partial_\tau f \tag{2.7b}$$

By injecting this into the Eqs.(2.4), we see that we can remove the drift term by choosing

$$\hat{p}_t = \int_0^t ds\, V_s \tag{2.8}$$

This passage has not to be seen as a purely mathematical convenience: from this moment on we will be in the reference frame of $\hat{p}_t$ and we will not consider all the changes of prices due to

external information, but only those whose origin is endogenous. Alternatively, the drift term accounts for the part of the price that is effectively determined by the news feed (which we argued plays a minor role) and all the considerations we'll make from this moment on about the price variation are only due to the role that market orders play in the determination of the trade price.
It is now easy to verify that we should impose the following conditions:

$$\phi^{(s)}(-y) = -\phi^{(s)}(y) \tag{2.9a}$$
$$\phi(y_\tau, \tau) = 0 \quad \forall\tau, \quad y_\tau = p_\tau - \hat{p}_\tau \tag{2.9b}$$
$$\phi(y, \tau) \nrightarrow \infty, \quad for \quad |y| \to \infty \tag{2.9c}$$

where $\phi^{(s)}(y)$ represents the stationary time independent solution.

2.2 Stationary solution

In the stationary state p_t is a constant. Once again, the equations are referred to the averages of the quantities: the process is intimately stochastic, and the trade price changes with time also at equilibrium, but not its average. After the considerations done in the previous paragraph, the differential equation becomes:

$$\partial_\tau \phi(y, \tau) = D\partial_{yy}\phi(y, \tau) - \nu\phi(y, \tau) + \lambda sign(p_\tau - \hat{p}_\tau - y) \tag{2.10}$$

To remove all the time dependence from the equation and obtain the stationary state, the relation $p_t = \hat{p}_t \to y_t = 0$ must hold. Also, by making use of a symmetry argument, since $\phi^{(s)}$ is symmetric in y, then also the *sign* function has to be symmetric in y. We therefore conclude that $\hat{p}_t \to p_t \quad$ as $\quad t \to \infty$. Thanks to the condition 2.9a we can solve this problem in $\mathbb{R}^{*+}$ without loss of generality. The new equation then becomes:

$$D\partial_{yy}\phi^{(s)}(y) - \nu\phi^{(s)}(y) = \lambda \tag{2.11}$$

Its solution, combined with the boundary conditions 2.9b, 2.9c reads

$$\phi^{(s)}(y) = -\frac{\lambda}{\nu}(1 - e^{-\gamma y}) \tag{2.12}$$

Where $\gamma := \sqrt{\frac{\nu}{D}}$. An interesting limit is for $\nu, \lambda \to 0$ under the assumption that $\frac{\lambda\gamma}{\nu}$ is still well defined *i.e.* it doesn't diverge or go to zero.

$$\phi^{(s)}(y) \approx -\frac{\lambda\gamma}{\nu} y := -\mathcal{L} y \tag{2.13}$$

We can define a diffusion current as $J := D\partial_y \rho_{\mathrm{A}}(y) = D\mathcal{L}$. This limit is particularly important for it actually doesn't depend deeply on the assumptions of the model: we can find a locally linear order book (LLOB) under different initial setups, as proved in the same paper. In the next we will always be in the linear regime with the idea that $\mathcal{L}$ is a fixed quantity for our model and it represents the liquidity of the market. We see that this nomenclature makes particularly sense according to the definition of market breadth that asks a large number of participants and hence a small impact. This should however be regarded as a first order approximation, because the liquidity is definitely a space dependent parameter and refers much more to volume of orders then a to slope: in the LLOB limit there is a bijection between these two concepts, that is however not general for different shapes of the order book. It should be noticed that the liquidity strongly depends on the time of the day and if one is to confront such parameter to the real data, it should be be confronted with a time average. In Fig.(2.1) it is reported the shape of the solution compared to our numerical simulation.

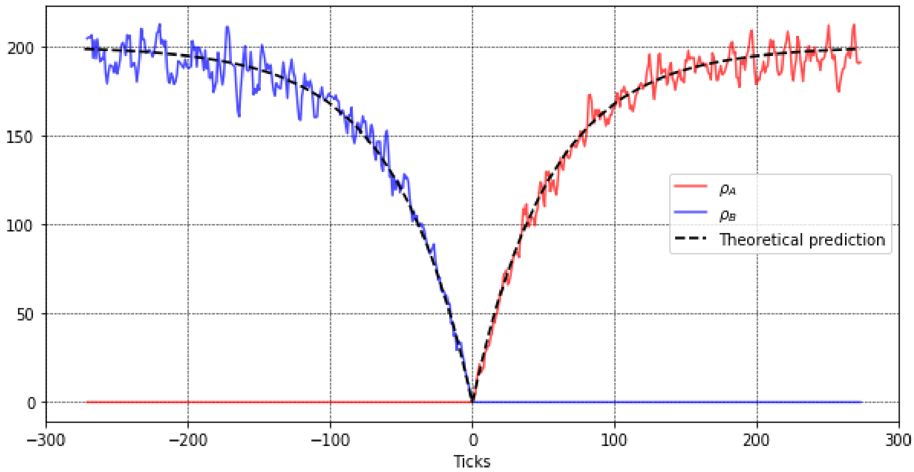


FIGURE 2.1: *Donier model - Stationary solution compared to the result the computer simulation*: the dotted line represents the theoretical curve as expressed in Eq.(2.12); the blue and red lines are the densities of the ask and bid respectively as resulted from our simulation. The parameters used are $\nu\tau = 1/3 \cdot 10^{-5}$, $\lambda = 200\nu$, $\gamma = (18.25 \cdot 10^{-3})ticks^{-1}$, where τ indicates the time step.

2.3 Square root impact

The great result of the model under analysis is that it is able to predict the square root impact, giving it a microscopical interpretation. In the linear regime the equation we want to solve is:

$$\partial_t \phi(y,t) = D\partial_{yy}\phi(y,t) + m_t \delta(y - y_t) \tag{2.14}$$

where we restored the notation $t = \tau$. The metaorder is modeled by and extra term that is placed exactly at the trade price[2] and, due to the fact that the model is continuous in time, it will follow it as it will change its value consequently to the action of the order itself. The goal is to find the explicit expression of y_t. Assume that the metaorder is placed on the stationary book, so that $\phi(y,0) = \phi^{(s)}(y)$. By moving to Fourier space we can rewrite

$$\partial_t \tilde{\phi}(k,t) = -Dk^2 \tilde{\phi}(k,t) + m_t e^{iky_t} \tag{2.15}$$

By writing $\tilde{\phi}(k,t) := \tilde{F}(k,t)e^{-Dk^2 t}$, we find

[2]Trade price, mid price, best quotes here are all the same concept. Due to the continuity in space and the presence of the diffusive term, the spread will always be equal to zero. However, if we are consuming the liquidity – that means that we would open the spread in the real and discrete order book – the density in a certain range of prices will be vanishing, making the continuity approximation still plausible.

$$\tilde{F}(k,t) = \tilde{F}(k,0) + \int_0^t ds\, m_s\, e^{Dk^2 s + iky_s} \tag{2.16}$$

and so, by antitransforming $\phi(y,t)$:

$$\phi(y,t) = \phi^{(s)}(y) + \int_0^t ds\, \frac{m_s}{\sqrt{4\pi D(t-s)}}\, e^{-\frac{(y-y_s)^2}{4D(t-s)}} \tag{2.17}$$

Finally, exploiting Eq.(2.9b) and the expression of Eq.(2.13), we obtain the following self-consistent relation:

$$y_t = \frac{1}{\mathcal{L}} \int_0^t ds\, \frac{m_s}{\sqrt{4\pi D(t-s)}}\, e^{-\frac{(y_t-y_s)^2}{4D(t-s)}} \tag{2.18}$$

We now assume $m_t = m_0\,, \forall t$ and focus our study to two limiting regimes. Let's consider the argument of the exponential:

$$\begin{aligned} \frac{(y_t-y_s)^2}{4D(t-s)} &= \frac{1}{4D(t-s)} \left(\frac{1}{\mathcal{L}} \int_0^t dt'\, \frac{m_0}{\sqrt{4\pi D(t-t')}}\, e^{-\frac{(y_t-y_{t'})^2}{4D(t-t')}} - \int_0^s ds'\, \frac{m_0}{\sqrt{4\pi D(s-s')}}\, e^{-\frac{(y_s-y_{s'})^2}{4D(s-s')}} \right)^2 \\ &\propto \left(\frac{m_0}{D\mathcal{L}}\right)^2 = \left(\frac{m_0}{J}\right)^2 \end{aligned} \tag{2.19}$$

So what governs the behavior of the exponential is the ratio between m_0 and J, regardless of the sign of m_0 and it takes the name of *participation rate* that, in other words, expresses the size of the order with respect to that of the market.

Small participation rate: $|m_0| \ll J$

Considering the small participation rate $\alpha = \frac{|m_0|}{J} \ll 1$ we have that $e^{-\frac{(y_t-y_s)^2}{4D(t-s)}} \to 1$ and we recover the expression of Eq.(1.2) regarding the price formation, where $y_0 = 0$ because we start from the stationary condition and $\mathcal{G}(t) \sim t^{-\frac{1}{2}}$. From this we hence conclude that this model doesn't solve the diffusivity puzzle because memory vanishes too quickly. Further comments about this will be presented in Chapter 4. By denoting with $\mathcal{Q}_t = |m_0|t$, we then obtain the explicit expression of y_t:

$$y_t = \frac{1}{\mathcal{L}} \int_0^t ds \, \frac{m_0}{\sqrt{4\pi D(t-s)}} = \frac{m_0}{\mathcal{L}} \sqrt{\frac{t}{\pi D}} = \epsilon \sqrt{\frac{|m_0| Q_t}{\pi J \mathcal{L}}} \tag{2.20}$$

Large participation rate: $|m_0| \gg J$

We rewrite Eq.(2.18) with a change of variable: $u = t - s$ and perform a saddle point approximation. The exponential will kill everything in this limit and the dominant term will be given by $\min_t(y_t - y_{t-u})$. We hence make an expansion for $u \to 0$.

$$\begin{aligned} y_t &= \frac{1}{\mathcal{L}} \int_0^t du \, \frac{m_0}{\sqrt{4\pi D u}} \, e^{-\frac{(y_t - y_{t-u})^2}{4Du}} \approx \frac{1}{\mathcal{L}} \int_0^\infty du \, \frac{m_0}{\sqrt{4\pi D u}} \, e^{-\frac{\dot{y}_t^2 u}{4D}} = \\ &= \frac{m_0}{\mathcal{L}|\dot{y}_t|\sqrt{\pi}} \int_0^\infty dv \, v^{-\frac{1}{2}} e^{-v} = \frac{m_0}{\mathcal{L}|\dot{y}_t|} \end{aligned} \tag{2.21a}$$

We also extended the integral up to infinity – which is a good approximation since we the dominant term is around zero – to recover the definition of the Γ function.
To solve this equation we finally notice that ϵy_t must be an increasing function of the time and write

$$\frac{m_0}{\mathcal{L}} = \epsilon \dot{y}_t y_t = \epsilon \frac{1}{2} \partial_t y_t^2 \tag{2.22}$$

And then by a simple integration we find the final result that is summarized in Eq.(2.24b). An interesting comment is that this result can be recovered also with a geometrical interpretation that can be written in the following terms:

$$m_0 t = \epsilon \int_0^{y_t} dx \, |\phi^{(s)}(x)| = \epsilon \int_0^{y_t} dx \, \mathcal{L} x = \epsilon \frac{\mathcal{L} y_t^2}{2} \tag{2.23}$$

Let us here summarize the results obtained for the impact: we note that in the large participation rate regime the impact doesn't depend on the participation rate itself, provided it is large enough, while it does in the small participation rate regime. Eating the book at a slow rate gives an additional contribution to the volume that comes from the diffusion current[3] that is not present when $|m_0| \gg J$.

[3] It is particularly relevant to stress that it comes from the current and not the diffusion per se. In fact $J \propto \partial_y \phi^{(s)}$: the presence of a gradient in the book is necessary to observe such effect.

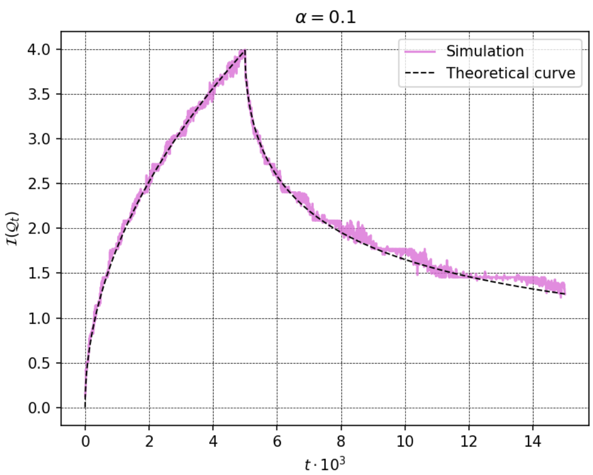
$$\mathcal{I}(\mathcal{Q}_t) = \sqrt{\frac{\alpha \mathcal{Q}_t}{\pi \mathcal{L}}}, \quad for \quad \alpha \ll 1 \tag{2.24a}$$

$$\mathcal{I}(\mathcal{Q}_t) = \sqrt{\frac{2\mathcal{Q}_t}{\mathcal{L}}}, \quad for \quad \alpha \gg 1 \tag{2.24b}$$

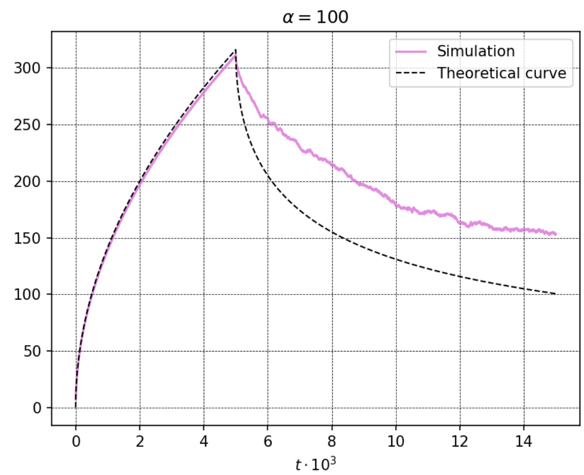
In Fig.(2.2) we report the comparison of the analytical result and the simulation. For the decay part of the impact – so from the end of the metaorder at time T on – we used as analytical line the linear propagator model that predicts:

$$\frac{\mathcal{I}(t > T)}{\mathcal{I}(T)} = \frac{\sqrt{t} - \sqrt{t - T}}{\sqrt{T}} \tag{2.25}$$

This expression works for $\alpha \ll 1$ for which in fact we recovered the expression of the linear propagator of Eq.(1.2) while we observe a slower decay for $\alpha \gg 1$.
Writing the price as in Eq.(1.2) is of course the results of many assumptions and comes from empirical and heuristic arguments, but it is good that we obtain such expression at least in the low participation limit: the regime $\alpha \gg 1$ is not a particularly interesting one since traders try to reveal as slow as they can, as commented above.



(A) Slow execution rate



(B) Fast execution rate

FIGURE 2.2: *The impact in the Donier model - comparison between theory and simulation*: in purple the result of our numerical simulation and the theoretical line for small and large participation rate. For the rising part it is the expression as reported in Eq.(2.24), while for the descent part it is the expression of Eq.(2.25) and we evidently observe that it works only in the small participation rate.

2.4 The numerical simulation

We here give a brief description of the implementation of the numerical simulation of the system considered.

The densities of the ask and bid are stored in two vectors of size 2000 where we indicate the numbers of orders at a given price. Our simulation is discrete both in space and time, but we chose some scales small enough to let such discretization be negligible. The size we chose represents a good trade off between computational performance and the negligibility of the tick size.
By defining $D = \frac{p}{2\tau}$, being τ the time step and $0 \leq p \leq 1$, we select through a random binomial of parameter p at each price the number of particles that have to move. Among these we further select with a random binomial of parameter $\frac{1}{2}$ the ones that have to move rightwards, while the remaining will move leftwards. Reflecting boundary conditions are imposed.
Regarding the cancellation term, we select at each tick and time step a number of orders with a random binomial of parameter $\nu\tau$ that are removed, while for the deposition we add at each tick and time step $int(\lambda\tau) + Binomial(int(\lambda\tau) - \lambda\tau, 1)$ orders. When two orders of opposite kind occupy the same tick the reaction makes them vanish.
In the linear regime we do not implement directly the deposition and cancellation, in order to make the algorithm faster. The only survived term is then the diffusion and, in order to ensure the slope of the book, we impose the current at the boundaries. More specifically, at each time step we add at the edges of our vector $int(J\tau) + Binomial(int(J\tau) - J\tau)$ orders.

What is particularly interesting to notice is that we are dealing with a discrete order book in space and time that we described through a set o continuous partial differential equations. Our simulation is however again discrete in space and time being a further approximation of the equations per se. However such approximation goes back towards the direction of the "physical" order book. We could have performed a numerical solution of the dynamical equations, but we made this particular choice to have a simulation the closest possible to the actual limit order book. In this perspective, seeing that the theoretical result of the continuous model matches with the one of a discrete numerical simulation, allows us to confirm the claim that the approximation of continuity is safe, provided that we give up the spread dynamics.

2.5 Comments and remarks

This model gives a micro-structural explanation of the square root impact starting from a small set of particularly reasonable hypothesis. In this sense we chose it as our initial point for it represents a simple enough way to describe the behavior of the latent order book without any further external behavioral assumption or input. The possibility of predicting the square root impact from a zero intelligence model in particular gives a great support to the initial claim.
Furthermore, we noticed that such law is in fact a direct consequence of having a linear shape of the limit order book around the origin as claimed in [50], a condition that is valid under a broad set of assumptions and hence looks particularly solid.

If on one hand these are the great achievements of this model, we here briefly discuss what are some aspects left behind and how some of them have already been solved.
In next chapter we will build a model based on the LLOB limit. Also, all the main calculations were performed in this limit in which $\lambda, \nu \to 0$. Benzaquen [6] showed that keeping $\lambda, \nu \neq 0$, *i.e.* introducing a memory effect to the model, it is possible to predict a permanent impact as well, which is experimentally measured. For non vanishing cancellation and deposition it exists a typical time scale in which the order book is completely renewed and, in this sense, memory is lost. By placing a large metaorder (that is therefore executed over a long amount of time) the price is pushed in one direction and the book will have in the end lost his memory about the initial shape and so the price won't relax to its original value, generating in fact a permanent impact. In this sense, when considering the linear regime, we know we are giving up this effect that we should however be able to recover by introducing back cancellation and deposition.
Another aspect is that of the correlation between different order books and the consequent study of *cross impact*. It is customary not to buy shares of a single asset, but rather a full portfolio. The question that then raises is how does the purchase of different correlated assets influences the total impact? This problem is completely not considered in the present model and it won't be neither in the one we propose, but other researchers have focused already on how to model a simple extension of inter-asset correlations, giving us the idea that it will be possible to extend this model to a multidimensional version of it [8, 40, 41].

Finally we discussed how this model doesn't solve the diffusivity puzzle. Benzaquen [5, 6] presented two possible extensions able to solve it for this model and we will discuss them more in depth in Chapter 4.

Chapter 3

Latent liquidity revealing in the limit order book

In this chapter it is presented the original work of this project. We start from the idea that, in order to model properly the dynamics of the limit order book, latent liquidity should be taken into account, as already claimed. The model of Donier *et.al.* represents a simple enough starting point yet able to predict the square root impact that should be considered a *conditio sine qua non*. We therefore propose a simple mechanism to describe how the latent liquidity reveals into the real and observable limit order book.

3.1 A mechanism for latent liquidity revealing

In order to describe the communication between the limit order book and the latent one revealing into it, we define the revealed and unrevealed order books together with the revealed and unrevealed limit order densities for the bid and ask sides of the book $\rho^{(\mathrm{r})}_{\mathrm{A/B}}(x,t)$ and $\rho^{(\mathrm{u})}_{\mathrm{A/B}}(x,t)$. In Fig.(3.1) a small sketch of the setup of this model. We denote with D_{u} and D_{r} the diffusion coefficients in the unrevealed and revealed order books respectively. While the diffusion in the unrevealed book signifies the reassessments of agents intentions as discussed in the previous chapter, the idea of diffusion in the revealed order book certainly deserves a discussion. As we already mentioned, once a limit order is placed in the revealed order book, if one wants to change its position before it gets executed, he has to cancel it and place it somewhere else. So one may argue that there should be no diffusion in the revealed book such that revealed order reassessments must go through the unrevealed book before they can be posted again. However, we believe that

- unrevealing one's order because one is no longer confident on one's choice and waiting for an arbitrary amount of time to reveal it back

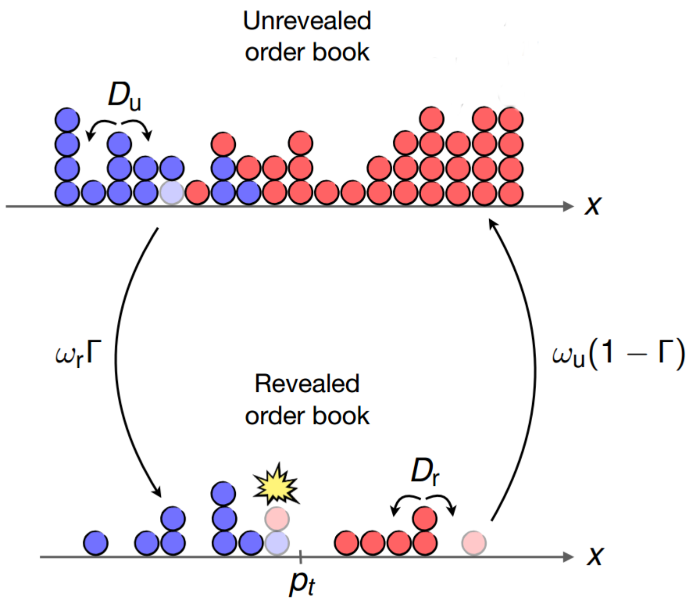


FIGURE 3.1: *Schematics of the unrevealed and revealed order books* : diffusion in the two books, reaction and the currents coupling them are indicated by the arrows.

- cancelling an order knowing that it will immediately be posted back at a revised price

are two distinct processes and to properly model them we leave the possibility for a nonzero diffusivity D_r in the revealed order book. In addition, note that trading fees and priority queues discourage traders from changing posted orders. In this perspective one is tempted to expect $D_r < D_u$ in the general case, but one might as well argue that the presence of HFT can considerably increase the value of D_r by that inverting such an inequality. Be that as it may, the limit $D_r \to 0$ will very likely be an interesting one to address as it represents the large ticks assets. Furthermore we posit that unrevealed orders are revealed with probability $\Gamma(k(x-p_t)) \in [0,1]$ at a rate ω_r and unrevealed with probability $1-\Gamma(k(x-p_t))$ at rate ω_u, where p_t denotes the trade price and k^{-1} is a characteristic price scale. Note that while it may be reasonable that, when unrevealed, orders don't land on the same price x they took off from – by that moving away the earlier decision, say $(x)_r \to (x+\Delta x)_u$, with $\text{sign}(\Delta x) = \text{sign}(x - p_t)$ –, we shall not consider such a possibility in the present study for reasons of analytical tractability restricting our interest to the case $\Delta x = 0$. Naturally, buy/sell order matching $\mathcal{A} + \mathcal{B} \to \varnothing$ only takes place in the revealed order book. For simplicity of notation we will imply the space/time dependence of $\rho^{(u/r)}_{A/B}(x,t) \equiv \rho^{(u/r)}_{A/B}$. Assuming $\Gamma(y)$ to be continuous and sufficiently regular on $\mathbb{R}^*$, one may write for the ask side:

$$\partial_t \rho_A^{(r)} = D_r \partial_{xx} \rho_A^{(r)} + \omega_r \Gamma(k(x - p_t)) \rho_A^{(u)} - \omega_u \left[1 - \Gamma(k(x - p_t))\right] \rho_A^{(r)} - \kappa \rho_A^{(r)} \rho_B^{(r)} \quad (3.1a)$$
$$\partial_t \rho_A^{(u)} = D_u \partial_{xx} \rho_A^{(u)} - \omega_r \Gamma(k(x - p_t)) \rho_A^{(u)} + \omega_u \left[1 - \Gamma(k(x - p_t))\right] \rho_A^{(r)} \quad (3.1b)$$

and for the bid side:

$$\partial_t \rho_B^{(r)} = D_r \partial_{xx} \rho_B^{(r)} + \omega_r \Gamma(k(p_t - x)) \rho_B^{(u)} - \omega_u \left[1 - \Gamma(k(p_t - x))\right] \rho_B^{(r)} - \kappa \rho_A^{(r)} \rho_B^{(r)} \quad (3.2a)$$
$$\partial_t \rho_B^{(u)} = D_u \partial_{xx} \rho_B^{(u)} - \omega_r \Gamma(k(p_t - x)) \rho_B^{(u)} + \omega_u \left[1 - \Gamma(k(p_t - x))\right] \rho_B^{(r)} \quad (3.2b)$$

Note that we are writing these equations in absence of a drift term, hence posing $V_t = 0$. We should however recall that we can include it and then reconduce ourselves to this form by performing the change of variable described in Chapter 2, obtaining the same equations provided the substitution $p_t \to y_t$ and $x \to y$.
In the limit $\kappa \to \infty$, $\rho_A^{(r)}(x,t)$ and $\rho_B^{(r)}(x,t)$ do not overlap such that one may instead consider the difference $\phi_r(x,t) := \rho_B^{(r)}(x,t) - \rho_A^{(r)}(x,t)$ and absorb the reaction terms without loss of information. Note however that the unrevealed order books are perfectly allowed to overlap. The trade price p_t is then defined as:

$$\lim_{\epsilon \to 0} \left[\phi_r(p_t + \epsilon, t) \phi_r(p_t - \epsilon, t)\right] < 0 \,. \quad (3.3)$$

For the sake of simplicity we shall set $\omega_r = \omega_u = \omega$. In Section 4.1.2 we study the effects of relaxing this hypothesis. Subtracting Eq. (3.1a) to Eq. (3.2a) and injecting $\rho_B^{(r)} = \phi_r \Theta(p_t - x)$, $\rho_A^{(r)} = -\phi_r \Theta(x - p_t)$, one obtains the following set of equations, central to our study:

$$\partial_t \rho_B^{(u)} = D_u \partial_{xx} \rho_B^{(u)} - \omega \left\{ \Gamma(k(p_t - x)) \rho_B^{(u)} - \left[1 - \Gamma(k(p_t - x))\right] \Theta(p_t - x) \phi_r \right\} \quad (3.4a)$$
$$\partial_t \rho_A^{(u)} = D_u \partial_{xx} \rho_A^{(u)} - \omega \left\{ \Gamma(k(x - p_t)) \rho_A^{(u)} + \left[1 - \Gamma(k(x - p_t))\right] \Theta(x - p_t) \phi_r \right\} \quad (3.4b)$$
$$\partial_t \phi_r = D_r \partial_{xx} \phi_r + \omega \left\{ \Gamma(k(p_t - x)) \rho_B^{(u)} - \Gamma(k(x - p_t)) \rho_A^{(u)} - \left[1 - \Gamma(k|x - p_t|)\right] \phi_r \right\} \quad (3.4c)$$

Eqs.(3.4) must be complemented with a set of boundary conditions. In particular we impose that $\lim_{x \to \infty} \partial_x \rho_A^{(u)} = - \lim_{x \to -\infty} \partial_x \rho_B^{(u)} = \mathcal{L}$ (in the spirit of the result obtained in the LLOB limit) and that $\rho_A^{(u)}$ does not diverge when $x \to -\infty$, respectively $\rho_B^{(u)}$ when $x \to \infty$. In addition, whenever $D_r \neq 0$, one must add $\phi_r(0) = 0$ and that $\phi_r(x)$ does not diverge when $|x| \to \infty$. The case $D_u = 0$ doesn't appear to be particularly appealing since we believe that removing the diffusion from the unrevealed order book while leaving it in the real order book doesn't correspond to an observable state nor to any particular limit. In the following we will also give a more solid argument in support to this claim.

3.2 Stationary states of the order book

In this section we compute analytically and numerically the stationary order books as function of the different parameters, and discuss interesting limit cases.
Setting $\partial_t \rho_B^{(u)} = \partial_t \rho_A^{(u)} = \partial_t \phi_r = 0$, and letting $\xi = x - p_t$[1] into Eqs. (3.4) one obtains for all $\xi \in \mathbb{R}^*$,

$$\begin{aligned} \rho_B^{(u)}(\xi) &= \rho_A^{(u)}(-\xi) \quad (3.5a) \\ \phi_r(\xi) &= -\phi_r(-\xi) \quad (3.5b) \end{aligned}$$

This allows to solve the problem on $\mathbb{R}^{+*}$. As a direct implication of this conditions, we have that $\rho_B^{(u)}(0^+) = \rho_A^{(u)}(0^+)$, $\partial_\xi \rho_B^{(u)}(0^+) = -\partial_\xi \rho_A^{(u)}(0^+)$. Furthermore, restricting to a revealing probability distribution satisfying $\Gamma(y \leq 0) = 1$, the system one must solve for $\xi > 0$ reduces to:

$$\begin{aligned} 0 &= D_u \partial_{\xi\xi} \rho_B^{(u)} - \omega \rho_B^{(u)} \quad (3.6a) \\ 0 &= D_u \partial_{\xi\xi} \rho_A^{(u)} - \omega \left\{ \Gamma(k\xi) \rho_A^{(u)} + [1 - \Gamma(k\xi)] \phi_r \right\} \quad (3.6b) \\ 0 &= D_r \partial_{\xi\xi} \phi_r - \omega \left\{ \Gamma(k\xi) \rho_A^{(u)} + [1 - \Gamma(k\xi)] \phi_r - \rho_B^{(u)} \right\} \quad (3.6c) \end{aligned}$$

Note that orders falling on the "wrong side" are consuming liquidity – so they represent market orders – but they do it at the revealing price rather then at the best quote. This is a difference with what actually happens in the limit order book. The reason why we chose this syntax is again to keep analytical tractability. We also performed some numerical simulations in which we posed the revealing to be at the best quote as it should be and we didn't observe any major difference in the statistical properties discussed in the following, making this a reasonable hypothesis. The aspect that such approximation completely fails to describe is the spread dynamics, since the spread can in principle become negative. However, we gave up its description the moment we decided to use a model continuous in space, so this shouldn't be regarded as a further approximation.

[1]Note that p_t actually doesn't depend on time at equilibrium and so ξ correctly doesn't show the temporal expression.

3.2.1 Analytical and numerical solutions

Here we provide solution of the system of Eqs. (3.6) for four distinct cases of interest $D_\mathrm{u} = 0$, $D_\mathrm{r} = 0$, $D_\mathrm{r} = D_\mathrm{u}$, and $D_\mathrm{r} \neq D_\mathrm{u}$. As a reminder we here synthesize the boundary conditions that we will refer to in the following

$$\begin{aligned}
\lim_{\xi\to\infty} \partial_\xi \rho_\mathrm{A}^{(\mathrm{u})} &= \mathcal{L} && (3.7a)\\
\rho_\mathrm{A}^{(\mathrm{u})}(0^+) &= \rho_\mathrm{B}^{(\mathrm{u})}(0^+) && (3.7b)\\
\partial_\xi \rho_\mathrm{A}^{(\mathrm{u})}(0^+) &= -\partial_\xi \rho_\mathrm{B}^{(\mathrm{u})}(0^+) && (3.7c)\\
\rho_\mathrm{B}^{(\mathrm{u})} &\not\to \infty, \quad for \quad \xi \to \infty && (3.7d)\\
|\phi_\mathrm{r}| &\not\to \infty, \quad for \quad \xi \to \infty \quad (if \quad D_\mathrm{r} \neq 0) && (3.7e)\\
\phi_\mathrm{r}(0) &= 0, \quad (if \quad D_\mathrm{r} \neq 0) && (3.7f)
\end{aligned}$$

$D_\mathrm{u} = 0$

As we mentioned already this limit doesn't seem particularly interesting for it doesn't correspond to a physically relevant state. We here want to show that also from the mathematical point of view the stationary solution allows us to safely neglect this regime. From Eq.(3.6a) it follows that:

$$\rho_\mathrm{B}^{(\mathrm{u})}(\xi) = 0, \quad \forall \xi > 0 \tag{3.8}$$

Then, from Eqs.(3.6b 3.6c) we obtain:

$$\begin{aligned}
\phi_\mathrm{r}(\xi) &= \mathcal{A}\xi + \mathcal{B}, \quad \mathcal{A}, \mathcal{B} \in \mathbb{R} && (3.9a)\\
\rho_\mathrm{A}^{(\mathrm{u})}(\xi) &= \frac{\phi_\mathrm{r}(\xi)[\Gamma(k\xi) - 1]}{\Gamma(k\xi)} && (3.9b)
\end{aligned}$$

So, by further assuming that $\lim_{y\to\infty} \Gamma(y) = 0$, then we have that $\lim_{\xi\to\infty} \rho_\mathrm{A}^{(\mathrm{u})} = -\lim_{\xi\to\infty} \phi_\mathrm{r}(\xi) \left[\Gamma(k\xi)\right]^{-1}$. To get the proper behavior at infinity of both $\rho_\mathrm{A}^{(\mathrm{u})}$ and ϕ_r we must assume that $\mathcal{A} = 0$ and

$\Gamma(y) \propto y^{-1}$. However, if we also want to impose that $\phi_r(0) = 0$, implying $\mathcal{B} = 0$, we notice that there is no interesting solution that solves this problem with the proper boundary conditions. In this perspective, this limit will not be further inspected.

$D_r = 0$

We introduce the following notation: $\ell_u := \sqrt{\frac{D_u}{\omega}}$. Eq.(3.6a) is simply solved considering the boundary condition Eq.(3.7d)

$$\rho_B^{(u)}(\xi) = ce^{-\xi/\ell_u} \tag{3.10}$$

By summing Eq.(3.6a) to Eq.(3.6c) and subtracting Eq.(3.6b), one finds $\partial_{\xi\xi}\rho_A^{(u)} = \partial_{\xi\xi}\rho_B^{(u)}$ and hence

$$\rho_A^{(u)} = \rho_B^{(u)} + \mathcal{A}\xi + \mathcal{B}, \quad \mathcal{A}, \mathcal{B} \in \mathbb{R} \tag{3.11}$$

The solution of Eq.(3.6c) then becomes straightforward. By imposing the boundary conditions we get

$$\rho_B^{(u)}(\xi) = \frac{\mathcal{L}\ell_u}{2} e^{-\xi/\ell_u} \tag{3.12a}$$

$$\rho_A^{(u)}(\xi) = \mathcal{L}\xi + \frac{\mathcal{L}\ell_u}{2} e^{-\xi/\ell_u} \tag{3.12b}$$

$$\phi_r(\xi) = \frac{\mathcal{L}\ell_u}{2} e^{-\xi/\ell_u} - \frac{\mathcal{L}\xi\Gamma(k\xi)}{1 - \Gamma(k\xi)} . \tag{3.12c}$$

For its graphical representation confront Fig.(3.2a). Note that, by indicating with $\Gamma^{(n)}$ the first non vanishing derivative in zero, then $\lim_{\xi \to 0^+} \phi_r(\xi) = \mathcal{L}\left[\frac{\ell_u}{2} + \frac{n!}{k\Gamma^{(n)}(0^+)\xi^{n-1}}\right] \neq 0$, in general. Since the function ϕ_r is odd, it means it is discontinuous in the origin. This has to be seen as a direct consequence of the absence of diffusion in the real order book that allows the reaction to happen only through the revelation of new limit orders on the opposite side. Also, it should be noticed that if $\Gamma'(0^+) = 0$ the density of the orders diverges in zero that is a non admissible solution. For this reason $\Gamma(y \leq 0) = e^{-y^2}$ wouldn't be a good choice for our model.
The solution for $D_r = 0$ is general and doesn't need an explicit expression for the function $\Gamma(y \geq 0)$. For the rest of the work we should however specify it to be:

$$\Gamma(y) = \begin{cases} 1 & \forall y \leq 0 \\ e^{-y} & \forall y > 0 \end{cases} \tag{3.13}$$

Note that $\Gamma'(0^+) = -1 \neq 0$, consistent with what we just stated. The study of the effect of a scale-invariant power law decaying Γ could also yield interesting results that aren't however considered in the present study.

$D_\mathrm{r} = D_\mathrm{u}$

We should now focus on the analytical solution of the system 3.6 in the particular case $D_\mathrm{r} = D_\mathrm{u}$. Again, summing Eq.(3.6a) to Eq.(3.6c) and subtracting Eq.(3.6b), one obtains the following equation, where we still don't need to assume $D_\mathrm{u} = D_\mathrm{r}$:

$$\partial_{\xi\xi}\left[\left(\frac{\ell_\mathrm{r}}{\ell_\mathrm{u}}\right)^2 \phi_\mathrm{r}(\xi) - \rho_\mathrm{A}^{(\mathrm{u})}(\xi)\right] = -\partial_{\xi\xi}\rho_\mathrm{B}^{(\mathrm{u})}(\xi) \tag{3.14}$$

From which we get by imposing the usual boundary conditions 3.7:

$$\rho_\mathrm{B}^{(\mathrm{u})} = c e^{-\xi/\ell_\mathrm{u}} \tag{3.15a}$$

$$\rho_\mathrm{A}^{(\mathrm{u})} = \rho_\mathrm{B}^{(\mathrm{u})} + \left(\frac{\ell_\mathrm{r}}{\ell_\mathrm{u}}\right)^2 \phi_\mathrm{r} + \mathcal{L}\xi \tag{3.15b}$$

$$\partial_{\xi\xi}\phi_\mathrm{r} = \frac{1}{\ell_\mathrm{r}^2}\left\{\Gamma(k\xi)\left[\rho_\mathrm{B}^{(\mathrm{u})} + \left(\frac{\ell_\mathrm{r}}{\ell_\mathrm{u}}\right)^2 \phi_\mathrm{r} + \mathcal{L}\xi\right] + [1-\Gamma(k\xi)]\,\phi_\mathrm{r} - \rho_\mathrm{B}^{(\mathrm{u})}\right\} \tag{3.15c}$$

Which makes the problem diagonal. This is the starting point for the numerical solution in the case $D_\mathrm{r} \neq D_\mathrm{u}$. Getting specific to the case $\ell_\mathrm{r} = \ell_\mathrm{u}$, we first focus on the case in which $k\ell_\mathrm{u} \neq 1$. To solve the third equation we notice it is a linear differential equation in ϕ_r.

$$\partial_{\xi\xi}\phi_\mathrm{r} - \frac{1}{\ell_\mathrm{u}^2}\phi_\mathrm{r} = \frac{1}{\ell_\mathrm{u}^2}\left[\mathcal{L}\xi e^{-k\xi} - c e^{-\xi/\ell_\mathrm{u}} + c e^{-(k+1/\ell_\mathrm{u})\xi}\right] \tag{3.16}$$

The solution will be given by the particular solution plus the sum of the three general solutions, one for each of the three addends on the right handside and will therefore be written in the following form:

$$\phi_{\mathrm{r}} = \sum_i \mathbb{P}_1^{(i)} e^{-h_i \xi}, \quad h_i \in \{k + 1/\ell_{\mathrm{u}}, 1/\ell_{\mathrm{u}}, k\} \tag{3.17}$$

Where $\mathbb{P}_1^{(i)}$ indicates some polynomial of order less or equal to one. By doing so one gets:

$$\phi_{\mathrm{r}}(\xi) = \frac{\mathcal{L}}{(k\ell_{\mathrm{u}})^2 - 1}\left[\xi + \frac{2k\ell_{\mathrm{u}}^2}{(k\ell_{\mathrm{u}})^2 - 1}\right] e^{-k\xi} + \frac{c}{k\ell_{\mathrm{u}}(k\ell_{\mathrm{u}}+2)} e^{-(1/\ell_{\mathrm{u}}+k)\xi} + \left(\frac{c\xi}{2\ell_{\mathrm{u}}} + \mathcal{A}\right) e^{-\xi/\ell_{\mathrm{u}}} \tag{3.18}$$

Where the constant $\mathcal{A}$ has to be determined imposing $\phi_{\mathrm{r}}(0) = 0$, giving:

$$\mathcal{A} = -\left(\frac{2k\ell_{\mathrm{u}}^2\mathcal{L}}{[(k\ell_{\mathrm{u}})^2 - 1]^2} + \frac{c}{k\ell_{\mathrm{u}}(k\ell_{\mathrm{u}}+2)}\right) \tag{3.19}$$

We now want to determine the last unknown, c. To do so we need to exploit the continuity of the derivative in the origin of the unrevealed book. First we study:

$$\lim_{\xi \to 0^+} \phi_{\mathrm{r}}'(\xi) = -\frac{\mathcal{L}}{(k\ell_{\mathrm{u}}+1)^2} + \frac{k}{2(k\ell_{\mathrm{u}}+2)}c \tag{3.20}$$

From Eq.(3.15b), by imposing the continuity of the derivative in zero, one obtains

$$\frac{2c}{\ell_{\mathrm{u}}} = \mathcal{L} + \phi_{\mathrm{r}}'(0) \tag{3.21}$$

giving:

$$c = \frac{2k\ell_{\mathrm{u}}^2(k\ell_{\mathrm{u}}+2)^2}{(k\ell_{\mathrm{u}}+1)^2(3k\ell_{\mathrm{u}}+8)}\mathcal{L} \tag{3.22}$$

By defining $g(\zeta) := \dfrac{2\zeta^2(2+\zeta)^2}{[(1+\zeta)^2(8+3\zeta)]}$ we can write the final solution in the following form:

$$\rho_B^{(u)}(\xi) = \frac{\mathcal{L}}{k} g(k\ell_u) e^{-\xi/\ell_u} \tag{3.23a}$$

$$\rho_A^{(u)}(\xi) = \mathcal{L}\xi + \frac{\mathcal{L}}{k} g(k\ell_u) e^{-\xi/\ell_u} + \phi_r(\xi) \tag{3.23b}$$

$$\begin{aligned} \phi_r(\xi) = \mathcal{L} \Bigg(& \frac{1}{(k\ell_u)^2 - 1} \left[\xi + \frac{2k\ell_u^2}{(k\ell_u)^2 - 1} \right] e^{-k\xi} + \frac{g(k\ell_u)}{k^2 \ell_u (k\ell_u + 2)} e^{-(1/\ell_u + k)\xi} \\ & + \left[\frac{g(k\ell_u)}{2k\ell_u} \xi - \frac{2k\ell_u^2}{[(k\ell_u)^2 - 1]^2} - \frac{g(k\ell_u)}{k^2 \ell_u (k\ell_u + 2)} \right] e^{-\xi/\ell_u} \Bigg) \end{aligned} \tag{3.23c}$$

In the case $k\ell_u = 1$ the solution will have the following form instead:

$$\phi_r = \sum_i \mathbb{P}_2^{(i)}(\xi) e^{-h_i \xi}, \quad h_i \in \{1/\ell_u, k + 1/\ell_u\} \tag{3.24}$$

The steps needed to obtain the solution the same we just illustrated. The result then reads:

$$\rho_B^{(u)} = \frac{9\mathcal{L}\ell_u}{22} e^{-\xi/\ell_u} \tag{3.25a}$$

$$\rho_A^{(u)} = \mathcal{L}\xi + \frac{3\mathcal{L}\ell_u}{22} e^{-2\xi/\ell_u} - \mathcal{L}\left(\frac{\xi^2}{4\ell_u} + \frac{\xi}{22} - \frac{3\ell_u}{11} \right) e^{-\xi/\ell_u} \tag{3.25b}$$

$$\phi_r = \frac{3\mathcal{L}\ell_u}{22} e^{-2\xi/\ell_u} - \mathcal{L}\left(\frac{\xi^2}{4\ell_u} + \frac{\xi}{22} + \frac{3\ell_u}{22} \right) e^{-\xi/\ell_u} \tag{3.25c}$$

This is a continuous limit of the solution for $k\ell_u$ and this sense it doesn't represents a particularly interesting limit. A comparison of the analytical solution with the simulation we performed is showed in Fig.(3.2c)

$D_r \neq D_u$

For $D_r \neq D_u$ the set of Eqs. (3.6) must be solved numerically. We make use of the finite difference method [26]. Recalling the result of Eqs.(3.15) we define the two following quantities:

$$f(y) := \frac{e^{-y}}{\ell_u^2} + \frac{1-e^{-y}}{\ell_r^2} \tag{3.26a}$$
$$g(y) := \frac{[1-e^{-y}]\, e^{-y/(k\ell_u)}}{\ell_r^2} \tag{3.26b}$$
$$\tag{3.26c}$$

Eq.(3.6a) can still be solved analytically, with its value in zero, c, to be fixed. Eqs.(3.6b, 3.6c) will be written as:

$$\frac{2c}{\ell_u} = \left(\frac{\ell_r}{\ell_u}\right)^2 \partial_x \phi(0^+) + \mathcal{L} \tag{3.27a}$$
$$\partial_{\xi\xi}\phi_r = f(k\xi)\phi_r - g(k\xi)c + \frac{\mathcal{L}\xi e^{-k\xi}}{\ell_r^2} \tag{3.27b}$$

We then perform a discretization of space $s.t$ $\xi_k = \xi_1 + (k-1)h$ with $h \to 0$. We impose the boundary conditions: $\phi_r(\xi_0 = 0)$, $\phi_r(\xi_{n+1}) = 0$ where ξ_n is the largest value of ξ we are considering and should be large enough in order not to have too strong effects coming from the boundaries[2]. We then obtain:

$$\frac{2c}{\ell_u} = \left(\frac{\ell_r}{\ell_u}\right)^2 \frac{\phi_r(\xi_1)}{h} + \mathcal{L} \tag{3.28a}$$
$$\frac{1}{h^2}[\phi_r(\xi_{i+1}) + \phi_r(\xi_{i-1}) - 2\phi_r(\xi_i)] - f(k\xi_i)\phi_r(\xi_i) + g(k\xi_i)c = \frac{\mathcal{L}\xi_i e^{-k\xi_i}}{\ell_r^2} \tag{3.28b}$$

That can be rewritten in a matrix form:

[2]Imposing $\phi_r(\xi_{n+1}) = 0$ is a stronger condition then just imposing that ϕ_r doesn't diverge at infinity, however we observe that it is fulfilled in the two limits in which we obtained an analytical solution and we hence extend such result also to this more general limit.

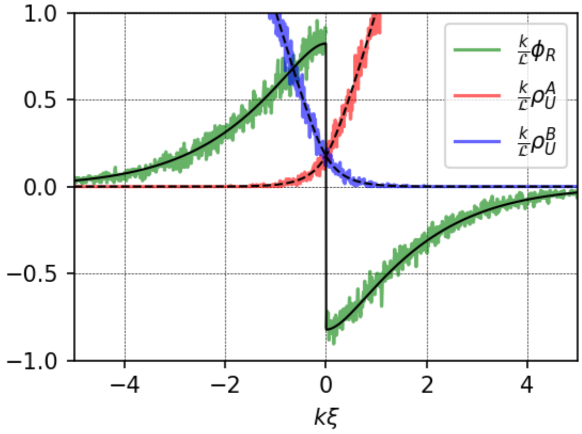
$$\begin{pmatrix} -\frac{2}{\ell_u} & \frac{1}{h}\left(\frac{\ell_r}{\ell_u}\right)^2 & 0 & \dots & \dots & 0 \\ g(k\xi_1) & -\frac{2}{h^2}-f(k\xi_1) & \frac{1}{h^2} & 0 & \dots & 0 \\ \vdots & \frac{1}{h^2} & \ddots & \ddots & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \dots & \vdots \\ g(k\xi_n) & 0 & \dots & \dots & \frac{1}{h^2} & -\frac{2}{h^2}-f(k\xi_n) \end{pmatrix} \begin{pmatrix} c \\ \phi_r(\xi_1) \\ \\ \vdots \\ \\ \phi_r(\xi_n) \end{pmatrix} = \mathcal{L} \begin{pmatrix} -1 \\ \xi_1 e^{-k\xi_1}/\ell_r^2 \\ \\ \vdots \\ \\ \xi_n e^{-k\xi_n}/\ell_r^2 \end{pmatrix} \tag{3.29}$$

And so by inverting it we obtain the vector of the unknowns.

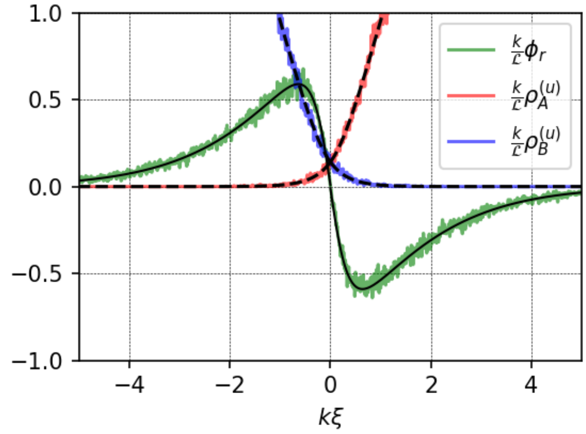
In Fig.(3.2) we confront the three limits just inspected and compare the solution to the theoretical result. As on can see, in the situation $D_r < D_u$ the revealed order book's shape is somewhat in between the cases $D_r = 0$ and $D_r = D_u$, being continuous but with a steeper slope at $\xi = 0$. One might say that a little diffusion in the revealed order book suffices to regularizes the singularity at the trade price.
As it can be seen from Eq. (3.6a) (or from Eqs. (3.12a) and (3.23a) in the particular cases $D_r = 0$ and $D_r = D_u$), in all cases ℓ_u denotes the typical scale of the overlap of the unrevealed books. This is consistent with the idea that ℓ_u is the typical displacement by diffusion of an unrevealed order in the vicinity of the trade price during a time interval ω^{-1}, that is before it gets revealed. Also, it can be seen from Eqs. (3.12c) and (3.23c) in the particular cases $D_r = 0$ and $D_r = D_u$ that the typical horizontal extension of the revealed order book, the order book depth, is given by $\max(k^{-1}, \ell_u)$, consistent with the decay of the revealing probability function Γ and the horizontal extension of the unrevealed books. Note however that, as shall be argued in Sect. 3.3, k^{-1} must always be of order or larger than ℓ_u for stability reasons, and therefore $\max(k^{-1}, \ell_u) \sim k^{-1}$. In the following we shall thus call k^{-1} the order book depth. Finally note that when ℓ_r (equivalently D_r) is decreased while keeping all other parameters constant, the slope of the revealed order book around the origin increases, by that concentrating further the available liquidity around the trade price.

3.2.2 The LLOB limit

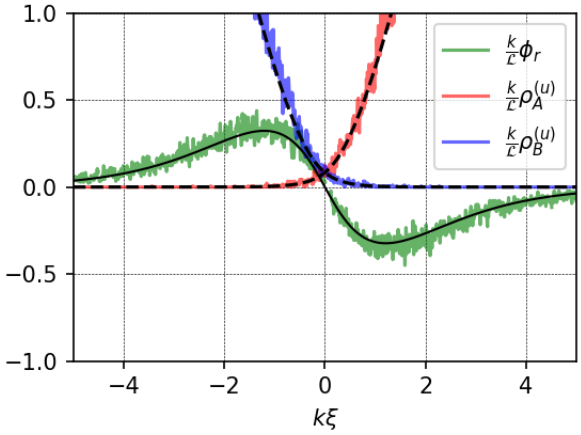
Our model being built upon the locally linear order book model (LLOB) by Donier *et. al* [22] and so we should be able to recover it for certain values of the parameters. Since there is only one diffusion coefficient in the LLOB model, the latter should correspond to the $D_r = D_u$ case.



(A) $\ell_{\mathrm{r}} = 0: \quad k\ell_{\mathrm{u}} = 0.375$



(B) $\ell_{\mathrm{r}} \neq \ell_{\mathrm{u}}: \quad k\ell_{\mathrm{u}} = 0.375, \quad \ell_{\mathrm{u}}/\ell_{\mathrm{r}} = 3.214$



(C) $\ell_{\mathrm{r}} = \ell_{\mathrm{u}}: \quad k\ell_{\mathrm{u}} = 0.375$

FIGURE 3.2: *Stationary solution and comparison to the simulation* : rescaled stationary order books as function of rescaled price. The solid black lines indicate the theoretical rescaled revealed order density $k\phi_{\mathrm{r}}/\mathcal{L}$ while dashed black lines signify the theoretical unrevealed order densities $k\rho^{(\mathrm{u})}_{\mathrm{A/B}}/\mathcal{L}$. The results of the numerical simulation are plotted with color lines on top of the analytical curves.

Then, the LLOB spirit would incline us towards revelation immediacy (no lag effects), which translates into $\omega \to \infty$ or equivalently $\ell_{\mathrm{u}} \to 0$, that is no overlap of the unrevealed books. More rigorously, nondimensionalizing Eqs. (3.23) as $\tilde{\phi}_{\mathrm{r}}, \tilde{\rho}^{(\mathrm{u})} = \frac{k}{\mathcal{L}}\phi_{\mathrm{r}}, \frac{k}{\mathcal{L}}\rho^{(\mathrm{u})}$ and $\tilde{\xi} = k\xi$, one can see that taking the limit $k\ell_{\mathrm{u}} \to 0$ yields for all $\tilde{\xi} > 0$, $\tilde{\rho}^{(\mathrm{u})}_{\mathrm{B}}(\tilde{\xi}) \to 0$, and $\tilde{\rho}^{(\mathrm{u})}_{\mathrm{A}}(\tilde{\xi}) + \tilde{\rho}^{(\mathrm{r})}_{\mathrm{A}}(\tilde{\xi}) = \tilde{\xi} := \tilde{\rho}^{\mathrm{LLOB}}_{\mathrm{A}}(\tilde{\xi})$ that is precisely the LLOB result. To summarize, provided the latent order book of Donier *et. al* is defined as the sum of the unrevealed and revealed books, the LLOB limit is recovered for $k\ell_{\rm u} \ll 1$. The latter condition indicates that the typical displacement $\ell_{\rm u}$ of unrevealed orders in the vicinity of the price must remain small compared to the order book depth k^{-1}. Note that the condition $k \to \infty$, that is revealing precisely and exclusively at the trade price, is thus not required to recover the LLOB limit, as a first intuition could suggest.

3.2.3 Numerical simulation

In order to test our results we performed a numerical simulation of the present model. We define four vectors $\rho_{\rm A}^{(\rm u)}, \rho_{\rm A}^{(\rm u)}, \rho_{\rm A}^{(\rm u)}, \rho_{\rm A}^{(\rm u)}$ of size 2000 and implement the diffusion as already commented in Sect.2.4. The reflecting boundary conditions are imposed on both books, while the current J only on the unrevealed book. Then, some orders in the unrevealed book are drawn from a binomial distribution of parameter $\omega\tau\Gamma(k(x-p_{\rm t}))$ for the ask side (resp. $\Gamma(k(p_{\rm t}-x))$ for the bid side) and moved to the revealed book. Here $p_{\rm t}$ denotes the mid-price. Equivalently revealed orders are moved to the unrevealed book, only with parameter $\omega\tau(1-\Gamma(\cdot))$. Whenever two revealed orders are found at the same price, they are cleared from the book.

3.3 Market stability and calibration to real data

In this paragraph, we address the question of market stability, as given by the amount of liquidity in the revealed order book. We calibrate our model to real order book data and discuss the results in the light of the stability map provided by our model.

3.3.1 Market stability

Imposing that the order densities $\rho_{\rm A}^{(\rm u)}, \rho_{\rm A}^{(\rm r)}, \rho_{\rm B}^{(\rm u)}, \rho_{\rm B}^{(\rm r)}$ must be non negative, consistent with a physically meaningful solution, restricts the possible values of $k, \ell_{\rm u}$.
In the $D_{\rm r} = 0$ case, taking Eq. (3.12c) together with Eq. (3.13) gives $\phi_{\rm r}(0^+) = \mathcal{L}\left[\ell_{\rm u}/2 - 1/k\right]$. Restricting to $\phi_{\rm r}(0^+) \le 0$ (which is tantamount to $\rho_{\rm A}^{(\rm r)}(0^+), \rho_{\rm B}^{(\rm r)}(0^+) \ge 0$) yields $k\ell_{\rm u} \le 2$. Note however that this condition is not sufficient to say that the order densities are everywhere positive, but they are positive only around the origin. In the case of a decreasing exponential revealing, we can express the necessary and sufficient condition of full positiveness, so that $\phi_{\rm r}(\xi) < 0, \forall \xi > 0$. Let's first assume $k\ell_{\rm u} > 1$.

$$\lim_{\xi\to\infty} \phi_{\rm r}(\xi) = \lim_{\xi\to\infty} \mathcal{L} e^{-\xi/\ell_{\rm u}} \left(\frac{\ell_{\rm u}}{2} - \frac{\xi e^{(1/\ell_{\rm u} - k)\xi}}{1 - e^{-k\xi}} \right) = \lim_{\xi\to\infty} \frac{\mathcal{L}\ell_{\rm u} e^{-\xi/\ell_{\rm u}}}{2} > 0 \tag{3.30}$$

From which we get the necessary condition of positiveness:

$$\phi_{\mathrm{r}}(\xi) < 0 \quad \forall \xi \in \mathbb{R}^{+*} \Rightarrow k\ell_{\mathrm{u}} \leq 1 \tag{3.31}$$

To get the converse relation we now assume $k\ell_{\mathrm{u}} \leq 1$

$$\phi_{\mathrm{r}}(\xi) \leq \mathcal{L}\left(\frac{\ell_{\mathrm{u}} e^{-\xi/\ell_{\mathrm{u}}}}{2} - \frac{e^{-k\xi}}{k}\right) := \psi(\xi) \tag{3.32}$$

Where $\psi(\xi)$ is an increasing function of its argument. Given that $\lim_{\xi \to \infty} \psi(\xi) = 0$, the following relation follows

$$\phi_{\mathrm{r}}(\xi) < 0 \quad \forall \xi \in \mathbb{R}^{+*} \iff k\ell_{\mathrm{u}} \leq 1 \tag{3.33}$$

Note that in the case $D_{\mathrm{r}} = 0$, in order to get the double implication of full positiveness we need to express the value of Γ for $y \geq 0$, but we don't if we want to get the positiveness condition around the origin for which is actually enough to specify the value of $\Gamma'(0^+)$. For $1 \leq k\ell_{\mathrm{u}} \leq 2$ the order book displays a "hole" along the price axis, but is well defined around the origin. Since we are most interested in the revealed liquidity in the vicinity of the trade price we choose as the *stability condition* the one that ensures the positiveness of the order book around the origin (*i.e.* $k\ell_{\mathrm{u}} \leq 2$), with no qualitative and only little quantitative effect on our main conclusions. An interesting quantity to look at in this regime is the maximum amplitude of the real order book density that scales as:

$$\max_{\xi} |\phi_{\mathrm{r}}(\xi)| = -\phi_{\mathrm{r}}(0^+) = \frac{\mathcal{L}}{k}\left(1 - \frac{k\ell_{\mathrm{u}}}{\zeta_{\mathrm{c}}}\right), \quad \zeta_{\mathrm{c}} = 2 \tag{3.34}$$

For $D_{\mathrm{u}} = D_{\mathrm{r}}$ the stability condition is imposed by the sign of the slope in zero. Recalling the result of Eqs.(3.20,3.22) and denoting by $\zeta = k\ell_{\mathrm{u}}$ for simplicity of notation, we can express it as:

$$\lim_{\xi \to 0^+} \partial_\xi \phi_{\mathrm{r}}(\xi) = \mathcal{L}\frac{\zeta^3 + 2\zeta^2 - 3\zeta - 8}{(1+\zeta)^2(8+3\zeta)} \tag{3.35}$$

From which we get $\zeta_{\mathrm{c}} = \frac{1}{3}\left[-2 + \left(73 - 6\sqrt{87}\right)^{\frac{1}{3}} + \left(73 + 6\sqrt{87}\right)^{\frac{1}{3}}\right] \approx 1.875$. Note that in this case the solution, provided $k\ell_{\mathrm{u}} > 1$, is also asymptotically unstable:

$$\lim_{\xi \to \infty} \phi_{\mathrm{r}}(\xi) = \frac{\mathcal{L}g(k\ell_{\mathrm{u}})\xi e^{-\xi/\ell_{\mathrm{u}}}}{2k\ell_{\mathrm{u}}} > 0 \tag{3.36}$$

but from Eq.(3.25c) we see that $\lim_{\xi\to\infty} \phi_{\mathrm{r}}(\xi) < 0$ for $k\ell_{\mathrm{u}} = 1$. Arguing that in this case that the maximum of the density can be approximated by

$$\max_{\xi} \left|\phi_{\mathrm{r}}'(0^+)\xi e^{-k\xi}\right| = \frac{|\phi_{\mathrm{r}}'(0^+)|}{ek} \tag{3.37}$$

then we obtain the following scaling, after having performed an expansion for $k\ell_{\mathrm{u}} \approx \zeta_{\mathrm{c}}$:

$$\max_{\xi} \left|\phi_{\mathrm{r}}(\xi)\right| \sim \frac{\mathcal{L}}{ek} \frac{\zeta_{\mathrm{c}}(3\zeta_{\mathrm{c}}^2 + 4\zeta_{\mathrm{c}} - 3)}{(1+\zeta_{\mathrm{c}})^2(8+3\zeta_{\mathrm{c}})} \left(1 - \frac{k\ell_{\mathrm{u}}}{\zeta_{\mathrm{c}}}\right), \quad \zeta_{\mathrm{c}} \approx 1.875\,. \tag{3.38}$$

What should be noticed is that the scaling with $k\ell_{\mathrm{u}}$ of the max is the same as that of the slope.

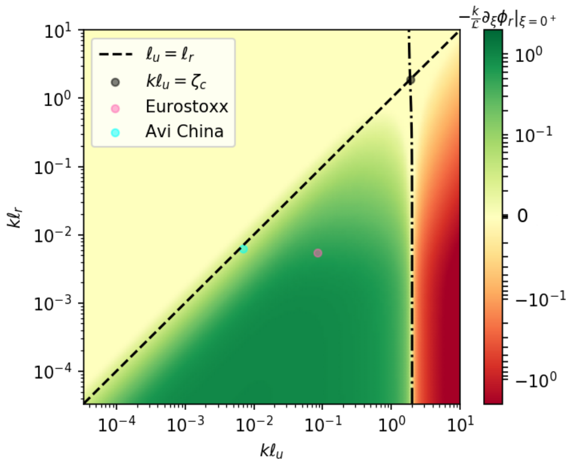
Fig.(3.3a) displays $-\frac{k}{\mathcal{L}}\phi_{\mathrm{r}}'(0^+)$ as function of the dimensionless parameters $k\ell_{\mathrm{u}}$ and $k\ell_{\mathrm{r}}$. The dash-dotted line corresponding to $\phi_{\mathrm{r}}'(0^+) = 0$ splits the parameter space into a stable region (green) and an unstable region (red). The analytical values of ζ_{c} obtained above for $\ell_{\mathrm{u}} = \ell_{\mathrm{r}}$ and $\ell_{\mathrm{r}} \to 0$ are recovered. As one can see, the role played by ℓ_{r} with respect to the position of the critical line ζ_{c} is quite marginal. A more complete study of the value of $\zeta_{\mathrm{c}}(\ell_r/\ell_u)$ is reported if Fig.(3.3c). It is easy to observe that this instability happens when there is a pronounced lag effect. For the sake of completeness, Fig.(3.3b) displays a measure of the overlap between the unrevealed books in the parameter space $(k\ell_{\mathrm{u}}, k\ell_{\mathrm{r}})$. While vanishing in the region $k\ell_{\mathrm{u}} \ll \zeta_{\mathrm{c}}$, the overlap is quite large in the vicinity of the critical line $k\ell_{\mathrm{u}} \lesssim \zeta_{\mathrm{c}}$ and is of the order of k^{-1}, indicating a large volume of unrevealed orders in the vicinity of the price. Combined with a vanishing level of liquidity, the increased level of activity around the origin shall induce important fluctuations of the trade price, as shown in Fig.(3.4). Here we study the numerical volatility of the trade price p_{t} as function of $k\ell_{\mathrm{u}}$ in the $\ell_{\mathrm{r}} = \ell_{\mathrm{u}}$ limit using two different estimators: Rogers-Satchell and Parkinson [45, 43]. By denoting with $p_{\mathrm{H}}, p_{\mathrm{L}}, p_{\mathrm{O}}, p_{\mathrm{C}}$ the high, low, open and close prices respectively, such estimators are defined as :

$$\sigma_{\mathrm{RS}}^2 = \mathbb{E}\left[(p_{\mathrm{H}} - p_{\mathrm{O}})(p_{\mathrm{H}} - p_{\mathrm{C}}) + (p_{\mathrm{L}} - p_{\mathrm{O}})(p_{\mathrm{L}} - p_{\mathrm{C}})\right] \tag{3.39a}$$

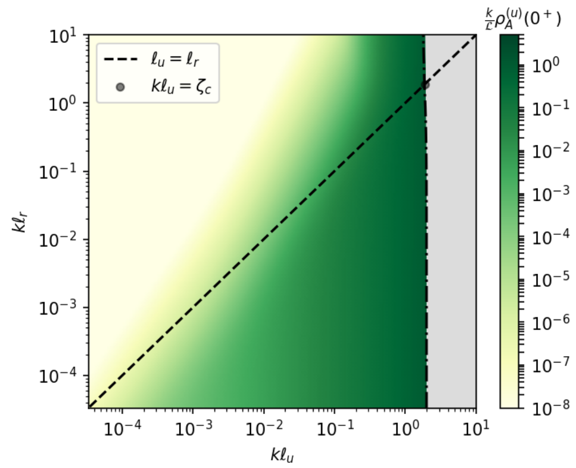
$$\sigma_{\mathrm{P}}^2 = \frac{1}{4\ln(2)}\mathbb{E}\left[(p_{\mathrm{H}} - p_{\mathrm{L}})^2\right] \tag{3.39b}$$

where $\mathbb{E}$ represents a time average. For comparison, we also plotted the numerical volatility of the fair price p_t^{f}, here defined as the value that equilibrates total (revealed and unrevealed) supply and demand:

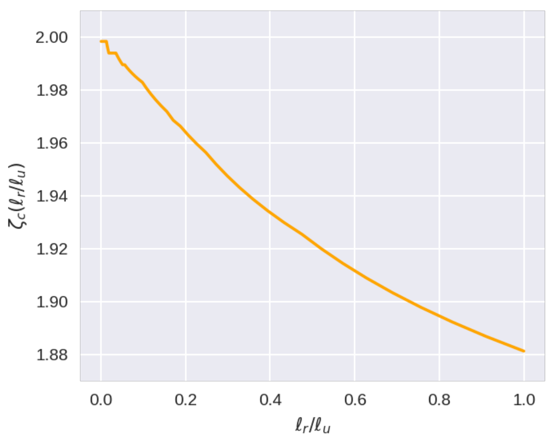
$$\int_0^{p_t^{\mathrm{f}}} \mathrm{d}\xi \left[\rho_{\mathrm{A}}^{(\mathrm{r})}(\xi,t) + \rho_{\mathrm{A}}^{(\mathrm{u})}(\xi,t)\right] = \int_{p_t^{\mathrm{f}}}^{\infty} \mathrm{d}\xi \left[\rho_{\mathrm{B}}^{(\mathrm{r})}(\xi,t) + \rho_{\mathrm{B}}^{(\mathrm{u})}(\xi,t)\right] \tag{3.40}$$



(A)



(B)



(C)

FIGURE 3.3: *Parametric study of the stability of the stationary order books* : (A) Stability map - density plot (symlog scale) of the rescaled slope of the revealed order density at the origin, as function of $k\ell_u$ and $k\ell_r$. The dashed line indicates $\ell_r = \ell_u$, the dash-dotted line indicates the critical line $\phi_r'(0) = 0$, the gray dot indicates the analytical solution $\zeta_c \approx 1,875$ in the $\ell_r = \ell_u$ case. The two colored dots indicate the values of the parameters as inferred from a fit to real data (see below). (B) Overlap of the unrevealed order books - density plot (log scale) of the y-intercept of the unrevealed order book densities at the origin. Note that such a quantity is direct measure of the overlap. (C) Numerical study of the critical value ζ_c.

As one can see, for $k\ell_u \ll \zeta_c$ the volatility of the trade price coincides with its fair price counterpart, consistent with the idea that for small values of ℓ_u the coupling between the revealed and unrevealed books is almost instantaneous ($\omega \to \infty$) and therefore the mid price tends to

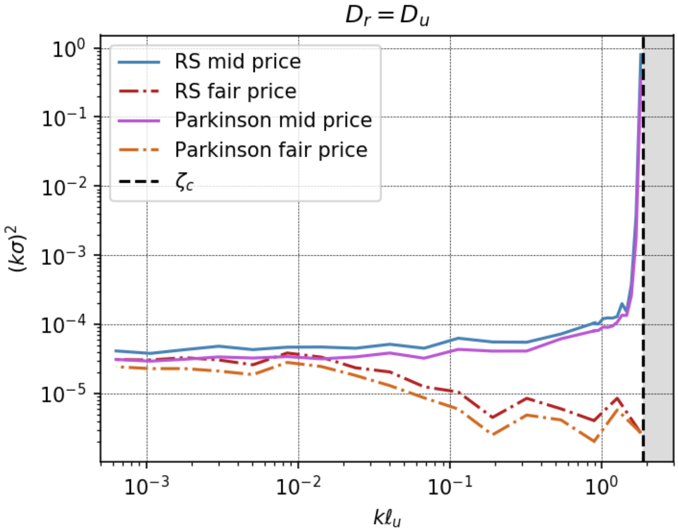
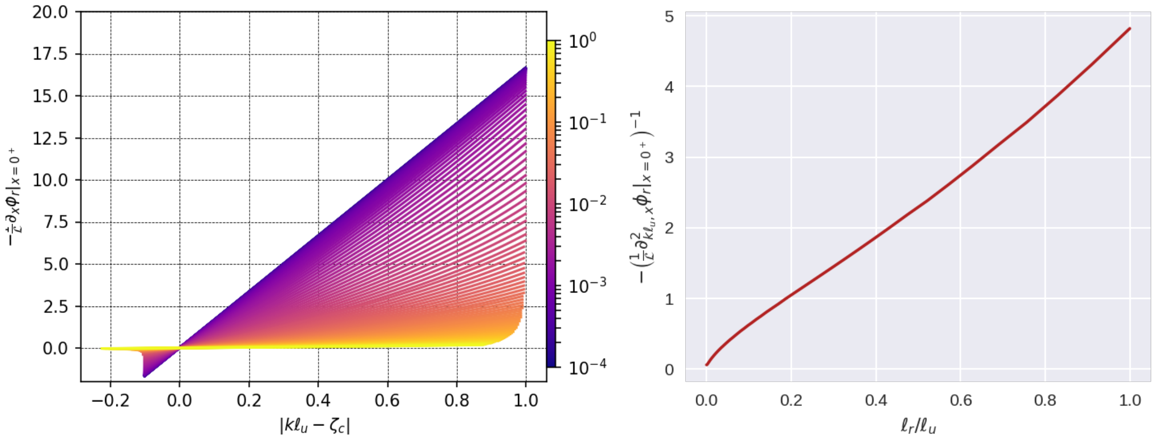


FIGURE 3.4: *Numerical study of the volatility*: plot of rescaled numerical squared volatility of the trade price and the fair price as defined in Eq. (3.40), in the $\ell_r = \ell_u$ case. We here display two different estimations: Rogers-Satchell and Parkinson.

follow the fair price. In the vicinity of the critical line the volatility of the fair price decreases, but not significantly, while the volatility of the trade price strongly diverges as we approach the vanishing liquidity limit $k\ell_u \to \zeta_c$. Note that while the trade price can no longer be defined when a liquidity crisis arises, the fair price as defined in Eq. (3.40) is always well behaved. We also investigated the volatility in the $\ell_r = 0$ limit and obtained similar qualitative results, only with weaker volatility levels consistent with liquidity concentration around the trade price.

Fig.(3.5a) displays the behavior of the slope of the revealed book in the vicinity of the transition. At given ℓ_r/ℓ_u the slope scales linearly with $|k\ell_u - \zeta_c|$. In addition Fig.(3.5b) shows that the slope angular coefficient scales as ℓ_u/ℓ_r, and thus finally $|\phi_r'(0^+)| \sim |k\ell_u - \zeta_c|(\ell_u/\ell_r)$, consistent with the observation that the slope increases as we the value of D_r decreases.
Finally in Fig.(3.6) the plot of how the total volume vanishes when approaching the critical condition. We recall that in that regime the full positiveness of the density is not guaranteed and hence the total available volume should be defined as

$$\mathcal{V}_r := \left| \int_0^\infty \mathrm{d}\xi \, \phi_r(\xi) \mathbb{H}(-\phi_r(\xi)) \right| \tag{3.41}$$



(A) (B)

FIGURE 3.5: *Behavior of the slope at the critical point* : (A) Plot of the slope of the revealed order book at the slope as function of the distance to the critical point $|k\ell_\mathrm{u} - \zeta_\mathrm{c}|$, for different values of $\ell_\mathrm{r}/\ell_\mathrm{u}$. (B) Study of the constant of proportionality of the linear characteristic of (A).

Because of this we can't claim that the total available volume scales as the maximum of the real order book density, but we expect it to vanish faster, as observed from the simulation.

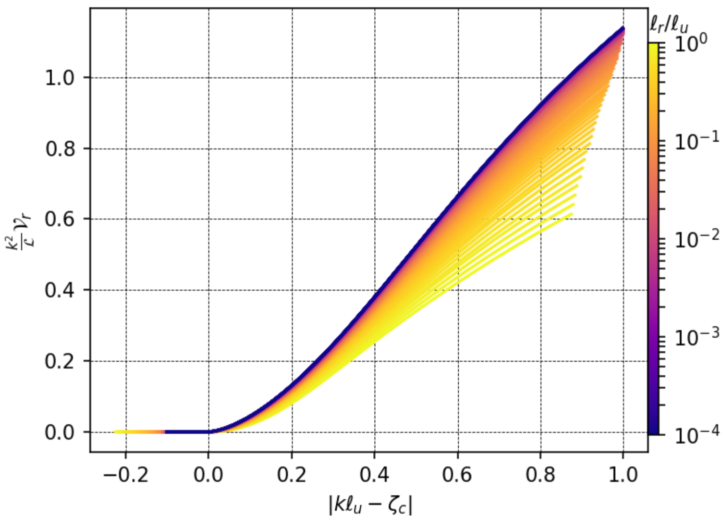


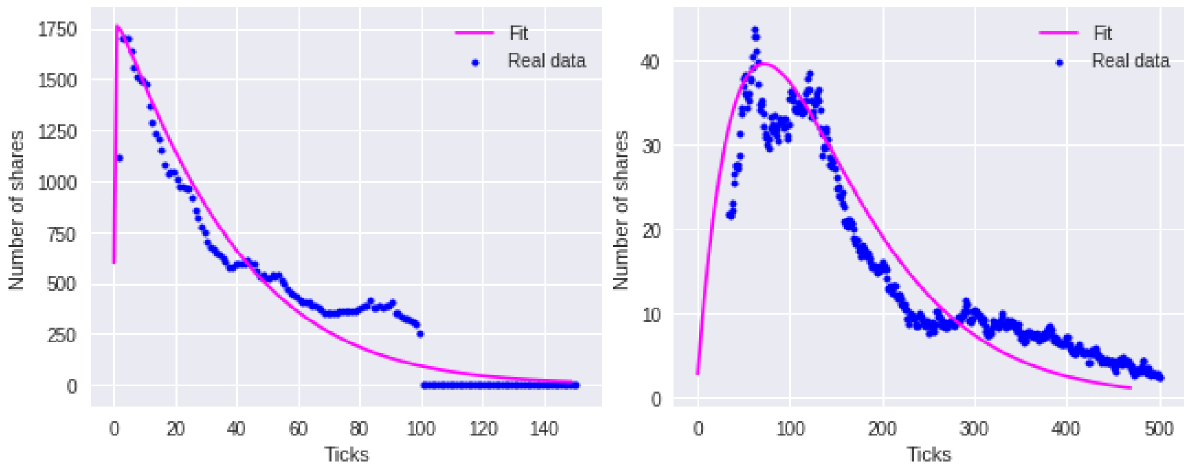
FIGURE 3.6: *Behavior of the total available volume at the critical point*: plot of the total available volume in the revealed order book as a function of the distance to the critical point $|k\ell_\mathrm{u} - \zeta_\mathrm{c}|$ for different values of $\ell_\mathrm{r}/\ell_\mathrm{u}$.

At this point, let us summarize our results. The market is most stable when the unrevealed diffusivity is small or revelations are immediate. In this case, there is a good level of revealed liquidity and the trade price follows the fair price. However, when lag effects become important, and more particularly when ℓ_u (namely the typical displacement of an unrevealed order in the vicinity of the trade price over a time ω^{-1}) becomes of the order of the revealed order book depth, the particles falling on the wrong side will consume the liquidity and the trade price is allowed large excursions from the fair price. As for the effect of diffusion in the real order book, a small ℓ_r concentrates the liquidity around the origin by that providing a wall to price fluctuations, while a large ℓ_r induces a weaker revealed order book slope around the origin facilitating larger price excursions.

3.3.2 Order book data

We here intend to calibrate our model to the real order books[3]. We considered two assets, representing a large (Euro Stoxx) and a small (AviChina) tick. We took a snap shot of the order book every ten minutes, on regular trading hours and averaged over the year 2017. The outputs of our fitting is shown in Fig.(3.7) and gives us four parameters: $\mathcal{L}$, k, ℓ_u and ℓ_r.
What we can see from the fits is that the ratio between $\frac{D_\mathrm{r}}{D_\mathrm{u}}$ changes by approximately a factor 200, coherent with our initial expectations regarding the diffusivity, but also that k is smaller for small ticks, since a larger number of ticks is required to make the same change of relative price as in a large tick market. Regarding the estimation of our parameters, what we see is that at equilibrium we are not able to separate ω from $D_{\mathrm{u}/\mathrm{r}}$ but we can only estimate their ratio. To see their individual contribute we should study the dynamical problem instead. Potentially, given all the parameters, our model allows us to study the evolution of the system and not only its stationary state, so this point has to be considered as quite important as it would allow us to tune completely our parameters and simulate the evolution of the system. To this end, we can think to estimate the parameters in an undirected way. The time $\frac{1}{\omega}$ represents some sort of reaction time of the traders. In the real market we can't think to define the same ω for all the traders, however, in our approximation, we should consider the average reaction time. In [11] Bonart *et. al* carried a study in which it is evidenced the reaction times of the HFT that in 2015 was approximately $10^{-4.5}s$. By extending this study also to low frequency traders, one should be in the end able to have a rough estimation of the value of ω. Another possible way to estimate it would be to measure the average number of orders that are revealed/unrevealed (under the assumption $\omega_\mathrm{r} = \omega_\mathrm{u}$) at price ξ per unit time, giving us a measure of $\Gamma(k\xi)\omega$ that can then be averaged over different values of ξ. Regarding D_r

[3] The data are courtesy of Capital Fund Management in Paris and their elaboration was performed by Antoine Fosset.



(A) Eurostoxx: large tick (B) Avi China: small tick

FIGURE 3.7: *Fit of the stationary book to real data*: in the plot we confront the result of the theoretical curve predicted by our model (pink line) and the average shape of the book (blue dots) for (A) large ticks: Eurostoxx - $\mathcal{L} = 88.3$, $k = 0.045$, $\ell_u/\ell_r \approx 14.4$; (B) small ticks: Avi China - $\mathcal{L} = 1.45$, $k = 0.0135$, $\ell_r/\ell_u \approx 1$.

instead, if one had access to all the transactions and the i.d. of the traders it would be possible to measure it directly and as a consequence to measure indirectly also ω, D_u that would be known at this point from the fitting parameters.

3.4 Price impact

In this section we study how our model reacts in the presence of a metaorder. Following Donier *et al.* [22] we introduce a metaorder as an additional current of buy/sell particles falling precisely at the trade price. The equations governing the system in the presence of a metaorder is left unchanged for the unrevealed order book (Eqs. (3.4a) and (3.4b)), while the right hand side of Eq. (3.4c) must be complemented with the extra additive term $m_t\delta(x - p_t)$ representing the metaorder, where m_t denotes its execution rate. In the following we will restrict to $m_t = m_0 \quad \forall t$ and $m_0 > 0$, with all the results that can be easily extended to the case $m_0 < 0$. In order to extract the dimensionless parameters governing the dynamic system, we introduce $\tilde{x} = kx, \tilde{t} = \omega t, \tilde{\rho} = \frac{k}{\mathcal{L}}\rho, \tilde{\phi}_{\rm r} = \frac{k}{\mathcal{L}}\phi_{\rm r}$ and write the equations in a dimensionless form:

$$\partial_{\tilde{t}}\tilde{\rho}_{\rm B}^{(\rm u)} = (k\ell_{\rm u})^2\partial_{\tilde{x}\tilde{x}}\tilde{\rho}_{\rm B}^{(\rm u)} - \left\{\Gamma(\tilde{p}_{\tilde{t}} - \tilde{x})\tilde{\rho}_{\rm B}^{(\rm u)} - [1-\Gamma(\tilde{p}_{\tilde{t}} - \tilde{x})]\Theta(\tilde{p}_t - \tilde{x})\tilde{\phi}_{\rm r}\right\} \tag{3.42a}$$

$$\partial_{\tilde{t}}\tilde{\rho}_{\rm A}^{(\rm u)} = (k\ell_{\rm u})^2\partial_{\tilde{x}\tilde{x}}\tilde{\rho}_{\rm A}^{(\rm u)} - \left\{\Gamma(\tilde{x} - \tilde{p}_{\tilde{t}})\tilde{\rho}_{\rm A}^{(\rm u)} + [1-\Gamma(\tilde{x} - \tilde{p}_{\tilde{t}})]\Theta(\tilde{x} - \tilde{p}_t)\tilde{\phi}_{\rm r}\right\} \tag{3.42b}$$

$$\partial_{\tilde{t}}\tilde{\phi}_{\rm r} = (k\ell_{\rm r})^2\partial_{\tilde{x}\tilde{x}}\tilde{\phi}_{\rm r} - \left\{\Gamma(\tilde{x} - \tilde{p}_{\tilde{t}})\tilde{\rho}_{\rm A}^{(\rm u)} - \Gamma(\tilde{p}_{\tilde{t}} - \tilde{x})\tilde{\rho}_{\rm B}^{(\rm u)} + [1-\Gamma(|\tilde{x} - \tilde{p}_{\tilde{t}}|)]\tilde{\phi}_{\rm r}\right\} + \frac{m_0}{\mathcal{J}}\delta(\tilde{x} - \tilde{p}_{\tilde{t}}) \,, \tag{3.42c}$$

with $\mathcal{J} = \frac{\mathcal{L}\omega}{k^2}$ the typical overall revealing current (with $\frac{\mathcal{L}}{k^2}$ is the typical available volume in the revealed order book). The set of Eqs. (3.42) must now be complemented by a dynamic boundary condition at $x = p_t$. Integrating the Eq. (3.42c) over an infinitesimal interval around $p_{\rm t}$ one obtains:

$$\partial_{\tilde{x}}\tilde{\phi}_{\rm r}(\tilde{p}_{\tilde{t}}^{+}) - \partial_{\tilde{x}}\tilde{\phi}_{\rm r}(\tilde{p}_{\tilde{t}}^{-}) + \frac{m_0}{J_{\rm r}} = 0 \tag{3.43}$$

with $J_{\rm r} = D_{\rm r}\mathcal{L}$. Weighing the first and third terms on the right hand side of Eq. (3.42c) yields a relevant dimensionless number $m_0[\mathcal{J}(k\ell_{\rm r})^2]^{-1} = \frac{m_0}{J_{\rm r}}$. In the case $\ell_{\rm r} = 0$ we should consider $\frac{m_0}{\mathcal{J}}$ instead and it has to be compared to one, magnitude of the revealing current. Note that when $\ell_{\rm r} \neq 0$, due to the stability condition $(k\ell_{\rm r})^2 \lesssim 1$, so it is enough to compare such term to $\frac{m_0}{\mathcal{J}}$ without loss of information being that the revealing current will always be large compared to the diffusive term.
In order to compute the price impact $\mathcal{I}(\mathcal{Q}_t) = \langle p_t - p_0 | \mathcal{Q}_t = m_0 t\rangle$ we performed some numerical simulations of our model in the presence of a metaorder in several limit cases. We explored in particular $\ell_{\rm r} = \ell_{\rm u}$ and $\ell_{\rm r} = 0$, in both high and low participation rate regimes, for different values of $k\ell_{\rm u}$. The results are shown in Figs.(3.8 and 3.9).

Before engaging in presenting the results of the numerical simulations, note that there exists a regime where the calculations can be brought a little bit further analytically, that is when we can give a geometrical interpretation to the problem. When the book is almost static on the time scale of the metaorder execution *i.e.* $m_0 \gg J_{\rm u}$ (resp. $m_0 \gg \mathcal{J}$ for $\ell_{\rm r} = 0$), one has

$$\int_0^t {\rm d}s\, m_0 = -\int_0^{p_t} {\rm d}\xi\, \phi_{\rm r}^{(\rm s)}(\xi) \tag{3.44}$$

with $\xi = x - p_0$ and $\phi_{\rm r}^{(s)}$ the stationary solution of the book. For $D_{\rm r} = D_{\rm u}$, the integral is quite easy to perform. For simplicity of notation we indicate with $a = [(k\ell_{\rm u})^2 - 1]^{-1}$, $b = 2ak\ell_{\rm u}^2$, $c = g(k\ell_{\rm u})[k^2\ell_{\rm u}(k\ell_{\rm u}+2)]^{-1}$, $d = g(k\ell_{\rm u})[2k\ell_{\rm u}]^{-1}$:

$$m_0 t = \frac{\mathcal{L}a}{k^2}\left\{[k(p_t+b)+1]e^{-kp_t}-(kb+1)\right\} - \frac{\mathcal{L}\ell_{\mathrm{u}}c}{1+k\ell_{\mathrm{u}}}\left(1-e^{-(k+1/\ell_{\mathrm{u}})p_t}\right)+ \\ \mathcal{L}\ell_{\mathrm{u}}dp_t e^{-p_t/\ell_{\mathrm{u}}} - \mathcal{L}[d\ell_{\mathrm{u}}^2+\ell_{\mathrm{u}}(ab+c)]\left(1-e^{-p_t/\ell_{\mathrm{u}}}\right) \quad (3.45)$$

Consider now the case $\ell_{\mathrm{r}} = 0$ instead:

$$m_0 t = -\int_0^{p_t} d\xi \phi_{\mathrm{r}}^{st}(\xi) = -\int_0^{p_t} d\xi \left(\frac{\mathcal{L}\ell_{\mathrm{u}}}{2}e^{-\xi/\ell_{\mathrm{u}}} - \frac{\mathcal{L}\xi e^{-k\xi}}{1-e^{-k\xi}}\right) \quad (3.46)$$

The first addend is trivial and will be explicitly reported only in the final solution. Let's focus on the second term instead. First we perform an integration by parts:

$$\int d\xi\, \mathcal{L}\xi\left(\frac{e^{-k\xi}}{1-e^{-k\xi}}\right) = \frac{\mathcal{L}\xi}{k} log\left(1-e^{-k\xi}\right) - \frac{\mathcal{L}}{k}\int d\xi\, log\left(1-e^{-k\xi}\right) \quad (3.47)$$

To solve the last integral we now perform a change of variable: $y = e^{-k\xi}$, where we notice that $0 \leq y \leq 1$ in our domain of interest, *i.e.* $\xi \geq 0$.

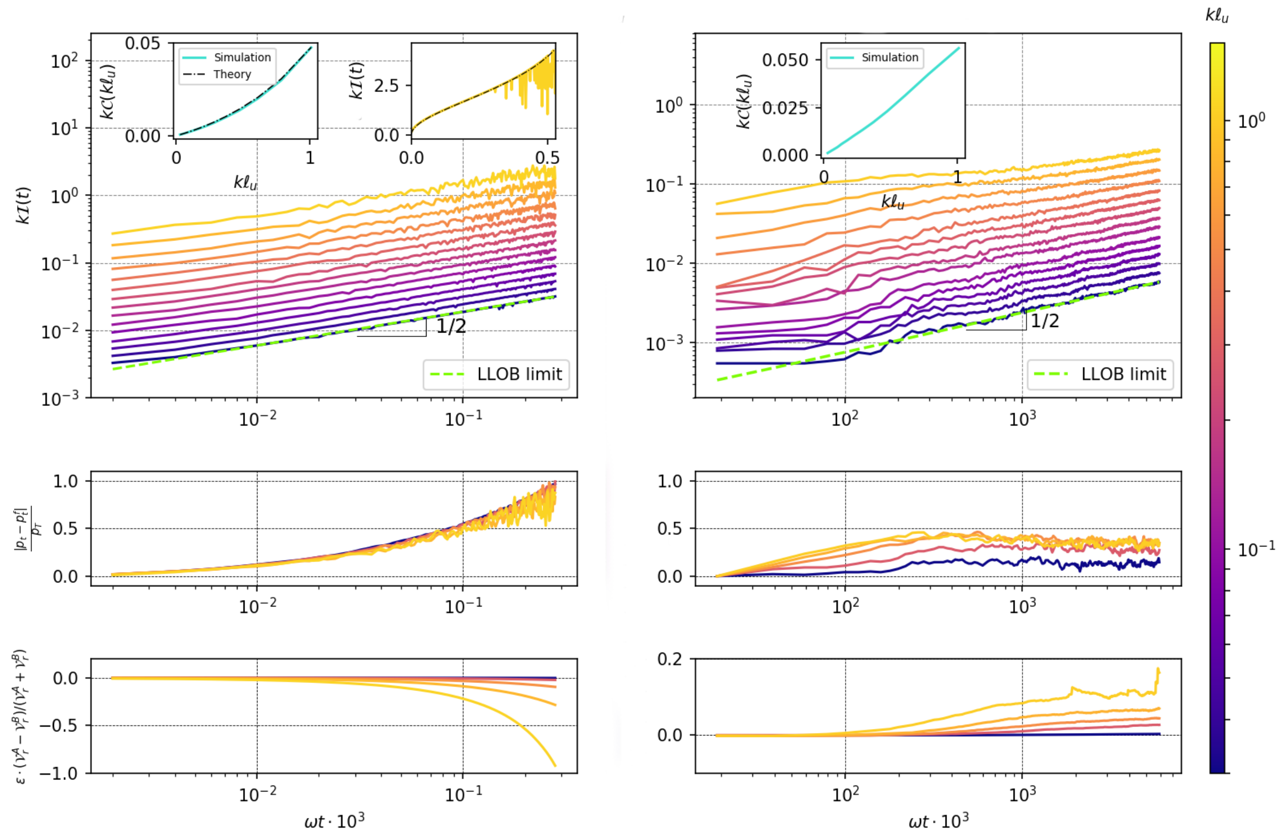
$$\int d\xi\, log\left(1-e^{-k\xi}\right) = -\frac{\mathcal{L}}{k^2}\int dy\, \frac{log(1-y)}{y} = -\frac{\mathcal{L}}{k^2}\int dy\, \frac{1}{y}\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{n}(-y)^n = \\ = \frac{\mathcal{L}}{k^2}\sum_{n=0}^{\infty}\frac{y^n}{n^2} = \frac{\mathcal{L}}{k^2}Li_2\left(e^{-k\xi}\right) \quad (3.48)$$

Where the definition of polylogarithm of order n – Li_n – was used [36]. Finally we obtain the following solution

$$m_0 t = -\frac{\mathcal{L}\ell_{\mathrm{u}}^2}{2}\left(1-e^{-p_t/\ell_{\mathrm{u}}}\right) + \frac{\mathcal{L}p_{\mathrm{t}}}{k}\log\left(1-e^{-kp_{\mathrm{t}}}\right) - \frac{\mathcal{L}}{k^2}\left[\mathrm{Li}_2(e^{-kp_{\mathrm{t}}}) - \mathrm{Li}_2(1)\right] \quad (3.49)$$

Both solutions have to be inverted numerically in order to obtain the price trajectory p_t.
Let us now comment the more general results we obtained numerically for both limit cases $\ell_{\mathrm{u}} = \ell_{\mathrm{r}}$ and $\ell_{\mathrm{r}} = 0$.

$\boldsymbol{\ell_u = \ell_r}$



(A) Fast execution rate : $m_0 \gg J_r$

(B) Slow execution rate : $m_0 \ll J_r$

FIGURE 3.8: *Price impact for* $\ell_r = \ell_u$: the dashed green lines indicates the LLOB limits, $\mathcal{I}_{\text{LLOB}}(t) = \sqrt{\alpha \mathcal{Q}_t/(\pi\mathcal{L})}$ for the slow regime and $\mathcal{I}_{\text{LLOB}}(t) = \sqrt{2\mathcal{Q}_t/\mathcal{L}}$ for the fast regime. The top left insets on each plot indicate the factor c as function of $k\ell_u$ defined as $\mathcal{I}(t) = c\,\mathcal{I}_{\text{LLOB}}(t)$.The top right inset of subplot (A) shows an extreme regime with very high execution rate, the dash-dotted line indicates the theoretical prediction as given by the numerical inversion of Eq.(3.45). The smaller subplots underneath display relative price difference between the trade price and the fair price as defined in Eq.(3.40) and the relative revealed volume imbalance.

The main plots in Figs.(3.8) display robust square root price trajectories, regardless of the values of $k\ell_u$. For $m_0 \gg J_r$ the price trajectory matches the theoretical prediction given above inverting Eq.(3.45). As expected from exponential vanishing liquidity at $x - p_0 > k^{-1}$

the impact eventually diverges for very extreme regimes, as shown in the top right inset in Fig.(3.8a). For $k\ell_u \ll 1$, one recovers the LLOB limit in both fast and slow regimes, also as expected. For non vanishing values of $k\ell_u$, the impact increases with increasing $k\ell_u$. In particular for $m_0 \gg J_r$ and small times (*i.e.* small volumes), one can perform an expansion of ϕ_r around the origin to the first non vanishing order obtaining:

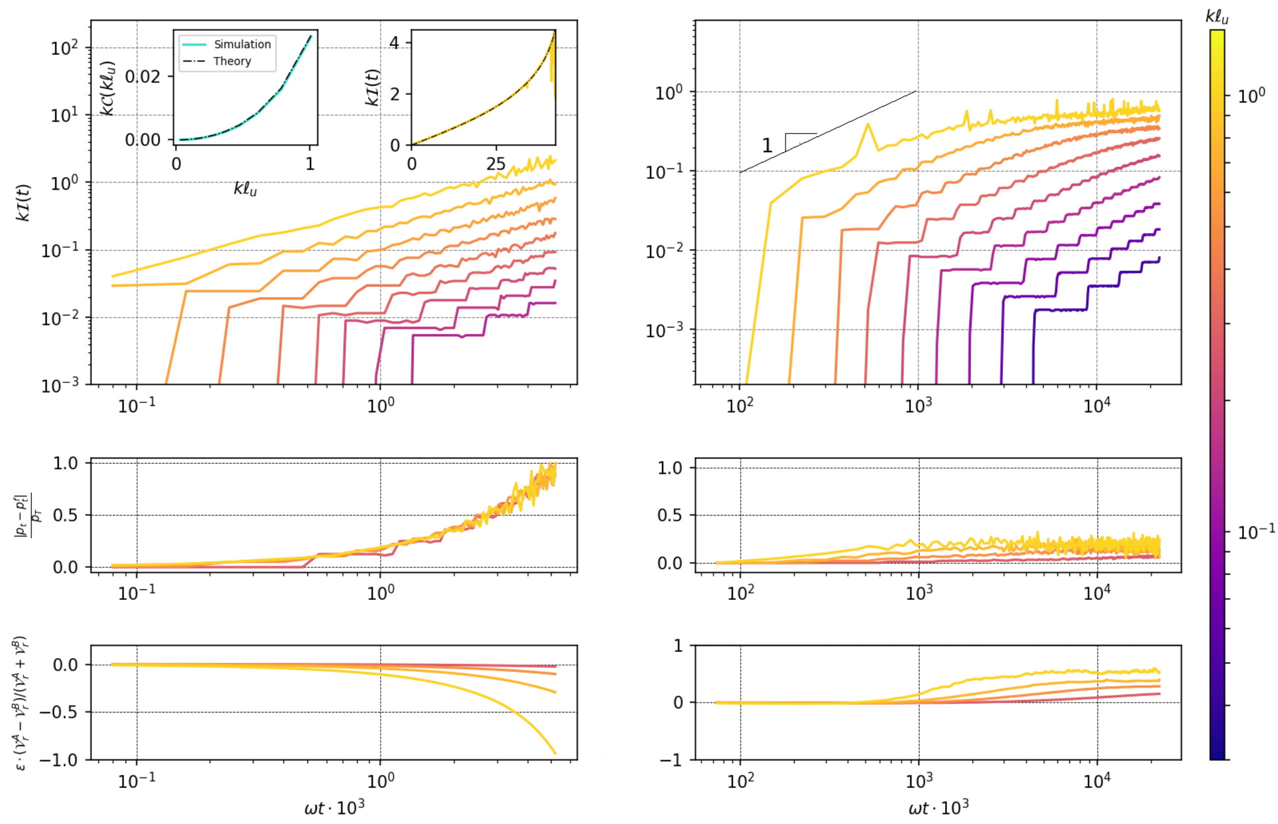
$$\mathcal{I}(\mathcal{Q}_t) = \sqrt{\frac{\mathcal{Q}_t}{|\partial_x \phi_r(0^+)|}} \sim \sqrt{\mathcal{Q}_t}\left(1 - \frac{k\ell_u}{\zeta_c}\right)^{-1/2}. \tag{3.50}$$

The plots in the second row display the relative distance between the trade price and the fair price as function of time. In the fast execution regime all curves fall on top of each other and $|p_T - p_T^f| \approx |p_T|$, consistent with the idea that the book (in particular the unrevealed one) does not have time to reassess during the execution and, as a consequence, the fair price varies at a much slower rate than the trade price. A different scenario takes place in the small execution rate regime. We observe that the relative distance between trade and fair prices stabilizes. In other terms, the unrevealed order book evolves at a speed that is comparable to that of the metaorder and the fair price follows the trade price quite accurately.
The evolution of relative volume imbalance (third row) allows to draw similar conclusions. In the fast execution limit, the imbalance diverges as the unrevealed order book has no time to refill the revealed order book (this effect is all the more evident as we approach the vanishing liquidity limit $k\ell_u \to \zeta_c$), while in the slow execution limit the imbalance is much smaller. Most importantly, note that in this limit the imbalance becomes positive. This is consistent with the fact that when the trade price moves slowly, the revealing probability Γ is shifted with it and new orders reveal on top of the existing ones to supply the metaorder, while the orders left behind progressively unreveal.

$\ell_r = 0$

The limit $\ell_r \to 0$ corresponds to $\frac{m_0}{J_r} \to \infty$, so in some sense one could say that we are always in a high participation rate regime. However the absence of diffusion means that the system can only evolve through the revealing and unrevealing currents. As mentioned above, in this case the relevant dimensionless number becomes $\frac{m_0}{\mathcal{J}}$, that will be referred to as participation rate in the following.
Fig(3.9a) displays price trajectories in the fast execution regime, here $m_0 \gg \mathcal{J}$. The metaorder is faster than the revealing current and the price trajectory is given by inverting Eq. (3.49) as we did already for $\ell_r = \ell_u$.

$$\mathcal{I}(\mathcal{Q}_t) = \frac{k\mathcal{Q}_t}{\mathcal{L}}\left(1 - \frac{k\ell_u}{\zeta_c}\right)^{-1} \tag{3.51}$$



(A) Fast execution rate : $m_0 \gg \mathcal{J}$

(B) Slow execution rate : $m_0 \ll \mathcal{J}$

FIGURE 3.9: *Price impact for* $\ell_r = 0$: the top left insets on subplot (A) indicates the factor c as function of $k\ell_u$ here defined as $\mathcal{I}(t) = ct$ for $\ell_r = 0$. The top right insets show an extreme regimes with very high execution rate for which the impact diverges. The second row displays relative price difference between the trade price and the fair price as defined in Eq.(3.40). The third row displays the relative revealed volume imbalance.

which is linear impact, consistent with $\phi_r(0) \neq 0$. Analogous to the case $\ell_u = \ell_r$, in extreme regimes the impact eventually diverged as shown in the top right inset of Fig.(3.9a). Regarding the imbalance and the relative distance between the trade price and the fair price, the interpretation is analogous to the $\ell_r = \ell_u$ case.

Fig.(3.9b) displays the price trajectories in the low participation rate regime $m_0 \ll \mathcal{J}$. Here the impact is genuinely linear at short times but crossovers to concave after a typical time $t^\star \sim \mathcal{V}_r / m_0$. This interesting regime can be easily understood as follows. At short times the metaorder is executed against the orders present in the stationary (locally constant) revealed order book, but as time passes liquidity must be revealed from the (linear) unrevealed order book. As $k\ell_{\mathrm{u}}$ is increased $t^\star$ is decreased consistent with that the larger $k\ell_{\mathrm{u}}$, the smaller the revealed liquidity $\mathcal{V}_{\mathrm{r}}$ and thus the sooner the unrevealed liquidity takes over the initial state. Note that linear impact at short times (equiv. small volumes) has been reported in the literature [55]. Also note that recovering precisely the square root in this regime is quite challenging because the numerical simulation is by essence discontinuous by that inducing artificial spread effects, and obtaining smooth results requires a lot of averaging. In the second row, we observe that the relative distance between trade and fair prices stabilizes after the typical time $t^\star$. In other terms, for $t > t^\star$ the unrevealed order book evolves at a rate comparable to that of the metaorder, and the fair price follows the trade price with some constant lag. The third row displays similar conclusions.

3.5 Comments and remarks

The present model is, to our knowledge, the first one to describe a link between the latent liquidity and the real limit order book, providing the possibility to infer parameters of the unrevealed books from the revealed ones.

Among the great achievements we number the theoretical prediction of a square root impact that we discussed is observable provided there exists dome diffusion in the real order book. On the other hand, if $D_{\mathrm{r}} = 0$, we observe a linear and then concave impact, coherent with what experimentally observed, allowing us to think of an extension at two time scale liquidity with different diffusion coefficients that will be further discussed in Chapter 4 and should be able to give a microstructural interpretation of the impact in both regimes, concave and linear.

Furthermore this model predicts some stability conditions and it is able to give a interpretation of how an endogenous crisis is originated: in particular we observed that, as we approach the critical condition, the liquidity will vanish and the volatility will increase, generating large jumps of the trade price. Consistent with the work of Joulin *et.al* [29], this model is then capable to address the origin of endogenous price jumps to liquidity dry outs.

From the parameter fitting to real data we were also able to observe a good agreement between the theoretically predicted results and the measurements, obtaining a behavior of the fitting parameters in good agreement with what expected.

In these terms the model we proposed represents a good starting point to describe the statistical behavior of the limit order book, suggesting how a market could be made more stable acting, for example, on the tick size.

If on one hand we observe very satisfactory results, we are also aware that this model lacks a number of characteristics that should have in order to efficiently describe the limit order book. In particular, as a inheritance of the LLOB model, ours doesn't solve the diffusivity puzzle and doesn't keep into account the varied behavior of the different agents, claiming the possibility to define some parameters $k, \omega, D_{\mathrm{u/r}}$ to accurately model them all. In the same spirit, we recall that the LLOB, and hence our model, doesn't keep into account the interaction between different assets and therefore cross impact. Further discussion on these last aspects will be treated more in detail in Chapter 4.

Chapter 4

Extensions of the model

In this final chapter we intend to discuss about how some approximations we made during the the main work can be relaxed – giving an overall more general result – as well as discussing what could be the direction in which to move to make improvements to this model, considering some aspects that where neglected so far.

4.1 Relaxing some constraints

4.1.1 The expression of the revealing probability

As we mentioned already, this is a key element of our model and we had to make some not completely justified assumption in order to keep the equations analytically treatable.
In the most straightforward revealing mechanism a trader who ended up on the wrong side of the book would very likely to be willing to reveal his position, but because of trading rules, the market order will always be placed at the best quote and not at an arbitrary price among the orders of opposite sign. More explicitly, if we put a buy order at price x with $x > a_t$, in our model it will be placed and executed at price x, while in the real market it would be executed at price a_t. We run some simulation with this new revealing mechanism and, as expected, we didn't observe any major difference from the results we discussed. All the laws we found regarding the impact and the shape of the book are left essentially unchanged.
Another important aspect to be considered is that no trader ever wants to make an inconvenient deal, so if he finds himself on the opposite side of the book, instead of revealing at once, he might want to reevaluate his position. To this end we have also studied a slightly different shape of the revealing probability that models in a very simple way this behavior:

$$\Gamma(y) = e^{-|y|} \tag{4.1}$$

In order to keep the equations treatable we assumed Γ to be symmetric, but there is actually no real reason to claim that, since the function is decaying on the positive and negative axis as a consequence of two different mechanisms. We will exploit again the symmetry of the problem to solve it only on $\mathbb{R}^{*+}$ and move into the reference frame of p_t with $\xi = x - p_t$. The system of equations to solve is therefore:

$$D_{\mathrm{u}}\partial_{\xi\xi}\rho_{\mathrm{B}}^{(\mathrm{u})} = \omega e^{-k\xi}\rho_{\mathrm{B}}^{(\mathrm{u})} \tag{4.2a}$$

$$D_{\mathrm{u}}\partial_{\xi\xi}\rho_{\mathrm{A}}^{(\mathrm{u})} = \omega\left[e^{-k\xi}\rho_{\mathrm{A}}^{(\mathrm{u})} + (1-e^{-k\xi})\phi_{\mathrm{r}}\right] \tag{4.2b}$$

$$D_{\mathrm{r}}\partial_{\xi\xi}\phi_{\mathrm{r}} = -\omega\left[e^{-k\xi}\rho_{\mathrm{B}}^{(\mathrm{u})} - e^{-k\xi}\rho_{\mathrm{A}}^{(\mathrm{u})} - (1-e^{-k\xi})\phi_{\mathrm{r}}\right] \tag{4.2c}$$

We will proceed again to its solution in the cases of $D_{\mathrm{r}} = 0$ and $D_{\mathrm{r}} = D_{\mathrm{u}}$.

$D_{\mathrm{r}} = 0$

Eq.(4.2c) can be rewritten using the other two as $\partial_{\xi\xi}\rho_{\mathrm{A}}^{(\mathrm{u})} = \partial_{\xi\xi}\rho_{\mathrm{B}}^{(\mathrm{u})}$ that combined with the usual boundary conditions gives

$$\rho_{\mathrm{A}}^{(\mathrm{u})}(\xi) = \rho_{\mathrm{B}}^{(\mathrm{u})}(\xi) + \mathcal{L}\xi \tag{4.3}$$

Plugging this into the third equation we find

$$\phi_{\mathrm{r}}(\xi) = -\frac{\mathcal{L}\xi e^{-k\xi}}{1-e^{-k\xi}} \tag{4.4}$$

Which leaves us only Eq.(4.2a) to solve. Let's assume the following shape of the solution: $\rho_{\mathrm{B}}^{(\mathrm{u})}(\xi) := f\left(\mathcal{A}e^{-k\xi/2}\right)$ where $\mathcal{A}$ is some constant to be determined. Since the exponential is bijective, this can be done without loss of generality. By calling $z = \mathcal{A}e^{-k\xi/2}$ and choosing $\mathcal{A} = 2\sqrt{\frac{\omega}{D_{\mathrm{u}}k^2}} := \frac{2}{k\ell_{\mathrm{u}}}$, the Eq.(4.2a) can be rewritten as

$$z^2\partial_{zz}f(z) + z\partial_z f(z) - z^2 f(z) = 0 \tag{4.5}$$

That is the definition of the modified Bessel function of order zero, $\mathcal{I}_0(z)$ [4]. So we found that

$$\rho_B^{(u)}(\xi) = \mathcal{C}\mathcal{I}_0\left(\frac{2}{k\ell_u}e^{-k\xi/2}\right) \tag{4.6}$$

To find the constant $\mathcal{C}$ we use the condition $\partial_\xi \rho_A^{(u)}|_{\xi=0^+} = -\partial_\xi \rho_B^{(u)}|_{\xi=0^+}$. We write the modified Bessel function of order n as the following series

$$\mathcal{I}_n(y) = \sum_{m=0}^{\infty} \frac{1}{m!(m+n)!}\left(\frac{y}{2}\right)^{2m+n} \tag{4.7}$$

From which is simple to verify the property $\mathcal{I}_{-n}(y) = \mathcal{I}_n(y)$ and that $\partial_y \mathcal{I}_0(y) = \mathcal{I}_1(y)$. So, since

$$-2\partial_\xi \rho_B^{(u)}|_{\xi=0^+} = \mathcal{L} \tag{4.8}$$

we can express the final solution on the positive ξ axis as follows.

$$\rho_B^{(u)} = \frac{\mathcal{L}\ell_u}{2\mathcal{I}_1\left(\frac{2}{k\ell_u}\right)} \cdot \mathcal{I}_0\left(\frac{2}{k\ell_u}e^{-\frac{k\xi}{2}}\right) \tag{4.9a}$$

$$\rho_A^{(u)} = \rho_B^{(u)} + \mathcal{L}\xi \tag{4.9b}$$

$$\phi_r = -\frac{\mathcal{L}\xi e^{-k\xi}}{1-e^{-k\xi}} \tag{4.9c}$$

Two important limits should be studied:

$$\lim_{\xi\to 0^+} \phi_r(\xi) = -\frac{\mathcal{L}}{k} \neq 0 \tag{4.10}$$

That, combined with the property $\phi_r(-\xi) = -\phi_r(\xi)$ implies that the function $\phi_r(\xi)$ is not continuous in zero.

$$\lim_{\xi\to\infty} \rho_B^{(u)}(\xi) = \frac{\mathcal{L}\ell_u}{2\mathcal{I}_1\left(\frac{2}{k\ell_u}\right)} \neq 0 \tag{4.11}$$

So this function has a finite limit at infinity, but this limit is not equal to zero. In Fig.(4.1a) the comparison of the analytical solution with our simulation.

$D_u = D_r$

Let us now solve the system (4.2) for $D_r = D_u$. By using the first two equations we can rewrite the third as

$$\partial_{\xi\xi}\phi_r(\xi) = \partial_{\xi\xi}\rho_A^{(u)}(\xi) - \partial_{\xi\xi}\rho_B^{(u)}(\xi) \tag{4.12}$$

That becomes, by adding the usual boundary conditions

$$\rho_A^{(u)}(\xi) = \rho_B^{(u)}(\xi) + \phi_r(\xi) + \mathcal{L}\xi \tag{4.13}$$

By injecting this expression back into Eq.(4.2c), we then obtain

$$\partial_{\xi\xi}\phi_r(\xi) = \frac{1}{\ell_u}\left(\mathcal{L}\xi e^{-k\xi} + \phi_r(\xi)\right) \tag{4.14}$$

We now rewrite this equation, without loss of generality, as $\phi_r(\xi) = \mathcal{A}(\xi)e^{-k\xi}$. The equation to solve is now

$$\ell_u^2\partial_{\xi\xi}\mathcal{A}(\xi) - 2k\ell_u^2\partial_\xi\mathcal{A}(\xi) + [(k\ell_u)^2 - 1]\mathcal{A}(\xi) = \mathcal{L}\xi \tag{4.15}$$

The particular solution reads

$$\mathcal{A}_p(\xi) = \frac{\mathcal{L}}{(k\ell_u)^2 - 1}\left[\xi + \frac{2k\ell_u^2}{(k\ell_u)^2 - 1}\right], \quad if \quad k\ell_u \neq 1 \tag{4.16a}$$

$$\mathcal{A}_p(\xi) = -\frac{\mathcal{L}}{4}(k\xi^2 + \xi) + \gamma, \quad \gamma \in \mathbb{R}, \quad if \quad k\ell_u = 1 \tag{4.16b}$$

That gives the following shape, once we add the general solution

$$\phi_r(\xi) = \left[\mathcal{C}_1 e^{(k-1/\ell_u)\xi} + \mathcal{C}_2 e^{(k+1/\ell_u)\xi} + \mathcal{A}_p(\xi)\right] e^{-k\xi} \tag{4.17}$$

where the constants $\mathcal{C}_1, \mathcal{C}_2$ are to be fixed with the boundary conditions

$$\lim_{\xi\to\infty}\phi_r(\xi) = c \in \mathbb{R} \tag{4.18a}$$

$$\phi_r(0) = 0 \tag{4.18b}$$

To solve the first two equations we can exploit the results we already obtained. In particular Eq.(4.2a) is the same as in the case $D_r = 0$ and is decoupled from the rest, so the solution will be of the same form as Eq.(4.6). This leaves us to impose the proper boundary conditions.

Solution for $k\ell_u \neq 1$

The constant $\mathcal{C}$ should be obtained by imposing the continuity of the first derivative in the origin of the unrevealed order book, *i.e.* $\partial_\xi \rho_A^{(u)}|_{\xi=0^+} = -\partial_\xi \rho_B^{(u)}|_{\xi=0^+}$. The result gives

$$\rho_B^{(u)}(\xi) = \mathcal{L}\frac{k\ell_u^2(k\ell_u+2)}{2(k\ell_u+1)^2\mathcal{I}_1\left(\frac{2}{k\ell_u}\right)}\mathcal{I}_0\left(\frac{2}{k\ell_u}e^{-\frac{k\xi}{2}}\right) \tag{4.19a}$$

$$\rho_A^{(u)}(\xi) = \rho_B^{(u)}(\xi) + \phi_r(\xi) + \mathcal{L}\xi \tag{4.19b}$$

$$\phi_r(\xi) = \frac{\mathcal{L}}{(k\ell_u)^2-1}\left[\left(\xi + \frac{2k\ell_u^2}{(k\ell_u)^2-1}\right)e^{-k\xi} - \frac{2k\ell_u^2}{(k\ell_u)^2-1}e^{-\xi/\ell_u}\right] \tag{4.19c}$$

Also here we notice that

$$\lim_{\xi\to\infty} \rho_B^{(u)}(\xi) \neq 0 \tag{4.20}$$

But in this case $\phi_r(0) = 0$ and therefore the function is continuous in the origin. In Fig.(4.1b) the comparison with the simulation.

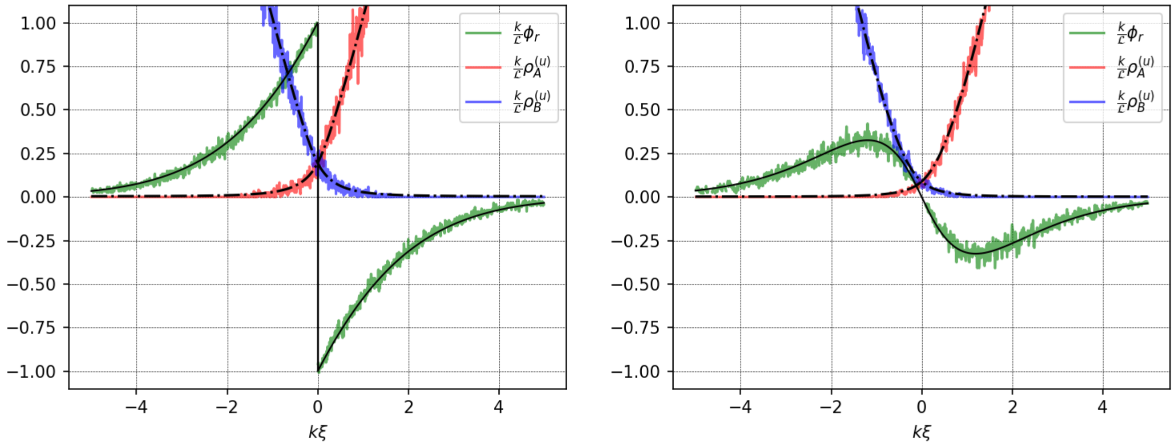
Solution for $k\ell_u = 1$

$$\rho_B^{(u)} = \frac{3\mathcal{L}}{8k\cdot\mathcal{I}_1(2)}\cdot\mathcal{I}_0\left(2e^{-\frac{k\xi}{2}}\right) \tag{4.21a}$$

$$\rho_A^{(u)} = \rho_B^{(u)} + \phi_r + \mathcal{L}\xi \tag{4.21b}$$

$$\phi_r = -\frac{\mathcal{L}}{4}(k\xi^2+\xi)e^{-k\xi} \tag{4.21c}$$

The resonant case has to be treated separately from the mathematical point of view, but is in fact a "continuous limit" and it shares the same shape as the case $k\ell_u = 1$ also in this case.



(A) $\ell_r = 0, k\ell_u = 0.375$ (B) $\ell_r = \ell_u, k\ell_u = 0.375$

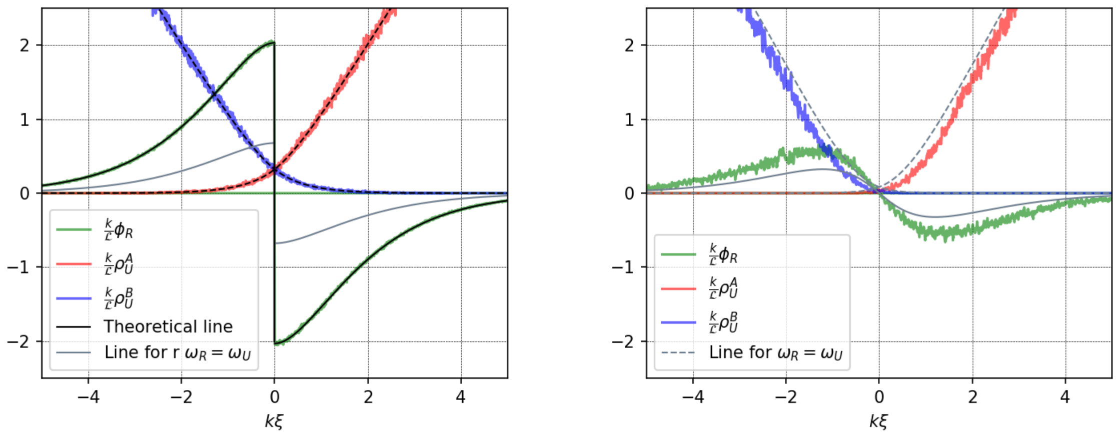
FIGURE 4.1: *Stationary solution and comparison to the simulation for* $\Gamma(y) = e^{-k|y|}$: rescaled stationary order books as function of rescaled price. The solid black lines indicate the theoretical rescaled revealed order density $\frac{k}{\mathcal{L}}\phi_r$ while dashed black lines signify the theoretical unrevealed order densities $\frac{k}{\mathcal{L}}\rho^{(u)}_{A/B}$. The results of the numerical simulation are plotted with color lines on top of the analytical curves.

To conclude this section, we observe that the shape of the solution looks very similar to the one deeply discussed in the main body of the work, leading us to to think that even if we chose Γ decreasing on both sides but with two different speeds, the qualitative shape of the solution will remain still the same. In some sense, choosing $\Gamma(y \leq 1) = 1$ is a particular case of such condition and therefore we studied analytically the two extreme regimes in which the decay length is the same on both sides or infinitely slow on one. Together with the numerical study of the revealing at the best quote, all these results allow us to claim that all the approximations we made on the choice of Γ to keep the analytical tractability do not affect largely the statistical properties inspected. Also, choosing Γ to decay exponentially allowed us to get good results in terms of fitting to the real data, so, even if considering a power law decay could still be an interesting aspect to deepen, we don't expect significant improvements in this direction.

4.1.2 Different revealing and unrevealing currents

We mentioned in the most general framework how the revealing and unrevealing currents have in principle two different intensities $\omega_r \neq \omega_u$. We chose however to consider them equal again for analytical tractability. We claimed that considering different intensities could

change the analytical expression of the books, without essentially change its shape.
To study the effect of having $\omega_\mathrm{r} \neq \omega_\mathrm{u}$ we performed a numerical simulation whose result is reported in Fig.(4.2).



(A) (B)

FIGURE 4.2: *Stationary solution and comparison to the simulation for $\omega_r \neq \omega_u$:* rescaled stationary order books as function of rescaled price for $\omega_\mathrm{r} = 3\omega_\mathrm{u}$. The solid black lines indicate the theoretical rescaled revealed order density $\frac{k}{\mathcal{L}}\phi_\mathrm{r}$ while dashed black lines signify the theoretical unrevealed order densities $\frac{k}{\mathcal{L}}\rho^{(\mathrm{u})}_{\mathrm{A/B}}$. The gray line indicates the shape of the order book for $\omega_\mathrm{r} = \omega_\mathrm{u}$. The results of the numerical simulation are plotted with color lines on top of the analytical curves.

What we see is that the effect of two different currents consists in increasing or decreasing the total liquidity, without changing drastically the shape of the book. Under this perspective we claim that all the results we obtained in the specific case can be extended in a straightforward manner to this more general case. In particular, we expect to observe some eventually more complicated stability condition and to confirm the square root impact.
It is to be noted that an analytical solution is easy to be found in the case $D_\mathrm{r} = 0$ and the difference with the earlier solution effectively reduces just to a multiplicative factor for what concerns the real order book, while the unrevealed ones are left unchanged. By denoting with $\ell_\mathrm{u}^{(\mathrm{r/u})} = \sqrt{\frac{D_\mathrm{u}}{\omega_\mathrm{r/u}}}$ we obtain:

$$\phi_\mathrm{r}(\xi) = \frac{\omega_\mathrm{r}}{\omega_\mathrm{u}} \left[\frac{\mathcal{L}\ell_\mathrm{u}^{(\mathrm{r})}}{2} e^{-\xi/\ell_\mathrm{u}^{(\mathrm{r})}} - \frac{\mathcal{L}\xi e^{-k\xi}}{1 - e^{-k\xi}} \right] \tag{4.22}$$

4.2 Possible future research

The model of Donier *et.al.* [22] on which we built ours upon definitely represents a good starting point for a zero intelligence model but is, however, a "first order approximation" for many aspects. We interpreted all the building blocks of our model in a *mean field* way. We are however aware that the market micro-structure is more complicated than that. In particular, Benzaquen [6, 5] described two extensions to the original model:

- *Multi time scale liquidity*: we have to think that the traders on the market can have substantially different time scales of operation. We mentioned at the beginning that the market makers are HFT and hence move inside the market at a much higher speed then the other traders. In particular, if we loosely define as $\tau \sim P(\tau)$ the reaction time of traders, then it definitely makes sense to think of $P(\tau)$ as a meanless fat tail distribution. In [6] it was considered a continuum of books ϕ_ν with $P(1/\nu)$ [1] a fat tailed distribution.
- *Time fractional diffusion* : in the same spirit, we have to remember the diffusion is an effective description of a much more complicated process. Traders will make a move on average every time step that will be different for each of them, again distributed according to a fat tail, leading to a fractional diffusion model.

As we saw already, Donier model was unable to solve the diffusivity puzzle having a memory kernel that was decaying too fast (see Eqs.(1.5, 1.6, 2.20)) and our model shares with it the same behavior. Introducing a fat tail distribution on time parameters as just explained allows to give more memory to the system and eventually solve the diffusivity puzzle, both with the multi time scale liquidity and with the time fractional diffusion. In this respect it hence looks quite straightforward that adding the latter will be enough to solve the diffusivity puzzle also inside our model. We will give a brief discussion on that at the end of this section.

The solution of the diffusivity puzzle in the two mentioned ways provides a relation between the exponents of the probability distributions that are in fact not easy to verify experimentally. Having more effects that are *individually* able to solve the diffusivity puzzle looks appealing for it will allow more freedom on the constraints needed to observe an actually diffusive price.

In this spirit we attempted to give a further solution of the diffusivity puzzle by assuming different values of ω for each trader and $P(\omega^{-1})$ to be fat tailed. Since we don't have an explicit expression of the memory kernel describing the price formation, this study is completely numerical.

[1] Recall that the parameter ν was defined in Donier model but we dropped it when we set in the linear regime.

We extended the model by defining different pairs of revealed/unrevealed order books, each one labeled by its own ω and occupied with a probability $P(\omega)$. In particular, let $\tau = \omega^{-1}$, then

$$P_\tau(y) \propto \frac{1}{y^{1+\alpha}} \to P_\omega(y) \propto y^{\alpha-1}, \quad 0 \leq \alpha \leq 1 \tag{4.23}$$

Each pair of book is coupled to the others only via the reaction term, *i.e.* we don't allow the traders to change their value of ω during the simulation.
In order to study the problem we had to see how this model reacted to a correlated market order flow:

$$\mathbb{E}[m_t m_{t'}] \sim \frac{1}{\sqrt{t-t'}} \tag{4.24}$$

To generate such order flow we defined its correlation matrix. In particular let $\vec{\mathcal{G}}$ be a vector of independent gaussian random numbers with zero mean and unitary variance and $\vec{m}$ the vector of correlated market orders. We write it as $\vec{m} = A\vec{\mathcal{G}}$, then we obtain:

$$\Sigma_{ij} := \mathbb{E}[m_i m_j] = \mathbb{E}[A_{ik}\mathcal{G}_k A_{j\ell}\mathcal{G}_\ell] = A_{ik}A_{j\ell}\delta_{k\ell} = A_{ik}A^T_{kj} = (AA^T)_{ij} \tag{4.25}$$

So we can define the matrix A as the Choleski decomposition of the covariance matrix. To write it we use the following relation valid for a fractional brownian motion $B_H(t)$ [10]:

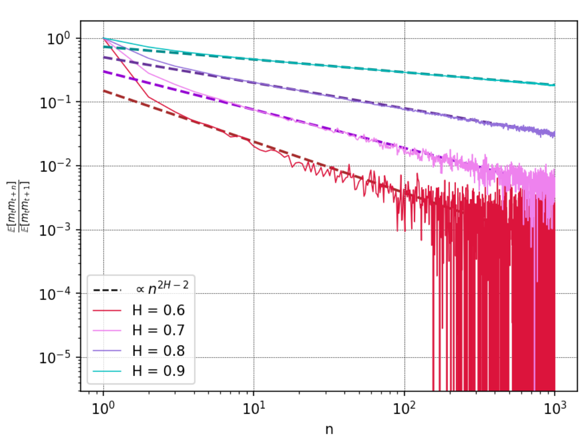
$$\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right) \tag{4.26}$$

We are looking for the covariance matrix of the increments:

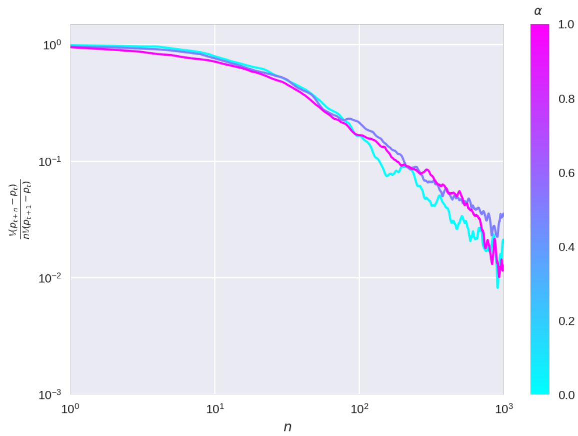
$$\begin{aligned}\mathbb{E}[(B_H(t+\tau) - B_H(t))(B_H(s+\tau - B_H(s)))] &= \\ &= \begin{cases} (2H-1)H\tau^2|t-s|^{2H-2} + o(\tau^3), & for \quad t \neq s \\ \tau^2 + o(\tau^3), & for \quad t = s \end{cases}\end{aligned} \tag{4.27}$$

Then, by performing a time discretization such that $t_{i+1} = t + \tau$ with $\tau \to 0$, we obtain the matrix $\Sigma = AA^T$, from which we obtain $\vec{m}$. Note that to have the correlation measured experimentally we should choose $H = \frac{3}{4}$. In Fig.(4.3a) we test the procedure just described to create correlated order flows and in Fig.(4.3b) we show the rescaled signature plot of the trade price obtained from the extension to our model. What we observe is that changing the value of α doesn't change the signature plot that is still mean reverting and not flat[2]. We hence conclude that the proposed modification doesn't represent an alternative solution to

[2]We call *signature* the following quantity: $\frac{\mathbb{V}(p_{t+n} - p_t)}{n}$



(A)



(B)

FIGURE 4.3: *Numerical simulation for the diffusivity puzzle*: (A) check of the correlation of the market order flow for different Hurst coefficients. The dotted line represents the theoretical behavior, the wiggly colored curves are the result of the renormalized random correlated orders as described above. (B) Renormalized signature plot of the trade price for different values of α.

the diffusivity puzzle. On one hand this was an interesting and simple modification to inspect because it represented an alternative way to give more memory to the system. On the other hand we can also have an intuition of the reason why it doesn't work: we noticed that the role of ω is determinant in the definition of the total available liquidity, but that it doesn't affect directly the trade price dynamics as it does the diffusive term instead.

To conclude, we can still claim that describing a continuum of books labeled by different values of ω makes sense, but it should be integrated with a more complete multi time scale framework. In fact, in our simulation we assumed that HFT and LFT (high and low ω respectively) have the same diffusion constants and revealing probability. Based on our observation on the MM behavior, we can propose instead a different scenario. Being HFT we suggest a large value of ω, but we also know they will post their orders only very close to the mid price, implying a large value of k and finally we expect a very low value of $D_r \to 0$ for them since MM provide liquidity and don't diffuse, never posting market orders. We remind that the presence of the MM is subject to the market orders unbalance, letting us to think of introducing a fundamental feedback term, as well.
From this observations on the MM behavior, we then suggest as a possible improvement of the present model a multi time scale version of it in which it is studied the relation between diffusion (in the revealed and unrevealed), revealing length and reaction time for each trader.

Conclusions

In the present work we addressed the problem of the modeling of the limit order book, as one of the main tools used in order to describe the financial market. We gave some basic definition about its structure and described some important statistical observation of the behavior of the traders.
Consequently we decided to approach the problem through a zero intelligence model, and exploiting the concept of latent liquidity, presenting the main results achieved by the existing models. We then proposed a mechanism to describe how the latent liquidity reveals into the real order book finding a number of interesting results:

- The possibility to fit our model to real data, whose behavior is in good agreement with our initial intuition and the prediction of a shape of the book consistent with the one observed
- The theoretical prediction of square root impact
- The identification of some stability conditions of the market and the possibility to address a financial crisis to a liquidity dry out of completely endogenous origin.

Finally we argued how some important assumption we made could be relaxed, making the results more general in a very straightforward way, suggesting how the present model could be linked to the already existing work in order to enclose some effects completely neglected by it but observable in the real limit order book.

Bibliography

[1] Robert Almgren et al. “Direct estimation of equity market impact”. In: *Risk* 18.7 (2005), p. 5862.

[2] Philip W Anderson et al. “More is different”. In: *Science* 177.4047 (1972), pp. 393–396.

[3] Emmanuel Bacry et al. “Market impacts and the life cycle of investors orders”. In: *Market Microstructure and Liquidity* 1.02 (2015), p. 1550009.

[4] Árpád Baricz. *Generalized Bessel functions of the first kind*. Springer, 2008.

[5] Michael Benzaquen and Jean-Philippe Bouchaud. “A fractional reaction–diffusion description of supply and demand”. In: *The European Physical Journal B* 91.2 (2018), p. 23.

[6] Michael Benzaquen and Jean-Philippe Bouchaud. “Market impact with multi-timescale liquidity”. In: (2017).

[7] Michael Benzaquen, Jonathan Donier, and Jean-Philippe Bouchaud. “Unravelling the trading invariance hypothesis”. In: *Market Microstructure and Liquidity* 2.03n04 (2016), p. 1650009.

[8] Michael Benzaquen et al. “Dissecting cross-impact on stock markets: An empirical analysis”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.2 (2017), p. 023406.

[9] Nataliya Bershova and Dmitry Rakhlin. “The non-linear market impact of large trades: Evidence from buy-side order flow”. In: *Quantitative finance* 13.11 (2013), pp. 1759–1778.

[10] Francesca Biagini et al. *Stochastic calculus for fractional Brownian motion and applications*. Springer Science & Business Media, 2008.

[11] Julius Bonart and Martin D Gould. “Latency and liquidity provision in a limit order book”. In: *Quantitative Finance* 17.10 (2017), pp. 1601–1616.

[12] Julius Bonart and Fabrizio Lillo. “A continuous and efficient fundamental price on the discrete order book grid”. In: *arXiv preprint arXiv:1608.00756* (2016).

[13] Richard M Bookstaber and Mark E Paddrik. “An agent-based model for crisis liquidity dynamics”. In: (2015).

[14] Jean-Philippe Bouchaud. “The subtle nature of financial random walks”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 15.2 (2005), p. 026104.

[15] Jean-Philippe Bouchaud, J Doyne Farmer, and Fabrizio Lillo. “How markets slowly digest changes in supply and demand”. In: *arXiv preprint arXiv:0809.0822* (2008).

[16] Jean-Philippe Bouchaud, Julien Kockelkoren, and Marc Potters. "Random walks, liquidity molasses and critical response in financial markets". In: *Quantitative finance* 6.02 (2006), pp. 115–123.

[17] Jean-Philippe Bouchaud and Marc Potters. *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge university press, 2003.

[18] Jean-Philippe Bouchaud et al. "Fluctuations and response in financial markets: the subtle nature of 'random'price changes". In: *Quantitative finance* 4.2 (2004), pp. 176–190.

[19] X. Brokmann et al. "Slow decay of impact in equity markets". In: *Market Microstructure and Liquidity* 1.2 (2015), p. 1550007.

[20] Frédéric Bucci et al. "Co-impact: Crowding effects in institutional trading activity". In: *arXiv preprint arXiv:1804.09565* (2018).

[21] Francesco Corradi, Andrea Zaccaria, and Luciano Pietronero. "Liquidity crises on different time scales". In: *Physical Review E* 92.6 (2015), p. 062802.

[22] Jonathan Donier et al. "A fully consistent, minimal model for non-linear market impact". In: *Quantitative finance* 15.7 (2015), pp. 1109–1121.

[23] Robert Engle, Robert Ferstenberg, and Jeffrey Russell. "Measuring and modeling execution cost and risk". In: (2006).

[24] Thierry Foucault, Ohad Kadan, and Eugene Kandel. "Limit order book as a market for liquidity". In: *The review of financial studies* 18.4 (2005), pp. 1171–1217.

[25] Roman Gayduk and Sergey Nadtochiy. "Liquidity effects of trading frequency". In: *Mathematical Finance* (2017).

[26] Christian Grossmann, Hans-Görg Roos, and Martin Stynes. *Numerical treatment of partial differential equations*. Vol. 154. Springer, 2007.

[27] Olivier Guéant. "Optimal market making". In: *Applied Mathematical Finance* 24.2 (2017), pp. 112–154.

[28] Joel Hasbrouck. *Empirical market microstructure: The institutions, economics, and econometrics of securities trading*. Oxford University Press, 2007.

[29] Armand Joulin et al. "Stock price jumps: news and volume play a minor role". In: *arXiv preprint arXiv:0803.1769* (2008).

[30] Andrei Kirilenko et al. "The flash crash: The impact of high frequency trading on an electronic market". In: *Available at SSRN* 1686004 (2011).

[31] Kęstutis Kubilius, Yuliya Mishura, and Kostiantyn Ralchenko. *Parameter Estimation in Fractional Diffusion Models*. 2018.

[32] Albert S Kyle. "Continuous auctions and insider trading". In: *Econometrica: Journal of the Econometric Society* (1985), pp. 1315–1335.

[33] Albert S Kyle and Anna A Obizhaeva. "Market microstructure invariance: Empirical hypotheses". In: *Econometrica* 84.4 (2016), pp. 1345–1404.

[34] Aimé Lachapelle et al. "Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis". In: *Mathematics and Financial Economics* 10.3 (2016), pp. 223–262.

[35] Charles-Albert Lehalle. "Market microstructure knowledge needed for controlling an intra-day trading process". In: *arXiv preprint arXiv:1302.4592* (2013).

[36] Leonard Lewin. "Polylogarithms and associated functions". In: *North-Holland* (1981).

[37] Fabrizio Lillo and J Doyne Farmer. "The long memory of the efficient market". In: *Studies in nonlinear dynamics & econometrics* 8.3 (2004).

[38] Xiaofei Lu and Frédéric Abergel. "Limit order book modelling with high dimensional Hawkes processes". In: (2017).

[39] Iacopo Mastromatteo, Bence Toth, and Jean-Philippe Bouchaud. "Agent-based models for latent liquidity and concave price impact". In: *Physical Review E* 89.4 (2014), p. 042805.

[40] Iacopo Mastromatteo et al. "Trading Lightly: Cross-Impact and Optimal Portfolio Execution". In: (2017).

[41] Michael Benzaquen Jean-Philippe Bouchaud Luis Carlos García del Molino Iacopo Mastromatteo. "The Multivariate Kyle model: More is different". In: *arXiv:1806.07791* (2018).

[42] Efstathios Panayi et al. "Designating market maker behaviour in Limit Order Book markets". In: *Econometrics and Statistics* (2016).

[43] Michael Parkinson. "The extreme value method for estimating the variance of the rate of return". In: *Journal of business* (1980), pp. 61–65.

[44] Adam Ponzi, Fabrizio Lillo, and Rosario N Mantegna. "Market reaction to a bid-ask spread change: A power-law relaxation dynamics". In: *Physical Review E* 80.1 (2009), p. 016112.

[45] Leonard CG Rogers, Stephen E Satchell, and Y Yoon. "Estimating the volatility of stock prices: a comparison of methods that use high and low prices". In: *Applied Financial Economics* 4.3 (1994), pp. 241–247.

[46] Robert J Shiller. *Do stock prices move too much to be justified by subsequent changes in dividends?* 1980.

[47] D E Taranto, Giacomo Bormetti, and Fabrizio Lillo. "The adaptive nature of liquidity taking in limit order books". In: *Journal of Statistical Mechanics: Theory and Experiment* 2014.6 (2014), P06002.

[48] N Torre and M Ferrari. "The market impact model". In: *Horizons, The Barra Newsletter* 165 (1998).

[49] Bence Tóth, János Kertész, and J Doyne Farmer. "Studies of the limit order book around large price changes". In: *The European Physical Journal B* 71.4 (2009), p. 499.

[50] Bence Tóth et al. "Anomalous price impact and the critical nature of liquidity in financial markets". In: *Physical Review X* 1.2 (2011), p. 021006.

[51] Bence Toth et al. "How does the market react to your order flow?" In: *Quantitative Finance* 12.7 (2012), pp. 1015–1024.

[52] Bence Toth et al. "Why is equity order flow so persistent?" In: *Journal of Economic Dynamics and Control* 51 (2015), pp. 218–239.

[53] Paul Wilmott. *Paul Wilmott introduces quantitative finance*. John Wiley & Sons, 2007.

[54] Hai-Chuan Xu et al. "Limit-order book resiliency after effective market orders: spread, depth and intensity". In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.7 (2017), p. 073404.

[55] Elia Zarinelli et al. "Beyond the square root: Evidence for logarithmic dependence of market impact on size and participation rate". In: *Market Microstructure and Liquidity* 1.02 (2015), p. 1550004.