

A million metaorder analysis of market impact on the Bitcoin

J. Donier¹ and J. Bonart²

¹*Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie (Paris 6).*

²*CFM-Imperial Institute of Quantitative Finance, Department of Mathematics,
Imperial College, 180 Queen's Gate, London SW7 2RH*

(Dated: September 22, 2015)

Squate-root law: Lecture 7

We present a thorough empirical analysis of market impact on the Bitcoin/USD exchange market using a complete dataset that allows us to reconstruct more than one million metaorders. We empirically confirm the “square-root law” for market impact, which holds on four decades in spite of the quasi-absence of statistical arbitrage and market making strategies. We show that the square-root impact holds during the whole trajectory of a metaorder and not only for the final execution price. We also attempt to decompose the order flow into an “informed” and “uninformed” component, the latter leading to an almost complete long-term decay of impact. This study sheds light on the hypotheses and predictions of several market impact models recently proposed in the literature and promotes heterogeneous agent models as promising candidates to explain price impact on the Bitcoin market – and, we believe, on other markets as well.

1. Introduction

Most financial markets have undergone rapid changes in the past 10 years. The implementation of limit order books, the electronization of trade exchanges, the rise of high-frequency trading, and the introduction of algorithmic and automated executions are the most emblematic features of the new financial economy. Both supporters (Brogaard et al. 2014, Hendershott 2011) and critics (Budish and Cramton 2013, Kirilenko et al. 2014) of this evolution hold the view that the recent speed revolution has had a profound impact on the functioning of financial markets. While this is obviously true from a technological and microstructural point of view, it is less clear from an economists’ perspective: In the end, a stock exchange is a market place where buyers and sellers meet and agree on a price; and it is not straightforward to assess the relevance of technological changes to this process.

It is however possible to directly compare financial markets which strongly differ in their respective degrees of technological sophistication and latency. The recent development of crypto-currencies such as the Bitcoin, traded against classical currencies on automated exchange platforms, provides us with a unique benchmark for a comparison with high-speed markets. Indeed, the Bitcoin exchange is an example of a proto-typical financial market, maintained – at the time of the study – in a rudimentary competitive state by high trading fees¹ which inhibit the development of substantial market making or arbitrage.

An important aspect of a financial market is how it absorbs the new information conveyed by a trade or a sequence of trades into the market price (Kyle 1985). Such incorporation of new information may not be instant due to market frictions (Beja and Goldman 1980); nor is the submission of the trader’s order, since order splitting and inventory considerations create a serial dependence of trades (Farmer and Lillo 2004, Toth et al. 2015). During the execution of a large sequential order (meta-order), spread over a certain time period, the difference between the impacted price and the initial price can be quantified by measuring the market impact of the meta-order.

In this paper we undertake a thorough empirical analysis of market impact on the Bitcoin/USD exchange market. Previous studies (Almgren et al. 2005, Bershova and Rakhlin 2013, Gomes and Waelbroeck 2015, Mastromatteo et al. 2014a, Moro et al. 2009, Toth et al. 2011a) found that average market impact of a meta-order of total volume Q approximately follows the “square-root law”,

$$I(Q) \approx \pm Y \sigma \left(\frac{Q}{V_D} \right)^\delta, \quad (1)$$

where $I(Q)$ quantifies the average difference between the impacted price and the initial price (with the positive sign corresponding to buy orders and vice versa), V_D and σ the daily traded volume and daily volatility of the stock. The exponent δ has been consistently found to be approximately 1/2. When the meta-order terminates so the pressure it was exerting on the price stops, the price is usually observed to revert (partially or totally) towards the initial (unimpacted) price (Brokmann et al. 2014, Gomes and Waelbroeck 2015). With our high-quality dataset we shall

¹ 0.6% fees per transaction for the vast majority of the main exchange MtGox’s users

confirm that above impact law holds for the Bitcoin/USD exchange market. Therefore, the rise of algorithmic trading may not have had as a profound effect on the functioning of markets as is often advocated. On the Bitcoin, both the response of the market to trades and the serial dependence of orders are similar to what is observed on mature liquid financial markets, in such a way that market impact can be specified by the same empirical law.

This observation is central: The precise mechanism which is behind the peculiar square-root law appears to be universal across markets which significantly differ with respect to their trade characteristics (latency, daily traded volume, volatility), microstructural parameters (tick or lot sizes) and fee structure (i.e. the high fees on the Bitcoin). This latter point is especially important. Statistical arbitrage is not profitable on the Bitcoin on scales below the fee level of 60 bps (120 bps for a round-trip!) and price efficiency is therefore not ensured in this region. The resulting large bid-ask spread frequently exceeds the peak impact of typical meta-orders; concepts such as price efficiency and arbitrage are hence not suitable to explain market impact on the Bitcoin exchange.

Therefore, we conclude that competitive equilibria between different agents, for instance amongst market makers or between informed liquidity takers and uninformed liquidity providers (Glosten and Milgrom 1985), should not be used as the fundamental starting point for the theoretical modelling of market impact on the Bitcoin. Nor should one use the martingale conditions for the market price, as the notion of “price” is not precisely defined under the scale of 60 bps (\sim spread/fees). In fact, as we shall see in the following, impact describes *how trades dig into the opposite (supply/demand) side* – before some post-trade mean-reversion occurs on this side – rather than how the *market price* itself is affected. Although these two definitions are equivalent when the spread is tight, they do not coincide in the case of the Bitcoin. The fact that very similar impact laws are found on mature financial markets suggests that this observation holds for these markets, as well.

1.1. Comparison with related literature

The square-root impact formula quantifies how market prices are affected by trades. This has implications for *market stability*: Price impact may lead to unexpected price swings (Lehalle and Lasnier 2012) and stock crashes (Kyle and Obizhaeva 2012) as well as to well-known phenomena such as stock pinning (Avellaneda et al. 2012). Second, the control of *execution costs*, i.e. market impact, is of great practical interest to financial institutions. Almgren and Chriss (2001), for instance, determined optimal execution strategies based on the assumption that impact is *linear* (i.e. that $\delta = 1$). This is the simplest theoretical setting that excludes round-trip execution strategies (zero terminal inventory) with negative execution costs (profit making). Since then, a huge effort has been made to model non-linear market impact, motivated by a flurry of empirical studies. These approaches can be broadly split into three categories:

- Concave market impact of metaorders can be reproduced by *propagator models*, in which each trade is assumed to have a transient impact which decreases according to some time-dependent kernel. Summing up the impacts of individual trades leads to the impact function of the metaorder which is then found to be concave. The possible absence of dynamical arbitrage in these settings (Gatheral 2010) allows one to solve convex optimization problems for optimal execution and find optimal liquidation strategies (Alfonsi and Schied 2013) within this framework. While these models yield fairly realistic results and are analytically tractable, they are however purely phenomenological and do not provide a mechanism to explain impact.
- *Equilibrium models* assume a competitive equilibrium between liquidity providers and takers. They can be regarded as an extension of the original argument in Glosten and Milgrom (1985) from a single market order to a sequence of trades. Equilibrium models typically use two constraints to fix the bid and the ask during the metaorder execution, usually using the martingale condition and a subtle fair pricing argument (Donier 2012, Farmer et al. 2013), which states that the transient impact of a metaorder anticipates its permanent impact on the price so that neither the informed trader nor the market maker should expect profits for any metaorder (on average). In the same vein, Jaisson (2015) shows how non-linearities can emerge from anticipations when the order flow is correlated by a Hawkes process, even though impact is linear *ex post*. These arguments, together with strong correlations in the order flow finally leads to an – asymptotically – concave metaorder impact. However, as such mechanisms are rather unlikely to be in force on the Bitcoin market, they seem not appropriate to explain the seemingly universal shape of impact.
- The third class of models, initiated in a different context in Bak et al. (1997), are *statistical models* of supply and demand (Donier et al. 2015, Mastromatteo et al. 2014a,b, Toth et al. 2011a), that may also be seen as *heterogeneous agents models* (Donier and Bouchaud 2015a). The dynamics they give to the supply and demand that underlie the order book is such that both generically increase quadratically with distance from the mid-price, in turn leading to an exact and universal square-root impact at all scales.

Empirical studies have been mostly concerned with the *peak impact* of meta-orders, i.e. the impact measured between the extremal points of the execution path. Further studies have shown that impact is to a large part transient: after execution the price falls from its peak to some intermediate level, that some studies (Bershova and Rakhlin 2013, Gomes and Waelbroeck 2015, Moro et al. 2009) argue to be close to 2/3 of the peak impact, in agreement with equilibrium models (Farmer et al. 2013). In Brokmann et al. (2014) the authors find that this high permanent level may be due to the correlation between the trader’s execution decision and the residual order flow² and to the price signal that triggered the decision to trade: By taking into account these effects they argue that the “bare” permanent impact (or mechanical permanent impact) is much lower and possibly even zero. Finally, Gomes and Waelbroeck (2015) conduct a separate impact study of informed trades and cash-flow (uninformed) trades, to find that the latter have no permanent effect on the price (although the transient impact are similar for both types of trades). Hitherto, these results are not well accounted for nor understood.

Amongst the empirical studies of market impact in the past literature, the vast majority has relied on partial datasets. Usually, market participants have only access to their own proprietary data (Bershova and Rakhlin 2013, Brokmann et al. 2014, Toth et al. 2011a) which leads to an unavoidable conditioning to their trading strategies (even though in some cases many different strategies are collated together so part of the conditioning may average out). Two notable exceptions do however exist: In Moro et al. (2009) hidden metaorders are directly inferred from brokerage codes, while Zarinelli et al. (2014) have unprecedented access to the start times, end times and volumes of a huge amount of metaorders stemming from Ancerno’s clients³.

In this paper, we use a dataset which allows on the contrary to identify *each* trade with a *unique* trader, thereby leading to a *complete picture of the market*. Our dataset is large enough to study market impact as a function of volume, participation rate and even as a function of the behaviour of the residual market. This allows us to retrieve pseudo-random metaorders, i.e. metaorders that are uncorrelated from the residual order flow, either because they do not convey any information or because the information that triggered them is not shared by the residual market.

1.2. Brief outline

After a presentation of the dataset (Sec. 2), we introduce the main definitions and methodology (Sec. 3) with a focus on the methods we use to identify distinct meta-orders. The price impact of metaorders is introduced in Sec. 4. The first major result of our study is that, despite its prototypical micro-structure, the square root law holds on the Bitcoin. Moreover, in Sec. 5 we show that not only the peak impact of metaorders, but rather the *whole impact trajectory* follows a square-root. We show unprecedented pictures of the mid-price, bid and ask trajectories during and after impact and discuss permanent and transient market impact as a function of the execution speed and its dependence on the residual order flow.

We are able to differentiate the meta-orders with respect to their correlation with the residual order flow. The permanent impact is identified as the information content of meta-orders: Pseudo-random metaorders (i.e. metaorders uncorrelated with the residual order flow) have little permanent impact, in agreement with previous studies Brokmann et al. (2014), Gomes and Waelbroeck (2015).

Finally, in Sec. 6 we summarize our main findings and in the last section we conclude and discuss the implications of our empirical findings for the most common impact models so far proposed in the literature.

2. Bitcoin market at a glance and data

2.1. Bitcoin : a prototypical market

Bitcoin is a crypto-currency introduced by an anonymous programmer in 2008 (Nakamoto 2008), designed to allow exchanges of money without the need of a central authority (e.g. state or bank) to enforce trust. The money is issued progressively through a process called *mining* – by analogy to gold mining – to the people who give their computing power to help build the transactions ledger (the *blockchain*). As a currency, it has some intrinsic value

² After a buy (resp. sell) metaorder, the order flow is on average biased towards buy orders (resp. sell orders). This correlation with future order flow artificially inflates its measured impact.

³ This study reports a log impact as a better overall fit to metaorder impact, although they actually mention some caveats due to the low level of details of their dataset.

(even though still quite imprecise) and can be exchanged against usual currencies on organized markets. The Bitcoin market started in 2008 and literally exploded in 2013 with a peak market capitalization above 10B dollars, a daily number of transactions over 100k and a daily traded volume above 100M dollars at the end of 2013.

Similarly to financial markets, trading takes place on an order book, where Bitcoins can be exchanged against other currencies. Notably, the role of the order book is quite important since a large fraction of the liquidity is not hidden, but actually posted in the order book. A more quantitative analysis indeed shows that typically 30 – 40% of the volume traded during the day is already present in the order book in the morning. This is to be compared with a ratio below 1% on more traditional financial markets, say stocks (Wyart et al. 2008). The tick size, i.e. the minimal price increment between two consecutive prices at which it can be exchanged, is of USD 10^{-5} on MtGox so that the order book can be considered as a continuous price grid.

Bitcoin market microstructure is quite unique for several more fundamental reasons. First, because of the very high level of fees (compared to other markets) of 0.6% per transaction on MtGox, resulting in an average spread of $\approx 0.6\%$. This, to a large extent, hobbles high frequency arbitrage/market making strategies on such a market, who only take part in a few percent of the transactions: the Bitcoin market is essentially a market between end users. Second, all relevant information is concentrated on one asset and one exchange (MtGox) on the time span considered, with very little notion of a “fundamental price”. This is a very unique example of a single-asset economy, with little correlations with any other asset on the planet (for the time being).

2.2. Data

This study was realized using a database of all 13M trades that happened on MtGox Bitcoin-USD exchange between August 2011 and November 2013, in which traders are uniquely identified⁴. Thus, we have a complete description of the market as a whole, in contrast to most hedge funds’ proprietary data (reflecting their trading decision) which does not allow for a global analysis of trading decisions and market impact. As a consequence, we do not face the problem of the potentially strong conditioning of impact paths on the particular strategy of the agent. We rather have an overall insight on how agent’s decisions entangle. Since there is very little brokerage intermediation in this market, we have direct knowledge of the actual initiators of each trade (only as an anonymized numerical ID code). One can thus assume that with very good approximation *all metaorders can be fully identified*. This is a most valuable property for studying impact, since it allows for the explanatory variable – the metaorder – to be fully characterized. This data quality is probably very difficult to obtain on other financial markets, because of brokerage intermediation, multiplicity of venues, and multiplicity of correlated instruments on which a trade can be executed. In addition, our dataset is the largest so far with such precision, and with full knowledge of the trajectories for the 1M identified metaorders.

3. Definitions and methodology

3.1. Definitions

When a trader (she) wishes to bring an excess of supply or demand Q on the market, she is confronted to the question of how to achieve it in a reasonable way. If the volume she wants to sell or buy is small, then she will probably do it all at once if enough liquidity is available on the order book. However, if she wants to invest or sell back a quantity that exceeds the available offer or demand (think of large speculators or professional service providers), an instantaneous execution might destabilize the market and incur her larger costs than planned: she therefore has to split her large order into chunks, so that the imbalance can be slowly digested by the market (Bouchaud et al. 2009). We refer to this *total* quantity Q as a *metaorder*, denoting T its duration and $\mu := Q/T$ its execution speed, and study the quantities

- $\mathcal{I}_{\text{path}}(r, Q, \mu)$, defined as the impact on the price of the first $r\%$ of a metaorder of size Q and execution speed μ . By definition $r \in [0, 1]$ is the part of the volume already executed, and $\mathcal{I}_{\text{path}}(0, Q, \mu) = 0$ is the initial price

⁴ The data we use was provided to us by an anonymous source, and is available upon request under certain conditions. The public part of this data file has been checked to match otherwise known price and volume data at www.bitcoincharts.com, and the private part to match all private trading data that are to the knowledge of the authors. Apart from these consistency checks, no insurance is given that all data is accurate. However the high precision of our findings suggests that they are.

(gauged to zero);

- $\mathcal{I}(Q, \mu) := \mathcal{I}_{\text{path}}(1, Q, \mu)$ is the price at the end of the metaorder, referred to as the *peak impact* of the metaorder;
- $\mathcal{I}_{\text{exec}}(Q, \mu) := \int_0^1 \mathcal{I}_{\text{path}}(r, Q, \mu) dr$ is the average execution price;
- $\mathcal{I}^\infty(Q, \mu)$ is the average price long after the end of the metaorder and can be decomposed into a predictable part $\mathcal{I}_{\text{info}}^\infty(Q, \mu)$ and the response to the metaorder in question $\mathcal{I}_{\text{mec}}^\infty(Q, \mu)$. The latter is the most relevant quantity regarding price formation, since the former represents *alpha* biases that has *a priori* no link with the market's mechanical reaction to the order.

Note that all these quantities are implicitly averaged over all residual noise or variables. The effect of such metaorders on the price is qualitatively well-known. While the metaorder is executed, the pressure exerted on the price tends to make it rise so $\mathcal{I}_{\text{path}}(r, Q, \mu)$ is an increasing function of r . When the metaorder is completed the price reverts to the *permanent* level $\mathcal{I}^\infty(Q, \mu) \leq \mathcal{I}(Q, \mu)$. Previous empirical studies have found the *peak impact* to be approximately square root of the volume and the amplitude of the *post-impact decay* to be about 1/3 on average (so that $\mathcal{I}^\infty(Q, \mu) \approx \mathcal{I}_{\text{exec}}(Q, \mu) \approx \frac{2}{3}\mathcal{I}(Q, \mu)$ for square root impact, as required by equilibrium models since this ensures that the price paid for the execution is fair *ex-post*). More recent studies (Brokmann et al. 2014, Gomes and Waelbroeck 2015) have shown that when subtracting the *predictable part* $\mathcal{I}_{\text{info}}^\infty$ from the price, the reversion goes all the way back to zero – so after waiting a long enough time the *mechanical impact* $\mathcal{I}_{\text{mec}}^\infty(Q, \mu)$ of trades on the price is zero.

One can also define the *execution rate* μ_V , defined as the ratio between the volume of the metaorder Q and the total volume traded by the market during the same period, V_M ,

$$|Q| = \mu_V V_M. \quad (2)$$

We will thus study the peak impact $\mathcal{I}(Q, \mu)$, the average executed price $\mathcal{I}_{\text{exec}}(Q, \mu)$ and the permanent impact $\mathcal{I}^\infty(Q, \mu)$, in order to identify any dependence on T (or equivalently μ) that goes beyond the usual rule of thumb (1) which predicts impact to be independent of the execution speed. Throughout this study, price is taken to be the *traded price*. Since we only consider aggressive metaorders (cf. Sec. 3.3.2), this amounts to studying the ask for buy metaorders and the bid for sell metaorders.

3.2. Metaorder decomposition and properties

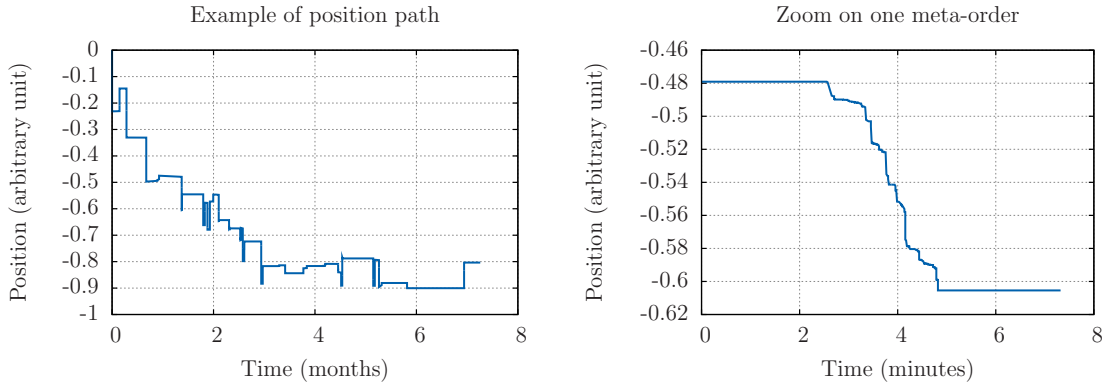


FIG. 1: Typical position vs time for one of the 1% most active traders in terms of volume. (*left*) position path during approximately 12 months. One can see that metaorders are clearly identifiable and alternate with long periods of inactivity. (*right*) Zoom on a 2 minute sell metaorder composed by a dozen trades (zoom 1 : 50000).

The first operation needed in order to study market impact is to spot the metaorders: due to the extreme irregularity and heterogeneity between the traders' typical position paths, usual time series decomposition methods (Toth et al. 2010) are not relevant here. In order to identify these large buy and sell metaorders for this particular data, in such a way that no conditioning in the start/end sequences of the metaorders is introduced (most intuitive techniques may create mean-reversion biases), we used the following: For each trader, we defined the start of a metaorder to

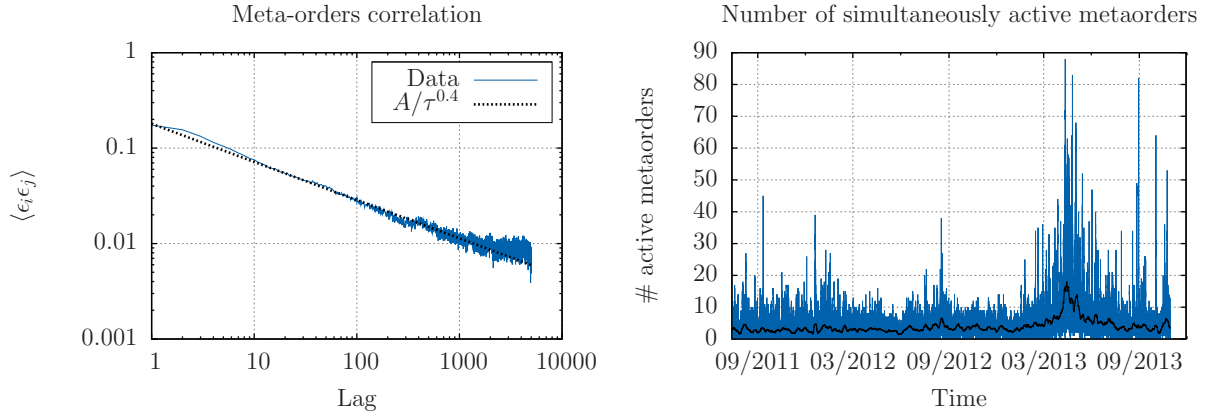


FIG. 2: (left) Autocorrelation function of the sign of *metaorders*, ordered by starting time. Like financial markets, not only trades but metaorders are strongly auto-correlated. (right) Number of simultaneously active metaorders vs time. The typical value is around 5, and we can clearly observe clustering.

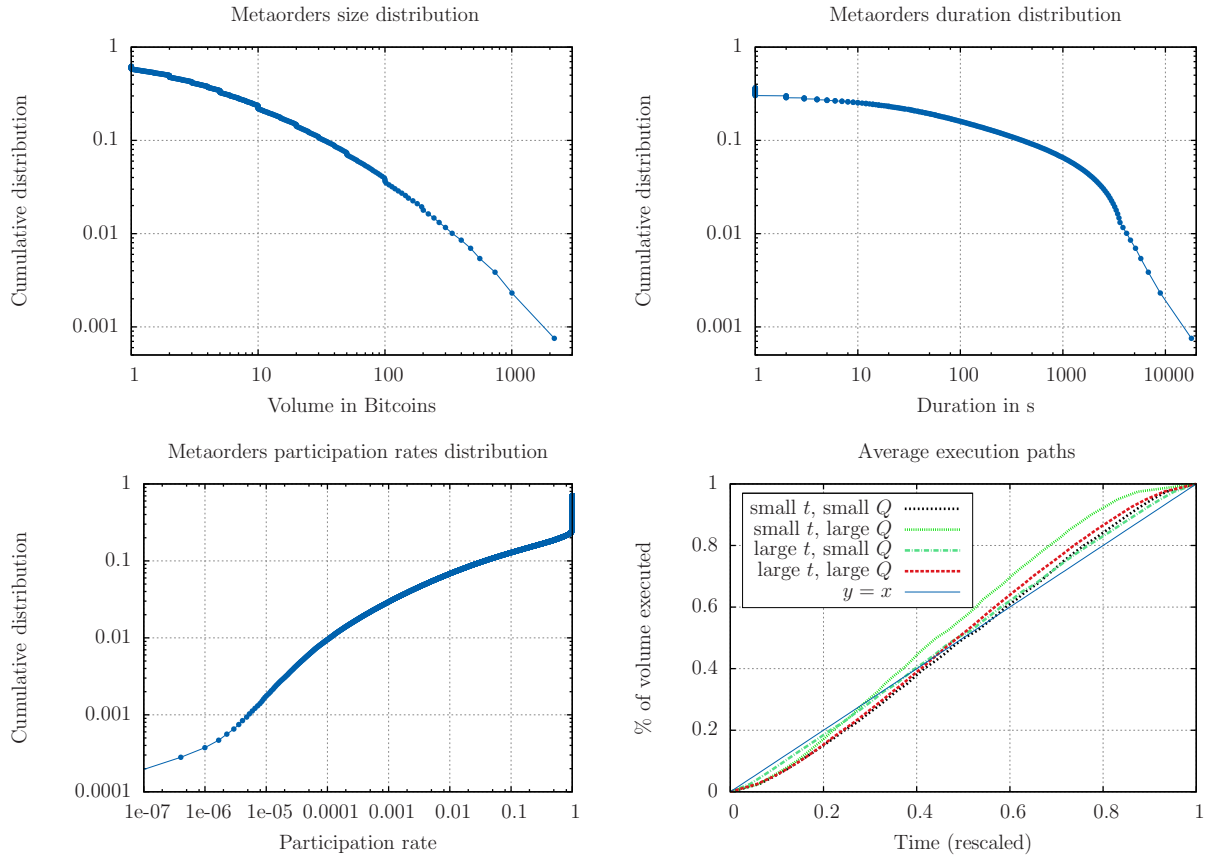


FIG. 3: (top and bottom left) Metaorder size, duration and participation rate distributions for the whole market, with no clear power-law fit for any. Note that the durations are much shorter than usual metaorder durations on financial markets. (bottom right) Percentage of volume executed vs time elapsed since the start of the metaorder, which appears to grow roughly linearly for all ranges of volume and duration (here the volume threshold between small and large has been fixed at 200BTC and the time threshold at 100s).

# of child trades	1	$2 \leq n \leq 4$	$5 \leq n \leq 9$	$10 \leq n$
% of metaorders	61%	29%	6.5%	3.5%

TABLE I: Number of child orders per metaorder.

coincide with the first aggressive⁵ order placed after a given period of inactivity⁶. We define the end point of the metaorder either as the last order before a (new) inactivity period, or as the point where the trader trades in the opposite direction. While the introduction of a time scale to define inactivity periods could seem to be arbitrary, it is the case in practice that the metaorders are so clearly distinguishable that the sensitivity to this time scale is minimal (Fig. 1): for one particular trader, metaorders are very clearly separated from each other. This way, from over 14 million trades we recover over 1 million metaorders of variable sizes/durations. A few percent of the trades are not assigned to a metaorder, corresponding to mean-reverting trades which by definition have a conditioning bias. Table 3.3.2 presents the repartition of metaorders in terms of number of child orders. As expected, more than half are one-shot trades whereas 10% percent are composed of more than 5 trades. However, as a result of the tick being small, more than 20% of the one-shot metaorders cross several price levels.

Fig. 2 shows the autocorrelation function of metaorder signs which is slowly decaying with a power-law exponent around 0.4 – which happens to be similar to the autocorrelation exponent of trades themselves – as well as the number of metaorders that are simultaneously active in the market at any point in time, which present clear clustering with a typical value around 5. The salient statistical characteristics of metaorders are presented in Fig. 3. It presents the metaorder sizes and durations distributions, which are crucial ingredients in many market impact and price models (Donier 2012, Farmer et al. 2013, Gabaix et al. 2003). Contrary to what is usually required in such models, none of these distributions are clear power-laws – and particularly not on *all* time scales⁷. This challenges such impact models, since in these models the impact function is closely related to metaorders distributions – and a power-law with exponent 3/2 is required to reproduce the square-root impact. This is a strong hint that the distribution of order sizes is not a fundamental input to explain the shape of market impact, since it is neither clearly a power-law nor universal. Most importantly, above figure shows that the average execution profile is linear in time. This property ensures that μ is well defined in the sense that it is on average constant during execution (one also needs to check that the executed volume in the market is constant, which is the case as shown e.g. by Fig. 4). This will allow us to properly compare points within impact trajectories, $\mathcal{I}_{\text{path}}(r, Q, \mu)$, with peak impacts $\mathcal{I}(rQ, \mu)$.

3.3. Conditioning checks

It has been shown in the previous section that average execution style is linear in time, so that we can really study the effect on the price of metaorders with constant execution speed on average⁸. In this section, we proceed to some other checks concerning the behaviour of the rest of the market while metaorders are being executed. Indeed, the usual argument against the square-root law assumes them to be conditioned to the rest of the market in such a way, that the apparent concavity of the marginal impact is only an artefact. In particular, anticipating Sec. 5, we look into the buy/sell market order flow while a metaorder is executed, to check whether dynamical effects exist in order flow that could result in concavity: for example, a higher correlation with market imbalance at the beginning of the metaorder would result in sharper price changes at the start of the trajectories⁹. Since we have the start and end times of every metaorders as well as their execution rates, we can compute at any point in time the number/volume of active buy/sell metaorders in the market. In Fig. 4 we compute these quantities during the execution of chosen metaorders (here all metaorders of duration about 2 minutes) and plot the average paths over all these metaorders to look for such dynamical effects. The data shows that such a synchronisation of metaorders is almost non-existent,

⁵ We only consider aggressive orders since they are a much better proxy for system perturbations. Limit orders in the contrary may have been posted long in advance without a specific intention nor view on the price: By ignoring them we therefore limit any adverse selection that might bias our study. Besides, the fact that the execution schedule is roughly linear strongly suggests that we do not bias the results by making this choice.

⁶ The relevant scale for this inactivity period has been empirically determined to be about one hour.

⁷ If however one estimates a tail exponent using the Hill estimator (Hill et al. 1975) for volumes greater than 10BTC, one finds a tail exponent of -2 for the probability density function.

⁸ as opposed e.g. to Zarinelli et al. (2014) in which they find strongly front-loaded executions.

⁹ This would imply some sort of synchronization between the agents, either exogenous (e.g. the agents react to the same news thus starting their metaorders at the same time, but some stopping before the others) or endogenous (e.g. arbitrageurs are able to detect metaorders and push the price up at its beginning to sell back later when the price has risen).

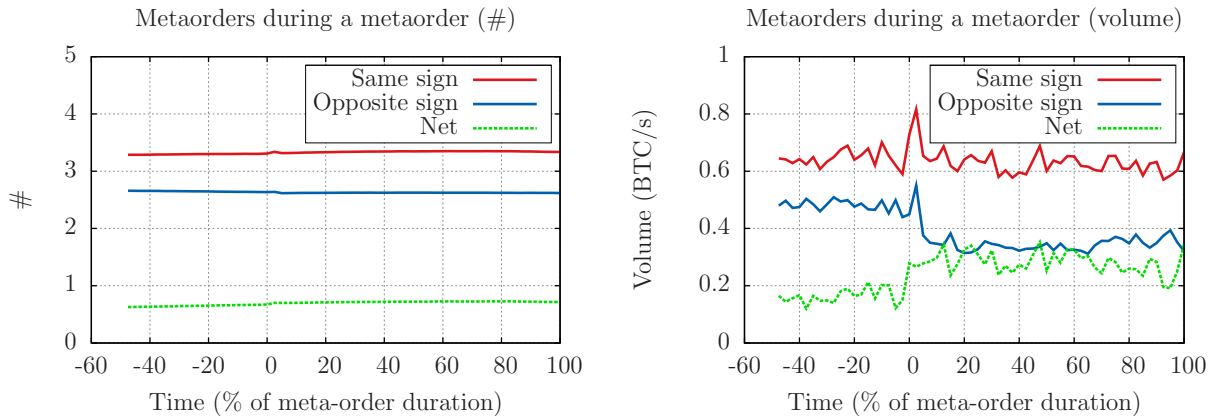


FIG. 4: Here we selected all metaorders with duration of approx. 2 minutes. (*left*) We plotted the number of active metaorders in the same direction than the metaorder considered after subtracting it (green), in the opposite direction (red) and the difference (blue). (*right*) The same plot, in terms of volume. On both plots, time is normalized so that the metaorders start at 0 and end at 100.

either in terms of number of metaorders or in terms of volume, so that concavity of impact is not generated by some kind of synchronization between the agents¹⁰. Since the study of impact paths in Sec. 5 will show that the square-root law holds for the whole *trajectories* and not only for the peak impact, one can assert that concavity is *true* in the sense that it does not stem from conditioning but is really the way the market absorbs an excess of supply or demand.

4. The square root impact law for the Bitcoin/USD exchange market

4.1. The square root law of peak impact for individual metaorders

For each of the 1M metaorders we identified above, we considered the impact defined as

$$\mathcal{I}(Q, \mu) = \mathcal{I}_{\text{path}}(r = 1, Q, \mu), \quad (3)$$

i.e. the difference between the first and the last executed price, quantifying the reaction of the market to the trader's order. Note that here impact is measured as the peak price (with the initial price gauged to zero). The result is shown in Fig. 5: In spite of the very special features of the Bitcoin market, a concave impact law (depicted as a straight line)

$$\langle \mathcal{I}(Q, \mu) \rangle_{\mu} \approx \tilde{Y} Q^{\delta}, \quad (4)$$

fits the data points very well from the smallest scales and over 4 decades with $\delta \approx 0.5$ and $\tilde{Y} \approx 4.5 \cdot 10^{-2}$. Normalizing by Bitcoin average volatility and daily volume gives a *Y-ratio* (as defined in Eq. 1) of $Y \approx 0.9$, close to the value reported on “mature” financial markets, e.g. futures or stocks (Brokmann et al. 2014, Toth et al. 2011a). For a more in-depth study of the *Y-ratio*, see Sec. 4.4.3 below. Thus, peak impact for the Bitcoin is consistent with the square-root law despite the prototypical nature of the Bitcoin market. This confirms the view expressed in Toth et al. (2011a) and Mastromatteo et al. (2014a) that market impact depends neither on microstructure (with e.g. arbitrage and high-frequency trading), nor on a clear metaorder distribution. As such, it directly challenges the explanation of concave market impact in terms of rational equilibrium theories.

¹⁰ One can note the interesting (unrelated) fact that market order volume in the opposite direction slightly decreases while a metaorder is being executed. Also, during a buy (resp. sell) metaorder the other metaorders present on the market tend to go in the same direction. This confirms the fact that metaorders are “informed” on average. A more in-depth discussion on this notion of information can be found in Donier and Bouchaud (2015a).

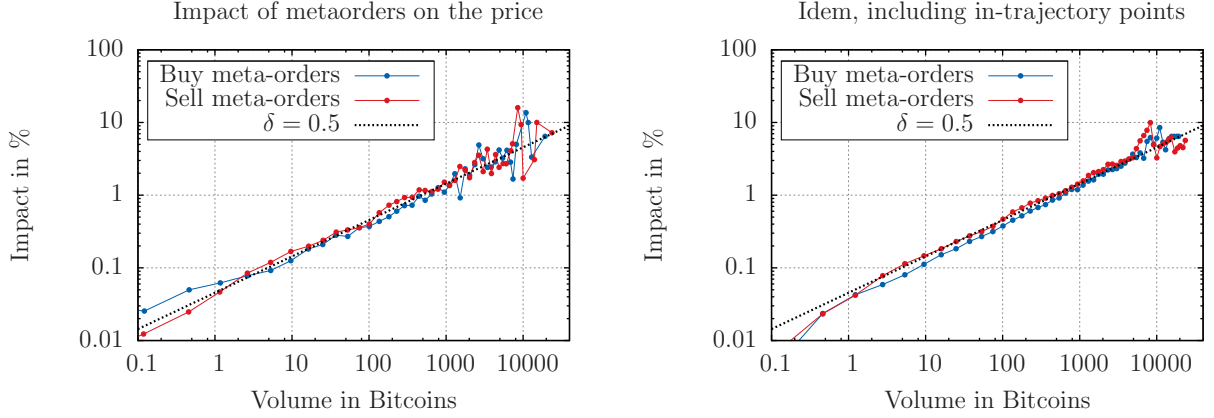


FIG. 5: Market impact $\langle \mathcal{I}(Q, \mu) \rangle_\mu$ (averaged over all execution rates μ), follows the same square root law as is observed by banks and hedge funds on financial markets (plots in log-log scale). Each point represents the average impact of all metaorders in a given range of volume. The impact exponent δ is found to be very close to 0.5, and the Y -ratio is around 0.9. One should emphasize that this power-law behaviour appears at the smallest scales and holds over 4 decades. (*left*) Only end points of metaorders (~ 1 M data). (*right*) 41 point per metaorder (every 2.5% quantile of volume), giving 27M data points. Part of these points being degenerate, one can assess the number of effective points around a few millions.

4.2. Square root impact trajectories

We now turn to the study of impact trajectories, i.e. the quantity $\mathcal{I}_{\text{path}}(r, Q, \mu)$ for given Q , μ and r varying from 0 to 1. Unless mentioned otherwise, we average over all other quantities in the following (like e.g. daily volatility, daily traded volume etc.). Fig. 6 shows the results, putting into light two facts of particular interest. The first is the answer to question whether the impact trajectories follows the same law as peak impacts. On the Bitcoin, where execution paths are roughly executed linearly in time, the agreement is remarkable, meaning that while the metaorder is not finished the market makes no difference between a metaorder that will stop soon and a metaorder that will continue¹¹. We find empirically that

$$\mathcal{I}_{\text{path}}(r, Q, \mu) = \mathcal{I}(rQ, \mu). \quad (5)$$

Hence, if t is the time elapsed since the start of the execution, and since μ is relatively constant for the metaorders in our dataset, one can write impact as

$$\mathcal{I}_{\text{path}}(r, Q, \mu) = \mathcal{I}(t, \mu) = f(\mu)t^\delta := \tilde{f}(\mu)(\mu t)^\delta \quad (6)$$

where $\delta \approx 1/2$. This in particular allows one to include in-trajectory points in the measurement of peak impact, cf. Fig. 5 (right) as they would have been the peak impact of the metaorder, had it stopped before. Note that in the above equation we implicitly average over any other variable so that f only depends on μ (with an assumption of independence).

Second, one observes that the value of the \tilde{Y} -factor appears to be noisy as impact lines do not superimpose, but are all perfectly parallel, across different Q s. A possible cause is the time variations of \tilde{Y} during the period considered (see Fig. 6), leading to a conditioning effect as soon as the distributions of Q is not similar during high- and low- \tilde{Y} periods. This suggests that the price during a metaorder execution can be written as

$$I(t, \mu) = \tilde{Y}(t)\sqrt{\mu t} + \sigma W_t, \quad (7)$$

where $\tilde{Y}(t)$ accounts for a (slowly) time-varying liquidity¹² and σ represent some additional market noise.

¹¹ This is at variance with Zarinelli et al. (2014) where average executions are not linear in time but rather front-loaded execution, resulting in very particular price trajectories.

¹² By writing this integrated form one implicitly assumes that $\tilde{Y}(t)$ is constant throughout the execution and varies at scales that are slower.

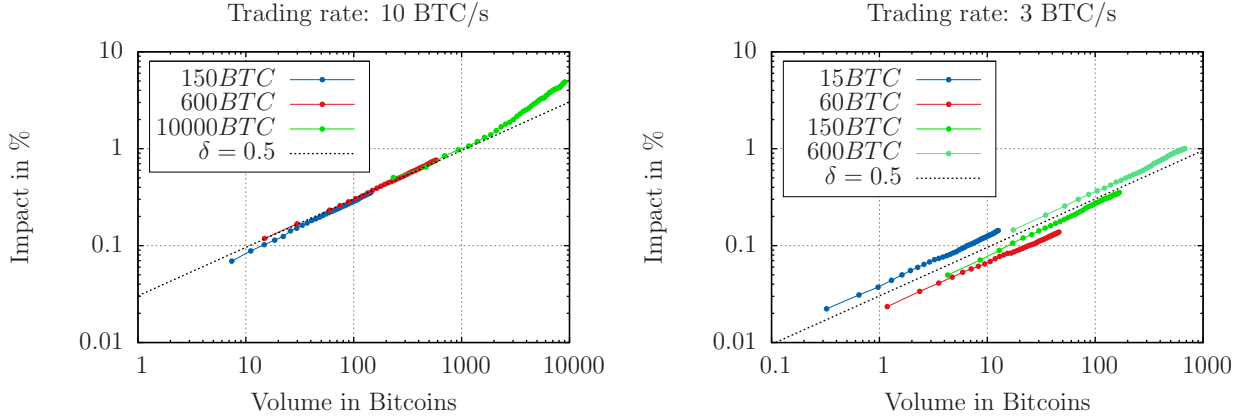


FIG. 6: Impact paths $\mathcal{I}^{\text{path}}(p, Q, \mu)$ in decimal loglog plot, for different metaorder volumes (cf. legends), for (left) $\mu = 10 \text{ BTC/s}$ and (right) $\mu = 3 \text{ BTC/s}$, and for $r \in [0, 1]$ for each couple (Q, μ) . The first value has intentionally been chosen high, so that it survives the criticism raised in Section 5.5.2 that on average other metaorders in the same direction are observed, which tends to artificially increase impact measures. One can observe a liquidity breakdown leading to an asymptotically linear impact when important pressures are maintained for too long on the same side of the order book.

4.3. The “Y-ratio”

Until now, much emphasis has been put in the literature on the study of the dependence in Q but very few studies have been realized on the pre-factor \tilde{Y} – or equivalently the *Y-ratio* defined in Eq. 1. We devote this section to a temporal analysis of both these pre-factors. For each metaorder i , we compute its individual \tilde{Y}_i as the ratio $\mathcal{I}(Q_i)/\sqrt{|Q_i|} \times \text{sign}(Q_i)$. For each day, we compute the daily average \tilde{Y} as a volume-weighted average of all individual ratios. In parallel, we compute the daily realized volatility σ_D and traded volume V_D , and we compare the daily \tilde{Y} ’s to the corresponding $\sigma_D/\sqrt{V_D}$ ratios for every day in our dataset. For each day, we also plot the actual *Y-ratio*. Results are presented on Fig. 7 and show that after such rescaling, the *Y-ratio* becomes nearly time independent. Its distribution is plotted on Fig. 8 and is well approximated by a Gaussian distribution $\mathcal{N}(Y_0, \Sigma_Y)$ with mean $Y_0 = 0.9$ and standard deviation $\Sigma_Y = 0.35$.

From these results, we can draw two conclusions of particular interest. First, it validates the scaling form of Eq. 1 proposed in Grinold and Kahn (2000), Torre and Ferrari (1997), Toth et al. (2011a). Indeed, the non-stationariness of the impact pre-factor \tilde{Y} is well encoded in the ratio $\sigma_D/\sqrt{V_D}$. Besides, the residual a-dimensional *Y-ratio* is shown to be of order unity with a standard deviation of the same order of magnitude so that it essentially lies in the interval $[0, 2]$. In this light, Eqs. 1 and 7 can be merged together so that impact reads

$$I(t, \mu) = \sigma \left[(Y_0 + \sigma_Y \eta) \sqrt{\frac{\mu t}{V_D}} + W_t \right], \quad (8)$$

where σ_Y accounts for some possible noise on the value of Y_0 . The relationship between the *liquidity noise* σ_Y and the *market noise* σ^{13} – and their dynamics – have not been investigated here, although it is a topic of interest for further studies.

The other conclusion is more macroscopic and relates to the study of market stability. Indeed, the scaling form for \tilde{Y} holds particularly well during extreme market events, such as the major crash that occurred on April 10, 2013, as evidenced in Donier and Bouchaud (2015b) who show that the impact pre-factor \tilde{Y} is a relevant proxy for market liquidity, even at scales much larger than that of impact. This definitely relates the microscopic aspect of price formation to its macroscopic characteristics such as its propensity to crash – see Kyle and Obizhaeva (2012) for similar discussions on financial markets. In this light, the understanding of how trades impact prices appears more crucial than ever to understand, detect – and perhaps even control? – market instabilities.

¹³ Both of which contribute to the Σ_Y evoked above.

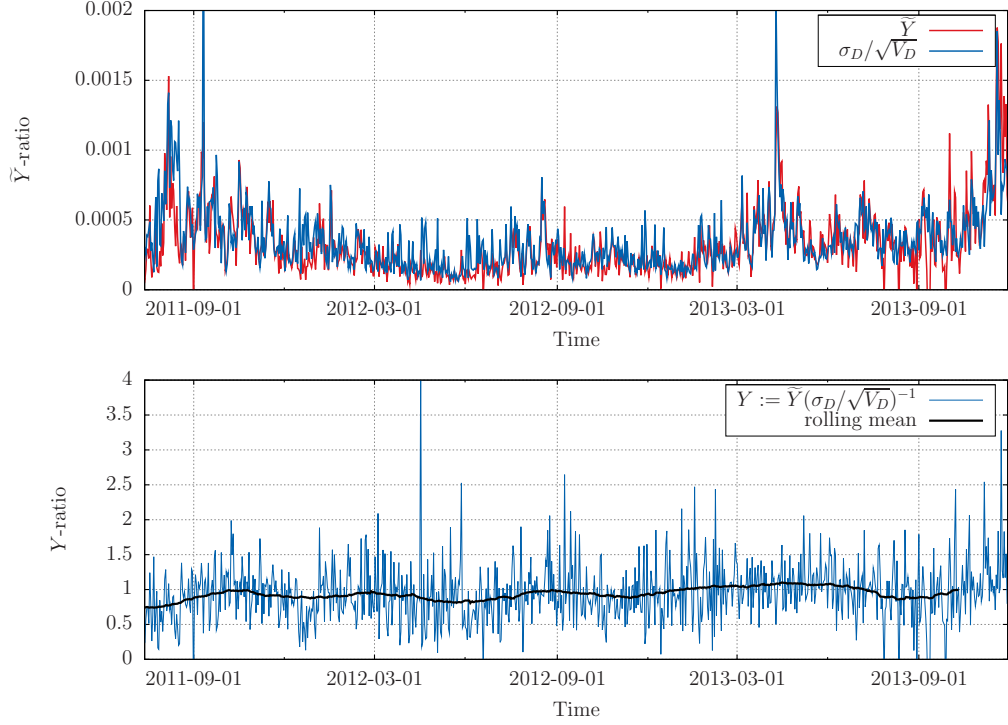


FIG. 7: (top) Raw impact pre-factor \tilde{Y} vs time. We also plot the usual normalization $\sigma_D\sqrt{V_D}$ to show that it accounts for the major part of the non-stationariness, particularly during extreme market events (e.g. April 10, 2013 major crash). (bottom) Y -ratio as defined in Eq. 1, which oscillates around its mean value $Y_0 \simeq 0.9$.

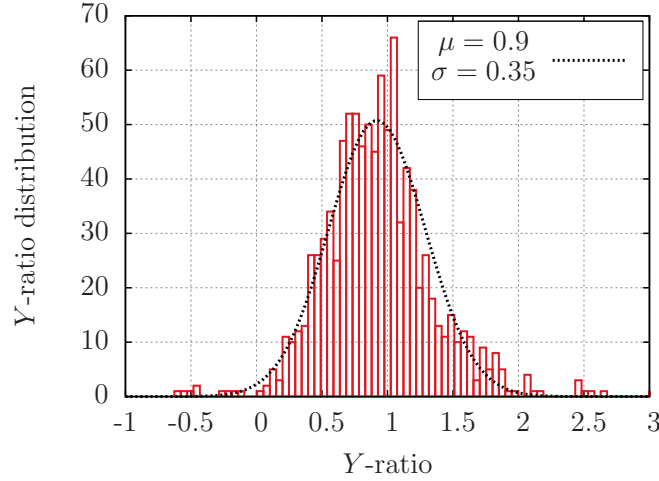


FIG. 8: Y -ratio distribution and its Gaussian approximation $\mathcal{N}(Y_0, \Sigma_Y)$ with mean $Y_0 = 0.9$ and $\Sigma_Y = 0.35$.

5. Impact, execution speed and correlations with the order flow

5.1. Impact trajectories of the bid and the ask

The large fees on Bitcoin allow for a separate study of the bid and the ask during the execution and prevents the data to be too noisy due to high-frequency arbitrage. For the sake of simplicity, we assume that metaorders are buy orders, so that the ask denotes the opposite side of the order book, on which the trader executes his/her metaorder. The facts of interest are the following :

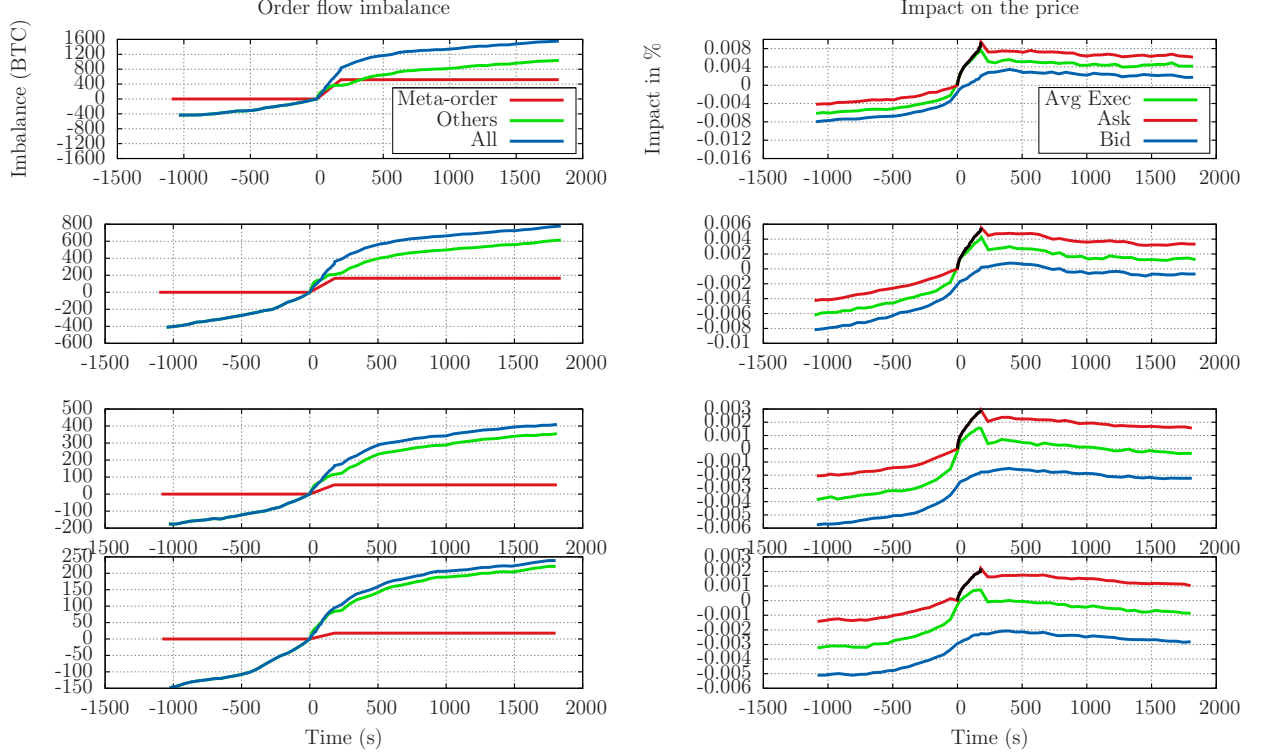


FIG. 9: (left) Order flow imbalances measured as the cumulated volumes of signed market orders during the execution of metaorders, for different metaorder volumes. For each panel, the metaorder starts at time 0 as shown by its executed volume in red. The blue line is the signed cumulated order flow of the whole market, whereas the green curve subtracts the metaorder from it to keep only the residual order flow. (right) Corresponding impact paths. In black, the actual execution path. In red, the ask path, and in light blue, the bid path. In green, we also plot the average execution price of the market. One can see that even for large volumes, impact is at most 1 or 2 spreads, not mentioning the fees that are of same order of magnitude than the spread: impact costs are dominated by friction costs, challenging martingale and fair pricing conditions.

- Before the execution, the spread is roughly constant, and the execution direction is on average positively correlated with both the bid and the ask evolution before the execution (Fig. 9).
- During the execution, the ask rises sharply – following the same square root as the execution price – whereas the bid follows more linearly. This is very reminiscent of what is found theoretically in Donier et al. (2015).
- After this quick reversion, the bid and the ask remain roughly constant at a non-zero permanent level, in spite of the order flow pressure from the rest of the market that continues some time after the metaorder.
- These observations hold when we condition the metaorder to be trend-following or mean-reverting, as shown on Fig. 10. In particular, the fact that impact is still square root at small scales for trend-following metaorders is non-trivial and reveals that *trading speed matters*¹⁴.
- In any case, impact does not exceed very much the order of magnitude of the spread and the fees, which questions any interpretation based on “fair-price” arguments, all the more so that impact is square root even at the smallest scales.

¹⁴ It contradicts in particular equilibrium models that only consider the aggregate order flow as a relevant quantity, since in these models *trading speed* is transparent.

Consistent with the observations of Fig. 4, one observes that on average metaorders are positively correlated with the remaining order flow. We will see below that this plays a crucial role in the fact that the permanent component of impact is non-zero in these plots.

Finally, one notices that the market VWAP in green lies around the mid-price in the absence of the metaorder, and is naturally biased towards the execution side during the execution – all the more so that the execution is aggressive.

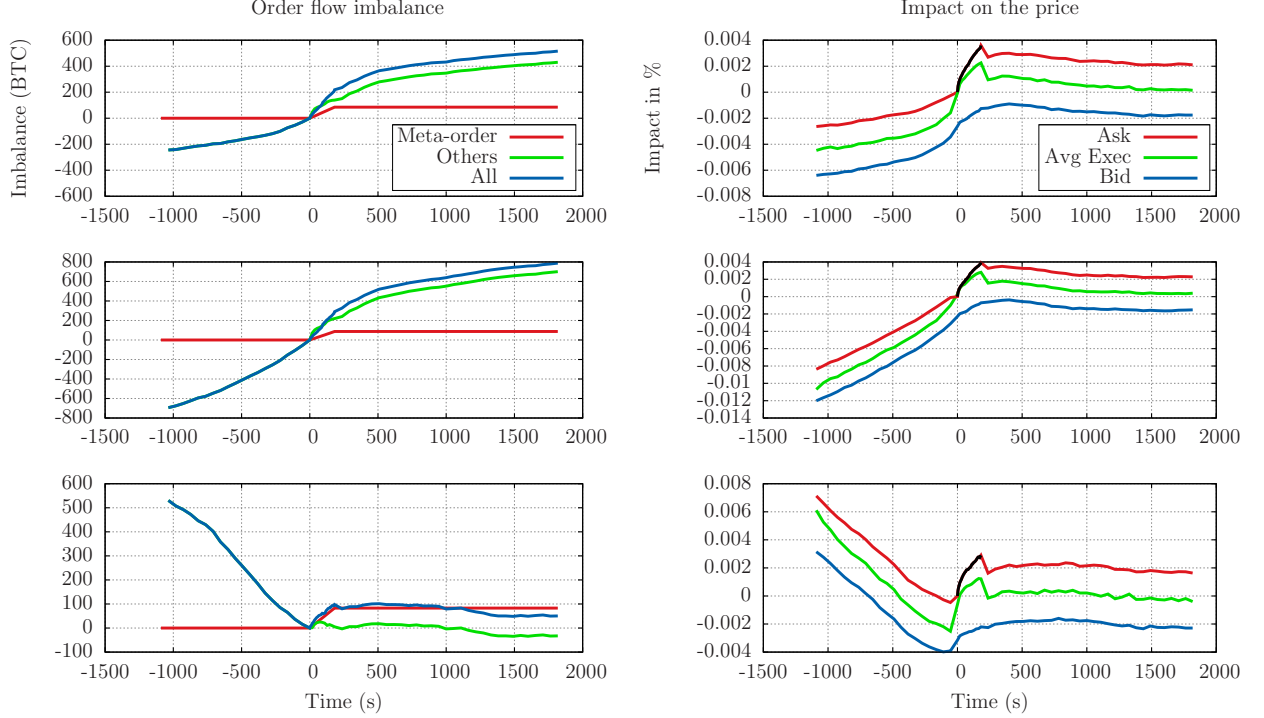


FIG. 10: Same pictures as in Fig. 9 for various types of conditioning (*top*) Unconditioned metaorders. (*middle*) Trending metaorders. (*bottom*) Mean-reverting metaorders.

5.2. Impact and execution speed

The question of dependence of impact on the execution speed μ has been addressed very recently by Zarinelli et al. (2014) even though practitioners have been looking into it for at least a few years. In this section we study the *impact surface* described by the bivariate function $\mathcal{I}_{\text{exec}}(Q, \mu_V)$ ¹⁵ For each fixed μ_V the dependence on Q is found to be square root:

$$\langle \mathcal{I}^{\text{exec}}(Q, \mu_V) | \mu_V \rangle \sim \sqrt{Q}. \quad (9)$$

However, the dependence on μ is somewhat surprising. First, for very high participation rates (close to 100% of the volume during the period), the impact becomes unusually high. This is not surprising since the market breaks down in this regime¹⁶. More importantly however, while one could expect that impact monotonically decreases as the execution rate decreases, the observed impact actually *increases* again (cf. Fig. 11), at variance with intuition and previous findings on financial markets (Zarinelli et al. 2014). The best fit to the empirical data reads $I^{\text{exec}}(Q, \mu_V) \sim Q^\delta / \mu_V^{\delta'}$, with $\delta \approx 0.5$ and $\delta' \approx 0.4$ (cf. Fig. 11). The slower the execution, the larger the measured impact: this strange

¹⁵ We prefer here the average quantity $\mathcal{I}_{\text{exec}}(Q, \mu_V)$ to the peak impact $\mathcal{I}(Q, \mu_V)$ since latter is much noisier. The mean execution price is anyway far more relevant in practice since it is directly related to *execution costs*. We also preferred the execution rate μ_V to the execution speed μ as a fundamental variable since for this particular study this quantity is more relevant and yields cleaner pictures.

¹⁶ this fact is of importance since such an observation would probably be impossible on financial markets, where participation rates rarely exceed 20 – 30% – for this very reason.

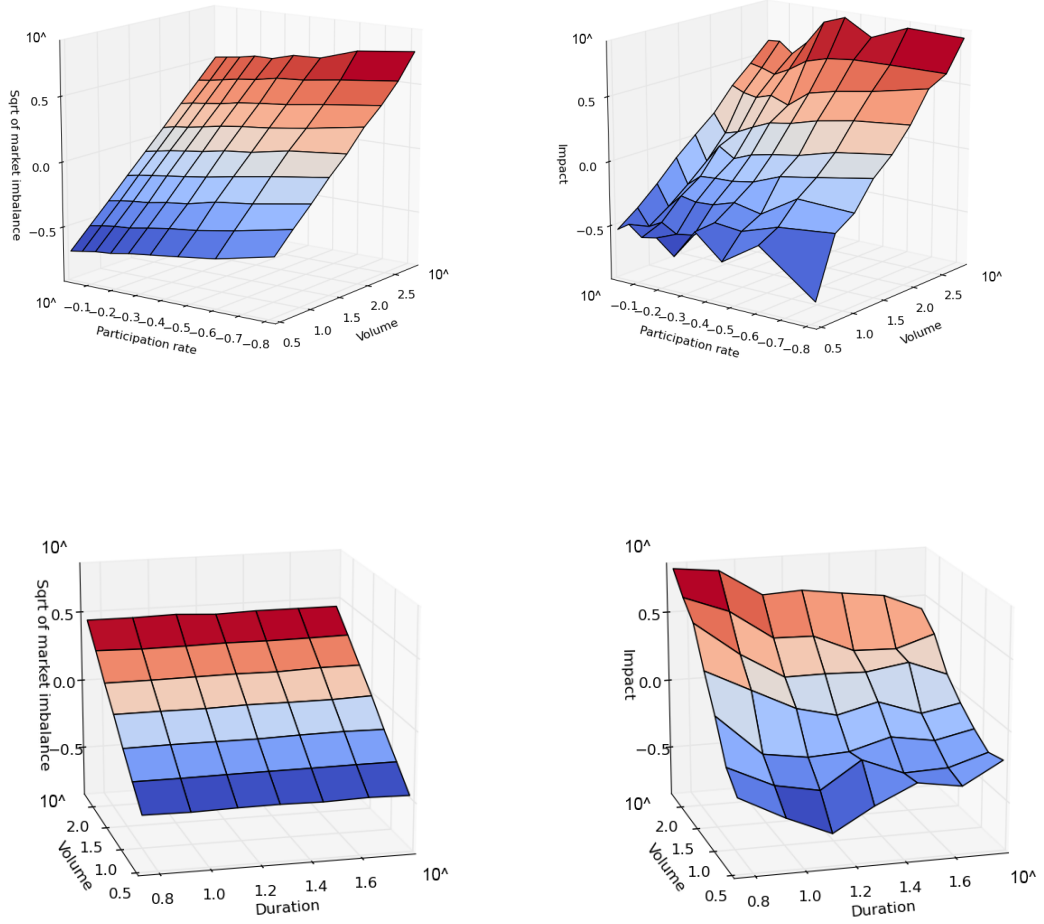


FIG. 11: The impact surface (in log-log-log scale) for typical metaorders (top) which are correlated with the total order flow. During the execution span of the metaorder, the market reacts rather to total market imbalance as to the individual market order: For all participation rates market impact is almost perfectly proportional to the square root of the total order flow as defined in the main text. Bottom images: Impact of an isolated metaorder (for which the residual market order flow remains neutral). The total market imbalance corresponds to the volume of the metaorder; hence, the participation rate is not a suitable measure. Rather, the execution time should be regarded. The figure clearly displays an impact that decreases when the execution time increases.

dependence on μ_V stresses the difference between mechanical and informational impact. Clearly, a slow execution gives other market participants the opportunity to detect the same signal and the information content of the metaorder realizes during its execution. On the other hand, fast execution leads mostly to mechanical impact as the *alpha* realizes itself afterwards. Following this first non-intuitive discovery – which is very specific to the Bitcoin –, for each of the data points of Fig. 11 we plotted the whole market order imbalance $\text{sign}(Q) \cdot V_M^{\text{signed}}$ ¹⁷. The similarity of both pictures (see Fig. 11 top images) is striking: during metaorders, impact is a very nice square root of *global market imbalance*. This leads to the following conclusion: Market impact is not a reaction to *individual* metaorders, but to the *whole* order flow. This seems rather natural since orders are anonymous, hence the aggregated order flow should be the only relevant quantity.

¹⁷ $V_M^{\text{signed}} = \sum_i v_i \epsilon_i$ where $\epsilon_i = \pm 1$ according to whether the trade is triggered by the buyer (resp. seller) and v_i is its volume.

5.3. Permanent impact and correlation

One question remains however: How to study the mechanical impact of *one isolated* metaorder? We can answer this question by searching the data for metaorders that are not correlated with the rest of the market, i.e. in the course of which the residual order flow does not trend nor anti-trend: we selected all metaorders that account for more than 75% of the market net imbalance on $[t_0, t_0 + 10T]$ where t_0 is the starting time of the metaorder and T is its duration, so that the rest of the market can be considered neutral during the measurements – hence the terminology of *isolated* metaorder. Fig. 12 compares the impact of randomly chosen metaorders (which are positively correlated to the rest of the market), that we will refer to as “informed”, and isolated metaorders (that we will refer to as “uninformed”) and clearly shows that impact of isolated metaorders decreases far below the 2/3 threshold. Note that the above selection labels about 3% of the metaorders as isolated. This illustrates the fact that impact is built up by an “informational”

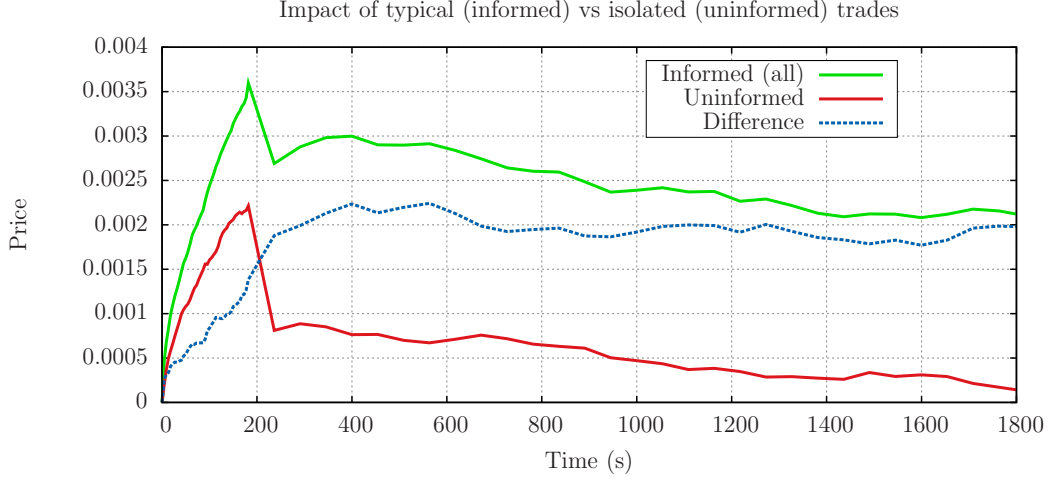


FIG. 12: Impact of “informed” metaorders vs. “uninformed” metaorders (i.e. isolated metaorders) on the opposite best price. While the former have a permanent impact due to their correlation with the residual order flow, the latter do not affect the price in the long run – or very few.

component, that reveals itself during and after the execution and results in an apparent permanent impact¹⁸ (in the case of isolated orders this component is by definition zero or very small) and a “mechanical” component whose shape after execution is consistent with a decay all the way to zero, in agreement with the findings of Gomes and Waelbroeck (2015). This suggests that the mechanical peak impact is found by removing the permanent part of the impact:

$$\mathcal{I}_{\text{mec}}^{\infty}(Q, \mu) \approx \mathcal{I}(Q, \mu) - \mathcal{I}^{\infty}(Q, \mu) . \quad (10)$$

One might ask whether the “informed/uninformed” terminology is appropriate here, since we never wonder about external information. Actually, the fact that isolated metaorders have no permanent impact, and therefore no “informational content” on average, indicates that the permanent component of the impact – that one may refer to as the *informational content* of the metaorder – should be interpreted as a *correlation* between the metaorder and future traders’ behaviours and not as some information about any “fundamental price” (see Donier and Bouchaud (2015a)). Concerning the dependence of impact on the execution speed, the same kind of impact surface can be plotted after conditioning on isolated metaorders only, showing that after accounting for the bias the impact actually *decreases*, as expected intuitively, when the execution speed decreases:

$$\mu_1 < \mu_2 \Rightarrow \mathcal{I}_{\text{mec}}^{\infty}(Q, \mu_1) < \mathcal{I}_{\text{mec}}^{\infty}(Q, \mu_2) .$$

However, the measure is too noisy to quantify precisely this dependence. Such a behaviour is strikingly reminiscent of what is obtained within the reaction-diffusion framework – and less consistent with equilibrium models which so far do not deal with the issue of execution speed.

¹⁸ This is Hasbrouck’s (Hasbrouck 1991) definition of “information”.

6. Summary of main results

We have presented a very detailed analysis of market impact on the Bitcoin exchange market. For the sake of a clear understanding let us summarize here our main results:

1. Large orders are split into small trades which confirms the wide-spread belief that large orders have to be executed incrementally. On our MtGox dataset over 1 million metaorders are clearly identifiable, corresponding to 14M trades.
2. Metaorders size and duration distributions are not power-law. However, on Bitcoin they have the very nice property of being executed at a constant average rate.
3. The market impact of a metaorder executed incrementally and linearly in time can be reduced in a dependence on the execution speed and on the time elapsed since the start of the execution and written $\mathcal{I}_{\text{path}}(r, Q, \mu) = \mathcal{I}(t, \mu)$. The square-root law is clearly confirmed as it even holds *trajectory-wise* and is exact even at the smallest scales, which once again invalidates a broad class of equilibrium theories. Besides it has been shown that this cannot be explained by biases: the square-root law describes how market digests local excesses in supply /demand.
4. The impact pre-factor \tilde{Y} is subject to large fluctuations, that are well explained by the ratio $\sigma_D/\sqrt{V_D}$ where σ_D is the daily volatility and V_D the daily volume. The residual *Y-ratio*, as defined in Eq. 1, fluctuates around a mean-value of order unity: to be more precise, its distribution is well approximated by a Gaussian distribution of mean $Y_0 = 0.9$ and standard deviation $\Sigma_Y = 0.35$. The full stochastic impact formula we propose is thus given by Eq. 8.
5. We have presented strong evidence that the market reacts to the total order flow and not to distinct metaorders. This implies that metaorders are not detected by other market participants, as assumed in some equilibrium models (Farmer et al. 2013). More importantly, the impacts of different metaorders are in fact *dependent* and *do not add up linearly*.
6. To subtract the effect of the order flow, and to measure the marginal effect of *one isolated* metaorder, we have selected metaorders that were not correlated to the market's local direction. This allowed us to show that for slow execution speeds the average execution cost $\mathcal{I}_{\text{exec}}(Q, \mu)$ decreases – even though measures are too noisy to make a proper fit – and that the mechanical component of the permanent impact $\mathcal{I}_{\text{mec}}^{\infty}(Q, \mu)$ is close to zero. Such a behaviour is consistent with kernel models and statistical order book models that naturally generalize them.
7. The marginal impact of these isolated metaorders drops far below the 2/3 level predicted by equilibrium theories. The permanent impact of isolated trades is probably zero, as concluded in Brokmann et al. (2014) and Gomes and Waelbroeck (2015).

7. Conclusion

Using a remarkable dataset which covers all transactions on the MtGox Bitcoin/USD exchange, we have conducted a comprehensive analysis of the market impact of over one million metaorders on Bitcoin. The Bitcoin market has two important features which motivate such a study: First, it corresponds to a single-asset economy, a quite unique example. Second, since each transaction is charged fees of 0.6% of its amount (60 bps), the existence of market makers and arbitrageurs is the exception rather than the rule.

The fact that the square-root law for impact holds in these conditions allows us to assess some underlying hypotheses of current impact models. Due to the large fees, neither statistical arbitrage nor high-frequency market making can be at the origin of the square-root law (on the Bitcoin) and theoretical explanations of this stylized fact must be found elsewhere. Accordingly, the notion of fair pricing is irrelevant in this market, as it seems not reasonable to assume that investors behave in such a way that they break even on their impact (of order a fraction of a percent) when they have to pay fees that are much larger. Nor should one rely on any martingale conditions for the price, as the notion of “price” is not precisely defined under the scale of 60 bps due to the large spread. The fact that the square-root impact law holds on the Bitcoin market so precisely at all scales clearly demonstrates that equilibrium and arbitrage mechanisms are not the underlying mechanism of market impact. Furthermore, our study suggests that impact should be regarded as how trades dig into the opposite side rather than how they affect the “price” itself, which is not equivalent when the spread is not tight.

To conclude, this study incites us to think that impact is driven by a generic fundamental and local mechanism that emerges together with the simple phenomenon of supply and demand and pre-exist any notion of arbitrage (although it might remain compatible with it in the end as behaviours adapt to it). A promising avenue in our view, is the use of heterogeneous agent models (Donier and Bouchaud 2015a), in which supply and demand are the output of interactions in a complex system that cannot be reduced to a few representative agents, and which seem to reproduce impact pictures remarkably well (Donier et al. 2015).

Acknowledgements

We warmly thank Jean-Philippe Bouchaud for countless discussions and for carefully reading the manuscript. We also thank N. Kornman and A. Tilloy for their precious insights on the Bitcoin and finally C.A. Lehalle and H.Waelbroeck for very useful discussions.

-
- Alfonsi, A. and Schied, A. (2013). Capacitary measures for completely monotone kernels via singular control. *SIAM Journal on Control and Optimization*, 51(2):1758–1780.
- Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.
- Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18:5752.
- Avellaneda, M., Kasyan, G., and Lipkin, M. D. (2012). Mathematical models for stock pinning near option expiration dates. *Communications on Pure and Applied Mathematics*, 65(7):949–974.
- Bak, P., Paczuski, M., and Shubik, M. (1997). Price variations in a stock market with many agents. *Physica A: Statistical Mechanics and its Applications*, 246(3):430–453.
- Beja, A. and Goldman, M. B. (1980). On the dynamic behavior of prices in disequilibrium. *The Journal of Finance*, 35:235.
- Bershova, N. and Rakhlin, D. (2013). The non-linear market impact of large trades: Evidence from buy-side order flow. *Quantitative Finance*, 13(11):1759–1778.
- Bouchaud, J.-P., Farmer, J. D., and Lillo, F. (2009). How markets slowly digest changes in supply and demand. *Handbook of financial markets: dynamics and evolution*, 1:57.
- Brogaard, J., Hendershott, T., and Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, 27:2267–2306.
- Brokmann, X., Serie, E., Kockelkoren, J., Bouchaud, J.-P., et al. (2014). Slow decay of impact in equity markets. *Available at SSRN 2471528*.
- Budish, E. and Cramton, P. (2013). The high-frequency trading arms race: Frequent batch auctions as a market design response.
- Donier, J. (2012). Market impact with autocorrelated order flow under perfect competition. *arXiv preprint arXiv:1212.4770*.
- Donier, J., Bonart, J. F., Mastromatteo, I., and Bouchaud, J.-P. (2015). A fully consistent, minimal model for non-linear market impact. *Quantitative finance*, 15(7):1109–1121.
- Donier, J. and Bouchaud, J.-P. (2015a). From walras’ auctioneer to continuous time double auctions: A general dynamic theory of supply and demand. *arXiv preprint arXiv:1506.03758*.
- Donier, J. and Bouchaud, J.-P. (2015b). Why do markets crash? bitcoin data offers unprecedented insights. *Available at SSRN 2583743*.
- Farmer, J. D., Gerig, A., Lillo, F., and Waelbroeck, H. (2013). How efficiency shapes market impact. *Quantitative Finance*, 13(11):1743–1758.
- Farmer, J. D. and Lillo, F. (2004). The long memory of the efficient market. *Studies in Nonlinear Dynamics and Econometrics*, 8:1–33.
- Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937):267–270.
- Gatheral, J. (2010). No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759.
- Glosten, L. and Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100.
- Gomes, C. and Waelbroeck, H. (2015). Is market impact a measure of the information value of trades? market response to liquidity vs. informed metaorders. *Quantitative Finance*, 15(5):773–793.
- Grinold, R. C. and Kahn, R. N. (2000). Active portfolio management.
- Hasbrouck, J. (1991). Measuring the information content of stock trades. *The Journal of Finance*, 46:179–206.
- Hendershott, T. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1):1–33.
- Hill, B. M. et al. (1975). A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174.
- Jaisson, T. (2015). Market impact as anticipation of the order flow imbalance. *Quantitative Finance*, (ahead-of-print):1–13.
- Kirilenko, A., Kyle, A., Mehrdad, S., and Tugkan, T. (2014). The flash crash: The impact of high frequency trading on an electronic market.

- Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–35.
- Kyle, A. S. and Obizhaeva, A. (2012). Large bets and stock market crashes. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2023776.
- Lehalle, C.-A. and Lasnier, M. (2012). Mathematical models for stock pinning near option expiration dates. what does the saw-tooth pattern on us markets on 19 july 2012 tell us about the price formation process. *Technical report, CA Chevreux Quantitative Research*.
- Mastromatteo, I., Toth, B., and Bouchaud, J.-P. (2014a). Agent-based models for latent liquidity and concave price impact. *Physical Review E*, 89(4):042805.
- Mastromatteo, I., Toth, B., and Bouchaud, J.-P. (2014b). Anomalous impact in reaction-diffusion financial models. *Physical review letters*, 113(26):268701.
- Moro, E., Vicente, J., Moyano, L. G., Gerig, A., Farmer, J. D., Vaglica, G., Lillo, F., and Mantegna, R. N. (2009). Market impact and trading profile of hidden orders in stock markets. *Physical Review E*, 80(6):066102.
- Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Consulted*, 1(2012):28.
- Torre, N. and Ferrari, M. (1997). Market impact model handbook, BARRA inc., Berkeley. <http://www.barra.com/newsletter/nl166/miminl166.asp>, but see also Grinold and Kahn (2000).
- Toth, B., Lemperiere, Y., Deremble, C., De Lataillade, J., Kockelkoren, J., and Bouchaud, J.-P. (2011a). Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2):021006.
- Toth, B., Lemperiere, Y., Deremble, C., de Lataillade, J., Kockelkoren, J., and Bouchaud, J.-P. (2011b). Anomalous price impact and the critical nature of liquidity in financial markets. *Phys Rev X*, 1:021006.
- Toth, B., Lillo, F., and Farmer, J. D. (2010). Segmentation algorithm for non-stationary compound poisson processes. *The European Physical Journal B-Condensed Matter and Complex Systems*, 78(2):235–243.
- Toth, B., Palit, I., Lillo, F., and Farmer, J. D. (2015). Why is equity flow so persistent. *Journal of Economic Dynamics and Control*, 51:218–239.
- Wyart, M., Bouchaud, J.-P., Kockelkoren, J., Potters, M., and Vettorazzo, M. (2008). Relation between bid–ask spread, impact and volatility in order-driven markets. *Quantitative Finance*, 8(1):41–57.
- Zarinelli, E., Treccani, M., Farmer, J. D., and Lillo, F. (2014). Beyond the square root: Evidence for logarithmic dependence of market impact on size and participation rate. To appear.