Dissecting cross-impact on stock markets: An empirical analysis

M Benzaquen, I Mastromatteo, Z Eisler and J-P Bouchaud

Capital Fund Management, 23 rue de l'Université, 75007 Paris

E-mail: zoltan.eisler@cfm.fr

9 September 2016

Abstract. The vast majority of recent studies in market impact assess each product individually, and the interactions between their order flows are disregarded. This strong approximation may lead to an underestimation of trading costs and possible contagion effects. Transactions mediate a significant part of the interaction between different instruments. In turn, liquidity shares the sectorial structure of market correlations, which can be encoded as a set of eigenvalues and eigenvectors. We introduce a multivariate linear propagator model that successfully describes such a structure, and reproduces well the response and a significant fraction of the covariance matrix of returns. We explain in detail the various dynamical mechanisms that contribute to these quantities. We also define two simplified models with substantially less parameters to reduce overfitting, and show that they have superior out-of-sample performance.

1. Introduction

Price impact in financial markets – the effect of transactions on the observed market price – is of both scientific and practical relevance [1]. A long series of studies has concentrated on its various aspects in the past decades [2–9]. The metrics used in this body of work are calculated individually on each product, and possibly averaged across them afterwards. The interactions between their order flows are typically disregarded. This is a very strong approximation, given that a financial instrument is rarely traded on its own. Most investors construct diversified portfolios by buying and selling tens or even hundreds of assets at the same time. Some of these might be similar, or even almost equivalent to each other (companies in the same industrial sector, dual-listed shares, etc.). In these cases it is immediately clear that to treat each of them separately leads to a gross error, and often an underestimation of trading costs. Intuition tells us that in two related products the order flow of one may reveal information, or communicate excess supply/demand regarding the other. The question is how important such effects are, both qualitatively and quantitatively.

The "self-impact" of a product's order flow on its own price, as studied in the literature, is an important component of price dynamics. In comparison, is "cross-impact" a detectable effect? If it is, is it strong enough to significantly contribute to cross-correlations between stocks? This question was already raised in the seminal work of Hasbrouck and Seppi [10]. It is particularly interesting, because despite the importance of cross-correlations in risk management, their microstructural origin is not clear. Many partial, competing explanations exist, for a review of recent economics literature on the subject see in Ref. [11]. When choosing their quotes, liquidity providers use correlation models calibrated from real data. It would be a circular argument to then say that such correlations emerge only due to the market makers' behavior. A dynamical explanation is more plausible. We believe that correlations observed between real quotes have a stable, logical structure, because if two stocks get out of line, liquidity takers act on such a mispricing. As they consume liquidity, market makers adjust their pricing to avoid building up a large inventory: this is price impact. As the relative price reaches the market consensus the arbitrage opportunity disappears, and order flows become balanced. Several structural, equilibrium theories exist with such dynamics, but the underlying models often have many parameters which cannot be directly fitted to data. Only the qualitative predictions can be observed, but these are nevertheless very important for practical purposes [12].

In this manuscript we are going to argue in favor of such a dynamical picture, where transactions mediate a significant part of the interaction between different instruments, and price impact is an integral part of price formation. We will demonstrate quantitatively that correlations and liquidity are intertwined. Refs. [13, 14] revisit the evidence for cross-impact by analyzing the cross-correlation structure of price changes and order flows. Our study complements such a perspective by focusing on the underlying interactions rather than on correlations. Based on a variant of the well-known propagator technique [6], calibrated on anonymous data, we will show that liquidity displays a sectorial structure related to the one of market correlations, that we will be able to describe through decomposition in eigenvalues and eigenvectors. This is in the spirit of the principal component analysis approach advocated in Ref. [10], and the analysis provided in Ref. [12] from an econometric point of view.

For the sake of simplicity in this paper we will use the language of stocks, and we will also limit our datasets to these. However, the techniques introduced below can be applied to many other markets. Moreover, note that we focus here on the impact of the aggregated order flow, rather than the one of a *meta-order* (a sequence of trades in the same direction submitted by the same actor). Even though the propagator formalism that we employ predicts inaccurately the impact of a meta-order, it provides qualitatively reliable estimates of market impact (see Section 5 of Ref. [15]). Thus, we believe that the cross-interaction network that we find should generalize to the meta-order case with little variation.

The paper is structured as follows. Section 2 introduces basic notations and our dataset. Section 3 defines a few fundamental quantities related to returns and price

impact, and summarizes that main stylized facts that we observe. Section 4 provides a non-parametric multivariate propagator model, which is then fitted to the data. Section 5 analyzes simpler, lower-dimensional models that can more efficiently capture the structure of cross-impact; and compares their in-sample and out-of-sample performance. Finally, Section 6 concludes.

2. Data and notations

We conduct our empirical analysis on a pool of $N = 275$ US stocks as representative as possible in terms of liquidity, market capitalisation and tick size. The large number of assets and their diversity allows for great statistical significance, and to investigate the scaling of our results when the number of products becomes large. The data consists of five-minute binned trades and quotes information from January 2012 to December 2012, extracted from the primary market of each stock (NASDAQ or NYSE). Furthermore, we only focus on the continuous trading session, removing systematically the first hour after the open and the last 30 minutes before the close. In this way we avoid artifacts arising from the particularities of trading activity in these periods. Out-of-sample tests will be carried out on an equivalent dataset from 2013.

For each five-minute window whose end point is t and for each asset i, we compute the log-return $x_t^i = X_t^i - X_{t-1}^i$, where $X_t^i = \log p_t^i$ and where p_t^i denotes the price of stock i at time t. In addition, we compute the trade imbalance $\varepsilon_t^i = n_t^{i,\text{buy}} - n_t^{i,\text{sell}}$, where $n_t^{i,\text{buy/sell}}$ denotes respectively the number of buyer- and seller-initiated market orders of stock i in bin t. We choose this proxy for volume imbalance because the strong fluctuations in the size of the trades are only moderately compensated by the information that they provide [6].

We normalise x_t^i and ε_t^i by their standard deviation computed over the entire trading period. As a result, both time series display zero mean and unit variance. This choice of normalisation has the benefit of making the problem *extensive* in the following sense: For any linear model that one infers, (such as the one presented in Section 4.1), the results obtained for a larger bin size (say, one hour) can always be recovered from the results obtained at a finer scale. Moreover, extensivity allows the predictions of the model not to depend on the estimation of the local normalization. One does not need to build estimators for volatility and volume in the next five-minutes bin in order to exploit these results. This would not have been the case had we used a local normalization for the fluctuations of the returns and the volumes. Indeed, we have checked that the choice of a local normalization, while spoiling extensivity, yields qualitatively similar results.

Also note that we have chosen to use real time to measure t as opposed to counting it on a trade-by-trade basis. This is because in the regime of large N that we consider, there would be too many trades, and our dataset would become unmanageable [13, 14]. Finally, the choice of a five-minute bin size allows us to abstract away from microstructure effects which are not the subject of the present *mesoscopic* study. All along this manuscript time shall be seen as dimensionless, five minutes being the unit.

3. Market impact and price fluctuations

In this section, we define the multivariate correlation functions relevant to the problem at hand, and investigate their relations.

3.1. The correlation structure of returns

The covariance matrix of returns is one of the central objects in quantitative finance, and is of paramount importance in a number of applications such as portfolio construction and risk management [16, 17]. Let us recall first some of its most prominent properties.

We denote by Σ_τ^{ij} the return covariance of contracts i and j at scale τ, defined as

$$\Sigma_\tau^{ij} = \mathbb{E}[(X_{t+\tau}^i - X_t^i)(X_{t+\tau}^j - X_t^j)] \tag{1}$$

Figure 1(a) displays a plot of the mean diagonal $\Sigma_\tau^{\text{diag}} = N^{-1}\sum_i \Sigma_\tau^{ii}$ and off-diagonal $\Sigma_\tau^{\text{off}} = (N^2-N)^{-1}\sum_{i\neq j}\Sigma_\tau^{ij}$ return covariances rescaled by τ. As one can see, the diagonal terms of the return covariance matrix are on average a factor ~ 5 larger than the off-diagonal ones. Microstructural effects are almost absent in Σ_τ^{ij} even at $\tau = 1$: we only observe a weak decrease of the variance at short lags in the signature plot, and the ratio between covariance and variance – that determines the so-called Epps effect [18, 19] – is almost flat in τ. This is consistent with price efficiency, because the time scale for arbitrage effects is expected to be well below our five-minute window [20]. Finally, one can define the customary return correlation matrix as $\Sigma_\tau^{ij}(\Sigma_\tau^{ii}\Sigma_\tau^{jj})^{-1/2}$.

Figure 2(a) displays a plot of Σ_τ^{ij} at $\tau = 1$ from which we subtracted its mean (≈ 0.21) for better readability, and in which the contracts have been sorted by industrial sector, as indicated by the labels. As one can see, Σ_τ^{ij} displays a strong sectorial structure, in line with previous studies [21]. The behaviour of the covariance matrix is well understood in its eigenbasis. Indeed, Σ_τ^{ij} is a real symmetric matrix, so it be diagonalised as

$$\Sigma_\tau^{ij} = \sum_a O_\tau^{ia}\Lambda_\tau^a O_\tau^{ja}. \tag{2}$$

O_τ^{ia} is an orthogonal matrix, its columns correspond to the eigenvectors of Σ_τ^{ij}, and Λ_τ^a is a vector made of the corresponding eigenvalues. Figure 2(b) displays the histogram of the eigenvalues Λ_τ^a at $\tau = 1$. We have assessed their stability by verifying that $\Lambda_\tau^a \propto \tau$, as it was the case for the average quantities displayed in Fig. 1(a). Interestingly, we find the eigenvectors O^{ia} to be stable in time, indicating that the directional structure of the market is consistent across scales ranging from some minutes to one day, while its associated fluctuations increase linearly. The value of the largest eigenvalue $\Lambda_1^0 \approx 62$, indicates that $\Lambda_1^0/N \approx 23\%$ of the total variance of the system can be explained by this mode, in good agreement with Ref. [10]. Often referred to as the market mode, it corresponds to a collective – and rather homogeneous – mode, as can be seen on Fig. 2(c). The next few modes after the market mode, individually, explain a considerably smaller part of the variance. Their structure supports an economic interpretation in terms of

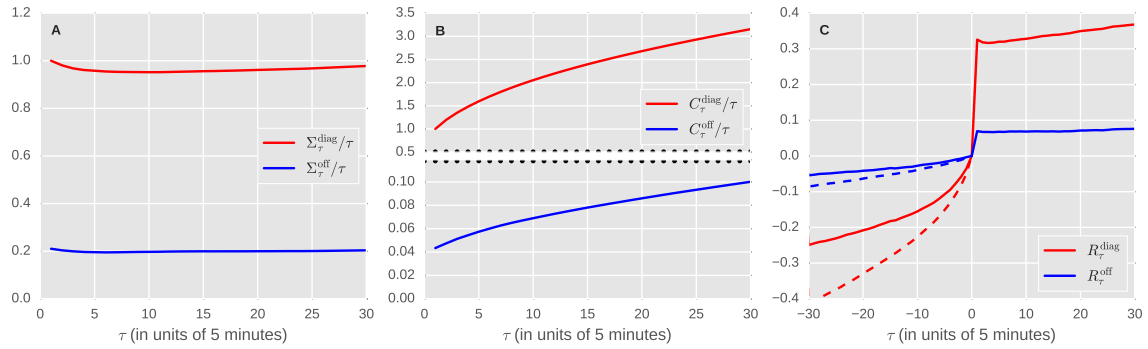

Figure 1. Plots of average diagonal and off-diagonal (a) returns covariance (see Eq. (1)), (b) sign covariance (see Eq. (4)), and (c) response function (see Eq. (5)). The dashed lines for the response indicate the prediction of the model at negative lags.

industrial sectors (see Fig. 2(c) and Ref. [21]). The subsequent modes fall into a noise band that is roughly described by a Marčenko-Pastur distribution [22, 23] (see red curve on Fig. 2(b)), due to the fact that the number of stocks is very large with respect to the number of observations that would be required to obtain a statistically accurate estimation of all the modes.

3.2. The correlation structure of the trade signs

In order to investigate the relation between returns and trade sign imbalance, it is natural to define a covariance matrix for the signs, and to compare its structure with the one built out of the returns. Accordingly, we define the lagged covariance of signs as

$$c_\tau^{ij} = \mathbb{E}[\varepsilon_{t+\tau}^i \varepsilon_t^j] \ . \tag{3}$$

Its behaviour is radically different from that of returns. While returns are uncorrelated ($\Sigma_\tau^{ij} \sim \tau$ after a few trades) thereby ensuring market efficiency, signs are long-range correlated with an exponent $c_\tau^{ij} \sim \tau^{-\gamma}$ with $\gamma \sim 0.5$ (see Appendix). This result stems from the fact that in limit order markets investors split their trading decisions into smaller pieces in order to avoid excessive costs, because instantaneously available liquidity at good prices is small [7]. This yields the famous anomalous response puzzle: Prices diffusive despite being driven by trades which themselves are superdiffusive.

A well-known solution to this problem is that of the linear propagator model (or surprise model), postulating that trades in an expected direction impact the price less than those in the unexpected one [6, 7, 24]. While this model has been thoroughly explored in one dimension, its richer multi-dimensional counterpart has not been considered yet. A multivariate framework allows to precisely formulate a number of questions that are central in our study, and that cannot be addressed in a one-

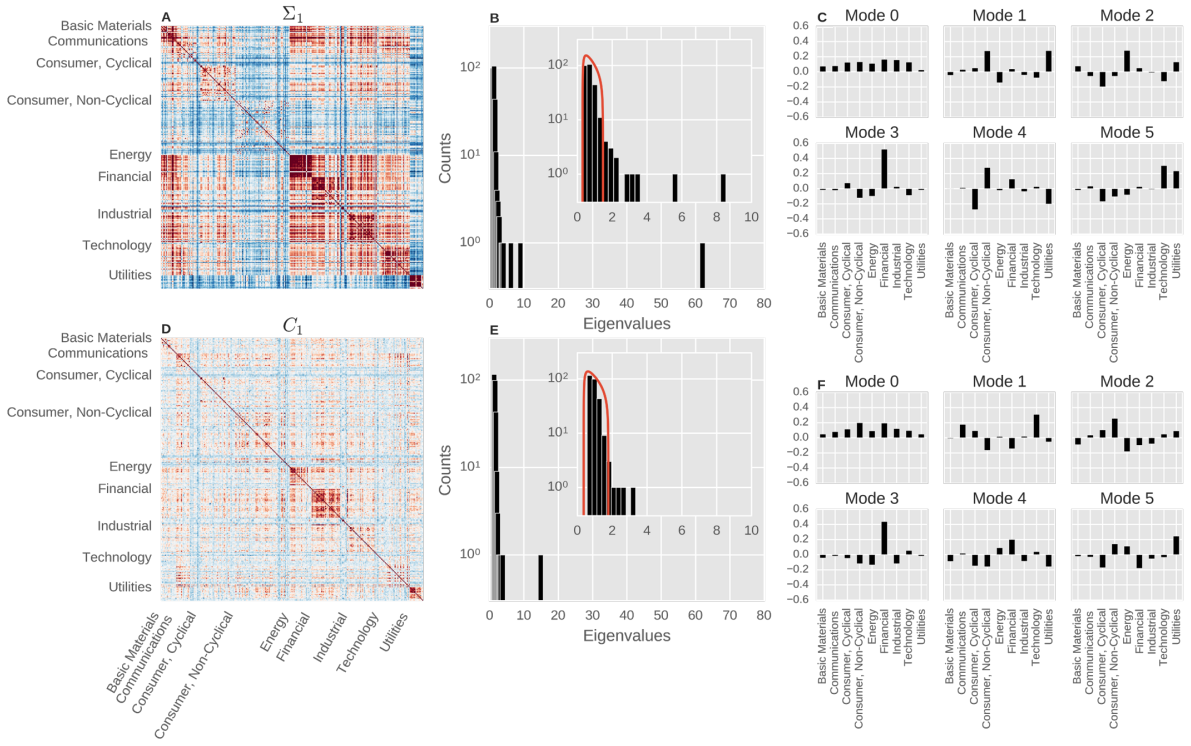

Figure 2. (a) Plot of the returns covariance matrix Σ_τ^{ij} at lag $\tau = 1$. (b) Histogram of eigenvalues of Σ_1^{ij}. (c) Composition of the eigenvectors (weights per sector). (d-f) Same plots for the sign covariance matrix C_τ^{ij}.

dimensional setting: What is the role of the trade sign process in shaping the cross-sectional structure of the return correlations? Is there such a thing as a market mode for signs (our proxy for liquidity)? Are there liquidity sectors? In order to further push this parallel, it is useful to define the equal-time covariance C_τ^{ij} of the cumulated trade sign process, which is analogous to Σ_τ^{ij}, defined as

$$C_\tau^{ij} = \mathbb{E}[(\mathcal{E}_{t+\tau}^i - \mathcal{E}_t^i)(\mathcal{E}_{t+\tau}^j - \mathcal{E}_t^j)] \ , \tag{4}$$

where, as for the prices, we have defined $\varepsilon_t^i = \mathcal{E}_t^i - \mathcal{E}_{t-1}^i$. Figure 1(b) displays a plot of the mean diagonal C_τ^{diag} and off-diagonal C_τ^{off} sign covariances rescaled by τ. Similarly to Σ_τ^{ii}, the diagonal terms are on average larger than the off-diagonal ones, only this time by a factor ~ 30. After a short sublinear regime, the results show superlinear time dependence at large t, consistent with the superdiffusive nature of signs in one dimension. Figure 2 shows that, in contrast with the covariance of prices, the covariance of signs displays no or very weak sectorial structure. Although the first mode of the sign covariance also corresponds to a market mode (delocalized and rather homogeneous), it is weaker. Additionally, one has a small number of "sectorial" modes out of the noise band [25], that even in this case are coherent in time, showing a time-overlap close to 1. Despite this, only the market mode is aligned with the market mode of returns. All the other modes show small overlap with their return counterparts (see Fig. 6(b) for a quantitative discussion on the fraction of common modes).

3.3. Price response

Do trades shape the return covariance matrix? Or does it result from other mechanisms such such as quote revisions, that do not involve trading volume? In order to address such questions, one needs to look into yet another quantity, the market response R_τ^{ij} defined as:

$$R_\tau^{ij} = \mathbb{E}[(X_{t+\tau}^i - X_t^i)\varepsilon_t^j] \ . \tag{5}$$

This measures the average price change of contract i at time $t+\tau$, after experiencing a sign imbalance ε_t^j in contract j. Figure 1(b) displays a plot of the mean diagonal R_τ^{diag} and off-diagonal R_τ^{off} responses. The diagonal terms are on average larger than the off-diagonal ones by a factor ~ 5. This is consistent with the ratio of the corresponding diagonal/off-diagonal factors for the price and sign covariances, and with the results of Refs. [13, 14]. The response at positive times is roughly constant, consistently with the hypothesis of an efficient price. In other words, the current sign does not predict future returns. The behavior at negative lag indicates that the current return allows a long-range prediction of the sign imbalance, an effect that has been extensively investigated in Refs. [26, 27].‡ It is worth mentioning that, other than the expected amplitude difference, the off-diagonal response shows the same temporal behaviour than its diagonal counterpart.

4. A simple model for cross-impact

In this section, we present and analyse the implications of the multivariate propagator model, which shall allow us to explain within a coherent framework the stylised facts discussed above.

4.1. The multivariate propagator model

As we shall see the simplest linear model (*i*) describing the cross-sectional structure of covariance matrices, (*ii*) accounting for their dynamical structure, and (*iii*) assuming future signs are not affected by past returns, is the multivariate *propagator model*:

$$X_t^i = \sum_{j,t'<t} G_{t-t'}^{ij}\varepsilon_{t'}^j + W_t^i \ . \tag{6}$$

This assumes that the price variations of contract i are a linear regression on the past sign imbalances of all assets j. The matrix G_τ^{ij} is customarily called the propagator, as it describes the effect of the trade sign imbalance of contract j at time t on the price of contract i at time $t+\tau$§. The quantities W_t^i are defined by $w_t^i = W_t^i - W_{t-1}^i$, where the

‡ We will disregard in the following the behavior of returns at negative lags, and only focus on the positive part of the curve, that is equivalent to assuming no price-sign correlation, that is approximately correct for small tick stocks, and breaks down at large tick stocks due to microstructural effects [26–28].

§ Note that the model is *self-consistent*, in the sense that artificially splitting the same contract i in two fully correlated instruments i_1 and i_2 yields a completely equivalent dynamics for the returns $X_t^{i_1} = X_t^{i_2}$

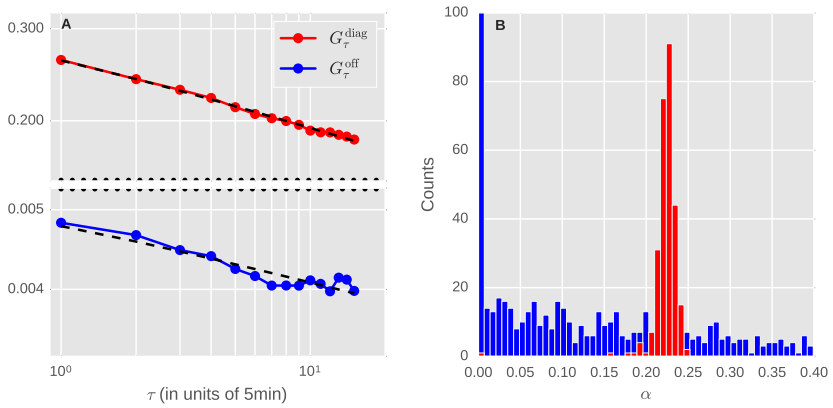

Figure 3. (a) Plot of the mean diagonal and off-diagonal propagators. (b) Corresponding histogram of fitted slopes β, as given by Eq. (9).

w_t^i are i.i.d. idiosyncratic noises with zero mean and covariance matrix given by

$$\mathbb{E}[w_t^i w_{t'}^j] = \sigma^{(W),ij}\delta_{t-t'} \,, \tag{7}$$

so that the covariance of the process W_t^i is linear in time, and is given by

$$\Sigma_\tau^{(W),ij} = \mathbb{E}[(W_{t+\tau}^i - W_t^i)(W_{t+\tau}^j - W_t^j)] = \sigma^{(W),ij}\tau \,. \tag{8}$$

We have fitted the propagator matrix G_τ^{ij} from data. Figure 3(a) displays a plot of the mean diagonal G_τ^{diag} and off-diagonal G_τ^{off} propagators that we have obtained under a non-parametric inversion of the model. See also Section 5 for a comparison of the different inversion techniques that we have adopted. The diagonal terms are on average larger than the off-diagonal ones by a factor ~ 50, see Figure 3(a). Both are consistent with a power-law decay in time, as expected from the one-dimensional case. Figure 3(b) shows fluctuations in the plot, although the slope of the diagonal components is well defined, that of the off-diagonal presents large fluctuations. However such fluctuations average away, as they seem to be structureless. More precisely, despite the large difference in magnitude between the diagonal and the off-diagonal entries of G_τ^{ij}, their time behaviour is compatible with

$$G_\tau^{ij} = G^{ij}\left(1 + \frac{\tau}{\tau_0}\right)^{-\beta} \,. \tag{9}$$

This constitutes a *factorized model* in which the temporal and cross-sectional parts are separated, and as we shall see this will facilitate the analysis by reducing the dimensionality of the problem. Fitting Eq. (9) to the diagonal and off-diagonal data yields $\beta^{\text{diag}} = 0.14$, $\beta^{\text{off}} = 0.09$, $\tau_0^{\text{diag}} = 0.30$ and $\tau_0^{\text{off}} = 0.32$. Figure 4(a) displays a plot

under any transformation of the type $\varepsilon^i = \varepsilon^{i_1} + \varepsilon^{i_2}$, provided that $G^{i_1 i_1} = G^{i_1 i_2} = G^{i_2 i_1} = G^{i_2 i_2}$, $G^{i_1 j} = G^{i_2 j}$ and $G^{j i_1} = G^{j i_2}$ for all j. This is due to our choice of extensive units for the volume. Due to our requirement of unit variance for the series of x_t^i and ε_t^i, in order to obtain consistency one obviously has to reintegrate units back into the problem.. We believe this self-consistency condition to be a necessary requirement for any satisfactory model for cross-impact.

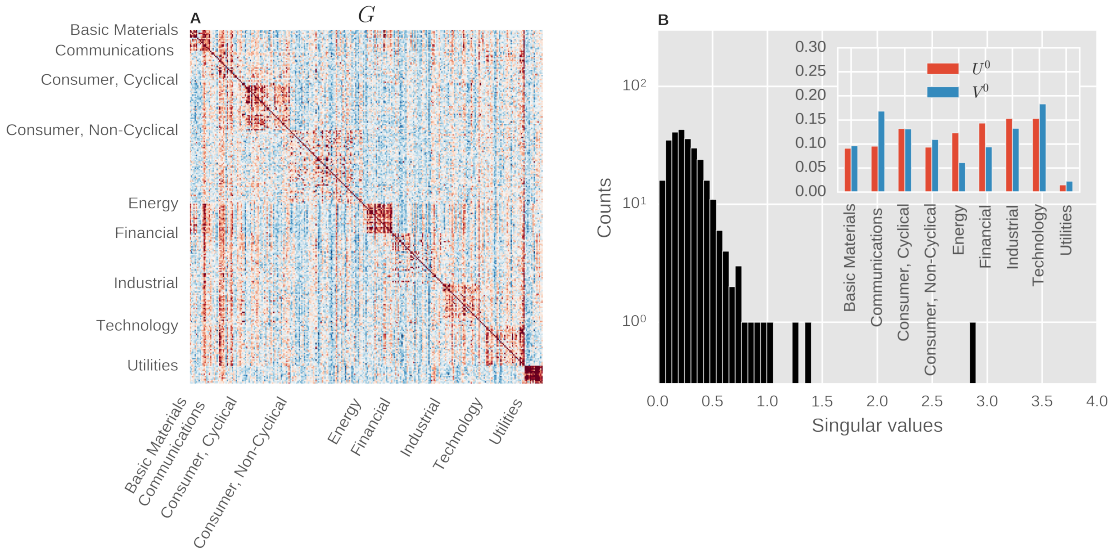

Figure 4. (a) Plot the propagator matrix G^{ij} as obtained from the factorised scheme. (b) Histogram of singular values and composition of the ground singular vectors.

of G^{ij} from which we subtracted its mean for better readability, and in which the stocks have been sorted by industrial sector, as indicated by the labels. As one can see, G^{ij} displays a stronger sectorial structure‖ than C_t^{ij}.

In order to address this issue more quantitatively, we introduce the singular-value decomposition of G^{ij}, defined as

$$G^{ij} = \sum_a U^{ia} S^a V^{ja}. \tag{10}$$

U^{ia} and V^{ia} are real orthogonal matrices, the columns of which correspond to the left/right singular vectors of G^{ij}, and where S^j is a vector made of the corresponding singular values. The interpretation of the decomposition is straightforward: For a given a the value S^a is the increase of a linear combination U^{ia} of stock prices after the combination of trades V^{ia}. Figure 4(b) displays the histogram of singular values S^a. Fig. 4(b)) shows, among other things, that a market-neutral net imbalance has a smaller impact on prices than a directional one. In fact, due to $U^{i0} \approx V^{i0} \approx N^{-1/2}$ (see the inset of Fig. 4(b)), trading one standard deviation of the imbalance of the market mode costs roughly three standard deviations of its price, while all the other modes have a consistently smaller impact.

‖ Also note the presence of vertical stripes in Figure 4(b), indicating that – while the choice of the standard deviation of returns for normalizing returns allows to obtain a homogeneous rows – using the standard deviation at $t = 1$ for the signs is not the best choice to obtain a uniform G^{ij}. Of course, this feature can be reabsorbed through a suitable definition of the units of ε_t^i.

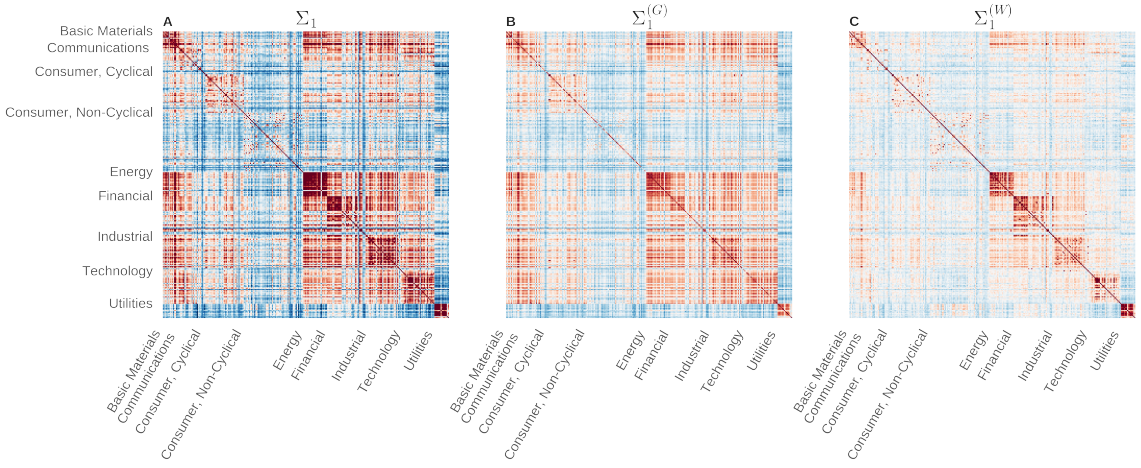

Figure 5. Plot of (a) the returns covariance matrix, (b) $\Sigma_\tau^{(G),ij}$, and (c) $\Sigma_\tau^{(W),ij}$, at lag $\tau = 1$ (see Eq. (11)). Note that the matrices having been substracted of their mean for better visibility of their structure.

4.2. Price covariation

Having found the propagator G_τ^{ij}, we can investigate its interplay with C_τ^{ij} in shaping the return correlation. In particular, within the propagator model one finds (see Appendix)

$$\Sigma_\tau^{ij} = \sum_{kl,t',t''} G_{\tau-t'}^{ik} c_{t'-t''}^{kl} G_{\tau-t''}^{jl} + \Sigma_\tau^{(W),ij} \quad = \quad \Sigma_\tau^{(G),ij} + \Sigma_\tau^{(W),ij}. \tag{11}$$

This is an extension of the result found in Ref. [12] in a linear equilibrium setting. The time-behavior of the first term $\Sigma_\tau^{(G),ij}$ captures the dynamics of the model, that is very similar to the one found in the one-dimensional model. In that case, even if at large times $c_\tau \sim \tau^{-\gamma}$ with $\gamma \approx 0.5$, the long range dependence of the resulting propagator $G_\tau \sim \tau^{-\beta}$ with $\beta \approx 0.25$ is able to compensate the long-range dependence of the imbalances and restore the diffusivity of price, which indeed requires $\beta = (1-\gamma)/2$ [1]. In this more general setting, as the time behavior of the model is found to be well-described by the factorized model (9), we are offered the same solution to reconcile the behavior of sign and return correlations. With these definitions, $\Sigma_\tau^{(G),ij}$ denotes the part of the return covariance explained by impact, while $\Sigma_\tau^{(W),ij}$ stands for its unexplained component.¶

For a visual interpretation, Fig. 5 displays plots of the three matrices that appear in Eq. (11). It can be clearly seen that the propagator model is successful in reproducing

¶ Equation (11) is a particularly robust one, as it does not explicitly depend upon the propagator. In fact, as long as the the propagator G_τ^{ij}, is compatible with the response R_τ^{ij} (see Appendix to see their relation), then $\Sigma_\tau^{(G),ij}$ depends upon the propagator *only through the response*. More precisely, one has $\Sigma_\tau^{(G),ij} = \sum_{kl,t',t''} G_{\tau-t'}^{ik} c_{t'-t''}^{kl} G_{\tau-t''}^{jl} = \sum_{kl,t',t''} R_{\tau-t'}^{ik} (c^{-1})_{t'-t''}^{kl} R_{\tau-t''}^{jl}$.

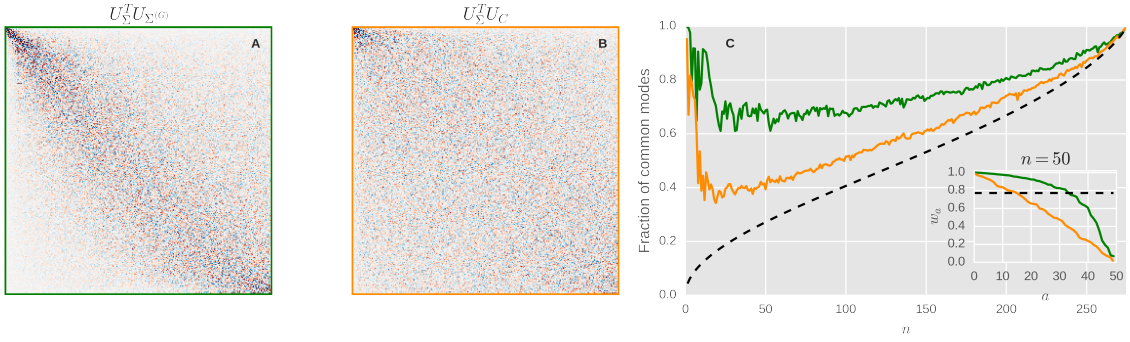

Figure 6. Plot of the overlap of eigen-rotation matrices of the returns covariance matrix with (a) $\Sigma_\tau^{(G),ij}$ (see Eq. (11)), and (b) the sign covariance matrix. (c) Plot of the fraction of common modes as defined in the text.

the sectorial structure of the covariance matrix. It is also interesting to mention that, while only $\approx 28\%$ of the diagonal variance can be explained by direct impact, this figure raises to $\approx 67\%$ for its off-diagonal counterpart, meaning that the propagator model is somewhat more efficient to explain the covariance than it is to account for the variance.

In order to assess whether the propagator model helps in understanding the direction of the risk modes of the market (and in particular, the composition of the sectors), we raise the following question: does $\Sigma_\tau^{(G),ij}$ explain more of the price covariance structure than C_τ alone? To answer this quantitatively we compare the overlap of the eigenvectors of Σ_τ with those of $\Sigma_\tau^{(G)}$ and with those of C_τ.

More precisely, if we denote the eigenvectors of these matrices by U_Σ, U_C and $U_{\Sigma^{(G)}}$, we have computed the overlap matrices $U_\Sigma^T U_C$ (see Figure 6(a)) and $U_\Sigma^T U_{\Sigma^{(G)}}$ (see Figure 6(b)). As one can see with the naked eye, the eigenmodes of Σ_τ have significantly larger overlap with $\Sigma_\tau^{(G)}$ than there is with C_τ. Figure 6(c) displays a plot that captures quantitatively the latter statement and is constructed as follows: (i) we crop each of the overlap matrices at $n \in [\![1, N]\!]$, (ii) compute their singular values $\{w_a\}_{a\in[\![1,n]\!]}$ and sort them in decreasing order (the inset shows the singular value spectra at $n = 50$), and (iii) compute the so-called fraction of common modes $\left(\prod_{a=1}^n w_a\right)^{1/n}$ and plot it as a function of n. The dashed black line represents the theoretical expectation for the noise level [25]. This measure represents the volume of the common subspace spanned by the first n eigenvectors, and is a very strict measure of similarity, which is why the results in Figure 6(c) are rather remarkable: the directional structure of the return covariance matrix can be predicted very well by only using liquidity variables!

4.3. Direct and cross-impact

Lot of covariance and part of the variance come from impact, but how to measure the direct influence of impact versus its cross-sectional component? For the response

function R_τ^{ij}, one can simply use the following relations:

$$R_\tau^{ii} = \sum_{\tau'} \Big(\underbrace{G^{ii}_{\tau-\tau'} c^{ii}_{\tau'}}_{\text{(a) (81\%)}} + \underbrace{\sum_{k\neq i} G^{ik}_{\tau-\tau'} c^{ki}_{\tau'}}_{\text{(b) (19\%)}} \Big) \tag{12}$$

$$R_\tau^{ij} = \sum_{\tau'} \Big(\underbrace{G^{ii}_{\tau-\tau'} c^{ij}_{\tau'}}_{\text{(c) (17\%)}} + \underbrace{G^{ij}_{\tau-\tau'} c^{jj}_{\tau'}}_{\text{(d) (7\%)}} + \underbrace{\sum_{k\neq i\neq j} G^{ik}_{\tau-\tau'} c^{kj}_{\tau'}}_{\text{(e) (76\%)}} \Big) , \tag{13}$$

where we have indicated the relative contribution of each term to the average self- or cross-response. The interpretation of each term in Eqs. (12) and (13).

- (a) *Self-response via direct impact.* This is the classical term considered in most works on impact: trading in product i impacts the price of i itself.
- (b) *Self-response mediated by cross-trading and cross-impact.* This term is induced by the order flow on all the stocks k that are correlated to i. This causes market makers to include this extra information in their price for i.
- (c) *Cross-response mediated by cross-trading and direct impact.* The mechanism is similar to (a), except that the order flow on j now induces an imbalance on i, that translates into a price change via direct impact.
- (d) *Cross-response mediated by direct trading and cross-impact.* Here the market markers react to the order flow on j by updating their quotes on product i.
- (e) *Cross-response market mediated by cross-trading and cross-impact.* Trading in a stock j that is correlated with a large number of other stocks k. The market maker observes the order flow all of those, and adjusts his quote of i based on this aggregate information.

Regarding the average weights of the different terms, it is interesting to notice that while most of the self-response can be explained through direct impact, this is no longer the case for the cross-response. For the latter, the dominating mechanism is (e), implying that most of the cross-response is mediated by *delocalized modes* (such as the market mode, or large sectors).

Note that the same approach can be used in order to decompose the returns covariance matrix discussed in Section 4.2 in a cross-impact component and in a direct-impact one.

4.4. Finite size scaling

As the main goal of this study is the characterization of the interactions among a large number of stocks, the fact that we only consider a sample of 275 instruments (whereas the US stock market amounts to several thousands of them), might seem counterintuitive. Such a restriction to a smaller sample implies that the order flow for all those missing products is – to us – unobserved, even though the interaction (C and

G) between stocks is on average positive. This may therefore lead to an overestimation of the magnitude of the propagators, which would then depend on system size.

In order to empirically verify this effect, we have fitted a factorized model as in Eq. (9) on many random subsets of stocks of variable size N. One would naively expect that the typical strength of direct impact propagators (G^{ij} with $i = j$) be constant regardless of N. On the other hand, cross propagators ($i \neq j$) are expected to decrease as N^{-1} in order to avoid that they dominate over direct impact when $N \to \infty$.

The results are shown in Fig. 7. We can see that indeed $\langle G^{ij} \rangle_{i=j}$ decreases only very slightly with N (fitting it to $k_1 N^{-\nu_1}$ yields $k_1 = 0.36$ and $\nu_1 = 0.04$), while we get an excellent fit of cross terms $\langle G^{ij} \rangle_{i \neq j} = k_2 (1 + \frac{N}{N_2})^{-1}$, with $k_2 = 0.06$ and $N_2 = 24$ (see dashed lines in Fig. 7).

We have also made a fit on the average of off-diagonal elements, conditioned such that i and j are in the same sector s: $\langle G^{ij} \rangle_{i \neq j; \mathrm{s}(i) = \mathrm{s}(j)} = k_3 (1 + \frac{N}{N_3})^{-\nu_3}$ gives $k_3 = 0.078$, $N_3 = 10.4$ and $\nu_3 = 0.54$. Cross-impact is naturally stronger in this case, than between two randomly selected stocks, since they are more likely to have a high correlation. Nevertheless, understanding how ν_3 should behave is more delicate, as it requires estimating how the sizes of the sectors scale in N.

5. Estimators of G: structure and statistical significance

5.1. The models

The model defined in Eq. (6) is a very general object, that is described by a propagator G^{ij}_{τ} of dimension $N^2 \times T$, plus a covariance matrix $\sigma^{(W),ij}$ of dimension N^2. Such an

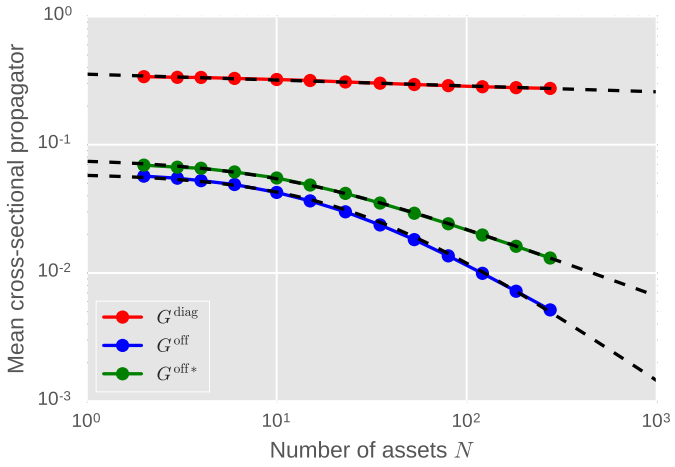

Figure 7. Plot of the mean diagonal (G^{diag}) and off-diagonal (G^{off}) propagators, as well as the mean of the off-diagonal elements where the traded and the impacted stock belong to the same sector ($G^{\mathrm{off}*}$). All curves are computed for N stocks, as a function of N, and each data point results from the average of 10^3 random subsample cross-sectional propagator inversions.

abundance of parameters results in the impossibility to estimate reliably the individual entries of G_τ^{ij} with the data in our possession. Only the aggregated statistics of G_τ^{ij} have been found statistically significant, see Fig. 4 and Table 1 below. We have thus decided of using a fully non-parametric estimation only in order to extract the main qualitative features of data. In order to estimate the interaction strengths G^{ij} and investigate their structure in a more robust fashion, we have considered lower dimensional models.

More precisely, we have used three models in order to fit the propagators and we have compared their performance:

Fully non-parametric The most general propagator model is specified the $N^2(T+1)$ parameters defining Eq. (6). This corresponds to the absence of any prior about the structure of the G_τ^{ij}.

Factorized A simpler model is obtained under the assumption $G_\tau^{ij} = G^{ij}\phi_\tau$, where ϕ_τ given by Eq. (9). The dimensionality of the model is then reduced to $2N^2 + T$.

Homogeneous The simplest, non-trivial linear model for cross-impact is obtained by assuming

$$G^{ij} = \delta^{ij} G^{\text{diag}} + (1-\delta^{ij}) G^{\text{off}}, \tag{14}$$

$$\sigma^{(W),ij} = \delta^{ij} \sigma^{(W),\text{diag}} + (1-\delta^{ij}) \sigma^{(W),\text{off}} , \tag{15}$$

so as to capture a single collective mode of the return covariance, corresponding to global market moves. The dimensionality of the model in this case is $4 + T$. The estimators for this model are reported in the Appendix, and yield $G^{\text{diag}} = 0.29$ and $G^{\text{off}} = 0.0046$, consistent with the average diagonal and off-diagonal values of the previous model (see Appendix A.2).

All these models can be calibrated by minimizing their negative log-likelihood under a Gaussian assumption for the residuals w_t^i:

$$-\ln \mathcal{L} = \frac{T}{2} \ln \det \sigma^{(W)} + \frac{1}{2} \sum_{i,j,t} w_t^i w_t^j \left(\sigma^{(W),-1}\right)^{ij} , \tag{16}$$

allowing to compute estimators for both the propagators G_τ^{ij} and the residual covariance matrix $\sigma^{(W),ij}$.[+] In this way, the estimated covariance matrix of the residual $\hat{\sigma}^{(W)}$ itself can be used in order to build metrics for the quality of the fit, and check how well the results generalize out-of-sample.

5.2. Discussion

The effort of fitting the propagator model under the different models described above can be justified by two different perspectives. On the one hand from the *statistical* point of view, it is desirable to avoid overfitting, so to have a robust model that generalizes

[+] Note that the Gaussian assumption can be relaxed, as the Generalized Method of Moments employed for example in Refs. [6, 8, 24] yields the same estimators that we have derived. Nevertheless, we choose the Gaussian assumption for the residuals in order to have closed-form results for the residuals and the likelihood function.

well out-of-sample. This enables us to predict the future covariation of prices given the imbalances. On the other hand, from the *informational* point of view, one might prefer to compress the structure of the interaction in a small number of informative parameters, rather than dealing with a larger set of more anonymous coefficients. In order to quantitatively address these points, we have chosen to inspect the behavior of the residuals and of the likelihoods in all the three models, by defining three types of scores. The first two scores assess how well one is able to describe the fluctuations along the diagonal and the off-diagonal parts of the return covariance matrix (thus, they specify particular axes of the matrix $\sigma^{(W)}$ in which we are interested:

$$\mathcal{R}^{\text{diag}} = \frac{\sum_i \hat{\sigma}^{(W),ii}}{\sum_i \sigma_0^{ii}} = \frac{1}{N}\sum_i \hat{\sigma}^{(W),ii}, \tag{17}$$

$$\mathcal{R}^{\text{off}} = \frac{\sum_{i\neq j} |\hat{\sigma}^{(W),ij}|}{\sum_{i\neq j} |\sigma_0^{ij}|}. \tag{18}$$

Alternatively, the likelihood function automatically considers the fluctuations along the eigenmodes of $\sigma^{(W)}$, as its value is uniquely fixed by the eigenvalues of the residual covariance:

$$\mathcal{R}^{\ln \mathcal{L}} = -\frac{\ln \mathcal{L}}{NT} = \frac{1}{2}\left(1 - \frac{1}{N}\log\det \hat{\sigma}^{(W)}\right). \tag{19}$$

Table 1 compares these scores for an in-sample period (2012) and an out-of-sample one (2013), in order to assess how each of the model generalizes to yet unseen data. We find that:

- All the in-sample scores improve by increasing the complexity of the models, as expected due to the fact that the models are nested. On the contrary, the good in-sample performance of the fully non-parametric model does not generalize out-of-sample. The scores displayed by the lower dimensional models are roughly the same in and out-of-sample, thus validating the practical use of the factorized and homogeneous propagator models.
- The quality in the reconstruction of the covariance of returns, measured by $\mathcal{R}^{\text{off}}$, is better than the one of the variance. While the factorized model explains around 20% of the variance, it accounts for more than 50% of the covariance. This is compatible with the findings of Ref. [10] using a model with purely permanent impact.
- While both the factorized model and the homogeneous one have good out-of-sample performance, it is interesting to notice that thes factorized model has a consistently better $\mathcal{R}^{\text{off}}$ score. This is consistent with our previous results (see Fig. 6), indicating that the directional structure of the matrix G^{ij} is statistically significant, and allows to explain a consistent part of the return covariation.

The good out-of-sample performance also indicates that heterogeneities in the temporal behaviour of the propagator discussed in Section 4.1 are weak enough for these models to generalize well across years.

Table 1. Table of scores for the three models described in Sec. 5.

	in-sample (2012)			out-of-sample (2013)		
Model	$\mathcal{R}^{\text{diag}}$	$\mathcal{R}^{\text{off}}$	$\mathcal{R}^{\ln \mathcal{L}}$	$\mathcal{R}^{\text{diag}}$	$\mathcal{R}^{\text{off}}$	$\mathcal{R}^{\ln \mathcal{L}}$
Non-parametric	0.437	0.185	0.466	2.08	1.312	1.187
Factorised	0.748	0.374	0.744	0.79	0.454	0.762
Homogeneous	0.819	0.484	0.841	0.786	0.628	0.81

6. Conclusions

The treatment of cross-impact in the existing literature has been scarce at best [10, 12–14], despite its importance – in our opinion – to correctly estimate the liquidation costs of a diversified portfolio. In this analysis we have attempted to give a more complete picture of such effects by decomposing them using a simple, linear propagator approach. Our dynamical model explains rather well the off-diagonal elements of the correlation matrix, which makes us conclude that cross-correlations between different stocks are, to a large extent, mediated by trades. Market makers learn from correlated order flow on multiple instruments, and adjust their prices on a portfolio level. This allows them to better adapt to global movements in the market, and to reduce the adverse selection they suffer. Such an observation is underpinned by the good fit of our homogeneous model, where each stock reacts to the total, net order flow of the others. This focus of market makers on their net inventory is consistent with the idea that being uniformly long or short across stocks is much more risky than to be long-short by same gross amount in a diversified fashion.

In the present study we took an empirical, descriptive point of view regarding price reaction and market maker behavior. In particular, we have disregarded the strong implications that these results have in the context of optimal execution, that will be the object of a forthcoming paper [29].

We wish to thank R. Benichou, J. Bun, A. Darmon, L. Duchayne, S. Hardiman, J. Kockelkoren, J. de Lataillade, C.-A. Lehalle, F. Patzelt, E. Sérié and B. Tóth for very fruitful discussions.

References

[1] Bouchaud J P, Farmer J D and Lillo F 2008 *How markets slowly digest changes in supply and demand* (Elsevier: Academic Press)

[2] Hasbrouck J 1988 *Journal of Financial Economics* **22** 229–252

[3] Hasbrouck J 1991 *The Journal of Finance* **46** 179–207

[4] Jones C M, Kaul G and Lipson M L 1994 *Review of Financial Studies* **7** 631–651

[5] Dufour A and Engle R F 2000 *The Journal of Finance* **55** 2467–2498

[6] Bouchaud J P, Gefen Y, Potters M and Wyart M 2004 *Quantitative Finance* **4** 176–190

[7] Lillo F and Farmer J D 2004 *Studies in Nonlinear Dynamics & Econometrics* **8**

[8] Eisler Z, Bouchaud J P and Kockelkoren J 2012 *Quantitative Finance* **12** 1395–1419

[9] Bacry E and Muzy J F 2014 *Quantitative Finance* **14** 1147–1166

[10] Hasbrouck J and Seppi D J 2001 *Journal of financial Economics* **59** 383–411

[11] Boulatov A, Hendershott T and Livdan D 2013 *The Review of Economic Studies* **80** 35–72

[12] Pasquariello P and Vega C 2013 *Review of Finance* **19** 229–282

[13] Wang S, Schäfer R and Guhr T 2016 *arXiv:1603.01586*

[14] Wang S, Schäfer R and Guhr T 2016 *The European Physical Journal B* **89** 1–16

[15] Mastromatteo I, Toth B and Bouchaud J P 2014 *Physical Review E* **89** 042805

[16] Markowitz H 1952 *The Journal of Finance* **7** 77–91

[17] Meucci A 2009 *Risk and asset allocation* (Springer Science & Business Media)

[18] Epps T W 1979 *Journal of the American Statistical Association* **74** 291–298

[19] Tóth B and Kertész J 2009 *Quantitative Finance* **9** 793–802

[20] Tóth B and Kertész J 2006 *Physica A: Statistical Mechanics and its Applications* **360** 505–515

[21] Laloux L, Cizeau P, Bouchaud J P and Potters M 1999 *Physical Review Letters* **83** 1467

[22] Marčenko V A and Pastur L A 1967 *Mathematics of the USSR-Sbornik* **1** 457

[23] Bouchaud J P and Potters M 2011 *Financial applications of random matrix theory: a short review* (Oxford University Press)

[24] Bouchaud J P, Kockelkoren J and Potters M 2006 *Quantitative Finance* **6** 115–123

[25] Bouchaud J P, Laloux L, Miceli M A and Potters M 2007 *The European Physical Journal B* **55** 201–207

[26] Taranto D E, Bormetti G, Bouchaud J P, Lillo F and Tóth B 2016 *arXiv:1602.02735*

[27] Taranto D E, Bormetti G, Bouchaud J P, Lillo F and Tóth B 2016 *arXiv:1604.07556*

[28] Dayri K and Rosenbaum M 2015 *Market Microstructure and Liquidity* **1** 1550003

[29] Mastromatteo I, Benzaquen M, Eisler Z and Bouchaud J P 2016 (in preparation)

[30] Sérié E 2010 Unpublished report, Capital Fund Management, Paris, France

[31] Eisler Z, Bouchaud J P and Kockelkoren J 2011 *Available at SSRN 1888105*

Appendix A. Models

It is important to mention that while the results below are presented with integrated response functions R_τ^{ij} and propagators G_τ^{ij}, all propagator inversions have been done with differential response functions r_τ^{ij}. This is consistent with the idea that such quantities have a decaying asymptotic behaviour in contrast with their integrated counterparts and thus suffer less from the cut-off effect [30, 31]. The integrated propagator was then computed by using the relation $G_\tau^{ij} = \sum_{\tau'} g_{\tau'}^{ij}$. Figure A1 displays the diagonal and off-diagonal means of the lagged sign correlation function, the lagged return correlation function and the differential response function.

Appendix A.1. Fully non-parametric model

The inversion of the model in the fully non-parametric case yields a well-known matrix equation for the propagator:

$$\hat{R}_\tau^{ij} = \sum_{\tau',k} \hat{G}_{\tau'}^{ik} c_{\tau-\tau'}^{kj} \,. \tag{A.1}$$

The estimator of the residuals is also straightforward to compute:

$$\hat{\sigma}^{(W),ij} = \frac{1}{T} \sum_t w_t^i w_t^j \,. \tag{A.2}$$

The total number of parameters to estimate under this method is $N^2(T+1)$, while the computational bottleneck results from the inversion of the block-Toeplitz matrix c_τ^{ij} appearing in Eq. (A.2).

Appendix A.2. Factorized model

The assumption of a propagator of the form

$$G_\tau^{ij} = G^{ij} \phi_\tau \,, \tag{A.3}$$

where ϕ_τ is given by Eq. (9), results in a simpler estimation of $N^2 + T$ parameters for the kernel and N^2 parameters for the residuals. The estimator for the propagator is found by solving:

$$\hat{G}^{ij} = A \left(B^T \right)^{-1} \,, \tag{A.4}$$

where one has preliminarily defined:

$$A^{ij} = \sum_\tau \hat{R}_\tau^{ij} \phi_\tau \tag{A.5}$$

$$B^{ij} = \sum_{\tau,\tau'} \hat{c}_{\tau-\tau'}^{ij} \phi_\tau \phi_{\tau'} \,, \tag{A.6}$$

whereas the estimator of the variance is given by the earlier expression (A.2).

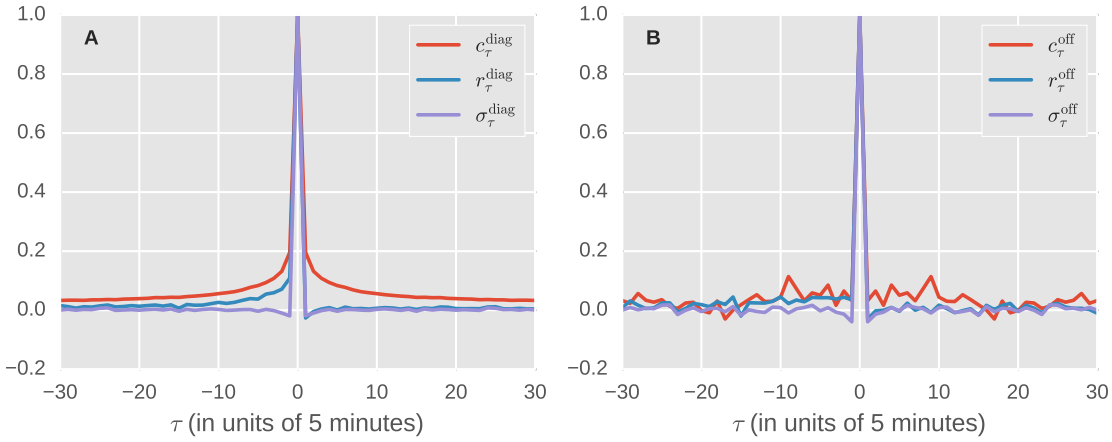

Figure A1. Plot of the diagonal (a) and off-diagonal (b) means of the lagged sign correlation function, the returns lagged correlation function and the differential response function.

Appendix A.3. The homogeneous model

The estimator for the propagator reads

$$\hat{G}^{\text{diag}} = \frac{1}{N}\left(\frac{A^{\text{M}}}{B^{\text{M}}} + (N-1)\frac{A^{\text{M}} - A^{\text{I}}}{B^{\text{M}} - B^{\text{I}}}\right) , \tag{A.7}$$

$$\hat{G}^{\text{off}} = \frac{1}{N}\left(\frac{A^{\text{M}}}{B^{\text{M}}} - \frac{A^{\text{M}} - A^{\text{I}}}{B^{\text{M}} - B^{\text{I}}}\right) , \tag{A.8}$$

where one has preliminarily defined the market (M) and idiosyncratic (I) means:

$$A^{\text{M}} = \mathbb{E}[A^{ij}] = \frac{1}{N^2}\sum_{ij} A^{ij} , \tag{A.9}$$

$$A^{\text{I}} = \mathbb{E}[A^{ii}] = \frac{\text{Tr}[A]}{N} , \tag{A.10}$$

and equivalently for B^{M} and B^{I}. The estimator of the variance is given by (we give the inverse of the estimator for simplicity):

$$\left(\hat{\sigma}^{(W)}\right)^{-1,\text{diag}} = \frac{1}{N}\left(\lambda^{\text{M}} + (N-1)\lambda^{\text{I}}\right) , \tag{A.11}$$

$$\left(\hat{\sigma}^{(W)}\right)^{-1,\text{off}} = \frac{1}{N}\left(\lambda^{\text{M}} - \lambda^{\text{I}}\right) . \tag{A.12}$$

We have also introduced

$$\lambda^{\text{M}} = \frac{1}{N}\left[\mathbb{E}[\sigma_0] - \frac{A^{\text{M}2}}{B^{\text{M}}}\right]^{-1} , \tag{A.13}$$

$$\lambda^{\text{I}} = \frac{N-1}{N}\left[1 - \mathbb{E}[\sigma_0] + \frac{\left(A^{\text{M}} - A^{\text{I}}\right)^2}{B^{\text{M}} - B^{\text{I}}}\right]^{-1} . \tag{A.14}$$