Chaos

Information dynamics of price and liquidity around the 2017 Bitcoin markets crash

Cite as: Chaos **32**, 043123 (2022); https://doi.org/10.1063/5.0080462
Submitted: 01 December 2021 • Accepted: 15 March 2022 • Published Online: 18 April 2022

Vaiva Vasiliauskaite, Fabrizio Lillo and Nino Antulov-Fantulin

Information dynamics of price and liquidity around the 2017 Bitcoin markets crash

Cite as: Chaos **32**, 043123 (2022); doi: 10.1063/5.0080462
Submitted: 1 December 2021 · Accepted: 15 March 2022 ·
Published Online: 18 April 2022

Vaiva Vasiliauskaite,[1,a)] Fabrizio Lillo,[2] and Nino Antulov-Fantulin[1]

AFFILIATIONS

[1]Computational Social Science, Swiss Federal Institute of Technology in Zürich (ETH), CH-8092 Zürich, Switzerland
[2]Dipartimento di Matematica, Università di Bologna, IT-40126 Bologna, Italy and Scuola Normale Superiore, Pisa IT-56126, Italy

[a)]**Author to whom correspondence should be addressed:** vvasiliau@ethz.ch

ABSTRACT

We study information dynamics between the largest Bitcoin exchange markets during the bubble in 2017–2018. By analyzing high-frequency market microstructure observables with different information-theoretic measures for dynamical systems, we find temporal changes in information sharing across markets. In particular, we study time-varying components of predictability, memory, and (a)synchronous coupling, measured by transfer entropy, active information storage, and multi-information. By comparing these empirical findings with several models, we argue that some results could relate to intra-market and inter-market regime shifts and changes in the direction of information flow between different market observables.

Cryptocurrencies are a novel financial instrument, whose uniqueness lies in a novel distributed ledger technology that serves as a public database of executed transactions. They are also characterized by high price fluctuations, price bubbles, and sudden price crashes. Cryptocurrencies can be traded (sold and bought) at many independently operating venues (exchange markets); however, the price of a cryptocurrency eventually synchronizes as the traders take advantage of mismatches between prices, as observed by a single asset in several markets, or exploiting several assets within one market. Therefore, the universal price of a cryptocurrency is also determined in a distributed fashion, depending on trading decisions within each individual market and trades that occur in between markets. To understand the importance of different types of dynamics that come into play while determining the next price of a cryptocurrency, we analyze temporal patterns of information flow within and across markets. In particular, we analyze minute-level market data that would capture price discrepancies among markets and allow us to detect any potentially anomalous information flows during a turbulent market event such as the Bitcoin price crash that occurred in 2017.

Bitcoin is a cryptocurrency that was originally designed as a medium of exchange;[1] however, there is still no strong consensus as to whether it is a currency, a commodity, or an asset.[2–4] Cryptocurrencies are usually transmitted and created via distributed peer-to-peer networks with well-defined cryptographic protocols that record the system's state via a public ledger (blockchain). Thus, blockchain (or another type of ledger that maintains knowledge of distributed consensus) is at the heart of digital currencies. Blockchain technology promises benefits such as proving the existence of an asset as well as keeping track of its current and all past ownerships, both in a distributed manner. It also plays a role in crypto-asset price discovery.[5]

Electronic coins, such as Bitcoin, are "chains" of digital signatures: an owner transfers the coin by digitally signing a hash of the previous transaction and the public key of the next owner, adding them to the end of the coin. The exchange of cryptocurrencies to fiat money [U.S. dollar (USD), Euro (EUR), and Great British Pound (GBP). etc.] occurs in cryptocurrency exchange markets, which are based on electronic double auctions, operating using limit order books. Here, real-time market data, such as transaction volume, amount of bid and ask orders, and exogenous information, such as news and social media mentions, are sources of information that feed into trading decisions. Within the crypto-market ecosystem, two distinguishable types of information sources exist: intra-market sources and inter-market sources. The first source is related to internal dynamics of a market's limit order book, e.g., its

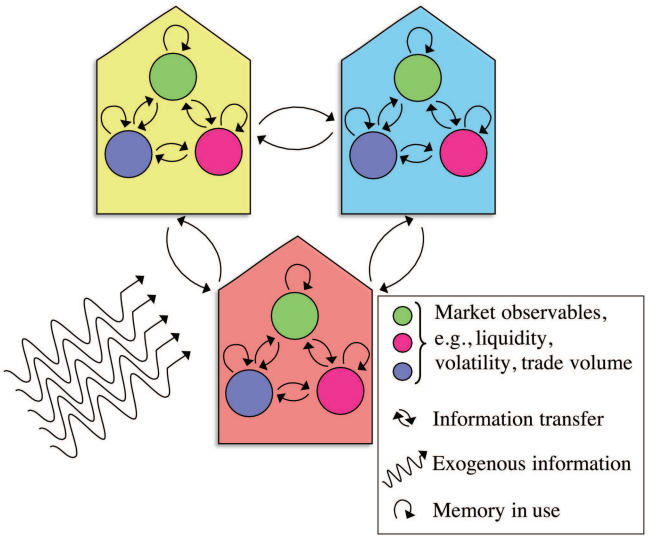

FIG. 1. Information is shared between markets (house icons) and within markets. Different nodes indicate separate market observables, discussed in Sec. I. Past information about one market observable (e.g., volatility) can be used by the same market observable (a self-loop), as well as by other market observables (edges between different color nodes in the same market). Information is also shared across markets (edges among market icons), and the system as a whole is also exposed to unknown exogenous information (curly arrows).

liquidity and volatility, which may influence trading decisions[6] as well as one another.[7] The second source is related to communication between different markets. Therefore, any two crypto-markets are "connected:" explicitly, when there is a mutually traded currency (or an arbitrage opportunity) or implicitly, when prices of currencies among several markets are correlated.

The blockchain ensures that, in the long run, the price of a cryptocurrency develops synchronously across exchanges, i.e., the law of one price holds.[5] However, even without the presence of blockchain, one would expect price synchronization across different venues due to arbitrage.[8] Hence, the Bitcoin price is processed collectively in a distributed manner. Its dynamics is affected both by internal feedback and by exogenous information (e.g., public news), as illustrated in Fig. 1. It is not clear, however, what type of information is the most relevant for the price formation process. Often, these multiple influences lead to very large fluctuations (high volatility in technical terms) of exchange rates between cryptocurrencies and fiat money, price bubbles, and sudden price crashes. These properties make cryptocurrencies exchange rates a unique laboratory to empirically study the collective dynamics leading to market instabilities, which are well known to be ubiquitous in all financial markets.

Although at low frequencies, the prices of Bitcoin in different markets (and, oftentimes, prices of other cryptocurrencies in relation to Bitcoin) develop in apparent synchrony, the law of one price does not hold at very high frequencies.[8,9] Furthermore, at sufficiently high frequency, one observes lagged relationships between the prices of the asset(s), i.e., past information about the change in the price of one cryptocurrency is informative for predicting the future price change of another cryptocurrency. Such mechanisms are detectable with information transfer measures: linear relationships can be detected via Granger causality,[10] while more general nonlinear dependencies can be detected via transfer entropy,[11,12] that is one of constituents of a process's entropy. Entropy itself relates to computation: at each point in time, *computation* of the next state of the process. When a system is composed of multiple interacting units, information transfer could be thought of as communication or signaling and the system as a whole as processing information to determine its collective behavior at each time step via distributed computation. A study of *information dynamics* aims to decompose this computation into unique elements, namely, transferred, stored, and modified information, and their changes in space and time.[13,14]

Spatiotemporal patterns of information dynamics (in particular, information transfer and multi-information) within the system of interacting markets were observed to increase in the financial crisis periods, meaning that the system appears to be more synchronized.[15] Such synchronization of a system could be a precursor of a "phase transition"[16,17]—a dramatic dynamical shift that occurs due to exogenous or endogenous events that perturb the system. Inefficiencies and delays of transfer of information among system's constituents open possibilities for arbitrage, risk, and distress spillover across markets. They may also be precursors of price bubbles. Although the information-theoretic measures have been shown to signify dramatic dynamical changes in the system, both where the phase transition can be pinpointed exactly[18,19] as well as when they are discussed in a more qualitative manner,[20] the observed dynamical patterns are difficult to interpret. This is particularly challenging if the data are incomplete, unrepresentative, noisy, e.g., sampled at a low,[21] inconsistent frequency, with not all relevant system's constituents taken into account, using erroneous measurement tools.

Therefore, in this paper, our aim is twofold. First, we study several econometric models that couple market microstructure variables. Our aim is to find out whether a particular coupling and its (sudden or slow) change are detectable with information dynamics tools. We then analyze the persistence of information dynamical patterns that signify a particular regime shift. It is important to clarify that in the following we are not attempting to model regime shift dynamics where the different regimes and the transitions between them must be inferred from the data. On the contrary, based on the observed price dynamics of Bitcoin described below, we test the hypothesis that the periods before and after the price peak are characterized by different statistical and information dynamic properties, suggesting the presence of two distinct regimes.[22] Specifically, we investigate several financial variables across markets during one turbulent event in the crypto-market ecosystem, namely, the Bitcoin bubble in 2017–2018. We remind that the bubble occurred in December 2017 when the Bitcoin price went up from around 6463$ on the 1st of November, 2017 to, at that time, an all-time-high of 19 716$ on the 17th of December, see the left panel of Fig. 2. After this crash date, the price dropped to 11 414$ before the end of the year, descending for weeks and months afterward. Furthermore, ripple effects of the Bitcoin bubble also penetrated other cryptocurrencies, and much research has been put into studying the drivers of the crypto crash, e.g., see Ref. 23. This price bubble is also unique in that in 2017, Bitcoin was still a dominant cryptocurrency with the

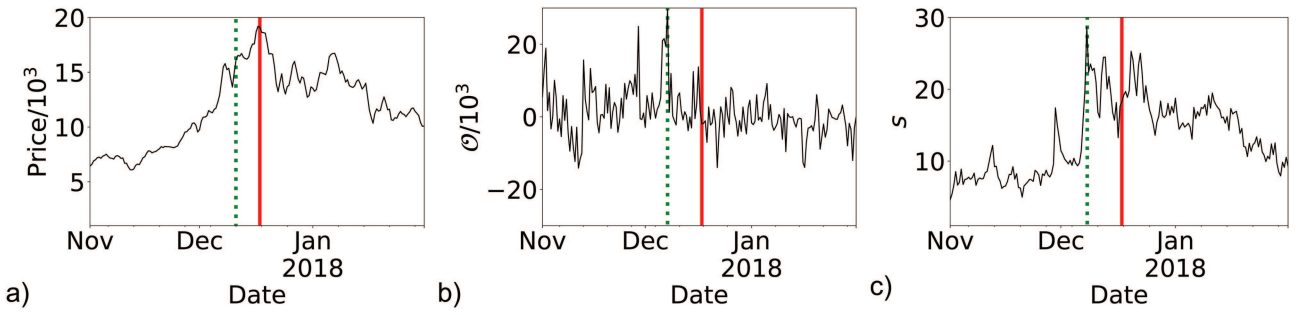

FIG. 2. (a) Average mid-price of Bitcoin (black). (b) Realized order imbalance expressed in quote currency, aggregated across 12-h windows and all considered trading venues (black). (c) Average 12-h spread (black). Red lines indicate the peak of the bubble, and the green dotted lines indicate the dates at which each observable reached the maximum value (for price, we show the date of the maximum value of the price returns). From left to right, each dotted green line indicates the following dates: December 10, 2017; December 7, 2017; December 8, 2017. The red line is positioned at December 17, 2017.

largest market capitalization. At the time of writing, alternative coins have become more prominent. Assuming that high-frequency market data related to Bitcoin traded against dollar and Tether captures the majority of the dynamics that took place at that time, we concentrate our attention on information dynamics within and across markets using this particular set of data.

The paper is organized as follows. First, in Sec. I, we describe the data used, as well as market microstructure observables, related to price as well as liquidity in markets, that we use for the inference of information dynamics measures in Sec. I. Section II includes a description of information-theoretical tools, an experimental setup for data analysis, and analysis of models. We report the data analysis results in Sec. III, and conclude the paper with a discussion in Sec. IV.

I. DATA

The data for this paper are tick-level trading and order book data obtained from Ref. 24. Data of trades are provided at a millisecond frequency, and for observables based on this data, we aggregate the information to a minute-level frequency. The limit order book data consists of one snapshot in each minute. Since the snapshots are not taken at exactly the same second of a minute in different markets, we align the data from the limit order books and from trades to ensure the lack of non-causal information flows. We describe this procedure in Appendix A. Note that the alignment procedure ensures that causal order of events is respected when we treat snapshots of a limit order book as a discrete-time process that is studied at a frequency that is at most 1 min.

For the proceeding analysis, we restrict our attention to the most liquid and the largest cryptocurrency—Bitcoin (BTC), traded against either a fiat currency of U.S. dollar (USD) or Tether (USDT), a *stablecoin*, designed to be worth $1.00 at all times. More specifically, we consider Bitcoin traded against USD in the following venues: Gemini, BTC-e, Bitstamp, Coinbase, Kraken, HitBTC, Bitfinex, and against USDT in Binance, Bittrex, and Poloniex. The time period under our study ranges from November 1, 2017 to February 1, 2018. This period involves a price bubble observed in Bitcoin as well as other cryptocurrencies. The left panel of Fig. 2 shows the price dynamics.

A. Microstructural variables

In order to study the price and liquidity dynamics in the different venues, we introduce three market microstructure variables.

a. *Price* quantifies the value of an asset. Here, we define price as the mid-price at each snapshot of the order book, namely, $p_t = \frac{p_t^a + p_t^b}{2}$, where p^a, p^b are, respectively, the best ask and bid price at time t. We define the price increments (hereafter termed returns for simplicity) as $r_t = p_t - p_{t-1}$. Price dynamics is strongly asymmetric around market crashes and shows a slow (power law) relaxation of volatility.[25]

b. *Order imbalance* quantifies the demand of liquidity takers. Specifically, suppose we have a set of executed trades, indexed over arbitrary integer index $i : i \in \mathbb{Z}$, and each trade also has an associated time stamp $t_i \in \mathbb{R}^+$, a sign $\epsilon_i \in \{-1, 1\}$, and a volume $v_i \in \mathbb{R}^+$. If the trade is initiated by a buyer, then its sign is $\epsilon_i = +1$, while $\epsilon_i = -1$ for seller-initiated trades. We define order imbalance for the time interval $(t - \delta, t]$ as a sum of the signed volumes of all trades within the interval

$$\mathcal{O}_t = \sum_{i|t-\delta<t_i\le t} \epsilon_i v_i. \tag{1}$$

Similarly, $\mathcal{O}$ can be expressed in quote currency: $\mathcal{O}_{\$,t} = \sum_{i|t-\delta<t_i\le t} \epsilon_i v_i p_i$ if p_i is the transaction price. For the relation between order flow and price returns, see, for example, 26 and 27.

c. *Spread* is a symmetric measure of market's liquidity (as small spread indicates that trades are easily executable) and is defined as $s_t = p_t^a - p_t^b$. Spread has been shown to display nonsymmetric dynamics around market crashes.[28]

In Fig. 2, we show the dynamics of price, order imbalance, and spread for time period from November 1, 2017 to February 1, 2018.

II. METHODS

A. Information dynamics for stochastic processes

Consider a system, composed of N stochastic processes $\{X^\alpha\}_{\alpha\in\Omega}$, together forming a multivariate process $\mathbf{X}^\Omega$. Each stochastic process X^α is a collection of random variables $\{X_t^\alpha\}_{t\in\mathbb{N}^+}$ with Q being the total number of observations. For each random variable X_t, its realized value is defined as x_t. Note that Ω denotes

the finite set of markets in this paper. In the definitions of information dynamics measures, we will use subscripts X, Y to refer to information dynamics among some two stochastic processes (e.g., $X = X^{\alpha}$, $Y = X^{\beta}$). These information dynamics measures will be defined using length-l and length-k collections of random variables, for each time t: $\mathbf{Y}_{t-\delta}^{(l)} = \{Y_{t-\delta-l}, Y_{t-\delta-l+1}, \ldots, Y_{t-\delta}\}$, $\mathbf{X}_{t-1}^{(k)} = \{X_{t-1-k}, X_{t-1-k+1}, \ldots, X_{t-1}\}$, where δ is a time delay.

a. Information-theoretic measures for static variables. The fundamental quantity in this work is the Shannon entropy of a random variable X, defined as $H(X) = -\sum_x p(x) \log p(x)$. Here, x is an instance of a random variable, and the sum is over of all possible values that x can take. Conditional entropy of X, given Y is the average uncertainty that remains about X after learning the values of Y: $H(X|Y) = -\sum_{x,y} p(x,y) \log p(x|y)$. Mutual information between X and Y measures the average amount of information that is communicated in one random variable about another. Y: $I(X,Y) = H(X) - H(X|Y)$. The conditional mutual information between X and Y when Z is known is defined as $I(X,Y|Z) = H(X|Z) - H(X|Y,Z)$.

These information-theoretic measures for static variables can also be adapted to analyze stochastic processes. When temporal information is incorporated, one can quantify: how much information is shared between system's units at each point in time; how much information about the current state of X is conveyed in the past states of Y; and how much information about the current state of X is conveyed in the past states of X. These questions can be addressed with information dynamics measures[13,14] that are summarized below.

b. Multi-information. Multi-information (MI) is defined as a measure of the deviation from the independence of the components in the system[14,29]

$$I_{\mathbf{X}^{\Omega}} = \left(\sum_{\alpha=1}^{N} H(X_t^{\alpha})\right) - H(\mathbf{X}_t^{\Omega}). \tag{2}$$

Large values of MI are a signature of high synchronous interconnectivity of a system.

c. Transfer entropy. Transfer entropy (TE) encapsulates the "distributed nature" of computation. Schreiber,[11] and, independently, Paluš *et al.*[12] defined TE as the amount of information that a source process Y, and, in particular, its past state $\mathbf{Y}_{t-\delta}^{(l)}$ provides about a target's state X_t in the context of the target's immediate past state $\mathbf{X}_{t-1}^{(k)}$[13]

$$\begin{aligned} T_{Y\to X}^{(k,l,\delta)} &= I(\mathbf{Y}_{t-\delta}^{(l)}; X_t | \mathbf{X}_{t-1}^{(k)}) \\ &= H(X_t|\mathbf{X}_{t-1}^{(k)}) - H(X_t|\mathbf{X}_{t-1}^{(k)}, \mathbf{Y}_{t-\delta}^{(l)}). \end{aligned} \tag{3}$$

The superscript indicates the parameters used for the delay δ as well as the lengths l and k of collections of the random variables. This formulation of information transfer quantifies pairwise relationships between variables. When a system is composed of more than two stochastic processes, (3) is known as *apparent transfer entropy*, as it does not account for unobserved sources, potentially leading to over-estimation of the total amount of entropy transferred within the system.

d. Conditional transfer entropy. To discount redundant joint influences of two sources Y, Z on a target X, (3) is generalized to a measure of transfer entropy from Y to X given that Z can also provide information about X,

$$\begin{aligned} T_{Y\to X|Z}^{(k,l,m,\delta)} &= I(\mathbf{Y}_{t-\delta}^{(l)}; X_t | \mathbf{X}_{t-1}^{(k)}, \mathbf{Z}_{t-1}^{(m)}) \\ &= H(X_t|\mathbf{X}_{t-1}^{(k)}, \mathbf{Z}_{t-1}^{(m)}) - H(X_t|\mathbf{X}_{t-1}^{(k)}, \mathbf{Y}_{t-\delta}^{(l)}, \mathbf{Z}_{t-1}^{(m)}). \end{aligned} \tag{4}$$

Finally, the conditional TE can be generalized to account for a multivariate set of processes with the next definition.

e. Collective transfer entropy. To find the total amount of information that was collectively transferred to a target X from all potential sources in the system, we use *collective TE*, T_X, which accounts for redundancies and synergies among information that sources provide to a target.[14] To compute T_X, we consider a set $\mathbf{X}^{\Omega}$. The true effect of one source variable Y to the target variable X in the universe of $\mathbf{X}^{\Omega}$ is computed through conditional transfer entropy: $T_{Y\to X|\mathbf{X}_X^{\Omega\setminus\{Y\}}}$, where $\mathbf{X}_X^{\Omega\setminus\{Y\}} = \{Z \in \mathbf{X}^{\Omega}\setminus\{X,Y\}\}$. However, $\sum_Y T_{Y\to X|\mathbf{X}_X^{\Omega\setminus\{Y\}}}$ would not be equal to the total amount of information transferred to X, as only unique information[30] would be accounted for. To calculate the collective TE (see Ref. 14 for more details), we sum incrementally conditional TE terms. Let us consider an ordered set $\mathbf{X}_X^{\Omega} = \{Z \in \mathbf{X}^{\Omega}\setminus\{X\} : Z^1, \ldots, Z^{\beta}, \ldots, Z^{N-1}$ and its subset $\mathbf{X}_X^{\Omega,\beta} = \{Z^{\alpha} \in \mathbf{X}_X^{\Omega} | \alpha \le \beta\}$. Collective transfer entropy is then defined as

$$T_X = \sum_{\beta=1}^{N-1} T_{Y^{\beta}\to X|\mathbf{X}_X^{\Omega,\beta-1}}. \tag{5}$$

f. Active information storage. Overall, the predictability of the next state of a process X is characterized by its entropy, whose non-overlapping constituents are the collective transfer entropy T_X, the *active information storage* (AIS), $A_X^{(k)}$, defined as *memory of the process that is actively in use*,[14]

$$A_X^{(k)} = I(\mathbf{X}_{t-1}^{(k)}; X_t) = H(X_t) - H(X_t|\mathbf{X}_{t-1}^{(k)}). \tag{6}$$

AIS is information in the past that contributes to the computation in the next state of X.

g. Local information dynamics. The measures defined above can also be considered in a point-wise fashion as they are expectation

values of local measures at each observation at time τ as follows:[14,18]

$$I_{\mathbf{X}^\Omega} = \langle i_{\mathbf{X}^\Omega}(\tau)\rangle_\tau = \frac{1}{Q}\sum_{\tau=1}^{Q}\log\frac{p(x_\tau^1, x_\tau^2, \ldots, x_\tau^N)}{\prod_{\alpha=1}^{N} p(x_\tau^\alpha)},$$
$$T_{Y\to X}^{(k,l,\delta)} = \langle t_{Y\to X}^{(k,l,\delta)}(\tau)\rangle_\tau = \frac{1}{Q}\sum_{\tau=1}^{Q}\log\frac{p(x_{\tau+1}|\mathbf{x}_\tau^{(k)}, \mathbf{y}_\tau^{(l)})}{p(x_{\tau+1}|\mathbf{x}_\tau^{(k)})}, \quad (7)$$
$$A_X^{(k)} = \langle a_X^{(k)}(\tau)\rangle_\tau = \frac{1}{Q}\sum_{\tau=1}^{Q}\log\frac{p(x_{\tau+1}, \mathbf{x}_\tau^{(k)})}{p(x_{\tau+1})p(\mathbf{x}_\tau^{(k)})}.$$

B. Experimental setup

a. Information dynamics. We use Java Information Dynamics Toolkit—JIDT[31] to compute all information dynamics quantities. For the TE estimator, we choose Kraskov, Stögbauer, Grassberger (KSG) K-nearest neighbor estimator,[32] which is minimally parameterized and has been shown to be robust for a wide range of data. To ensure our estimates are reliable, we studied sample-size bias of transfer entropy using the approach of Ref. 33 as described in Appendix C. We found that $K = 4$ for nearest neighbors estimator produces reliable results for our dataset. All values of information dynamics are reported in units of nats.

Throughout the paper, we considered the case where $l = 1$. To ensure that we do not over-estimate information transfer for *active information storage*—the memory of a process in use—of X, we chose $k = \underset{\kappa\in[1,60]}{\text{argmax}}(A_X^{(\kappa)})$ for results in Sec. III A and $k = \underset{\kappa\in[1,10]}{\text{argmax}}(A_X^{(\kappa)})$ for results in Sec. III B. We then used this k when computing $T_{Y\to X}$ as well as T_X. Lastly, as was demonstrated in Ref. 34, when two processes are coupled via non-zero delay u, $T_{Y\to X}^{(\delta)}$ is maximized for $\delta = u$. Therefore, we consider $T_{Y\to X}^{(\delta)}$ for $\delta \in [1, 10]$ and select δ for which $T_{Y\to X}^{(\delta)}$ is maximal. Such a procedure ensures that non-instantaneous coupling across markets is also captured. For the results in Sec. III B, we do not do this and consider a fixed $\delta = 1$.

b. Rolling windows. For results discussed in Sec. III A, we considered weekly subsets of the time series, moving each sliding window by 3 days with respect to the previous one. In such a way, we achieve a twofold advantage. First, we reveal temporal patterns in information dynamics. Second, we ensure that the data are locally stationary (we expect the time series associated with markets around the crash to be globally non-stationary; however, at sufficiently small time windows intervals, we observe stationary time series). In Sec. III B, we considered two non-overlapping time series windows: the data preceding the price peak date (December 17, 2017) and the remainder.

c. TE significance test. All reported results are significant, where significance is tested against the null hypothesis that no information transfer exists. For all results, we used 100 surrogate time series to evaluate significance, and, unless otherwise stated, we chose the significance level of 0.05. The surrogate time series are generated by permuting time indices of appropriate time series, see Ref. 31 for details. For the results of Sec. III A in each time window, a Benjamini–Yekutieli procedure[35] was also performed.

d. Stationarity test. For each observable's time series in all rolling windows, we considered the Augmented Dickey–Fuller (ADF) test[3637] whose null hypothesis is a presence of a unit root, and we used significance level of 0.05. To determine the lag length, we used an information minimization criterion (AIC, in particular) and considered values from a range $[0, 1, \ldots, 38]$. We found that we can reject the null hypothesis that there is a unit root with a 5% significance level for the test statistics for all of the time windows of s, $\mathcal{O}$ time series. Contrary, we found the presence of unit roots in price time series, so we used price returns (i.e., equal to applying difference operation $r_t = p_t - p_{t-1}$), after which all time windows passed the ADF test.

C. Statistics of information dynamics

In the data analysis part, we will consider information dynamics, described in Sec. II, first using one of the observables for each market. In doing so, we assume that the strongest coupling occurs through the same observable (spread in α has a stronger relation to spread in β than to price returns in β), see Fig. 3 for an illustration. Second, we also consider the extent to which market variables are internally coupled via different observables (e.g., quantifying how much spread in α has an effect on price returns in α). See Fig. 4 for an illustration. To interpret the results, we aggregate information dynamics measures into several spatiotemporal global and local metrics.

For each information dynamics metric, namely, MI, AIS, and TE, the argument w denotes the time series window, based on which the reported values are computed. We will also denote one microstructure observable (e.g., s), with the same capital letter that defines a random variable for a market, e.g., $T_{X^\alpha\to X^\beta}$ denotes transfer entropy from a market α to a market β considering the same market microstructure observable while $T_{X^\alpha\to Y^\beta}$ would indicate that different microstructure observables are considered.

a. Global metrics. To quantify the global, system-describing information dynamics, we aggregate results obtained for all markets within a time series window. In particular, in Sec. III A we study the extent to which a microstructural variable in one market affects the same microstructural variable in another market. Therefore, the extent of this effect is captured with a sum of transfer entropies (3) between all pairs of markets, namely, the *total apparent transfer entropy*

$$T_X^{\text{app,sys}} = \sum_{\alpha,\beta|\alpha\neq\beta} T_{X^\alpha\to X^\beta} \quad (8)$$

and the *total collective transfer entropy*

$$T_X^{\text{coll,sys}} = \sum_{\alpha} T_{X^\alpha}. \quad (9)$$

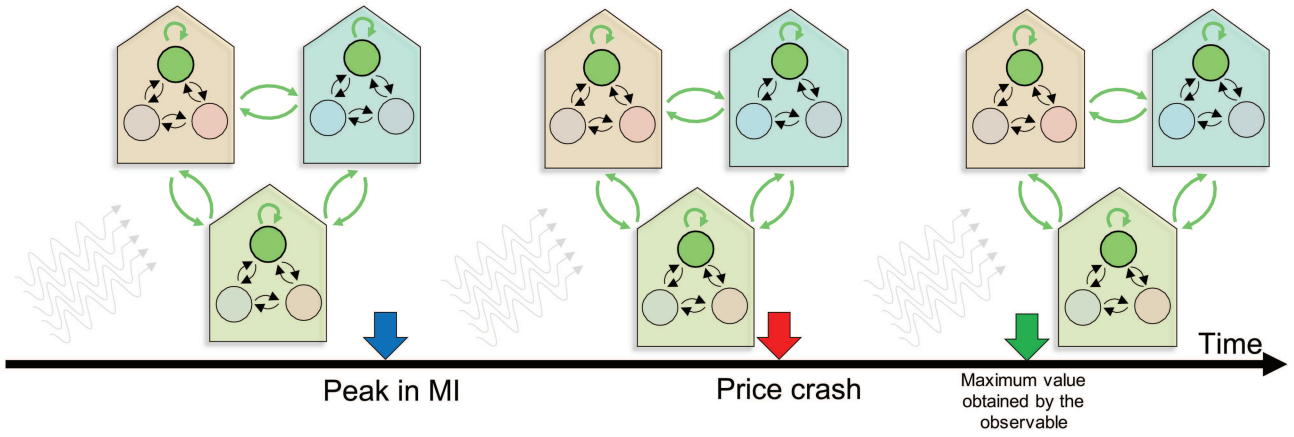

FIG. 3. Illustration of analysis described in Sec. III A. We choose one observable (illustrated with green node) for each market. We then divide the observable's time series into overlapping sliding windows and compute information dynamical quantities (MI, TE) between markets using this one observable. The connectivity between markets via green observable's "channel" is illustrated with green edges linking different markets. We also compute the AIS for each market; this is illustrated with a self-loop for a green node. Given a set of individual market's, or pairs of markets information dynamics quantities, we then aggregate these quantities into global statistics for each time window. The timeline illustrates three critical points: the price peak date, the date when we observe the maximum value obtained by the (green) observable, and the peak in MI for a given same observable. In Fig. 7, these points are shown in vertical lines of corresponding colors.

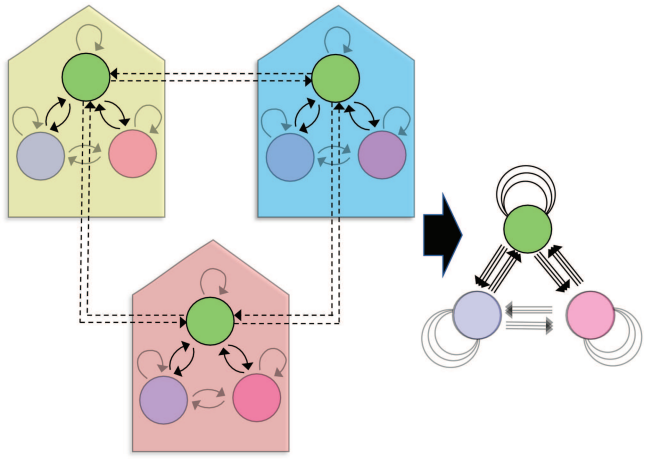

FIG. 4. Illustration of analysis described in Sec. III B. Each market is characterized by a set of market microstructure observables (nodes of the same color illustrates the same observable). In the section, we discuss the information transfer between observables. TE between the same observable is illustrated as an edge between nodes of the same color and results into self-loops in a resulting multigraph where nodes represent microstructure observables. We allow for multiple loops where each edge represents a link based on a single pair of markets. TE between different observables is illustrates as the flow from one color nodes to another color nodes. In a resulting directed multigraph, these connections are edges between nodes. We again allow for multiple edges where each indicates information transfer from one observable in a market to another observable in the same market, calculated for all markets. The illustration emphasizes how connectivity in the triadic multidigraph is obtained for a "green observable."

These two metrics reflect on the total amount of information transfer within the system.

Similarly, we also study *markets' synchronization within one variable* via the multi-information: in the definition (2) we consider $\mathbf{X}^{\Omega}$ as a union of random variables that describe a single market observable in N markets, denoting it as $I^{\text{sys}}_{\mathbf{X}^{\Omega}}$ in the further discussion.

Lastly, we report the *average active information storage*, A^{sys}_{X}, defined as

$$A^{\text{sys}}_{X} = \frac{1}{N}\sum_{\alpha=1}^{N} A_{X^{\alpha}} = \langle A_{X^{\alpha}} \rangle_{\alpha} \tag{10}$$

i.e., it is the average AIS per market, at a each time series window.

b. Market-specific statistics. To analyze the market-specific information dynamics measures, we will consider averages of a measure obtained from different time windows $w \in [1, W]$. The amount of information transfer received by each individual market α, termed an *average collective transfer entropy* is defined as

$$T^{\text{coll}}_{X^{\alpha}} = \frac{1}{W}\sum_{w=1}^{W} T_{X^{\alpha}}(w) = \langle T_{X^{\alpha}} \rangle_{w}, \tag{11}$$

and an individual *market's average AIS* is defined as

$$A_{X^{\alpha}} = \frac{1}{W}\sum_{w=1}^{W} A_{X^{\alpha}}(w) = \langle A_{X^{\alpha}} \rangle_{w}. \tag{12}$$

c. Comparison between inter- and intra- market connectivity. In Sec. III B, we consider the amount of information transfer between microstructure observables within one market. Therefore, we compute the apparent information transfer of the form $T_{X^\alpha \to Y^\alpha}$, and contrast the amount of information transfer within the market to the amount of information sharing across markets. To make a fair comparison, we normalize the average values of TE per possible link in a transfer entropy network where nodes are market observables.

For the comparative analysis of inter- and intra- market interactions, consider a directed graph that consists of three nodes that represent our observables. We map information transfer results onto this triadic graph in which multiple edges are allowed between node pairs. Furthermore, each node is allowed to have multiple self-loops. The directed multi-graph representation depicts information transfer among different observables in a market (a somewhat endogenous information flow). In opposition to this type of internal information sharing, we also allow for inter-market connectivity, as discussed in Sec. III A (this can be thought of as exogenous information flow). Such links are represented as loops in the directed multi-graph as information is transferred from an observable in one market to the same observable in another. By appropriate edge averaging, we will turn this directed multi-graph into a directed graph that depicts averaged information transfer.

Each loop represents a significant transfer entropy link from one market to another, therefore, at most, there can be $N(N-1)$ self-loops for each node-observable, since each market can have $2(N-1)$ TE links ($N-1$ incoming and $N-1$ outgoing). We calculate the averaged TE for an observable X between all pairs of markets α, β as

$$\omega_X^{\text{self}} = \langle T_{X^\alpha \to X^\beta} \rangle_{\alpha,\beta} = \frac{1}{N(N-1)} \sum_{\alpha=1}^{N} \sum_{\beta=1|\beta\neq\alpha}^{N} T_{X^\alpha \to X^\beta}. \tag{13}$$

We also compute the average strength of incoming links, ω^{in}, and outgoing links ω^{out}. For each market observable, there can be at most two incoming edges (from the other two observables) and at most two outgoing edges, therefore, the normalization factor is $\frac{1}{2N}$,

$$\omega_X^{\text{in}} = \langle T_{Y^\alpha \to X^\alpha} \rangle_{\alpha,Y} = \frac{1}{2N} \sum_{Y\neq X} \sum_{\alpha=1}^{N} T_{Y^\alpha \to X^\alpha}, \tag{14}$$

$$\omega_X^{\text{out}} = \langle T_{X^\alpha \to Y^\alpha} \rangle_{\alpha,Y} = \frac{1}{2N} \sum_{Y\neq X} \sum_{\alpha=1}^{N} T_{X^\alpha \to Y^\alpha}. \tag{15}$$

Here, X represents one type of observable, e.g., s and Y represents a different observable, e.g., r.

D. Models of non-linear information transfer in complex systems

Here, we define two case models of stochastic processes coupled via lagged non-linear interaction. By analyzing the information-theoretic, spatiotemporal signatures of these models, we will be able to explain in the information dynamics trends observed using real market microstructure data.

a. Vector auto-regressive model of regime shifts. Let us first consider two coupled autoregressive processes,

$$\begin{aligned} X_t &= \alpha_1 X_{t-1} + \beta_1 \varepsilon_{1,t} + K(t)\varepsilon_t, \\ Y_t &= \alpha_2 Y_{t-1} + \beta_2 \varepsilon_{2,t} + K(t)\varepsilon_t + C(t)\,|X_{t-1}|^d, \end{aligned} \tag{16}$$

i.e., we have a time-delayed directional coupling $X \to Y$, whose time-varying strength is $C(t)$. In addition to this, let us assume the presence of a common, time-varying hidden driver whose strength is $K(t)$. ε terms denote independent Gaussian noise. Here, d is a constant, which we use to vary the linearity of the causal link from X to Y. Note that the hidden driver term $K(t)\varepsilon_t$ is present in both variables without any delay. This model is useful to analyze causal interaction among system's sub-units when they are "of the same type," i.e., representing the same market observable in some pair of markets.

We consider several potential systemic changes, and local information dynamics signatures associated with them. A regime shift itself is defined as a significant change in the strength of a certain type of coupling between two variables, observed at a certain point in time t_C. In particular, here we consider a time-varying coupling strength f modeled via logistic function,

$$f(t) = \frac{s}{1 + e^{-b(t-t_C)}} + C. \tag{17}$$

Here, s relates to the maximum value of the function and b to the sharpness of the transition. For example, a *regime shift of a causal driver* may be modeled as a change of a coupling strength term $C(t)$.

As Fig. 5(a) shows, *in case of a regime shift in a causal driver, we observe a significant change in the absolute values of local information dynamical measures.*[38] This situation is characterized by large TE and large AIS in the high-coupling regime and no significant change in MI. In Appendix D, we also show that a change in MI is also possible, when $K(t)$'s absolute value is large: its value is high when AIS and TE values are low and vice versa.

Similarly, the case of a regime shift of a hidden (common, simultaneous) driver can be modeled via variation in the coupling strength $K(t)$. The result, shown in Fig. 5(b) suggests that *the change in a hidden driver's strength is signified by high MI when coupling is strong, and vice versa when coupling is weak.* Note that changes in AIS and TE are possible: the former is possible when $K(t)$ is sufficiently high-valued, while the latter is possible when $C(t)$ is high-valued.

Lastly, we studied a case when $\beta_1 = \beta_2 = \beta(t)$, defined using (17) i.e., the system's overall uncertainty is a time functional, while $C(t) = K(t) = 0.3 = \text{const}$. In this case, we find that all three information dynamic measures can be mutually low-valued on the side of transition that models high-uncertainty regime, and mutually low-valued on the other, where the system's intrinsic uncertainty is low, see Fig. 5(c).

In Appendix D, we also show that the results are robust for a wide range of parameter values s, b, C. Although the analysis ensures

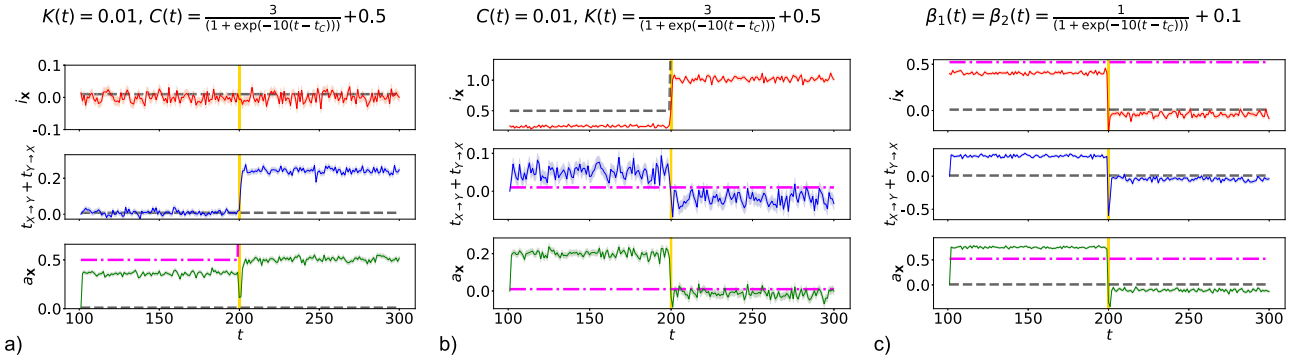

FIG. 5. Causally coupled autoregressive processes and local information-theoretic signatures of the response of a system with three types of regime shifts: the change in (a) strength of a causal driver, (b) strength of a hidden driver, and (c) intrinsic uncertainty. Figures report mean at each time step and the standard error in the mean, obtained from 100 independent simulations. The yellow vertical line indicates t_C, purple line shows causal coupling strength $C(t)$, and the gray dashed line shows the strength $K(t)$ of a hidden driver. The parameters $\alpha_1 = \alpha_2 = 0.2$, $d = 0.5$ in all figures. In (a) and (b), $\beta_1 = \beta_2 = 1$, while in (c), $\beta_1 = \beta_2$ are functions of time, defined using (17) and depicted in the figure in gray dashed and magenta dash-dot lines, while $K(t)$ and $C(t)$ are constants, equal to 0.01 and 0.5, respectively.

that the results shown in Fig. 5 are not coincidental, further statistical proof is needed to prove that signatures are persistent in a general spectrum of nonlinearly coupled autoregressive systems. We also note that it is more than possible that a combination of these effects comes into play in real data, and such convolution and its signature is not studied any further in the current work.

b. GARCH model for price returns and spread. Next, we analyze a model of coupling between different market microstructure variables, using a variant of generalized autoregressive conditional heteroskedasticity.[7,39] Our model considers possibly bidirectional coupling between price returns r and spread s when spread is coupled with volatility (variance) of price. Therefore, we have coupled time series of the form

$$\begin{aligned} r_t &= \sigma_t \varepsilon_{1,t}, \quad \varepsilon_{1,t} \sim \mathcal{N}(0,1), \\ \sigma_t^2 &= w + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 + \gamma s_{t-1}^2, \\ s_t^2 &= a s_{t-1}^2 + b \sigma_{t-1}^2 + c \varepsilon_{2,t}^2 \quad \varepsilon_{2,t} \sim \mathcal{N}(0,1). \end{aligned} \tag{18}$$

Compared to the standard GARCH(1,1)[39] (where "(1,1)" indicates that there is one variance dependence term, and the second indicates that there is one ARCH term ε, w denotes constant base level of volatility, α, β determine the influence of past squared returns and past volatility), we have several additional parameters, a, b, c, γ. The new parameters have the following economic interpretation: a is the (bare, i.e., not volatility driven) persistence of the spread, which is known to be strongly autocorrelated (especially for small tick stocks);[28] c measures the dispersion of spread and ε_t is the associated innovation noise; b links past volatility to future spread. It might be connected with asymmetric information, since market makers adjust spread according to volatility (see, for example[7]); γ links past spread with future volatility. If the spread is large, it is likely that the limit order book is sparse (as empirically shown in Ref. 40). But a sparse, i.e., illiquid, book leads to a more volatile price, since any order can create a large price change.[41]

Note that in this formulation, a link from price to spread is indirect, and in fact price returns affect spread at the two-step delay. Therefore, in the following figures, we will scan through $\delta \in [1, 2]$ and use the one for which the observed TE is larger. Depending on

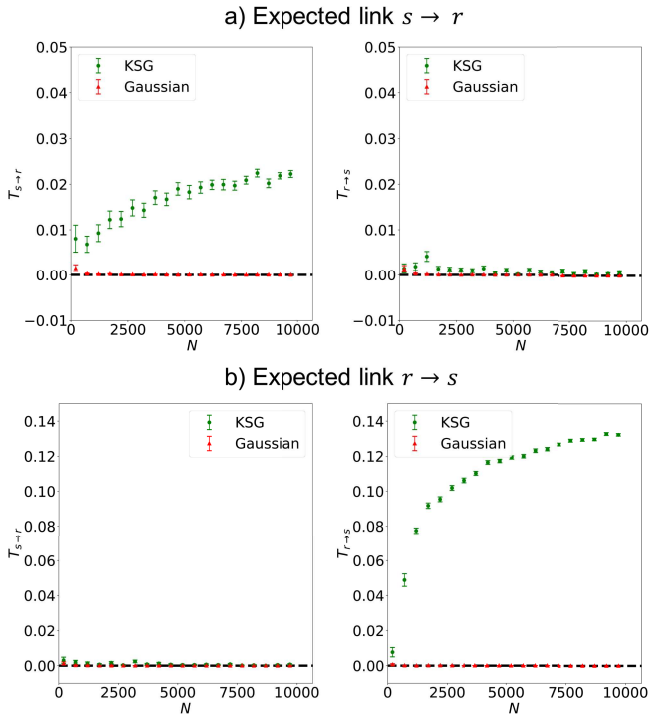

FIG. 6. (a) TE from spread to returns (left) and reverse (right), $\alpha = 0.1$, $\beta = 0.4$, $a = 0.8$, $b = 0.0$, $c = 0.1$, $\gamma = 0.9$. (b) TE from spread to returns (left) and reverse (right), $\alpha = 0.1$, $\beta = 0.1$, $a = 0.1$, $b = 0.9$, $c = 0.1$, $\gamma = 0.0$.

parameter values b, γ we can have bi- or uni-directional couplings $T_{s\to r}$ or $T_{r\to s}$. Each datapoint in the proceeding results is considered at a significance level of 0.01, and results are obtained from 1000 independent simulations.

Parameters $\alpha, \beta, a, b, c, w, \gamma$ affect stationarity and damping in the system, therefore, we need to choose such parameter values that system would be stationary throughout. To find a stationary combination of the parameters, we perform a moment analysis of the two processes (see Appendix E for more details). In Fig. 6, we show that expected links from spread to returns (top) as well as from returns to spread (bottom) are detectable using KSG estimator for TE, whereas Gaussian estimator cannot detect the non-linearly coupled GARCH variables. We also note that at small time series lengths, spurious bidirectional coupling may be observed, and the observed coupling strength converges to a true one when the time series are sufficiently long.

Since for stationary time series transfer entropy is an expectation value of the logarithm of ratios of conditional probability density functions, defined as per (7), in GARCH model (18), these probability density functions can be estimated using numerical integration and samples of a joint process $\{r_t, s_t\}$. In case of causal link from spread to returns, TE is defined as an expectation value of $\log \frac{p(r_t|r_{t-1},s_{t-1})}{p(r_t|r_{t-1})}$. Here, we have an exact probabilistic model, namely,

$$p(r_t = x|r_{t-1}, s_{t-1}, \sigma_{t-1}) = \mathcal{N}(x; 0, w + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}^2),$$

where $\mathcal{N}(x; \mu = 0, \sigma^2 = w + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}^2)$ is a normal probability density function parameterized with μ, σ. Similarly, one can obtain the analytical expression for $p(r_t|r_{t-1})$ via marginalization. Similarly, transfer entropy from returns to spread is the expectation value of $\log \frac{f(s_t|s_{t-1},s_{t-2},r_{t-1},r_{t-2})}{f(s_t|s_{t-1},s_{t-2})}$. In Appendix E, we show that the numerator can be estimated by marginalization over $\phi(y) = f(s_t = y|s_{t-1}, s_{t-2}, r_{t-2}, \sigma_{t-2})$:

$$\phi(y) = \frac{2y}{\Gamma(1/2)\sqrt{2c(y^2 - c_t^*)}} \exp\left(-\frac{y^2 - c_t^*}{2c}\right) \mathbf{1}_{[\sqrt{c_t^*},\infty)}(y).$$

Here, $c_t^* = as_{t-1}^2 + b(w + \alpha r_{t-2}^2 + \beta\sigma_{t-2}^2 + \gamma s_{t-2}^2)$ and $\mathbf{1}_{\mathscr{A}}$ stands for an indicator function defined over a subset $\mathscr{A}$, of real numbers. Again, $f(s_t|s_{t-1}, s_{t-2})$ can be obtained by marginalization.

In Appendix E, we show that the values of information transfer that we obtained empirically and report in Fig. 6 are in agreement with the transfer entropy obtained using numerical integration.

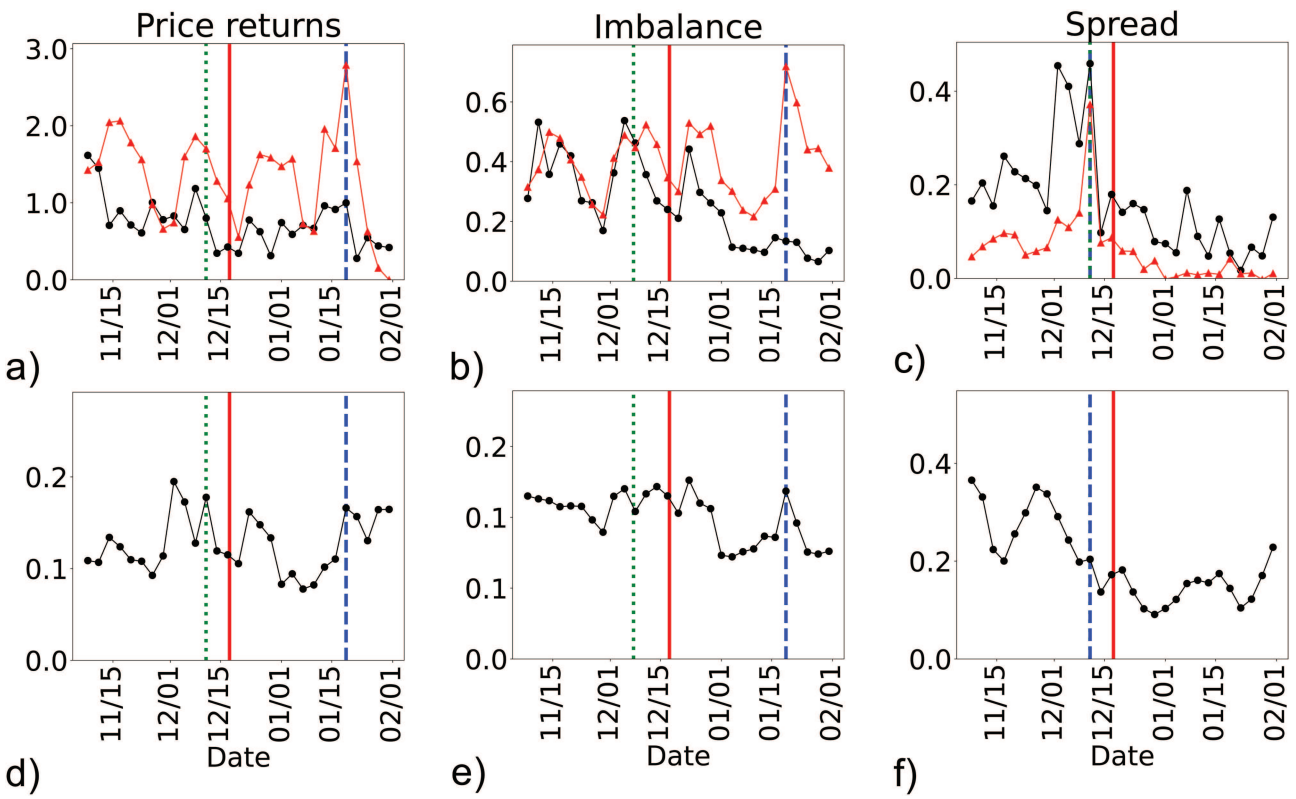

FIG. 7. Top row: multi-information I^{sys} (denoted with red triangles) and the total apparent transfer entropy $T^{\text{app,sys}}$ (denoted with black circles) across all pairs of markets using the following time series: (a) price returns, (b) order imbalance, and (c) spread. The blue dashed vertical lines indicate "critical points": the date of the peak value obtained by each variable is indicated by green dotted line; the red lines indicate the peak of the bubble, and the blue dashed lines highlight the dates at which multi-information was maximal. Bottom: average active information storage A^{sys} for (d) price returns, (e) order imbalance, and (f) spread.

III. EMPIRICAL RESULTS

In this section, we report results of information sharing across markets when a single observable is considered, and when information communication between different observables is taken into account.

A. Inter-market information dynamics for a single observable

First, we consider information dynamics individually for each of the observables: r, s, $\mathcal{O}$. Here, we look at the following: the total apparent transfer entropy, $T_r^{\text{app,sys}}, T_s^{\text{app,sys}}, T_{\mathcal{O}}^{\text{app,sys}}$; multi-information, $I_r^{\text{sys}}, I_s^{\text{sys}}, I_{\mathcal{O}}^{\text{sys}}$; and average active information storage, $A_r^{\text{sys}}, A_s^{\text{sys}}, A_{\mathcal{O}}^{\text{sys}}$. When constructing each of these measures, only significant information dynamics values are considered (the significance is discussed in Sec. II B). The results obtained for each sliding window are reported in Fig. 7: top figures show the total apparent transfer entropy as well as multi-information; bottom figures report average active information storage. Overall, we observe that week by week, there were large fluctuations in the amount of information sharing and synchronization across markets. The patterns of information dynamics for each microstructural variable differ, therefore, we discuss them individually.

a. Spread. The temporal evolution of information processing in spread suggests that a "critical point" exists nearing price peak and the peak in spread itself. Such a point in time is thought to be characterized by a peak in MI among agents that constitute the system.[42–44] The authors of Ref. 20 claim that stock market crashes exhibit a peak in MI at the point in time when one would expect a significant regime change to take place. The peak in MI observed in Fig. 7 for spread is similar to the case simulated in Fig. 5(b), where a mutual hidden driver increased in strength, making the system more synchronous.

Among spread variables, we see a large amount of total apparent transfer entropy, as well as large average AIS, persisting for weeks preceding the price peak, similar to a situation simulated in Fig. 5(a), where the importance of information sharing increased to a larger value after t_C in contrast with before. Therefore, we suggest that *spread system variables transit from a strong-coupling regime before the price peak to a weak-coupling regime after the crash.* Our intuition is further supported by similar results observed by considering $T_s^{\text{coll,sys}}$, shown in left of Fig. 8. The figure shows that the absolute values of the total collective transfer entropy $T_s^{\text{coll,sys}}$ are much smaller than the values of apparent transfer entropy. However, we still observe a similar shift from a strong-coupling regime to a weak-coupling regime, centered around the crash date.

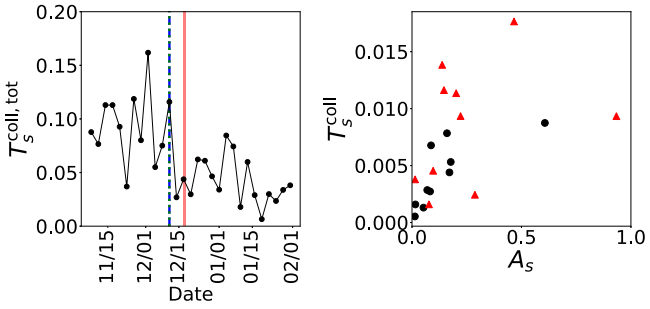

FIG. 8. Total collective TE. Left: change of total collective TE over time. Vertical lines have the same meaning as in Fig. 7. Right: relationship between average collective TE and average market's AIS. Black circles show this relationship averaging over time series windows before the price peak date; red triangles show the relationship averaging over the time series windows after the price peak date.

In Table I, we analyze AIS and total collective TE for individual markets in the two market regimes (leading up to and following the price peak). The averages are calculated from either the observed AIS values in the windows that precede the crash or follow it. Markets with small A_s are those whose future values of spread are not predictable from the past values of spread (e.g., Gemini, Coinbase): their order book spreads are responsive to either exogenous information sources or inter-market dynamics or are generally intrinsically unpredictable. Other markets such as BTC-e use a lot of information from their memory to compute the next market spread.

TABLE I. Average market spread's AIS and average market spread's collective TE, obtained by averaging over windows before the price peak date (first and third columns) and averaging over windows after the price peak date (second and fourth columns).

α	$A_{s^\alpha} \times 10^{-2}$ before December 17th	$A_{s^\alpha} \times 10^{-2}$ after December 17th	$T_{s^\alpha}^{\text{coll}} \times 10^{-3}$ before December 17th	$T_{s^\alpha}^{\text{coll}} \times 10^{-3}$ after December 17th
BTC-e	30.1	23.1	9.34	0.87
Binance	15.18	3.32	17.65	0.13
Bitfinex	4.13	2.35	1.59	0.16
Bitstamp	7.12	5.58	11.62	0.68
Bittrex	8.15	5.87	9.34	0.53
Coinbase	9.17	4.88	3.79	0.05
Gemini	11.4	9.54	4.54	0.29
HitBTC	23.0	9.58	2.44	0.44
Kraken	10.32	5.6	11.35	0.78
Poloniex	7.42	5.03	13.84	0.27

Lastly, right figure of Fig. 8 shows that those markets that were shown to have large T_s, also are those with the largest A_s. On the other hand, spread in markets such as Gemini, Coinbase, Bitfinex is driven by exogenous information as both T_s^{coll} and A_s is small. Therefore, given information available from market microstructure observables, they are the least predictable. Table I shows that Bitfinex, Gemini, and Coinbase are exchanges that experienced the smallest influx of information from other exchanges prior to crash. We also observed that these markets have small average AIS, suggesting that they are "independent" from other exchanges and could primarily be driven by exogenous information sources. On the contrary, Binance, Poloniex absorb the largest amount of information from other markets. They were also found to have large average AIS. Such markets could be labeled as "information absorbers."

b. Price returns and order imbalance. Figure 7 shows that information dynamics patterns for $\mathcal{O}$, r are different from those of s. First, for both observables, temporal evolution of MI is similar: somewhat symmetric around the crash date, minimal at the crash date, and fluctuating from high MI to low MI with a period of around 2–3 weeks. These observations suggest that similar processes were driving the system at either side of the price bubble. The peak of the bubble is characterized by a clear, although not dramatic, reduction in both TE and MI, and to a lesser extent, reduction in AIS. We observe that for $\mathcal{O}$ (and to a smaller extent, for r), all three measures—AIS, TE, MI—follow the same pattern, that is, when TE is large, AIS, MI are also large. We were able to simulate this scenario in Fig. 5(c) by changing the autoregressive parameters responsible for the amount of uncertainty in the signals. When the amount of noise, unique for each system's constituents, is varied, the three information dynamics measures either increase or decrease in magnitude simultaneously as a response, being high-valued when the uncertainty is low. *The minimal values of information dynamical quantities at the peak of the bubble suggest that the system was unpredictable, chaotic and not synchronized.*

B. Information flow between market microstructure observables

Here, we consider information sharing between different market microstructure observables and compare it to inter-market information sharing. We study two non-overlapping time series windows: the time series preceding the price peak date (November 1, 2017 to December 17, 2017) and the remainder (December 17, 2017 to January 31, 2018).

We report the results in Table II. Altogether, they suggest that *the different periods before and after the price peak are characterized by significantly different properties of information dynamics.* More specifically, we found that all significant changes in the strength of information dynamics where higher in the period before the price peak, as opposed to the period after the price peak. We find that the spread's inter-market connectivity changed significantly, as opposed to the other two observables, suggesting that after the price peaked, spread became responsive to changes in other market observables or it responded to unobserved variables. Considering that there are less and weaker links in the multigraph that illustrates the post-peak market microstructure (Fig. 9, right) as opposed to pre-peak market microstructure (Fig. 9, left), we consider the former case less likely. We also found that ω^{in} decreased for spread, and ω^{out}, A^{sys}, I^{sys} decreased for price returns.

IV. DISCUSSION

In this paper, we studied patterns of information dynamics in the Bitcoin system, describing it with market microstructure observables: price returns, order imbalance, and spread—three measures that quantify the state of market makers and market takers.

TABLE II. Intra- and inter-market information dynamics before and after the price peak. Top and middle blocks of the table: amount of active information storage, multi-information, apparent information transfer between market microstructure observables per possible link. The first column shows an average active information storage for the same time series window. The second column reports I^{sys} which is the average mutual information for each observable and each other observable within a market, averaged over markets, whereas the penultimate column reports the multi-information for a given observable across markets. All values are in nats. Bottom block of the table: results of Welch's t-test,[45] which tests the hypothesis that two populations for each observable have equal means before the price peak and after (the test cannot be performed on I^{sys}, since only one data point exists for each time period). With an asterisk we highlighted the result for which the p-value of the test was below 0.05, indicating that we have evidence against the null hypothesis of equal population means.

Period	α	A^{sys}	I^{sys}	ω^{in}	ω^{out}	$I^{\text{sys,self}}$	ω^{self}
Before Dec. 17th	s	0.367 ± 0.292	0.079 ± 0.09	0.281 ± 0.015	0.008 ± 0.007	0.369	0.009 ± 0.009
	$\mathcal{O}$	0.112 ± 0.052	0.029 ± 0.031	0.064 ± 0.003	0.007 ± 0.01	0.413	0.006 ± 0.007
	r	0.468 ± 0.082	0.097 ± 0.081	0.233 ± 0.01	0.014 ± 0.015	1.857	0.033 ± 0.023
After Dec. 17th	s	0.168 ± 0.183	0.034 ± 0.039	0.102 ± 0.006	0.004 ± 0.006	0.05	0.001 ± 0.002
	$\mathcal{O}$	0.076 ± 0.034	0.028 ± 0.028	0.046 ± 0.004	0.006 ± 0.008	0.418	0.005 ± 0.005
	r	0.342 ± 0.062	0.054 ± 0.035	0.168 ± 0.009	0.005 ± 0.006	2.545	0.031 ± 0.027
	s		*	*		n.a.	*
	$\mathcal{O}$					n.a.	
	r	*	*		*	n.a.	
			Intra-market			Inter-market	

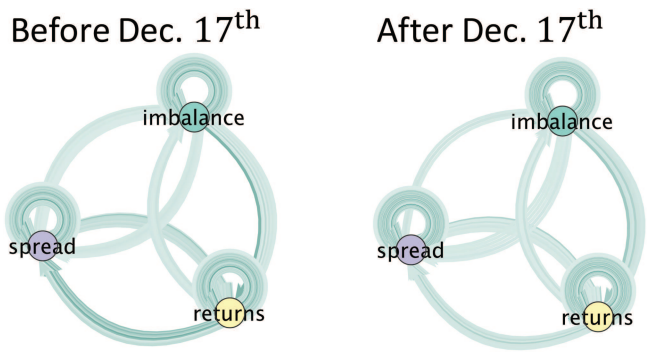

FIG. 9. Directed multi-graphs of information flow between the three market observables before the price peak (left) and after the price peak (right). Saturation of edges indicates the strength of information transfer. Self-loops represent transfer entropy from one market to another when considering the same observable.

Our analysis contrasted micro-level information processing within important Bitcoin trading venues during two distinguishable states: while the price was, on average, ascending and during the period in which the price was, on average, dropping. We found persisting intra-market connectivity at the minute level frequency. Furthermore, we found that at high frequency, *markets are interconnected via order book spread, and the interconnectivity is enhanced prior to price peak, suggesting a potential regime shift*. On the contrary, the inter-market connectivity via price returns and order imbalance appeared symmetric around the price peak. We also observed the dip in all information dynamic measures at the time of the crash, suggesting that the system was asynchronous and unpredictable. We also observed information flow between different market observables. By contrasting the two market regimes, we found that the system shifted from a strongly interlinked state to a sparsely connected one.

To supplement these findings, we analyzed several simulated econometric models. With the model of autoregressive non-linearly coupled variables that have time-varying drivers, we found several simplified mechanism that could explain the empirical observations. With the generalized autoregressive conditional heteroskedasticity model (GARCH) coupled with spread dynamics, we studied a link between returns and spread, as well as potential of detecting it using means of information transfer. Our findings suggest that different types of regime changes can be distinguished when a collection of information dynamics measures is used together.

All in all, the results indicate that prior to a price peak, the Bitcoin system was in a strongly coupled state. In particular, this is clear when liquidity marker—spread—is considered. The drop in all forms of information at the point at which the price is maximal suggests a reduction in predictability of the system. It may also indicate that such a system is susceptible to various types of perturbations; however, we leave such hypothesis for further research. We also make a note that the results of Fig. 7 (right column) for s are reminiscent of a second order phase transition[42,43] and could be a predictor of a financial crash.[44] Of course other explanations are possible for this empirical observation and future studies could further clarify its microscopic origin.

Our findings, although interesting, have limited explanatory capacity due to the granularity of order-book data, the presence of noise, and effects due to mixtures of unobserved system drivers. We also assumed that the minute-level snapshots of the limit order book data are sufficiently high-frequency to reveal true interactions across and within markets. Of course, thess data are an approximation of, in reality, practically continuous-time dynamics of the markets. Finally, since we limited our analysis to Bitcoin traded against USD(T), an analysis of larger market microstructure datasets, consisting of a bigger variety of currency pairs and forming larger information transfer networks, would undoubtedly provide richer insights as to why we observe increased market co-integration at the time of a price bubble. Lastly, it would be important to repeat our analysis using other similar datasets, concerning market microstructure during price bubbles. Observing similar results would suggest that information dynamics framework may be a generally applicable tool for early warning signals of financial crashes.

ACKNOWLEDGMENTS

This work is supported by the European Union—Horizon 2020 Program under the scheme "INFRAIA-01-2018-2019—Integrating Activities for Advanced Communities," Grant Agreement No. 871042, "SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics" (http://www.sobigdata.eu).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

All the authors designed the research. V.V. conducted experiments, analyzed the data, and derived mathematical models. F.L. and N.A.-F. derived mathematical models and supervised the work. All authors interpreted the results and wrote the manuscript.

DATA AVAILABILITY

The data that support the analysis of Sec. II D and Appendixes D and E of this study have been synthetically generated by the authors and the results can be reproduced using the equations and parameters described in the article.

APPENDIX A: DATA ALIGNMENT

As stated in the main text, the limit order books are snapshots sampled at approximately minute level frequency. However, the exact time of a snapshot is not the zeroth second of a minute, the intervals between consecutive snapshots in one market are not identical due to different technological reasons. Furthermore, for a given minute, the times at which the snapshots are taken in different markets need not be exactly the same. The trading data on the other hand is stored in continuous time, as each trading event is being registered with its exact execution time. Derivatives obtained from each type

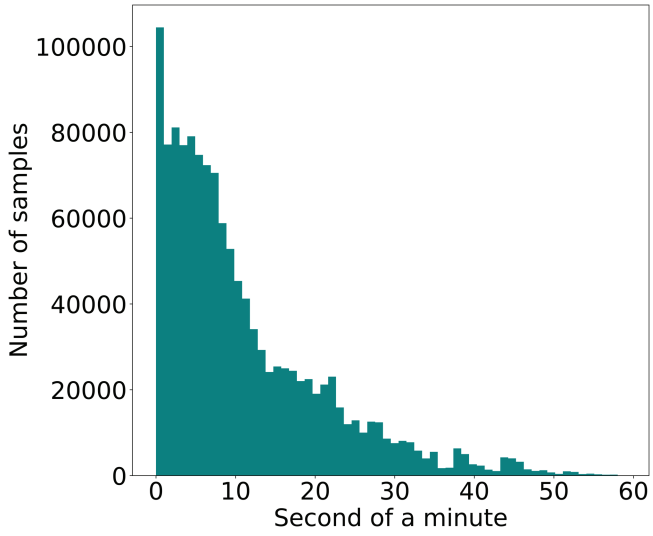

FIG. 10. Histogram of the exact capture second of each minute snapshot of the limit order books used in this paper.

of data, therefore, have to be carefully aligned for the consequent analysis to always respect the order of time.

To illustrate the data alignment problem, consider some two consecutive LOB snapshots in the BTC-e market: the first limit order book snapshot is taken on November 1, 2017 at t_1 =00:00:03.021 and the second one is taken approximately 1 min later, at t_2 =00:01:02.798. Further consider the Binance market, where the first limit order book snapshots are taken at t'_1 =00:00:01.267 and t'_2=00:01:01.274. One can already see that $t_1 \neq t'_1$ and that $t_2 - t_1 \neq 60$s precisely.

Let us first consider the data alignment for analysis in Sec. III A, where we consider information transfer between markets given one observable, e.g., spread.

First, for the observable based on trading data, we simply resample the trades at the minute level so that the trades that happened between 00:00:00.000 and 00:00:59.999 would represent the order imbalance at time τ = 00:01:00.000.

For observables based on limit order books, we assume that the snapshot is *representative of a limit order book that would be observed at the zeroth second of the minute*, namely, at precisely HH:mm:00.000 for every hour HH and minute mm. In Fig. 10, we show the histogram of the exact snapshot seconds within a minute of each limit order book. It is clear that the majority of the limit order books were captured at first halves of a minute. Since we chose to use information dynamics that assumes discrete time series, the limit order book snapshot times are first floored to the closest minutes, after which spread and price returns observables are extracted.

For the example snapshots, we consider here t_1 and t_2 would be rounded to the closes smaller minutes: $\tau_1 = \lfloor t_1 \rfloor$ = 00:00:00.000, $\tau_2 = \lfloor t_2 \rfloor$ = 00:01:00.000 from where we obtain s_{τ_1}, s_{τ_2}, and r_{τ_2}.

We note that the problems would occur only when there is a systematic delay between the exact snapshot times for a given minute in a pair of markets. If for some pair of α, β, $t^\alpha < t^\beta$ in the majority of minutes, one may observe a spurious link $\beta \to \alpha$. To test whether this could be the case, we computed the differences between the exact snapshot times at each minute between pairs of markets. We report several observed distributions and note that the distributions do not appear skewed and the mean of the distributions is close to 0, see Fig. 11.

We could also observe spurious information transfer if mapping to the closest minute is not done systematically. For instance, information transfer from BTC-e (whose first snapshots are taken at t_1, t_2 and are later rounded to τ_1, τ_2) to Binance (whose first snapshots are taken at $t'_1 < t_1, t'_2 < t_2$ and are later rounded to τ_1, τ_2). One

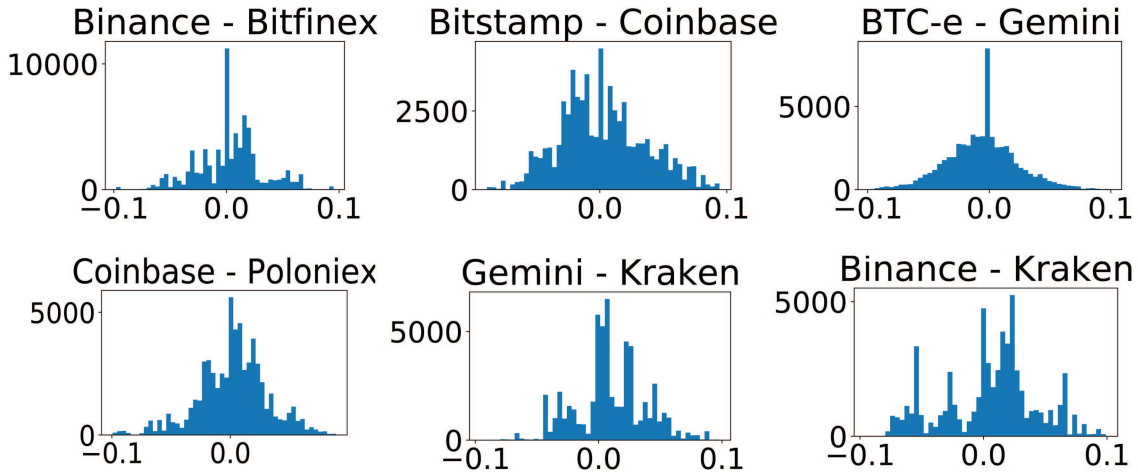

FIG. 11. Differences between capture seconds at the same minute for a pair of markets. The title of each subfigure indicates $\alpha - \beta$ for the histogram that shows $t^\alpha - t^\beta$. The *x*-axis is in units of 1 min.

may think that TE in the direction from Binance to BTC-e would be spurious because the market representation of Binance at either τ relates to an earlier true date t. However, in estimating transfer entropy, we measure the amount of reduction of uncertainty about the state of some variable X at time τ_2, given the state of another variable Y at an earlier time τ_1, over and above the information about X at time τ_2 that is already available from its own past at time τ_1. Therefore, issues would arise only if we mapped, e.g., t_1 to τ_1 and t_1' to τ_2 but not if we use flooring operator. Note that this type of rounding, overall, ensures that no acausal, spurious links can occur when the system is probed at a frequency of at least 1 min. Therefore, our results are limited to measuring information dynamics at the inter-minute granularity.

For the second part of the analysis, presented in Sec. III B, we consider information transfer between different observables. Therefore, we need to ensure that we do not create acausal links. To compare the information flow from spread at t_2 to order imbalance within that minute, we must only consider the trades that occurred in the interval $(t_1, t_2]$, ((00:00:03.021,00:01:02.798] in the example) when calculating $\mathcal{O}$. Similarly, the price returns at t_2 must be the price difference between t_2 and t_1. Therefore, we consider the time index of order books, where the frequency is approximately 1 min, and aggregate trading data based on that.

APPENDIX B: KOLMOGOROV–SMIRNOV TESTS FOR OBSERVABLES BEFORE AND AFTER PRICE PEAK

Table III reports the Kolmogorov–Smirnov test results for all observables in each market analyzed in the main text. The table shows that the two partitions of the data, namely, before the price peak date, December 17th, 2017, and afterward, have different mean values, and for all Kolmogorov–Smirnov tests at 1 min frequency and almost all tests at 12 h frequency (except for s in Coinbase and r in HitBTC and Bitstamp), the null hypothesis that the means are the same is rejected, suggesting that different types of dynamics prevail during the two periods.

APPENDIX C: CHOICE OF K IN K-NEAREST NEIGHBORS

At the heart of transfer entropy is the estimation of differential conditional mutual information, which, in turn, relies on the quality of probability distributions of our variables. Mutual information involves averages of logarithms of P, the underlying probability distribution. Since, for small P, $\log P \to \infty$, the ranges of values for our variables X, Y where P is small and hence cannot be sampled and estimated reliably from data contribute disproportionately to the value of information.[33]

TABLE III. Average of observed values for observables and results of the two-sided Kolmogorov–Smirnov test for each observable and market analyzed in the main text, performed on the two samples of the data: values observed before December 17th, 2017 and after. The null hypothesis of the tests is that two independent samples are drawn from the same continuous distribution, therefore, if p-value is high, then we cannot reject the hypothesis. For a p-value threshold of 0.05, at high frequency, we reject the hypothesis for all observables in each market, whereas at the low frequency, we cannot reject the hypothesis for a few cases, namely, s in Coinbase and r in HitBTC and Bitstamp. Columns "Before" and "After" show the average of observed values for the two time periods, "KS" and "p-value" columns report the Kolmogorov–Smirnov statistic and significance, respectively.

		s				r				$\mathcal{O}$			
	α	Before	After	KS	p-value	Before	After	KS	p-value	Before	After	KS	p-value
1 min frequency	BTC-e	26.08	36.47	0.24	0.0	0.17	−0.11	0.03	0.0	0.13	−0.04	0.22	0.0
	Binance	18.96	15.27	0.11	0.0	0.19	−0.07	0.09	0.0	0.22	0.06	0.16	0.0
	Bitfinex	2.6	3.29	0.32	0.0	0.21	−0.03	0.03	0.0	0.13	−1.68	0.06	0.0
	Bitstamp	10.86	18.64	0.26	0.0	0.16	−0.03	0.05	0.0	0.9	−0.21	0.07	0.0
	Bittrex	10.29	21.94	0.27	0.0	0.18	−0.04	0.03	0.0	0.14	−0.13	0.05	0.0
	Coinbase	0.98	1.0	0.04	0.0	0.22	−0.04	0.05	0.0	2.77	0.81	0.11	0.0
	Gemini	3.74	6.22	0.12	0.0	0.21	−0.06	0.07	0.0	0.42	−0.04	0.07	0.0
	HitBTC	13.57	13.42	0.15	0.0	0.19	−0.04	0.04	0.0	0.0	−0.04	0.06	0.0
	Kraken	10.99	13.13	0.15	0.0	0.25	−0.02	0.06	0.0	−0.06	−0.23	0.06	0.0
	Poloniex	10.26	17.34	0.27	0.0	0.19	−0.05	0.03	0.0	−0.0	−0.28	0.05	0.0
12 h frequency	BTC-e	26.94	44.01	0.75	0.0	122.53	−41.44	0.22	0.0	92.86	−29.87	0.42	0.0
	Binance	19.22	15.47	0.28	0.0	124.78	−48.25	0.2	0.01	158.92	42.88	0.39	0.0
	Bitfinex	2.63	3.31	0.46	0.0	133.09	−29.04	0.21	0.01	91.29	−1206.33	0.21	0.04
	Bitstamp	10.93	18.93	0.68	0.0	123.08	−15.31	0.18	0.03	647.54	−152.98	0.37	0.0
	Bittrex	10.32	21.96	0.75	0.0	116.57	−30.16	0.18	0.03	98.48	−93.8	0.29	0.0
	Coinbase	0.99	1.0	0.08	0.95 >	130.5	−36.43	0.21	0.0	1995.38	585.12	0.43	0.0
	Gemini	3.97	6.5	0.54	0.0	129.14	−41.82	0.21	0.01	302.27	−26.91	0.34	0.0
	HitBTC	13.59	13.55	0.35	0.0	126.85	−30.76	0.17	0.04	2.62	−29.66	0.26	0.0
	Kraken	11.01	13.8	0.36	0.0	175.86	−12.71	0.25	0.0	−43.55	−166.86	0.23	0.02
	Poloniex	10.26	17.45	0.67	0.0	124.65	−42.06	0.22	0.0	−1.88	−202.68	0.28	0.0

Information estimators that use continuity of real valued data to overcome issues related to undersampling have proved to be the most successful for a wide range of databases and applications. Among those, one of the most successful is the Kraskov, Stögbauer, and Grassberger estimator,[32] which we will refer to as KSG. The KSG estimates information transfer based on the statistics of distances between neighboring data points. To implement this, KSG uses the $\max(\Delta x, \Delta y, \Delta z)$ metric to define the distance between two points that are $(\Delta x, \Delta y, \Delta z)$ away from each other in the joint space $\{x, y, z\}$. For each given test point, n_z is the neighbor count strictly within ε in the z marginal space, and n_{xz} and n_{yz} are the neighbor counts strictly within ε in the joint $\{x, z\}$ and $\{y, z\}$ spaces, respectively. Conditional mutual information is then defined in terms of these variables,[13]

$$I(X, Y|Z) = \psi(K) - E\left[\psi(n_{xz}) - \psi(n_{yz}) + \psi(n_z)\right], \qquad \text{(C1)}$$

where ψ is a digamma function and averaging is over the samples.

As pointed out in Ref. 33, for any information estimator, it is essential to ensure that such estimators (there a KSG estimator for MI rather than conditional MI was considered) is minimally biased toward the sample size, and the choice of K is optimal: that is, for independent sub-samples of the dataset the variance of obtained result is minimal.

The details of this approach for the choice of K are thoroughly described in Ref. 33. Here, we simply note that we tested several choices of nearest neighbors for a pair of sample time series for which we identified significant transfer entropy with estimator using $K = 4$ nearest neighbors. As Fig. 12 shows, this choice of K has a small bias for the sample size: even if we slice the time series in four to five parts, the value of TE does not increase significantly. Furthermore, the variance for values of individual subsamples is also minimal. Note for $K = 1$, we have a strong sample-size dependent bias as well as large variance.

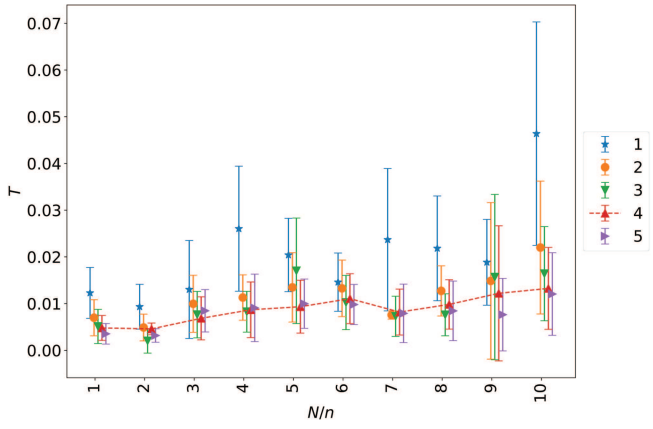

FIG. 12. Bias and variance of transfer entropy for subsamples of size n for spread time series from Bitstamp to HitBTC where we identified a significant information transfer of 0.0052 nats using $K = 4$ nearest neighbors.

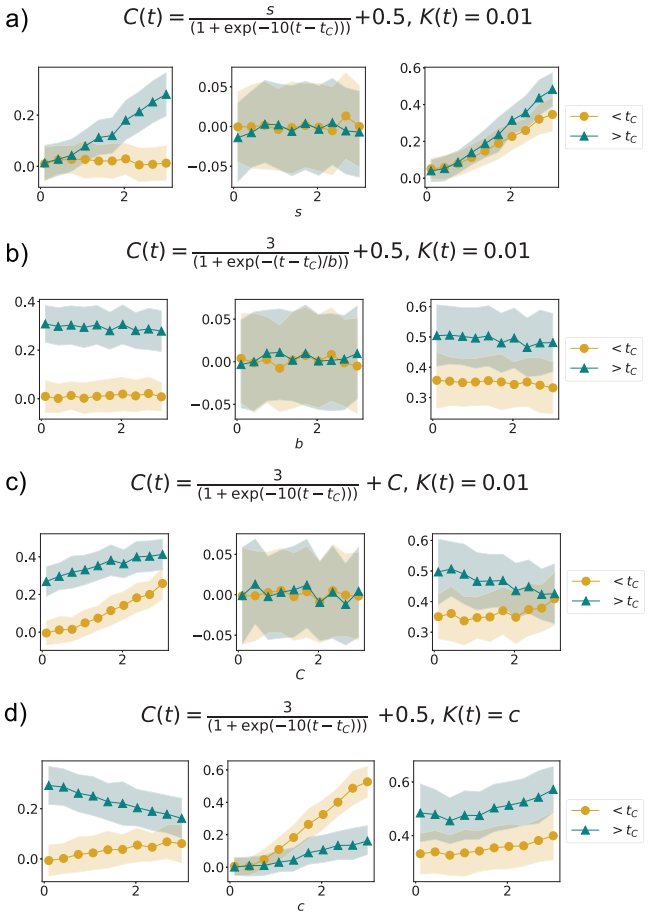

FIG. 13. Expected values of information dynamical variables when $C(t)$ is a sigmoid functional of time, (17), parameterized by (a) s, (b) b, (c) C, (d) in the presence of a constant hidden driver of strength c. Information dynamic measures are significantly different in the two regimes, as defined by the values of the sigmoid function.

APPENDIX D: INFORMATION TRANSFER METRICS AROUND REGIME SHIFTS IN AUTOREGRESSIVE TIME SERIES

Here, we perform a parameter sensitivity analysis for the results, shown in Fig. 5. Note that in Figs. 13 and 14 we do not consider the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ as variables and set them to $\alpha_1 = \alpha_2 = 0.2$ and $\beta_1 = \beta_2 = 1$. In all figures, we also chose $d = 0.5$. We note that for a definite conclusion regarding these information dynamical signatures, a thorough analysis of the relation to these parameters should also be conducted.

In Fig. 13, we show that a signature of high TE and AIS with no change in MI indicates strong causal coupling [illustrated in Fig. 5(a) of the main text], and this statement is robust for a wide range of parameters. It is worth pointing out, however, that at high levels of K, a significant amount of MI is also detectable in a weak-coupling regime.

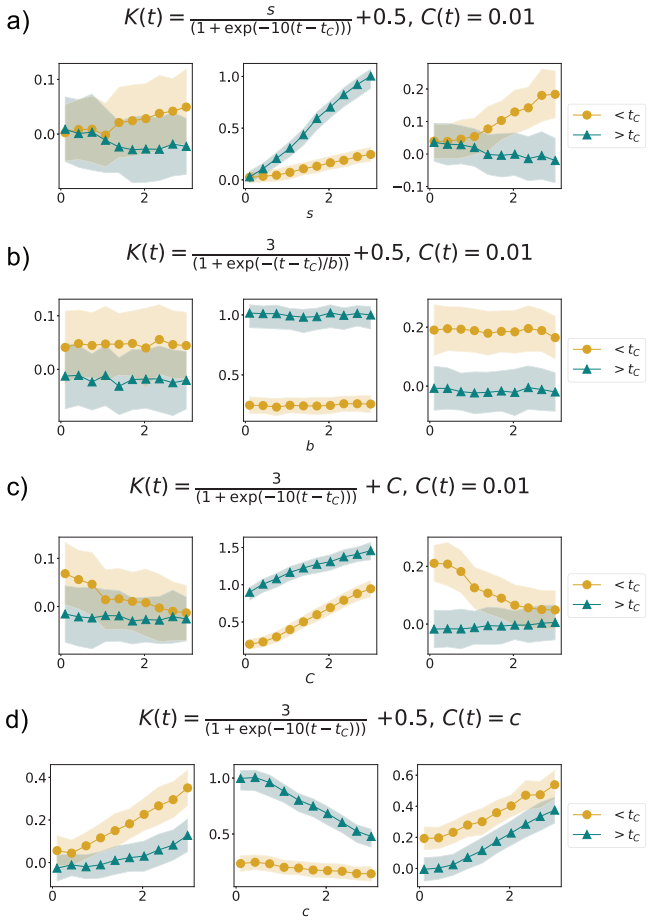

FIG. 14. Expected values of information dynamical variables when K is a sigmoid functional of time, parameterized by (a) s and (b) b, (c) c, (d) in the presence of a constant causal driver of strength c. Parameter b quantifies the abruptness of the transition, whereas s quantifies the magnitude of change in the two regimes centered at t_C. MI is high in the high-coupling regime, and the reverse is true in the low-coupling regime.

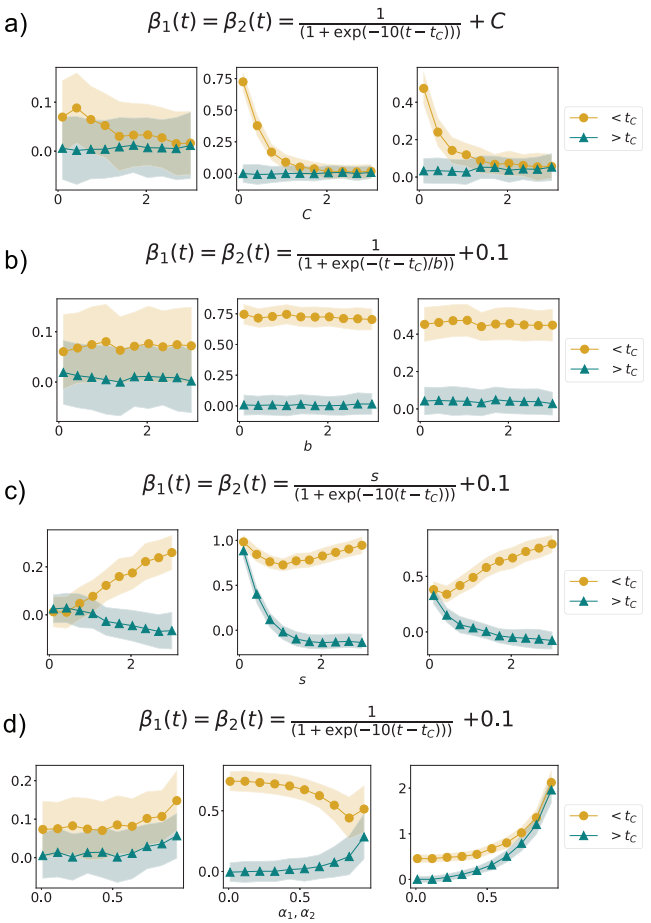

FIG. 15. Expected values of information dynamical variables when β_1, β_2 are sigmoid functionals of time, parameterized by (a) C, (b) b, (c) s, (d) autoregressive parameters α_1, α_2. Parameter b quantifies the abruptness of the transition, whereas s quantifies the magnitude of change in the two regimes centered at t_C. MI, TE, AIS are high in the high-uncertainty regime (high β_1, β_2 regime), and the reverse is true in the low-uncertainty regime. Here, we have $C(t) = 1 =$ cost, and $K(t) = 1 =$ cost.

The case of a shift in the hidden driver [illustrated in Fig. 5(b) of the main text] is signed with a high MI value in a strong-coupling regime and vice versa in a weak-coupling regime. In Fig. 14, we show that for a wide range of parameter values that characterize the strength and steepness of the regime transition function, as well as the strength of causal coupling, MI is significantly larger in the strong-coupling regime. Although for some parameter values AIS is anti-correlated with MI, it is not a robust signifier based on the sensitivity analysis. We also note that, not surprisingly, the strength of the coupling is directly related to the magnitude of MI.

Lastly, Fig. 15 shows that a change in overall system's uncertainty [the amount of uncorrelated random noise, that was illustrated in Fig. 5(c) of the main text] is detectable with a signature of having low TE, MI, and AIS in a high-uncertainty regime and vice versa in a low-uncertainty regime.

APPENDIX E: MODELING THE NON-LINEAR RELATIONSHIP BETWEEN PRICE AND SPREAD

Consider VAR toy example,

$$\begin{aligned} X_t &= \alpha_1 X_{t-1} + \beta_1 \varepsilon_{1,t} + K(t)\varepsilon_t, \\ Y_t &= \alpha_2 Y_{t-1} + \beta_2 \varepsilon_{2,t} + K(t)\varepsilon_t + C(t)\,(X_{t-1})^d, \end{aligned} \tag{E1}$$

with $d = 1$ and $d = 2$. For simplicity, we set $\alpha_1, \alpha_2 = 0$, $\beta_1, \beta_2 = 0.1$, $C(t) = 1$, and $K(t) = 0$. In Fig. 16, we show that when the non-linear relation between the variables is present, KSG estimator can detect a large amount of information transfer. In case of linear relation, we show a convergence between two measures in the large N limit, while a non-linear relationship is only detectable with a KSG estimated TE.

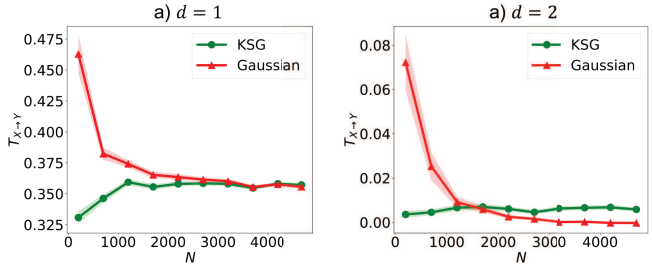

FIG. 16. (a) $d = 1$, i.e., linear relationship between X, Y, (b) $d = 2$, i.e., non-linear relationship between X, Y.

A more realistic example considers possibly bidirectional coupling between price returns r and spread s when spread is coupled with volatility (variance) of price. Therefore, we have coupled time series of the form (18).

Note that in this formulation, a link from price to spread is indirect, and in fact, price returns affect spread at the two-step delay. Therefore, in the following figures, we will scan through $\delta \in [1, 2]$ and choose the delay for which the observed TE is larger. Depending on parameter values b, γ we can have bi- or uni-directional couplings $T_{s\to r}$ or $T_{r\to s}$. Each datapoint in the proceeding results is considered at a significance level of 0.01, and results are obtained from 100 independent simulations.

Parameters $\alpha, \beta, a, b, c, w, \gamma$ affect stationarity and damping in the system; therefore, we need to choose such parameter values that system would be stationary throughout. To find a stationary combination of the parameters, we need to consider the moments of the two processes.

First, to compute the unconditional mean values σ^2 and s^2 of volatility and spread, let us consider

$$\begin{aligned}\sigma^2 = Var[r_t] = E[r_t^2] &= E[E[r_t^2|\mathcal{F}_{t-1}]]\\ &= E[w + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}^2]\\ &= w + (\alpha+\beta)\sigma^2 + \gamma s^2\end{aligned}$$

and similarly,

$$\begin{aligned}s^2 = E[s_t^2] &= E[as_{t-1}^2 + b\sigma_{t-1}^2 + c\epsilon_t^2]\\ &= as^2 + b\sigma^2 + c.\end{aligned}$$

Solving the system, we obtain

$$\sigma^2 = \frac{w(1-a) + \gamma c}{(1-\alpha-\beta)(1-a) - \gamma b}, \tag{E2}$$

$$s^2 = \frac{b\sigma^2 + c}{1-a}. \tag{E3}$$

From the first equation, we see that unconditional volatility increases with γ, c, and b, i.e., there is a contribution of illiquidity to volatility. From the second equation, for large volatility ($\sigma^2 \gg b/c$), we obtain the well known proportionality $s \propto \sigma$ between liquidity and volatility. The stationary condition is obtained by imposing that these two quantities are positive, hence, since $a < 1$, it must be

$$(1-\alpha-\beta)(1-a) - \gamma b > 0. \tag{E4}$$

Volatility can be eliminated from the Eq. (18) and by recursive substitution, we get

$$\sigma_t^2 = \frac{w}{1-\beta} + \frac{\alpha}{\beta}\sum_{i=1}^{+\infty}\beta^i r_{t-i}^2 + \frac{\gamma}{\beta}\sum_{i=1}^{+\infty}\beta^i s_{t-i}^2$$

$$s_t^2 = \frac{bw}{1-\beta} + as_{t-1}^2 + \frac{b\gamma}{\beta^2}\sum_{i=2}^{+\infty}\beta^i s_{t-i}^2 + \frac{ba}{\beta^2}\sum_{i=2}^{+\infty}\beta^i r_{t-i}^2 + c\epsilon_t^2,$$

i.e., the volatility is the sum of an Exponentially Weighted Moving Average (EWMA) of past square returns and of an EWMA of past square spreads and the same holds true for squared spread. Squaring the last two equations in (18) and taking expectations, one gets

$$\begin{aligned}E[\sigma^4] = w^2 &+ 3\alpha^2 E[\sigma^4] + \beta^2 E[\sigma^4] + \gamma^2 E[s^4]\\ &+ 2w\alpha\sigma^2 + 2w\beta\sigma^2 + 2w\gamma s^2\\ &+ 2\alpha\beta E[\sigma^4] + 2\alpha\gamma E[\sigma^2 s^2] + 2\beta\gamma E[\sigma^2 s^2]\end{aligned}$$

and

$$E[s^4] = a^2 E[s^4] + b^2 E[\sigma^4] + 3c^2 + 2abE[\sigma^2 s^2] + 2acs^2 + 2bc\sigma^2,$$

where we have used the fact that $E[\epsilon_t^4] = E[\varepsilon_t^4] = 3$ (because they are Gaussian), and we introduced in the equations the unconditional means σ^2 and s^2 in Eq. (E2).

In order to close the system, we need an expression for $E[\sigma^2 s^2]$. This can be obtained by taking the product of the last two equations in (18) and taking the expectation, obtaining

$$\begin{aligned}E[\sigma^2 s^2] = was^2 &+ wb\sigma^2 + wc + \alpha aE[\sigma^2 s^2]\\ &+ \alpha bE[\sigma^4] + \alpha c\sigma^2 + a\beta E[\sigma^2 s^2] + \beta bE[\sigma^4]\\ &+ \beta c\sigma^2 + \gamma aE[s^4] + \gamma bE[\sigma^2 s^2] + \gamma cs^2.\end{aligned}$$

We now have a system of three equations and three unknowns ($E[\sigma^4]$, $E[s^4]$, and $E[\sigma^2 s^2]$) whose solution gives the second

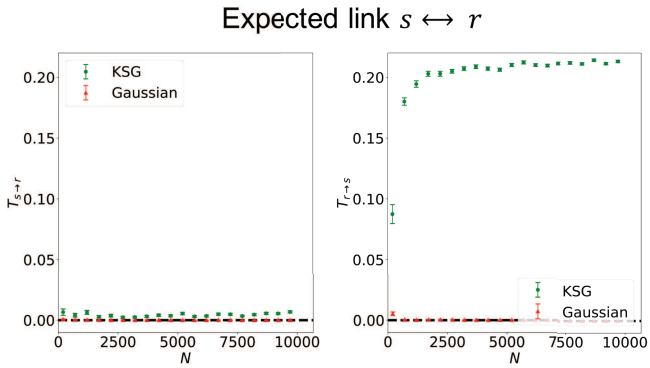

FIG. 17. TE from spread to returns (left) and reverse (right), $\alpha = 0.1$, $\beta = 0.5, a = 0.1, b = 0.5, c = 0.1, \gamma = 0.5$.

moments. As usual, these moments will be finite in the parameter region where they are positive.

The solutions for fourth moments are cumbersome to write in full form. However, we checked that for the parameter combinations that we will study next the expectation of $\sigma^2, s^2, \sigma^4, s^4$ are all well-defined. In particular, we considered the three following cases:

- Information transfer from spread to volatility and, therefore, to price returns; no information flow from price returns to spread: $\alpha = 0.1, \beta = 0.4, a = 0.8, b = 0.0, c = 0.1, \gamma = 0.9, \omega = 0.1$.
- Information transfer from volatility and, therefore, price returns to spread; no information flow from spread to price returns: $\alpha = 0.1, \beta = 0.1, a = 0.1, b = 0.9, c = 0.1, \gamma = 0.0, \omega = 0.1$.
- Bidirectional coupling from price returns to spread and vice versa: $\alpha = 0.1, \beta = 0.5, a = 0.1, b = 0.5, c = 0.1, \gamma = 0.5, \omega = 0.1$.

First set of parameter values gives $\sigma^2 \approx 0.13, s^2 = 0.24$, $\sigma^4 \approx 0.01 + 0.04\sigma^2, s^4 \approx 0.04 + 0.02s^2 + 0.24\sigma^2$. All values are positive. For the second set of parameters, we get $\sigma^2 \approx 1.1$, $s^2 \approx 0.5$, $\sigma^4 \approx 0.2 + 1.4s^2 + 0.2\sigma^2$, $s^4 \approx 0.08 + 0.44s^2$. Lastly, for the third set of parameters, we find $\sigma^2 \approx 1.27$, $s^2 \approx 0.82$, and $\sigma^4 \approx 0.2 + 1.4s^2 + 0.2\sigma^2$, $s^4 \approx 0.08 + 044s^2$. Therefore, all parameter values are suitable.

1. Simulation results

The three cases of coupling between spread and price returns are shown in Figs. 6, and 17. The first case, shown in Fig. 6(top) considers $\gamma = 0.9$ and, therefore, coupling from spread to returns is strong. The figure shows that Gaussian estimate of TE is always close to zero, whereas KSG estimate of TE is clearly larger in the direction from spread to price returns (left figure) as opposed to the reverse (right figure). Note that in some simulations significant non-zero coupling is observed, when Gaussian estimate is used. However, on average, the magnitude of information transfer estimated with Gaussian estimate is smaller in contrast with KSG estimated TE.

The second case, shown in Fig. 6(bottom) considers a situation where $\gamma = 0.0$, therefore, the link from spread to returns should not be observed, whereas parameter b, that indicates the strength of a reciprocal coupling is set to a large value $b = 0.9$. Here, transfer entropy detects a significant information transfer only from price returns to spread. Similar to the previous case, the Gaussian estimate is also found significant and different from zero in some of the simulations; however, the magnitude of TE is smaller than that obtained with the KSG estimate.

The last figure, Fig. 17, considers a case when both $\gamma \neq 0$ and $b \neq 0$. Although the link from spread to returns is much weaker than the reverse, we, nevertheless, observe significant flow in both directions.

In all cases considered, TE estimates show no apparent size-dependent value drift, and, therefore, we assume there is no sample size-related bias, as long as $N \geq 1000$.

2. Theoretical values for transfer entropy between price returns and spread

The mathematical simplicity of our model (18) allows us to estimate the TE using numerical integration techniques. By comparing the simulation results, obtained using a data-dependent estimator of TE, with the results obtained analytically, we can attest the usability of such estimator for the type of data that could arise from the model considered.

a. Information transfer from spread to price returns

To proceed, note that transfer entropy from spread to price returns is the difference of two conditional entropy ($H(\cdot|\cdot)$ terms, namely,

$$T_{s\to r} = H(r_t|r_{t-1}) - H(r_t|r_{t-1}, s_{t-1}).$$

It can also be rewritten in terms of expectation of $\frac{f(r_t|r_{t-1},s_{t-1})}{f(r_t|r_{t-1})}$, where we sample points $\{r_t, r_{t-1}, s_{t-1}\}$, $\{r_t, s_t\}$ obtained from (18). Using a natural logarithm, log yields a result in units of nats, and $f(\cdot)$ denotes probability density function (p.d.f.).

First, let us consider the case when $b = 0$ (Fig. 6) we derive the theoretical distribution densities for model-based transfer entropy from spread to returns. To obtain the p.d.f. $f(r_t|r_{t-1}, s_{t-1})$, one can sample values of $f(r_t|r_{t-1}, s_{t-1}, \sigma)$ and integrate over the hidden volatility variable

$$f(r_t|r_{t-1}, s_{t-1}) = \int f(\sigma) f(r_t|r_{t-1}, s_{t-1}, \sigma)\, d\sigma.$$

Here, we assume independence of σ. Further marginalizing, we obtain the conditional p.d.f. of price returns

$$f(r_t|r_{t-1}) = \int\int f(\sigma) f(s) f(r_t|r_{t-1}, s, \sigma)\, d\sigma\, ds.$$

Finally, in case of our model defined in (18), we know the exact probabilistic model

$$f(r_t = x|r_{t-1}, s_{t-1}, \sigma_{t-1}) = \mathcal{N}(x; 0, w + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}^2),$$

where $\mathcal{N}(x; \mu = 0, \sigma^2 = w + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma s_{t-1}^2)$ denote the normal probability density function parameterized with μ, σ.

Now, to compute the quantity $T_{s\to r}$, we perform Monte Carlo sampling

$$f(r_t = x|r_{t-1}, s_{t-1}) = E_{\sigma_i\sim f(\sigma)}[f(r_t = x|r_{t-1}, s_{t-1}, \sigma_i)] \approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{N}(x; 0, w + \alpha r_{t-1}^2 + \beta\sigma_i^2 + \gamma s_{t-1}^2).$$

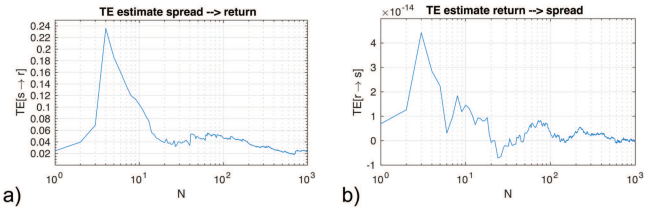

FIG. 18. Model-based TE from spread to returns (a) and reverse (b), with expected link $s \to r$. Parameters: $\alpha = 0.1, \beta = 0.4, a = 0.8, b = 0.0, c = 0.1, \gamma = 0.9$.

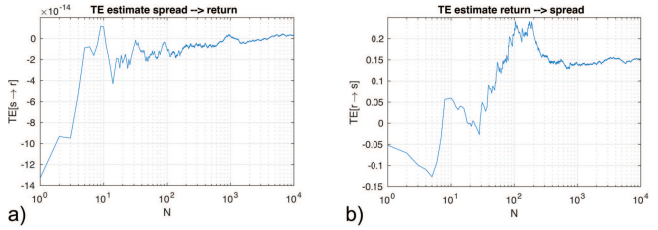

FIG. 19. Model-based TE from spread to returns (a) and reverse (b) with expected link $r \to s$. Parameters: $\alpha = 0.1, \beta = 0.1, a = 0.1, b = 0.9, c = 0.1, \gamma = 0.0$.

$$f(r_t = x|r_{t-1}) = E_{\sigma_i \sim f(\sigma), s_i \sim f(s)}[f(r_t = x|r_{t-1}, s_i, \sigma_i)]$$
$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{N}(x; 0, w + \alpha r_{t-1}^2 + \beta\sigma_i^2 + \gamma s_i^2).$$

On the left panel of Fig. 18 and right panel of Fig. 19, we show agreement between the transfer entropy obtained from such analytical integration and the result we found in Fig. 6 using a K-nearest neighbor data-driven estimator.

b. Information transfer from price returns to spread

Similar to the case before, transfer entropy from returns to spread is the difference of two conditional entropies,

$$T_{r\to s} = H(s_t|s_{t-1}, s_{t-2})] - H(s_t|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2})],$$

which can also be rewritten in terms of expectation value $\log \frac{f(s_t|s_{t-1},s_{t-2},r_{t-1},r_{t-2})}{f(s_t|s_{t-1},s_{t-2})}$.

Here, we need to estimate a conditional p.d.f. of spread given last two spread values. We do this again via marginalization of hidden variables,

$$f(s_t|s_{t-1}, s_{t-2}) = \int\int f(s_t, r_{t-1}, r_{t-2}|s_{t-1}, s_{t-2}) dr_{t-1} dr_{t-2}.$$

The conditional p.d.f. of spread given last two spread and price return values, $f(s_t|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2})$, is obtained by marginalizing over a hidden volatility variable,

$$f(s_t|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2}) \tag{E5}$$
$$= \int f(s_t, \sigma_{t-2}|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2}))\, d\sigma_{t-2}. \tag{E6}$$

For simplicity, let us denote the conditional density of spread with $\phi(y)$

$$f(s_t = y|s_{t-1}, s_{t-2}, r_{t-2}, \sigma_{t-2}) = \phi(y).$$

Then, the value of spread at time t is connected to the conditional variables as

$$s_t^2 = \underbrace{as_{t-1}^2 + b(w + \alpha r_{t-2}^2 + \beta\sigma_{t-2}^2 + \gamma s_{t-2}^2)}_{c_t^*} + c\epsilon^2,$$

where ϵ^2 is chi-square distribution. Note that spread can be rewritten as

$$s_t = \sqrt{c_t^* + c\epsilon^2},$$

where $\sqrt{c\epsilon^2}$ is the Nakagami distribution. Then, we have

$$\phi(x) = \frac{2\sqrt{0.5}}{\Gamma(0.5)\sqrt{c}} \exp\left(\frac{-x^2}{2c}\right),$$

and $c_t^* + c\epsilon^2$ is shifted Gamma distribution. Under the monotone transformation of random variable $y = g(x) = \sqrt{c_t^* + x}$, we can compute the p.d.f. of s_t as

$$\phi(y) = f_x(g^{-1}(y)) \left|\frac{d}{dy}(g^{-1}(y))\right|,$$

where

$$f_x(x) = \frac{1}{\Gamma(1/2)\sqrt{2c}} x^{-1/2} \exp\left(-\frac{x}{2c}\right)$$

and

$$g^{-1}(y) = y^2 - c_t^*.$$

Finally, we get conditional probability density for spread $f(s_t = y|s_{t-1}, s_{t-2}, r_{t-2}, \sigma_{t-2}) = \phi(y)$,

$$\phi(y) = \frac{2y}{\Gamma(1/2)\sqrt{2c(y^2 - c_t^*)}} \exp\left(-\frac{y^2 - c_t^*}{2c}\right) \mathbf{1}_{[\sqrt{c_t^*},\infty)}(y),$$

where $c_t^* = as_{t-1}^2 + b(w + \alpha r_{t-2}^2 + \beta\sigma_{t-2}^2 + \gamma s_{t-2}^2)$ and $\mathbf{1}_{\mathscr{A}}$ stands for an indicator function defined over a subset of real numbers, $\mathscr{A}$. Now, we can get a numerical Monte Carlo estimate of the model-based transfer entropy as

$$f(s_t = y|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2})$$
$$= E_{\sigma_i \sim f(\sigma_{t-2})}[f(s_t|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2}, \sigma_i))$$
$$= \frac{1}{N}\sum_{i=1}^{N} \phi(y; c, c_t^* = as_{t-1}^2 + b(w + \alpha r_{t-2}^2 + \beta\sigma_i^2 + \gamma s_{t-2}^2))$$

and $f(s_t = y|s_{t-1}, s_{t-2})$ is an expectation

$$f(s_t = y|s_{t-1}, s_{t-2}) = E_{r_i, r_j \sim f(r_{t-1}, r_{t-2})}[f(s_t = y|s_{t-1}, s_{t-2}, r_{t-1}, r_{t-2})).$$

In Figs. 18 and 19, we show agreements of transfer entropy obtained from such analytical integration, in comparison to transfer entropy values from returns to spread with KSG estimates shown in corresponding panels of Fig. 6. For the case shown in Fig. 17, where we have bidirectional flow, namely, $b > 0$ and $\gamma > 0$, deriving theoretical density distributions requires special attention since we need to account for dependence between spread and volatility in marginalization step, (E5), that falls out of the scope of current work.

REFERENCES

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Decentralized Bus. Rev. 21260 (2008); available at https://www.debr.io/article/21260-bitcoin-a-peer-to-peer-electronic-cash-system.

[2] D. G. Baur, K. Hong, and A. D. Lee, "Bitcoin: Medium of exchange or speculative assets?," J. Int. Financ. Markets Inst. Money **54**, 177–189 (2018).

[3]F. Glaser, K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering, "Bitcoin—Asset or currency? Revealing users' hidden intentions," in *ECIS 2014*, April 15, 2014 (SSRN, 2014); available at https://ssrn.com/abstract=2425247.
[4]M. Gronwald, "Is bitcoin a commodity? On price jumps, demand shocks, and certainty of supply," J. Int. Money Finance **97**, 86–92 (2019).
[5]D. G. Baur and T. Dimpfl, "Information transmission across cryptocurrency markets and the role of the blockchain," SSRN 3573367 (2020).
[6]J. Donier and J.-P. Bouchaud, "Why do markets crash? Bitcoin data offers unprecedented insights," PLoS One **10**, e0139356 (2015).
[7]L. R. Glosten and P. R. Milgrom, "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders," J. Financ. Econ. **14**, 71–100 (1985).
[8]I. Makarov and A. Schoar, "Trading and arbitrage in cryptocurrency markets," J. Financ. Econ. **135**, 293–319 (2020).
[9]P. Pagnottoni and T. Dimpfl, "Price discovery on bitcoin markets," Digital Finance **1**, 139–161 (2019).
[10]C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," Econometrica: J. Econ. Soc. **37**, 424–438 (1969).
[11]T. Schreiber, "Measuring information transfer," Phys. Rev. Lett. **85**, 461 (2000).
[12]M. Paluš, V. Komárek, Z. C. V. Hrnčř, and K. Šěrbová, "Synchronization as adjustment of information rates: Detection from bivariate time series," Phys. Rev. E **63**, 046211 (2001).
[13]T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier, *An Introduction to Transfer Entropy* (Springer, 2016).
[14]J. T. Lizier, *The Local Information Dynamics of Distributed Computation in Complex Systems* (Springer Science & Business Media, 2012).
[15]B. Afsharizand, P. H. Chaghoei, A. A. Kordbacheh, A. Trufanov, and G. Jafari, "Market of stocks during crisis looks like a flock of birds," Entropy **22**, 1038 (2020).
[16]T. Bossomaier, L. Barnett, A. Steen, M. Harré, S. d'Alessandro, and R. Duncan, "Information flow around stock market collapse," Acc. Finance **58**, 45–58 (2018).
[17]T. D. Peron and F. A. Rodrigues, "Collective behavior in financial markets," Europhys. Lett. **96**, 48004 (2011).
[18]J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, "The information dynamics of phase transitions in random Boolean networks," Artif. Life **11**, 374–381 (2008).
[19]J. T. Lizier, S. Pritam, and M. Prokopenko, "Information dynamics in small-world Boolean networks," Artif. Life **17**, 293–314 (2011).
[20]M. Harré and T. Bossomaier, "Phase-transition–like behaviour of information measures in financial markets," Europhys. Lett. **87**, 18009 (2009).
[21]A. Abeliuk, Z. Huang, E. Ferrara, and K. Lerman, "Predictability limit of partially observed systems," Sci. Rep. **10**, 1–10 (2020).
[22]We also note that the analysis in Appendix B indicates that the time of a maximum price date partitions the data into two significantly different subsets w.r.t. empirical distributions of our observables.
[23]T. Dimpfl and F. J. Peter, "Group transfer entropy with an application to cryptocurrencies," Physica A **516**, 543–551 (2019).
[24]Kaiko is a cryptocurrency market data provider for institutional investors and enterprises, see kaiko.com.
[25]F. Lillo and R. N. Mantegna, "Power law relaxation in a complex system: Omori law after a financial market crash," Phys. Rev. E **68**, 016119 (2003).
[26]J.-P. Bouchaud, J. Farmer, and F. Lillo, "How markets slowly digest changes in supply and demand," in *Handbook of Financial Markets: Dynamics and Evolution*, Handbook of Finance (North-Holland, 2009), pp. 57–160.
[27]F. Lillo, "Order flow and price formation" arXiv:2105.00521 (2021).
[28]A. Ponzi, F. Lillo, and R. N. Mantegna, "Market reaction to a bid-ask spread change: A power-law relaxation dynamics," Phys. Rev. E **80**, 016112 (2009).
[29]G. Tononi, O. Sporns, and G. M. Edelman, "A measure for brain complexity: Relating functional segregation and integration in the nervous system," Proc. Natl. Acad. Sci. U.S.A. **91**, 5033–5037 (1994).
[30]P. L. Williams and R. D. Beer, "Nonnegative decomposition of multivariate information" arXiv:1004.2515 (2010).
[31]J. T. Lizier, "JIDT: An information-theoretic toolkit for studying the dynamics of complex systems," Front. Rob. AI **1**, 11 (2014).
[32]A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," Phys. Rev. E **69**, 066138 (2004).
[33]C. M. Holmes and I. Nemenman, "Estimation of mutual information for real-valued data with error bars and controlled bias," Phys. Rev. E **100**, 022404 (2019).
[34]M. Wibral, N. Pampu, V. Priesemann, F. Siebenhühner, H. Seiwert, M. Lindner, J. T. Lizier, and R. Vicente, "Measuring information-transfer delays," PLoS One **8**, e55809 (2013).
[35]Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," Ann. Stat. **29**, 1165–1188 (2001).
[36]S. E. Said and D. A. Dickey, "Testing for unit roots in autoregressive-moving average models of unknown order," Biometrika **71**, 599–607 (1984).
[37]Implementation of ADF test in Python[46] with version v0.14.0.dev0.
[38]Note that the results reported in Fig. 5 are local information dynamics values, and each local value is estimated for each individual run. For non-stationary time series, an appropriate treatment would be to obtain each local value from an ensemble of time series available.[47] However, in financial time series, usually only one trajectory is available. What we aimed to show with our controlled vector autoregressive experiments of regime shifts in Fig. 5 was that certain regime shifts are detectable even from a single time series realization, and the information-theoretic signature is visible for a wide range of such non-stationary time series (see Appendixes A–E).
[39]T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," J. Econ. **31**, 307–327 (1986).
[40]F. Lillo and J. Doyne Farmer, "The key role of liquidity fluctuations in determining large price changes," Fluct. Noise Lett. **5**, L209–L216 (2005).
[41]J. Farmer, L. Gillemot, F. Lillo, S. Mike, and A. Sen, "What really causes large price changes?," Quant. Finance **4**, 383–397 (2004).
[42]H. Matsuda, K. Kudo, R. Nakamura, O. Yamakawa, and T. Murata, "Mutual information of ising systems," Int. J. Theor. Phys. **35**, 839–845 (1996).
[43]L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, "Information flow in a kinetic ising model peaks in the disordered phase," Phys. Rev. Lett. **111**, 177203 (2013).
[44]A. Johansen, D. Sornette, and O. Ledoit, "Predicting financial crashes using discrete scale invariance" arXiv:cond-mat/9903321 (1999).
[45]B. L. Welch, "The generalization of 'student's' problem when several different population variances are involved," Biometrika **34**, 28–35 (1947).
[46]S. Seabold, and J. Perktold, "Statsmodels: Econometric and statistical modeling with python", in *9th Python in Science Conference* (2010).
[47]G. Gómez-Herrero, W. Wu, K. Rutanen, M. C Soriano, G. Pipa, and R. Vicente, "Assessing coupling dynamics from an ensemble of time series," Entropy **17**, 1958–1970 (2015).