



HAL
open science

Riding Wavelets: A Method to Discover New Classes of Financial Price Jumps

Cecilia Aubrun, Rudy Morel, Michael Benzaquen, Jean-Philippe Bouchaud

► **To cite this version:**

Cecilia Aubrun, Rudy Morel, Michael Benzaquen, Jean-Philippe Bouchaud. Riding Wavelets: A Method to Discover New Classes of Financial Price Jumps. 2024. hal-04735506

HAL Id: hal-04735506

<https://hal.science/hal-04735506v1>

Preprint submitted on 14 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Riding Wavelets: A Method to Discover New Classes of Financial Price Jumps

Cecilia Aubrun^{a,b,1}, Rudy Morel^{c,a,d,1,2}, Michael Benzaquen^{a,b,e}, and Jean-Philippe Bouchaud^{a,e,f}

This manuscript was compiled on October 14, 2024

We introduce an unsupervised classification framework that leverages a multi-scale wavelet representation of time-series and apply it to stock price jumps. In line with previous work, we recover the fact that time-asymmetry of volatility is the major feature that separates exogenous, news-induced jumps from endogenously generated jumps. Local mean-reversion and trend are found to be two additional key features, allowing us to identify new classes of jumps. Using our wavelet-based representation, we investigate the endogenous or exogenous nature of co-jumps, which occur when multiple stocks experience price jumps within the same minute. Perhaps surprisingly, our analysis suggests that a significant fraction of co-jumps result from an endogenous contagion mechanism.

price jumps | classification | endogeneity | mean-reversion | trend | wavelets | co-jumps

Extreme events and cascades of events are widespread occurrences in both natural and social systems (1). Examples include earthquakes, volcanic eruptions, hurricanes, epileptic crises (2, 3), epidemic spread, financial crashes (4–6), economic crises (7, 8), book sales shocks (9, 10), riot propagation (11, 12) or failures in socio-technical systems (13). Understanding the origin of such events is essential for forecasting and possibly stabilizing their dynamics.

A widely studied question is the reflexive, self-exciting nature of those shocks. The concept of *financial market reflexivity* was introduced by Soros in (14), to describe the idea that price dynamics are mostly endogenous and arise from internal feedback mechanisms, as was first surmised by Cutler, Poterba and Summers in 1988 (15) (see also (16)). Extreme events, in particular, often arise from feedback mechanisms within the system’s structure (1, 17, 18). Quantifying the extent of endogeneity in a complex system and distinguishing events caused by external shocks from those provoked endogenously, and more generally identifying different classes of events, are crucial questions.

Prior research has proposed to differentiate between endogenous and exogenous dynamics by analyzing the profile of activity around the shock (9, 10, 19, 20), in particular in the context of financial markets (21–23). It has been observed that endogenous shocks are preceded by a growth phase mirroring the post event power-law relaxation, in contrast to exogenous shocks that are strongly asymmetric. The universality of this result is quite intriguing as they have been observed in various contexts: intra-day book sales on Amazon (9, 10), daily views of YouTube videos (20) and intra-day financial market volatility and price jumps (23, 24). Meanwhile, Wu *et al.* (25) differentiate exogenous and endogenous bursts of comment posting on social media using the analysis of collective emotion dynamics and time-series distributions of comment arrivals.

Furthermore, in complex systems, events can propagate along two directions: temporally and towards other elements of the system. Financial markets offer an attractive setting for studying multi-dimensional shocks due to the abundance of available data, the frequent occurrence of financial shocks and price jumps and the inter-connectivity of markets. In fact, a recent study by Lillo *et al.* (26, 27) demonstrates the frequent occurrence of “co-jumps”, defined as simultaneous jumps of multiple stocks (as illustrated in Fig. 1) and establishes a correlation between their prevalence and the inter-connectivity of different markets.

In this paper, we address the problem of classifying financial price jumps (and co-jumps), in particular measuring their self-exciting character, by analyzing their time-series using wavelets. We introduce an unsupervised classification based on an embedding $\Phi(x)$ of each jump time-series of returns $x(t)$ into a low dimensional-space more appropriate to clustering. Such embedding, composed of wavelet scattering coefficients (see (28) and below), relies on wavelet coefficients of the time-series at the time of the jump $t = 0$ and wavelet coefficients of volatility. Such coefficients are

Significance Statement

Cascades of events and extreme occurrences have garnered significant attention across diverse domains like seismology, neuroscience, economics, finance, and other social sciences. Such events may arise from internal system dynamics (endogenous) or external shocks (exogenous). Devising rigorous methods to distinguish between them is vital for professionals and regulators to create early warning systems and effective responses. Understanding these dynamics could improve the stability and resilience of crisis-prone socio-economic systems. We show how wavelets can be used for the unsupervised separation of shocks in financial time-series, based on time-asymmetry around the shock. Additionally, we highlight the significant role contagion mechanisms play in financial markets.

Author affiliations: ^aChair of Econophysics and Complex Systems, École polytechnique, 91128 Palaiseau Cedex, France; ^bLadHyX UMR CNRS 7646, École polytechnique, 91128 Palaiseau Cedex, France; ^cÉcole Normale Supérieure, 45 rue d’Ulm, 75005 Paris; ^dFlatiron Institute, 162 5th Av, New York, NY10010, US; ^eCapital Fund Management, 23 Rue de l’Université, 75007 Paris, France; ^fAcadémie des Sciences, 23 Quai de Conti, 75006 Paris, France

C.A. made significant contributions to the implementation, particularly in data processing and classification, and also substantially to the writing. R.M. made significant contributions to the implementation, specifically in wavelet-based coefficients and classification, and also to the writing. M.B. provided valuable input for presenting results and editing the manuscript. J.P.B. provided guidance on experimental design, classification methodology, and interpretation of results, and made significant contributions to the writing.

CFM is an asset management firm.

¹C.A. and R.M. contributed equally to this work

²To whom correspondence should be addressed. E-mail: rmorel@flatironinstitute.org

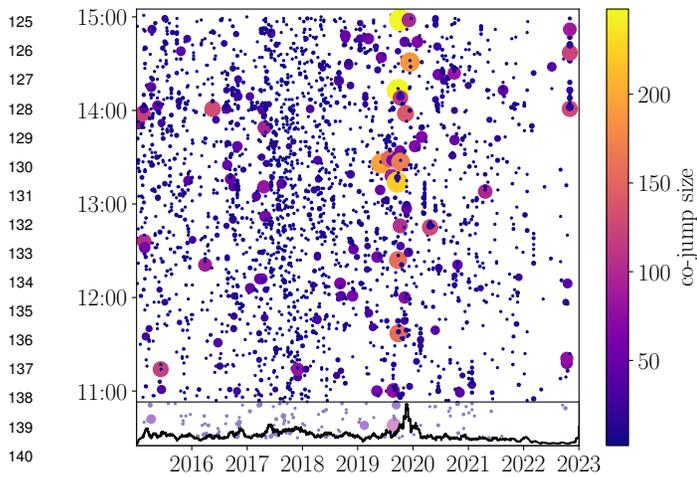


Fig. 1. Visualization of our co-jumps dataset (295 US stocks, 8 years) (as in (26, 27, 29)). The horizontal axis corresponds to the day of the co-jump and the vertical axis gives the time of day. Each circle denotes a co-jump. The size and color of the circle encode the number of stocks jumping simultaneously (in the same minute). Inset: number of jumps on a rolling window of 30 days. The maximum is reached in October 2019 with 2003 jumps.

particularly suitable to characterize (among other properties) the asymmetry of time-series at multiple scales.

Through a Principal Component Analysis we retrieve the fact that time-asymmetry of volatility indeed plays an important role for classification. However, our analysis identifies two further crucial features for characterizing the nature of price jumps: local mean-reversion and local trend. Specifically, mean-reverting jumps are such that pre-jump and post-jump returns are of opposite signs, whereas trend-aligned and trend-anti-aligned jumps occur on a sequence of returns of same sign before and after the jump, but either aligned with the jump itself, or of opposite sign.

For each jump, our analysis provides a measure of the volatility asymmetry, the mean-reversion and the trend. We propose a visualization of our dataset of price jumps in the form of *two* 2D projections. For both projections, one direction characterizes price jumps based on volatility asymmetry, or “exogeneity score”. The second direction characterizes jumps either in terms of mean-reversion, or in terms of alignment with the local trend behavior. One can then measure the endogeneity of price co-jumps, revealing that many jumps/co-jumps are *not* related to news and arise only due to endogenous dynamics. This is consistent with the observed power-law distribution of the number of firms affected by a co-jump, indeed predicted by a simple branching (or contagion) process.

Surprisingly, we uncover that a significant number of large co-jumps (affecting a large number of stocks), which might have been assumed to be caused by a common factor and thus share analogous dynamics, actually have uncorrelated returns both pre- and post-jump. This again suggests that such jumps are mostly of endogenous origin and result from a contagion mechanism.

The outline of our paper is as follows. Section 1 describes our dataset of price jumps resulting from Marcaccioli *et al.* (23), reviews their supervised classification method based on news labels, and investigates its limitation. Section 2

presents our unsupervised classification of univariate jump time-series based on wavelet coefficients. Such classification identifies three main directions in the dataset, the time-asymmetry, the mean-reversion and the trend. Finally, section 3 is devoted to the characterization of the endogeneity of co-jumps.

1. Supervised classification through endogeneity

Prior work has identified endogeneity as an important feature for the classification of jumps in financial markets (23, 24). Given the time-series of a jump, the main challenge is to efficiently measure such endogeneity.

One can for example look at contemporaneous news labels to determine whether or not a jump is exogenous. Indeed, news labels may serve as ground truth to learn a classification model on the activity profile around a shock. To exemplify, Fig. S1 (Supplementary Materials) from the work of Marcaccioli *et al.* (23), illustrates the time asymmetry difference between endogenous and exogenous jumps.

In this section, we first introduce the jump detection method, which allows us to build our dataset. Then, we present the supervised classification based on news labels introduced in (23) and show its limitations. This will motivate an alternative approach in section 2.

A. Jump detection. We refer to (23, 24, 30) for a detailed description of the method to detect price jumps. The detection relies on an estimator of “jump-score” $x(t) = r(t)/(f(t)\sigma(t))$, which is the ratio of 1-minute returns time-series $r(t)$ and de-seasonalized local volatility $f(t)\sigma(t)$ where $\sigma(t)$ is an estimator of local volatility and $f(t)$ an estimator of the intra-day periodicity (the so-called “U-shape”). Throughout this paper, our statistical analyses will focus on $x(t)$, or on its “jump-aligned” version $\bar{x}(t) := x(t)\text{sign}(x(0))$, where $x(0)$ is the return corresponding to the jump. In other words, $\bar{x}(t)$ is the rescaled return profile in the direction of the jump.

Under the null hypothesis of Gaussian residuals (no jump hypothesis) $|x(t)|$ converges towards a Gumbel distribution. A statistical test then allows us to reject the null hypothesis. The resulting method comes down to detecting price movements deviating by more than 4 standard deviations from their average value (here equal to zero).

The jump detection is performed on time-series describing individual stocks dynamics but also on averaged time-series across stocks belonging to the same sector. Hence, we obtain price jumps of individual stocks but also sectoral price jumps.

Similarly to Marcaccioli *et al.* (23), we find that price jumps are clustered in time. We assume that jumps taking place within the same “time-cluster” subsequent to an initial jump are merely replicas of the initial jump. They are likely to be either of the same dynamics (as they occurred for the same reason) or endogenously induced by the first jump of the cluster. We thus discard jumps that follow an initial jump if the inter-time is statistically unlikely under a null hypothesis of independent jumps (23). This leads to the same detection method as in (23) which allows to retrieve an exponential distribution for the inter-time between two consecutive jumps (see part II.D of (23)).

From such a collection of price jumps, we can then extract “co-jumps”. A co-jump is simply defined as a set of jumps in different stocks occurring in the same minute. Here we

249 avoid tackling the question of lagged jumps and consider only
250 simultaneous jumps (up to the minute resolution).

251 The price behavior before and after a jump can be used
252 to classify the jump. In light of Marcaccioli *et al.*'s findings
253 (23), which indicate that volatility can begin to rise up to
254 75 minutes prior to the jump, we adopt a time window of 2
255 hours centered around the jump occurrence at time $t = 0$.
256 Consequently, for each jump we extract a time-series of 119
257 rescaled returns $x(t)$, corresponding to 1 hour preceding the
258 jump and 1 hour following the jump.

259 We implement such detection on 301 US stocks from
260 January 2015 to December 2022, considering only what
261 happened between 10:30 and 15:00 in order to avoid special
262 jumps due to the high activity at the beginning (due to people
263 reacting to the overnight news) and at the end of the day
264 (due to market closing). In order to discard major market
265 shocks, we also remove all co-jumps involving more than 250
266 stocks, and days on which the FED made an announcement
267 (1 per month*). We end up with 43 628 jumps, of which
268 18 802 belong to one of the 2905 co-jumps, and the remainder
269 (24 826) are single jumps.

270 **B. Classification based on news labels.** In an attempt to
271 characterize the endogeneity of a jump, one can gather the
272 date and time of news associated to each stock we consider†
273 and of the main US announcements‡. According to such
274 news labels, we might label as “news-related” a jump which
275 happened within 3 minutes of a news and label as “non-news-
276 related” any other jump. That would lead to a puny $\approx 4.3\%$ of
277 the jumps being classified as “news-related” and is illustrated
278 in Fig. S1 (Supplementary Materials). Hence, as previously
279 argued in (15, 16, 23, 24, 26), it appears that individual price
280 jumps and more surprisingly co-jumps themselves are often
281 *not* related to news announcements.

282 However, it is clear that some news may affect a whole
283 economic sector and lead to a co-jump without appearing in
284 our considered set of news. An example would be an OPEC
285 announcement that affects oil prices and in turn ricochets
286 onto stocks prices, without any of them explicitly showing up
287 in the news feed. Another vivid example is the impeachment
288 of the US president D. Trump in September 2019§. Our
289 “news-related” label is blind to such events. One objective
290 of our study will be to propose a possible classification of
291 co-jumps that does not rely on the news feed, see section 3.
292

293 **C. Classification based on the volatility profile.** In (23),
294 Marcaccioli *et al.* built a supervised classification of uni-
295 variate jumps into exogenous and endogenous classes. The
296 classification relies on parameters derived from fitting $|x(t)|$
297 to the following functional form (10):
298

$$299 |x(t)| = \mathbf{1}_{t < t_c} \frac{N_{<}}{|t - t_c|^{p_{<}}} + \mathbf{1}_{t > t_c} \frac{N_{>}}{|t - t_c|^{p_{>}}} + d \quad [1]$$

300 and on a measure of the asymmetry of the jump, defined as:

$$301 A_{\text{jump}} = \frac{\mathcal{A}_{>} - \mathcal{A}_{<}}{\mathcal{A}_{>} + \mathcal{A}_{<}} \quad [2]$$

306 * see FOMC Calendars

307 † source: Bloomberg

308 ‡ source: economic-calendar

309 §For example, the largest co-jump is related to Nancy Pelosi announcement of a formal
310 impeachment inquiry into US President Donald Trump. On 2019-09-24, at 14:13, 248 stocks saw
their price jump in the same minute.

311 where $\mathcal{A}_{</>} := \sum_{t < 0 / t > 0} |x(t) - \min_{t < 0 / t > 0} (x(t))|$ and
312 d denotes the baseline volatility. Such an indicator means
313 that when the activity is stronger before (resp. after)
314 the jump, one has $A_{\text{jump}} < 0$ (resp. $A_{\text{jump}} > 0$). The
315 classification is then obtained as a logistic regression of
316 the news label (endogenous/exogenous) by the parameters
317 ($\mathcal{A}_{\text{jump}}, p_{<}, p_{>}, N_{<}, N_{>}, t_c$). Exogenous jumps appear as
318 strongly asymmetric jumps with little activity ahead of the
319 jump, i.e. $A_{\text{jump}} > 0$, whereas self-exciting endogenous jumps
320 are much more symmetric with $A_{\text{jump}} \approx 0$ (23).

321 The above approach, based on news labels, presents several
322 limitations:

- 323 • As discussed above, news labels might miss some relevant
324 economic news, so the resulting price jumps might be
325 wrongly labeled as “non news-related”.
326
- 327 • Exogenous jumps could have two types of dynamics: if
328 the exogenous shock is a complete surprise, there should
329 indeed be no activity before the jump. However, if the
330 announcement is planned or if there was some news
331 leakage, there might be a growth of activity before the
332 jump. In this case, one would wrongly classify a news-
333 related jump as endogenous based on its approximately
334 symmetric activity profile.
335

336 In light of such limitations and in order to uncover new classes
337 of jumps, we opt in the rest of the paper for an unsupervised
338 classification which significantly improves upon the method
339 of (23) while still leaving open some ambiguities, as we will
340 see below.

341 Although news labels do not reveal the whole truth about
342 the endogenous nature of a jump, we will henceforth still
343 call “news-related” jumps that occurred within 3 minutes of
344 a news present in our database and “non news-related” all
345 the others.
346

347 2. Classification of single jumps using wavelets

348 The rescaled returns time-series around a jump $x(t) \in \mathbb{R}^T$
349 is inherently noisy. Relevant features $\Phi(x) \in \mathbb{R}^q$ must be
350 extracted to effectively distinguish different classes of jumps.
351 Such features should be selected carefully, in particular, they
352 should include time-asymmetry measures. Indeed, authors
353 in (9, 10, 19, 20, 23) show that the jumps mostly differ in
354 their time-asymmetry: endogenous jumps tend to be more
355 symmetric around the jump than exogenous ones. But what
356 are the other possibly relevant features?
357

358 In this section, we embrace a signal processing approach
359 to discover important features of univariate jumps and unveil
360 new classes of jumps that are prevalent in the data.
361

362 **A. Wavelet and scattering coefficients.** Wavelet filters have
363 been used to analyze, classify and detect transient events,
364 see e.g. (31–35). A complex wavelet filter $\psi(t)$ is a filter
365 whose Fourier transform $\hat{\psi}(\omega) = \int \psi(t) e^{-i\omega t} dt$, is real. It is
366 localized both in time and Fourier domains, see Fig. 3. It has
367 a fast decay away from $t = 0$ and a zero-average $\int \psi(t) dt = 0$.
368 We write $\psi(t) = \text{Re } \psi(t) + i \text{Im } \psi(t)$ where $\text{Re } \psi(t)$ and $\text{Im } \psi(t)$
369 are its real and imaginary parts. They are respectively even
370 and odd functions:
371

$$372 \text{Re } \psi(-t) = \text{Re } \psi(t) \quad \text{and} \quad \text{Im } \psi(-t) = -\text{Im } \psi(t). \quad [3]$$

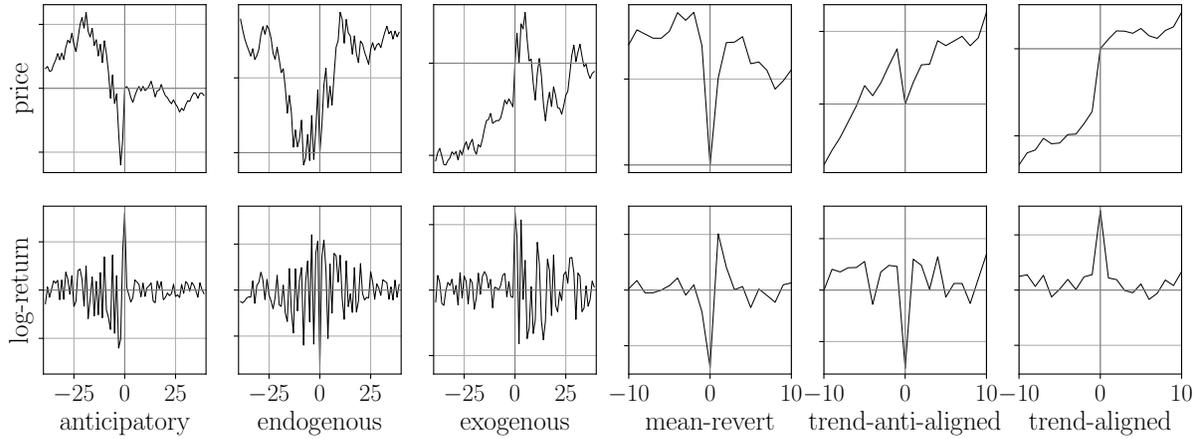


Fig. 2. Classes of price jumps (synthetic examples). Each column shows an example of a class of jumps (price and log-return time-series). The three first classes (anticipatory, endogenous, exogenous) are separated by measuring volatility asymmetry. The three last classes (mean-reverting, trend-anti-aligned, trend-aligned) are identified by analyzing the signed returns around the jump. See Fig. S9 for examples of true observed jumps.

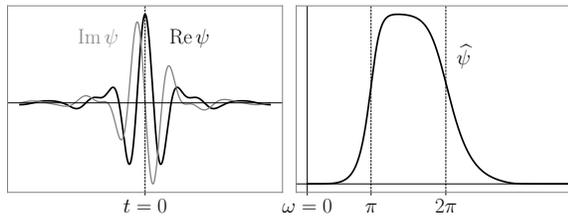


Fig. 3. Filter used to analyze jump time-series. Left: complex Battle-Lemarié wavelet $\psi(t)$ as a function of t . Right: Fourier transform $\hat{\psi}(\omega)$ as a function of ω .

The wavelet coefficients $W_j x(t)$ compute the variations of the signal x around t at scale 2^j , for $j = 1, \dots, J$ with

$$W_j x(t) := x \star \psi_j(t) \quad \text{where} \quad \psi_j(t) = \psi(2^{-j}t). \quad [4]$$

where \star denotes the convolution: $x \star y(t) := \int x(t-u)y(-u)du$.

The sign of the jump $\text{sign}(x(0))$ and its amplitude $|x(0)|$ vary, but they are not necessarily informative for their classification. To remove this source of variability we consider the *jump-aligned* time-series

$$\bar{x}(t) = \text{sign}(x(0))x(t) \quad [5]$$

and we further normalize the wavelet coefficients (4) by the corresponding “volatility” σ_j of the full time-series, defined as $\sigma_j^2 = \langle |x \star \psi_j(t)|^2 \rangle_t$, where $\langle \cdot \rangle_t$ denotes the empirical average over time t .

From Eq. (3), one can see that if x is an even signal i.e. $x(-t) = x(t)$ then $\text{Im} W_j x(t) \equiv 0$. This property is key to detect asymmetry of a signal at different scales.

Volatility information can be extracted by taking a modulus. The time-series $|W_j x(t)|$ provides the volatility of the signal at scale 2^j . This volatility can be asymmetrical in $t = 0$. In order to quantify it, we again consider the wavelet coefficients at $t = 0$

$$W_{j_2} |W_{j_1} x|(t) := |x \star \psi_{j_1}| \star \psi_{j_2}(t). \quad [6]$$

Our representation for univariate jumps in this paper is thus composed of wavelet coefficients (4) at $t = 0$ and

scattering coefficients (6) at $t = 0$

$$\Phi(x) = \left(W_j \bar{x}(0), W_{j_2} |W_{j_1} x|(0) \right). \quad [7]$$

For a time-series of size T , it contains less than $(\log_2 T)^2/2$ coefficients which represents few coefficients. In our case, $T = 119$ and we chose $J = 6$, which yields 42 coefficients (21 real parts and 21 imaginary parts). The normalized scattering features $\Phi(x)$ (Eq. (7)) are invariant to sign changing and to dilation

$$\Phi(-x) = \Phi(x) \quad \text{and} \quad \Phi(\lambda x) = \Phi(x).$$

which means we do not aim at discriminating jumps neither based on their sign nor on their amplitude.

In order to classify price jumps, we are interested in Principal Component directions of the 42-dimensional vector $\Phi(x)$ in the dataset. This method, called kernel PCA (36), relies on the linear separation power of our scattering coefficients $\Phi(x)$. We considered several directions, i.e. combinations of scattering coefficients, and found three salient features: the time-asymmetry of the volatility, the mean-reversion and the trend behavior of the price around the jump.

B. First Direction D_1 : Volatility asymmetry.

B.1. Three types of jumps. The first PCA direction (called D_1 henceforth) is a linear combination of the 15 coefficients $\text{Im} W_{j_2} |W_{j_1} x|(0)$ in Eq. (7), which characterizes time-asymmetry of the *volatility profile* at multiple scales 2^{j_2} , confirming a previous analysis that postulated this asymmetry to be relevant (23). Such a linear combination allows one to embed each jump time-series into a one dimensional space, which quantifies the endogenous nature of each jump. In fact, Fig. 4 (see also Fig. S2 in Supplementary Materials) displays average profiles $|x(t)|$ along the “endogenous direction” D_1 . One can visually verify that such a representation discriminates jumps according to the asymmetry of their profiles as measured by $\mathcal{A}_{\text{jump}}$ (Eq. (2)): the D_1 direction continuously separates asymmetric jumps with dominant activity before the shock from asymmetric jumps with dominant activity

after the shock; see Fig. 4 and Figs. S2,S3 in Supplementary Materials.

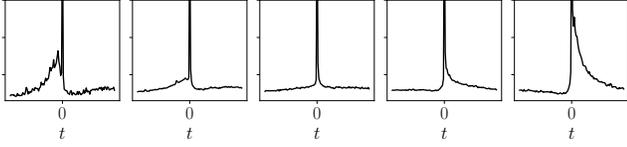


Fig. 4. Average absolute profiles $|x(t)|$ of jumps along direction D_1 (sliced into five bins, delimited by quantiles 0.1, 0.25, 0.35, 0.9). From left to right: anticipatory jumps, endogenous jumps and exogenous jumps.

From this analysis, three types of jumps can thus be defined:

- Symmetric jumps, with an pre-shock excitation activity that approximately mirrors the post-shock relaxation activity. These were called “endogenous jumps” in (23): increased activity before the jump is in fact responsible for the jump itself, with some decay of activity thereafter. The symmetry of the profile for endogenous jumps is in fact predicted by a Hawkes process description of the self-exciting mechanism, see (10, 23).
- Asymmetric jumps with dominant activity after the shock. These were called “exogenous jumps” in (23): the market reacts *after* unexpected news, but not before.
- Asymmetric jumps with dominant activity *before* the shock. This type of jumps, which we call “anticipatory”, was quite unexpected and was not discussed in (23).

In order to validate the above analysis, we created synthetic time-series with volatility profiles of varying time-asymmetry and applied our classification method. Results of this benchmark case are shown in Supplementary Materials (Fig. S4 and Section 1), and fully confirm that the D_1 direction indeed separates jumps according to their asymmetry $\mathcal{A}_{\text{jump}}$.

B.2. Discussion. Using the above classification, we find that a large proportion ($\sim 50\%$) of our sample exhibit positive asymmetry and thus should, at least naively, be considered as exogenous jumps. This seems in contradiction with the results of (23), where exogenous jumps were found to be a minority, and with a fraction of jumps associated to a news found to be 4.3%, as already quoted above. Two arguments can explain such a difference.

- The main one is the fact that our analysis includes all jumps involved in a sector jump (corresponding to 24% of all jumps) whereas those jumps were discarded in (23). Sector jumps are such that many stocks of the same industry jump simultaneously. While some of these jumps are likely due to major exogenous shocks – like macro-economic announcements – that affect a whole economic sector or even the whole market, we argue in section 3 that these jumps can actually be induced by a jump of one particular stock of the sector, which is deemed as “news” in and by itself. In any case, taking these sector jumps into account mechanically increases the count of jumps with a positive D_1 score. In the present study, we chose to keep these co-jumps and

study their statistics, to which we will specifically turn in section 3.

- As already noted above, the classification of single jump profiles in (23) relies on the goodness of fit of power law function Eq. (1), and as such, was only conducted on a smaller sample for which such a fit is acceptable (~ 5000 jumps out of ~ 37000 jumps).

The appearance of “anticipatory jumps”, where the asymmetry parameter $\mathcal{A}_{\text{jump}}$ (see Eq. (2)) is negative, came somewhat as a surprise to us. One possible interpretation is that these jumps are in fact also endogenous, with a pre-shock self-exciting dynamics and very little “after-shocks”. Indeed, if such jumps are immediately deemed endogenous by the market, it might make sense that activity quickly reverts back to normal. This would simply mean that it is necessary to extend the Hawkes framework, which currently predicts a symmetric profile, to adequately describe all endogenous shocks. Such extensions could involve generalizing the marks distributions within the Hawkes model, as discussed in (37–40).

Another possibility is that such events correspond to news/exogenous events whose *timing* is expected by the market, or if there was some leakage ahead of the news, both leading to increased activity before the actual release time. Now, if the actual news content turns out to be insignificant, it would again make sense that the market activity quickly wanes off. We in fact do find a very small fraction of news-related jumps with $D_1 < 0$, see in Fig. 6, bottom graph.

C. Second Direction D_2 : Mean-Reversion.

C.1. Capturing Mean-Reversion. We observed that coefficients $\text{Im } W_{j_1} \bar{x}(0)$ (7) for fine scales, i.e. small j_1 , are consistently chosen by the leading PCA directions. They amount to multiplying the jump-aligned time-series $\bar{x}(t)$ by the imaginary filter $\text{Im } \psi_1(t)$ (see Fig. 3) and averaging over t . Such coefficients capture the asymmetry of the return profile shortly before and shortly after the jump, and define what we will call below direction D_2 .

A typical time-series that maximizes this coefficient is thus characterized by a positive value of $\bar{x}(-1)$ and a negative value of $\bar{x}(1)$. In other words, large positive values along the D_2 coordinate capture mean-reverting return profiles, i.e. positive (resp. negative) returns before a positive (resp. negative) jump that become negative (resp. positive) immediately after the jump.

Large negative values along the D_2 coordinate, on the other hand, also capture mean-reverting return profiles, but in this case mean-reversion starts with (or is triggered by?) the jump itself, and not after the jump.

Now that we identified a potentially discriminating direction using PCA, we transition to a simpler filter tailored to capture short time mean-reversion, depicted in Fig. 5. This filter is then applied to the jump-aligned time-series $\bar{x}(t)$

$$\tilde{D}_2(x) := \bar{x} \star \psi_{\text{MR}}(0), \quad [8]$$

where the tilde indicates that we have simplified the true second PCA direction and only retained the component spanned by ψ_{MR} .

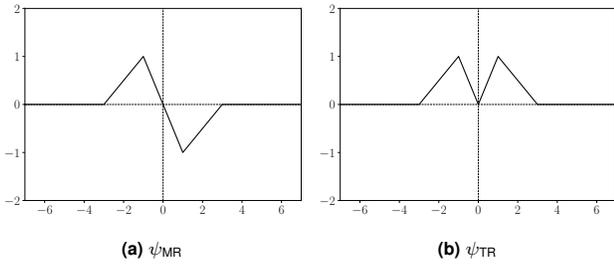


Fig. 5. Handcrafted filters for measuring the mean-reversion (filter ψ_{MR}) or the trend (filter ψ_{TR}) character of a jump. Average profiles along resulting mean-reversion and trend directions are shown in Fig. 7 and Fig. 8.

C.2. A 2D representation of jumps. Based on the first volatility asymmetry direction D_1 and the mean-reversion direction \tilde{D}_2 , we are in a position to propose the 2D representation of jumps shown in Fig. 6 (top), in which the horizontal axis corresponds to D_1 and the vertical axis corresponds to the mean-reversion index \tilde{D}_2 . Visually, news-related jumps are mostly to the right of the projection, corresponding to increased volatility after the jump, as expected. This is confirmed by the plot in Fig. 6 (bottom).

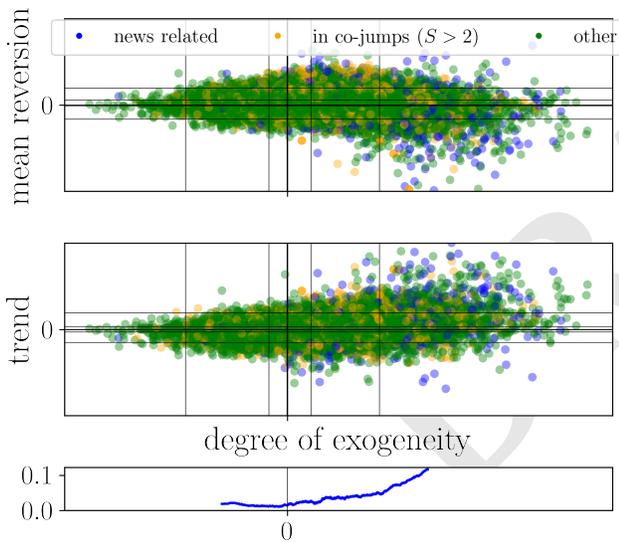


Fig. 6. Top graph: Projection of jumps in our dataset onto the endogenous direction D_1 (horizontal axis) and mean-reverting direction \tilde{D}_2 (vertical axis). Middle graph: Projection of our dataset on the endogenous direction D_1 (horizontal axis) and trend direction \tilde{D}_3 (vertical axis). Each point represents a jump, the blue color corresponds to news-related jumps according to the classification of Section B, the oranges are jumps involved in a co-jump of size greater than 2 and non news related and the greens are all the other jumps. The vertical and horizontal lines represent the following quantiles: 0.05, 0.35, 0.65, 0.95. Bottom graph: ratio of “news-related” jumps along the endogenous direction D_1 , based on a direct classification using the news feed (rolling ratio every 2000 jumps). This ratio clearly increases as we move to positive values of D_1 .

In order to illustrate the discriminating power of such coefficient, Fig. 7 displays the average profiles of $\bar{x}(t)$ along the \tilde{D}_2 axis. One can see that jumps with a high coefficient \tilde{D}_2 (rightmost graph) are characterized by a strong pre-jump trend aligned with the jump, followed by a change of sign in the next minute after the jump (as also shown in Supplementary Materials, Fig. S5).

The leftmost graph, on the other hand, shows relatively mild pre-jump trends opposite to the jump, followed by stronger trends in the direction of the jump, not very different from the cases corresponding to quantiles between 0.1 and 0.5. In our dataset, 60% of the jumps have a positive mean-reversion score $\tilde{D}_2 > 0$; we refer to Fig. S6 (Supplementary Materials) for the full distribution of \tilde{D}_2 .

To confirm this observation and ascertain that it is not attributable to spurious effects in the data processing, we looked deeper into these jumps. To get a better understanding of the mechanisms at play, we investigated what happens at tick-by-tick scale in the Limit Order Book. We show in Fig. S7 (Supplementary Materials) two illustrative examples. We again observe, at a different time resolution, a strong mean-reversion behavior induced by order placement. Note that both exogenous, or endogenous jumps can have such mean reverting behavior, as clear from the 2D representation Fig. 6.

In fact, a mean reverting behavior can be expected both following an exaggerated response to a news release, or after a self-initiated jump with no discernible catalyst. This is confirmed by Fig. S8 (Supplementary Materials) which shows positive average values of \tilde{D}_2 for all levels of endogeneity D_1 , except for strongly exogenous jumps (large values of $D_1 > 0$), where the mean-reversion disappears ($\tilde{D}_2 \approx 0$).

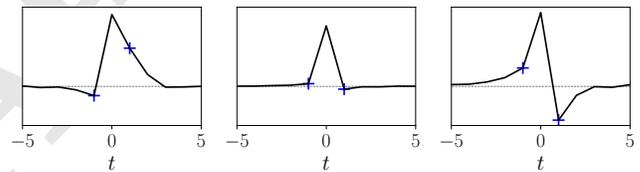


Fig. 7. Mean-reverting profiles. Average jump-aligned return profiles $\bar{x}(t) = \text{sign}(x(0))x(t)$ along the mean-reverting direction \tilde{D}_2 (sliced into four bins, delimited by quantiles 0.1 & 0.9). Left-most graph: price jumps mean-revert on previous trends. Right-most graph: prices mean-revert after the jump.

Note finally that mean-reversion is characterized by a V-shape price profile (see Fig. S7 in Supplementary Materials), which has recently been used as a criterion to detect price jumps in time-series ((41)).

D. Third Direction D_3 : Trend. In the previous section, we have defined a filter ψ_{MR} that detects mean-reversion, but is by construction orthogonal to trends, i.e. post-jump returns continuing in the same direction as pre-jump returns. This feature can be naturally captured by the trend filter ψ_{TR} shown in Fig. 5, which is orthogonal to the mean-reversion filter ψ_{MR} . This filter is then applied to the jump-aligned profile $\bar{x}(t)$ to get the following trend score

$$\tilde{D}_3(x) := \bar{x} \star \psi_{TR}(0). \quad [9]$$

A large positive value of $\tilde{D}_3(x)$ therefore describes a persistent trend aligned with the direction of the jump. If such jumps exist, we refer to them as “trend-aligned” jumps. A large negative value of $\tilde{D}_3(x)$ indicates that the jump goes against the pre- and post-jump trend. If such jumps exist, we refer to them as “trend-anti-aligned” jumps.

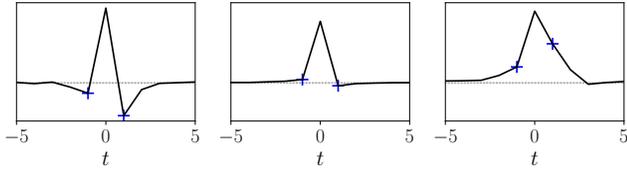


Fig. 8. Trending profiles. Average jump-aligned return profiles $\bar{x}(t) = \text{sign}(x(0)) x(t)$ along the trend direction \tilde{D}_3 (again sliced into four bins, delimited by quantiles 0.1 & 0.9). Left-most graph: anti-aligned trends. Right-most graph: aligned trends.

Fig. 8 shows that both classes of jumps do indeed exist: the average profiles in the first and last quantiles in Fig. 8 do conform to expectations. Furthermore we directly observe many stylized examples such as the one reported in Fig. S9 (Supplementary Materials). As for the mean-reversion indicator, we can represent all jumps in 2D plane based on D_1 and \tilde{D}_3 (see the bottom graph in Fig. 6). Visually, trending news-related jumps appear to be mostly aligned with the jump (top-right corner), although anti-aligned trends can also be spotted for moderate values of D_1 . Different profiles of $\bar{x}(t)$ corresponding to the grid are shown in Fig. S10 (Supplementary Materials).

E. Preliminary Conclusions. Let us summarize the results obtained by our unsupervised approach so far. First, our proposed 2D projections provide an embedding of a jump according to three meaningful, intuitive properties: its endogenous nature, its mean-reversion character or its trend character. On top of the separation between exogenous and endogenous jumps, our clustering method revealed new classes of jumps, some of which we did not expect *a priori*: anticipatory jumps, mean-reverting jumps, trend-aligned and trend-anti-aligned jumps. Identifying additional interpretable classes of jumps might be possible by considering more expressive wavelet-based embeddings such as Scattering Spectra recently used in the context of financial time-series (42, 43). However, our attempts so far seemed to mostly recover directions which overlap with the volatility time-asymmetry and mean-reverting directions.

3. Classification of co-jumps

As mentioned above, a “co-jump” is defined as a collection of jumps across several stocks, occurring in the same minute. The number S of assets involved in the co-jump is referred to as the “size” of the co-jump. Co-jumps reveal interconnectivity and contagion in financial markets (26, 27, 44). As such, studying them – in particular their possible endogenous nature – is a crucial question for investors and regulators alike. This section aims at investigating whether co-jumps are created through endogenous dynamics or exogenous shocks.

To assemble our co-jump dataset we consider the same dataset of jumps as in the previous section. We end up with 2905 co-jumps, the size of which varies from 2 stocks to 248 stocks. The co-jumps size distribution, is shown in Fig. 9a. Quite remarkably, the tail of this distribution is well fitted by a truncated power-law $S^{-1-\tau} \exp(-\varepsilon^2 S)$ with exponent $\tau = 1/2$, with a cut-off parameter $\varepsilon = 0.077$, see Fig. 9a, inset. Such a value for τ can be rationalized within the framework of classical critical branching processes (45), as if co-jumps

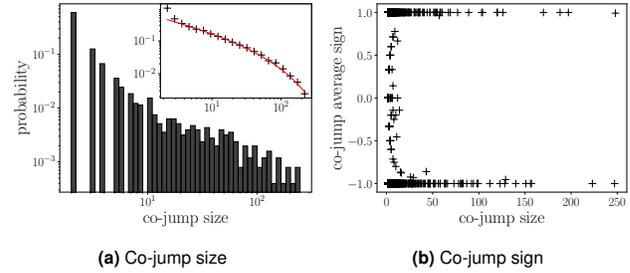


Fig. 9. Statistics on co-jumps. (a) Main: Distribution of co-jumps size i.e. number of stocks involved in a co-jump. Inset: Cumulative distribution of co-jumps size for co-jumps with $\min(D_1) < 0$ and $\min(D_1) < \overline{D}_1 - 1\sigma$, defining the LL and LR regions in Fig. 12. Inset: cumulative distribution in log-log coordinates, with a fit obtained corresponding to a truncated power-law $S^{-1-\tau} \exp(-\varepsilon^2 S)$ for the probability distribution function. We find that the best fit parameters are $\tau = 0.5$ and $\varepsilon = 0.077$. (b) Average sign of jumps involved in a co-jump, showing that most co-jumps are composed of jumps in the same direction.

were the result of a contagion mechanism, with ε the distance to the critical point. Such a power-law behaviour was in fact already noted in previous works: in Ref. (24) on a US data set from 2004 to 2006, in (27) from 2001 to 2013 and in (29) from 2013 to 2018.

Furthermore, the signs of the jumps involved in a co-jump are, most of the time, all aligned, i.e. different stocks jump in the same direction, as shown in Fig. 9b.

The first stage of co-jump characterization is to classify jumps according to their endogeneity coordinate along the D_1 direction. In Fig. 10, we highlight the coordinates of three particular co-jumps in the 2D projections introduced in the previous section. Each color point is a stock involved in one of the three co-jumps. Let us comment on each of these three cases in turn:

- The blue co-jump, with 29 stocks involved, has most of its elements in the right side of the 2D projection, suggesting an exogenous, news driven shock. However, one of the jump is below the 0.35 quantile and therefore appears endogenous. This might be a mis-classification because of the inherent noise in our D_1 exogeneity score. An alternative interpretation, in line with a contagion mechanism, might be that this particular stock jumped for no particular reason and this created a surprise to which other stocks reacted.
- The pink co-jump, with 19 stocks involved, staunchly belongs to the anticipatory class – which we believe to be of endogenous nature, as explained above. Co-jumps with a negative or positive but moderate maximum value of the D_1 score can thus be deemed endogenous.
- The yellow co-jump, with 9 stocks involved, has most of its elements in the intermediate “endogenous” region, except one which is classified as exogenous. This might be either again a mis-classification because of the inherent noise in our D_1 exogeneity score, or else a stock that was not part of the anomalous pre-jump activity but is drawn into the jump through contagion.

From these cursory observations, one may propose three natural indicators for classifying co-jumps:

1. The average value of the individual exogeneity score \overline{D}_1 over all jumps belonging to a given co-jump, see Fig. 11.

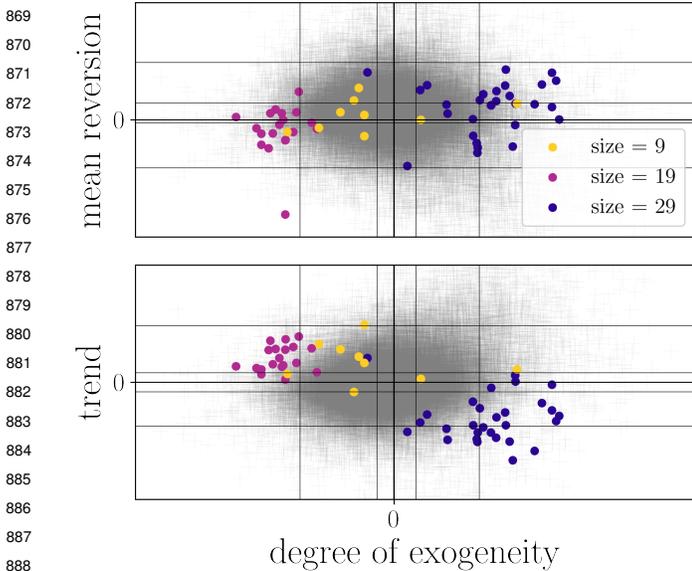


Fig. 10. Projections of 3 co-jumps along our 2D projections. Yellow co-jump: one jump is exogenous and the others are more endogenous. Pink co-jump: all jumps of the co-jumps are endogenous and are trend-aligned. Blue co-jump: Most jumps appear to be exogenous except one. Those jumps are also trend-anti-aligned.

2. The maximum value of the individual exogeneity score D_1 over all jumps belonging to a given co-jump: if the most exogenous jump is still deemed endogenous, the whole co-jump is classified as endogenous (see distribution in Fig. S11, Supplementary Materials).
3. The minimum value of the individual exogeneity score D_1 over all jumps belonging to a given co-jump: if the most endogenous jump is still deemed exogenous, the whole co-jump is classified as exogenous (see distribution in Fig. S12, Supplementary Materials).

Fig. 12 represents the normalized minimum value of exogeneity score D_1 over all jumps of a given co-jump as a function of the normalized average value of exogeneity score D_1 over all jumps of a given co-jump (co-jump indicator 3. as a function of co-jump indicator 1.). The normalization is such that Fig. 12 can be read in units of standard deviation of the exogeneity score D_1 for co-jumps of same size, i.e. σ is the average of the standard deviation of the score D_1 over co-jumps with same size. The size and color of a point depict the size of the co-jump. The gray shaded region represents jumps with insignificant differences between the mean and the minimum value of the D_1 score.

Co-jumps with negative minimum and average values of exogeneity score D_1 (lower left quadrant of Fig.12, LL) can be deemed endogenous, whereas co-jumps with positive minimum and average values of exogeneity score D_1 (upper right quadrant of Fig.12, UR) can be deemed exogenous.

The lower right quadrant (LR) represent the most intriguing co-jumps. Indeed, according to their average score D_1 those co-jumps should naively be classified as exogenous, however they contain at least one strongly endogenous co-jump. It might be that those endogenous jumps, whose pre-activity starts while most other stocks are still quiet, are interpreted by the market in and by themselves as news. Such

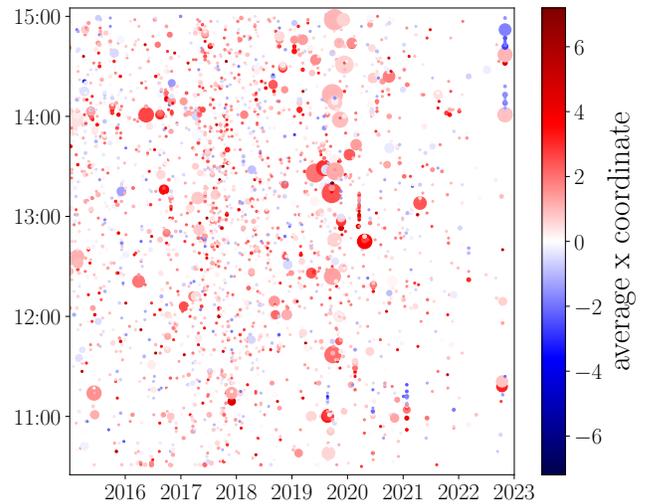


Fig. 11. Exogeneity score $\overline{D_1}$ of co-jumps in our dataset, obtained by averaging the exogeneity score D_1 of each jump involved in a co-jump. Large co-jumps tend to have a higher average score (in red) but, surprisingly, there are many large co-jumps with pale color that would be classified as endogenous. See discussion in the text.

surprise triggers all other jumps – which therefore appear as exogenous, with no special pre-jump activity but without being related to any news.

Note that the largest co-jumps are in the LR region; our interpretation in terms of a contagion mechanism would then naturally explain the truncated power-law distribution of size $S^{-3/2}$ shown in Fig. 9a.

There are obviously also large sector wide co-jumps that are truly news-related – upper-right quadrant of Fig. 12. For instance, the significant co-jumps highlighting the year 2019 mostly exhibit a negative average (exogenous) and are related to the announcements during the US vs China trade war.

Conversely, some co-jumps (20% of our sample) involve only jumps exhibiting a symmetric or anticipatory profile (LL region of Fig. 12). Those co-jumps are usually $S = 2$ stocks co-jumps (76%), but their size can go up to $S = 87$ stocks.

Hence, the most striking conclusion of this section is that many large co-jumps are in fact explained by endogenous dynamics and propagate across stocks, rather than being due to impacting external news. A(n) (in)famous example of such propagation is the flash crash of May 6th 2010, where the S&Pmini crashed in less than 30min, due to a sell algorithm set with an excessively high execution rate. This crash triggered a price drop in other US stocks. Here, our results suggest that this synchronization phenomenon is not such a rare event and actually happens quite often (26, 44).

This finding is further supported by examining the correlation of the individual jump time-series composing a co-jump (see section 3 in Supplementary Materials). Naively, one would expect large co-jumps to be exogenous, i.e. induced by news. As a result, the stocks involved in the co-jump should all share the same profile around the jump, as in the left graph of Fig. S13 (Supplementary Materials) for example. In fact, Fig. S14 (Supplementary Materials) shows that there remain many co-jumps whose constituting univariate jump profiles are weakly correlated (see Section 3 in Supplementary Materials for more details). We also refer the reader to

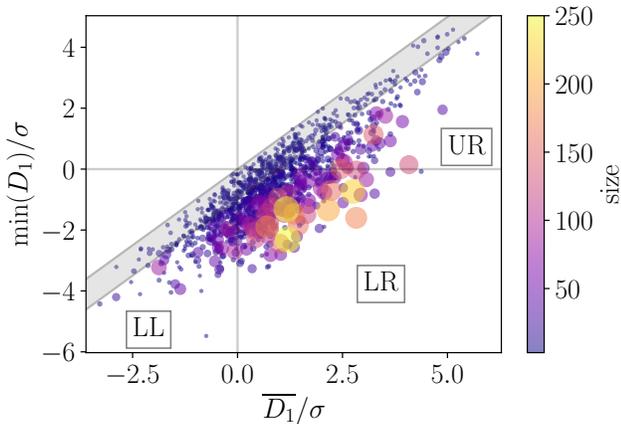


Fig. 12. Minimum value of exogeneity score D_1 over all jumps of a given co-jump as a function of the average value $\overline{D_1}$ of exogeneity score D_1 over all jumps of a given co-jump (co-jump indicator 3 as a function of co-jump indicator 1). Both indicators have been normalized by the average standard deviation of the exogeneity score D_1 of co-jumps with the same size σ . The size and color of a point depict the size of the co-jump. The grey area represents the zone between $\min(D_1) = \overline{D_1}$ and $\min(D_1) = \overline{D_1} - 1$, corresponding to co-jumps where the difference between the minimum and the average D_1 score is less than 1σ . Here, we only consider co-jumps with a size strictly greater than 2. LL, LR & UR stand for lower left, lower right and upper right.

additional statistics on co-jumps in Supplementary Materials. For example, Fig. S15 (Supplementary Materials) shows that the sector jumps are not all exogenous, as discussed in Section 2B.

4. Conclusion

Thanks to an unsupervised approach based on wavelet scattering coefficients, we have identified three main directions along which price jumps can be classified. The first, well-known direction relates to the time-asymmetry of the volatility of the price around the jump which emerges as dominant in a PCA sense, and results in three classes of jumps, exogenous (with activity after the jump) symmetric and “anticipatory”. We have argued that the last two cases correspond to endogenous events. Thanks to this classification we have shown that a large portion of the jumps are endogenous or anticipatory jumps, confirming – but also making much more precise – the main conclusions of (15, 16, 23, 24).

We also evidenced that mean-reversion and trend are important features for classification. This allowed us to identify three additional classes of jumps, “mean-reverting”, “trend-aligned” and “trend-anti-aligned” which concerns a significant portion of the dataset.

Extending our analysis to co-jumps, we have gathered several pieces of evidence that a large proportion of these co-jumps should also, quite surprisingly, be classified as endogenous in the sense that they seem to originate from the contagion of one single endogenous jump triggering the jump of possibly many others, or even the whole market (15, 16). One signature of such a scenario is the power-law distribution of co-jump sizes, which is indeed close to that predicted by a critical branching (contagion) process. Such a broad, power-law distribution of co-jump sizes was noted previously for different datasets in (24, 27, 29). Further work should focus on higher frequency data that would allow one to dissect more

precisely the contagion mechanism and ascertain that many large co-jumps are indeed *not* triggered by exogenous news, but related to the close-knit nature of financial markets that may bring them close to critical fragility, as argued many times in the past, see e.g. (18, 46–48) and refs. therein., and (49–51) for dissenting views.

Unlike parametric fit of the time-series, the wavelet scattering embedding is defined and can be computed for any time-series. As such, our study could be transposed to other fields as well, such as fracture surfaces (52, 53) for example.

ACKNOWLEDGMENTS.

We would like to thank Riccardo Marcaccioli for sharing his work on the detection of jumps in financial data and ideas on how to classify events based on their asymmetry. We are also very grateful to Guillaume Maitrier for helping us visualize and understand what happened in the Limit Order Book. We also thank Marcello Rambaldi and Iacopo Mastromatteo for their help with news data. Finally we would like to thank Maria Flora, Jérôme Garnier-Brun, Samy Lakhali and Yuling Yao for insightful discussions.

This research was conducted within the Econophysics & Complex Systems Research Chair, under the aegis of the Fondation du Risque, the Fondation de l’Ecole polytechnique, the Ecole polytechnique and Capital Fund Management. R.M. acknowledges funding from the French government as part of the “Investissements d’avenir” program ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

1. D Sornette, Endogenous versus exogenous origins of crises. *Extrem. events nature society* pp. 95–119 (2006).
2. I Osorio, MG Frei, D Sornette, J Milton, YC Lai, Epileptic seizures: quakes of the brain? *Phys. Rev. E* **82**, 021919 (2010).
3. D Sornette, I Osorio, Prediction. *chapter “Epilepsy: The Intersect. Neurosci. Biol. Math. Phys. Eng. Ed. Osorio I., Zaveri H.P., Frei M.G., Arthurs S., CRC Press. Taylor & Francis Group, pp. 203-237* (2010).
4. V Filimonov, D Sornette, Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Phys. Rev. E* **85**, 056108 (2012).
5. SJ Hardiman, N Bercot, JP Bouchaud, Critical reflexivity in financial markets: a hawkes process analysis. *The Eur. Phys. J. B* **86**, 1–9 (2013).
6. SJ Hardiman, JP Bouchaud, Branching-ratio approximation for the self-exciting hawkes process. *Phys. Rev. E* **90**, 062807 (2014).
7. P Bak, K Chen, J Scheinkman, M Woodford, Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics. *Ricerche economica* **47**, 3–30 (1993).
8. J Moran, JP Bouchaud, May’s instability in large economies. *Phys. Rev. E* **100**, 032307 (2019).
9. D Sornette, F Deschâtres, T Gilbert, Y Ageon, Endogenous versus exogenous shocks in complex networks: An empirical test using book sale rankings. *Phys. Rev. Lett.* **93**, 228701 (2004).
10. F Deschâtres, D Sornette, Dynamics of book sales: Endogenous versus exogenous shocks in complex networks. *Phys. Rev. E* **72**, 016112 (2005).
11. GO Mohler, MB Short, PJ Brantingham, FP Schoenberg, GE Tita, Self-exciting point process modeling of crime. *J. Am. Stat. Assoc.* **106**, 100–108 (2011).
12. L Bonnasse-Gahot, et al., Epidemiological modelling of the 2005 french riots: a spreading wave and the role of contagion. *Sci. reports* **8**, 1–20 (2018).
13. J Moran, FP Pijpers, U Weitzel, JP Bouchaud, D Panja, Temporal criticality in socio-technical systems. *arXiv preprint arXiv:2307.03546* (2023).
14. G Soros, The alchemy of finance: Reading the mind of the market by george soros (1994-05-06). (year?).
15. DM Cutler, JM Poterba, LH Summers, *What moves stock prices?* (National Bureau of Economic Research Cambridge, Massachusetts) Vol. 487. (1988).
16. RC Fair, Events that shook the market. *The J. Bus.* **75**, 713–731 (2002).
17. P Båk, *How nature works: The science of self-organized criticality* (copernicus, new york). (1996).
18. P Bak, M Paczuski, Complexity, contingency, and criticality. *Proc. Natl. Acad. Sci.* **92**, 6689–6696 (1995).
19. D Sornette, A Helmstetter, Endogenous versus exogenous shocks in systems with memory. *Phys. A: Stat. Mech. its Appl.* **318**, 577–591 (2003).
20. R Crane, D Sornette, Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci.* **105**, 15649–15653 (2008).
21. D Sornette, Y Malevergne, JF Muzy, What causes crashes? *RISK-LONDON-RISK MAGAZINE LIMITED*. **16**, 67–72 (2003).
22. D Sornette, Y Malevergne, JF Muzy, Volatility fingerprints of large shocks: endogenous versus exogenous in *The Application of Econophysics: Proceedings of the Second Nikkei Econophysics Symposium*. (Springer), pp. 91–102 (2004).
23. R Marcaccioli, JP Bouchaud, M Benzaquen, Exogenous and endogenous price jumps belong to different dynamical classes. *J. Stat. Mech. Theory Exp.* **2022**, 023403 (2022).
24. A Joulin, A Lefevre, D Grunberg, JP Bouchaud, Stock price jumps: news and volume play a minor role. *Wilmott Mag.* **46** (2008).

1117	25. Q Wu, Y Sano, H Takayasu, M Takayasu, Classification of endogenous and exogenous bursts in collective emotions based on weibo comments during covid-19. <i>Sci. Reports</i> 12 , 3120 (2022).	1179
1118		1180
1119	26. G Bormetti, et al., Modelling systemic price cojumps with hawkes factor models. <i>Quant. Finance</i> 15 , 1137–1156 (2015).	1181
1120		1182
1121	27. LM Calcagnile, G Bormetti, M Treccani, S Marmi, F Lillo, Collective synchronization and high frequency systemic instabilities in financial markets. <i>Quant. Finance</i> 18 , 237–247 (2018).	1183
1122	28. J Bruna, S Mallat, Invariant scattering convolution networks. <i>IEEE transactions on pattern analysis machine intelligence</i> 35 , 1872–1886 (2013).	1184
1123		1185
1124	29. C Aubrun, M Benzaquen, JP Bouchaud, Multivariate quadratic hawkes processes—part i: theoretical analysis. <i>Quant. Finance</i> 23 , 741–758 (2023).	1186
1125	30. K Boudt, C Croux, S Laurent, Robust estimation of intraweek periodicity in volatility and jump detection. <i>J. Empir. Finance</i> 18 , 353–367 (2011).	1187
1126		1188
1127	31. S Probert, Y Song, Detection and classification of high frequency transients using wavelet analysis in <i>IEEE Power Engineering Society Summer Meeting.</i> , (IEEE), Vol. 2, pp. 801–806 (2002).	1189
1128		1190
1129	32. DI Kim, TY Chun, SH Yoon, G Lee, YJ Shin, Wavelet-based event detection method using pmu data. <i>IEEE Transactions on Smart grid</i> 8 , 1154–1162 (2015).	1191
1130	33. D Ardila, D Sornette, Dating the financial cycle with uncertainty estimates: a wavelet proposition. <i>Finance Res. Lett.</i> 19 , 298–304 (2016).	1192
1131		1193
1132	34. L Rueda, A Cardenas, S Kelouwani, K Agbossou, Transient event classification based on wavelet neuronal network and matched filters in <i>IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society.</i> (IEEE), pp. 832–837 (2018).	1194
1133		1195
1134	35. E Cuoco, M Razzano, A Utina, Wavelet-based classification of transient signals for gravitational wave detectors in <i>2018 26th European Signal Processing Conference (EUSIPCO)</i> . (IEEE), pp. 2648–2652 (2018).	1196
1135		1197
1136	36. B Schölkopf, A Smola, KR Müller, Kernel principal component analysis in <i>International conference on artificial neural networks.</i> (Springer), pp. 583–588 (1997).	1198
1137		1199
1138	37. P Blanc, J Donier, JP Bouchaud, Quadratic hawkes processes for financial prices. <i>Quant. Finance</i> 17 , 171–188 (2017).	1200
1139		1201
1140	38. A Helmstetter, D Sornette, Foreshocks explained by cascades of triggered seismicity. <i>J. Geophys. Res. Solid Earth</i> 108 (2003).	1202
1141	39. A Helmstetter, D Sornette, JR Grasso, Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws. <i>J. Geophys. Res. Solid Earth</i> 108 (2003).	1203
1142		1204
1143	40. K Kanazawa, D Sornette, Ubiquitous power law scaling in nonlinear self-excited hawkes processes. <i>Phys. review letters</i> 127 , 188301 (2021).	1205
1144	41. M Flora, R Renò, V-shapes. Available at SSRN 4260832 (2022).	1206
1145	42. R Morel, G Rochette, R Leonarduzzi, JP Bouchaud, S Mallat, Scale dependencies and self-similarity through wavelet scattering covariance. <i>arXiv preprint arXiv:2204.10177</i> (2022).	1207
1146		1208
1147	43. R Morel, S Mallat, JP Bouchaud, Path shadowing monte-carlo. <i>arXiv preprint arXiv:2308.01486</i> (2023).	1209
1148	44. A Gerig, High-frequency trading synchronizes prices in financial markets. <i>arXiv preprint arXiv:1211.1919</i> (2012).	1210
1149		1211
1150	45. TE Harris, , et al., <i>The theory of branching processes.</i> (Springer Berlin) Vol. 6, (1963).	1212
1151	46. JP Bouchaud, The endogenous dynamics of markets: Price impact, feedback loops and instabilities. <i>Lessons from credit crisis</i> pp. 345–74 (2011).	1213
1152	47. A Fosset, JP Bouchaud, M Benzaquen, Endogenous liquidity crises. <i>J. Stat. Mech. Theory Exp.</i> 2020 , 063401 (2020).	1214
1153		1215
1154	48. JP Bouchaud, The self-organized criticality paradigm in economics & finance. <i>arXiv preprint arXiv:2407.10284</i> (2024).	1216
1155	49. V Filimonov, D Sornette, Apparent criticality and calibration issues in the hawkes self-excited point process model: application to high-frequency financial data. <i>Quant. Finance</i> 15 , 1293–1314 (2015).	1217
1156		1218
1157	50. S Wheatley, A Wehrli, D Sornette, The endo–exo problem in high frequency financial price fluctuations and rejecting criticality. <i>Quant. Finance</i> 19 , 1165–1178 (2019).	1219
1158		1220
1159	51. A Wehrli, D Sornette, The excess volatility puzzle explained by financial noise amplification from endogenous feedbacks. <i>Sci. reports</i> 12 , 18895 (2022).	1221
1160		1222
1161	52. S Vernède, L Ponson, JP Bouchaud, Turbulent fracture surfaces: A footprint of damage percolation? <i>Phys. Rev. Lett.</i> 114 , 215501 (2015).	1223
1162	53. S Lakhjal, L Ponson, M Benzaquen, JP Bouchaud, Wrapping and unwrapping multifractal fields. <i>arXiv preprint arXiv:2310.01927</i> (2023).	1224
1163		1225
1164		1226
1165		1227
1166		1228
1167		1229
1168		1230
1169		1231
1170		1232
1171		1233
1172		1234
1173		1235
1174		1236
1175		1237
1176		1238
1177		1239
1178		1240