Copulas and time series with long-ranged dependencies

Rémy Chicheportiche* and Anirban Chakraborti†
Chaire de finance quantitative, École Centrale Paris, 92 295 Châtenay-Malabry, France
(Received 20 November 2013; published 8 April 2014)

We review ideas on temporal dependencies and recurrences in discrete time series from several areas of natural and social sciences. We revisit existing studies and redefine the relevant observables in the language of copulas (joint laws of the ranks). We propose that copulas provide an appropriate mathematical framework to study nonlinear time dependencies and related concepts—like aftershocks, Omori law, recurrences, and waiting times. We also critically argue, using this global approach, that previous phenomenological attempts involving only a long-ranged autocorrelation function lacked complexity in that they were essentially monoscale.

DOI: 10.1103/PhysRevE.89.042117 PACS number(s): 02.50.−r, 05.45.Tp, 87.85.Ng

I. INTRODUCTION

A thorough understanding of the occurrences and statistics of extreme events is crucial in fields like seismicity, finance, astronomy, physiology, and so on [1,2]. The analyses of extreme events plays a pivotal role every time an addressed problem has a stochastic nature, since the rare extreme events can have rather strong or drastic consequences. One theoretical motivation for studying extreme events in a particular field like finance is to account for the observed fat tails of log-returns (deviation from the normal distribution in the tails) of stock prices [3]. A more practical motivation is that the extreme events such as "market crashes" or "shocks" pose a substantial risk for investors, even though these events are rare and do not provide enough data for reliable statistical analyses [4]. It has been observed that common financial shocks are relatively smaller in magnitude, duration, and number of stocks affected. However, the extremely large and infrequent financial crashes, such as the Black Monday crash, have significant "aftershocks" that can last for many months. This observation is very similar to the "dynamic relaxation" of the aftershock cascade following an earthquake. Hence, it is meaningful to also ask the general scientific question: How is the dynamics of a "complex" system, such as an earthquake fault [5–12] or a financial market [13–17], affected when the system undergoes an extreme event? The statistics of return intervals between extreme events is a powerful tool to characterize the temporal scaling properties of the observed time series and to estimate the risk for such hazardous events like earthquakes or financial crashes. Evaluating the return time statistics of extreme events in a stochastic process is one of the classical problems in probability theory.

Earlier, from an analysis of the probability density functions (PDF) of waiting times for earthquakes, Bak *et al*. [5] had suggested a unified scaling law combining the Gutenberg-Richter law, the Omori law, and the fractal distribution law in a single framework. This global approach was later extended by Corral [18,19], who proposed the existence of a universal scaling law for the PDF of recurrence times between earthquakes in a given region. This is useful because, due to the scaling properties, it is possible to analyze the statistics of return intervals for different thresholds by studying only the behavior of small fluctuations occurring very frequently, which have much better statistics and reliability than those of the rare extreme large fluctuations. It also reveals a spatiotemporal organization of the seismicity, as suggested by Saichev and Sornette [10].

In this paper, we review the ideas on temporal dependencies and recurrences in discrete time series common to several areas of natural sciences (earthquakes, etc.) and social sciences (financial markets). We revisit the existing studies, cited above, and redefine the relevant observables in the mathematical language of "copulas." We propose that copulas is a very general and appropriate framework to study nonlinear time dependencies and related concepts—like aftershocks, Omori law, recurrences, and waiting times. Our overall aim is to study several properties of recurrence times and the statistic of other observables (waiting times, cluster sizes, records, aftershocks) described in terms of the diagonal copula. We hope that these studies can shed light on the n-point properties of the process. We also critically argue that that previous phenomenological attempts involving only a long-ranged autocorrelation function lacked complexity in that they were essentially monoscale.

A. The copula

As a tool to study the (possibly highly nonlinear) correlations among random variables, "copulas," i.e., joint distributions of the ranks (see formal definition below), have long been used in actuarial sciences and finance to describe and model cross-dependencies of assets, often in a risk management perspective [20–22]. Although the widespread use of simple analytical copulas to model multivariate dependencies is more and more criticized [23,24], copulas remain useful as a tool to investigate empirical properties of multivariate data [24].

More recently, copulas have also been studied in the context of serial dependencies in univariate time series, where they find yet another application range: just as Pearson's ρ coefficient is commonly used to measure both linear cross-dependencies and temporal correlations, copulas are well-designed to assess non-linear dependencies both transversally or serially [25–27]; we will speak of "self-copulas" in the latter case.

B. Notations

We consider a time series $\{X_t\}_{t=1\dots T}$ of length T as a realization of a discrete stochastic process. The joint

*remy.chicheportiche@graduates.centraliens.net
†Present address: School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi - 110067, India; anirban.chakraborti@ecp.fr

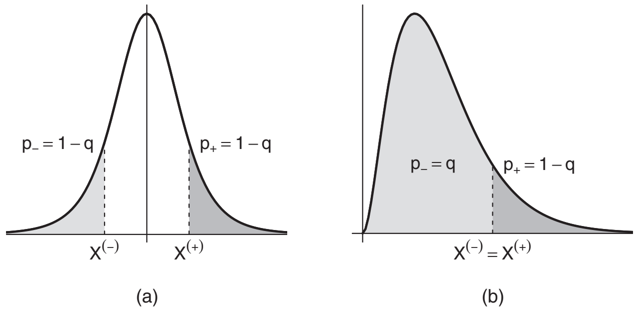


FIG. 1. Two possible definitions of events: (a) p_- and p_+ are probabilities of extremes (negative and positive, respectively), so $F(X^{(+)}) = 1 - F(X^{(-)}) = q$; (b) p_+ is a probability of extreme and $p_- = 1 - p_+$, so $F(X^{(+)}) = F(X^{(-)}) = q$.

cumulative distribution function (CDF) of n occurrences $(1 \leqslant t_1 < \cdots < t_n < T)$ of the process is

$$\mathcal{F}_{t_1,\ldots,t_n}(\mathbf{x}) = \mathbb{P}\left[X_{t_1} < x_{t_1},\ldots,X_{t_n} < x_{t_n}\right]. \tag{1}$$

We assume that the process is stationary with a distribution F, and a translational-invariant joint distribution $\mathcal{F}$ with long-ranged dependencies, as is typically the case, e.g., for seismic and financial data.

A realization of X_t at date t will be called an "event" when its value exceeds a threshold $X^{(\pm)}$, a "negative event" when $X_t < X^{(-)}$ and a "positive event" when $X_t > X^{(+)}$. The probability p_- of such a "negative event" is $F(X^{(-)})$ and, similarly, the probability that X_t is above a threshold $X^{(+)}$ is the tail probability $p_+ = 1 - F(X^{(+)})$.

If a unique threshold $X^{(+)} = X^{(-)}$ is chosen, then obviously $p_+ = 1 - p_-$. This is appropriate when the distribution is one-sided, typically for positive-only signals, and one wishes to distinguish between two regimes: extreme events (above the unique threshold) and regular events (below the threshold). This case is illustrated schematically in Fig. 1(a). When the distribution is two-sided, it is more convenient to define $X^{(+)}$ as the q-th quantile of F and $X^{(-)}$ as the $(1-q)$-th quantile for a given $q \in [\frac{1}{2},1]$, so $p_+ = p_- = 1 - q$. This allows us to investigate persistence and reversion (antipersistence) effects in *signed* extreme events, while excluding a neutral zone of regular events between $X^{(-)}$ and $X^{(+)}$; see Fig. 1(b).

When the threshold for the recurrence is defined in terms of quantiles like above (a *relative* threshold), stationarity is not needed theoretically but much wanted empirically, as already stated, otherwise the height of the threshold might change every time. In contrast, when the threshold is set as a number (an *absolute* threshold), there is no issue on the empirical side, but the theoretical discussion makes sense only under stationary marginal.

The next section recalls several two-point and many-point properties of stationary processes and discusses associated measures of dependence in light of the copula. This rather theoretical content is followed in Sec. III by applications to financial data. The definition and some properties of copulas are recalled in the appendix, and the Gaussian case with long-ranged correlations is treated.

II. DEPENDENCIES IN DISCRETE-TIME PROCESSES

We consider the case where the discrete times t_n in the definition (1) are equidistant ("regularly sampled").

A. Two-point dependence measures

Typical measures of dependencies in stationary processes are two-point expectations that only involve one parameter: the lag ℓ separating the points in time. For example, the usefulness of the linear correlation function

$$\rho(\ell) = \mathbb{E}[X_t X_{t+\ell}] - \mathbb{E}[X_t]\,\mathbb{E}[X_{t+\ell}] \tag{2}$$

is rooted in the analysis of Gaussian processes, as those are completely characterized by their covariances, and multilinear correlations are reducible to all combinations of two-point expectations, according to Isserlis' theorem [28,29] (also known later as Wick's theorem [30]). Some nonlinear dependencies, like the tail dependence, for example [21,22], are, however, not expressed in terms of simple correlations but involve the whole bivariate copula as follows:

$$C_\ell(u,v) = \mathcal{F}_{t,t+\ell}(F^{-1}(u),F^{-1}(v)), \tag{3}$$

where $(u,v) \in [0,1]^2$. C_ℓ can be understood as the distribution of the marginal ranks $U = F(X_t), V = F(X_{t+\ell})$ and contains the full information on bivariate dependence that is invariant under increasing transformations of the marginals. For example, the conditional probability

$$p_{++}^{(\ell)} = \mathbb{P}[X_{t+\ell} > X^{(+)} | X_t > X^{(+)}], \tag{4}$$

which is a measure of *persistence* of the "positive" events, can be written in terms of copulas, together with all three other cases of conditioning,

$$p_{++}^{(\ell)} = [2p_+ - 1 + C_\ell(1-p_+,1-p_+)]/p_+, \tag{5a}$$

$$p_{--}^{(\ell)} = C_\ell(p_-,p_-)/p_-, \tag{5b}$$

$$p_{-+}^{(\ell)} = [p_- - C_\ell(p_-,1-p_+)]/p_-, \tag{5c}$$

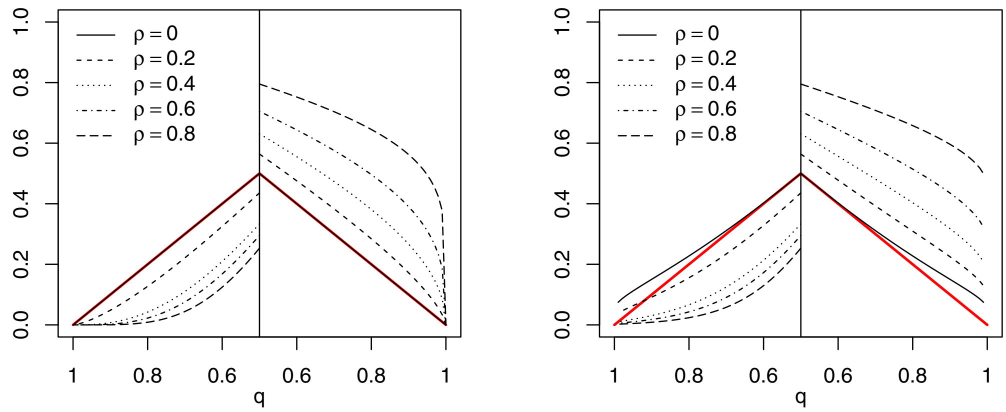
$$p_{+-}^{(\ell)} = [p_- - C_\ell(1-p_+,p_-)]/p_+, \tag{5d}$$

where $p_{\pm\pm}^{(\ell)}$ and $p_{\pm\mp}^{(\ell)}$ are defined similarly to Eq. (4), with the convention "from *first* to *second*," i.e., the first sign stands for the conditioning. When $X^{(+)} = X^{(-)} = 0$ and $\ell = 1$, this is exactly the definition of Boguná and Masoliver [31], with, accordingly, $p_- = p_+ = F(0)$; see Fig. 1. Note also that $p_{\pm\pm}^{(\ell)}$ and $p_{\pm\mp}^{(\ell)}$ are straightforwardly related to the "tail dependence coefficients" [32].

As an example, consider the Gaussian bivariate copula of the pair $(X_t,X_{t+\ell})$, whose whole ℓ dependence is in the linear correlation coefficient $\rho(\ell)$. Figure 2(a) illustrates the conditional probabilities (5) as a function of the threshold, when $p_+ = p_- = 1 - q$. A similar plot for the Student copula (with $\nu = 5$ degrees of freedom) is shown in Fig. 2(b): The fatter tails of the joint distribution are responsible for the abnormal behavior of the conditional probabilities in the region $q = 1$. When $q = 0.5$, the coefficients (5) are all equal to

$$\frac{1}{2} + \frac{1}{\pi}\arcsin\rho(\ell)$$

for any elliptical copula [32].



(a) Gaussian copula (b) Student copula ($\nu = 5$)

FIG. 2. (Color online) Conditional probabilities $p_{\pm\mp}^{(\ell)}$ (left panels) and $p_{\pm\pm}^{(\ell)}$ (right panels) for different values of $\rho(\ell)$ with thresholds at $p_+ = p_- = 1 - q$. In the case of the independence copula (shown in bold red), the conditional probabilities are all equal: $p_{\pm\mp} = p_{\pm\pm} = 1 - q$.

Aftershocks

Omori's law characterizes the ℓ dependence of $p_{++}^{(\ell)}$, i.e., the average frequency of events occurring ℓ time steps after a main event. This was first stated in the context of earthquake occurrences [33], where this time dependence is a power law as follows:

$$p_{++}^{(\ell)} = \lambda \ell^{-\alpha}. \tag{6}$$

Notice that any dependence on the threshold must be hidden in λ according to this description. The average *cumulated* number N_ℓ of these aftershocks until ℓ is thus

$$\langle N_\ell \rangle_+ = \lambda \frac{\ell^{1-\alpha}}{1-\alpha}, \tag{7}$$

with in fact $\lambda \equiv p_+$ since, when $\alpha \to 0$, N_ℓ has no time dependence, i.e., it counts independent events, and $p_{++}^{(\ell)}$ must thus tend to the unconditional probability.

In order to give a phenomenological grounding to this empirical law also later observed in finance [15,34], Lillo and Mantegna [35] model the aftershock volatilities in financial markets as a decaying scale $\sigma(\ell)$ times an independent stochastic amplitude r_ℓ with CDF ϕ. As a consequence, $p_{++}^{(\ell)} \sim 1 - \phi(X^{(+)}/\sigma(\ell))$ and the power-law behavior of Omori's law results from (i) the power-law marginal $\phi(r) \sim r^{-\gamma}$ and (ii) the scale decaying as a power law $\sigma(t) \sim t^{-\beta}$, so relation (6) is recovered with $\alpha = \beta\gamma$. The nonstationarity described by σ is only introduced in a *conditional* sense and might be appropriate for aging systems or financial markets, but we believe that Omori's law can be accounted for in a stationary setting and without necessarily having power-law-distributed amplitudes.

The scaling of $p_{++}^{(\ell)}$ with the magnitude of the main shock is encoded in the prefactor $\lambda \equiv p_+$, which, for example, accounts for the exponentially distributed magnitudes of earthquakes (Gutenberg-Richter law [36]). The linear dependence of $p_{++}^{(\ell)}$ on p_+ shall be reflected in the diagonal of the underlying copula,

$$C_\ell(p,p) = p^2\,\ell^{-\alpha}, \tag{8}$$

a prediction that can be tested empirically.

Note that Omori's law is a measure involving only the two-point probability. In the next subsection, we show which additional information many-point probability can reflect.

B. Multipoint dependence measures

Although the n-point expectations of Gaussian processes reduce to all combinations of two-point expectations (2), their full dependence structure is not reducible to the bivariate distribution in the general case. Furthermore, when the process is not Gaussian, even the multilinear correlations are irreducible. In the general case, the whole multivariate CDF is needed, but many measures of dependence that we introduce below only involve the diagonal n-point copula[1] as follows:

$$\mathcal{C}_n(p) = \mathcal{F}_{t+[[1,n]]}(F^{-1}(p),\ldots,F^{-1}(p)), \tag{9}$$

which measures the joint probability that all $n \geqslant 1$ consecutive variables $X_{t+1},\ldots,X_{t+n}$ are below the upper p-th quantile of the stationary distribution, $p \in [0,1]$, and $t + [[1,n]]$ is a shorthand for $\{t+1,\ldots,t+n\}$. Clearly, $\mathcal{C}_1(p) = p$ and we set by convention $\mathcal{C}_0(p) \equiv 1$.

Empirically, the n-point probabilities are very hard to measure due to the large noise associated with such rare joint occurrences. However, there exist observables that embed many-point properties and are more easily measured, such as the length of sequences (clusters) of thresholded events and the recurrence times of such events, which we study next.

1. Recurrence intervals

The probability $\pi(\tau)$ of observing a recurrence interval τ between two events is the conditional probability of observing a sequence of $\tau - 1$ "nonevents" bordered by two events as follows:

$$\pi(\tau) = \mathbb{P}[\mathcal{X}_{0;\tau}^{(+)} | X_0 > X^{(+)}] \tag{10}$$

where

$$\mathcal{X}_{t;\tau}^{(+)} \equiv \{X_{t+[[1,\tau-1]]} < X^{(+)}, X_{t+\tau} > X^{(+)}\} \tag{11}$$

[1]We use a calligraphic $\mathcal{C}$ in order to make it clearly distinct from the bivariate copula discussed in the previous section.

designates a sequence of "nonevents" starting in t and terminated by a "positive event" at $t+\tau$. (We focus on positive events, but the recurrence of negative events can be studied with the substitution $X \to -X$, and the case of recurrence in amplitudes with the substitution $X \to |X|$). After a simple algebraic transformation flipping all ">" signs to "<," it can be written in the language of copulas as

$$\pi(\tau) = \frac{\mathcal{C}_{\tau-1}(1-p_+) - 2\,\mathcal{C}_\tau(1-p_+) + \mathcal{C}_{\tau+1}(1-p_+)}{p_+}. \quad (12)$$

The cumulative distribution

$$\Pi(\tau) = \sum_{n=1}^{\tau} \pi(n) = 1 - \frac{\mathcal{C}_\tau(1-p_+) - \mathcal{C}_{\tau+1}(1-p_+)}{p_+} \quad (13)$$

is more appropriate for empirical purposes, being less sensitive to noise. These exact expressions make clear—almost straight from the definition—that (i) the distribution of recurrence times *depends only on the copula* of the underlying process and not on the stationary law, in particular, its domain or its tails (this is because we take a relative definition of the threshold as a quantile); (ii) *nonlinear* dependencies[2] are highly relevant in the statistics of recurrences, so linear correlations can, in the general case, by no means explain alone the properties of $\pi(\tau)$; and (iii) recurrence intervals have a *long memory* revealed by the $(\tau+1)$-point copula being involved, so only when the underlying process X_t is Markovian will the recurrences themselves be memoryless.[3] Hence, when the copula is known [Eq. (A1) in the appendix for Gaussian processes], the distribution of recurrence times is characterized by the exact expression in Eq. (12).

The average recurrence time is found straightforwardly by summing the series

$$\mu_\pi = \langle\tau\rangle = \sum_{\tau=1}^{\infty} \tau\,\pi(\tau) = \frac{1}{p_+}, \quad (14)$$

and is *universal* whatever the dependence structure. This result was first stated and proven by Kac in a similar fashion [38]. It is intuitive as, for a given threshold, the whole time series is the succession of a fixed number p_+T of recurrences whose lengths τ_i necessarily add up to the total size T, so $\langle\tau\rangle = \sum_i \tau_i/(p_+T) = 1/p_+$. Note that Eq. (14) assumes an infinite range for the possible lags τ, which is achieved either by having an infinitely long time series or, more practically, when the translationally invariant copula is periodic at the boundaries of the time series, as is typically the case for artificial data which are simulated using numerical Fourier transform methods. Introducing the copula allows us to emphasize the validity of the statement even in the presence of nonlinear long-term dependencies, as Eq. (14) means that the average recurrence interval is *copula* independent.

More generally, the m-th moment can be computed as well by summing $\tau^m\pi(\tau)$ over τ as follows:

$$\langle\tau^m\rangle = \frac{1 + \sum_{\tau=1}^{\infty}[|\tau+1|^m - 2\tau^m + |\tau-1|^m]\,\mathcal{C}_\tau(1-p_+)}{p_+}. \quad (15)$$

In particular, the variance of the distribution is

$$\sigma_\pi^2 \equiv \langle\tau^2\rangle - \mu_\pi^2 = \frac{2}{p_+}\sum_{\tau=1}^{\infty}\mathcal{C}_\tau(1-p_+) - \frac{1-p_+}{p_+^2}. \quad (16)$$

It is not universal, in contrast with the mean, and can be related to the average unconditional waiting time (see below). Notice that in the independent case, and because of discreteness effects, the variance $\sigma_\pi^2 = (1-p_+)/p_+^2$ is not exactly equal to the squared mean $\mu_\pi^2 = 1/p_+^2$, as would be the case for the exponential distribution of the recurrence intervals of a *continuous*-time Poisson process.

It is important to notice that the main ingredient in the distribution of recurrence times (13) is the copula (i.e., the serial dependence structure) rather than the stationary distribution F, a finding already noted by Olla [39], but which the current description highlights. The sensitivity to the extreme statistics of the process is in fact hidden in p_+, but what matters more is the (possibly multiscale) dependence structure $\mathcal{C}_\tau$.

2. Conditional recurrence intervals and clustering

The dynamics of recurrence times is as important as their statistical properties, and in fact impacts the empirical determination of the latter.[4] It is now clear, both from empirical evidences and analytically from the discussion on Eq. (12), that recurrence intervals have a long memory. In dynamic terms, this means that their occurrences show some clustering. The natural question is then as follows: "Conditionally on an observed recurrence time, what is the probability distribution of the next one?" This probability of observing an interval τ' immediately following an observed recurrence time τ is

$$\mathbb{P}[\mathcal{X}^{(+)}_{\tau;\tau'}|\mathcal{X}^{(+)}_{0;\tau}, X_0 > X^{(+)}]. \quad (17)$$

Again, flipping the ">" to "<" allows us to decompose it as

$$\frac{\mathcal{C}_{\tau-1;\tau'-1} - \mathcal{C}_{\tau;\tau'-1} - \mathcal{C}_{\tau-1;\tau'} + \mathcal{C}_{\tau;\tau'}}{\mathcal{C}_{\tau-1} - 2\mathcal{C}_\tau + \mathcal{C}_{\tau+1}} - \frac{\pi(\tau+\tau')}{\pi(\tau)}, \quad (18)$$

where the $(\tau+\tau')$-point probability

$$\mathcal{C}_{\tau;\tau'}(p) = \mathcal{F}_{[[0,\tau+\tau']]\setminus\{\tau\}}(F^{-1}(p),\ldots,F^{-1}(p))$$

shows up, where $[[0,\tau+\tau']]\setminus\{\tau\}$ is the sequence $[[0,\tau+\tau']]$, where τ has been dropped. Of course, this exact expression has no practical use, again because there is no hope of empirically measuring any many-point probabilities of extreme events with a meaningful signal-to-noise ratio. We rather want to stress that nonlinear correlations and multipoint dependencies are

[2]"Nonlinear dependencies" designates any type of serial dependence that is not grasped by the linear autocorrelation like, for example, the correlation of absolute values, the conditional persistence or reversion probabilities Eq. (5), the tail dependencies, and so on.

[3]Renewal processes are also able to produce independent consecutive recurrences [16,37].

[4]Distribution testing for $\pi(\tau)$ involving goodness-of-fit tests [17] should be discarded because those are not designed for dependent samples and rejection of the null cannot be relied upon. See Chicheportiche and Bouchaud [42] for an extension of goodness-of-fit tests when some dependence is present.

relevant and that a characterization of clustering based on the autocorrelation of recurrence intervals is an oversimplified view of reality.

3. Waiting times

The conditional mean residual time to next event, when sitting τ time steps after a (positive) event, is

$$\langle w|\tau\rangle = \sum_{w=1}^{\infty} w\,\pi(\tau+w) = \frac{1}{p_+}\mathcal{C}_\tau(1-p_+). \tag{19}$$

One is often more concerned with unconditional waiting times, which is equivalent to asking what the size w of a sequence of "nonevents" starting now will be, *regardless of what happened previously*. The distribution $\phi(w) = \mathbb{P}[\mathcal{X}^{(+)}_{0;w+1}]$ of these waiting times is equal to

$$\phi(w) = \mathcal{C}_w(1-p_+) - \mathcal{C}_{w+1}(1-p_+), \tag{20}$$

and its expected value is

$$\mu_\phi = \langle w\rangle = \sum_{w=1}^{\infty} \mathcal{C}_w(1-p_+), \tag{21}$$

consistently to what would be obtained by averaging $\langle w|\tau\rangle$ over all possible elapsed times in Eq. (19). From Eq. (16), we have the following relationship between the variance of the distribution π of recurrence intervals and the mean waiting time:

$$\sigma_\pi^2 = \mu_\pi\,[2\mu_\phi + 1] - \mu_\pi^2. \tag{22}$$

4. Sequences lengths

The serial dependence in the process is also revealed by the distribution of sequences sizes. The probability that a sequence of consecutive negative events,[5] starting just after a "nonevent," will have a size n is

$$\psi(n) = \frac{\mathcal{C}_n(p_-) - 2\,\mathcal{C}_{n+1}(p_-) + \mathcal{C}_{n+2}(p_-)}{p_-\,(1-p_-)} \tag{23}$$

and the average length of a random sequence

$$\mu_\psi = \langle n\rangle = \sum_{n=1}^{\infty} n\,\psi(n) = \frac{1}{1-p_-} \tag{24}$$

is universal, just like the mean recurrence time. This property rules out the analysis of Boguná and Masoliver [31], who claim to be able to distinguish the dependence in processes according to the average sequence size.

5. Record statistic

We conclude this theoretical section on multipoint nonlinear dependencies by mentioning that the diagonal n-point copula $\mathcal{C}_n$ can be alternatively understood as the distribution of the maximum of n realizations of X in a row, since

$$\mathbb{P}\big[\max_{\tau\leqslant n}\{X_\tau\} < F^{-1}(p)\big] = \mathbb{P}\big[X_{[[1,n]]} < F^{-1}(p)\big]$$

is equal to $\mathcal{C}_n(p)$. Thus, studying the statistics of such "local" maxima in short sequences (see, e.g., Ref. [40]) can provide information on the multipoint properties of the underlying process. The CDF of the running maximum, or *record*, is $\mathcal{C}_t(F(x))$ and the probability that $t > 1$ will be a record-breaking time is the joint probability

$$R(t) = \mathbb{P}\big[\max_{\tau<t}\{X_\tau\} < X_t\big],$$

which is *irrespective of the marginal law*.

A summary of all the quantities introduced above is provided in Table I for the case of a white noise (WN), whose joint distribution Eq. (1) factorizes into a product over the t_i's.

[5]We consider the case of "negative" events, i.e., those with $X_t < X^{(-)}$ because it expresses simply in terms of diagonal copulas. The mirror case with "positive" events has the exact same expression but $\mathcal{C}_n$ must be inverted around the median. For a symmetric F, this distinction is irrelevant.

TABLE I. Different probabilities introduced with thresholds defined as $F(X^{(+)}) = q = 1 - F(X^{(-)})$ for the white noise process.

$\pi_+(\tau)$	$\langle\tau\rangle_+$	$\phi_+(w)$	$\langle w\rangle_+$	$\psi_-(n)$	$\langle n\rangle_-$	R(t)
$(1-q)\,q^{\tau-1}$	$\frac{1}{1-q}$	$(1-q)\,q^w$	$\frac{q}{1-q}$	$q\,(1-q)^{n-1}$	$\frac{1}{q}$	$\frac{1}{t}$

III. FINANCIAL SELF-COPULAS

We illustrate some of the quantities introduced above on series of daily index returns. The properties of the time series used are summarized in Table II.

A. Conditional probabilities and two-point dependencies

We reproduce the study of Boguná and Masoliver [31] on the statistic of price changes conditionally on previous return sign and extend the analysis to any threshold $|X^{(\pm)}| \geqslant 0$ and to remote lags. In addition to the time series of the five stock indices presented in Table II, we look at electroencephalogram (EEG) data from Ref. [41]. We first illustrate on Fig. 3 the conditional probabilities $p^{(\ell)}_{\pm\pm}$ (filled symbols) and $p^{(\ell)}_{\pm\mp}$ (empty symbols) with varying threshold $q = F(X^{(+)}) = 1 - F(X^{(-)})$, for $\ell = 1$. To study the departure from the independent case, it is more convenient to subtract the WN contribution to get the corresponding *excess* probabilities.

TABLE II. Description of the data set used: Time series of logarithmic returns $X_t = \ln(P_t/P_{t-1})$ of daily closing prices of international stock indices.

Stock Index	Country	From	To	T
S&P-500	USA	Jan. 02, 1970	Dec. 23, 2011	10 615
KOSPI-200	S. Korea	Jan. 03, 1990	Dec. 26, 2011	5 843
CAC-40	France	Jul. 09, 1987	Dec. 23, 2011	6 182
DAX-30	Germany	Jan. 02, 1970	Dec. 23, 2011	10 564
SMI-20	Switzerland	Jan. 07, 1988	Dec. 23, 2011	5 902

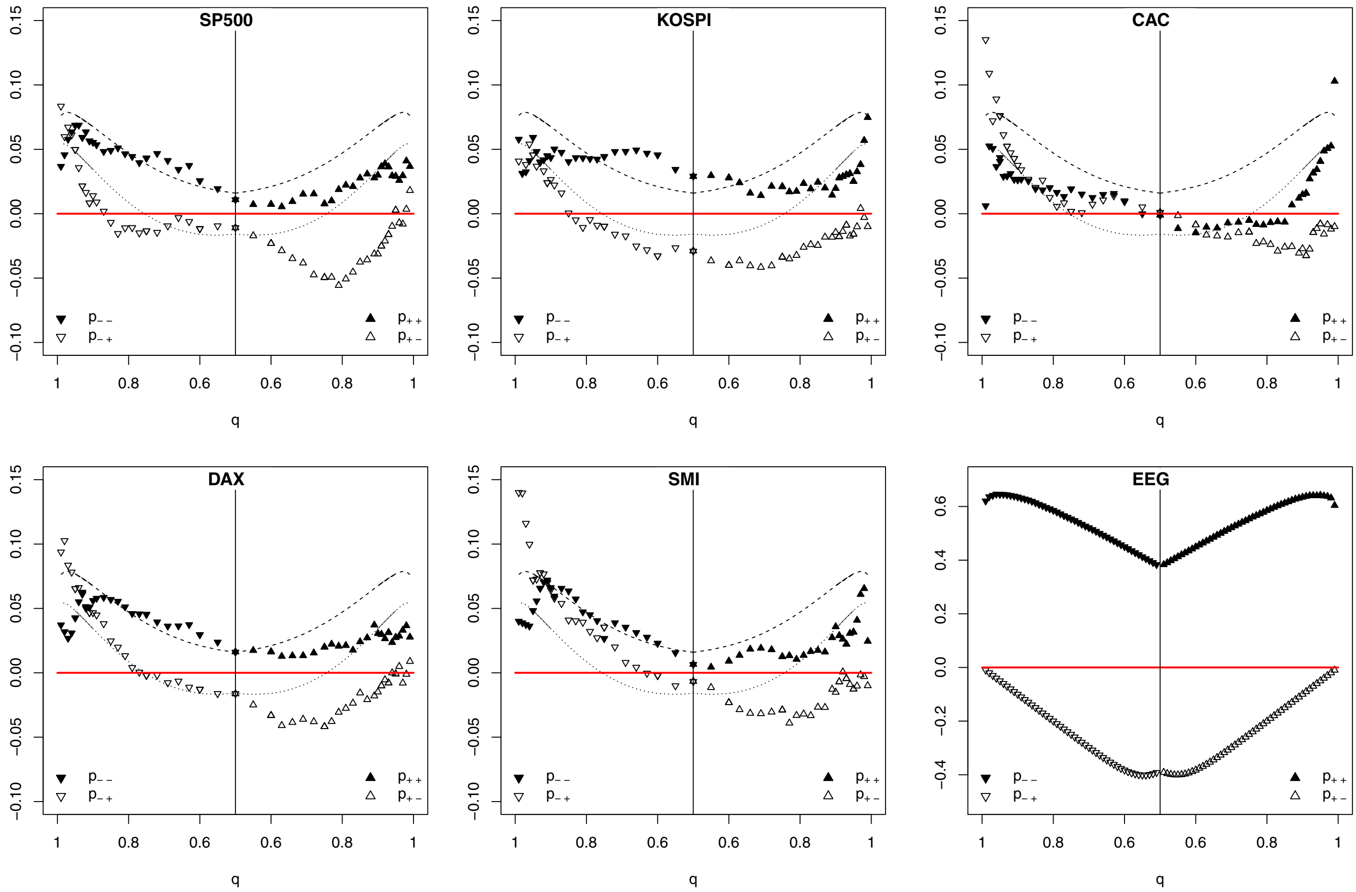


FIG. 3. (Color online) Conditional extreme probabilities at $\ell = 1$ (the white noise contribution $p_{\pm} = 1 - q$ has been subtracted, hence, the possibly negative values; compare with Fig. 2). Filled symbols are for persistence, and empty symbols for reversion. Upward triangles are conditioned on positive jumps, and downward triangles are conditioned on negative jumps.

First, the EEG data [Fig. 3(EEG)] exhibit a very strong and symmetric persistence; reversion on the other side is shut down for extreme events (like for WN) and is more suppressed than WN for intermediate values. As of the plots relative to financial indices, several features can be immediately observed: Positive events (upward triangles) trigger more subsequent positive (filled) than subsequent negative (empty) events; negative events (downward triangles) trigger more subsequent negative (filled) than subsequent positive (empty) events, except in the far tails, $q \gtrsim 0.9$, where reversion is stronger than persistence after a negative event. Both of these effects dominate the WN benchmark, but the latter effect is, however, much stronger. This overall behavior is similar for the time series of returns of all the stock indices studied. The shapes of $p_{\pm\pm}$ and $p_{\pm\mp}$ versus q are not compatible with the Student copula benchmarks (correlation $\rho = 0.05$ and degrees of freedom $\nu = 5$) shown in dashed and dotted lines, respectively. Notice that, due to its nontrivial tail correlations, see Ref. [24], the Student copula does generate increased persistence with respect to WN, lower reversion in the core, and higher reversion in the tails. But, empirically, the reversion is asymmetric and typically stronger when conditioning on negative events rather than on positive events, a property reminiscent of the leverage effect, which cannot be accounted for by a pure (symmetric) Student copula. Some of the indices exhibit more pronounced reversion and persistence effects. Interestingly, the CAC-40 returns have a regime $0.5 \leqslant q \lesssim 0.9$ close to a WN (with, in particular, a value of $p^{(1)}_{\pm\pm} = p^{(1)}_{\pm\mp}$ very close to 0 at $q = 0.5$, revealing an inefficient conditioning, i.e., as many positive and negative returns immediately following positive or negative returns), but the extreme positive events $q \gtrsim 0.9$ show a very strong persistence, and the extreme negative events a very strong reversion.

Chicheportiche and Bouchaud [42] study in detail the p and ℓ dependence of $[C_\ell(p,p) - p^2]$ and $[C_\ell(p,1-p) - p\,(1-p)]$—which are straightforwardly related to $p^{(\ell)}_{\pm\pm}$ and $p^{(\ell)}_{\pm\mp}$, respectively—and find that the self-copula of stock returns can be modeled with a high accuracy as the product $X_t = \upsilon_t \epsilon_t$ of a residual ϵ_t and a log-normally distributed scale (the "volatility") υ_t with log-decaying correlation, in agreement with multifractal volatility models. We give an overview of the results in Fig. 4, for the aggregated copula of all stocks in the S&P500 in 2000–2004. It is possible to show precisely how every kind of dependence present in the underlying process (discussed in Ref. [43]) reflects itself in $p^{(\ell)}_{++}$ for different q's: Short-ranged linear anticorrelation accounts for the central part ($p \approx 0.5$) departing from the WN prediction, long-ranged amplitude clustering is responsible for the "M" and "W" shapes that reveal excess persistence and suppressed reversion, and the leverage effect can be observed in the asymmetric heights of the "M" and "W".

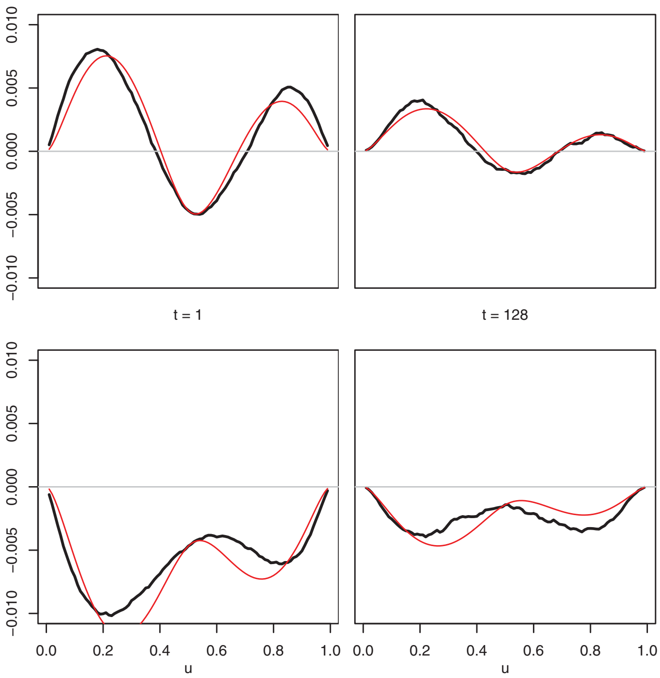


FIG. 4. (Color online) The self-copula for close-by ($\ell = 1$ day) and remote ($\ell = 128$ days) lags, with the product copula subtracted. Top: Diagonal $C_\ell(u,u) - u^2$. Bottom: $C_\ell(u,1-u) - u(1-u)$. The value determined empirically on stock returns is in bold black, and a fit with the model of Ref. [42] is shown in thin red. Adapted from Ref. [42].

B. Recurrence intervals and many-point dependencies

Even the elementary, two-point measures of self-dependence studied until now show that nonlinearities and multiscaling—in the simple sense of a memory ranging over (possibly infinitely) many scales in the past—are two ingredients that must be taken into account when attempting to describe financial time series; we now examine their many-point properties. As an example, we compute the distribution of recurrence times of returns above a threshold, $X^{(+)} = F^{-1}(1-p_+)$.

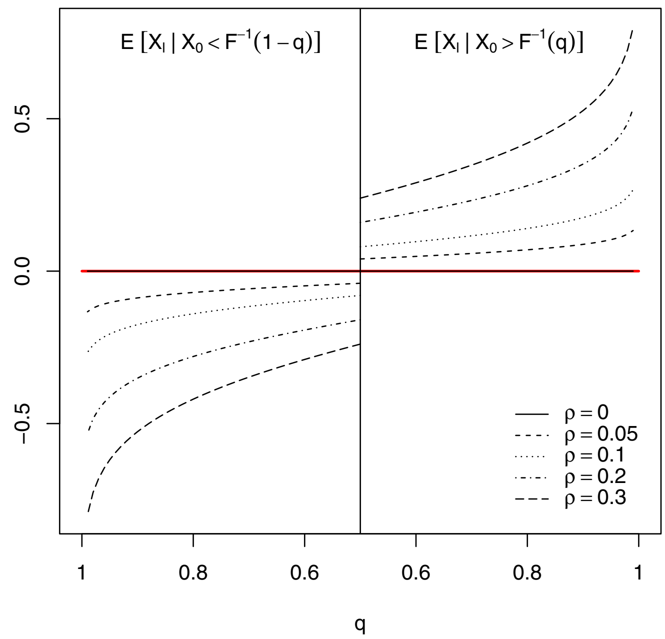


FIG. 6. (Color online) Conditional expected shortfall of a Gaussian pair (X_0, X_ℓ) for different values of $\rho(\ell)$. Left: Lower tail; Right: Upper tail. The value at $q = 0.5$ is $\sqrt{2/\pi}\,\rho(\ell)$.

Figure 5 shows the tail cumulative distribution $1 - \Pi(\tau)$ of the recurrence intervals of DAX returns, at several thresholds $p_+ = 1/\langle\tau\rangle$—the distribution for other indices is very similar. In the log-log representation used, an exponential distribution (corresponding to independent returns) would be concave and rapidly decreasing, while a power law would decay linearly. The empirical distributions fit neither of those, and Ludescher *et al.* [44] suggested a parametric fit of the form

$$1 - \Pi(\tau) = [1 + b\,(a-1)\,\tau]^{(a-2)/(a-1)}. \tag{25}$$

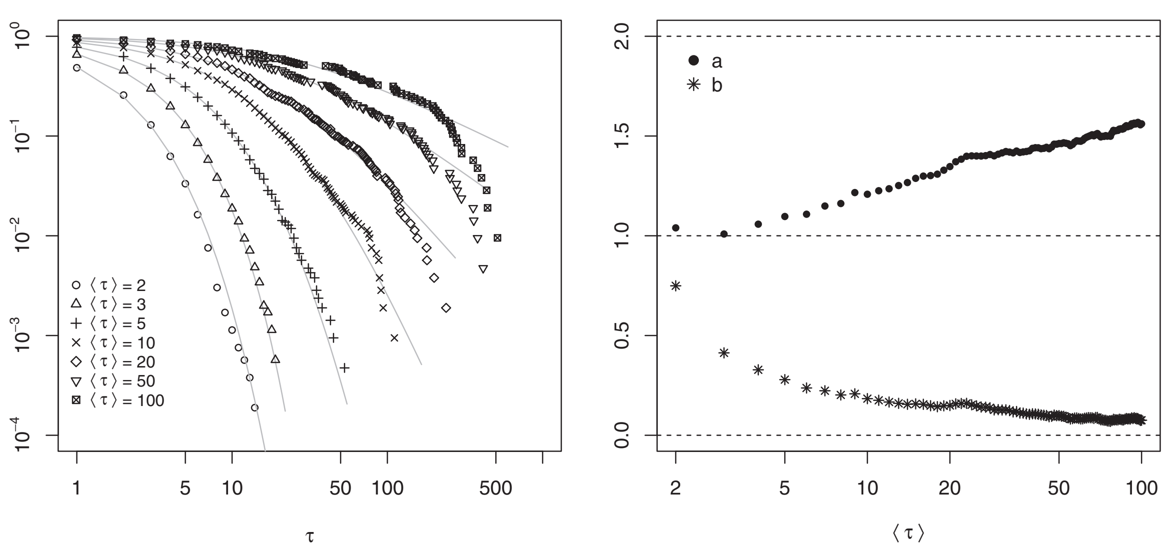


FIG. 5. DAX index returns. Left: Tail probability $1 - \Pi(\tau)$ of the recurrence intervals, at several thresholds $p_+ = 1/\langle\tau\rangle$, in log-log scale. Gray curves are best fits to Eq. (25) suggested in Ref. [44]. Right: Estimated parameters a and b of the best fit.

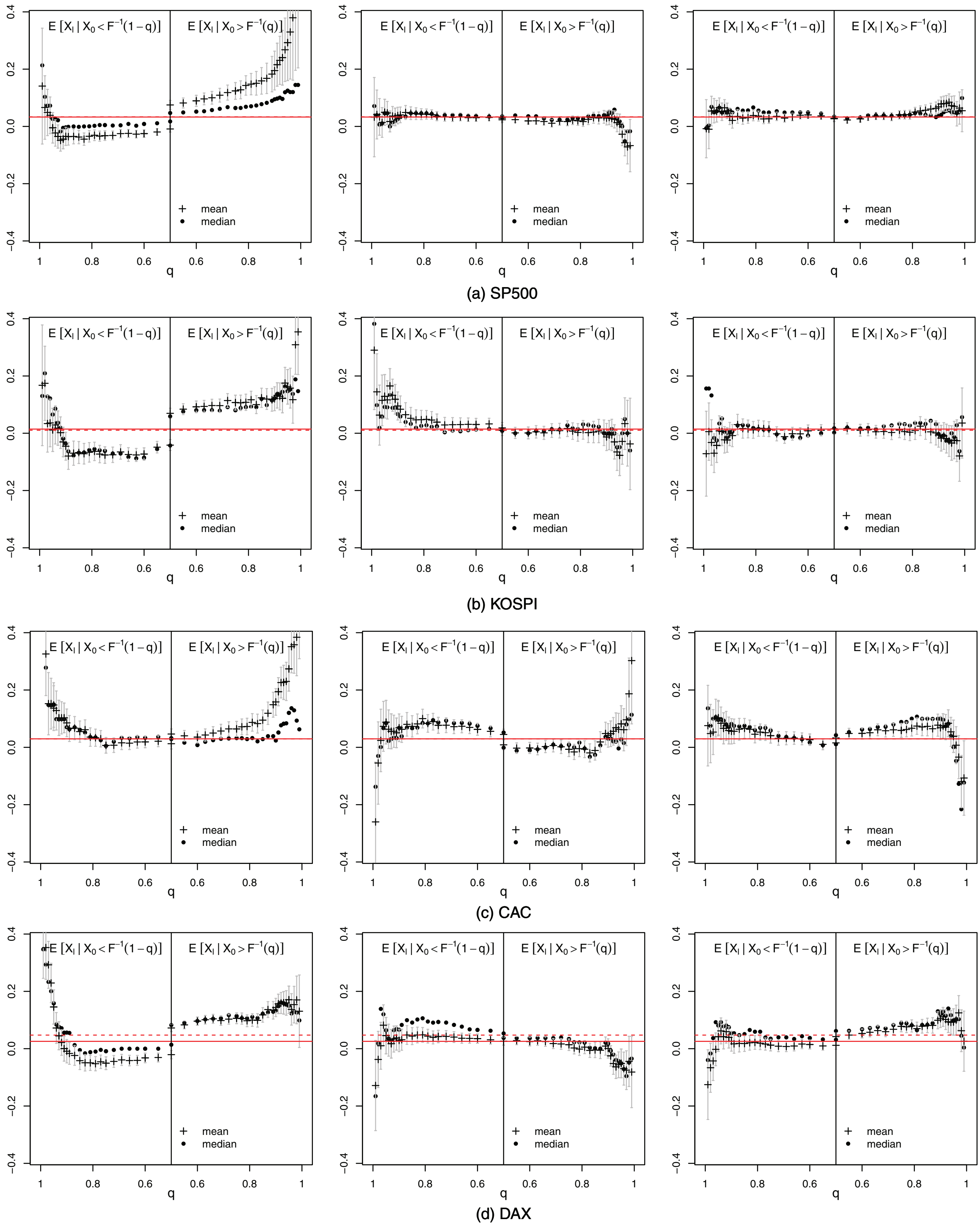


FIG. 7. (Color online) Conditional extreme amplitudes, at lags $\ell = 1,5,20$, from left to right. The upper-right and lower-left quadrants express persistence, while the upper-left and lower-right quadrants reveal reversion. For a scale-free dependence structure, one would expect the magnitudes to decrease with the lag ℓ but the global shape to be conserved. What we instead observe is important changes of configuration at different lags: For example, the strong reversion of negative tail events at $\ell = 1$ vanishes at farther lags and even turns into strong persistence for the CAC and DAX indices. That is to say, these indices tend to mean-revert after a negative event at the daily frequency but to trend on the weekly scale. Similarly, the strong persistence of positive events at $\ell = 1$ converts to a strong reversion in the tails at $\ell = 20$ for the European indices (CAC, DAX); a weaker reversion is observed at intermediate scale ($\ell = 5$) for most indices (including US and Korean).

However, important deviations are present in the tail regions for thresholds at $X^{(+)} \gtrsim F^{-1}(0.9)$, i.e., $\langle\tau\rangle \gtrsim 1/(1-0.9) = 10$: as a consequence, there is no hope that the curves for different threshold collapse onto a single curve after a proper rescaling [45], as is the case e.g., for seismic data. A more fundamental determination of the form of $\Pi(\tau)$ should rely on Eq. (13) and a characterization of the τ-point copula.

Similarly to the statistic of the recurrence intervals, their dynamics must be studied carefully. We have shown that the conditional distribution of recurrence intervals after a previous recurrence is very complex and involves long-ranged nonlinear dependencies, so a simple assessment of recurrence times autocorrelation may not be informative enough for a deep understanding of the mechanisms at stake.

IV. DISCUSSION

A. Conditional expected shortfall

In addition to caring for frequencies of conditional events, one can try to characterize their magnitudes. This of course no longer fits in the framework of copulas (that "count" joint events) but can instead be quantified by a multivariate generalization of the expected shortfall [(ES) or tail conditional expectation]. For a single random variable with cdf F, the expected shortfall is the average loss when conditioning on large events,

$$\mathrm{ES}(p_-) = \mathbb{E}[X_t | X_t < X^{(-)}] = \frac{1}{p_-}\int_{-\infty}^{F^{-1}(p_-)} x\, dF(x)$$
$$= \frac{1}{p_-}\int_0^{p_-} F^{-1}(p)\, dp.$$

In the same spirit, for bivariate distributions, the mean return conditionally on preceding return "sign" is defined as follows:

$$\langle X\rangle_-^{(\ell)} = \mathbb{E}[X_t | X_{t-\ell} < X^{(-)}], \tag{26a}$$
$$\langle X\rangle_+^{(\ell)} = \mathbb{E}[X_t | X_{t-\ell} > X^{(+)}]. \tag{26b}$$

As an example, consider the Gaussian bivariate pair $(X_t, X_{t+\ell})$, whose whole ℓ dependence is in the linear correlation coefficient $\rho(\ell)$. Figure 6 shows the conditional expected shortfall that can be computed exactly from Eqs. (26) and is proportional to the inverse Mills ratio [46] as follows:

$$\langle X\rangle_\pm = \pm\rho(\ell)\frac{\Phi'(X^{(\pm)})}{p_\pm},$$

where Φ denotes the CDF of the univariate standard normal distribution.

This Gaussian prediction is to be compared with an empirical assessment of the same quantity for series of returns of stock indices. Figure 7 displays the behavior of $\langle X\rangle_\pm$ versus q (we also show the median $\mathrm{med}(X)_\pm$) at lags corresponding to 1 day ($\ell = 1$), 1 week ($\ell = 5$), and 1 month ($\ell = 20$). The conditional amplitudes $\langle X\rangle_\pm$ measure "how large" a realization will be on average after an event at a given threshold, whereas the conditional probabilities $p_{\pm\pm}$ and $p_{\pm\mp}$ quantify "how often" repeated such events occur. Mind the *unconditional* mean and median values, both above zero and distinct from each other. At $\ell = 1$, the reversion of *extreme* events is revealed again by the change of monotonicity from $q \approx 0.8$ onward and more strongly for $q > 0.9$ where $\langle X\rangle_-$ has an opposite sign than the preceding return; this corroborates the observation made on conditional probabilities above. Beyond the next day, the general picture is that dependencies tend to vanish and the empirical measurements get more concentrated around the WN prediction. However, tail effects are strongly present, with unexpectedly a typical behavior opposite to that of $\ell = 1$: weekly and monthly reversion of extreme positive jumps. See the caption for a detailed discussion of the specificities of each stock index at every lag ℓ.

B. Conclusion

We report several properties of recurrence times and the statistic of other observables (waiting times, cluster sizes, records, and aftershocks) in light of their description in terms of the diagonal copula and hope that these studies can shed light on the n-point properties of the process by assessing the statistics of simple variables rather than positing an *a priori* dependence structure.

The exact universality of the mean recurrence interval imposes a natural scale in the system. A scaling relation in the distribution of such recurrences is only possible in the absence of any other characteristic time. When such additional characteristic times are present (typically in the nonlinear correlations), no such scaling is expected, in contrast with time series of earthquake magnitudes.

We also stress that recurrences are intrinsically multipoint objects related to the nonlinear dependencies in the underlying time series. As such, their autocorrelation is not a reliable measure of their dynamics, for their conditional occurrence probability is largely history dependent.

Ultimately, recurrences may be used to characterize risk in a new fashion. Instead of (or in addition to) caring for the amplitude and probability of adverse events at a given horizon, one should be able to characterize the risk in a dynamical point of view. In this sense, an asset A_1 could be said to be "more risky" than another asset A_2 if its distribution of recurrence of adverse events has "bad" properties that A_1 does not share. This amounts to characterizing the disutility by "when?" shocks are expected to happen, in addition to the usual "how often?" and "how large?".

It would be interesting to study many-point dependencies in continuous-time processes, where the role of the n-point copula is played by a counting process. The events to be counted can either be triggered by an underlying continuous process crossing a threshold or more directly be modeled as a self-exciting point process, like a Hawkes process. A typical financial application could be found in transaction times in a limit order book.

ACKNOWLEDGMENTS

We thank F. Abergel, D. Challet, and G. Tilak, for helpful discussions and comments. R.C. acknowledges financial support by Capital Fund Management, Paris.

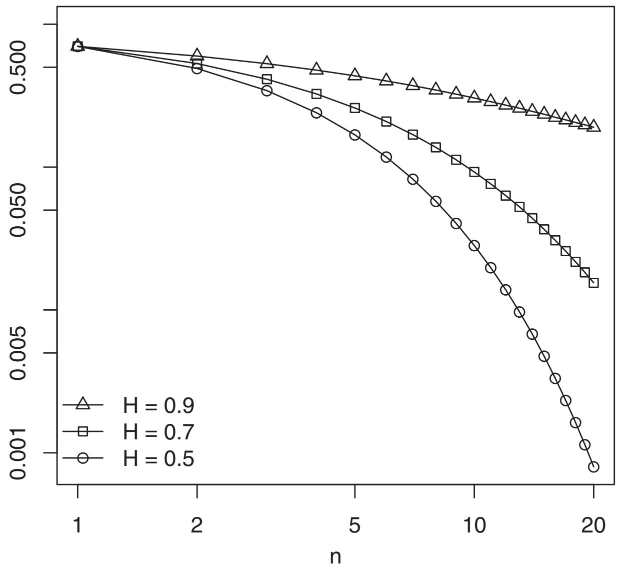


FIG. 8. Copula $\mathcal{C}_n(p)$ of the fractional Gaussian noise (FGN) with different Hurst exponents and at threshold $p_+ = 0.7$ (log-log scale).

APPENDIX A: SIMPLE COPULAS AND SKLAR'S THEOREM

Sklar's (rather trivial) theorem [47] states that any multivariate distribution $\mathcal{F}_{[[1,n]]}(x_1,\ldots,x_n)$ can be written in terms of univariate marginal distribution functions $F_i(x_i)$ $(i = 1,\ldots,n)$ and a "copula" function $C(u_1,\ldots,u_n)$ on $[0,1]^n$ which, by definition, characterizes the dependence structure between the variables. In practice, constructing the copula is achieved letting $u_i = F_i(x_i)$ for every variable i. This is expressed mathematically by Eq. (3) for bivariate distributions and can be generalized straightforwardly [see Eq. (9) for the *diagonal* of the n-point copula].

As an example, the Gaussian diagonal copula is

$$\mathcal{C}_n(p) = \Phi_\rho(\Phi^{-1}(p),\ldots,\Phi^{-1}(p)), \tag{A1}$$

where Φ^{-1} is the univariate inverse CDF and Φ_ρ denotes the multivariate CDF with $(n \times n)$ covariance matrix ρ, which is Tœplitz with symmetric entries

$$\rho_{tt'} = \rho(|t - t'|), \quad t,t' = 1,\ldots,n. \tag{A2}$$

The WN product copula $\mathcal{C}_n(p) = p^n$ is recovered in the limit of vanishing correlations $\rho(\ell) = \delta_{\ell 0}$, and other examples include the exponentially correlated Markovian Gaussian noise, the logarithmically correlated multifractal Gaussian noise, and the power-law correlated (thus scale-free) fractional Gaussian noise.

The latter is defined by the correlation function

$$\rho(\ell) = \tfrac{1}{2}[|\ell+1|^{2H} - 2\ell^{2H} + |\ell-1|^{2H}],$$

where the exponent H, called the "Hurst index," ranges from 0 (persistence) to 1 (antipersistence), with $H = 1/2$ characterizing the Gaussian white noise [48]. Figure 8 displays $\mathcal{C}_n(p_+ = 0.7)$ versus n for different Hurst indices $H = 0.5,0.7,0.9$. The asymptotic behavior at large n cannot be displayed here because of numerical restrictions, but the small n properties are more relevant for characterizing short-time conditional dynamics.

Similarly to Eq. (A1), the Student copula [illustrated in Fig. 2(b) for the bivariate case] is defined as

$$\mathcal{C}_n(p) = T_{\rho,\nu}(T^{-1}(p),\ldots,T^{-1}(p)),$$

where T^{-1} is the univariate inverse Student-t CDF and $T_{\rho,\nu}$ denotes the multivariate CDF with scale matrix ρ and ν degrees of freedom; see, e.g., Ref. [49] for a modern treatment.

APPENDIX B: PROOFS OF FORMULAS

1. Equation (12)

$$\pi(\tau) = \mathbb{P}[X_\tau > X^{(+)}, X_{[[1;\tau-1]]} < X^{(+)} | X_0 > X^{(+)}].$$

After a simple algebraic transformation flipping all ">" signs to "<", it can be written in the language of copulas as follows:

$$\begin{aligned}\pi(\tau) &= \frac{\mathbb{P}[X_\tau > X^{(+)}, X_{[[1;\tau-1]]} < X^{(+)}; X_0 > X^{(+)}]}{\mathbb{P}[X_0 > X^{(+)}]}\\ &= \frac{\mathbb{P}[X_{[[1;\tau-1]]} < X^{(+)}]}{p_+}\\ &\quad - \frac{\mathbb{P}[X_\tau < X^{(+)}, X_{[[1;\tau-1]]} < X^{(+)}]}{p_+}\\ &\quad - \frac{\mathbb{P}[X_{[[1;\tau-1]]} < X^{(+)}; X_0 < X^{(+)}]}{p_+}\\ &\quad + \frac{\mathbb{P}[X_\tau < X^{(+)}, X_{[[1;\tau-1]]} < X^{(+)}; X_0 < X^{(+)}]}{p_+},\\ \pi(\tau) &= \frac{\mathcal{C}_{\tau-1}(1-p_+) - 2\,\mathcal{C}_\tau(1-p_+) + \mathcal{C}_{\tau+1}(1-p_+)}{p_+}.\end{aligned} \tag{B1}$$

2. Equation (15)

$$\begin{aligned}\langle\tau^m\rangle &= \sum_{\tau=1}^{\infty}\tau^m\pi(\tau) = \frac{1}{p_+}\left\{\sum_{\tau=1}^{\infty}\tau^m\,\mathcal{C}_{\tau-1}(1-p_+) - 2\sum_{\tau=1}^{\infty}\tau^m\,\mathcal{C}_\tau(1-p_+) + \sum_{\tau=1}^{\infty}\tau^m\,\mathcal{C}_{\tau+1}(1-p_+)\right\}\\ &= \frac{1}{p_+}\left\{1 + \sum_{\tau=2}^{\infty}\tau^m\,\mathcal{C}_{\tau-1}(1-p_+) - 2\sum_{\tau=1}^{\infty}\tau^m\,\mathcal{C}_\tau(1-p_+) + \sum_{\tau=0}^{\infty}\tau^m\,\mathcal{C}_{\tau+1}(1-p_+)\right\}\\ \langle\tau^m\rangle &= \frac{1 + \sum_{\tau=1}^{\infty}[|\tau+1|^m - 2\tau^m + |\tau-1|^m]\,\mathcal{C}_\tau(1-p_+)}{p_+}.\end{aligned} \tag{B2}$$

3. Equation (17)

The probability of observing an interval τ' immediately following an observed recurrence time τ is, with the definition of Eq. (11),

$$\mathbb{P}[\mathcal{X}^{(+)}_{\tau;\tau'}|\mathcal{X}^{(+)}_{0;\tau},X_0 > X^{(+)}].$$

Again, flipping the ">" to "<" allows us to decompose it as

$$\begin{aligned}
\mathbb{P}[\mathcal{X}^{(+)}_{\tau;\tau'}|\mathcal{X}^{(+)}_{0;\tau},X_0 > X^{(+)}] &= \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n < \tau,1 \leqslant k < \tau']}{\mathbb{P}[X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n < \tau] - \mathbb{P}[X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n \leqslant \tau]} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n < \tau,1 \leqslant k \leqslant \tau']}{\mathbb{P}[X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n < \tau] - \mathbb{P}[X_n < X^{(+)},X_0 > X^{(+)},1 \leqslant n \leqslant \tau]} \\
&= \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},1 \leqslant n < \tau,1 \leqslant k < \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},0 \leqslant n < \tau,1 \leqslant k < \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},1 \leqslant n < \tau,1 \leqslant k \leqslant \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad + \frac{\mathbb{P}[X_{\tau+k} < X^{(+)},X_\tau > X^{(+)},X_n < X^{(+)},0 \leqslant n < \tau,1 \leqslant k \leqslant \tau']}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&= \frac{\mathcal{C}_{\tau-1;\tau'-1}(p) - \mathcal{C}_{\tau+\tau'-1}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} - \frac{\mathcal{C}_{\tau;\tau'-1}(p) - \mathcal{C}_{\tau+\tau'}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&\quad - \frac{\mathcal{C}_{\tau-1;\tau'}(p) - \mathcal{C}_{\tau+\tau'}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} + \frac{\mathcal{C}_{\tau;\tau'}(p) - \mathcal{C}_{\tau+\tau'+1}(p)}{\mathcal{C}_{\tau-1}(p) - 2\mathcal{C}_\tau(p) + \mathcal{C}_{\tau+1}(p)} \\
&= \frac{\mathcal{C}_{\tau-1;\tau'-1} - \mathcal{C}_{\tau;\tau'-1} - \mathcal{C}_{\tau-1;\tau'} + \mathcal{C}_{\tau;\tau'}}{\mathcal{C}_{\tau-1} - 2\mathcal{C}_\tau + \mathcal{C}_{\tau+1}} - \frac{\pi(\tau+\tau')}{\pi(\tau)},
\end{aligned} \tag{B3}$$

where

$$\mathcal{C}_{\tau;\tau'}(p) = \mathcal{F}_{[[0;\tau+\tau']]\backslash\{\tau\}}(F^{-1}(p),\dots,F^{-1}(p)).$$

[1] S. Kotz and S. Nadarajah, *Extreme Value Distributions* (Imperial College Press, London, 2000).
[2] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance* (Springer-Verlag, Berlin, 1997).
[3] A. Chakraborti, I. Muni Toke, M. Patriarca, and F. Abergel, Quant. Financ. **11**, 991 (2011).
[4] M. Politi, N. Millot, and A. Chakraborti, Physica A **391**, 147 (2012).
[5] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).
[6] A. Bunde, J. F. Eichner, S. Havlin, and J. W. Kantelhardt, Physica A **330**, 1 (2003).
[7] A. Corral, Phys. Rev. Lett. **92**, 108501 (2004).
[8] G. Molchan, Pure Appl. Geophys. **162**, 1135 (2005).
[9] E. G. Altmann and H. Kantz, Phys. Rev. E **71**, 056106 (2005).
[10] A. Saichev and D. Sornette, J. Geophys. Res. **112**, B04313 (2007).
[11] D. Sornette, S. G. Utkin, and A. Saichev, Phys. Rev. E **77**, 066109 (2008).
[12] R. Blender, K. Fraedrich, and F. Sienz, Nonlinear Proc. Geophys. **15**, 557 (2008).
[13] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley, Proc. Natl. Acad. Sci. USA **102**, 9424 (2005).
[14] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, Phys. Rev. E **73**, 026117 (2006).
[15] P. Weber, F. Wang, I. Vodenska-Chitkushev, S. Havlin, and H. E. Stanley, Phys. Rev. E **76**, 016109 (2007).
[16] N. Sazuka, J.-I. Inoue, and E. Scalas, Physica A **388**, 2839 (2009).
[17] F. Ren and W.-X. Zhou, New J. Phys. **12**, 075030 (2010).
[18] A. Corral, Phys. Rev. E **68**, 035102 (2003).
[19] A. Corral, Physica A **340**, 590 (2004).
[20] P. Embrechts, F. Lindskog, and A. McNeil, in *Handbook of Heavy Tailed Distributions in Finance*, edited by S. T. Rachev (Elsevier, Amsterdam, 2003), pp. 329–384, Chap. 8.
[21] P. Embrechts, A. J. McNeil, and D. Straumann, in *Risk Management: Value at Risk and Beyond* (Cambridge University Press, Cambridge, 2002), pp. 176–223.
[22] Y. Malevergne and D. Sornette, *Extreme Financial Risks: From Dependence to Risk Management* (Springer-Verlag, Berlin, 2006).
[23] T. Mikosch, Extremes **9**, 3 (2006).

[24] R. Chicheportiche and J.-P. Bouchaud, Int. J. Theor. Appl. Financ. **15**, 1250019 (2012).
[25] B. K. Beare, Econometrica **78**, 395 (2010).
[26] R. Ibragimov and G. Lentzas, Harvard Institute of Economic Research Discussion Paper No. 2160, 2008, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1185822.
[27] A. J. Patton, in *Handbook of Financial Time Series* (Springer-Verlag, Berlin, 2009), pp. 767–785.
[28] L. Isserlis, Biometrika **11**, 185 (1916).
[29] L. Isserlis, Biometrika **12**, 134 (1918).
[30] G.-C. Wick, Phys. Rev. **80**, 268 (1950).
[31] M. Boguná and J. Masoliver, Eur. Phys. J. B **40**, 347 (2004).
[32] R. Chicheportiche, Ph.D. thesis, École Centrale Paris, 2013, arXiv:1309.5073.
[33] F. Omori, J. Coll. Sci. **7**, 111 (1895).
[34] A. M. Petersen, F. Wang, S. Havlin, and H. E. Stanley, Phys. Rev. E **82**, 036114 (2010).
[35] F. Lillo and R. N. Mantegna, Phys. Rev. E **68**, 016119 (2003).
[36] B. Gutenberg and C. F. Richter, *Seismicity of the Earth and Associated Phenomena* (Hafner, New York, 1965).
[37] M. S. Santhanam and H. Kantz, Phys. Rev. E **78**, 051113 (2008).
[38] M. Kac, Bull. Am. Math. Soc. **53**, 1002 (1947).
[39] P. Olla, Phys. Rev. E **76**, 011122 (2007).
[40] J. F. Eichner, J. W. Kantelhardt, A. Bunde, and S. Havlin, Phys. Rev. E **73**, 016130 (2006).
[41] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, Phys. Rev. E **64**, 061907 (2001).
[42] R. Chicheportiche and J.-P. Bouchaud, J. Stat. Mech: Theor. Exp. (2011) P09003.
[43] J. Perelló, J. Masoliver, and J.-P. Bouchaud, Appl. Math. Financ. **11**, 27 (2004).
[44] J. Ludescher, C. Tsallis, and A. Bunde, Europhys. Lett. **95**, 68002 (2011).
[45] R. Chicheportiche and A. Chakraborti, arXiv:1302.3704 [physics.data-an] (2013).
[46] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. (Oxford University Press, Oxford, 2011).
[47] A. Sklar, Publ. Inst. Statist. Univ. Paris **8**, 229 (1959).
[48] B. B. Mandelbrot and J. W. Van Ness, SIAM Rev. **10**, 422 (1968).
[49] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge University Press, Cambridge, 2003).