



CENTRE DE RECERCA MATEMÀTICA

This is a preprint of: *Scaling in the timing of extreme events*
Journal Information: *CRM Preprints*,
Author(s): Á. Corral.
Volume, pages: 1-32,

DOI:[--]



CENTRE DE RECERCA MATEMÀTICA

Preprint núm. 1197
October 2014

Scaling in the timing of extreme events

Á. Corral

SCALING IN THE TIMING OF EXTREME EVENTS

ÁLVARO CORRAL

ABSTRACT. Extreme events can come either from point processes, when the size or energy of the events is above a certain threshold, or from time series, when the intensity of a signal surpasses a threshold value. We are particularly concerned by the time between these extreme events, called respectively waiting time and quiet time. If the thresholds are high enough it is possible to justify the existence of scaling laws for the probability distribution of the times as a function of the threshold value, although the scaling functions are different in each case. For point processes, in addition to the trivial Poisson process, one can obtain double-power-law distributions with no finite mean value. This is justified in the context of renormalization-group transformations, where such distributions arise as limiting distributions after iterations of the transformation. Clear connections with the generalized central limit theorem are established from here. The non-existence of finite moments leads to a semi-parametric scaling law in terms of the sample mean waiting time, in which the (usually unknown) scale parameter is eliminated but not the exponents. In the case of time series, scaling can arise by considering random-walk-like signals with absorbing boundaries, resulting in distributions with a power-law “bulk” and a faster decay for long times. For large thresholds the moments of the quiet-time distribution show a power-law dependence with the scale parameter, and isolation of the latter and of the exponents leads to a non-parametric scaling law in terms only of the moments of the distribution. Conclusions about the projections of changes in the occurrence of natural hazards lead to the necessity of distinguishing the behavior of the mean of the distribution with the behavior of the extreme events.

1. INTRODUCTION

The theory of self-organized criticality (SOC), introduced by Bak and collaborators, provides an appealing explanation for the distribution of sizes (and presumably durations) of many naturally occurring catastrophic events [1, 2]. Earthquakes, landslides, volcanic eruptions, rainfall, hurricanes, etc. [1, 3–5], have been found to display power-law distributions of event sizes, which implies the absence of characteristic scales for those sizes. This means that if you ask, for instance: “how big are earthquakes in Japan?”, this innocent question has no possible answer [6, 7]. In mathematical terms, this implies that after performing a scale transformation of a function describing the size of the events, the function has to remain unchanged, and the only univariate function with this property (for any dilatation or contraction) is the power law.

SOC explains this power-law behavior as coming from a self-organization process that leads the system to the critical point of a phase transition, where scale invariance is ensured. Other approaches, purely probabilistic, support the power-law tail in some cases, as described by the generalized Pareto distribution [8]. Nevertheless, the exponent of the power law in self-organized-critical systems takes such values that it is associated to an infinite expected value for the size of the events, which is unphysical; therefore, in practice, one cannot ignore finite-size effects.

When one is worried not about the size but about the time between events, the situation is a bit more involving. In fact, the waiting time is defined as the time between consecutive events when these events are above a certain threshold in size. In the usual case this approach takes place in a slow time scale, in which events happen instantaneously (in comparison with the waiting times, earthquakes are a good example). When the threshold is high enough, the events can be considered as extreme events, but there is no formal difference between these and ordinary events. Models of self-organized criticality are not very useful at this point, because, although the most common SOC models lead to a Poissonian occurrence of events and therefore to an exponential distribution of waiting times (in contrast to some observational data [9]), other SOC models yield correlated avalanches [10, 11]. One concludes from here that SOC is not about the temporal occurrence of events (in the slow time scale) but about their size (or duration).

But it is interesting anyhow to stretch the notion of scale invariance to the waiting-time problem. One deals now with two variables (size and waiting time), and in this case the scale transformation \mathcal{T} is written as

$$(1) \quad \mathcal{T}[f(x, y)] = cf(x/a, y/b),$$

with $f(x, y)$ a bivariate function, and $a, b, c > 0$, the scale factors of the transformation. When a, b, c are greater than one the transformation dilates the function in its three directions; for instance, taking $a = b = c = 100$ takes a function in the scale of, let us say, meters to the scale of centimeters.

The condition of scale invariance comes in this case from imposing

$$(2) \quad \mathcal{T}[f(x, y)] = cf(x/a, y/b) = f(x, y),$$

for all a . So, we are looking for a fixed point for \mathcal{T} and the only solution [6] can be written in the form

$$(3) \quad f^*(x, y) = \frac{1}{x^\alpha} H\left(\frac{x}{y^\beta}\right)$$

with H an arbitrary univariate function called scaling function and b and c not free but given by $c = 1/a^\alpha$ and $b = a^{1/\beta}$. We will refer to this functional form of $f(x, y)$ as a scaling law. Note that for the univariate problem, described just by x , the scaling function has to be a constant, and then a power law is left. In fact, a power law is a particular case of a scaling law, but not the opposite, and so

it is useful to keep the distinction between both concepts if one is involved with multivariate functions. The scaling law can be expressed in other, alternative, equivalent forms, as for instance,

$$(4) \quad f^*(x, y) = \frac{1}{x^\alpha} H_2 \left(\frac{y}{x^{1/\beta}} \right) = \frac{1}{y^{\alpha\beta}} H_3 \left(\frac{y}{x^{1/\beta}} \right),$$

etc., just remember that the scaling function H is arbitrary.

However, one can immediately realize that, when events above a low threshold in size are compared to events with higher size, the analysis of the waiting-time does not only involve a scale transformation. Indeed, an additional coarse-graining transformation is operating. This comes from the fact that events below the threshold need to be eliminated, in order to measure the time between consecutive events above the threshold. So, the transformation of the data comprises a coarse-graining followed by a scale transformation [12, 13]. This is the essence of a renormalization-group transformation [6].

On the other hand, one can analyze SOC models and SOC-like systems in a complementary way, not in the slow time scale in which almost instantaneous events happen but in the fast time scale inside of the events. Then, one takes a threshold not in size but in intensity (size per unit time) and measures the analogous of the waiting time, the time the signal is below threshold, which in this context is called quiet time. This approach was pursued in Ref. [14], where a totally different behavior was found, with a power-law distribution of quiet times (in agreement then with some observational results [9]). Naturally, this framework is valid not only for SOC-like systems but for any time series from which events are defined by a process of thresholding.

This paper has as a goal to approach both kind of problems, waiting times defined in a slow time scale and quiet times defined in a fast time scale, from the point of view of scaling theory. In the next section the renormalization-group approach to the waiting-time sequence will be revisited and the precise form of the resulting scaling law will be presented for two broad class of processes, in analogy with the central limit theorem. A different scaling form for the quiet-time distribution will be used in the third section, generalizing the result for the diffusion of a Brownian particle. This yields however important consequences for the projection of extreme-event recurrence. The result is extended to the interesting case in which the right tail of the distribution is a power law, with such an exponent that the expected value of the variable is infinite. The paper illustrates in which way one has to extrapolate the statistics of ordinary events in order to infer the recurrence of extreme events in a system that shows scaling.

2. SCALING IN MARKED POINT PROCESSES

A point process, for our purposes, is a set of point events, i.e., instantaneous events, that occur in time [15–17]. The simplest point process is the Poisson process, in which the events take place at a constant rate, independently of any

other event. A marked point process is a point process in which the points, i.e., the events, carry some “mark”, which in our case corresponds to their size or dissipated energy.

As mentioned in the introduction, when one considers events of a large enough size, the original point process transforms or renormalizes to a new point process. Events below the threshold in size are eliminated whereas events above the threshold survive. This coarse graining or decimation is called thinning [15] or rarefaction [18] in the theory of point processes. As a special class of marked point processes we will consider in this section what is called marked renewal processes [16], in which the size of events will be independent on other sizes and on the time of occurrence of the events, and additionally, the waiting times will be independent on previous waiting times (and constitute then what is called a renewal processes [15, 16], in which only the time of occurrence of the last event determines the time of the next one).

By construction, sizes are independent random variables, thus, the elimination of events below the threshold in size is equivalent to a random thinning. In this case, the waiting-time density transforms after thinning to

$$(5) \quad \tilde{f}_s(\omega) = \frac{p\tilde{f}(\omega)}{1 - (1-p)\tilde{f}(\omega)}$$

where p is the probability of surviving to the thinning, i.e., the probability of being above the threshold s , $\tilde{f}(\omega)$ is the Laplace transform of the original waiting-time probability density $f(t)$, and $\tilde{f}_s(\omega)$ is the Laplace transform of the waiting-time density for events of size larger than s , see Refs. [12, 13, 15, 18]. The Laplace transform arises because the waiting-time density for events above the threshold is the convolution of a random number of densities for all events. Then, going to Laplace space simplifies considerably the equations, as convolutions transform into simple products there.

The subsequent step of the renormalization transformation is a scale transformation. This can be done in two different ways, a trivial one and a non-trivial one, as we will show in the next two subsections.

2.1. Linear rescaling and trivial Poissonian fixed point. Let us considering the following scale transformation for the waiting-time density for events above size s , $f_s(t) \rightarrow p^{-1}f_s(t/p)$, or, in Laplace space,

$$(6) \quad \tilde{f}_s(\omega) \rightarrow \tilde{f}_s(p\omega).$$

This is because we expect that thinning increases the time scale in a $1/p$ factor, and so, in order to compensate for this fact, we perform a contraction of the function in the x -axis with a scale factor $a = p < 1$, whereas we perform a dilatation of the function in the y -axis with a scale factor $c = 1/p > 1$, in order to keep normalization. We will see in the next subsection that this makes sense only if the mean waiting time $\langle t \rangle$ is finite. The combined effect of the thinning

(5) and scale transformation yields the complete renormalization transformation \mathcal{R} , which is then

$$(7) \quad \mathcal{R}[\tilde{f}(\omega)] = \frac{p\tilde{f}(p\omega)}{1 - q\tilde{f}(p\omega)}$$

with $q = 1 - p$. The solution of the fixed-point condition $\mathcal{R}[\tilde{f}^*(\omega)] = \tilde{f}^*(\omega)$, for all p , leads to the fixed-point solution or renormalization-invariant Laplace transform of the density $\tilde{f}^*(\omega)$. Defining a new variable $u = p\omega$ and separating the variables ω and u leads to the Laplace transform of an exponential distribution, i.e.,

$$(8) \quad \tilde{f}^*(\omega) = \frac{1}{1 + \ell\omega} \Rightarrow f^*(t) = \frac{1}{\ell} e^{-t/\ell},$$

with ℓ an arbitrary constant, see Refs. [12, 13, 15]. This exponential distribution corresponds, due to the independence of the events in the process, to the Poisson process. In words, the Poisson process is invariant under the renormalization transformation given by Eq. (7). This is obvious if one intuitively knows the properties of the Poisson process. But it is not only that the Poisson process is invariant under the transformation (7), the results also tells us that the Poisson process is the only marked renewal process invariant under such transformation [13, 15]. And, even more, it is easy to show that it is also an attractor for any marked renewal process with waiting-time density with a finite mean (and whose Laplace transform exists), see Ref. [13].

This leads us to the first scaling form or data collapse considered in this paper. Remember that in order to compare the thinned or coarse-grained distribution with the original one we perform a scale transformation of the form $f_s(t) \rightarrow p^{-1}f_s(t/p)$. This means that the fixed-point solution, i.e., the resulting probability density for very high thresholds in size fulfills the following scaling law

$$(9) \quad f_s(t) = pF(pt),$$

where the scaling function F is a decreasing exponential, but we will see that F can take more general forms in other cases. Therefore, if plotting $f_s(t)/p$ versus pt yields a data collapse, i.e., a single curve, independently on the value of p , this implies that the scaling law is fulfilled.

Other equivalent forms of the scaling law can be obtained from the fact that the parameter p gives the probability that an event has energy E larger than the threshold s (given that has energy larger than a reference level 0), i.e.,

$$(10) \quad p = \text{Prob}[E > s | E > 0] \simeq \frac{N_s}{N_0}$$

where N_s is the number of events with energy above s . Thus, the scaling $t \rightarrow pt$ is equivalent to $N_0 t \rightarrow N_s t$ but this is only valid if the window width T is fixed. If time windows of different width are compared one needs to correct for this fact;

so, the rescaling is equivalent to

$$(11) \quad R_0 t = \frac{N_0}{T_0} t \rightarrow R_s t = \frac{N_s}{T_s} t$$

where R_s is the rate of events above s , i.e., the number of events above s per unit time. The situation of comparing different time windows is common when dealing with real data of natural catastrophes, as different sizes have different windows of completeness (see caption of Fig. 23 in Ref. [19]). Then, the scaling law (9) can be alternatively written as

$$(12) \quad f_s(t) = R_s F(R_s t) = \frac{1}{\langle t \rangle} F\left(\frac{t}{\langle t \rangle}\right),$$

where we have introduced the mean waiting time $\langle t \rangle = 1/R_s$ (which of course depends on the threshold s although this is not reflected in the notation). Note that the scaling function F appearing here is not exactly the same as the one in Eq. (9), as two multiplicative constants are reabsorbed in F . In this case the scaling law can be fulfilled by plotting the dimensionless waiting-time density $f_s(t)/R_s = \langle t \rangle f_s(t)$ versus the dimensionless time $R_s t = t/\langle t \rangle$ (in other words, plotting the waiting-time density in units of R_s versus the waiting time in units of $\langle t \rangle$).

In order to end with the different scaling forms, if there is scale invariance in the sizes (as it happens in SOC systems, in earthquakes, etc.), then $p \propto 1/s^\phi$, so, an additional form for the scaling law is

$$(13) \quad f_s(t) = \frac{1}{s^\phi} F\left(\frac{t}{s^\phi}\right),$$

where again, some multiplicative constant has been dropped deliberately. This form of the scaling law is the “most natural one” in the sense that it is written in terms of the two variables of the problem, t and s . The equivalence with the scale-invariant form of Eq. (3) is achieved by realizing that $\alpha = 1/\beta = \phi$.

These scaling laws, Eqs. (9), (12), and (13), are very useful in practice, as they have been shown to show up in several natural hazards, as for instance, earthquakes [20]. However, the scaling function F is not provided by an exponential function alone; rather, it contains a decreasing power-law for small to intermediate times, with exponent $\nu < 1$. Other authors, using a simplified model of seismicity, have argued for a different behavior, with no scaling law [21]. In any case, the simple approach in this section, based in a process with independent times and sizes, is clearly not valid for earthquakes and other natural hazards, in which important dependences exist [22].

Other examples of a general scaling law under thinning with non-exponential scaling function include fractures [23,24], forest fires [25], and, curiously, printing requests [26]. In the case of fires only Eqs. (9) and (12) are valid, and not Eq. (13), as the fire size is not power-law distributed, for the particular data analyzed in Ref. [25]. The scaling is also found in a different approach for earthquakes

[27–29], but there not only simple thinning is performed and the scaling under thinning should arise from the underlying scaling explained in Ref. [20]. Other references provide only very indirect evidence of the scaling, as in nanofractures [30] and in tsunamis [31], the problem being that no different thresholds in size are considered, although the resulting scaling functions seem to be the same as for earthquakes. The recurrence of words in texts, despite yielding a function very similar to the earthquake case, has nothing to do with thinning or renormalization [32]. Finally, approaches in which the threshold is imposed not in size but in intensity will be considered in the next section [33–35], whereas the approach of Ref. [36] is for a transformation of the original signal.

2.2. Nonlinear rescaling and non-trivial Poissonian fixed point . We explain in this subsection how the linear scaling with p in Eq. (6) is not the only reasonable possibility. Let us consider instead a constant $r > 1$ and perform the rescaling as $f_s(t) \rightarrow p^{-r} f_s(t/p^r)$; this yields the (completed) renormalization transformation \mathcal{R} , using Eq. (5),

$$(14) \quad \mathcal{R}[\tilde{f}(\omega)] = \frac{p\tilde{f}(p^r\omega)}{1 - q\tilde{f}(p^r\omega)}.$$

The fixed point condition leads in this case, in the same way as in the previous subsection, to

$$(15) \quad \tilde{f}^*(\omega) = \frac{1}{1 + \ell\omega^{1/r}}.$$

It is easy to show [13] that, for $r > 1$, this corresponds to a waiting time density with a double power-law behavior, i.e.,

$$(16) \quad f^*(t) \propto \begin{cases} 1/t^{1-1/r} & \text{for } t \ll \ell^{1/r} \\ 1/t^{1+1/r} & \text{for } t \gg \ell^{1/r}, \end{cases}$$

which leads to an infinite mean, as the tail is a power law with an exponent $1 + 1/r < 2$. In the same way as for the trivial $r = 1$ case, this distribution is not only the only fixed point, for a fixed r , but it is also an attractor for distributions with power-law tail, of the form $f(t) \sim B/t^{1+1/r}$. The case $r < 1$ does not lead to well-defined probability distributions.

The non-linear rescaling with p^r is explained by the generalized central limit theorem. Indeed, as the times t_i have a distribution with a power-law tail, with exponent smaller than 2, the classic central limit theorem does not hold and one has instead that the total time T scales as

$$(17) \quad T = \sum_{i=1}^N t_i \propto N^r,$$

see Ref. [37] or the Appendix I here. Thus, after thinning by a factor p , the time window which restores the number of events N is the one that contains N/p

events, and this is the time window of duration T/p^r . In other words, we need to rescale time as $t \rightarrow t/p^r$. This is the origin of the non-linear rescaling.

In the same way as in the previous subsection, invariance under the renormalization transformation means that the invariant density is of the form

$$(18) \quad f_s(t) = p^r G(p^r t),$$

with G a scaling function given by the inverse Laplace transform of Eq. (15), which we have explained has a double power-law form.

Expressions alternative to the previous rescaling can be obtained as before, by using that

$$(19) \quad p^r \simeq \frac{N_s^r/T_s}{N_0^r/T_0} \propto \frac{1}{s^{\phi r}}$$

where N_s is the number of events with size (energy) above s and we have taken into account that, in practice, the time window can be different for events above size s than for events above the reference size 0.

Therefore, the scaling law (18) can be written as

$$(20) \quad f_s(t) = \frac{1}{\sum t_i/N_s^r} G\left(\frac{t}{\sum t_i/N_s^r}\right) = \frac{N_s^{r-1}}{\bar{t}_s} G\left(\frac{N_s^{r-1}t}{\bar{t}_s}\right),$$

where $\sum t_i$ refers to events above threshold s and \bar{t}_s the sample mean of the waiting time above s (different from the mean of the distribution, which is infinite). Note that the generalized central limit theorem implies that \bar{t}_s diverges with the number of data, but $\sum t_i/N_s^r$ does not. In the case $r = 1$, then $\bar{t}_s \simeq \langle t \rangle$ and we recover the scaling of the previous subsection. If we put the size-dependence explicitly,

$$(21) \quad f_s(t) = \frac{1}{s^{\phi r}} G\left(\frac{t}{s^{\phi r}}\right).$$

In this one and in Eq. (20) we have deliberately ignored some multiplicative constants. Comparison with the scale-invariant solution of the scale transformation, Eq. (3), shows that $\alpha = 1/\beta = \phi r$.

A marked renewal process of the sort of the one in this subsection is neither implausible nor artificial. Indeed, consider that a hidden signal triggers instantaneous events (as earthquakes) when it reaches a fixed threshold and that at that point the signal is immediately reset just below the threshold. If the sizes of the events are independent and the signal cannot reach an absorbing boundary below it (this can be achieved in practice by putting the boundary of a random walk at $-\infty$, or by taking the exponentiation of the random walk, which never reaches the absorbing boundary at zero), then this leads to a marked renewal process with waiting-time density given by the so-called Lévy-Smirnov distribution, see Eq. (27), which has a right power-law tail with exponent $3/2$ [38]. Iterative application of the thinning transformation leads to a different distribution, which

keeps the power-law tail with exponent $3/2$ for large times but develops a new power law in the regime of small t , with exponent $1/2$.

2.3. Paralellism with the generalized central limit theorem. Several references have pointed out the relationship between the central limit theorem and renormalization [39–41], including the generalized case [42]. From our perspective, we can provide a clear connection between the renormalization under random thinning explained in the previous subsections and the generalized central limit theorem. The key is to consider a point process in which events are not removed randomly but deterministically, surviving a proportion p , in such a way that out of $1/p$ events, the $1/p$ -th survives and the rest are removed (with $1/p$ integer, for instance, if $p = 1/2$ the events are removed alternatively). Then, for a process in which the waiting times are independent (i.e., a renewal process), this deterministic thinning transformation can be written as

$$(22) \quad f_s(t) = [f(t)]^{\star 1/p},$$

where $f(t)$ is the probability density of the original point process, $f_s(t)$ the corresponding one after thinning (associated to events above size s), and $\star 1/p$ denotes convolution $1/p$ times. In Laplace space, the combined thinning plus rescaling transformation becomes

$$(23) \quad \mathcal{R}[\tilde{f}(\omega)] = [\tilde{f}(p^r \omega)]^{1/p},$$

which is the equivalent of Eq. (14), and then, the fixed point condition for \mathcal{R} , taking logarithms, fulfills

$$(24) \quad \ln \tilde{f}^*(\omega) = \frac{1}{p} \ln \tilde{f}^*(p^r \omega)$$

This means that $\ln \tilde{f}^*(\omega)$, which is a cumulant generating function, has to fulfill the scale-invariance condition. If we impose the fixed point condition for all p the solution is then a power law,

$$(25) \quad \ln \tilde{f}^*(\omega) = -A\omega^{1/r},$$

and for the Laplace transform of the density we find,

$$(26) \quad \tilde{f}^*(\omega) = e^{-A\omega^{1/r}}.$$

For the case $r = 1$ we recognize the Laplace transform of a Dirac's delta function, $f^*(t) = \delta(t - A)$, whereas for $r = 2$ we get

$$(27) \quad f^*(t) = e^{-A^2/(4t)} \frac{A}{2\sqrt{\pi} t^{3/2}},$$

which is sometimes called Lévy-Smirnov distribution. For any $r > 1$ the fixed-point distribution has a power-law tail, for large times, with exponent $1 + 1/r < 2$ (with infinite mean therefore). As one can see in Appendix I, these distributions are also attractors, with different domains of attraction. The case $r = 1$ yields the classic law of large numbers [43], in the form of the attractiveness of the

Dirac's delta function, whereas $r > 1$ leads to a certain case of the generalized central limit theorem [37].

3. SCALING IN THE QUIET TIMES OF TIME SERIES

The previous section was devoted to point processes in which the size of the events was larger than some threshold. From the point of view of SOC, we were looking at the system in the slow time scale. Here we analyze the equivalent problem in the fast time scale (if there are two time scales), in order to see similarities and differences. Our approach will be also valid for non-SOC systems, with only one time scale, as usual time series. The procedure consists simply in introducing a threshold in the value of the signal (intensity or activity in the SOC language) and define the quiet time as the time the signal is below threshold. Several papers have used this approach before [33–35]. Nevertheless, the results will be of broader applicability.

3.1. Non-parametric scaling form. In Appendix II we consider an intensity signal modeled by a Brownian noise between two absorbing boundaries (the zero-intensity state and the threshold) and show how the quiet-time distribution fulfills a scaling law. Let us generalize here that scaling law, replacing the Brownian exponent $3/2$ by a generic and undefined exponent ν , and considering a generic scaling function G , then we write

$$(28) \quad f_\tau(t) = \frac{1}{\tau^\nu} G\left(\frac{t}{\tau}\right),$$

where the scale parameter is now called τ , and enters as a subindex of the probability density $f_\tau(t)$ to indicate the dependence on the threshold (as τ depends on it). The density is defined for $t > m$ and is zero otherwise. The scaling function G has to decrease, for large arguments, fast enough, which in our case means faster than $1/t^6$, as we are interested, at most, in the calculation of the third moment (i.e., we ask not only for the finiteness of the expected value of t^3 , which is $\langle t^3 \rangle$, but also for the finiteness of the second moment of t^3 , which is $\langle t^6 \rangle$). This is the “worst” scenario and in some case G can decay slower than $1/t^6$, the main requirement being that the moments involved in the equations are finite, as well as their variances.

As shown in Appendix II, the scaling function G can be a bit complicated and it is more convenient to write the scaling law in a slightly modified form, separating a power-law behavior for intermediate values from the rest of the scaling function,

$$(29) \quad f_\tau(t) = \frac{K}{\tau} \left(\frac{\tau}{t}\right)^\nu F\left(\frac{t}{\tau}\right)$$

for $t > m$, where we assume that the function F tends to a positive constant at zero, and K is a normalization constant. This scaling form also comprises the scaling laws of the previous section, taking $0 < \nu < 1$ for the case of earthquakes

[20], and $\nu = 0$ for the trivial Poisson fixed point of uncorrelated point processes. As, for small arguments, the scaling function F tends to a constant and decreases “fast enough” for large arguments, the scale parameter τ separates at least two regimes, being the one for $t \gg \tau$ the corresponding for extreme events; in other words, τ sets the scale of extreme events. It is worth mentioning that here we use a broader definition of scaling function than the one of Ref. [44], as there it is enforced that the scale parameter τ only appears in the argument of the scaling function, and not outside. The use of that prescription or the use of our scaling form is a matter of choice and has no fundamental implications.

Thus, in the case $\nu < 1$ we can take $m = 0$ and verify that K is a constant indeed and then it is reabsorbed into the scaling function (i.e., $K = 1$). However, for $\nu > 1$ one finds that K is not a true constant but proportional to $(m/\tau)^{\nu-1}$, and then m cannot be zero. The scaling law for $\nu > 1$ turns out to be

$$(30) \quad f_\tau(t) = \frac{1}{m} \left(\frac{m}{t} \right)^\nu F \left(\frac{t}{\tau} \right) \text{ for } t > m > 0.$$

Both cases, $\nu < 1$ and $\nu > 1$, can be summarized into a unique scaling law,

$$(31) \quad f_\tau(t) = \frac{1}{\theta} \left(\frac{\theta}{t} \right)^\nu F \left(\frac{t}{\tau} \right) \text{ for } t > m,$$

with $\theta = \tau$ and $m = 0$ for $\nu < 1$ and with $\theta = m$ and $m > 0$ for $\nu > 1$. Note that for $\nu > 1$ we can write

$$(32) \quad f_\tau(t) = \frac{1}{m} \left(\frac{m}{\tau} \right)^\nu G \left(\frac{t}{\tau} \right),$$

although for $\nu < 1$ we still have $f_\tau(t) = G(t/\tau)/\tau$.

The scaling law becomes more apparent substituting the scale parameter τ by its scaling with the threshold L (assuming it happens, as for a Brownian particle), i.e., $\tau \propto L^\phi$. In this way

$$(33) \quad f_\tau(t) = \frac{1}{L^\phi} G \left(\frac{t}{L^\phi} \right),$$

for $\nu < 1$, and

$$(34) \quad f_\tau(t) = \frac{1}{m} \left(\frac{m}{L^\phi} \right)^\nu G \left(\frac{t}{L^\phi} \right),$$

for $\nu > 1$.

If one is interested in describing the distribution, instead of by its probability density, by its cumulative or survivor function, defined by $S_\tau(t) = \int_t^\infty f_\tau(t') dt'$, one has that

$$(35) \quad S_\tau(t) \propto \begin{cases} I(t/\tau) & \text{for } \nu < 1, \\ (m/\tau)^{\nu-1} I(t/\tau) & \text{for } \nu > 1, \end{cases}$$

with I a new scaling function. Anyhow, for us the most useful quantity will be the density, $f_\tau(t)$.

At this point it becomes interesting to see the consequences of the scaling law on the moments of the distribution (moments about the origin, $t = 0$). One can easily calculate the mean for $\tau \gg m$ [6],

$$(36) \quad \langle t \rangle = \int_m^\infty t f_\tau(t) dt \propto \begin{cases} \tau & \text{for } \nu < 1, \\ m(\tau/m)^{2-\nu} & \text{for } 1 < \nu < 2, \\ m & \text{for } 2 < \nu, \end{cases}$$

and in the same way the second moment,

$$(37) \quad \langle t^2 \rangle = \int_m^\infty t^2 f_\tau(t) dt \propto \begin{cases} \tau^2 & \text{for } \nu < 1, \\ m^2(\tau/m)^{3-\nu} & \text{for } 1 < \nu < 3, \\ m^2 & \text{for } 3 < \nu, \end{cases}$$

The idea is that the scaling forms we have seen up to now are parametric, in the sense that depend on unknown parameters ν and τ . The exponent ν may be calculated with careful methods [45], but the scale parameter needs a particular parameterization of the scaling function G , which may be arbitrary. An alternative, free of these restrictions, is to use a non-parametric scaling form.

For the case $\nu < 1$ we know that $\tau \propto \langle t \rangle$ (see Eq. (36)) and then we recover the scaling form proposed in the previous section [20],

$$(38) \quad f_\tau(t) = \frac{1}{\langle t \rangle} G\left(\frac{t}{\langle t \rangle}\right),$$

where the constant of proportionality between τ and $\langle t \rangle$ is reabsorbed in G . We could also have used that $\tau \propto \sqrt{\langle t^2 \rangle}$, but the scaling with $\langle t \rangle$ is preferred as the computation of the first moment has a smaller error than that of the second moment.

For the new case $1 < \nu < 2$ we cannot apply the previous scaling, and instead, we have, on the one hand, dividing Eqs. (37) and (36),

$$(39) \quad \tau \propto \frac{\langle t^2 \rangle}{\langle t \rangle},$$

and, on the other hand, isolating from Eq. (36) and substituting in the previous one,

$$(40) \quad \left(\frac{m}{\tau}\right)^\nu \propto \frac{\langle t \rangle}{m} \left(\frac{m}{\tau}\right)^2 \propto m \frac{\langle t \rangle^3}{\langle t^2 \rangle^2}.$$

Substituting this into the scaling law (32) we get

$$(41) \quad f_\tau(t) = \frac{\langle t \rangle^3}{\langle t^2 \rangle^2} G\left(\frac{\langle t \rangle}{\langle t^2 \rangle} t\right),$$

see Refs. [4, 46]. Figure 1 shows an example of this rescaling.

An intermediate, semi-parametric scaling form is also possible, if one knows the exponent ν but not the scale parameter τ , in this case one may write, from

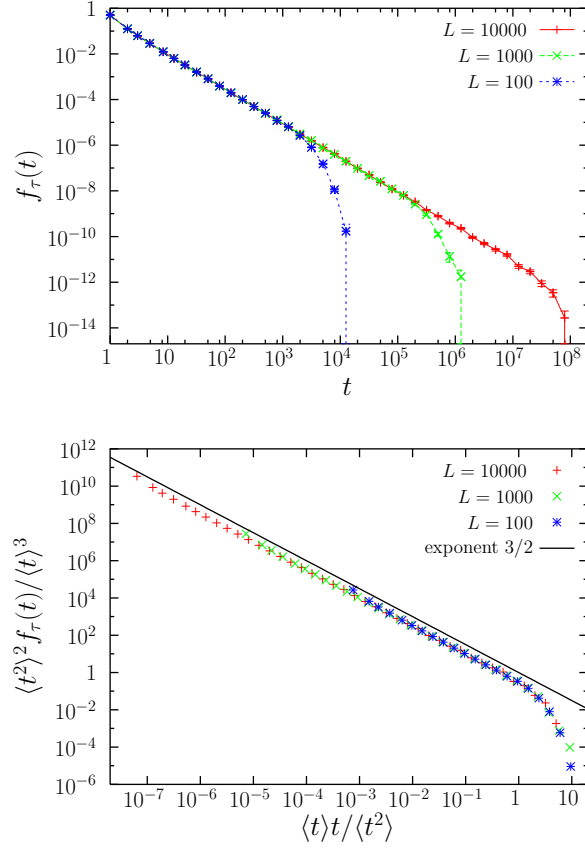


FIGURE 1. (a) Probability densities for the first-passage time of a random walk between two absorbing boundaries. Every time step the walker has probability $1/2$ of increasing its position by 1 and the same probability to decrease it by 1 . The initial position is one unit below the threshold. The time steps take place every 0.5 time units (in order that the first-passage times take values $1, 2, 3$, etc.) Different values of the threshold L are considered. (b) Same distributions under the non-parametric rescaling given by Eq. (41). The data collapse is an indication of the fulfillment of the scaling law.

Eq. (36), $\tau \propto m(\langle t \rangle / m)^{1/(2-\nu)}$ and from here,

$$(42) \quad f_\tau(t) = \frac{1}{m} \left(\frac{m}{\langle t \rangle} \right)^{\nu/(2-\nu)} G \left(\left(\frac{m}{\langle t \rangle} \right)^{1/(2-\nu)} \frac{t}{m} \right),$$

see Refs. [5, 47].

For completeness, it is interesting to analyze the scaling behavior for $2 < \nu < 3$. In this case the mean does not depend on the scale parameter, and we have to

replace the mean with the third moment, which scales as $\langle t^3 \rangle \propto m^3(\tau/m)^{4-\nu}$, valid for $1 < \nu < 4$. Using also the scaling for $\langle t^2 \rangle$ and dividing both,

$$(43) \quad \frac{\langle t^3 \rangle}{\langle t^2 \rangle} \propto \tau,$$

and isolating from the scaling of $\langle t^2 \rangle$,

$$(44) \quad \left(\frac{m}{\tau}\right)^\nu \propto \frac{\langle t^2 \rangle}{m^2} \left(\frac{m}{\tau}\right)^3,$$

and substituting the scaling of τ ,

$$(45) \quad \left(\frac{m}{\tau}\right)^\nu \propto m \frac{\langle t^2 \rangle^4}{\langle t^3 \rangle^3},$$

and substituting in the scaling law (32),

$$(46) \quad f_\tau(t) = \frac{\langle t^2 \rangle^4}{\langle t^3 \rangle^3} G\left(\frac{\langle t^2 \rangle t}{\langle t^3 \rangle}\right).$$

valid in principle for $1 < \nu < 3$ but useful in practice for $2 < \nu < 3$. A semi-parametric scaling form, if one knows ν but not τ , can be found from the fact that $m/\tau \propto (m^2/\langle t^2 \rangle)^{1/(3-\nu)}$ and then,

$$(47) \quad f_\tau(t) = \frac{1}{m} \left(\frac{m^2}{\langle t^2 \rangle}\right)^{\nu/(3-\nu)} G\left(\left(\frac{m^2}{\langle t^2 \rangle}\right)^{1/(3-\nu)} \frac{t}{m}\right).$$

The different non-parametric scaling functions, for the different ranges of ν , are summarized in Table 1.

Scaling has important consequences for the projection of extreme events. We have presented our scaling analysis for the waiting time or quiet time t of events, but of course the results are more general, being suitable for other variables different than t . Considering the scaling form (29), valid for $\nu < 1$, we have seen that $\langle t \rangle \propto \tau$; in consequence, an increase in the mean has associated a proportional increase in the scale of extreme events, quantified by τ . But for the scaling form (30), linked to $\nu > 1$, the mean does not scale linearly with the scale τ of extreme events. In the case $1 < \nu < 2$ we have $\langle t \rangle \propto \tau^{2-\nu}$ (see Eq. (36)); this means, for instance, that a two-fold increase in the mean of the distribution implies an increase of a factor $2^{1/(2-\nu)}$ in the scale of the largest, extreme events (a quadruplication if $\nu = 3/2$ and much higher if ν approaches 2). This is the case of the energy released by hurricanes [5], for which the mean of the distribution increases with sea-surface temperature, but the increase of the largest energies is higher. On the other hand, if $\nu > 2$, the mean and the scale of extreme events become unrelated, as $\langle t \rangle \propto m$ (i.e., a constant). As a recent publication states, “*much of today’s research represents climate change in terms of changes to the annual or seasonal averages*”, whereas the subject of major concern is the change corresponding to the most devastating, extreme events [48].

TABLE 1. Non-parametric rescaling of the axes coming from the scaling law $f_\tau(t) = \theta^{-1}(\theta/t)^\nu F(t/\tau)$ with $\theta = \tau$ for $\nu < 1$ and with $\theta = m$ for $\nu > 1$. The scaling function F decays fast enough in order that the moments up to $\langle t^3 \rangle$ and their variances are not infinite. The equivalent rescaling for the survivor function $S_\tau(t) = (\theta/\tau)^{\nu-1} I(t/\tau)$ is also included.

	abscissa	ordinate for $f_\tau(t)$	ordinate for $S_\tau(t)$
$\nu < 1$	$t/\langle t \rangle$	$\langle t \rangle f_\tau(t)$	$S_\tau(t)$
$1 < \nu < 2$	$\langle t \rangle t / \langle t^2 \rangle$	$\langle t^2 \rangle^2 f_\tau(t) / \langle t \rangle^3$	$\langle t^2 \rangle S_\tau(t) / \langle t \rangle^2$
$1 < \nu < 3$	$\langle t^2 \rangle t / \langle t^3 \rangle$	$\langle t^3 \rangle^3 f_\tau(t) / \langle t^2 \rangle^4$	$\langle t^3 \rangle^2 S_\tau(t) / \langle t^2 \rangle^3$

3.2. Case with infinite mean generalized. In the previous section (subsec. 2.2) we saw how, in some peculiar cases, marked renewal processes do not converge under thinning to the Poisson process but to a process in which the waiting-time density shows a double (decreasing) power-law behavior, with an exponent between 0 and 1 for short times, and another exponent between 1 and 2 for long times (and with the sum of both exponents equal to 2). This behavior in the right tail is associated to an infinite mean-waiting time $\langle t \rangle$.

In this subsection we study the proper scaling of such distributions in a more general framework, considering distributions (no matter if for the waiting time, the quiet time, or any other variable showing scaling) as those given by Eqs. (31) and (32),

$$(48) \quad f_\tau(t) = \frac{1}{\theta} \left(\frac{\theta}{t} \right)^\nu F \left(\frac{t}{\tau} \right) = \frac{1}{\theta} \left(\frac{\theta}{\tau} \right)^\nu G \left(\frac{t}{\tau} \right),$$

defined for $t > m$, with $m = 0$ and $\theta = \tau$ if $\nu < 1$ and with $m > 0$ and $\theta = m$ if $\nu > 1$. Note that t is the variable and τ is the scale parameter. The difference with the previous subsection is that $f_\tau(t)$ has now a (right) power-law tail, with exponent between 1 and 2, i.e., for large arguments G decays as a power law with exponent $1 + \rho$, with $0 < \rho < 1$, and then $\langle t \rangle$ is infinite, as in subsec. 2.2. But in contrast to that subsection, and in the same way as in the previous one, F goes to a constant for small arguments, so G goes as another (decreasing) power law there, with exponent ν , with no particular relation between both exponents ($1 + \rho$ and ν).

In the same way as before, we are interested in relating the mean value of the variable with its scale parameter τ , with the difference that now the expected value $\langle t \rangle$ is infinite and we have to emphasize the role of the sample mean, $\bar{t} = \sum_i t_i / N$ (which is always finite if N is finite). The power-law tail of the distribution means that $f_\tau(t)$ goes, for large enough t as $B/t^{1+\rho}$. We can realize that a scale parameter is given by $\tau = B^{1/\rho}$. Applying the results of the generalized

central-limit theorem explained in Appendix I, we get that

$$(49) \quad \sum_{i=1}^N t_i \sim [-\Gamma(\rho)BN]^{1/\rho},$$

where the symbol “ \sim ” means that, for large N , $\sum_i t_i$ is distributed following a certain (Lévy-stable) distribution whose scale parameter is $[-\Gamma(\rho)BN]^{1/\rho}$. From here we get that the scale parameter τ of $f_\tau(t)$ can be related to $\sum_i t_i$ and therefore with the sample mean as

$$(50) \quad \tau = B^{1/\rho} \sim \frac{\sum_{i=1}^N t_i}{[-\Gamma(\rho)N]^{1/\rho}} \propto \frac{\bar{t}}{N^{1/\rho-1}}.$$

The case $\nu < 1$, for which $\theta = \tau$, can be obtained straightforwardly from Eq. (48), which leads to

$$(51) \quad f_\tau(t) = \frac{N^{1/\rho-1}}{\bar{t}} G\left(\frac{N^{1/\rho-1}t}{\bar{t}}\right)$$

and is of the same form as Eq. (20). For the new case $\nu > 1$ we need to consider $\theta = m$, and from Eq. (48), substituting τ , we get

$$(52) \quad f_\tau(t) = \frac{1}{m} \left(\frac{mN^{1/\rho-1}}{\bar{t}}\right)^\nu G\left(\frac{N^{1/\rho-1}t}{\bar{t}}\right).$$

Table 2 summarizes the results for $\nu < 1$ and $\nu > 1$.

TABLE 2. Semi-parametric rescaling of the axes coming from the scaling law $f_\tau(t) = \theta^{-1}(\theta/t)^\nu F(t/\tau)$ with $\theta = \tau$ for $\nu < 1$ and with $\theta = m$ for $\nu > 1$. The right tail of F decays slow enough in order that $\langle t \rangle$ is infinite. In contrast, the sample mean \bar{t} is finite if N is finite.

	abscissa	ordinate
$\nu < 1$	$N^{1/\rho-1}t/\bar{t}$	$\bar{t}f_\tau(t)/N^{1/\rho-1}$
$\nu > 1$	$N^{1/\rho-1}t/\bar{t}$	$(\bar{t})^\nu f_\tau(t)/N^{\nu(1/\rho-1)}$

Notice that, in contrast to the case of previous subsection, where the tail of the distribution was decaying fast enough, we do not arrive at non-parametric forms of the scaling laws. We can get ride of the scale parameter τ but the results still depend on the parameter ρ and, if $\nu > 1$, also on ν . The reason is that, for fixed number of data, the (sample) moments do not depend on the exponents, and therefore the exponents cannot be isolated from the moments. Indeed, in order to calculate the second sample moment t^2 we can use that t^2 is also power-law

distributed with exponent $1 + \rho/2$ (and a multiplying constant $B/2$) and we can apply in the same way the generalized central-limit theorem to get

$$(53) \quad \sum_{i=1}^N t_i^2 \sim \left[-\Gamma(\rho/2) \frac{B}{2} N \right]^{2/\rho},$$

and from here

$$(54) \quad \tau^2 = B^{2/\rho} \sim \frac{\bar{t}^2}{N^{2/\rho-1}}.$$

For higher order moments the results are analogous, and so we cannot isolate neither the exponent ν nor ρ from here (for a fixed N).

Both forms of the scaling law, for $\nu < 1$ and for $\nu > 1$, can be tested from simulations. We generate double-power-law distributed values of t by the rejection method, first generating a (simple) power law with exponent $1 + \rho$ for $t > 1$ and then rejecting the value of t if it yields a $t^\epsilon/(k + t^\epsilon)$ smaller than a uniform random number u . This leads to a double-power-law distribution, with scale parameter $\tau = k^{1/\epsilon}$ and with exponents $\nu = 1 + \rho - \epsilon$ and $1 + \rho$ for small and large t , respectively (if $\epsilon > 0$, if not, the situation is reversed). Figure ?? shows an example for $\nu > 1$.

4. SUMMARY

We have seen how scaling laws can arise in the waiting time or quiet time between extreme events. We consider two scenarios. The first one deals with point processes, for which events happen instantaneously in comparison with the waiting time between events. This corresponds to the slow time scale of driving in SOC systems. It is shown that scaling might arise as an attractive fixed point of a renormalization transformation, consisting of thinning of events below a threshold in size plus a rescaling of the time axis (so, it is considered that high enough thresholds lead to the fixed-point solution). For processes without correlations, the resulting scaling law contains an exponential scaling function (characteristic of a Poisson process) if the mean waiting time between the original events has a finite mean. In contrast, if the mean waiting time is infinite, a non-trivial scaling function arises, with a double-power-law behavior. Different forms of the scaling law are shown in terms of the probability p of surviving the thinning process, the mean waiting time for the renormalized events, or the distribution of events above size s . The results are critically dependent on whether $\langle t \rangle$ is finite or infinite. As a corollary, a slight change in the renormalization procedure (removing events deterministically instead of randomly) leads to a direct derivation of the law of large numbers and the generalized central-limit theorem (for the case of maximum asymmetry) [18].

The second scenario contemplates time series, and the variable of interest is the quiet time (the time the signal is below a threshold), representing again but in a different way the time between extreme events. In the SOC framework this

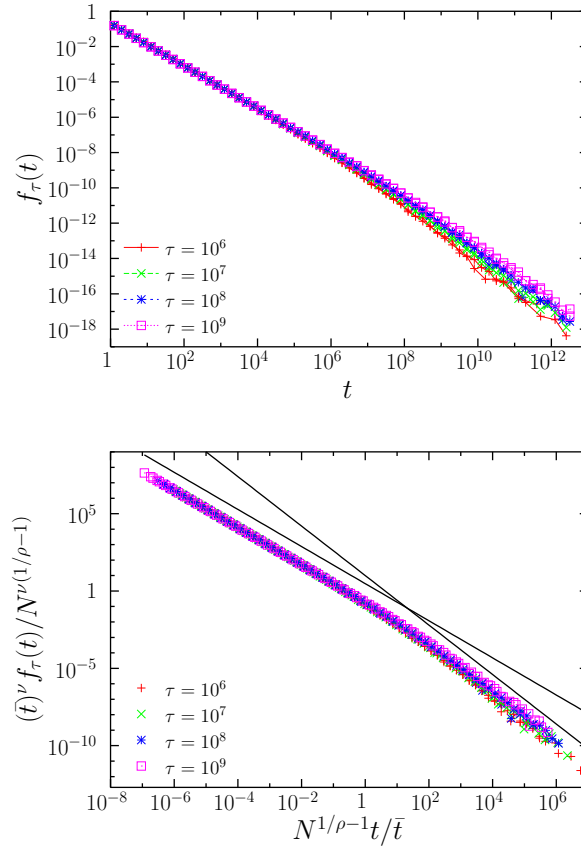


FIGURE 2. (a) Probability densities for simulated double-power-law distributions, with parameters $\nu = 1.2$, $1 + \rho = 1.6$, and $m = 1$, for different values of the scale parameter τ . Two different values of N are used for each τ , these are $N = 10^5$ and $N = 10^6$. (b) Same distributions after a semi-parametric rescaling of the axes, following Eq. (52). The collapse of data indicates the fulfillment of the scaling law. The straight lines are power laws with exponents 1.2 and 1.6 .

would correspond to look at the system in the fast time scale inside the avalanches. In this case scaling laws might show up from the first-passage time of diffusion processes that model the signal, in the limit of high enough thresholds. The scaling law yields scaling functions different than the ones from point processes, with a decaying power law for short times (except for the shortest times) and a much faster decay for long times. Using the dependence of the moments of the distribution on the scale parameter and on the power-law exponent we can replace the latter ones in terms of the former ones and obtain a non-parametric scaling law, depending only on the moments. The particular equation for the

scaling law depends on the value of the power-law exponent; in particular we have distinguished $\nu < 1$ (trivial scaling), $1 < \nu < 2$, and $1 < \nu < 3$. Finally, this result is generalized for the case in which the waiting-time distribution does not decay fast enough for long times but slow enough in such a way that the mean waiting time becomes infinite. No non-parametric scaling is possible then and one obtains a semi-parametric scaling in which the dependence on the scale parameter can be substituted by the dependence on the sample moments, but the exponents cannot be replaced. Additionally, important lessons for the projection of extreme events are derived from the different scaling of the mean of the distribution and the scale of extreme events.

5. ACKNOWLEDGEMENTS

This paper is a re-elaboration of a talk of the author at the EXEV14 workshop; invitation by A. Deluca, H. Kantz and N. W. Watkins is gratefully acknowledged. The author is indebted to N. R. Moloney for pointing to the references relating renormalization and the central limit theorem and to G. Pruessner for his influential lectures at the Joint CRM-Imperial College School in Complex Systems. Research expenses were founded by projects FIS2009-09508, from the disappeared Spanish MICINN, FIS2012-31324 from Spanish MINECO, and 2009SGR-164, from AGAUR. This work is also associated to 2014SGR-1307.

6. APPENDIX I: GENERALIZED CENTRAL LIMIT THEOREM FROM RENORMALIZATION

Remember that in the (generalized-) central-limit-theorem scenario one sums independent identically distributed random variables and then the sum is rescaled in order that the resulting distribution is comparable to the original distribution of the individual variables. We showed in subsec. 2.3 that this is exactly the same problem one finds for the waiting time of extreme events when these happen in a deterministic way, being the $1/p$ -th out of $1/p$ events an extreme event, and the rest, $1/p - 1$ events, being ordinary events, which are removed. In this way, the connection with a renormalization transformation is direct. Introducing $n = 1/p$, and assuming n is a natural number, we know from subsec. 2.3 that the cumulant generating function of the sum and rescaling of these n random variables (in our case waiting times, but we put ourselves here in a more general context) transforms as

$$(55) \quad \mathcal{R}[\ln \tilde{f}(\omega)] = n \ln \tilde{f}\left(\frac{\omega}{n^r}\right),$$

i.e., suffers a simple scale transformation under this operation, where $\ln \tilde{f}(\omega)$ is the cumulant generating function of the original random variables (which is the logarithm of the Laplace transform $\tilde{f}(\omega)$ of the probability density $f(t)$). In subsec. 2.3 we saw the fixed points of this transformation (i.e., the distributions which are stable under renormalization) and now we explore the attractiveness

of these fixed points, which is what provides the content for the central limit theorem.

6.1. Existence of all moments. The simplest case corresponds to the common situation of an initial probability density whose moments exist and are finite (non-infinite). If the cumulant generating function exists, it is well known that its expansion will be

$$(56) \quad \ln \tilde{f}(\omega) = -\langle t \rangle \omega + \mathcal{O}(\omega),$$

where $\mathcal{O}(\omega)$ refers to terms that are of higher order than linear in ω (just expand by Taylor the exponential factor $e^{-\omega t}$ in the Laplace transform, take the logarithm of the resulting series and expand again the logarithm). Applying the renormalization transformation to the expansion of $\ln \tilde{f}(\omega)$ one gets

$$(57) \quad \mathcal{R}[\ln \tilde{f}(\omega)] = n \left(-\langle t \rangle \frac{\omega}{n^r} + \mathcal{O}(\omega/n^r) \right),$$

and imposing linear rescaling, $r = 1$, leads, after successive iterations of the transformation (or, equivalently, after taking the limit $n \rightarrow \infty$) to

$$(58) \quad \mathcal{R}[\cdots \mathcal{R}[\ln \tilde{f}(\omega)] \cdots] \rightarrow -\langle t \rangle \omega,$$

where note that all non-linear terms vanish (we are assuming that $\langle t \rangle > 0$). We already saw that this is the logarithm of the Laplace transform of a Dirac's delta function, and therefore

$$(59) \quad \mathcal{R}[\cdots \mathcal{R}[f(t)] \cdots] \rightarrow \delta(t - \langle t \rangle)$$

(we use the same notation for the renormalization transformation operating over $f(t)$, over its Laplace transform, or over the cumulant generating function). Although we are in the context of the central limit theorem, this trivially-looking result is nothing else than a weak version of the law of large numbers [43], perhaps the most fundamental result in probability theory and statistics, ensuring that the arithmetic mean of a sample converges to the expected value of the distribution (if this expected value exists and is not infinite, and the same for the rest of the moments).

It is noteworthy that the equivalent of the law of large numbers when the number of variables that are added is not fixed but random, following a geometrical distribution, is the exponential distribution (as we saw in subsec. 2.1). On the other hand, if a fixed number of variables are summed but the mean is zero one obtains the classic central-limit result in the form of a Gaussian distribution, which is of no interest here, but see Ref. [37].

6.2. Power laws and power-law tails. The “non-trivial” case corresponds to a distribution $f(t)$ which is a power law, or at least that has a power-law upper tail. In the first case $f(t) = B/t^{1+\rho}$, for $t > m > 0$, with $\rho > 0$ and $B = \rho m^\rho$

(normalization condition). The Laplace transform of the density can be expressed in terms of the incomplete gamma function,

$$(60) \quad \tilde{f}(\omega) = B \int_m^\infty e^{-\omega t} t^{-\rho-1} dt = B \omega^\rho \Gamma(-\rho, m\omega).$$

Using the expansion of this function (see 6.5.29 of Ref. [49]),

$$(61) \quad \Gamma(\gamma, z) = \Gamma(\gamma) - z^\gamma \sum_{n=0}^{\infty} \frac{(-z)^n}{(\gamma+n)n!},$$

valid for $\gamma \neq 0, -1, -2, -3, \dots$, we can write

$$(62) \quad z^\rho \Gamma(-\rho, z) = z^\rho \Gamma(-\rho) - \sum_{n=0}^{\infty} \frac{(-z)^n}{(n-\rho)n!}$$

$$(63) \quad = \frac{1}{\rho} \left[\rho \Gamma(-\rho) z^\rho + \left(1 + \frac{\rho z}{1-\rho} + \mathcal{O}(z) \right) \right],$$

for $\rho \neq 0, 1, 2, \dots$. From here, taking the logarithm, and using its expansion around $z = 1$,

$$(64) \quad \ln [z^\rho \Gamma(-\rho, z)] = -\ln \rho + \rho \Gamma(-\rho) z^\rho + \frac{\rho z}{1-\rho} + \mathcal{O}(z),$$

which we can substitute in the expression of $\ln \tilde{f}(\omega)$, from Eq. (60), so

$$(65) \quad \ln \tilde{f}(\omega) = \ln \frac{B}{m^\rho} + \ln [m^\rho \omega^\rho \Gamma(-\rho, m\omega)]$$

$$(66) \quad = \rho \Gamma(-\rho) m^\rho \omega^\rho + \frac{\rho m \omega}{1-\rho} + \mathcal{O}(\omega),$$

where we have used that $B = \rho m^\rho$. Applying the renormalization transformation we arrive at

$$(67) \quad \mathcal{R}[\ln \tilde{f}(\omega)] = n \ln \tilde{f}\left(\frac{\omega}{n^r}\right)$$

$$(68) \quad = n \left(\rho \Gamma(-\rho) m^\rho \left(\frac{\omega}{n^r}\right)^\rho + \frac{\rho m \omega / n^r}{1-\rho} + \mathcal{O}(\omega / n^r) \right).$$

We can distinguish two situations here. On the one hand, if $\rho > 1$ in order to have a well-defined limiting distribution we need to take $r = 1$, which yields, in the same way as in the trivial case, the delta function corresponding to the law of large numbers,

$$(69) \quad \mathcal{R}[\dots \mathcal{R}[\ln \tilde{f}(\omega)] \dots] \rightarrow -\frac{\rho m \omega}{\rho - 1},$$

so,

$$(70) \quad \mathcal{R}[\dots \mathcal{R}[f(t)] \dots] \rightarrow \delta\left(t - \frac{\rho m}{\rho - 1}\right).$$

Notice that for the power-law distribution with $\rho > 1$ the expected value is $\langle t \rangle = \frac{\rho m}{\rho-1}$, but the other moments can diverge and still the law of large numbers holds in this case.

On the other hand, a different situation corresponds to $\rho < 1$, and in order to have a well-defined limit we need to take $r = 1/\rho$, and then

$$(71) \quad \mathcal{R}[\dots \mathcal{R}[\ln \tilde{f}(\omega)] \dots] \rightarrow \rho \Gamma(-\rho) m^\rho \omega^\rho,$$

which corresponds to a Laplace transform given by $\tilde{f}^*(\omega) = \exp(B\Gamma(-\rho)\omega^\rho)$. This yields Lévy-stable distributions with maximal asymmetry [37]. As far as the author knows, the Laplace transform only can be inverted for $\rho = 1/2$, yielding the sometimes called Lévy-Smirnov distribution, Eq. (27), with $A = -B\Gamma(-1/2) = 2\sqrt{\pi}B$. For general values of ρ (in the range $(0, 1)$) it can be shown that the limiting distribution behaves, for large t as $f^*(t) \propto 1/t^{1+\rho}$. So, when power-law distributed variables, with exponent $1 + 1/r < 2$, are summed, the way to rescale them is not by the number of terms n but by n^r . With this “strange” average we get convergence to a Lévy stable distribution, but not to a Dirac’s delta distribution, so, no generalized form of the law of large numbers holds. Rather, this result is considered to belong to the generalized central-limit theorem, although one is not dealing with the center of the distribution.

In the second “non-trivial” case we have a distribution which asymptotically (for large t) is a power law, $f(t) \sim B/t^{1+\rho}$, with B not necessarily equal to ρm^ρ . Still it is possible to use the expansion of the Laplace transform of $f(t)$, which is

$$(72) \quad \tilde{f}(\omega) = B\Gamma(-\rho)\omega^\rho + \sum_{n=0}^{\infty} \frac{a_n(-\omega)^n}{n!},$$

see 4.6.23 of Ref. [50]. As $a_0 = 1$ then, taking the logarithm,

$$(73) \quad \ln \tilde{f}(\omega) = B\Gamma(-\rho)\omega^\rho - a_1\omega + \mathcal{O}(\omega),$$

and therefore we are in the same situation as for the pure power-law case, Eq. (66). The convergence is towards $f^*(t) = \delta(t - a_1)$ if $\rho > 1$ (so, the law of large numbers holds) and towards $\tilde{f}^*(\omega) = \exp(-A\omega^\rho)$ if $\rho < 1$, with $A = -B\Gamma(-\rho)$ (notice that $A > 0$ and that the scale parameter of $f^*(t)$ is $A^{1/\rho}$). This justifies the non-linear rescaling of the sum of n random variables, so the sum “goes” as $(An)^{1/\rho}$ but it is broadly distributed. In other words, $\sum t_i/n^r$ follows a Lévy-stable distribution (with maximal asymmetry) with scale parameter $[-\Gamma(\rho)B]^{1/\rho}$.

7. APPENDIX II: QUIET TIME OF A BROWNIAN SIGNAL

In order to understand how scaling can show up in the times a time series is below a threshold, let us consider that the signal is given by the position of a Brownian particle diffusing between two absorbing boundaries. The particle starts below but very close to the threshold, which is an absorbing boundary (assuming it had just crossed the threshold from above), and yields a quiet time if

it reaches the threshold; on the contrary, if it reaches the zero-value intensity, the process (the SOC avalanche) dies out and no quiet time is computed. The quiet times generated in this way define a renewal process, in which the distribution of quiet times completely describes the process. A description of quiet times in the most complete framework would involve the use of copulas [51].

The first step is the solution of the diffusion equation, $\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}$, with D the diffusion coefficient. For simplicity we take the x -axis in such a way that the intensity threshold corresponds to $x = 0$, whereas the other absorbing boundary, at zero intensity, corresponds to $x = L$ (in this way, increasing intensity corresponds to decreasing x). In any case, L is the separation between the threshold and the zero-intensity absorbing boundary. It is well known [38, 52] that the solution to this diffusion problem is given by

$$(74) \quad C(x, t) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi x}{L} e^{-\left(\frac{n\pi}{L}\right)^2 D t}$$

with $A_n = \frac{2}{L} \int_0^L C(x, 0) \sin \frac{n\pi x}{L} dx$ and $n = 1, 2, \dots$ etc. The absorbing boundary conditions are represented by $C(0, t) = C(L, t) = 0$.

In this case the quiet time turns out to be just the first-passage at $x = 0$. If the concentration $C(x, t)$ is normalized to 1 in $t = 0$, the probability that the first-passage time is larger than a value t is

$$(75) \quad S(t) = \int_0^L C(x, t) dx,$$

which corresponds to the fraction of Brownian particles that have not left the system, and therefore, the probability density of the first-passage time is

$$(76) \quad f(t) = -\frac{dS(t)}{dt} = -\int_0^L \frac{\partial C(x, t)}{\partial t} dx = J(L, t) - J(0, t)$$

where $J(x, t) = -D \frac{\partial C}{\partial x}$ is the flow of particles. This $f(t)$ is the probability density of the first passage time to any of the two boundaries. The quiet-time distribution will be given by the out-flow of particles at $x = 0$ [38], which is then

$$(77) \quad f_L(t) = -J(0, t) = D \left. \frac{\partial C(x, t)}{\partial x} \right|_{x=0} = \frac{\pi D}{L} \sum_{n=1}^{\infty} n A_n e^{-t/\tau_n}$$

with the characteristic times $\tau_n = \frac{L^2}{\pi^2 D n^2}$. The subindex L in the density indicates the dependence with L , not the position of the boundary. Note that this distribution is not normalized, due to the out-flow of particles at the other boundary (the normalization factor is explained in Ref. [38], but it is not relevant for our purposes).

The behavior of the quiet-time density is obtained immediately in the limit of very large times, by the largest time scale τ_1 , so

$$(78) \quad f_L(t) \simeq \frac{\pi D A_1}{L} e^{-t/\tau_1}$$

for $t \gg \tau_1$, taking into account also the value of the constants A_n , which we will see do not alter this limiting behavior. For this we need a precise initial condition; in our case this is $C(x, 0) = \delta(x - x_0)$, with δ de Dirac's delta function and x_0 the initial position, very close to the boundary at $x = 0$ and fulfilling therefore $x_0 \ll L$. Substituting in A_n ,

$$(79) \quad A_n = \frac{2}{L} \sin \frac{n\pi x_0}{L},$$

from which it is immediate that $nA_n e^{-tn^2/\tau_1}$ for $n = 1$ is greater than any other term if $t \gg \tau_1$. The fact that, in this limit, the term with $n = 1$ is greater than the sum of the rest is made clear below.

Coming back to the quiet-time density for all times, we get, substituting the expression for A_n given by the initial condition,

$$(80) \quad f_L(t) = \frac{2\pi D}{L^2} \sum_{n=1}^{\infty} n \sin \frac{n\pi x_0}{L} e^{-n^2 t/\tau_1} \simeq$$

$$(81) \quad \simeq \frac{2}{\pi \tau_1} \int_1^{\infty} dn n \sin(n\pi \ell) e^{-n^2 t/\tau_1},$$

where we have approximated the sum by an integral and have introduced the rescaled distance of the initial position to the threshold, $\ell = x_0/L$, which verifies $\ell \ll 1$.

As functions of n , the scale of the sinus is given by $1/\ell$ and the scale of the Gaussian factor by $\sqrt{\tau_1/t}$. Then, at intermediate times, large enough such that $\sqrt{\tau_1/t} \ll 1/\ell$ but small enough to keep $t \ll \tau_1$ the variation of the Gaussian factor is much faster than that of the sinus and we can approximate the later as $\sin(n\pi \ell) \simeq n\pi \ell$, and the resulting integral can be straightforwardly solved, yielding,

$$(82) \quad f_L(t) \simeq \frac{2\ell}{\tau_1} \int_1^{\infty} dn n^2 e^{-n^2 t/\tau_1} =$$

$$(83) \quad = \frac{\ell}{\tau_1} \left(\frac{\tau_1}{t} \right)^{3/2} \int_{t/\tau_1 \simeq 0}^{\infty} du \sqrt{u} e^{-u} = \frac{\sqrt{\pi} \ell}{2\tau_1} \left(\frac{\tau_1}{t} \right)^{3/2}$$

with the change $n = \sqrt{\tau_1 u/t}$ and recognizing the gamma function of $3/2$, which is $\Gamma(3/2) = \sqrt{\pi}/2$. We repeat that this is valid in the range $\ell^2 \tau_1 \ll t \ll \tau_1$.

In the same way, we can go back to the case $t \gg \tau_1$ and compare the first term, $n = 1$, with the sum of the rest, to see that

$$(84) \quad \frac{2\ell}{\tau_1} e^{-t/\tau_1} \gg \frac{\ell}{\tau_1} \left(\frac{\tau_1}{t}\right)^{3/2} \Gamma(3/2, 2^2 t/\tau_1) \sim \frac{2\ell}{t} e^{-4t/\tau_1},$$

using the incomplete gamma function, $\Gamma(\gamma, z) = \int_z^\infty y^{\gamma-1} e^{-y} dy$ and its asymptotic behavior $\Gamma(\gamma, z) \sim z^{\gamma-1} e^{-z}$ when $z \rightarrow \infty$. Finally, at very short times $\sqrt{t/\tau_1} \ll \ell$, the variation of the sinus is very fast compared with the Gaussian factor and then the integral vanishes, approximately, so

$$(85) \quad f_L(t) \simeq \frac{1}{\tau_1} \int_1^\infty dn \sin(n\pi\ell) \cdot \text{constant} \simeq 0.$$

Summarizing,

$$(86) \quad f_L(t) \simeq \begin{cases} 0 & \text{for } t \ll \ell^2 \tau_1 = \frac{x_0^2}{\pi^2 D} \\ \frac{\sqrt{\pi}\ell}{2\tau_1} \left(\frac{\tau_1}{t}\right)^{3/2} & \text{for } \frac{x_0^2}{\pi^2 D} \ll t \ll \tau_1 = \frac{L^2}{\pi^2 D} \\ \frac{2\ell}{\tau_1} e^{-t/\tau_1} & \text{for } t \gg \tau_1 \end{cases}$$

where we can identify a simple scaling form. Nevertheless, as $\ell = x_0/L = x_0/\sqrt{\pi^2 D \tau_1} = \sqrt{m/\tau_1}$, with $m = x_0^2/(\pi^2 D)$ the minimum waiting time (below it, the density is zero), the scaling law turns out to be

$$(87) \quad f_L(t) \simeq \sqrt{\frac{m}{\tau_1^3}} G\left(\frac{t}{\tau_1}\right),$$

for $t \gg m$, with m fixed, or in terms of the threshold value L ,

$$(88) \quad f_L(t) \simeq \frac{\sqrt{D^3 m}}{L^3} G\left(\frac{Dt}{L^2}\right).$$

where the scaling function G absorbs missing multiplicative constants. The scaling function in Eq. (86) can be further approximated to give, in this particular case,

$$(89) \quad f_L(t) \simeq \frac{\sqrt{\pi}/4 + (t/\tau_1)^{3/2}}{(t/\tau_1)^{3/2}} \left(2\sqrt{\frac{m}{\tau_1^3}}\right) e^{-t/\tau_1},$$

for $t > m$. This function has the right asymptotic behavior for all temporal ranges, although other “fits” are possible. This simple model justifies the scaling ansatz for the quiet time distributions.

REFERENCES

- [1] P. Bak. *How Nature Works: The Science of Self-Organized Criticality*. Copernicus, New York, 1996.
- [2] G. Pruessner. *Self-Organised Criticality: Theory, Models and Characterisation*. Cambridge University Press, Cambridge, 2012.

- [3] B. D. Malamud. Tails of natural hazards. *Phys. World*, 17 (8):31–35, 2004.
- [4] O. Peters, A. Deluca, A. Corral, J. D. Neelin, and C. E. Holloway. Universality of rain event size distributions. *J. Stat. Mech.*, P11030, 2010.
- [5] A. Corral, A. Ossó, and J. E. Llebot. Scaling of tropical-cyclone dissipation. *Nature Phys.*, 6:693–696, 2010.
- [6] K. Christensen and N. R. Moloney. *Complexity and Criticality*. Imperial College Press, London, 2005.
- [7] A. Corral. Scaling and universality in the dynamics of seismic occurrence and beyond. In A. Carpinteri and G. Lacidogna, editors, *Acoustic Emission and Critical Phenomena*, pages 225–244. Taylor and Francis, London, 2008.
- [8] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer, London, 2001.
- [9] G. Boffetta, V. Carbone, P. Giuliani, P. Veltri, and A. Vulpiani. Power laws in solar flares: Self-organized criticality or turbulence? *Phys. Rev. Lett.*, 83:4662–4665, 1999.
- [10] R. Sánchez, D.E. Newman, and B. A. Carreras. Waiting-time statistics of self-organized-criticality systems. *Phys. Rev. Lett.*, 88:068302, 2002.
- [11] J. Davidsen and Paczuski M. $1/f^\alpha$ noise from correlations between avalanches in self-organized criticality. *Phys. Rev. E*, 66:050101(R), 2002.
- [12] A. Corral. Renormalization-group transformations and correlations of seismicity. *Phys. Rev. Lett.*, 95:028501, 2005.
- [13] A. Corral. Point-occurrence self-similarity in crackling-noise systems and in other complex systems. *J. Stat. Mech.*, P01022, 2009.
- [14] M. Paczuski, S. Boettcher, and M. Baiesi. Interoccurrence times in the Bak-Tang-Wiesenfeld sandpile model: A comparison with the observed statistics of solar flares. *Phys. Rev. Lett.*, 95:181102, 2005.
- [15] D. R. Cox and V. Isham. *Point Processes*. Chapman and Hall, London, 1980.
- [16] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York, 1988.
- [17] S. B. Lowen and M. C. Teich. *Fractal-Based Point Processes*. Wiley-Interscience, New Jersey, 2005.
- [18] B. V. Gnedenko and V. Y. Korolev. *Random Summation: Limit Theorems and Applications*. CRC Press, Boca Raton, 1996.
- [19] H. Kanamori and E. E. Brodsky. The physics of earthquakes. *Rep. Prog. Phys.*, 67:1429–1496, 2004.
- [20] A. Corral. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Phys. Rev. Lett.*, 92:108501, 2004.
- [21] A. Saichev and D. Sornette. “Universal” distribution of interearthquake times explained. *Phys. Rev. Lett.*, 97:078501, 2006.
- [22] A. Corral. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics*, 424:177–193, 2006.
- [23] J. Davidsen, S. Stanchits, and G. Dresen. Scaling and universality in rock fracture. *Phys. Rev. Lett.*, 98:125502, 2007.
- [24] J. Baró, A. Corral, X. Illa, A. Planes, E. K. H. Salje, W. Schranz, D. E. Soto-Parra, and E. Vives. Statistical similarity between the compression of a porous material and earthquakes. *Phys. Rev. Lett.*, 110:088702, 2013.
- [25] A. Corral, L. Telesca, and R. Lasaponara. Scaling and correlations in the dynamics of forest-fire occurrence. *Phys. Rev. E*, 77:016101, 2008.
- [26] U. Harder and M. Paczuski. Correlated dynamics in human printing behavior. *Physica A*, 361:329–336, 2006.

- [27] P. Bak, K. Christensen, L. Danon, and T. Scanlon. Unified scaling law for earthquakes. *Phys. Rev. Lett.*, 88:178501, 2002.
- [28] A. Corral. Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Phys. Rev. E*, 68:035102, 2003.
- [29] A. Corral. Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A*, 340:590–597, 2004.
- [30] J. Åström, P. C. F. Di Stefano, F. Pröbst, L. Stodolsky, J. Timonen, C. Bucci, S. Cooper, C. Cozzini, F.v. Feilitzsch, H. Kraus, J. Marchese, O. Meier, U. Nagel, Y. Ramachers, W. Seidel, M. Sisti, S. Uchaikin, and L. Zerle. Fracture processes observed with a cryogenic detector. *Phys. Lett. A*, 356:262–266, 2006.
- [31] E. L. Geist and T. Parsons. Distribution of tsunami interevent times. *Geophys. Res. Lett.*, 35:L02612, 2008.
- [32] A. Corral, R. Ferrer-i-Cancho, and A. Díaz-Guilera. Universal complex structures in written language. <http://arxiv.org>, 0901.2924, 2009.
- [33] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin. Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records. *Phys. Rev. Lett.*, 94:048701, 2005.
- [34] M. Baiesi, M. Paczuski, and A. L. Stella. Intensity thresholds and the statistics of the temporal occurrence of solar flares. *Phys. Rev. Lett.*, 96:051103, 2006.
- [35] L. Laurson, X. Illa, and M. J. Alava. The effect of thresholding on temporal avalanche statistics. *J. Stat. Mech.*, P01019, 2009.
- [36] K. Yamasaki, L. Muchnik, S. Havlin, A. Bunde, and H. E. Stanley. Scaling and memory in volatility return intervals in financial markets. *Proc. Natl. Acad. Sci. USA*, 102:9424–9428, 2005.
- [37] J.-P. Bouchaud and A. Georges. Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications. *Phys. Rep.*, 195:127–293, 1990.
- [38] S. Redner. *A Guide to First-Passage Processes*. Cambridge University Press, Cambridge, 2007.
- [39] G. Jona-Lasinio. Renormalization group and probability theory. *Phys. Rep.*, 352:439–458, 2001.
- [40] D. Sornette. *Critical Phenomena in Natural Sciences*. Springer, Berlin, 2nd edition, 2004.
- [41] J. P. Sethna. *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press, Cambridge, 2006.
- [42] I. Calvo, J. C. Cuchí, J. G. Esteve, and F. Falceto. Generalized central limit theorem and renormalization group. *J. Stat. Phys.*, 141:409–421, 2010.
- [43] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, New York, 2nd edition, 1971.
- [44] K. Christensen, N. Farid, G. Pruessner, and M. Stapleton. On the scaling of probability density functions with apparent power-law exponents less than unity. *Eur. Phys. J. B*, 62:331–336, 2008.
- [45] A. Deluca and A. Corral. Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophys.*, 61:1351–1394, 2013.
- [46] A. Rosso, P. Le Doussal, and K. J. Wiese. Avalanche-size distribution at the depinning transition: A numerical test of the theory. *Phys. Rev. B*, 80:144204, 2009.
- [47] G. Pruessner. Comment on “Avalanches and non-Gaussian fluctuations of the global velocity of imbibition fronts”. *Phys. Rev. Lett.*, 105:029401, 2010.
- [48] E. C. de Perez, F. Monasso, M. van Aalst, and P. Suarez. Science to prevent disasters. *Nature Geosci.*, 7:78–79, 2014.
- [49] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions*. Dover, New York, 1965.

- [50] N. Bleistein and R. A. Handelsman. *Asymptotic Expansions of Integrals*. Dover, New York, 1986.
- [51] R. Chicheportiche and A. Chakraborti. Copulas and time series with long-ranged dependences. *Phys. Rev. E*, 89:042117, 2014.
- [52] J. Crank. *The Mathematics of Diffusion*. Clarendon Press, Oxford, second edition, 1975.

CENTRE DE RECERCA MATEMÀTICA
CAMPUS DE BELLATERRA, EDIFICI C
E-08193 BARCELONA, SPAIN
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT AUTÒNOMA DE BARCELONA
E-08193 BARCELONA, SPAIN
E-mail address: `acorral@crm.es`

