




Universality of power-law exponents by means of maximum-likelihood estimation

Víctor Navas-Portella ^{1,2,3} Álvaro González ^{1,4} Isabel Serra,^{1,5} Eduard Vives ^{6,7} and Álvaro Corral^{1,2,8,9}

¹*Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Bellaterra, Catalonia, Spain*

²*Barcelona Graduate School of Mathematics, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain*

³*Facultat de Matemàtiques i Informàtica, Universitat de Barcelona, E-08007 Barcelona, Spain*

⁴*GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany*

⁵*Barcelona Supercomputing Center, Nexus II, 08034 Barcelona, Spain*

⁶*Departament de Matèria Condensada, Facultat de Física, Universitat de Barcelona, 08028 Barcelona, Catalonia, Spain*

⁷*Universitat de Barcelona Institute of Complex Systems (UBICS), Facultat de Física, Universitat de Barcelona, E-08028 Barcelona, Catalonia, Spain*

⁸*Complexity Science Hub Vienna, 1080 Vienna, Austria*

⁹*Departament de Matemàtiques, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain*



(Received 30 July 2019; published 3 December 2019)

Power-law-type distributions are extensively found when studying the behavior of many complex systems. However, due to limitations in data acquisition, empirical datasets often only cover a narrow range of observation, making it difficult to establish power-law behavior unambiguously. In this work we present a statistical procedure to merge different datasets, with two different aims. First, we obtain a broader fitting range for the statistics of different experiments or observations of the same system. Second, we establish whether two or more different systems may belong to the same universality class. By means of maximum likelihood estimation, this methodology provides rigorous statistical information to discern whether power-law exponents characterizing different datasets can be considered equal among them or not. This procedure is applied to the Gutenberg-Richter law for earthquakes and for synthetic earthquakes (acoustic emission events) generated in the laboratory: labquakes. Different earthquake catalogs have been merged finding a Gutenberg-Richter law holding for more than eight orders of magnitude in seismic moment. The value of the exponent of the energy distribution of labquakes depends on the material used in the compression experiments. By means of the procedure proposed in this manuscript, we find that the Gutenberg-Richter law for earthquakes and charcoal labquakes can be characterized by the same power-law exponent, whereas Vycor labquakes exhibit a significantly different exponent.

DOI: [10.1103/PhysRevE.100.062106](https://doi.org/10.1103/PhysRevE.100.062106)

I. INTRODUCTION

Generally speaking, a complex system can be understood as a large number of interacting elements whose global behavior cannot be derived from the local laws that characterize each of its components. The global response of the system can be observed at different scales and the vast number of degrees of freedom makes prediction very difficult. In this context, a probabilistic description of the phenomenon is needed in order to reasonably characterize it in terms of random variables. When the response of these systems exhibits lack of characteristic scales, it can be described in terms of power-law-type probability density functions (PDF), $f(x) \propto x^{-\gamma}$, where x corresponds to the values that the random variable that characterizes the response of the system can take, \propto denotes proportionality and the power-law exponent γ acquires values larger than 1. The power law is the only function which is invariant under any scale transformation of the variable x [1]. This property of scale invariance confers a description of the response of the system where there are no characteristic scales. This common framework is very usual in different disciplines [2,3] such as condensed matter physics [4], economics [5], linguistics [6], geoscience [7], and, in particular, seismology [8,9].

It has been broadly studied [10,11] that when different complex systems present common values of all their power-law exponents and share the same scaling functions they can be grouped into the same universality class. Therefore, it is important to determine these exponents rigorously, not only to properly characterize phenomena but also to provide a good classification into universality classes.

In practice, exponents are difficult to measure empirically. Due to experimental limitations that distort the power-law behavior, the property of scale invariance can only be measured in a limited range. Therefore, when a power-law distribution is fitted to empirical data is more convenient to talk about local or restricted scale invariance. In this context, the wider the range the fitted PDF spans, the more reliable and strong this property will be.

A paradigmatic example of power-law behavior in complex systems is the well-known Gutenberg-Richter (GR) law for earthquakes [12]. This law states that, above a lower cut-off value, earthquake magnitudes follow an exponential distribution; in terms of the magnitude PDF

$$f(m) = (b \log 10) 10^{-b(m-m_{\min})} \propto 10^{-bm}, \quad (1)$$

defined for $m \geq m_{\min}$, with m the magnitude (moment magnitude in our case), m_{\min} the lower cut-off in magnitude, b

is the so called b value, and log corresponds to the natural logarithm. The general relationship between seismic moment x and moment magnitude m is given by:

$$x = 10^{\frac{3}{2}m+9.1}, \quad (2)$$

measured in units of Nm [13]. The GR law is a power-law distribution when it is written as a function of the seismic moment x :

$$f(x) = \frac{2}{3} \frac{b}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-(1+\frac{2}{3}b)} = \frac{\gamma-1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\gamma}, \quad (3)$$

where we conveniently define $\gamma = 1 + \frac{2}{3}b$ ($b > 0$ and $\gamma > 1$) and x_{\min} corresponds to the value of the seismic moment of the cut-off magnitude m_{\min} [14] introduced in Eq. (2). Note that this PDF has a domain $x \in [x_{\min}, +\infty)$.

An earthquake catalog is an empirical dataset that characterizes each earthquake by an array of observations: time of occurrence, spatial coordinates, magnitude, and so on. The magnitude m_{\min} is usually associated to the completeness threshold, such as all earthquakes with $m \geq m_{\min}$ are recorded in the catalog [14]. For $m < m_{\min}$, some events are missing from the catalog due to the difficulties of detecting them (e.g., Refs. [15,16]), especially when the seismic activity rate suddenly increases, such as in aftershock sequences, in which earthquake waveforms tend to overlap each other difficulting detection [14,17–19]. This incompleteness distorts the power-law behavior below m_{\min} , whose value should be an upper bound to encompass these variations.

One has to keep in mind that there also exists an upper cut-off due to the finite-size effects [20], implying that, at a certain value of the magnitude, there are deviations from the power-law behavior. Consequently, strictly speaking, the range of validity of the GR law cannot be extended up to infinity [7,21]. By ignoring which is the model that conveniently fits the tail of the distribution, the power-law behavior has to be restricted to an upper cut-off x_{\max} and the PDF for the truncated GR law is written:

$$f(x) = \frac{1-\gamma}{x_{\max}^{1-\gamma} - x_{\min}^{1-\gamma}} x^{-\gamma}, \quad (4)$$

defined for $x \in [x_{\min}, x_{\max}]$.

Recent studies regarding the acoustic emission (AE) in compression experiments of porous glasses and minerals [22–26] or wood [27] have focused the attention on the energy distribution of AE events due to the similarities with the GR law for earthquakes [21]. According to the terminology which is used in some of these studies, we will name as labquakes those AE events that occur during the compression of materials.

Earthquake and labquake catalogs as well as other empirical datasets in complex systems only report a limited range of events, making it difficult to estimate parameters of the power-law PDF accurately. In this work we try to solve this problem by combining datasets with rigorous statistical tools, with the goal of finding a broader range of validity when the different datasets correspond to the same system. If different datasets can be combined and characterized by a unique power-law exponent, that means that the particular exponents of each dataset are statistically compatible. When different

phenomena share the same power-law exponents for the distributions of all their observables, they can be classified into the same universality class. Consequently, this methodology represents a statistical technique to discern whether different phenomena can be classified into the same universality class or not. In order to conveniently classify earthquakes and labquakes into the same universality class, all the observables should be taken into account and all the corresponding exponents should be compatible. In this work we will illustrate this methodology by focusing on the distribution of seismic moment for earthquakes and the AE energy for labquakes. The results will reveal whether they are candidates to be classified into the same universality class or not, but the final establishment of their belonging to the same universality class would require the study of all the possible observables.

The manuscript is structured as follows: in Sec. II we will describe the different statistical procedures that are used in order to conveniently merge datasets and to find a global power-law PDF. Although previous procedures to combine independent datasets have been already proposed [28], the approach presented in this paper is totally different. In Sec. III we apply this methodology to different datasets that are explained in Sec. III A and analyzed in Sec. III B for earthquake catalogs and in Secs. III C and III D for AE data obtained during the compression of two different materials. In Sec. IV we present our conclusions, and details supporting the methodology are presented in Appendixes A–D.

II. METHODS

Statistical methodology: Merging datasets

By considering n_{ds} datasets of N_i ($i = 1, \dots, n_{\text{ds}}$) observations each, one wants to fit a general power-law distribution with a unique global exponent for all of them. We assume that for the i th dataset, the variable \mathcal{X} (seismic moment if one works with the GR law for earthquakes or AE energy if one works with the GR law for labquakes) follows a power-law PDF $f_{\mathcal{X}}[x; \gamma_i, x_{\min}^{(i)}, x_{\max}^{(i)}]$ given by Eq. (4) from a certain lower cut-off $x_{\min}^{(i)}$ to an upper cut-off $x_{\max}^{(i)}$ with exponent γ_i and number of data n_i ($n_i \leq N_i$) in the range $[x_{\min}^{(i)}, x_{\max}^{(i)}]$. Note that one also can consider the untruncated power-law model for the i th dataset if $x_{\max}^{(i)} \rightarrow \infty$. By means of the methods explained in Refs. [7,29] (or, alternatively, Ref. [30]), one can state that data from the i th dataset does not lead to the rejection of the power-law hypothesis for a certain range. Note that, in the i th dataset, the variable \mathcal{X} can acquire values in a range typically spanning several orders of magnitude.

Generally, the procedure of merging datasets is performed by selecting upper and lower cut-offs $x_{\min}^{(i)}$ and $x_{\max}^{(i)}$ [$x_{\min}^{(i)} < x_{\max}^{(i)}$] for each dataset. Data outside these ranges are not considered. All the possible combinations of cut-offs $\{x_{\min}\}$ and $\{x_{\max}\}$ are checked with a fixed resolution (see below). The residual coefficient of variation (CV) test can be used to fix some upper cut-offs, thus reducing the computational effort. For more details about the CV test, see Appendix B.

Given a set of cut-offs, datasets can be merged by considering two different models:

(i) Model α : All datasets are merged by considering a unique global exponent Γ ($\gamma_i = \Gamma$ for all datasets).

(ii) Model β : All datasets are merged, but each one with its own exponent γ_i ($i = 1, \dots, n_{\text{ds}}$).

Note that model α is nested in model β and the difference in the number of parameters characterizing these models is $n_\beta - n_\alpha = n_{\text{ds}} - 1$. Since we are interested in merging datasets with a unique global exponent (model α), we need enough statistical evidence that this simpler model is suitable to fit the data. For a given set of cut-offs, the fit is performed by means of the following protocol:

(i) Maximum likelihood estimation (MLE) of model α : The log-likelihood function of model α can be written as:

$$\log \mathcal{L}_\alpha = \sum_{i=1}^{n_{\text{ds}}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}}[x_{ij}; \Gamma, x_{\min}^{(i)}, x_{\max}^{(i)}], \quad (5)$$

where x_{ij} corresponds to the n_i values of the variable \mathcal{X} that are in the range $x_{\min}^{(i)} \leq x_{ij} \leq x_{\max}^{(i)}$ in the i th dataset and Γ is the global exponent. The definition of likelihood is consistent with the fact that likelihoods from different datasets can be combined in this way [31]. At this step, one has to find the value $\hat{\Gamma}$ of the global exponent Γ that maximizes the log-likelihood expression in Eq. (5). For the particular expressions corresponding to the truncated and untruncated power-law PDF, see Eqs. (A2) and (A3) in Appendix A. If all the power-law distributions are untruncated, then this exponent can be easily found analytically [32] as

$$\hat{\Gamma} = 1 + \frac{\sum_{i=1}^{n_{\text{ds}}} n_i}{\sum_{i=1}^{n_{\text{ds}}} \frac{n_i}{\hat{\gamma}_i - 1}},$$

where the hats denote the values of the exponents that maximize the log-likelihood of the particular power-law distribution (model β) and the general one in Eq. (5). If truncated power-law distributions are considered, one has to use a numerical method in order to determine the exponent Γ that maximizes this expression [29,33].

(ii) MLE of model β : The log-likelihood function of model β can be written as:

$$\log \mathcal{L}_\beta = \sum_{i=1}^{n_{\text{ds}}} \sum_{j=1}^{n_i} \log f_{\mathcal{X}}[x_{ij}; \gamma_i, x_{\min}^{(i)}, x_{\max}^{(i)}], \quad (6)$$

using the same notation as in Eq. (5). For the particular expressions corresponding to the truncated and untruncated power-law PDFs, see Eqs. (A2) and (A3) in Appendix A. The values of the exponents that maximize Eq. (6) are denoted as $\hat{\gamma}_i$.

(iii) Likelihood-ratio test: We perform the likelihood ratio test (LRT) for the models α and β in order to check whether model α is good enough to fit data or not in comparison with model β . For more details about the LRT see Appendix A. If model α “wins,” then the procedure goes to step (iv). Otherwise, this fit is discarded and a different set of cut-offs $\{x_{\min}\}$ and $\{x_{\max}\}$ is chosen, and the procedure goes back to step (i). Note that model α can be a good model to fit if the particular exponents $\hat{\gamma}_i$ do not exhibit large differences among each other in relation to their uncertainty.

(iv) Goodness-of-fit test: In order to check whether it is reasonable to consider model α as a good candidate to fit data, we formulate the next null hypothesis H_0 : The variable \mathcal{X} is power-law distributed with the global exponent

$\hat{\Gamma}$ for all the datasets. In this work, we are going to use two different statistics in order to carry out the goodness-of-fit tests: the Kolmogorov-Smirnov distance of the merged datasets (KSDMD) and the composite Kolmogorov-Smirnov distance (CKSD). The KSDMD statistic can be used as long as datasets overlap each other, whereas the CKSD statistic does not require this condition. For more details about how these statistics are defined and how the p value of the test is found, see Appendix C. If the resulting p value is greater than a threshold value p_c (in the present work we are going to use $p_c = 0.05$ and $p_c = 0.20$), we consider this as a valid fit and it can be stated that the variable \mathcal{X} is power-law distributed with exponent $\hat{\Gamma}$ along all the different datasets for the different ranges $\{x_{\min}\}$ and $\{x_{\max}\}$. Otherwise, this fit will not be considered as valid, a different set of cut-offs is chosen and the procedure goes back to step (i).

When all the combinations of cut-offs have been checked, one may have a list of valid fits. In order to determine which one of them is considered the definitive fit, the following procedure is carried out:

(i) The fit that covers the largest sum of orders of magnitude: $\max \{ \sum_{i=1}^{n_{\text{ds}}} \log_{10} [\frac{x_{\max}^{(i)}}{x_{\min}^{(i)}}] \}$ is chosen. If the power-law fit is untruncated, then $x_{\max}^{(i)}$ can be substituted by the maximum observed value $x_{\text{top}}^{(i)}$. If there is a unique candidate with a maximum number of orders of magnitude, then this is considered as the definitive global fit; otherwise, the procedure goes to the next step (ii).

(ii) The fit with the broadest global range: $\max \{ \frac{\max [x_{\max}^{(i)}]}{\min [x_{\min}^{(i)}]} \}$ for $i = 1, \dots, n_{\text{ds}}$ is chosen. If there is a unique candidate, then this is considered as the definitive global fit; otherwise, the procedure goes to the next step (iii).

(iii) The fit with the maximum number of data $\mathcal{N} = \sum_{i=1}^{n_{\text{ds}}} n_i$ is considered as the definitive global fit.

By means of these three steps, a unique fit has been found for all the datasets analyzed in this work. Nevertheless, one could deal with datasets in which more conditions are needed in order to choose a definitive fit unambiguously.

At the end of this procedure, if a solution is found, then one is able to state that the datasets that conform the global fit correspond to phenomena that are candidates to be classified into the same universality class, at least regarding the observable \mathcal{X} . If no combination of cut-offs is found to give a good fit, then it can be said that there exists at least one catalog that corresponds to a phenomenon that must be classified in a different universality class. The performance of the methodology over synthetic power-law data with different sample sizes and relative difference in exponents δ_γ is shown in Appendix D.

III. APPLICATIONS

In this section we apply the methodology of merging datasets to different earthquake and labquake catalogs. First, we try to merge three earthquake catalogs. Second, a fourth catalog of charcoal labquakes is added in order to check whether these two phenomena are candidates to be classified into the same universality class or not. Finally, we apply the methodology to four catalogs of Vycor labquakes that cover different observation windows.

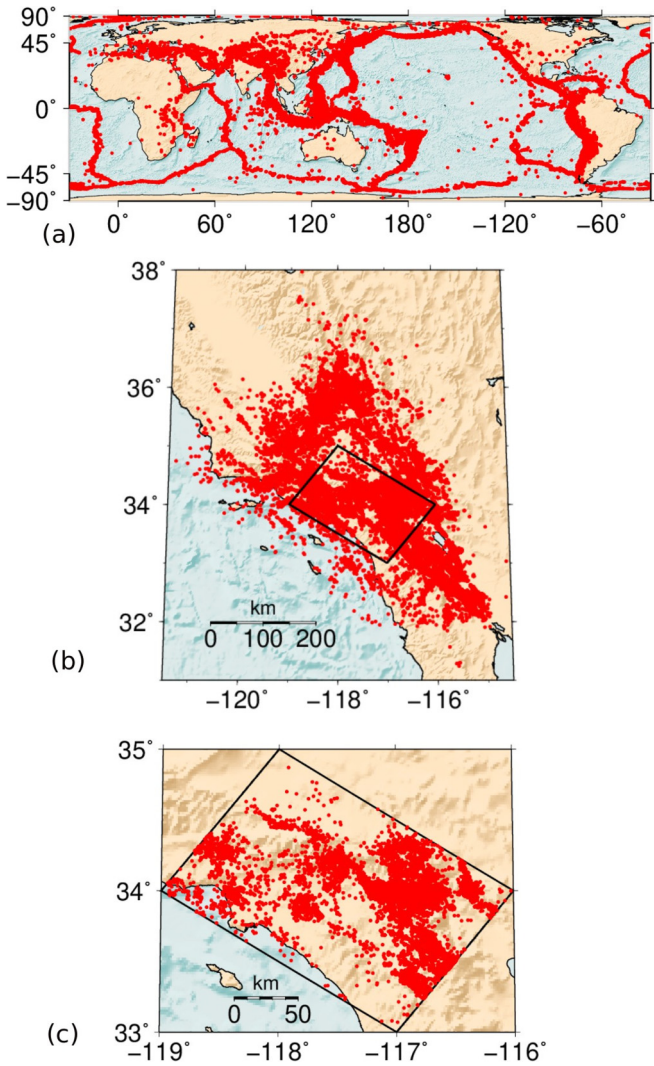


FIG. 1. Earthquake epicenters of the different catalogs. Horizontal and vertical axes correspond to longitude and latitude in degrees, respectively. (a) Full CMT catalog for the period 1977–2017 [34,35]. (b) Entire YHS A Catalog [36] of Southern California for the period 1981–2010. (c) Subcatalog corresponding to a region (YHS B) of the YHS catalog for the Los Angeles area for the period 2000–2010. Maps in cylindrical equal-area projection (a) and sinusoidal projection (b) and (c), produced with Generic Mapping Tools [37].

A. Datasets

1. Earthquake catalogs

We have selected catalogs that have different completeness magnitudes in order to cover different magnitude ranges. We hope that a convenient combination of these datasets will give us a larger range of validity of the GR law with a unique exponent. Let us briefly describe the catalogs that have been used in this work (see also Fig. 1):

(a) Global Centroid Moment Tensor (CMT) Catalog: This catalog compiles earthquakes worldwide since 1977 [34,35], from which we analyze the dataset until the end of 2017. This catalog reports the values of the moment magnitude as well as the seismic moment. Given that the seismic moment is provided with three significant digits, the resolution of the magnitude in the catalog is approximately $\Delta m \simeq 10^{-3}$.

(b) Yang-Hauksson-Shearer (YHS) A: It records earthquakes in Southern California with $m \geq 0$ in the period 1981–2010 [36]. This catalog does not report the seismic moment but a preferred magnitude that is approximately converted into seismic moment according to Eq. (2). The resolution of the catalog is $\Delta m = 0.01$.

(c) Yang-Hauksson-Shearer (YHS) B: It is a subset of YHS A that contains the earthquakes in the region of Los Angeles in the period 2000–2010 [36]. This LA region is defined by the following four vertices in longitude and latitude: $(119^\circ \text{ W}, 34^\circ \text{ N})$, $(118^\circ \text{ W}, 35^\circ \text{ N})$, $(116^\circ \text{ W}, 34^\circ \text{ N})$, and $(117^\circ \text{ W}, 33^\circ \text{ N})$ (see Fig. 1). This region has been selected because it is among the best monitored ones [15,38]. Furthermore, we selected this time period due to the better detection of smaller earthquakes than in previous years [38], which should reduce the completeness magnitude of the catalog [15]. The resolution of this catalog is the same as YHS A.

Figure 1 shows the epicentral locations of the earthquakes contained in each catalog. In the statistical analysis, in order to not to count the same earthquake more than once, the spatio-temporal window corresponding to the YHS B catalog has been excluded from the YHS A and the spatial [Fig. 1(b)] and temporal window corresponding to the full YSH catalogue (A+B) has been excluded from the CMT catalog.

2. Labquake catalogs

We performed uniaxial compression experiments of porous materials in a conventional test machine Z005 (Zwick/Roell). Samples with no lateral confinement were placed between two plates that approached each other at a certain constant displacement rate \dot{z} . Simultaneous to the compression, recording of an AE signal was performed by using a piezoelectric transducer embedded in one of the compression plates. The electric signal $U(t)$ was pre-amplified, band filtered (between 20 kHz and 2 MHz), and analyzed by means of a PCI-2 acquisition system from Euro Physical Acoustics (Mistras Group) with an AD card working at 40 megasamples per second with 18-bit precision [39]. Signal preamplification was necessary to record small AE events. Some values of the preamplified signal were so large that could not be detected correctly by the acquisition system. This fact led to signal saturation and, consequently, to an underestimated energy of the AE event [32]. Recording of data stopped when a big failure event occurred and the sample got destroyed.

An AE event (often called AE hit in specialized AE literature) starts at the time t_j when the signal $U(t)$ exceeds a fixed detection threshold and finishes at time $t_j + \tau_j$ when the signal remains below threshold from $t_j + \tau_j$ to at least $t_j + \tau_j + 200 \mu\text{s}$. The energy E_j of each event is determined as $E_j = \frac{1}{R} \int_{t_j}^{t_j + \tau_j} U^2(t) dt$, where R is a reference resistance of 10 k Ω . This AE energy corresponds to the radiated energy received by the transducer. At the end of an experiment, a catalog of events is collected, each of them characterized by a time of occurrence t , energy E , and duration τ .

Experiments were performed with two different materials:

(i) Charcoal: One experiment with a prismatic charcoal sample (basis of 8.47 mm \times 5.96 mm and height $H = 14.22$ mm) was performed at a constant rate $\dot{z} = 0.005$ mm/min with a preamplification of 40 dB and a value

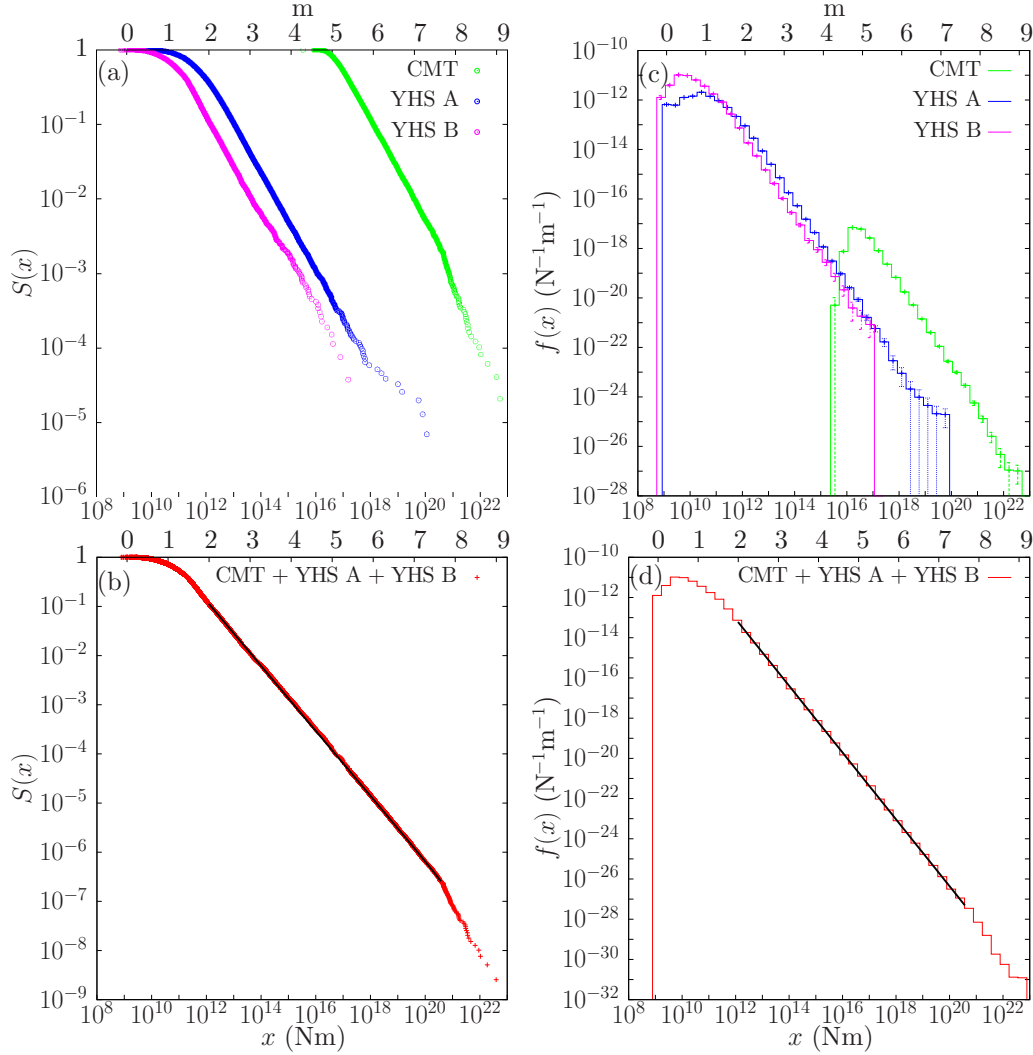


FIG. 2. Estimated survivor functions $S(x)$ of the seismic moment for each catalog (a) and for the merged catalogs (b). Estimated PDFs $f(x)$ of the Gutenberg-Richter law for each catalog (c) and for the merged catalogs (d). The merged histograms are plotted by following the procedure explained in Ref. [32]. Fits are represented by solid black lines. Top axes represent the same scale in moment magnitude. Note that earthquakes from the CMT and YSH A catalogs below their respective lower cut-offs x_{\min} have been excluded in the bottom plots.

of 24 dB for the detection threshold referring to the signal $U(t)$, not the preamplified signal (in such a way that after preamplification the threshold moves to 64 dB). This value of the threshold was set as low as possible in order to avoid parasitic noise. The sample corresponded to commercially available fine art fusains (HB 5 mm, NITRAM, Canada).

(ii) Vycor: We use the labquake catalogs of Ref. [32], in which the authors performed four experiments with cylindrical samples (diameters $\Phi = 4.45$ mm and heights $H = 8$ mm) of Vycor (a mesoporous silica glass with 40% porosity) at a constant rate $\dot{z} = 0.005$ mm/min. Before compression, samples were cleaned with a 30% solution of H_2O_2 during 24 h and dried at 130°C . The four experiments were performed with the following preamplification values: 60, 40, 20, and 0 dB, and the respective values of the detection threshold, 23, 43, 63, and 83 dB (in such a way that after preamplification the threshold always moves to 83 dB). This value of the threshold was set as low as possible in order to avoid parasitic noise.

An important difference between these two materials is the degree of heterogeneity. The synthetic mesoporous silica structure of Vycor is much more homogeneous than the one of the charcoal, which has been formed through different natural processes and may contain voids and macropores. These structural differences may lead to differences in the energy exponents [40].

B. Merging earthquake catalogs

In its usual form, the GR law fits a power-law model that contemplates a unique power-law exponent. Nevertheless, several studies have elucidated the existence of double-power-law behavior in the GR law for global seismicity [7,41,42]. Authors in Ref. [7] claim that a truncated power law with exponent $\gamma \simeq 1.66$ cannot be rejected up to $m_{\max} \simeq 7.4$ and a second power-law tail emerges from $m'_{\min} = 7.67$ with an exponent $\gamma' = 2.1 \pm 0.1$.

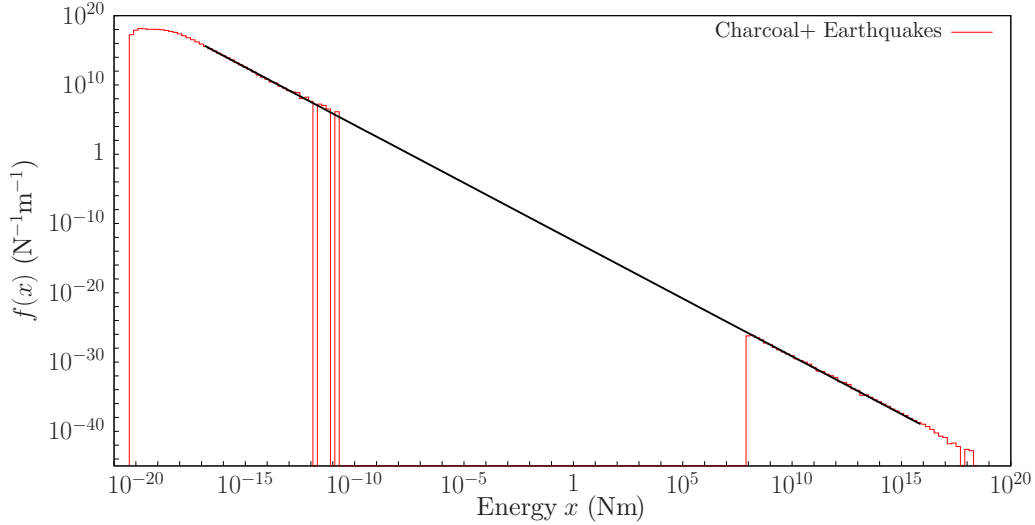


FIG. 3. Estimated PDF $f(x)$ of the radiated energy for the merged earthquake and charcoal labquake catalogs. The fit is represented by a solid black line. The methodology to merge the three earthquake catalogs is the same as in the rest of plots, whereas the addition of the charcoal catalog to this fit has been done *ad hoc* by conveniently rescaling both parts (those corresponding to charcoal and earthquakes, respectively). Each part has been divided by an effective number of events by assuming that the probability in each part corresponds to that obtained from a global power-law exponent with exponent $\hat{\Gamma}$ from $x_{\min}^{(0)}$ to $x_{\max}^{(3)}$. Note that events from the CMT, YSH A and YSH B catalogs below their respective lower cut-offs x_{\min} have been excluded in the plot.

We have checked these results by using the residual CV test. Furthermore, we have been able to establish that, by fixing the upper truncation at $m_{\max} = m'_{\min} = 7.67$, the truncated power-law hypothesis cannot be rejected (see Table I). Consequently, if one wants to fit a power-law PDF with a unique exponent, one has to exclude all those earthquakes with $m \geq m_{\max} = 7.67$. For the CMT catalog, we fix the upper cut-off $x_{\max} = 10^{\frac{3}{2}m_{\max}+9.1}$ and we work by following the same procedure described in the previous section. The rest of catalogs can be safely fitted by untruncated power-law PDFs because the CV test does not reject the hypothesis of a unique power-law tail and the magnitudes which are studied are considerably smaller than those in the CMT catalog.

Thus, we consider two untruncated power-law distributions and a third one which is truncated for the CMT catalog [see Figs. 2(a) and 2(c)]. For each decade, we sample five values of $x_{\min}^{(i)}$ equally spaced in logarithmic scale, and all the possible combinations of cut-offs $x_{\min}^{(1)}$, $x_{\min}^{(2)}$, and $x_{\min}^{(3)}$ are checked for a fixed upper-truncation $x_{\max}^{(3)}$. The labels (1), (2), and (3) correspond to the catalogs YSH B, YSH A, and CMT, respectively.

In Table II we present the results of the global fit for models α and β . The values of the statistics as well as the resulting p values are shown in Table III. It can be observed that, for this particular set of cut-offs, the CKSD test has a smaller p value and can be considered to be a more strict statistic with respect the KSDMD (in agreement with the results in Appendix D). No fit with a smaller p value has been found for the KSDMD statistic, and the fit shown in Table II has a p value that exceeds both $p_c = 0.05$ and $p_c = 0.20$. A b value very close to one holds for more than 8 orders of magnitude in seismic moment from $m_{\min} = 1.93$ to $m_{\max} = 7.67$ [see Figs. 2(b) and 2(d)]. Due to the upper-truncation, the value of the global exponent is not exactly the same as the value

of the harmonic mean of the particular exponents [32,43]. The results for the CKSD statistic do not show remarkable differences if the critical p value is set to $p_c = 0.20$ (see Table II). We consider that the definitive fit is the one whose goodness-of-fit test has been performed with a threshold value of $p_c = 0.20$. Given the number of data for each dataset and the relative difference among exponents (around 1%), these results are compatible with the ones obtained from synthetic catalogs shown in Appendix D.

C. Merging earthquakes and charcoal labquake catalogs

Motivated by the fact that the power-law exponents of earthquake catalogs and charcoal labquakes are very similar (see Table I), we apply this methodology by adding also the charcoal catalog.

At this step it is important to stress the fact that the seismic moment does not correspond with the radiated energy E_r by the earthquake, which would be the reasonable energy to compare with the AE energy. Whether the ratio of seismically radiated energy over the seismic moment is independent on the moment magnitude is still an unsolved question [44]. A constant ratio would imply that the static stress drop is constant for all the earthquakes. This ratio may depend on different earthquake parameters such as moment magnitude and the depth of the source [45–48]. On the contrary, some authors argue that the seismically radiated energy is in some cases underestimated and this ratio can be considered as constant [49–52]. For our study, we are going to consider this ratio as constant so that the values of the seismic moment should just be multiplied by a unique factor. The value of this unique factor is $\frac{E_r}{M} = 10^{-4.6}$, see Ref. [53], where M is the seismic moment (previously called x). In this case, x corresponds to the energy radiated in seismic waves by earthquakes E_r and the AE energy.

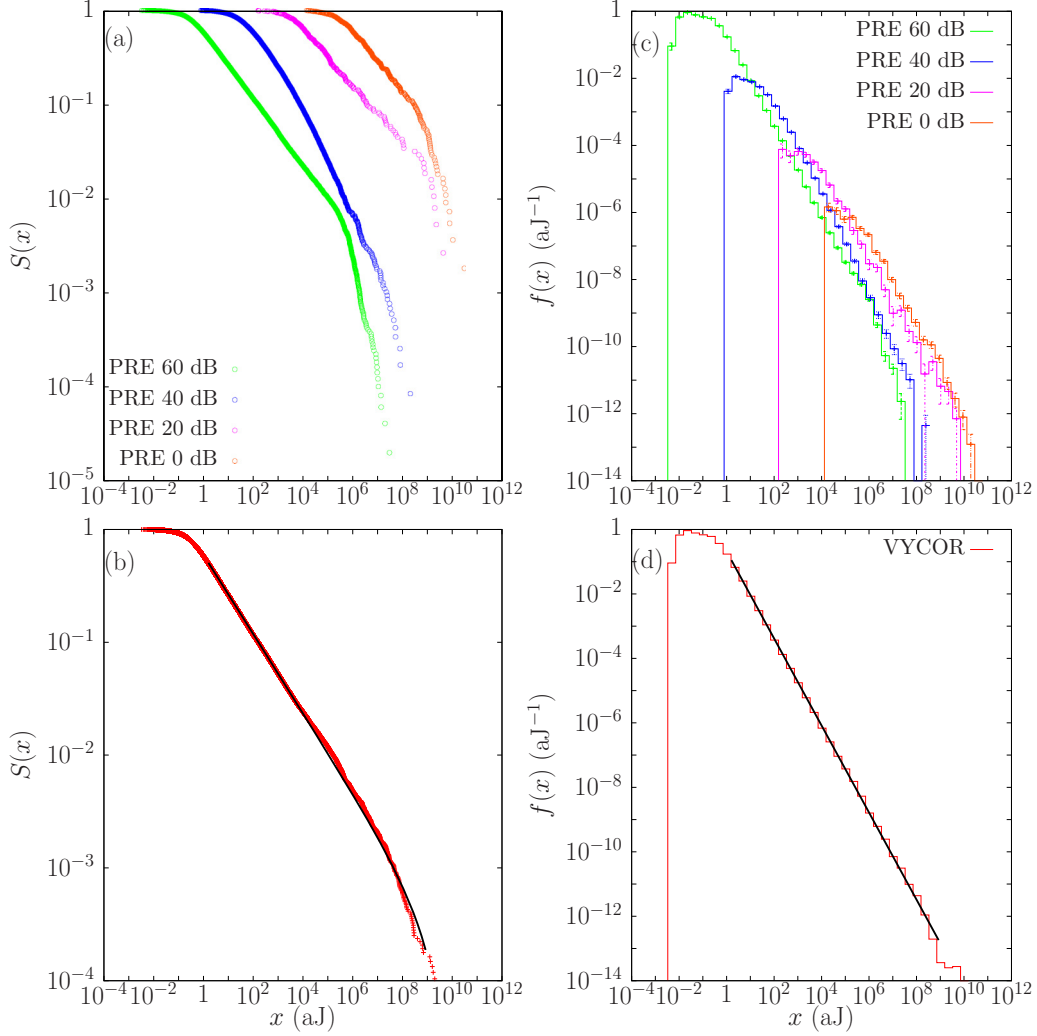


FIG. 4. Estimated survivor function $S(x)$ of the Gutenberg-Richter law for each Vycor catalog (a) and for the merged Vycor catalogs (b). Estimated PDFs $f(x)$ of the Gutenberg-Richter law for each Vycor catalog (c) and for the merged catalogs (d). The merged histograms are plotted by following the procedure explained in Ref. [32]. Fits are represented by solid black lines. In the bottom plots, events corresponding to the experiments performed at 40, 20, and 0 dB below their respective lower cut-offs as well as those corresponding to the experiments performed at 60, 40, and 20 dB above their respective upper cut-offs have been excluded.

It can be shown that, for both models α and β , multiplying the variable by a constant factor only introduces a constant term in the log-likelihood that does not change neither the

maximum nor the difference of the log-likelihoods. As the CKSD statistic is a weighted average of the particular KS distances of each dataset, it does not change either. Note

TABLE I. Results of fitting the GR law for each individual catalog. The total number of earthquakes in each catalog is given by N , whereas the number of data entering into the fit is n . The value m_{top} corresponds to the maximum observed value for each catalog. The GR law is valid for each catalog from $[m_{\text{min}}, m_{\text{max}}]$ with a particular b value. m_{max} has no upper limit for any of the fits except for the CMT catalog, in which $m_{\text{max}}^{(3)} = 7.67$ ($x_{\text{max}}^{(3)} = 4.03 \times 10^{20}$ Nm). For the charcoal catalog, x represents the energy collected by the transducer (instead of the seismic moment). No magnitude scale is provided in this case. Numbers in parentheses correspond to the error bar estimated with one standard deviation in the scale given by the last digit shown. The p value of the fits has been computed with 10^3 simulations and $p_c = 0.05$. For completeness, values of cut-offs in seismic moment x and power-law exponents $\hat{\gamma}$ have been included.

	N	n	m_{min}	$x_{\text{min}} \text{ (Nm)}$	m_{top}	$x_{\text{top}} \text{ (Nm)}$	\hat{b} value	$\hat{\gamma}$	p_{fit}
(0) Charcoal	101524	18625	—	4.93×10^{-18}	—	1.86×10^{-11}	0.984(8)	1.656(5)	0.088(9)
(1) YHS B	26330	3412	1.93	10^{12}	5.39	1.53×10^{17}	0.99(3)	1.66(2)	0.072(8)
(2) YHS A	152924	4353	3.17	7.08×10^{13}	7.20	7.94×10^{19}	0.98(1)	1.65(1)	0.080(9)
(3) CMT	48637	22336	5.33	1.24×10^{17}	9.08	5.25×10^{22}	0.982(7)	1.655(5)	0.36(2)

TABLE II. Results of fitting the models α and β to the earthquake catalogs when doing the goodness-of-fit test with the CKSD statistic and different values of p_c . If the goodness-of-fit test is performed with the KSDMD statistic, the same set of cut-offs as the fit done by using CKSD_{0.20} is found (see Table III). Same notation as in Table I. \hat{b}_g and $\hat{\gamma}$ correspond to the fitted global b value and global power-law exponent, respectively. The number of orders of magnitude $OM = \log_{10} [\frac{x_{\max}^{(i)}}{x_{\min}^{(i)}}]$ covered by each fit as well as for model α is also shown. The value $x_{\max}^{(i)}$ is replaced for $x_{\text{top}}^{(i)}$ for untruncated fits. Note that the range of the merged fit $m_{\max}^{(3)} - m_{\min}^{(1)} = 7.67 - 1.93 = 5.74$ units of magnitude, which correspond to 8.6 orders of magnitude in seismic moment.

$p_c = 0.05$	n	m_{\min}	x_{\min} (Nm)	OM	\hat{b} value	$\hat{\gamma}$	p_{fit}
(1) YHS B	3412	1.93	10^{12}	5.18	0.99(1)	1.633(7)	0.072(8)
(2) YHS A	3500	3.27	10^{14}	5.90	0.99(2)	1.66(1)	0.089(9)
(3) CMT	19003	5.40	1.58×10^{17}	3.40	0.98(8)	1.655(5)	0.26(1)
Model α	\mathcal{N}			ΣOM	\hat{b}_g value	$\hat{\Gamma}$	p_{fit}
	25915	1.93	10^{12}	14.48	0.991(6)	1.661(4)	0.079(9)
$p_c = 0.20$	n	m_{\min}	x_{\min} (Nm)	OM	\hat{b} value	$\hat{\gamma}$	p_{fit}
(1) YHS B	3412	1.93	10^{12}	5.18	0.99(1)	1.633(7)	0.072(8)
(2) YHS A	3500	3.27	10^{14}	5.90	0.99(2)	1.66(1)	0.089(9)
(3) CMT	10422	5.67	3.98×10^{17}	3	1.00(1)	1.663(7)	0.62(2)
Model α	\mathcal{N}			ΣOM	\hat{b}_g value	$\hat{\Gamma}$	p_{fit}
	17334	1.93	10^{12}	14.08	1.000(8)	1.667(5)	0.326(5)

that, as the data do not overlap, the KSDMD cannot be used. Therefore, the results shown in Table I would not change except for the values of the cut-offs.

The CV test does not reject the hypothesis of a unique power-law tail for the charcoal catalog and an untruncated power-law model is considered for this catalog. For each decade, five different values of $x_{\min}^{(i)}$ equally spaced in logarithmic scale, for a fixed upper-truncation $x_{\max}^{(3)}$, were checked and all the possible combinations of cut-offs $x_{\min}^{(0)}$, $x_{\min}^{(1)}$, $x_{\min}^{(2)}$, and $x_{\min}^{(3)}$ were checked for a fixed upper-truncation $x_{\max}^{(3)}$. The labels (0), (1), (2), and (3) correspond to the catalogs of the charcoal experiment, YSH B, YSH A, and CMT, respectively. In Table IV we present the results of the global fit for the CKSD statistic. The value of the global exponent is approximately in agreement with the harmonic mean of the particular exponents of the GR law for each catalog [32]. The results do not show remarkable differences if the critical p value is set to $p_c = 0.20$ (see Table IV). We consider that the definitive fit is the one whose goodness-of-fit test has been performed with a

threshold value of $p_c = 0.20$. The fit is shown in Fig. 3. These results are compatible with the ones obtained from synthetic catalogs shown in Appendix D.

D. Merging Vycor labquake catalogs

The values of the particular exponents γ_i of Vycor labquake catalogs differ remarkably from those found for earthquakes and charcoal labquakes [32]. No combination of the cut-offs has lead to a good fit if Vycor labquakes are merged with charcoal labquakes or earthquakes. Therefore, Vycor labquakes can be considered to be in a different universality class.

Due to the limitations due to saturation effects for high-energy AE events, it is convenient to check whether it is necessary an upper truncation for the power-law regime [see Figs. 4(a) and 4(c)]. The residual CV test reveals the upper truncations of the experiments performed at 60, 40, and 0 dB but it does not provide conclusive results for the experiment at 20 dB. In order to explore fewer combinations of cut-offs, thus reducing the computation time, we fix the upper truncations $x_{\max}^{(0)}$, $x_{\max}^{(1)}$, and $x_{\max}^{(3)}$, whereas $x_{\max}^{(2)}$ is considered as a free parameter. The labels (0), (1), (2), and (3) correspond to the experiments performed at 60, 40, 20, and 0 dB, respectively. Five intervals per decade equally spaced in logarithmic scale are sampled for each catalog and all the possible combinations of cut-offs, $x_{\min}^{(0)}$, $x_{\min}^{(1)}$, $x_{\min}^{(2)}$, $x_{\min}^{(3)}$, and $x_{\max}^{(3)}$, are checked for fixed upper-truncations $x_{\max}^{(0)}$, $x_{\max}^{(1)}$, and $x_{\max}^{(3)}$.

In Table V we present the results of the global fits for Vycor catalogs for the KSDMD and the CKSD statistics. In both cases, the global exponents are very similar but the fit performed with the KSDMD statistic maximizes the sum of orders of magnitude $\sum_{i=1}^{n_{\text{ds}}} \log_{10} [\frac{x_{\max}^{(i)}}{x_{\min}^{(i)}}]$. The same combination

TABLE III. Results of fitting model α to all the catalogs when doing the goodness-of-fit test with the KSDMD and the CKSD statistics for $p_c = 0.05$ (same cut-offs for all the catalogs) and for the set of cut-offs for which the p value is larger or equal than $p_c = 0.20$ for the CKSD statistic. KSDMD has a unique fit because the fit already exceeds $p_c = 0.20$. For the definition of D_e , see Appendix C. The remaining notation is as in previous tables.

	D_e	p_{fit}
KSDMD	0.018487	0.20(1)
CKSD _{0.05}	3.25082	0.079(9)
CKSD _{0.20}	2.778202	0.326(5)

TABLE IV. Results of fitting models α and β to the charcoal and earthquake datasets for two different values of p_c . Same notation as in previous tables. Note that x represents the radiated energy in seismic waves by earthquakes and the radiated AE energy.

$p_c = 0.05$	n	m_{\min}	x_{\min} (Nm)	OM	\hat{b} value	$\hat{\gamma}$	p_{fit}
(0) Charcoal	15906	—	6.31×10^{-18}	6.47	0.988(8)	1.658(5)	0.15(1)
(1) YHS B	1353	2.33	9.99×10^7	4.58	0.98(3)	1.66(2)	0.10(1)
(2) YHS A	234	4.47	1.59×10^{11}	4.10	0.98(6)	1.65(4)	0.62(2)
(3) CMT	7689	5.80	1.59×10^{13}	2.80	1.00(1)	1.667(9)	0.393(5)
Model α	\mathcal{N}			$\sum \text{OM}$	\hat{b}_g value	$\hat{\Gamma}$	p_{fit}
CKSD $D_e = 3.589973$	25182		6.3×10^{-18}	17.95	1.003(6)	1.669(4)	0.057(7)
$p_c = 0.20$	n	m_{\min}	x_{\min} (Nm)	OM	\hat{b} value	$\hat{\gamma}$	p_{fit}
(0) Charcoal	3555	—	6.31×10^{-17}	5.47	1.04(2)	1.69(1)	0.88(1)
(1) YHS B	1007	2.47	1.59×10^8	4.38	0.99(3)	1.66(2)	0.058(7)
(2) YHS A	234	4.47	1.59×10^{11}	4.10	0.98(6)	1.65(4)	0.62(2)
(3) CMT	3014	6.20	6.33×10^{13}	2.20	1.00(2)	1.67(2)	0.59(2)
Model α	\mathcal{N}			$\sum \text{OM}$	\hat{b}_g value	$\hat{\Gamma}$	p_{fit}
CKSD $D_e = 2.998867$	7810		6.3×10^{-17}	16.15	1.03(1)	1.688(8)	0.21(1)

of cut-offs would not be a good fit if the CKSD statistic were used instead. This fact can be justified by the performance of the method for synthetic data by applying the goodness-

of-fit test with both statistics. The goodness-of-fit test for the KSDMD statistic tends to exhibit nonrejectable p values for the same datasets for which the CKSD statistic clearly

TABLE V. Results of fitting models α and β to the Vycor Labquake catalogs when doing the goodness-of-fit test with the KSDMD and the CKSD statistics and different values of p_c . KSDMD has a unique fit because the fit already exceed $p_c = 0.20$. Same notation as in previous tables.

$p_c = 0.20$	n	x_{\min} (aJ)	x_{top} (aJ)	x_{\max} (aJ)	OM	$\hat{\gamma}$	p_{fit}
(0) PRE 60	24338	1.58	2.66×10^7	5.37×10^5	5.53	1.351(2)	0.080(6)
(1) PRE 40	5083	158.489	1.82×10^8	15860	2.00	1.34(1)	0.087(9)
(2) PRE 20	263	6309.57	3.76×10^9	10^9	5.20	1.30(2)	0.67(1)
(3) PRE 0	321	6.31×10^5	2.62×10^{10}	3.19×10^8	2.70	1.31(1)	0.25(1)
Model α	\mathcal{N}				$\sum \text{OM}$	$\hat{\Gamma}$	
KSDMD $D_e = 0.005911$	30005	1.58	2.62×10^{10}	10^9	15.43	1.350(2)	0.20(1)
$p_c = 0.05$	n	x_{\min} (aJ)	x_{top} (aJ)	x_{\max} (aJ)	OM	$\hat{\gamma}$	p_{fit}
(0) PRE 60	24338	1.58	2.66×10^7	5.37×10^5	5.53	1.351(2)	0.080(6)
(1) PRE 40	4179	251.189	1.82×10^8	15860	1.80	1.35(1)	0.19(1)
(2) PRE 20	181	2.51×10^4	3.76×10^9	2.51×10^9	5	1.29(3)	0.18(1)
(3) PRE 0	198	2.51×10^6	2.62×10^{10}	3.19×10^8	2.10	1.36(6)	0.36(2)
Model α	\mathcal{N}				$\sum \text{OM}$	$\hat{\Gamma}$	
CKSD $D_e = 3.362608$	28896	1.58	2.62×10^{10}	2.51×10^9	14.43	1.351(2)	0.052(7)
$p_c = 0.20$	n	x_{\min} (aJ)	x_{top} (aJ)	x_{\max} (aJ)	OM	$\hat{\gamma}$	p_{fit}
(0) PRE 60	20778	2.51	2.66×10^7	5.37×10^5	5.33	1.354(3)	0.51(2)
(1) PRE 40	3404	398.107	1.82×10^8	15860	1.60	1.36(2)	0.36(2)
(2) PRE 20	155	3.98×10^4	3.76×10^9	10^9	4.40	1.32(4)	0.20(1)
(3) PRE 0	159	3.98×10^6	2.62×10^{10}	3.19×10^8	1.90	1.34(7)	0.29(1)
Model α	\mathcal{N}				$\sum \text{OM}$	$\hat{\Gamma}$	
CKSD $D_e = 2.586434$	24496	2.51	2.62×10^{10}	10^9	13.23	1.354(2)	0.20(1)

rejects the null hypothesis. See Appendix D for more details. Although the fit performed with the KSDMD covers a larger number of orders magnitude, the definitive fit will be the one performed with the CKSD and $p_c = 0.20$ (see Table V). Given that the ranges covered by the other acceptable fits lead to the rejection of the global power-law hypothesis when the CKSD is used with $p_c = 0.20$, this fit is more restrictive than the other ones (see Appendix D) and it is the “safest” choice. The definitive fit is shown in Fig. 4(b) and 4(d) and covers about nine orders of magnitude in energy with a number of data $\mathcal{N} = 24\,496$ entering into the fit.

IV. CONCLUSIONS

We have presented a statistical procedure to merge different datasets in order to validate the existence of universal power-law exponents across different scales or phenomena. This methodology can be useful in the study of different complex systems in order to check whether the power-law exponents obtained via maximum likelihood estimation are statistically compatible among them or not. Therefore, the procedure presented in this paper provides a statistical tool that enables us to establish whether different complex systems can be classified into the same universality class or not.

In this work, the methodology has been applied to the Gutenberg-Richter law for earthquakes and labquakes. By merging earthquake catalogs, a global power law with a global exponent $\Gamma = 1.667$ holds for more than eight orders of magnitude in seismic moment (from $m_{\min} = 1.93$ to $m_{\max} = 7.67$ in moment magnitude). To our knowledge, this is the broadest fitting range that has been found for the Gutenberg-Richter law for earthquakes with a unique value of the exponent [54]. There are catalogs of tiny mining-induced earthquakes which exhibit a much smaller completeness magnitude [55] than the ones of natural seismicity used in this work. They were not considered here because they are not currently public and show b values significantly different [56] from the one found here, which would result in nonacceptable fits when merging them with the rest of catalogs, possibly pointing to a different universality class. Future works involving different earthquake catalogs can be carried out in order to find a broader fitting range of the Gutenberg-Richter law and also to check whether different regions have compatible power-law exponents or not. This kind of studies would be of interest in order to statistically strengthen the geological arguments that justify the difference in the b values observed in some regions [43].

Earthquake catalogs have been also merged with a charcoal labquake catalog with a global power-law exponent $\Gamma = 1.688$ suggesting that these different systems might be classified into the same universality class. Further investigations involving different observables, such as the distribution of waiting times (time between consecutive events), might be necessary in order to properly classify charcoal labquakes and real earthquakes into the same universality class.

A previous methodology for merging datasets was not able to find a good fit for the Gutenberg-Richter law for the four Vycor Labquake catalogs [32]. The previous procedure did not take into account the fact that the cut-offs of the merged catalogs could be different to those that limited the power-law regime for the individual catalogs. With the methodology

exposed in this paper, which allows to consider different cut-offs, these Vycor labquake catalogs have been merged and a GR law with exponent $\Gamma = 1.35$ for the energy AE events has been found to hold for nine orders of magnitude. As it was proposed in Ref. [40], the labquakes in this material have been found to belong to a different universality class than earthquakes and charcoal labquakes.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the “La Caixa” Foundation. V.N.-P. acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (MINECO, Spain) through the “María de Maeztu” Programme for Units of Excellence in R & D (Grant No. MDM-2014-0445). We also acknowledge financial support from the MINECO under Grants No. FIS2015-71851-P, No. FIS-PGC2018-099629-B-I00, and No. MAT2016-75823-R and from the Agència de Gestió d’Ajuts Universitaris i de Recerca (AGAUR) under Grant No. 2014SGR-1307 and the MINECO Juan de la Cierva research Contract No. FJCI-2016-29307 held by Á.G.

APPENDIX A: LIKELIHOOD RATIO TEST

The LRT is a statistical procedure used to check whether it is worth using a statistical model α (characterized by a set of n_α parameters) which is nested into a more general statistical model β (characterized by a set of n_β parameters). The null hypothesis of this test states that the simpler model α is good enough to describe the data and a more general model β does not provide a significant improvement. In order to perform this test, the likelihood ratio is defined as:

$$\mathcal{R} = \log \left(\frac{\hat{\mathcal{L}}_\beta}{\hat{\mathcal{L}}_\alpha} \right) = \log \hat{\mathcal{L}}_\beta - \log \hat{\mathcal{L}}_\alpha, \quad (\text{A1})$$

where \log is the natural logarithm and the hats denote that the likelihoods \mathcal{L}_α and \mathcal{L}_β are evaluated at those values of the set of parameters for which they reach their maximum values.

Although not all the regularity conditions exposed in Ref. [31] are satisfied when dealing with power-law PDFs, it has been numerically tested that the statistic $2\mathcal{R}$ follows a chi-squared distribution with $n_\beta - n_\alpha > 0$ degrees of freedom, at least for the largest percentiles [21]. Once the significance level of the test has been fixed, one is able to determine the critical value $2\mathcal{R}_c$. If the empirical value $2\mathcal{R}_e$ is below this threshold, then one can consider that there are not enough statistical evidences to say that the more complex model β is needed in order to describe the data. In this situation, due to its simplicity, model α is preferable to fit data. This procedure can be naively understood as a statistical version of the Occam’s razor criterion. Note that the rejection of the null hypothesis does not imply the rejection of model α as a fit to data, on the contrary, the “acceptance” of model α does not imply that it is a good fit to the data. The test is just a relative comparison between two models.

We present the expressions that are necessary to compute the statistic of the likelihood ratio test in this work. The log-

likelihood for the untruncated PDF in Eq. (3) is written:

$$\log \mathcal{L}_{\text{untrunc}}(x; x_{\min}, \gamma) = n \log \left(\frac{\gamma - 1}{x_{\min}^{1-\gamma}} \right) - \gamma \sum_{j=1}^n \log x_j, \quad (\text{A2})$$

whereas the log-likelihood for the truncated PDF in Eq. (4) is written:

$$\begin{aligned} \log \mathcal{L}_{\text{trunc}}(x; x_{\min}, x_{\max}, \gamma) \\ = n \log \left(\frac{1 - \gamma}{x_{\max}^{1-\gamma} - x_{\min}^{1-\gamma}} \right) - \gamma \sum_{j=1}^n \log x_j. \end{aligned} \quad (\text{A3})$$

These expressions are valid for model β and also for model α replacing γ by Γ .

APPENDIX B: RESIDUAL COEFFICIENT OF VARIATION TEST

A useful statistical tool to check whether an untruncated power-law distribution is a good model to fit data in a certain range is the so-called residual coefficient of variation (CV) test [57].

Let us suppose that we have a variable \mathcal{X} that acquires N values $\{x\}$. We define $x_{(j)}$ as the j th value of the variable when it is sorted in ascending order: $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(j)} \leq \dots \leq x_{(N)}$. The null hypothesis of the test states that there exists a power-law tail given by $x > x_{(k)}$. The test is based on the idea that for an untruncated power-law distribution the logarithmic coefficient of variation C_l is close to 1:

$$C_l = \frac{s_l}{m_l}, \quad (\text{B1})$$

where m_l corresponds to the mean of the logarithm of the rescaled variable $l = \log(x/x_{(k)})$:

$$m_l = \frac{1}{N - k} \sum_{j=k+1}^N \log \left[\frac{x_{(j)}}{x_{(k)}} \right], \quad (\text{B2})$$

and s_l^2 is the unbiased variance:

$$s_l^2 = \frac{1}{N - k - 1} \sum_{j=k+1}^N \left\{ \log \left[\frac{x_{(j)}}{x_{(k)}} \right] - m_l \right\}^2. \quad (\text{B3})$$

In order to check whether the value of C_l is close to 1 or not for a particular number of remaining data, one simulates many samples power-law distributed with the same number of data $N - k$ in order to extract a distribution for C_l . It is important to remark that the distribution of the statistic does not depend on the power-law exponent and, therefore, it is not necessary to estimate it previously. Once one determines the level of significance for the test (in this case 0.05), one can obtain the upper and lower critical values of the statistic by checking the percentiles of the distribution of simulated values (percentiles 2.5 and 97.5 in our case). If the empirical value of C_l lies between these two critical values, then one cannot reject the null hypothesis of power-law tail. Otherwise, if the empirical value is below percentile 2.5, then the power law is rejected in favor of a truncated log-normal [58]; if the empirical value is above percentile 97.5, then the power law is rejected but there

is no alternative. One proceeds by computing the residual C_l for increasing values of the index k , thus analyzing different possible starting points for the tail of the distribution.

APPENDIX C: GOODNESS-OF-FIT TEST

Determining whether the null hypothesis of considering a global exponent Γ is compatible with the values of the particular fits presents some differences with the standard Kolmogorov-Smirnov goodness-of-fit test presented in Ref. [29]. The test statistic used in this standard goodness-of-fit test is the Kolmogorov-Smirnov distance (KSD), which is defined as the maximum difference between the empirical and the theoretical cumulative distribution functions. In this Appendix we expose the two different statistics for the goodness-of-fit test that have been used for the global fit: The first statistic is essentially an adaptation of the standard KS test for the merged case, whereas the second method uses a statistic which is a composition of KS distances.

When the value of the global exponent $\hat{\Gamma}$ has been found through MLE, one can understand that each dataset contributes to the global PDF with a global exponent $\hat{\Gamma}$ in their particular ranges $[x_{\min}^{(i)}, x_{\max}^{(i)}]$ ($i = 1, \dots, n_{\text{cat}}$). The global distribution can be understood as a global power-law PDF with exponent Γ ranging from $X_{\min} = \min\{x_{\min}^{(i)}\}$ to $X_{\max} = \max\{x_{\max}^{(i)}\}$. In this situation, we need to redefine the KSD where we have merged several datasets (the formulas below for truncated power laws, the untruncated case is recovered in the limit case in which the upper cut-off goes to infinity).

(i) Kolmogorov-Smirnov distance of the merged datasets (KSDMD): This statistic can be used as long as datasets overlap each other. Datasets can be merged by pairs by giving a certain weight to each point from data. Depending on which overlapping or nonoverlapping region a point from our data comes from, each point will have a certain weight ω_j , $j = 1, \dots, \mathcal{N}$. For more details about the expression of these weights, see Ref. [32]. By sorting data in ascending order, it is easy to construct the empirical cumulative distribution function for the merged datasets by using the following expression:

$$F_e[x_{(k)}] = \frac{\sum_{j=1}^k \omega_j}{\sum_{j=1}^{\mathcal{N}} \omega_j}, \quad (\text{C1})$$

with $k = 1, \dots, \mathcal{N}$ and the subindex e refers to the empirical CDF. Note that, for a standard situation in which datasets are not merged, all the weights would be 1 [29]. Note also that the survivor function is $S_e(x_{(k)}) = 1 - F_e[x_{(k)}]$. Once we have constructed the merged $F_e(x)$, we can easily compute the KS distance by:

$$D_e^{(\text{KSDMD})} = \max \left| \left(\frac{x^{1-\hat{\Gamma}} - X_{\min}^{1-\hat{\Gamma}}}{X_{\max}^{1-\hat{\Gamma}} - X_{\min}^{1-\hat{\Gamma}}} \right) - F_e(x) \right|, \quad (\text{C2})$$

where the values of x are taken from the ranges $[x_{\min}^{(i)}, x_{\max}^{(i)}]$ for each dataset.

(ii) Composite KS distance (CKSD): This second statistic can be computed independently on whether the datasets are overlapping each other. This statistic is constructed from the

TABLE VI. Performance of the methodology for merging datasets explained in the main text for two untruncated power laws with exponents γ_1 and γ_2 , lower cut-offs $x_{\min}^{(1)} = 10^{12}$ and $x_{\min}^{(2)} = 10^{14}$ (arbitrary units) and sizes n_1 and n_2 . Five different groups are presented depending on the relative difference of the power-law exponents δ_γ . For these datasets, the fitted global exponent $\hat{\Gamma}$ is estimated and the LRT statistic $2\mathcal{R}_e$ is computed. The statistics $D_e^{(\text{CKSD})}$ and $D_e^{(\text{KSDMD})}$ together with their corresponding p values $p\text{-CKSD}$ and $p\text{-KSDMD}$ are presented for each combination of datasets. As a complement, the z statistic and the p value $p_{\text{norm-}z}$ from the z test are also shown. p values in bold correspond to those which are below $p_c = 0.05$.

n_1	γ_1	σ_1	n_2	γ_2	σ_2	δ_γ (%)	$\hat{\Gamma}$	$2\mathcal{R}_e$	$D_e^{(\text{CKSD})}$	$p\text{-CKSD}$	$D_e^{(\text{KSDMD})}$	$p\text{-KSDMD}$	z	$p_{\text{norm-}z}$
100	1.50	0.05	100	1.50	0.05	0	1.471	0.624	1.635	0.425	0.075	0.453	0.788	0.431
1000	1.50	0.02	1000	1.50	0.02	0	1.503	1.284	1.316	0.801	0.023	0.595	1.132	0.257
10000	1.50	0.005	10000	1.50	0.005	0	1.501	0.139	1.289	0.837	0.009	0.986	0.373	0.707
100	1.53	0.05	100	1.50	0.05	2	1.484	1.443	1.660	0.395	0.070	0.578	1.195	0.232
1000	1.53	0.02	1000	1.50	0.02	2	1.518	0.029	1.239	0.876	0.017	0.893	0.170	0.865
10000	1.53	0.005	10000	1.50	0.005	2	1.516	14.037	3.187	5×10^{-4}	0.013	0.024	3.745	1.81×10^{-4}
100	1.575	0.058	100	1.50	0.05	5	1.514	3.154	1.815	0.249	0.058	0.755	1.755	0.079
1000	1.575	0.018	1000	1.50	0.02	5	1.541	2.458	2.032	0.111	0.031	0.180	1.567	0.117
10000	1.575	0.006	10000	1.50	0.005	5	1.536	64.957	4.740	0	0.029	0	9.474	0
100	1.65	0.065	100	1.50	0.05	10	1.528	6.956	2.262	0.046	0.054	0.832	2.571	0.010
1000	1.65	0.02	1000	1.50	0.02	10	1.57	22.365	3.369	10^{-4}	0.044	0.015	4.690	2.73×10^{-6}
10000	1.650	0.007	10000	1.50	0.005	10	1.566	329.552	10.601	0	0.053	0	17.933	0
100	1.725	0.073	100	1.50	0.05	15	1.551	11.566	2.661	0.008	0.090	0.260	3.262	0.001
1000	1.725	0.023	1000	1.50	0.02	15	1.598	51.266	4.608	0	0.066	10^{-4}	7.026	2.12×10^{-12}
10000	1.725	0.007	10000	1.50	0.005	15	1.593	667.096	14.507	0	0.075	0	25.204	0

n_{cat} particular KS distances:

$$D_{e,i} = \max_{x_{\min}^{(i)} \leq x \leq x_{\max}^{(i)}} \left| \left[\frac{x^{1-\hat{\Gamma}} - x_{\min}^{(i)1-\hat{\Gamma}}}{x_{\max}^{(i)1-\hat{\Gamma}} - x_{\min}^{(i)1-\hat{\Gamma}}} \right] - F_{e,i}[x; x_{\min}^{(i)}, x_{\max}^{(i)}] \right|, \quad (\text{C3})$$

where the subindex i refers to the i th catalog and F_e is the empirical cumulative distribution function. In order to compute the statistic, we perform the following summation:

$$D_e^{(\text{CKSD})} = \sum_{i=1}^{n_{\text{cat}}} \sqrt{n_i} D_{e,i}, \quad (\text{C4})$$

where the factors $\sqrt{n_i}$ are due to the scaling of the KS distance with the number of data [59].

Once the statistic is computed, one needs to determine whether it is big or small in relation to the one found for data sampled from a PDF with the same parameters $\{x_{\min}^{(i)}\}$, $\{x_{\max}^{(i)}\}$, Γ , and $\{n_i\}$. Data are sampled from a power-law PDF in the range given by the i th catalog with probability $q_i = n_i/\mathcal{N}$, where $\mathcal{N} = \sum_{i=1}^{n_{\text{ds}}} n_i$. Note that the particular number in each simulated dataset is not necessarily the empirical one n_i but the total number of data in the global fit \mathcal{N} is maintained. Hence, one needs a first random number to choose the dataset i and therefore the range $[x_{\min}^{(i)}, x_{\max}^{(i)}]$ and a second one to generate the random power-law number in that range with exponent $\hat{\Gamma}$ [29]. When \mathcal{N} events have been generated according to this procedure, one finds the global exponent $\hat{\Gamma}_{\text{sim}}$ by maximizing the global log-likelihood and computes either $D_{\text{sim}}^{(\text{KSDMD})}$ or $D_{\text{sim}}^{(\text{CKSD})}$ for the merged simulated datasets.

An estimation of the standard deviation of the global exponent $\hat{\Gamma}$ can be computed from the standard deviation of $\hat{\Gamma}_{\text{sim}}$. By performing several realizations of the previous procedure, one can estimate the p value of the fit by computing the fraction of simulated datasets where the simulated statistic is larger than the empirical one.

APPENDIX D: PERFORMANCE ON SYNTHETIC DATASETS

Once the methodology for merging datasets has been presented together with the different goodness-of-fit tests, it is important to check the performance of the method on synthetic data. In order to carry out this analysis, two untruncated power-law-distributed datasets with exponents γ_1 and γ_2 , lower cut-offs $x_{\min}^{(1)}$ and $x_{\min}^{(2)}$, and sizes n_1 and n_2 , were generated. The sizes of both datasets were considered to be equal $n_1 = n_2$ so as to simplify the analysis.

The global exponent $\hat{\Gamma}$ of the merged catalogs was estimated according to the methodology exposed in the main text and the empirical LRT statistic $2\mathcal{R}_e$ was computed according to the methods exposed in Appendix B. Given that the difference on parameters between model α and β in this case is 1, the critical value of the test with a significance level equal to 0.05 is $2\mathcal{R}_c = 3.84$. If the empirical LRT statistic $2\mathcal{R}_e$ is found to be larger than the critical one, then one rejects the null hypothesis that the simpler model α is good enough to describe data and, consequently, more parameters are needed. Once the global exponent and the likelihood ratio statistic were found, the two different goodness-of-fit tests exposed in Appendix C were performed.

The analysis was performed for different dataset sizes as well as different power-law exponents γ_1 and γ_2 . The results of the method for synthetic datasets are shown in Table VI. Five groups are presented depending on the relative difference δ_γ between exponents. The p values of the CKSD and the KSDMD goodness-of-fit tests have been computed with 10^4 Monte Carlo simulations (see Appendix C for more details).

In order to compare our results with a test that checks whether two power-law exponents are significantly different or not, the p_{norm} value of the z test exposed in Ref. [54] is also shown. These p values can be computed by assuming that, for a sufficiently large sample sizes, the z statistic follows a normal distribution with zero mean and standard deviation equal to 1.

As one would expect, if both datasets have exactly the same power-law exponent, the null hypothesis that variable \mathcal{X} is power-law distributed for all the datasets with the global

exponent $\hat{\Gamma}$ is not rejected (this has to happen in the 95% of the cases, for a significance level equal to 0.05). When the CKSD goodness-of-fit test rejects the null hypothesis of a power-law distribution with a global exponent, the z test also rejects the null hypothesis of considering $\gamma_1 = \gamma_2$. The same does not apply for the KSDMD goodness of fit test, where some fits yield to nonrejectable p values, whereas the z test clearly rejects the null hypothesis. In this sense, one can consider that the KSDMD statistic is less strict than the CKSD statistic, which has therefore more power.

However, for the sample sizes which are involved in this work, both goodness-of-fit tests reject the null hypothesis for sufficiently large difference in the exponents. It can also be seen that the null hypothesis is rejected independently on the goodness-of-fit test for those fits in which the likelihood ratio statistic exceeds the critical value $2\mathcal{R}_c = 3.84$. This fact justifies the decision of performing the LRT before the goodness-of-fit test.

-
- [1] K. Christensen and N. R. Moloney, *Complexity and Criticality* (Imperial College Press, London, 2005).
 - [2] P. Bak, *How Nature Works: The Science of Self-organized Criticality* (Springer, New York, 1996).
 - [3] J. P. Sethna, K. A. Dahmen, and C. R. Myers, Crackling noise, *Nature* **410**, 242 (2001).
 - [4] E. K. H. Salje and K. A. Dahmen, Crackling noise in disordered materials, *Annu. Rev. Condens. Matter Phys.* **5**, 233 (2014).
 - [5] X. Gabaix, Power laws in economics: An introduction, *J. Econ. Perspect.* **30**, 185 (2016).
 - [6] I. Moreno-Sánchez, F. Font-Clos, and Á. Corral, Large-scale analysis of Zipf's law in English texts, *PLoS ONE* **11**, e0147073 (2016).
 - [7] Á. Corral and Á. González, Power law size distributions in geoscience revisited, *Earth Space Sci.* **6**, 673 (2019).
 - [8] Y. Y. Kagan, *Earthquakes: Models, Statistics, Testable Forecasts* (John Wiley & Sons, New York, 2013).
 - [9] Y. Y. Kagan, Earthquake size distribution: Power-law with exponent $\beta \equiv \frac{1}{2}$? *Tectonophysics* **490**, 103 (2010).
 - [10] H. E. Stanley, Scaling, universality, and renormalization: Three pillars of modern critical phenomena, *Rev. Mod. Phys.* **71**, S358 (1999).
 - [11] H. E. Stanley, L. A. N. Amaral, P. Gopikrishnan, P. Ch Ivanov, T. H. Keitt, and V. Plerou, Scale invariance and universality: Organizing principles in complex systems, *Physica A: Stat. Mech. App.* **281**, 60 (2000).
 - [12] T. Utsu, Representation and analysis of the earthquake size distribution: A historical review and some new approaches, *Pure Appl. Geophys.* **155**, 509 (1999).
 - [13] T. C. Hanks and H. Kanamori, A moment magnitude scale, *J. Geophys. Res.: Solid Earth* **84**, 2348 (1979).
 - [14] J. Woessner and S. Wiemer, Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty, *Bull. Seismol. Soc. Am.* **95**, 684 (2005).
 - [15] D. Schorlemmer and J. Woessner, Probability of detecting an earthquake, *Bull. Seismol. Soc. Am.* **98**, 2103 (2008).
 - [16] Á. González, The Spanish national earthquake catalog: Evolution, precision and completeness, *J. Seism.* **21**, 435 (2017).
 - [17] J. Davidsen and A. Green, Are Earthquake Magnitudes Clustered? *Phys. Rev. Lett.* **106**, 108502 (2011).
 - [18] S. Hainzl, Rate-dependent incompleteness of earthquake catalogs, *Seism. Res. Lett.* **87**, 337 (2016).
 - [19] A. Helmstetter, Y. Y. Kagan, and D. D. Jackson, Comparison of short-term and time-independent earthquake forecast models for Southern California, *Bull. Seismol. Soc. Am.* **96**, 90 (2006).
 - [20] Á. Corral, R. García-Millán, N. R. Moloney, and F. Font-Clos, Phase transition, scaling of moments, and order-parameter distributions in Brownian particles and branching processes with finite-size effects, *Phys. Rev. E* **97**, 062156 (2018).
 - [21] I. Serra and Á. Corral, Deviation from power law of the global seismic moment distribution, *Sci. Rep.* **7**, 40045 (2017).
 - [22] J. Baró, Á. 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).
 - [23] G. F. Nataf, P. O. Castillo-Villa, J. Baró, X. Illa, E. Vives, A. Planes, and E. K. H. Salje, Avalanches in compressed porous SiO₂-based materials, *Phys. Rev. E* **90**, 022405 (2014).
 - [24] V. Navas-Portella, Á. Corral, and E. Vives, Avalanches and force drops in displacement-driven compression of porous glasses, *Phys. Rev. E* **94**, 033005 (2016).
 - [25] N. Yoshimitsu, H. Kawakata, and N. Takahashi, Magnitude -7 level earthquakes: A new lower limit of self-similarity in seismic scaling relationships, *Geophys. Res. Lett.* **41**, 4495 (2014).
 - [26] I. G. Main, Little earthquakes in the lab, *Physics* **6**, 20 (2013).
 - [27] T. Mäkinen, A. Miksic, M. Ovaska, and M. J. Alava, Avalanches in Wood Compression, *Phys. Rev. Lett.* **115**, 055501 (2015).
 - [28] T. P. Hill and J. Miller, How to combine independent data sets for the same quantity, *Chaos* **21**, 033102 (2011).
 - [29] A. Deluca and Á. Corral, Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions, *Acta Geophys.* **61**, 1351 (2013).
 - [30] A. Clauset, C. R. Shalizi, and M. E. J. Newman, Power-law distributions in empirical data, *SIAM Rev.* **51**, 661 (2009).

- [31] Y. Pawitan, in *All Likelihood: Statistical Modelling and Inference Using Likelihood* (Oxford University Press, Oxford, 2013).
- [32] V. Navas-Portella, I. Serra, Á. Corral, and E. Vives, Increasing power-law range in avalanche amplitude and energy distributions, *Phys. Rev. E* **97**, 022134 (2018).
- [33] J. Baró and E. Vives, Analysis of power-law exponents by maximum-likelihood maps, *Phys. Rev. E* **85**, 066121 (2012).
- [34] T.-A. Chou, A. M. Dziewonski and J. H. Woodhouse, Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.* **86**, 2825 (1981).
- [35] G. Ekström, M. Nettles, and A. M. Dziewoński, The Global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. Inter.* **200-201**, 1 (2012).
- [36] W. Yang, E. Hauksson, and P. M. Shearer, Computing a large refined catalog of focal mechanisms for Southern California (1981–2010): Temporal stability of the style of faulting, *Bull. Seismol. Soc. Am.* **102**, 1179 (2012).
- [37] P. Wessel, J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian, The generic mapping tools version 6, *Geochem. Geophys. Geosyst.* **20** (2019).
- [38] K. Hutton, J. Woessner, and E. Hauksson, Earthquake monitoring in Southern California for seventy-seven years (1932–2008), *Bull. Seismol. Soc. Am.* **100**, 423 (2010).
- [39] Physical Acoustics. PCI-2: PCI-Based Two-Channel AE Board & System, <https://www.physicalacoustics.com/by-product/pci-2/>.
- [40] Y. Xu, A. G. Borrego, A. Planes, X. Ding, and E. Vives, Criticality in failure under compression: Acoustic emission study of coal and charcoal with different microstructures, *Phys. Rev. E* **99**, 033001 (2019).
- [41] J. F. E. Pacheco, C. H. Scholz, and L. R. Sykes, Changes in frequency-size relationship from small to large earthquakes, *Nature* **355**, 71 (1992).
- [42] M. R. Yoder, J. R. Holliday, D. L. Turcotte, and J. B. Rundle, A geometric frequency-magnitude scaling transition: Measuring $b = 1.5$ for large earthquakes, *Tectonophysics* **532**, 167 (2012).
- [43] Y. Kamer and S. Hiemer, Data-driven spatial b value estimation with applications to California seismicity: To b or not to b , *J. Geophys. Res.: Solid Earth* **120**, 5191 (2015).
- [44] H. Kanamori and E. E. Brodsky, The physics of earthquakes, *Rep. Prog. Phys.* **67**, 1429 (2004).
- [45] E. E. Brodsky and H. Kanamori, Elastohydrodynamic lubrication of faults, *J. Geophys. Res.: Solid Earth* **106**, 16357 (2001).
- [46] M. S. Vassiliou and H. Kanamori, The energy release in earthquakes, *Bull. Seismol. Soc. Am.* **72**, 371 (1982).
- [47] Y. Izutani and H. Kanamori, Scale-dependence of seismic energy-to-moment ratio for strike-slip earthquakes in Japan, *Geophys. Res. Lett.* **28**, 4007 (2001).
- [48] H. Kanamori and T. H. Heaton, Microscopic and macroscopic physics of earthquakes, in *Geocomplexity and the Physics of Earthquakes*, Geophysics Monograph Series, Vol. 120, edited by J. B. Rundle, D. L. Turcotte, and W. Klein (American Geophysical Union, Washington, DC, 2000), pp. 147–163.
- [49] S. Ide and G. C. Beroza, Does apparent stress vary with earthquake size? *Geophys. Res. Lett.* **28**, 3349 (2001).
- [50] J. H. Wang, The energy-magnitude scaling law for $M_s \leq 5.5$ earthquakes, *J. Seism.* **19**, 647 (2015).
- [51] S. L. Bilek, T. Lay, and L. J. Ruff, Radiated seismic energy and earthquake source duration variations from teleseismic source time functions for shallow subduction zone thrust earthquakes, *J. Geophys. Res.: Solid Earth* **109**, B09308 (2004).
- [52] A. Zollo, A. Orefice, and V. Convertito, Source parameter scaling and radiation efficiency of microearthquakes along the Irpinia Fault zone in Southern Apennines, Italy, *J. Geophys. Res.: Solid Earth* **119**, 3256 (2014).
- [53] P. Bormann, Are new data suggesting a revision of the current M_w and M_e scaling formulas? *J. Seism.* **19**, 989 (2015).
- [54] Y. Kagan, Universality of the seismic moment-frequency relation, in *Seismicity Patterns, Their Statistical Significance and Physical Meaning*, Pageoph Topical Volumes, edited by M. Wyss, K. Shimazaki, and A. Ito (Birkhäuser, Basel, 1999), pp. 537–573.
- [55] K. Plenkers, D. Schorlemmer, G. Kwiatak, and JAGUARS Research Group, On the probability of detecting picoseismicity, *Bull. Seismol. Soc. Am.* **101**, 2579 (2011).
- [56] G. Kwiatak, K. Plenkers, M. Nakatani, Y. Yabe, G. Dresen, and JAGUARS-Group, Frequency-magnitude characteristics down to magnitude -4.4 for induced seismicity recorded at Mponeng gold mine, *Bull. Seismol. Soc. Am.* **100**, 1165 (2010).
- [57] Y. Malevergne, V. Pisarenko, and D. Sornette, Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities, *Phys. Rev. E* **83**, 036111 (2011).
- [58] J. del Castillo and P. Puig, The best test of exponentiality against singly truncated normal alternatives, *J. Am. Stat. Assoc.* **94**, 529 (1999).
- [59] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992).