

Scaling laws in earthquake memory for interevent times and distances

Yongwen Zhang,^{1,2,3,*} Jingfang Fan^{4,2}, Warner Marzocchi⁵, Avi Shapira⁶, Rami Hofstetter,⁷ Shlomo Havlin,² and Yosef Ashkenazy¹

¹*Department of Solar Energy and Environmental Physics, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 84990, Israel*

²*Department of Physics, Bar-Ilan University, Ramat Gan 52900, Israel*

³*Data Science Research Center, Kunming University of Science and Technology, Kunming 650500, Yunnan, China*

⁴*Potsdam Institute for Climate Impact Research, 14412 Potsdam, Germany*

⁵*Department of Earth, Environmental, and Resources Sciences, University of Naples, Federico II, Complesso di Monte Sant'Angelo, Via Cinthia 21, 80126 Napoli, Italy*

⁶*National Institute for Regulation of Emergency and Disaster, College of Law and Business, Bnei Brak 511080, Israel*

⁷*Geophysical Institute of Israel, Lod 7019802, Israel*



(Received 5 November 2019; accepted 7 February 2020; published 5 March 2020)

Earthquakes involve complex processes that span a wide range of spatial and temporal scales. The limited earthquake predictability is partly due to the erratic nature of earthquakes and partly due to the lack of understanding of the underlying mechanisms of earthquakes. To improve our understanding and possibly the predictability of earthquakes, we develop here a lagged conditional probability method to study the spatial and temporal long-term memory of interevent earthquakes above a certain magnitude. We find, in real data from different locations, that the lagged conditional probabilities show long-term memory for both the interevent times and interevent distances and that the memory functions obey scaling and decay slowly with time, while, at a characteristic time (crossover), the decay rate becomes faster. We also show that the epidemic-type aftershock sequence model, which is often used to forecast earthquake events, fails in reproducing the scaling function of real catalogs as well as the crossover in the scaling function. Our results suggest that aftershock rate is a critical factor to control the long-term memory.

DOI: [10.1103/PhysRevResearch.2.013264](https://doi.org/10.1103/PhysRevResearch.2.013264)

I. INTRODUCTION

The mechanisms of earthquakes are still not fully understood and remain a great scientific challenge [1]. Still, there are some well known empirical laws regarding earthquakes: (i) the Gutenberg-Richter law, which determines the relation between the number of earthquakes N in a given region and a time period and the magnitude m as $N(m) \propto 10^{-bm}$ ($b \approx 1$) [2], and (ii) the Omori law, according to which the rate of triggered events is $\sim t^{-p}$, where t is the time since a triggering earthquake ($p \approx 1$ for large earthquakes) [3].

An important characteristic of some complex systems is the scaling behavior [4] where earthquakes have been found to follow scaling behavior [5–7]. Using the Gutenberg-Richter law and the exponent of the Omori law, Bak *et al.* [8] found that the probability density function (PDF) of interevent times for different magnitude thresholds and different spatial grid sizes can be rescaled into a single function. This suggests a universal scaling law for earthquakes. Corral [9] extended

this scaling to different regions and introduced a more general approach. Specifically, he introduced a unified function $f(x)$ to describe the distribution of interevent times as $D(\tau) = Rf(R\tau)$, where τ is the interevent time and $R = 1/\bar{\tau}$ is the average occurrence rate which depends on magnitude, space scale, and different locations. Corral also argued that the optimal fitted function of $f(x)$ is the generalized Γ distribution [10]. This scaling function follows a power law for small scales and decays exponentially at large scales. Some questions have been raised regarding the universal scaling with region size [11]. In the context of the epidemic-type aftershock sequence (ETAS) model, the scaling function has been found to depend on the ratio between correlated and independent events [11,12]. In addition, a different study suggested that multiple characteristic timescales, which are controlled by the parameters of the ETAS model, are relevant for the universal scaling behavior of the interevent time distribution in the ETAS model [13].

The distribution of earthquake events alone does not reflect all the information about the dynamics, and further time series analysis could improve our understanding of the underlying dynamics of earthquakes. For example, Livina *et al.* [14] studied the conditional probability of consecutive interevent times and found that these are correlated and not random, i.e., a short interevent time tends to follow a short one and a long interevent tends to follow a long one. Furthermore, detrended fluctuation analysis (DFA) of the interevent interval

*zhangyongwen77@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

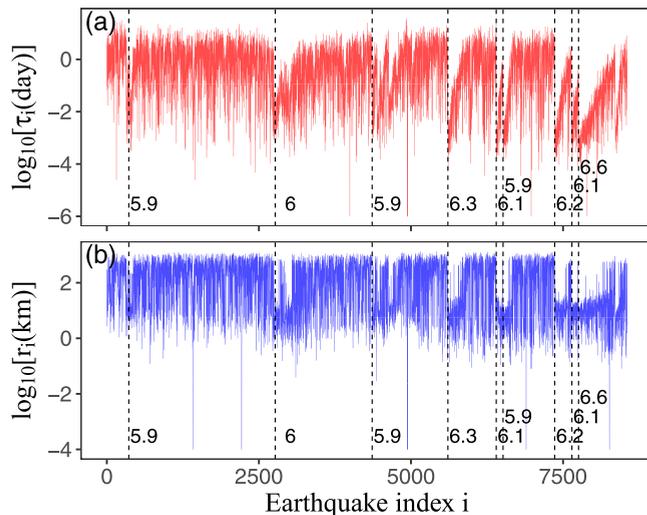


FIG. 1. Time series (1981–2017) of (a) interevent times and (b) their corresponding distances for the whole Italian catalog for a magnitude threshold of 3.0. The vertical dashed lines indicate earthquake events with magnitudes above 5.8.

time series revealed long-range (power-law) correlations [15]. In addition, memory has been found in the time series of earthquakes magnitudes [15,16]. The conditional probability method and DFA have been applied recently to study the memory in time series of consecutive interevent interval time series of real and ETAS model earthquake data [17].

The studies mentioned above [14–17] focused only on short-term memory, i.e., on successive interevent intervals of earthquakes. However, as shown below, there exists long-term memory in both interevent times and distances. Long-term memory in earthquakes has been found by previous studies using different techniques [18–20]. Here we study the long-term memory by considering the lagged conditional probabilities of interevent times and distances in real earthquake catalogs. Thus, we consider here not only the dependence of an interevent interval on the previous one (as in Refs. [14,17]) but also the conditional probability of an interevent interval depending on a previous (lagged) k th interevent intervals. Moreover, we study not only interevent times (as in previous studies) but also interevent distances.

II. MEMORY IN REAL SEISMIC CATALOGS

We start by analyzing the seismic catalog of Italy [21]. An interevent interval time τ_i is defined as the time interval between two consecutive earthquake events $\tau_i = t_{i+1} - t_i$ (in days) above a certain magnitude threshold. Following the Gutenberg-Richter law, the mean interevent time increases with the threshold magnitude. Similarly, we define an interevent distance r_i as the distance (in kilometers) between the locations of events $i + 1$ and i above a certain magnitude threshold. Figure 1 depicts the time series of interevent times and their corresponding distances for all of Italy. As can be seen, after the occurrence of a large earthquake, the interevent times decrease rapidly and then slowly increase, in agreement with the Omori law mentioned above [Fig. 1(a)]. Similarly, the interevent distances [Fig. 1(b)] also typically decrease fast

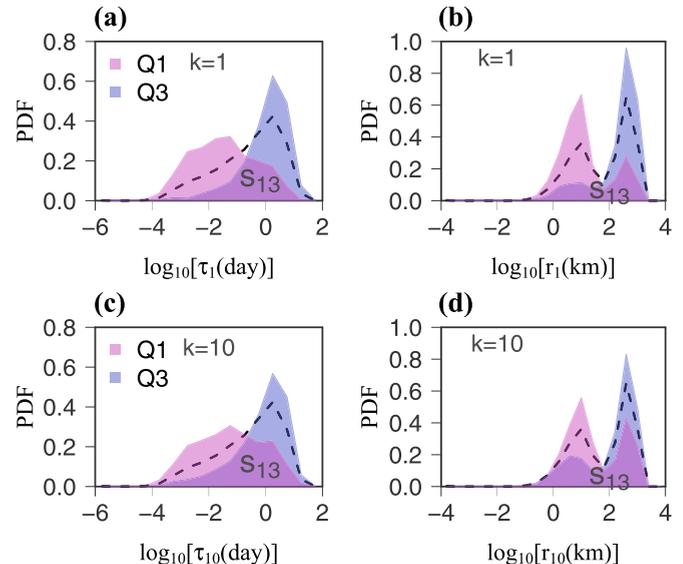


FIG. 2. Conditional PDF of (a) the interevent times τ_1 and (b) the interevent distances r_1 (for magnitude threshold $M_0 = 3.0$) for all of Italy; the common area between the conditional PDF of the first and third quantiles is $s_{13} = 0.43$ and $s_{13} = 0.42$ for times and distances, respectively. Note that no memory corresponds to $s_{13} = 1$ and full memory to $s_{13} = 0$. (c) and (d) Same as (a) and (b) but for lag $k = 10$ where the common area is now $s_{13} = 0.57$ and $s_{13} = 0.67$ for times and distances, respectively. A larger common area indicates less correlation. The black dashed curves indicate the PDFs for all τ (r). The PDFs are normalized in a logarithmic scale.

after a large earthquake but the following gradual increase observed in the interevent times [see Fig. 1(a)] is less apparent here. We find that similar behavior occur for two specific smaller areas in Italy (see Fig. S3 in [22]) and also for Japan and California [23] (Fig. S4 in [22]). Figure S5 in [22] depicts a scatter plot of interevent times vs interevent distances where two well separated blobs can be observed [see also Fig. 2(b)]. These two groups can be attributed to aftershocks (left blob, short distances) and main shocks (right blob, long distances) (see also [24,25]). Figure S5 in [22] shows that the interevent times and distances exhibit some dependence when considering different blobs (as shorter distances have shorter interevent time) but seem almost uncorrelated within each of the blobs.

To study possible long-term memory in the interevent times and distances of earthquakes, we introduce a lagged conditional PDF method. First, we sort all the interevent times (distances) in ascending order and then divide the sorted series into three equal quantiles. Thus, the first quantile Q1 contains the smallest 1/3 interevent times (distances) and the third quantile Q3 contains the largest 1/3 interevent times (distances). We define the conditional PDF of interevent times and distances as $\rho(\tau_k|\tau_0)$ [$\rho(r_k|r_0)$], where τ_0 (r_0) belongs to Q1 or Q3 and τ_k (r_k) is the lagged k th interevent time that follows τ_0 (r_0). Note that earlier studies [14,17] considered only the first lag ($k = 1$) and considered only interevent times (but not interevent distances). Figure 2(a) shows the lagged $k = 1$ conditional PDF $\rho(\tau_1|\tau_0)$ for the first (left peak) and third (right peak) quantiles. Both $\rho(\tau_1|\tau_0)$ for Q1 and Q3 are substantially different from the overall PDF $\rho(\tau)$ (indicated

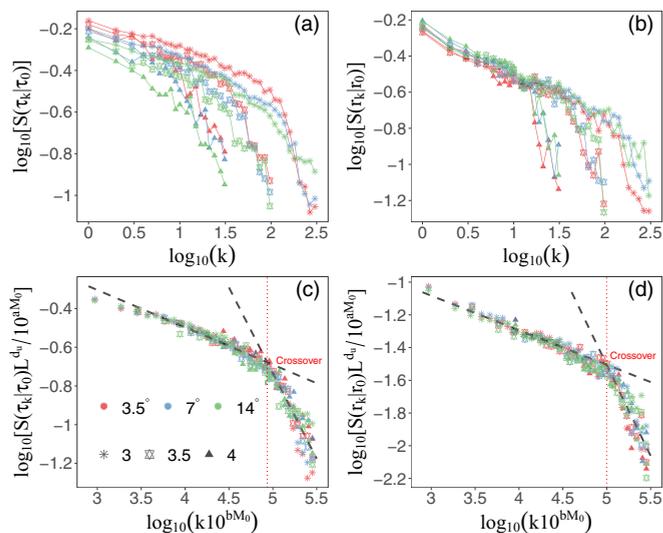


FIG. 3. Memory measure (a) $S(\tau_k|\tau_0)$ and (b) $S(r_k|r_0)$ as a function of the lag index k for interevent times and distances, respectively. Colors represent different grid sizes L . The shapes of the symbols represent different magnitude thresholds M_0 . The grid size of $L = 14^\circ$ covers the whole region of the Italian catalog. Rescaled memory measure for (c) interevent times and (d) distances. Black dashed lines are fitted power-law curves. Note that the narrow large- k regimes could be fitted also by exponential decay. The vertical red dotted lines indicate the location of the crossover.

by the dashed line). These results are consistent with the significant short-term memory of the nearest time intervals reported in previous studies [14,17]. In addition, we find significant memory for the interevent distances function $\rho(r_1|r_0)$ as shown in Fig. 2(b). Moreover and interestingly, we find significant long-term and long-range memory (correlations) even for large lags, e.g., for $k = 10$ [Figs. 2(c) and 2(d)] and for $k = 50$ and $k = 100$ (Fig. S6 in [22]). The significance of this type of long-term memory can be verified by comparing the conditional PDF to those of the randomly shuffled time series. Since the randomly shuffled time series contain no memory, $\rho(\tau_k|\tau_0)$ and $\rho(r_k|r_0)$ should be identical to the unconditional overall PDF of interevent times and distances, $\rho(\tau)$ or $\rho(r)$, as indeed shown in Fig. S7 in [22].

To quantify the level of memory expressed by the conditional PDF, we suggest as a measure the common area between the conditional PDFs of the smallest and largest quantiles Q1 and Q3. In the absence of memory the common area is one, while when the PDFs of the two quantiles are completely separated, the common area is zero. The common area is marked in Fig. 2 as s_{13} (dark area). Therefore, we define the level of memory to be $S(\tau_k|\tau_0) \equiv 1 - s_{13}$ in a range between 0 and 1 where large $S(\tau_k|\tau_0)$ indicates strong memory. Similarly, $S(r_k|r_0)$ represents the level of memory for the interevent distances.

Next we divide all of Italy into a grid of boxes of edge size L and construct the interevent times (distances) of the events within each box. Then, as described above, we obtain the memory measure $S(\tau_k|\tau_0)$ for all the boxes (grid points) of size L . The conditional PDF for the small size $L = 3.5^\circ$ is depicted in Fig. S8 in [22]. Figure 3(a) shows $S(\tau_k|\tau_0)$ for

different grid sizes L (different colors) and magnitude thresholds M_0 (different symbols). The memory in the small grid size (red or dark gray) is stronger than the memory of large grid size (green or light gray). The weaker memory for the larger grid size is due to the mixture of the weakly correlated events from remote locations with the nearby highly correlated events. The corresponding memory measure $S(r_k|r_0)$ for the interevent distances exhibits weaker dependence on the grid size [Fig. 3(b)]. Note that the memory measure S has two decay rates: It decays slowly for small k and faster for large k , for both the interevent times and interevent distances. Moreover, the crossover point for both is nearly the same.

Since the frequency of earthquakes decreases exponentially with magnitude in Fig. S1 in [22] (Gutenberg-Richter law), the interevent times grow exponentially with magnitude. Indeed, if we rescale k as $k \times 10^{bM_0}$ (where $b = 1$ as for the Gutenberg-Richter law) and multiply the memory measure S by $L^{d_u}/10^{aM_0}$, we obtain a single scaling function [Figs. 3(c) and 3(d)]. Thus, the rescaled memory measure $F(x)$ is

$$F(k \times 10^{bM_0}) = S(k)L^{d_u}/10^{aM_0}, \quad (1)$$

where $d_u = 0.14$ and $a = 0.09$ for the interevent times and $d_u = -0.08$ and $a = 0.24$ for the interevent distances. Thus, we are able to rescale the memory measure of different grid sizes and different magnitude thresholds into a single function (memory measure). The scaling parameters were obtained by minimizing the average of the standard variation in all bins in Figs. 3(a) and 3(b). The average of the standard variation for different parameter values is shown in Fig. S9 in [22].

It is apparent that the memory measure decreases for larger grid size L and $d_u > 0$ represents this decrease for the interevent times. For the interevent distances we find the opposite where $d_u < 0$, indicating that the memory measure actually increases with the grid size. The PDF for the large quantile Q3 is limited by the size L and as the grid size increases, the PDF of Q3 shifts to the right and the common area is reduced [Fig. 2(b)], indicating a stronger memory. We find a positive a , indicating that a larger magnitude threshold M_0 tends to have a stronger memory after k is rescaled. Note that the parameter a of interevent distances is larger than the corresponding parameter a of interevent times.

The scaling functions shown in Figs. 3(c) and 3(d) indicate a crossover between two distinct power-law relations. The scaling function is $F(x) \sim x^{-\gamma_1}$ for $x = k \times 10^{bM_0}$, where x is in the range of $[10^0, 10^5]$ and γ_1 is 0.19 and 0.21 for the interevent times and interevent distances, respectively. Both scaling functions exhibit a significant crossover at $x_c \approx 10^{5.0}$ and the approximate scaling exponent for large x (i.e., in range of $[10^5, 10^{5.5}]$) is $\gamma_2 = 0.88$ and 1.11 for interevent times and interevent distances, respectively. It is clear that $\gamma_1 \ll \gamma_2$ such that the decay for small scales is much slower than the decay for larger scales. Note that, due to the short range of large x $[10^5, 10^{5.5}]$, this range decay could be also fitted to an exponential decay. We cannot rule out that the crossover is due to a transition from a correlated (short-range) regime to an uncorrelated (long-range) regime where the transition is exponential.

Figure S10 in [22] shows the relation between the average (and rescaled average) time differences between two earthquakes and their lag k (and rescaled lag). It indicates that the

TABLE I. Estimated parameters a and d_u , power-law exponents of the scaling function γ_1 and γ_2 , and the crossover points x_c and t_c for the interevent times and distances for the Italy, Japan, and California earthquake catalogs.

Parameters	Italy	Japan	California
Time			
a	0.09	0.07	0.1
d_u	0.14	0.05	0.07
γ_1	0.19	0.09	0.2
γ_2	0.88	0.28	0.65
γ_2/γ_1	4.63	3.11	3.25
$\log_{10}(x_c)$	4.91	4.77	4.98
t_c	126 days	19 days	88 days
Distance			
a	0.24	0.2	0.2
d_u	-0.08	-0.07	-0.14
γ_1	0.21	0.23	0.35
γ_2	1.11	0.71	0.75
γ_2/γ_1	5.29	3.09	2.14
$\log_{10}(x_c)$	4.97	4.74	4.97
t_c	144 days	18 days	86 days

crossover x_c for the smaller grid size L corresponds to a longer time t_c . For all of Italy, the time t_c is around 130 days and it is the same for different magnitude thresholds.

To verify the generality of our scaling, we performed the same scaling analysis [using Eq. (1)] for the Japan and California earthquake catalogs and obtained good scaling. The scaling functions exhibit a crossover similar to that discussed above (Figs. S11 and 12 in [22]). The scaling parameters and exponents are slightly different for the different catalogs and are summarized in Table I. The rescaled memory measure F of the Japan catalog decays slower (as expressed in the smaller exponents γ_1 and γ_2) in comparison to the other locations, probably due to the high aftershock rate there. The rescaled lags of the crossover x_c are also listed in Table I. Note that the scaling curves of interevent times and distances have a similar crossover point x_c for the Italy, Japan, and California catalogs. Thus the crossover x_c seems to be unaffected by time and size scales of catalogs (see also Fig. S16 in [22]). It is likely that the crossover point x_c is related to the universal behavior of aftershocks. After a characteristic time, the sequences for independent and dependent events will overlap with high probabilities. This could destroy correlations leading to a fast decay of memory even when they exist in spatially separated but temporally overlapping sequences [11].

III. MEMORY IN THE ETAS MODEL

A good earthquake model should be able to reproduce the observed long-term and long-range memory features. Such a model could have the potential to significantly improve the forecasting capability of earthquakes. Thus, we next test the memory in the frequently used earthquake model, the epidemic-type aftershock sequence model (see Supplemental Material in [22]) [26,27]. The ETAS model can provide statistically some reliable forecasts of seismicity [28]. Figure S13 in [22] shows the conditional PDF of the interevent times and

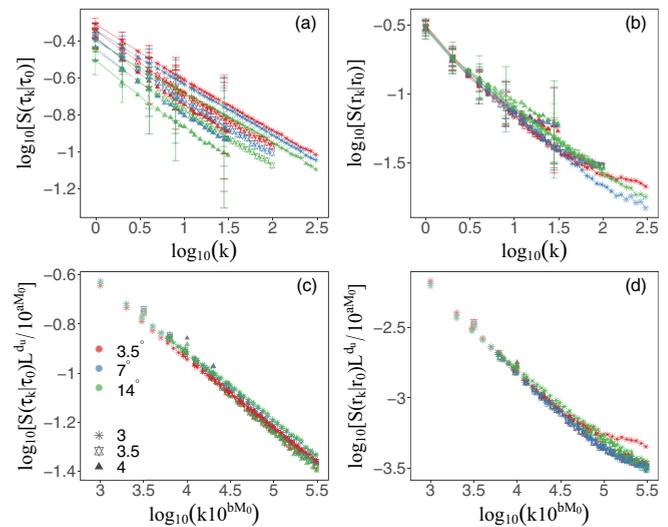


FIG. 4. Memory measure of the ETAS model catalogs for the area (34°N - 48°N , 6°E - 20°E) as a function of the lag index k : (a) $S(\tau_k|\tau_0)$ (interevent times) and (b) $S(r_k|r_0)$ (interevent distances). Colors and shapes are the same as in Fig. 3. Also shown is the rescaled memory measure for (c) interevent times and (d) distances. The model's parameters are estimated for the Italian catalog [17,29] (see also [22]). The memory measure S is averaged over 50 independent realizations and each realization includes 10^6 events. The error bars are the standard deviations.

distances and these exhibit much larger overlap in comparison to the memory in the data (Fig. 2), indicating weaker memory in the model. Figures 4(a) and 4(b) shows the memory measure for both interevent times and distances calculated from the ETAS model. As for the real catalogs [Figs. 4(c) and 4(d)], the memory measure of the model satisfies the scaling relation expressed in Eq. (1). The scaling parameters are $d_u = 0.16$ and $a = 0.14$ for the interevent times and $d_u = -0.04$ and $a = 0.55$ for the interevent distances. As expected, the grid size scaling leads to $d_u > 0$ for the interevent times and $d_u < 0$ for the interevent distances, similar to the real catalogs. Also, the magnitude threshold scaling parameter a is positive, as for the real catalogs; however, the value of a of the interevent distances in the model ($a = 0.55$) is two to three times larger than the value we obtain for the real catalogs ($a = 0.24$, Table I). Thus, the memory measure of the model is significantly smaller than that of the real catalog, even when using the optimal parameters for the ETAS model suggested in [17,29]. Moreover, the crossover power-law behavior at x_c that has been observed in the real catalogs (Fig. 3) cannot be seen in the model (Fig. 4). We thus conclude that the ETAS model does not reproduce the main memory features found in the present study for real catalogs.

To better understand which parameters control the power-law exponent and to find the optimal model's parameters that reproduce the characteristics of the real catalogs, we perform sensitivity tests by varying the parameters p and α for different choices of μ . We estimate the scaling exponents γ_1 and γ_2 for different regimes of k [similar to Figs. 3(c) and 3(d) and Table I]. The ratio between the high- and small- k exponents γ_2/γ_1 quantifies the level of crossover and the double

power-law behavior. For real catalogs the ratio γ_2/γ_1 is much larger than 1 and is about 5 for the interevent times and above 2 for the interevent distances. For the model, the results of the sensitivity tests for interevent times and distances are shown in Figs. S14 and 15 in [22]. It indicates that γ_1 and γ_2 depend on both p and α . They are almost unaffected by the choice of the background noise level μ . A small p and large α correspond to a small γ_1 and a large γ_2 at the bottom right corner; these exponents mimic the exponents of the data. Thus, the largest γ_2/γ_1 will be found at this corner too. Still, γ_2/γ_1 for the model is very different from the ratio for the real catalogs summarized in Table I; for the ETAS model, the largest γ_2/γ_1 is about 1.0, thus indicating the absence of crossover and double power-law behavior that was observed for the data (see Figs. S14 and 15 in [22]).

IV. CONCLUSION

In summary, we have proposed a lagged conditional PDF method and detected significant memory in both interevent times and distances for real data and the model. We proposed a

scaling function of this memory, for both interevent times and distances, in which the scaling curves of different magnitude thresholds and grid sizes collapse into a single curve. However, we found that the memory function in real data is very different from that of the ETAS model: The model's memory is weaker (stronger) in the short (long) timescale compared to the real catalogs. Moreover, the model does not exhibit the clear crossover observed in the real catalogs (see Figs. 3 and 4). The two power-law decays found here and the crossover between them may imply the existence of large and small aftershock productivity rates α , corresponding to short- and long-term memories of earthquakes.

ACKNOWLEDGMENTS

We thank Prof. Xiaosong Chen for helpful discussions. We thank the Italian Ministry of Foreign Affairs and International Cooperation, the Israeli Ministry of Science, Technology, and Space, the Israel Science Foundation, ONR, Japan Science Foundation, BSF-NSF, ARO, the EU H2020 project RISE, and DTRA (Grant No. HDTRA-1-10-1-0014) for financial support.

-
- [1] L. de Arcangelis, C. Godano, J. R. Grasso, and E. Lippiello, *Phys. Rep.* **628**, 1 (2016).
 - [2] B. Gutenberg and C. F. Richter, *Bull. Seismol. Soc. Am.* **34**, 185 (1944).
 - [3] T. Utsu, Y. Ogata, and R. S. Matsu'ura, *J. Phys. Earth* **43**, 1 (1995).
 - [4] D. Sornette, *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools* (Springer, Berlin, 2006), p. 395.
 - [5] T. Mäkinen, A. Miksic, M. Ovaska, and M. J. Alava, *Phys. Rev. Lett.* **115**, 055501 (2015).
 - [6] H. V. Ribeiro, L. S. Costa, L. G. A. Alves, P. A. Santoro, S. Picoli, E. K. Lenzi, and R. S. Mendes, *Phys. Rev. Lett.* **115**, 025503 (2015).
 - [7] J. Baró, Á. Corral, X. Illa, A. Planes, E. K. H. Salje, W. Schranz, D. E. Soto-Parra, and E. Vives, *Phys. Rev. Lett.* **110**, 088702 (2013).
 - [8] P. Bak, K. Christensen, L. Danon, and T. Scanlon, *Phys. Rev. Lett.* **88**, 178501 (2002).
 - [9] Á. Corral, *Phys. Rev. E* **68**, 035102(R) (2003).
 - [10] Á. Corral, *Phys. Rev. Lett.* **92**, 108501 (2004).
 - [11] S. Touati, M. Naylor, and I. G. Main, *Phys. Rev. Lett.* **102**, 168501 (2009).
 - [12] A. Saichev and D. Sornette, *Phys. Rev. Lett.* **97**, 078501 (2006).
 - [13] M. Bottiglieri, L. de Arcangelis, C. Godano, and E. Lippiello, *Phys. Rev. Lett.* **104**, 158501 (2010).
 - [14] V. N. Livina, S. Havlin, and A. Bunde, *Phys. Rev. Lett.* **95**, 208501 (2005).
 - [15] S. Lennartz, V. N. Livina, A. Bunde, and S. Havlin, *Europhys. Lett.* **81**, 69001 (2008).
 - [16] E. Lippiello, L. de Arcangelis, and C. Godano, *Phys. Rev. Lett.* **100**, 038501 (2008).
 - [17] J. Fan, D. Zhou, L. M. Shekhtman, A. Shapira, R. Hofstetter, W. Marzocchi, Y. Ashkenazy, and S. Havlin, *Phys. Rev. E* **99**, 042210 (2019).
 - [18] M. D. Zoback, R. M. Hamilton, A. J. Crone, D. P. Russ, F. A. McKeown, and S. R. Brockman, *Science* **209**, 971 (1980).
 - [19] M. T. Page and S. E. Hough, *Science* **343**, 762 (2014).
 - [20] J. Wang, I. G. Main, and R. M. Musson, *Geophys. J. Int.* **211**, 1005 (2017).
 - [21] The Italian earthquake catalog is provided by the Seismic Hazard Center at Istituto Nazionale di Geofisica e Vulcanologia (<http://terremoti.ingv.it/>).
 - [22] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.2.013264> for the details of the ETAS model and the following figures: Fig. S1, number of earthquakes versus the magnitude threshold for Italian earthquake catalog; Fig. S2, spatial distribution of Italian earthquake catalog; Fig. S3, time series of the interevent time and distance for the subareas; Fig. S4, time series of the interevent time and distance for Japan and California; Fig. S5, scatter plot of the interevent times versus distances; Fig. S6, conditional PDF for the lags $k = 50$ and $k = 100$; Fig. S7, conditional PDF of the shuffled data; Fig. S8, conditional PDF for $L = 3.5^\circ$; Fig. S9, average of the standard variation of the scaling curves; Fig. S10, average of time differences between two earthquakes with the lag k ; Fig. S11, memory measure for the Japan catalog; Fig. S12, memory measure for California catalog; Fig. S13, conditional PDF for the ETAS model; Figs. S14 and S15, sensitivity tests of γ_1 and γ_2 , respectively, for the ETAS model; Fig. S16, memory measure for the real catalog of Italy with the different time periods; and Fig. S17, memory measure for

- the simulated catalog generated by the ETAS model where the length of the simulated catalogs is the same as the real catalog of Italy.
- [23] The Japan catalog is the Japan Unified High-Resolution Relocated Catalog for Earthquakes from 2001 to 2012 (<http://www.hinet.bosai.go.jp/topics/JUICE/?LANG=en>) and the California catalog is provided by the California Integrated Seismic Network for 1970–2018 (<https://earthquake.usgs.gov/data/comcat/contributor/ci>).
- [24] I. Zaliapin, A. Gabrielov, V. Keilis-Borok, and H. Wong, *Phys. Rev. Lett.* **101**, 018501 (2008).
- [25] R. C. Batac and H. Kantz, *Nonlinear Process. Geophys.* **21**, 735 (2014).
- [26] Y. Ogata, *J. Am. Stat. Assoc.* **83**, 9 (1988).
- [27] Y. Ogata, *Ann. Inst. Stat. Math.* **50**, 379 (1998).
- [28] W. Marzocchi, M. Taroni, and G. Falcone, *Sci. Adv.* **3**, e1701239 (2017).
- [29] A. M. Lombardi, *Sci. Rep.* **5**, 8417 (2015).