

JGR Solid Earth

RESEARCH ARTICLE

10.1029/2019JB018530

Key Points:

- Two parameters determine the topology of triggering trees in the standard epidemic aftershock model
- The same parameters determine the classification of triggering trees into bursts and swarms
- Swarms can also represent false causal links in clusters of independent events

Supporting Information:

• Supporting Information S1

Correspondence to:

J. Baró, jbaro@crm.cat

Citation:

Baró, J. (2020). Topological properties of epidemic aftershock processes. Journal of Geophysical Research: Solid Earth, 125, e2019JB018530. https:// doi.org/10.1029/2019JB018530

Received 9 AUG 2019 Accepted 30 MAR 2020 Accepted article online 6 APR 2020

©2020. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Topological Properties of Epidemic Aftershock Processes

Jordi Baró¹

¹Centre for Mathematical Research, Barcelona, Spain

Abstract Earthquakes in seismological catalogs and acoustic emission events in lab experiments can be statistically described as point events in linear Hawkes processes, where the spatiotemporal rate is a linear superposition of background intensity and aftershock clusters triggered by preceding activity. Traditionally, statistical seismology interpreted these models as the outcome of epidemic branching processes, where one-to-one causal links can be established between mainshocks and aftershocks. Declustering techniques are used to infer the underlying triggering trees and relate their topological properties with epidemic branching models. Here, we review how the standard Epidemic Type Aftershock Sequence (ETAS) model extends from the Galton-Watson branching processes and bridges two extreme cases: Poisson and scale-free power law trees. We report the statistical laws expected in triggering trees regarding some topological properties. We find that the statistics of such topological properties depend exclusively on two parameters of the standard ETAS model: the average branching ratio n_{b} and the ratio between exponents α and b characterizing the production of aftershocks and the distribution of magnitudes, respectively. In particular, the classification of clusters into bursts and swarms proposed by Zaliapin and Ben-Zion (2013b, https://doi.org/10.1002/jgrb.50178) appears naturally in the aftershock sequences of the standard ETAS model depending on n_b and α/b . On the other hand swarms can also appear by false causal connections between independent events in nontectonic seismogenic episodes. From these results, one can use the memory-less Galton-Watson as a null model for empirical triggering processes and assess the validity of the ETAS hypothesis to reproduce the statistics of natural and artificial catalogs.

1. Introduction

The concept of aftershocks is traditionally associated with seismology (Utsu et al., 1995), but similar phenomena have been observed in other natural systems and are common in many mechanical processes in rocks, composites, and porous materials (Baró et al., 2013; Benioff, 1951; Davidsen et al., 2017; Hirata, 1987; Ribeiro et al., 2015). Aftershocks are identified in sequences of point events as a sudden increase of the activity causally linked to a previous event—usually stronger—called a mainshock. The empirical Omori-Utsu law (Utsu et al., 1995) describes the usual temporal evolution of the number of aftershocks ($\#_{AS}$) after time τ since a mainshock of magnitude m_{MS} as

$$\#_{\rm AS}(\tau|m_{\rm MS}) \approx \frac{\tilde{k}10^{m_{\rm MS}}}{(C+\tau)^p},\tag{1}$$

where *p* is a power law exponent usually close to 1, *C* a constant time representing a lower bound to the power law regime, and \tilde{k} a productivity factor. Additionally, aftershocks are spatially clustered, usually according to power law decay with distance to the mainshock $r := ||\mathbf{r} - \mathbf{r}_{MS}||$, where \mathbf{r} and \mathbf{r}_{MS} are the locations of the aftershock and the mainshock, respectively (Guo & Ogata, 1995).

Due to the complexity of the seismogenic mechanisms, statistical seismology considers all earthquakes mainshocks and aftershocks—as nonisolated stochastic events in space and time, rather than the outcome of deterministic mechanical processes. Mainshocks and aftershock sequences from different mainshocks coexist in the same regions and temporal windows. As consequence, all seismic events are customarily interpreted as point events in a single stochastic point process, determined by an intensity $\mu(t, \mathbf{r}, m)$ accounting for the instantaneous probability of finding an event defined by a mark—in this case, the earthquake magnitude *m*—at a spatial (\mathbf{r}) and temporal (t) location. Simple proportional hazard models consider events to be independent, with a space-time dependent μ (Varotsos et al., 1996). More sophisticated hazard models take into account the correlations between events caused by aftershock production (Michael, 1997; Ogata, 1998;





Figure 1. Schematic representation of a temporal sequence of aftershocks as events in a marked point process (top); the intensity can be inferred under the assumption of a linear Hawkes process (center); and interpreted as a branching process (bottom). Background events occur at depth = 0. Dark green circles represent leaves or singlets if depth = 0.

Turcotte et al., 2007; Vere-Jones & Ozaki, 1982; Zhuang et al., 2002). The most simple aftershock model is to consider a linear Hawkes self-exciting process (Hawkes & Oakes, 1974) incorporating the observations of statistical seismology. In a linear Hawkes process, the intensity can be expressed as the linear superposition of a background rate μ_0 and the individual contribution of all previous events {i} with a triggering term Ψ_i :

$$\mu(t, \mathbf{r}, m) = \mu_0(t, \mathbf{r}, m) + \sum_{i|t_i < t} \Psi_i(t, \mathbf{r}, m|t_i, \mathbf{r_i}, m_i),$$
(2)

where t_i , $\mathbf{r_i}$, and m_i are the time, position, and magnitude of the parent event *i*. Notice that the linear Hawkes model assumes an additive contribution for each event in the intensity, meaning that all events can generate aftershocks. This effective stochastic process is often interpreted as the result of an epidemic or branching process (Ogata, 1998; Saichev & Sornette, 2004; Turcotte et al., 2007; Vere-Jones & Davies, 1966). Figure 1 shows a schematic representation of a sequence of point events interpreted as the outcome of a history-dependent intensity, or Hawkes process, and its representation as a branching process. The branching process is strictly constituted by two distinct categories of events: background-independent events generated entirely by the back-

ground rate (μ_0 in equation (2)), and triggered events, caused by a unique preceding parent event *i*, through its individual and independent contribution to the intensity (Ψ_i in equation (2)), represented as arrows in Figure 1. Under the branching process assumption, earthquake catalogs are built as *branching forests* (bottom panel in Figure 1): linear superpositions of independent topological objects that here we call *triggering trees*. Each triggering tree (\mathcal{T}) is an ordered set of causally connected events initiated by a background event, the root of the tree (at depth = 0 in Figure 1). Such root can trigger a number $Z_{\mathcal{T}}(1)$ of events in a first generation of aftershocks (depth = 1 in Figure 1). In its turn, each event *i* of the $Z_{\mathcal{T}}(d)$ events in the *d*th generation of aftershocks can trigger a number $K_{\mathcal{T}}(i, d)$ of offspring events in the (d + 1)th generation of aftershocks. This (d + 1)th generation will total $Z_{\mathcal{T}}(d + 1) := \sum_{i \in [1, Z_{\mathcal{T}}(d)]} K_{\mathcal{T}}(i, d)$ events. We call *leaves* those events with no offspring ($K_{\mathcal{T}}(i, d) = 0$), extinguishing a branch (dark green circles in Figure 1). Singlets are background events which are also leaves, that is, have no offspring ($Z_{\mathcal{T}}(1) = 0$). The triggering tree is extinguished when all events in a maximum generational depth $D_{\mathcal{T}}$ are leaves.

Assuming the validity of the branching process approach, the identification of triggering trees would provide valuable topological information of the branching and enable the direct measurement of the triggering kernel Ψ_i (Davidsen et al., 2017; Zhuang et al., 2004). Hence, this representation of aftershock sequences is a useful approach, although one can argue whether it is strictly valid. While the stochastic point process resulting from the branching process can be represented as a linear Hawkes process, the linear Hawkes processes do not require the existence of explicit one-to-one causal links between events. In theory, univocal causal links in mechanical interactive systems can be defined from an energy stability point of view. This can be shown in micromechanical models such as the viscoelastic democratic fiber bundle model (Baró & Davidsen, 2018). In field studies, however, we have a limited capability to identify such a deterministic process. The bare statistical analysis of earthquake catalogs do not determine the explicit causal links but provide, instead, an assessment of its point process representation (2) where all terms contribute to μ with their specific weight. Advanced declustering techniques, either stochastic (Zhuang et al., 2004) or based on the nearest neighbor distance (Baiesi & Paczuski, 2004; Zaliapin & Ben-Zion, 2013a; Zaliapin et al., 2008) can be used to infer the most plausible one-to-one causal structure. This classification is never entirely free of uncertainty (Zhuang et al., 2002) but appears to be reliable when tested against synthetic catalogs (Zaliapin & Ben-Zion, 2013a). Notice that the topological concepts discussed in the following sections are only valid under the branching process assumption and have no correspondence to the more general point process description.

Recently, the analysis of seismological catalogs as branching processes revealed significant discrepancies between some reconstructed clusters or triggering trees, and the branching model expectations in Southern California (Zaliapin & Ben-Zion, 2013b). The same authors suggested a regional classification based upon such inconsistencies in the topological properties. Thereupon, aftershock sequences or clusters were

classified into two distinct categories: bursts and swarms. Burst-like clusters were defined as clusters of events with shallow generational depth (d), where most activity is accumulated in the first generation of aftershocks, mostly a consequence of significantly strong mainshocks. The burst-like activity was found to be compatible with the numerical results of the Epidemic Type Aftershock Sequence (ETAS) model, defined in the next section, by imposing the parameters fitted from field catalogs. Burst-like sequences were linked to tectonic settings with low heat flow. On the contrary, swarm-like clusters designate aftershock sequences with deeper generational depths, usually growing with the size of the swarm. The swarm-like activity was not predicted by the ETAS model with the fitted parameters and was found to match those regions with high heat flow. The results by Zaliapin and Ben-Zion (2013b) paved the road for a new analytical methodology based upon the topological statistics of triggering trees in local (Maghsoudi et al., 2016; Martínez-Garzón et al., 2018; Schoenball & Ellsworth, 2017; Trugman et al., 2017; Zaliapin & Ben-Zion, 2016b), regional (Vasylkivska & Huerta, 2017), and global (Zaliapin & Ben-Zion, 2016b) catalogs. This methodology will potentially lead to a better understanding of the seismogenic mechanisms behind triggering processes and improve the accuracy of stochastic point process models with potential applicability to hazard assessment(Field et al., 2014). Overall, the study of topological properties of triggering trees will foreseeably be gaining more popularity in the following years, thanks to the improved refinement of event detection techniques (Ross et al., 2017; Shelly et al., 2016). However, few works (Saichev & Sornette, 2004; Saichev et al., 2005) have addressed so far the topological properties expected from these hypothetical branching models, and a global picture of the model predictions remains incomplete.

Here, we will revisit some of the most common topological features used to characterize natural triggering for the standard ETAS model, defined here as a particular Galton-Watson branching model (section 1.1). We will illustrate numerically the results by Saichev et al. (2005) on the distribution of tree sizes (section 2.1) and briefly introduce a few general predictions regarding the topological magnitudes of family branching number and average leaf depths (section 2.2). Although most of these results are numerical, they can be useful as a benchmark to validate ETAS as a null hypothesis to natural and synthetic catalogs. The power law distributions found by Saichev et al. (2005) explain the diverse relationships between average leaf depths and tree sizes used to classify clusters (section 2.3). In particular, the conditions for an asymptotic equivalence between the standard ETAS model and random walks determine the transition between bursts and swarms, where both are solutions for aftershock sequences (section 3.1). Since the process is Galton-Watson, the expected average leaf depth given a cluster size is determined by the distribution of first-generation aftershocks, determined in turn by the ratio between two parameters of the model, α and b, and the average branching rate in the standard ETAS model. We will notice that swarm-like clusters can also be generated from a Galton-Watson process without productivity law, which can be constructed by misrepresenting causal links (section 3.2) and briefly discuss possible deviations from the Galton-Watson predictions in real processes (section 3.3).

1.1. Epidemic Aftershocks as Galton-Watson Branching Models

ETAS models are based on field and lab observations such as the empirical Omori-Utsu law (1) and the spatial distribution of aftershocks (Guo & Ogata, 1995; Ogata, 1998). Here we discuss the standard ETAS model (Ogata, 1988, 1998), simulating a Hawkes process where the triggering kernel is factorized in its dependencies as

$$\Psi_i(t, \mathbf{r}, m | t_i, \mathbf{r}_i, m_i) = \rho_m(m) \nu(m_i) \Psi_t(t - t_i) \Psi_r(||\mathbf{r} - \mathbf{r}_i||).$$
(3)

Both the temporal and spatial kernel are normalized to $\int_0^\infty \Psi_t(t - t_i)dt = 1$ and $\int_{\mathbb{R}^3} \Psi_r(||\mathbf{r} - \mathbf{r_i}||)d\mathbf{r} = 1$ and, therefore, have no effect on the topology of the triggering trees if we consider an infinite domain. The magnitudes (*m*) of the events are independent and identically distributed (i.i.d.) following the Gutenberg-Richter law (Gutenberg & Richter, 1944):

$$\rho_m(m) := 10^{-b(m-m_c)} / (b \log(10)), \tag{4}$$

where the magnitude of completeness m_c is an effective lower bound to the distribution. For the sake of simplicity, we consider m_c to be also the minimum magnitude able to generate aftershocks. We also assume that the offspring number K_i , representing the number of aftershocks generated by event *i* of magnitude m_i , is a Poisson number with a *specific branching ratio* $v(m_i)$ corresponding to the expected offspring number for events with magnitude m_i . Thus,

$$\mathbb{P}(K_i = k | v(m_i)) = \frac{v(m_i)^k e^{-v(m_i)}}{k!}.$$
(5)



The specific branching ratio is given by the aftershock production introduced as one of the factors constituting the triggering kernel in (3):

$$\nu(m_i) := \nu_c 10^{\alpha(m_i - m_c)}.$$
 (6)

This expression reproduces the mainshock-magnitude (m_i) dependence in the Omori-Utsu law (1) with a productivity exponent α usually found between 0.5 and 1 in the field (Utsu et al., 1995) and experiments (Baró et al., 2013; Davidsen et al., 2017). The term v_c normalizes the specific aftershock production for $m_i = m_c$. We can obtain an equivalent branching model in terms of v by performing a change of variables. Given the distribution $\rho(m_i)$ from (4),

$$\rho(\nu) = \rho((m_i(\nu)) \left| \frac{dm_i(\nu)}{d\nu} \right| = \begin{cases} 0 & \text{for } \nu < \nu_c \\ \frac{b}{\alpha} \frac{1}{\nu_c} \left(\frac{\nu}{\nu_c} \right)^{-\binom{b}{\alpha} + 1} & \text{for } \nu_c \le \nu \end{cases},$$
(7)

where $m_i(v)$ is obtained by inverting (6), α/b is bounded inside the range (0, 1], and v is sharply bounded by the condition $v > v(m_c) = v_c$. From now on we change the notation from v_c to the more convenient average branching ratio $n_b := \langle v \rangle = v_c \frac{b}{b-\alpha}$, equivalent to the expected offspring number for all events marginalizing out the parent magnitude m_i . The number of first-generation aftershocks for all events and all v is i.i.d. as

$$\mathbb{P}(K_i = k) = \int \rho(v) \mathbb{P}(K_i = k | v) dv = \frac{b}{\alpha} \frac{\left(n_b \left(1 - \frac{\alpha}{b}\right)\right)^{\overline{\alpha}}}{k!} \Gamma\left(k - \frac{b}{\alpha}, n_b \left(1 - \frac{\alpha}{b}\right)\right), \tag{8}$$

where $\Gamma(h, x) = \int_x^{\infty} t^{h-1} e^{-t} dt$ is the nonnormalized upper incomplete Gamma function, accounting for $v > v_c$. Therefore, from now on the standard ETAS model is simply considered as a particular case of a *Galton-Watson* (GW) *process* (Pitman, 2006), where all individual events *i*, background and triggered, have the same probability to trigger a number K_i of events defined by an offspring distribution $\mathbb{P}(K_i = k) := p_K(k)$, accounting for all variability in the i.i.d. values of m_i or v. In this case, for large *k*, this distribution can be approximated to a power law $p_K(k) \sim k^{-\gamma_k}$ with the exponent value $\gamma_k = b/\alpha + 1$ inherited from the *v* distribution $\rho(v)$ (7). Considering $0 < \alpha \le b$, this exponent value is constraint to $2 \le \gamma_k$. This result agrees with the reconstructed trees in seismology (Baiesi & Paczuski, 2004; Zaliapin & Ben-Zion, 2013a), where $b \approx \alpha$ and the distribution of the number of first-generation aftershocks—called degree distribution byBaiesi and Paczuski (2004)—is a power law with an exponent $\gamma_k \approx 2$.

A singular case of the ETAS model is found for $\alpha = 0$. In that case, the distribution (7) becomes a Dirac delta around v_c : $\lim_{r\to 0} \rho(v) = \delta(v - v_c)$, that is, $v(=v_c = n_b)$ is unique and all events have the same probability of generating aftershocks given by distribution (5) for all m_i . The ETAS model for $\alpha = 0$ is, thus, equivalent to a Poisson Galton-Watson (P-GW) branching process (Pitman, 2006).

Different modifications of the ETAS model have been proposed to account for more precise observations such as anisotropic spatial drift of the aftershock production and nonfactorizable magnitude dependencies (Ogata & Zhuang, 2006), generalized scaling forms (Davidsen & Baiesi, 2016; Vere-Jones, 2005) or more complex temporal decay forms (Baró & Davidsen, 2017; Davidsen et al., 2017). Such details are excluded from the following mathematical and numerical developments but will be recovered in the discussion section.

2. Results

Beyond the direct K_i distribution (8), we can evaluate the topological properties of the triggering trees resulting from the known or inferred causal links in an artificial or natural catalog. This section reports some predictions for the standard ETAS model regarding different topological magnitudes. The numerical results are obtained through the Monte Carlo generation of up to 10^7 aftershock sequences with $0 \le \alpha/b < 1$ and $0 < n_b < 1$. We use an arbitrary threshold $m_c = 1$. The background rate and the spatiotemporal kernel parameters are ignored since they play no role in this study.

2.1. Distribution of Tree Sizes

A fundamental concept for the topological characterization of the branching process is the *tree size* (S_{τ}) defined here as the total number of members in an extinguished tree: $S_{\tau} := \sum_{d=0}^{D_{\tau}} Z_{\tau}(d)$. Notice that, for $n_b > 1$, there is a nonzero probability of finding infinite trees $(P(D_{\tau} = \infty) > 0)$ with infinite S_{τ} . We only



Figure 2. (a) Tree size frequencies ($P(S_{\mathcal{T}} = s)$) and (b) scatter plot of $\langle d_l \rangle_{\mathcal{T}}$ against $S_{\mathcal{T}}$ of 10⁶ simulated P-GW trees for $n_b = 0.8, 0.9, 0.99$. (a) Dashed lines represent the expected Borel distribution ($p_S(s|n_b)$) 9 for the different n_b values. (b) Black line represents the expected relation $\langle \langle d_l \rangle_{\mathcal{T}} \rangle (s) \propto s^{0.5}$.

consider values $n_b \leq 1$ through all the discussion, imposing that all trees are extinguished at a finite time. The tree size is entirely determined by the K_i values of the events belonging to the tree. Since K_i values are i.i.d. no information is transfered between trees and S_{τ} is i.i.d. with a given $\mathbb{P}(S_{\tau} = s) := p_S(s)$. In particular, in the case of the P-GW with branching ratio n_b , obtained by imposing $\alpha = 0$, this distribution is known to be a Borel distribution (Pitman, 2006):

$$p_S(s|n_b) := \mathbb{P}(S_{\mathcal{T}} = s) = \frac{(n_b s)^{s-1} e^{-n_b s}}{s!}.$$
(9)

Figure 2a shows a numerical verification for different values of n_b and $\alpha = 0$. For large n_b and s, this distribution can be approximated by the exponentially tapered power law with an exponent $\gamma_s = 3/2$: $p_S(s|n_b) \sim s^{-3/2}e^{-s(1-n_b)^2/2}$.

Fifteen years ago, Saichev et al. (2005) found the mathematical expression for the distribution of tree sizes for the standard ETAS model. Their results can be numerically verified in artificial and natural catalogs by performing a maximum likelihood estimation of a power law exponent in different size intervals (Baró & Vives, 2012). An illustrative example is shown in Figure 3. The insets in Figure 3 show in log-log scale the distributions of S given $n_b = 0.995$ (a) and $n_b = 0.30$ (b) obtained for a set of α/b values, compared with the theoretical result for the P-GW expected for $\alpha = 0$ (thick orange line). Figure 3a shows the power law exponent (\hat{r}_s) estimated for $n_b = 0.99$ and $n_b = 0.30$ as function of α/b . As expected, the Borel distribution (9) is recovered for $\alpha \to 0$. The distribution for $0 < \alpha \leq b$ is different and can be summarized in the following observations: (i) For all values $0 \le \alpha/b < 1$ the original power law behavior in (9) is preserved below a characteristic s value $s_c \approx (1 - n_b)^{1/(\alpha/b-1)}$ (Saichev et al., 2005) for $n_b < 1$ and diverges for $n_b = 1$, but with a different exponent value with the general dependency $\gamma_s^{low} = \gamma_s^{low}(\alpha/b, n_b)$. (ii) This effective power law exponent observed below s_c increases from the $\gamma_s^{\text{low}}(\alpha/b = 0, n_b) = 1.5$, which is kept almost constant for $\alpha/b < 0.5$ toward higher values $\gamma_s^{\text{low}} \approx 2.0$ for $\alpha/b \approx 1$ (see Figure 3a). This can easily be observed by imposing a high n_b in the standard ETAS model. The thick golden line in Figure 3b represents the Borel distribution for $n_b = 0.99$, indistinguishable to a power law with γ_s^{low} within the range of observation. Since the characteristic scale is high ($s_c > 10^4$ for all α/b), the exponents are used as a proxy for γ_s^{low} . The present results for γ_s^{low} agree with Saichev et al. (2005), which predicted a sharp transition from $\gamma_s^{\text{low}} = 1.5$ for $\alpha/b < 0.5$ to $\gamma_s^{\text{low}} = 1 + \alpha/b$ for $0.5 \le \alpha/b \le 1$ (dashed green line in Figure 3a). (*iii*) The exponential regime for large trees found in the Borel distribution becomes fat tailed with a power law exponent $\gamma_s^{\text{high}} = 1 + b/\alpha$ for all the regime $0 < \alpha < b$, reminiscent of the large-scale regime in equations (7) and (8). This power law regime can be observed in the distributions for $n_b = 0.30$ represented in Figure 3c. The exponents for $n_b = 0.30$ are used as a proxy for γ_s^{high} since the selected estimation interval is considerably above s_c except for $\alpha = b$. (iv) For $\alpha = b$, both exponents coincide and s_c diverges. The power law behavior has a single exponent ($\gamma_s \approx 2.0$) with an infinite domain. This singular scale-free distribution is usually observed in field catalogs and incited the development of the more restrictive Branching Aftershock Sequence (BASS) model introduced by Turcotte et al. (2007). In this singular solution, the value of n_b only affects the proportion of singlets $\mathbb{P}(s = 0)$ (8), offsetting the distribution for all s > 0. This can be shown by comparing the faintest lines in Figures 3b and 3c.



Figure 3. (a) Estimated power law exponents $\hat{\gamma}_s$ for the distribution of tree sizes (*S*) assuming that for $s > s_{\min}$: $p_S(s|n_b) \sim s^{-\gamma_s}$. Results are shown for $n_b = 0.99$ (green triangles) and $n_b = 0.30$ (blue circles) and different values of a/b. Error bars show an estimate of σ of the likelihood function (Baró & Vives, 2012). The blue and green lines represent the results by Saichev et al. (2005). The dashed black line marks the value $\gamma_s = 3/2$ expected in a P-GW process. (b, c) Distribution of S_T for selected a/b values and (b) $n_b = 0.995$ and (c) $n_b = 0.30$ represented in integer exponential binning. The exponents are estimated by maximum likelihood within the interval $s_{\min} < S_T < \infty$ taking $s_{\min} = 10$ for $n_b = 0.995$ and $s_{\min} = 35$ for $n_b = 0.30$ (dashed vertical lines). Golden lines represent the result for a P-GW process with the same n_b (9).

2.2. Scalar Magnitudes Derived From p_K

Other common topological observables can be derived directly from the K_i distribution. Equation (8) establishes a relation between the ETAS parameters and the number of *singlets*, by imposing k = 0, that is, $\mathbb{P}(D_{\mathcal{T}} = 0) \equiv \mathbb{P}(S_{\mathcal{T}} = 0) = p_K(0)$. Given that the branching process is GW, the fraction or probability of singlets $p_K(0)$, or g(0) in Zaliapin and Ben-Zion (2013b), is equivalent to the average fraction of leaves ($\langle L_{\mathcal{T}}/S_{\mathcal{T}} \rangle$, where $S_{\mathcal{T}}$ is the cluster size and $L_{\mathcal{T}}$ the number of leaves). The $p_K(0)$ also determines a good approximation to the average *family branching number B* defined by Zaliapin and Ben-Zion (2013b) as "the average number of offspring over all earthquakes in the family that have at least one offspring." This definition is mathematically identical to $B := \langle \frac{S_{\mathcal{T}}-1}{S_{\mathcal{T}}-L_{\mathcal{T}}} \rangle$. For large clusters, $\frac{S_{\mathcal{T}}-1}{S_{\mathcal{T}}-L_{\mathcal{T}}} \approx (1 - L_{\mathcal{T}}/S_{\mathcal{T}})^{-1}$. Hence, in a GW process such as the standard ETAS: $B \approx (1 - p_K(0))^{-1}$. Notice that $p_K(0)$ and derived magnitudes depend on n_b and the ratio α/b but are essentially independent of the number of events $(S_{\mathcal{T}})$ in the standard ETAS model when the cluster is large enough.

2.3. Relations Between Generational Depth and Sizes

Finally, some topological properties of a GW process related to the generational depth of extinguished trees can be derived from the dualities between branching processes and random walks (Bennies & Kersting, 2000; Pitman, 2006). The concept of *Harris path* (Harris, 1951) is particularly useful for the topological analysis of aftershock sequences since it establishes a link between topological concepts such as size and depth with properties of a one-dimensional stochastic process. We consider a GW rooted tree \mathcal{T} with planar embedding. The Harris path follows a depth-first search of the vertexs of the tree counting the backtracking as new steps. The value d(j) is given by the generational depth of the vertex explored at time step *j*. Starting from the mainshock or root of the tree, d(j = 1) = 0, each vertex is explored $K_i + 1$ times before reaching d(j) = -1 at exactly $j = 2S_{\mathcal{T}}$. Leaves are identified as local maximums in the profile; that is, we find a leaf at time step *j* if d(j) = d(j - 1) + 1 and d(j + 1) = d(j) - 1. Each leaf is explored exactly once by the Harris path.

Let us consider, for a moment, that the tree is generated through a GW process with a geometric offspring distribution. This is given by $p_K(k|n_b) = n_b^{\ k}(n_b + 1)^{-k-1}$ in terms of n_b . At any time step *j*, the probability of finding a new child, adding a step $d \rightarrow d + 1$ in the Harris path, is independent of the number of previous offspring of the same parent *i* and, therefore, the Harris path is equivalent to a random walk (Harris, 1951).



Figure 4. (a) Relation between average leaf depths $(\langle d_l \rangle_T)$ and tree size (S_T) for $n_b = 0.99$. The scatter plot corresponds to trees sampled with $\alpha/b = 0.60$. Lines show the conditional averages $\langle \langle d_l \rangle_T | S_T \in (s, s + ds] \rangle$ in independent windows $(s = c^m, s + ds = c^{m+1}$ where c = 1.33 and $m \in (0, 1, 2, ...)$. Error bars represent σ of the conditional distribution. (b) estimated exponent \hat{r}_d of the power law relation (10) within $30 < S_T < 1,000$ for $n_b = 0.99$ (triangles) and $n_b = 0.50$ (circles) and different ratios α/b .

In particular, $2S_{\tau}$ is the return time of the random walk and the depth profile is a diffusion process that scales with the time step as $j^{1/2}$. All magnitudes associated with characteristic depths such as the maximum depth D_{τ} and the average leaf depths of each tree $(\langle d_l \rangle_{\tau})$ are expected to scale with the size of the tree as $S_{\tau}^{1/2}$. While the random walk analogy is only strictly valid in the geometric case, the Harris path of a GW with another offspring distributions is asymptotically equivalent to a random walk for large tree sizes as long as $n_b = 1$ or $n_b \leq 1$ and the variance is well defined ($0 < \sigma^2 < \infty$) (Aldous, 1991; Pitman, 2006). This approximation is valid for the P-GW process, but not for the ETAS model when $\alpha/b > 0.5$, rendering effective exponents values $\gamma_k < 3$ for large trees in equation (8) and, hence, infinite variance.

Figure 2b shows the numerical validation of the diffusive assumption in the case of the P-GW, that is, a standard ETAS with $\alpha = 0$. On average, the dependence between $\langle d_l \rangle_{\tau}$ and S_{τ} follows the relation $\langle \langle d_l \rangle_{\tau} \rangle (s) := \langle \langle d_l \rangle_{\tau} | S_{\tau} = s \rangle \propto s^{0.5}$, which stands even for significantly low values of n_b . Although not shown here, a proportional relationship is also found in other characteristic depths such as the average depth of all events $\langle \langle d \rangle | S_{\tau} = s \rangle \propto s^{0.5}$ and maximum depth $\langle D_{\tau} | S_{\tau} = s \rangle \propto s^{0.5}$. Notice that, similarly as how a stochastic process cannot diffuse faster than a ballistic trajectory, the maximum depth cannot be larger than the tree size, forcing the limit $\langle d_l \rangle_{\tau} < S_{\tau}$ which biases the average values for small trees (usually $S_{\tau} < 20$ as seen in Figure 2b).

The topological relations between depths and sizes are more complex in the ETAS model due to the power law distribution (8). Here we only introduce the numerical results and leave the mathematical derivation, if possible, as an open question for future works. Figure 4a shows the conditional average $\langle \langle d_l \rangle_T | S_T \in (s, s + ds] \rangle$ for different parameters α/b . A scatter plot of the data pairs $(\langle d_l \rangle_T, S_T)$ for $\alpha/b = 0.6$ and $n_b = 0.995$ is also shown as an example. The distribution of small trees exhibits the aforementioned upper bound $\langle d_l \rangle_T < S_T$, biasing the bivariate distribution for $S_T \leq 20$. For larger trees, an average power law relation is observed beyond statistical fluctuations:

$$\langle \langle d_l \rangle_{\mathcal{T}} \rangle(s) \propto s^{\gamma_d},$$
 (10)

with an exponent $\gamma_d = \gamma_d(\alpha/b, n_b)$. This power law relation gets distorted for low values of n_b , although one is still able to estimate a reliable power law relation for a shorter range. The inset in Figure 4b shows the exponent values γ_d fitted for trees with $30 < S_T < 1,000$. The exponent appears to be consistent with the typical GW process ($\gamma_d = 0.5$) for the relatively broad range $0 < \alpha/b < 0.3$ but drops toward lower exponent values for larger ratios. This first regime for low α/b values is consistent with the asymptotic limit since the effective power law exponent $\gamma_k > 3$ in the asymptotic form of equation (8). For $0.5 < \alpha/b < 1$, the drift from the diffusive relationship is apparent for all values of n_b . Close to $\alpha/b = 1$, the depth of triggering trees



appears to be independent of S_{τ} . The branching ratio n_b changes the range of the distribution in S_{τ} as well as the dependence on the bivariate distribution. For low n_b values, large events tend to have shallower trees than predicted by the diffusive relationship. As a consequence, the effective exponent γ_d depends on n_b for intermediate values of α/b but almost coincide for the extreme cases $\alpha < 0.3b$ and $\alpha \approx 1$. The exponent transition is steeper for low n_b values.

3. Discussion

The numerical results and nonformal arguments presented here highlight three direct measurements that stand out to characterize the topology of epidemic aftershock processes: the direct triggering distribution $p_K(k)$, the relation between $\langle d_l \rangle_{\tau}$ and S_{τ} , and the size distribution $p_S(s)$. Their joint analysis can verify whether the relation between $\langle d_l \rangle_{\tau}$ and S_{τ} corresponds to the expected behavior of the specific branching model hypothesis and an ETAS model with independently fitted parameters v_0 or n_b , b, and α . Scalar measurements such as the fraction of singlets or leaves, the family branching number (B), and the bulk average leaf depth ($\langle \langle d_l \rangle_{\tau} \rangle$), which entangle both dependencies in n_b and α/b , can mislead the validation or fitting of a branching model. for example, a high fraction of singlets ($p_K(0)$) or family branching number (B) can either be caused by a high α/b ratio or a low n_b (see equation (8)). However, the expected relationships between such scalar measurements, such as the fraction of leaves and singlets, can be used to develop preliminary tests to reject the standard ETAS hypothesis before undertaking any parameter fitting.

The ETAS model can potentially explain the topological variability of triggering trees reconstructed from seismological catalogs (Zaliapin & Ben-Zion, 2013b) and acoustic emission experiments (Davidsen et al., 2017). Provided the validity of the branching model approach, unequivocal relations exist between the topological properties and the parameters of aftershock production, determined by n_b and α/b in the ETAS model (8). Assuming independence between the values of ν —guaranteed in the present study because of the independence of m—all topological information is contained in the distribution of ν (as equation (7) in the ETAS). In general, the same relations can be extrapolated to other GW models with well defined $p_K(k)$. Beyond strict GW processes, the same results would be locally valid in models implementing spatiotemporal variations of the ETAS parameters such as the spatially variating ETAS model introduced by Nandan et al. (2017). In that case, the topological properties would depend on the local distribution of n_b and α/b . A global evaluation would report a spread in the ($\langle d_l \rangle_{\tau}, S_{\tau}$) space with a blurred power law relationship, as the ones reported by Zaliapin and Ben-Zion (2013b) in southern California.

3.1. Bursts and Swarms as Clusters of Causally Connected Events

The ETAS model is a peculiar case of GW processes where the aftershocks sampling is power law(Saichev et al., 2005), leading to a natural cluster classification in the terms presented by Zaliapin and Ben-Zion (2013b). The results shown in Figure 4 validate the hypothesis that swarms, like bursts, can appear as a consequence of event-event triggering processes, that is, aftershocks, represented as one-to-one causal links in branching processes (Zaliapin & Ben-Zion, 2013b). The topological properties of the trees used for the classification of swarms and bursts—and, in particular, the exponent γ_d —differ depending on the parameter ratio α/b and the branching ratio n_b . Such classification is noticeably sharp in the parameter space for low n_b (Figure 4b) but smooth for high branching ratios ($n_b \leq 1$). The two classes are found in the extreme cases $\alpha = b$ and $\alpha = 0$:

For $\alpha \approx b$, only strong events are likely to generate aftershocks. Because the branching ratio is fixed, most of the triggered activity for $\alpha \approx b$ is due to the few stronger events, which, as consequence, are more likely to be background events. The first-generation offspring of this strong event is unlikely to generate aftershocks of their own, rendering spray-like short tree sequences and star-shaped spatial structures characteristic of burst-like activity (Zaliapin & Ben-Zion, 2013b). As pointed out in Zaliapin and Ben-Zion (2013a), this result is also consistent with the tree size distribution in the ETAS with $\alpha/b \approx 1$ (Saichev et al., 2005) or BASS (Turcotte et al., 2007).

For $\alpha < 0.5b$, the P-GW limit is recovered and trees grow as swarms, forming relatively slender tree structures, with $\gamma_d \approx 0.5$. A clear example of swarm-like aftershocks is given by the reconstructed triggering trees from the ultrasonic signals recording during the triaxial compression of sandstones (Davidsen et al., 2017). In this specific case, an ad-hoc ETAS model can be fitted with an effective ratio $\alpha/b \sim 0.5$ (Baró & Davidsen, 2017). Notice that the distribution of tree sizes reported by (Zaliapin & Ben-Zion, 2016a) in hot areas fits a steep power law, which could match a $\gamma_s^{high} \approx 3$ if the data is above s_c , typically low for $\alpha \ll b$.



Although the transition is not sharp in the parameter space, the empirical α/b values are typically close to the ones in seismological catalogs. Hence the separation between $\alpha/b \approx 1$ and $\alpha/b < 1$, although fuzzy, is a natural choice to define the classification between burst-like and swarm-like clusters. The GW model does not expect values $\gamma_d > 0.5$. Significantly higher exponent values might indicate memory in the branching process, which cannot be modeled as a GW.

This result also sets an alternative classification of swarms and clusters based on the ETAS assumption. Size distributions of bursts will typically exhibit a power law with exponent $\gamma_s \approx 2$. Swarms, instead, will typically exhibit higher exponents for low n_b , or lower ones at small s, down to $\gamma_s = 1.5$, if n_b is close to one. This implicit relation between γ_s and γ_d is consistent with the data splitting in areas of low and high heat flow. The exponent of the tree size survival distribution ($\gamma_s - 1$) was observed to transition from an equivalent $\gamma_s \approx 2$ for low heat flow areas with shallow tree depths to $\gamma_s \approx 3$ for high heat flow with deep trees (Zaliapin & Ben-Zion, 2016a, 2016b).

3.2. Swarms as Clusters of Independent Events

Overall, the ETAS model establishes a clear relationship between the topological properties and the ratio between the Gutenberg-Richter exponent b and the productivity exponent α . Swarms, however, are resulting particularly problematic when validating the ETAS model through this relationship. Early results by Zaliapin and Ben-Zion (2013b) had shown how the ETAS model parameterized with field estimations failed to predict the observed topology of swarms in southern California, where $\alpha \approx b$. According to later analysis, such regions exhibit $\gamma_s \sim 3$ during fluid injection operations (Zaliapin & Ben-Zion, 2016b), also inconsistent with $\alpha \approx b$ in the ETAS model. This apparent contradiction could be explained by local variations in the α/b ratio, hidden below a resolution limit. Another plausible explanation is the misallocation of causal links during the cluster identification. Tree reconstruction techniques are based on the contraposition of observations against the predictions from a nonhomgeneous Poisson null hypothesis (Zaliapin et al., 2008). Such methods are able to deal with factorizable spatial and temporal variations in the background activity, that is, $\mu_0(t, \mathbf{r}) = \mu_r(\mathbf{r})\mu_t(t)$. Causal links are identified from a rescaled spatiotemporal distance and by comparing the results with either (i) a bimodal distribution fitting (Zaliapin & Ben-Zion, 2016a), (ii) the result of catalog surrogates (Davidsen et al., 2017), or (iii) the local spatial variations in rate (Zaliapin & Ben-Zion, 2013a, 2016b). However, these methods can have problems when both spatial and temporal variations occur simultaneously, that is, when $\mu_0(t, \mathbf{r})$ has a nonfactorizable dependence in both space and time (see supporting information). Such variations can be caused by exogenous geological or anthropogenic processes introducing short but intense seismogenic episodes compared long term tectonic processes. This can be the case, for example, of episodic volcanic (Roberts et al., 2016), natural geothermal (Cheng & Chen, 2018; Gaeta et al., 1998), or human-induced (Ellsworth, 2013; Martínez-Garzón et al., 2018; Schoenball & Ellsworth, 2017; Trugman et al., 2017) seismicity, or even tectonic seismicity in the presence of seasonal variations (Ueda & Kato, 2019). A clear observation of such bimodality arising from the separation of tectonic and episodic fluid induced microseismicty was reported by Vasylkivska and Huerta (2017) in Oklahoma earthquakes. Spatiotemporal correlations in such settings are not necessarily a consequence of a history dependence and might disrupt the performance of cluster identification techniques, which will overestimate triggering relationships. The random assignment of causal links in a set of uncorrelated events could lead to the generation of swarm-like clusters, which could occur due to the aforementioned limitations in the declustering method (see supporting information). Notice that the $\langle d_l \rangle_T - S_T$ relation observed in swarms is compatible with $\langle \langle d_l \rangle_{\mathcal{T}} | S_{\mathcal{T}} \rangle \sim S_{\mathcal{T}}^{0.5}$ (e.g., see Figure 4 in Zaliapin & Ben-Zion, 2013b). In particular, geothermal systems and human-induced seismicity reproducing P-GW processes are suspicious and might not always represent actual triggering, but rather exogenous variations in the background rate. Studies focused on well confined episodes of seismic activity, such as (Zaliapin & Ben-Zion, 2016b), are advised to validate the results.

3.3. Branching Models Beyond GW

Discrepancies between model and data might be corrected by modifications of the ETAS models implementing more sophisticated field observations. The presence of magnitude-magnitude correlations has been reported in earthquake catalogs (Lippiello et al., 2007, 2008). It is a matter of debate if such correlations are consequence of an observation incompleteness (Davidsen & Green, 2011; Davidsen et al., 2012) or might potentially have a physical origin (Maghsoudi et al., 2016). In any case, reconstructed triggering trees might be affected by such correlations, introducing deviations from the GW predictions. Depth-dependent *m* distributions would as well fall outside the GW category. In particular, recent observations suggest a simple distinction in the Gutenberg-Richter exponent for aftershocks (b_{AS}) and mainshocks (b) (Davidsen & Baiesi, 2016; Davidsen et al., 2017; Gu et al., 2013).

Finally, any empirical study on the topological properties should account for the fact that natural catalogs have a limited spatial and temporal range. Even if the branching approach is valid, the concept of leaf and root are ill-defined in spatially confined or finite time series (Zhuang et al., 2002). Events misclassified as mainshocks might be actually triggered by remote or ancestral events (van der Elst, 2017) and apparent leaves might generate aftershocks outside the observational range. The results can be especially biased for the analysis of short catalogs, considering the power law kernels in Ψ_t and Ψ_r of the ETAS model.

4. Conclusions

Here, we have considered the branching model as a valid representation of aftershock sequences and other triggering processes. The reconstruction of triggering trees accounting for all event-event correlations opens new perspectives to learn about the seismogenic mechanisms behind aftershocks and improve our current forecasting techniques. We have revisited, and added to, the expected topological properties for the standard ETAS, which is interpreted as a fat-tailed Galton-Watson process, extended from the Poisson Galton-Watson process (P-GW), which is recovered as a particular case. This list of properties had been proposed in previous works as helpful observables to characterize and classify aftershock sequences. Specifically, this analysis serves to distinguish between swarms and bursts (Zaliapin & Ben-Zion, 2013b, 2016a).

All the singular properties of the ETAS model within the GW category derive from the power law distribution of branching ratios v leading to a similar power law in $p_k(k)$. In particular, all topological properties depend only on two parameters: the average branching ratio n_b and the ratio between exponents α/b . Since the P-GW is recovered as a limiting case of the ETAS model and the BASS as the opposite limit, a transition must exist in the distribution of tree sizes (S_{τ}) . The n_b controls the characteristic size s_c marking the transition between two distinct power law exponents (γ_s^{low} and γ_s^{high}), which coincide to $\gamma_s^{\text{low}} = \gamma_s^{\text{high}} = 2$ for $\alpha = b$, and recovers the Borel distribution for $\alpha = 0$. Characteristic generational depths are strongly dependent on α/b once the power law tail has a significant statistical weight ($\gamma_k < 3$). The average leaf depth of a tree ($\langle d_l \rangle_T$) has a power law dependence with S_{τ} for high branching ratios or low exponent ratios α/b . The exponent of this power law relation coincides with $\gamma_d = 0.5$, typical in well-defined GW processes, for $\alpha = 0$, and decreases for values $b \leq 2\alpha$, vanishing to 0 at the limit $\alpha = b$. For low branching ratios this transition in the exponent gets sharper and occurs at lower values, for example, $b/\alpha \sim 0.3$ for $n_b = 0.5$. This numerical result interprets the separation between bursts and swarms as a phenomenological observation based upon the common values of $\alpha \approx b$ found in natural seismicity. A regional analysis of the ratio is required to validate the exponent dependence on the α/b ratio. In any case, the results in the ETAS model prove that the topological structure of swarms, as bursts, can be explained as an event-event triggering processes, that is, aftershocks, represented by one-to-one causal links. We also noticed that the topological properties of swarms are the natural result of associating false causal links in a catalog of independent events. Such false causal links could appear after an inappropriate use of the current cluster identification techniques.

The present results are intended to serve as a reference to validate the ETAS model as a null hypothesis to describe the triggering processes associated with tectonic and induced seismicity, as well as acoustic emission experiments (Baró et al., 2013; Benioff, 1951; Costa et al., 2016; Davidsen et al., 2017; Hirata, 1987; Ribeiro et al., 2015), and micromechanical models (Baró & Davidsen, 2018; Dieterich, 1994; Hainzl et al., 1999; Jagla & Kolton, 2010; Lyakhovsky et al., 2005; Yamashita & Knopoff, 1987; Zhang & Shcherbakov, 2016). A rejection of the ETAS model from the topological properties of the triggering trees might indicate more sophisticated epidemic processes, involving magnitude-magnitude correlations or depth dependencies. On the contrary, the validation of the ETAS model would set a step forward in the testing and development of new micromechanical models implementing seismogenic mechanisms of aftershocks and other event-event triggering mechanisms.

Acknowledgments Thanks to A. Corral, K. Wiese, and I.

Zaliapin for fruitful discussions. The author acknowledges the Associate Editor and two anonymous referees for their time and helpful remarks. Thanks to AXA Research Fund for financial support through the project RheMechFail. Data files required to generate the results are archived and available at (Barć, 2020).

References

Aldous, D. (1991). The continuum random tree II: An overview. Stochastic analysis, 167, 23-70.

Baiesi, M., & Paczuski, M. (2004). Scale-free networks of earthquakes and aftershocks. *Physical Review E*, 69, 66,106. https://doi.org/10. 1103/PhysRevE.69.066106

Baró, J., & Davidsen, J. (2017). Are triggering rates of labquakes universal? Inferring triggering rates from incomplete information. The European Physical Journal Special Topics, 226, 3211–3225. https://doi.org/10.1140/epjst/e2017-70072-4 Baró, J. (2020). Branching trees from standard Epidemic Aftershock Sequences (ETAS) model (no time, no space). Zenodo. https://doi.org/ 10.5281/zenodo.3628664

Baró, J., Corral, Á., Illa, X., Planes, A., Salje, E. K. H., Schranz, W., et al. (2013). Statistical similarity between the compression of a porous material and earthquakes. *Physical Review Letter*, 110, 88702. https://doi.org/10.1103/PhysRevLett.110.088702

Baró, J., & Davidsen, J. (2018). Universal avalanche statistics and triggering close to failure in a mean-field model of rheological fracture. *Physical Review E*, 97, 33002. https://doi.org/10.1103/PhysRevE.97.033002

Baró, J., & Vives, E. (2012). Analysis of power-law exponents by maximum-likelihood maps. *Physical Review E*, 85, 66,121. https://doi.org/ 10.1103/PhysRevE.85.066121

Benioff, H. (1951). Earthquakes and rock creep: (Part I: Creep characteristics of rocks and the origin of aftershocks). Bulletin of the Seismological Society of America, 41(1), 31–62.

Bennies, J., & Kersting, G. (2000). A random walk approach to Galton–Watson trees. Journal of Theoretical Probability, 13(3), 777–803. https://doi.org/10.1023/A:1007862612753

Cheng, Y., & Chen, X. (2018). Characteristics of Seismicity inside and outside the Salton Sea Geothermal Field. Bulletin of the Seismological Society of America, 108(4), 1877–1888. https://doi.org/10.1785/0120170311

Costa, L. S., Lenzi, E. K., Mendes, R. S., & Ribeiro, H. V. (2016). Extensive characterization of seismic laws in acoustic emissions of crumpled plastic sheets. EPL (Europhysics Letters), 114(5), 59002. https://doi.org/10.1209/0295-5075/114/59002

Davidsen, J., & Baiesi, M. (2016). Self-similar aftershock rates. *Physical Review E*, 94, 22,314. https://doi.org/10.1103/PhysRevE.94.022314
Davidsen, J., & Green, A. (2011). Are earthquake magnitudes clustered? *Physical Review Letters*, 106, 108502. https://doi.org/10.1103/
PhysRevLett.106.108502

Davidsen, J., Kwiatek, G., Charalampidou, E.-M., Goebel, T., Stanchits, S., Rück, M., & Dresen, G. (2017). Triggering processes in rock fracture. *Physical Review Letters*, 119, 68501. https://doi.org/10.1103/PhysRevLett.119.068501

Davidsen, J., Kwiatek, G., & Dresen, G. (2012). No evidence of magnitude clustering in an aftershock sequence of nano- and picoseismicity. *Physical Review Letters*, 108, 38,501. https://doi.org/10.1103/PhysRevLett.108.038501

Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research*, 99(B2), 2601–2618. https://doi.org/10.1029/93JB02581

Ellsworth, W. L. (2013). Injection-induced earthquakes. Science, 341(6142). https://doi.org/10.1126/science.1225942

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., et al. (2014). Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3)—The Time-Independent Model Uniform California Earthquake Rupture Forecast, Version 3 (UCERF3) The Time Independent Model. *Bulletin of the Seismological Society of America*, 104(3), 1122–1180. https://doi.org/10.1785/0120130164

Gaeta, F. S., De Natale, G., Peluso, F., Mastrolorenzo, G., Castagnolo, D., Troise, C., et al. (1998). Genesis and evolution of unrest episodes at Campi Flegrei caldera: The role of thermal fluid-dynamical processes in the geothermal system. *Journal of Geophysical Research*, *103*(B9), 20,921–20,933. https://doi.org/10.1029/97JB03294

Gu, C., Schumann, A. Y., Baiesi, M., & Davidsen, J. (2013). Triggering cascades and statistical properties of aftershocks. Journal of Geophysical Research: Solid Earth, 118, 4278–4295. https://doi.org/10.1002/jgrb.50306

Guo, Z., & Ogata, Y. (1995). Correlation between characteristic parameters of aftershock distributions in time, space and magnitude. Geophysical Research Letters, 22(8), 993–996. https://doi.org/10.1029/95GL00707

Gutenberg, B., & Richter, C. F. (1944). Frequency of earthquakes in california. Bulletin of the Seismological Society of America, 34(4), 185-188.

Hainzl, S., Zöller, G., & Kurths, J. (1999). Similar power laws for foreshock and aftershock sequences in a spring-block model for earthquakes. *Journal of Geophysical Research*, 104(B4), 7243–7253. https://doi.org/10.1029/1998JB900122

Harris, T. E. (1951). First passage and recurrence distributions (Tech. Rep.): RAND CORP SANTA MONICA CA.

Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, *11*(3), 493–503. Hirata, T. (1987). Omori's power law aftershock sequences of microfracturing in rock fracture experiment. *Journal of Geophysical Research*, *92*(B7), 6215–6221. https://doi.org/10.1029/JB092iB07p06215

Lippiello, E., de Arcangelis, L., & Godano, C. (2008). Influence of time and space correlations on earthquake magnitude. Physical Review Letters, 100, 38501. https://doi.org/10.1103/PhysRevLett.100.038501

Lippiello, E., Godano, C., & de Arcangelis, L. (2007). Dynamical scaling in branching models for seismicity. *Physical Review Letters*, 98, 98501. https://doi.org/10.1103/PhysRevLett.98.098501

Lyakhovsky, V., Ben-Zion, Y., & Agnon, A. (2005). A viscoelastic damage rheology and rate-and state-dependent friction. *Geophysical Journal International*, 161(1), 179–190.

Maghsoudi, S., Eaton, D. W., & Davidsen, J. (2016). Nontrivial clustering of microseismicity induced by hydraulic fracturing. *Geophysical Research Letters*, 43, 10,672–10,679. https://doi.org/10.1002/2016GL070983

Martínez-Garzón, P., Zaliapin, I., Ben-Zion, Y., Kwiatek, G., & Bohnhoff, M. (2018). Comparative study of earthquake clustering in relation to hydraulic activities at geothermal fields in California. *Journal of Geophysical Research: Solid Earth*, 123, 4041–4062. https://doi.org/ 10.1029/2017JB014972

Michael, A. J. (1997). Testing prediction methods: Earthquake clustering versus the poisson model. *Geophysical Research Letters*, 24(15), 1891–1894. https://doi.org/10.1029/97GL01928

Nandan, S., Ouillon, G., Wiemer, S., & Sornette, D. (2017). Objective estimation of spatially variable parameters of epidemic type aftershock sequence model: Application to california. *Journal of Geophysical Research: Solid Earth*, 122, 5118–5143. https://doi.org/10.1002/ 2016JB013266

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association, 83(401), 9–27. https://doi.org/10.1080/01621459.1988.10478560

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics, 50(2), 379–402. https://doi.org/10.1023/A:1003403601725

Ogata, Y., & Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics*, 413(1), 13–23. Retrieved from http:// www.sciencedirect.com/science/article/pii/S0040195105004889 (Critical Point Theory and Space-Time Pattern Formation in Precursory Seismicity) https://doi.org/10.1016/j.tecto.2005.10.016

Pitman, J. (2006). Combinatorial stochastic processes: Ecole d'eté de probabilités de saint-flour xxxii-2002. Berlin: Springer-Verlag. https:// doi.org/10.1007/b11601500

Ribeiro, H. V., Costa, L. S., Alves, L. G. A., Santoro, P. A., Picoli, S., Lenzi, E. K., & Mendes, R. S. (2015). Analogies between the cracking noise of ethanol-dampened charcoal and earthquakes. *Physical Review Letters*, 115, 25503. https://doi.org/10.1103/PhysRevLett.115.025503

Jagla, E. A., & Kolton, A. B. (2010). A mechanism for spatial and temporal earthquake clustering. Journal of Geophysical Research, 115, B05312. https://doi.org/10.1029/2009JB006974

- Roberts, N. S., Bell, A. F., & Main, I. G. (2016). Mode switching in volcanic seismicity: El Hierro 2011–Ş2013. *Geophysical Research Letters*, 43, 4288–4296. https://doi.org/10.1002/2016GL068809
- Ross, Z. E., Rollins, C., Cochran, E. S., Hauksson, E., Avouac, J.-P., & Ben-Zion, Y. (2017). Aftershocks driven by afterslip and fluid pressure sweeping through a fault-fracture mesh. *Geophysical Research Letters*, 44, 8260–8267. https://doi.org/10.1002/2017GL074634
- Saichev, A., Helmstetter, A., & Sornette, D. (2005). Power-law distributions of offspring and generation numbers in branching models of earthquake triggering. *Pure and Applied Geophysics*, *162*(6), 1113–1134. https://doi.org/10.1007/s00024-004-2663-6
- Saichev, A., & Sornette, D. (2004). Anomalous power law distribution of total lifetimes of branching processes: Application to earthquake aftershock sequences. *Physical Review E*, 70, 46123. https://doi.org/10.1103/PhysRevE.70.046123
- Schoenball, M., & Ellsworth, W. L. (2017). A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in oklahoma and southern kansas. *Journal of Geophysical Research: Solid Earth*, 122, 10,189–10,206. https://doi.org/10.1002/ 2017JB014850
- Shelly, D. R., Ellsworth, W. L., & Hill, D. P. (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 long valley caldera, california, earthquake swarm. *Journal of Geophysical Research:* Solid Earth, 121, 1776–1795. https://doi.org/10.1002/2015JB012719
- Trugman, D. T., Dougherty, S. L., Cochran, E. S., & Shearer, P. M. (2017). Source spectral properties of small to moderate earthquakes in southern Kansas. Journal of Geophysical Research: Solid Earth, 122, 8021–8034. https://doi.org/10.1002/2017JB014649

Turcotte, D. L., Holliday, J. R., & Rundle, J. B. (2007). Bass, an alternative to ETAS. Geophysical Research Letters, 34, L12303. https://doi.org/10.1029/2007GL029696

Ueda, T., & Kato, A. (2019). Seasonal variations in crustal seismicity in San-in district, southwest Japan. *Geophysical Research Letters*, 46, 3172–3179. https://doi.org/10.1029/2018GL081789

Utsu, T., Ogata, Y., & Matsu ura, R. s. (1995). The centenary of the Omori formula for a decay law of aftershock activity. Journal of Physics of the Earth, 43(1), 1-33.

van der Elst, N. J. (2017). Accounting for orphaned aftershocks in the earthquake background rate. *Geophysical Journal International*, 211(2), 1108–1118. https://doi.org/10.1093/gji/ggx329

Varotsos, P., Eftaxias, K., Vallianatos, F., & Lazaridou, M. (1996). Basic principles for evaluating an earthquake prediction method. *Geophysical Research Letters*, 23(11), 1295–1298. https://doi.org/10.1029/96GL00905

Vasylkivska, V. S., & Huerta, N. J. (2017). Spatiotemporal distribution of oklahoma earthquakes: Exploring relationships using a nearest-neighbor approach. *Journal of Geophysical Research: Solid Earth*, 122, 5395–5416. https://doi.org/10.1002/2016JB013918Vere-Jones, D. (2005). A class of self-similar random measure. *Advances in Applied Probability*, 37(4), 908–914.

re-Jones, D. (2005). A class of self-similar random measure. Advances in Applied Probability, 37(4), 908–914.

Vere-Jones, D., & Davies, R. (1966). A statistical survey of earthquakes in the main seismic region of New Zealand: Part 2—Time series analyses. New Zealand Journal of Geology and Geophysics, 9(3), 251–284.

Vere-Jones, D., & Ozaki, T. (1982). Some examples of statistical estimation applied to earthquake data. Annals of the Institute of Statistical Mathematics, 34(1), 189–207.

Yamashita, T., & Knopoff, L. (1987). Models of aftershock occurrence. Geophysical Journal International, 91(1), 13-26.

Zaliapin, I., & Ben-Zion, Y. (2013a). Earthquake clusters in Southern California I: Identification and stability. Journal of Geophysical Research: Solid Earth, 118, 2847–2864. https://doi.org/10.1002/jgrb.50179

Zaliapin, I., & Ben-Zion, Y. (2013b). Earthquake clusters in Southern California II: Classification and relation to physical properties of the crust. Journal of Geophysical Research: Solid Earth, 118, 2865–2877. https://doi.org/10.1002/jgrb.50178

Zaliapin, I., & Ben-Zion, Y. (2016a). A global classification and characterization of earthquake clusters. *Geophysical Journal International*, 207(1), 608–634. https://doi.org/10.1093/gji/ggw300

Zaliapin, I., & Ben-Zion, Y. (2016b). Discriminating characteristics of tectonic and human-induced seismicity. Bulletin of the Seismological Society of America, 106(3), 846–859. https://doi.org/10.1785/0120150211

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., & Wong, H. (2008). Clustering analysis of seismicity and aftershock identification. *Physical Review Letters*, 101, 18,501. https://doi.org/10.1103/PhysRevLett.101.018501

Zhang, X., & Shcherbakov, R. (2016). Power-law rheology controls aftershock triggering and decay. *Scientific Reports*, 6. https://doi.org/10. 1038/srep36668

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association, 97(458), 369–380. https://doi.org/10.1198/016214502760046925

Zhuang, J., Ogata, Y., & Vere-Jones, D. (2004). Analyzing earthquake clustering features by using stochastic reconstruction. Journal of Geophysical Research, 109, B05301. https://doi.org/10.1029/2003JB002879