View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Explore Bristol Research

Zhuang, J., Werner, M. J., & Harte, D. S. (2013). Stability of earthquake clustering models: Criticality and branching ratios. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 88(6), [062109]. 10.1103/PhysRevE.88.062109

Link to published version (if available):
10.1103/PhysRevE.88.062109

Link to publication record in Explore Bristol Research

PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html

Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact open-access@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

Stability of earthquake clustering models: Criticality and branching ratios

Jiancang Zhuang*
The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tokyo 190-8562, Japan and Department of Statistical Science, The Graduate University for Advanced Studies, Tokyo 190-8562, Japan

Maximilian J. Werner†
Department of Geosciences, Princeton University, New Jersey 08544, USA

David S. Harte‡
GNS Science, Wellington, New Zealand

(Received 27 December 2012; revised manuscript received 28 October 2013; published 4 December 2013)

We study the stability conditions of a class of branching processes prominent in the analysis and modeling of seismicity. This class includes the epidemic-type aftershock sequence (ETAS) model as a special case, but more generally comprises models in which the magnitude distribution of direct offspring depends on the magnitude of the progenitor, such as the branching aftershock sequence (BASS) model and another recently proposed branching model based on a dynamic scaling hypothesis. These stability conditions are closely related to the concepts of the criticality parameter and the branching ratio. The criticality parameter summarizes the asymptotic behavior of the population after sufficiently many generations, determined by the maximum eigenvalue of the transition equations. The branching ratio is defined by the proportion of triggered events in all the events. Based on the results for the generalized case, we show that the branching ratio of the ETAS model is identical to its criticality parameter because its magnitude density is separable from the full intensity. More generally, however, these two values differ and thus place separate conditions on model stability. As an illustration of the difference and of the importance of the stability conditions, we employ a version of the BASS model, reformulated to ensure the possibility of stationarity. In addition, we analyze the magnitude distributions of successive generations of the BASS model via analytical and numerical methods, and find that the compound density differs substantially from a Gutenberg-Richter distribution, unless the process is essentially subcritical (branching ratio less than 1) or the magnitude dependence between the parent event and the direct offspring is weak.

DOI: 10.1103/PhysRevE.88.062109 PACS number(s): 64.60.A−, 05.45.Tp, 02.50.Ey, 91.30.Dk

I. INTRODUCTION AND MOTIVATIONS

Clustering has been widely recognized as one of the most dominant features of seismicity and has been modeled with some success by branching processes (see, e.g., Kagan and Knopoff [1], Musmeci and Vere-Jones [2], Kagan [3], Ogata [4], and Ogata and Zhuang [5]). In particular, the epidemic-type aftershock sequence (ETAS) model (see, e.g., Ogata [4,6]) has been widely accepted as a standard model for describing clustering characteristics of earthquake occurrences. Several versions of this model have been proposed for and applied to the studies of seismicity analysis and forecasting in different regions of the globe (see, e.g., Kagan and Knopoff [7], Console *et al.* [8], Helmstetter *et al.* [9], Zhuang *et al.* [10], Zhuang [11], and Werner *et al.* [12]). Model parameters are usually estimated by a maximum likelihood estimation (MLE) (see, e.g., Daley and Vere-Jones [13]) or by the expectation-maximization (EM) algorithm (Veen and Schoenberg [14]). The obtained model parameter values can vary amongst study periods and regions, sometimes substantially (see, e.g., Ref. [15], Table 1), and it is generally recognized that parameter estimation remains a difficult issue [14] with potentially serious consequences (e.g., over- or underpredicting seismicity).

A strong constraint on reasonable estimates of model parameters and model forecasts is provided by the prerequisite that the model be stable. We therefore need a good understanding of the model stability conditions. More broadly, any generalizations of the model or, indeed, any explicit branching model for seismicity should be formulated with consideration of stability conditions. These conditions are closely related to the concepts of the criticality parameter and the branching ratio. These two concepts for the ETAS model have already been addressed in other papers (e.g., Refs. [16–18]). For the ETAS model, as we prove below, these two quantities are the same if the model is subcritical. But we also show that, for a more general branching process, these two concepts are not identical. Helmstetter and Sornette [15] discussed how the population increases with time when the ETAS model is subcritical, critical or supercritical. In the present paper, we focus on how the population changes in the magnitude dimension.

In the following sections, rather than discussing directly the properties of the ETAS model, we focus on a more general class of branching models in which the magnitude distribution of direct offspring depends on the parent's magnitude, and derive the formulas for the criticality parameter and the branching ratio. Next, we discuss the branching aftershock sequence (BASS) model [19,20] and another branching model proposed by Lippiello *et al.* [18] as instances of this general class. Finally, we analytically and numerically analyze a modified version of the BASS model, which is reformulated

*zhuangjc@ism.ac.jp
†mwerner@princeton.edu
‡d.harte@gns.cri.nz

 ©2013 American Physical Society

to allow for the possibility of stability, as an illustration of the theory.

II. BRANCHING MODELS OF SEISMICITY

In this article, we only consider the process of earthquake occurrences in the temporal and magnitude domain, which can be described by a marked point-process model with a conditional intensity (also called time-varying seismicity rate, or stochastic intensity)

$$\lambda(t,m) = \lim_{\Delta_t \downarrow 0} \lim_{\Delta_m \downarrow 0} \frac{1}{\Delta_t \Delta_m} \Pr\{N([t,t+\Delta_t) \times (m-\Delta_m/2, m+\Delta_m/2)) \geqslant 1 \mid \mathcal{H}_t\}, \quad (1)$$

where the history $\mathcal{H}_t$ represents the observed process of earthquake occurrences before time t and $N([a,b) \times (m_1,m_2))$ denotes the number of earthquakes with occurrence times in the interval $[a,b)$ and magnitudes in the interval (m_1,m_2). The conditional intensity of the ETAS model can be expressed as (Ogata [6])

$$\lambda_{\mathrm{ETAS}}(t,m) = s(m)\left[\mu + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i)\right], \quad (2)$$

where $s(m)$ is the probability density function (PDF) of the magnitudes, μ is a constant background seismicity rate, $\kappa(m)$ is the mean number of events triggered by an earthquake of magnitude m, and $g(t)$ is the PDF of the time lags between the occurrence times of a parent event and its direct offspring. The branching representation of this model is as follows: Once an event, say, i, occurs, it triggers a Poisson process with rate $\kappa(m_i)g(t-t_i)$ starting from t_i, and each offspring triggers its own Poisson process (offspring) independently. Another interpretation of this model is that the rate of earthquake occurrence at time t consists of the contributions from the background rate μ and all previous events.

In practice, $s(m)$ is usually set equal to the PDF form of the Gutenberg-Richter magnitude-frequency relation for earthquakes above a threshold magnitude m_c, i.e.,

$$s(m) = \beta e^{-\beta(m-m_c)}, \quad m \geqslant m_c, \quad (3)$$

where β is linked with the so-called Gutenberg-Richter's b value by $\beta = b \ln 10$; the productivity function $\kappa(m)$ takes the form of a positive exponential law, i.e.,

$$\kappa(m) = A\, e^{\alpha(m-m_c)}, \quad m \geqslant m_c, \quad (4)$$

and the time distribution has a PDF of the form

$$g(t) = \frac{p-1}{c}\left(1+\frac{t}{c}\right)^{-p}, \quad t > 0, \quad (5)$$

which corresponds to the Omori-Utsu formula for the frequency decay of aftershocks [21,22].

In this article, we consider a more general form of branching model than the ETAS model, where the magnitudes of the direct offspring may depend on the magnitude of the parent. That is, the conditional intensity can be written in the form

$$\lambda(t,m) = \mu s_0(m) + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i) s(m \mid m_i), \quad (6)$$

where μ, g, κ are the same as their counterparts in Eq. (2), $s_0(m)$ is the magnitude probability density for the background events, and $s(m \mid m_i)$ is the magnitude PDF of the direct offspring produced by an event of magnitude m_i. If we let $s_0(m) = s(m \mid m_i) = s(m)$, then the model in Eq. (6) is simply the ETAS model.

The above class of models (6) is motivated by the interesting question whether the magnitude of a triggered event depends on the magnitude of its parent event. Zhuang *et al.* [23] found, by using the stochastic reconstruction method, that there is an indication of such dependence in the Japan Meteorological Agency (JMA) catalog. Lippiello *et al.* [24] made similar conclusions by evaluating the magnitude differences between subsequent events in the Southern California Earthquake Data Center (SCEDC) earthquake catalog.

It is worthwhile to mention that spatiotemporal versions of the ETAS model and the other two models, which are also discussed in following sections, have already been implemented by introducing a spatial dependence between the parent events and their direct offspring in the triggering term (e.g., Ogata [4], Ogata and Zhuang [5], Console *et al.* [8], Helmstetter *et al.* [9], Zhuang *et al.* [25], Zhuang [11], Werner *et al.* [12], and Lippiello *et al.* [24]). In this article, without loss of generality, we only consider the magnitude and temporal components. The theory developed here also applies to the spatiotemporal versions of these models.

III. CRITICALITY

The criticality parameter characterizes the asymptotic behavior of the population size $G_n(m)$ at the nth generation when n is sufficiently large. Here we call the background events generation 0, and the direct offspring of generation n generation $n+1$. It is easy to see from (6) that the expected (magnitude) intensity function for the first generation (earthquakes that are direct offspring of background events) is

$$G_1(m) = \int_{\mathcal{M}} \kappa(m') s(m \mid m') G_0(m') dm',$$

where $\mathcal{M}$ represents the range of magnitudes, say $[m_c,\infty)$, and $G_0(m') = \mu s_0(m)$ is the background intensity. Similarly, the second generation is given by

$$G_2(m) = \int_{\mathcal{M}} \kappa(m') s(m \mid m') G_1(m') dm',$$

and the $(n+1)$th generation by

$$G_{n+1}(m) = \int_{\mathcal{M}} \kappa(m') s(m \mid m') G_n(m') dm' = \int_{\mathcal{M}} K^{[n+1]}(m;m') G_0(m') dm',$$

where $K^{[n]}$ $(n \geqslant 1)$ is defined by induction

$$K^{[1]}(m;m') = K(m;m') = \kappa(m') s(m \mid m'),$$
$$K^{[n+1]}(m;m') = \int_{\mathcal{M}} \kappa(m^*) s(m \mid m^*) K^{[n]}(m^*;m') dm^*.$$

Suppose that $a(m')$ and $b(m)$ are the left and right eigenfunctions of K corresponding to the maximum eigenvalue ϱ,

i.e.,

$$\varrho a(m') = \int_{\mathcal{M}} a(m)K(m;m')dm \tag{7}$$

and

$$\varrho b(m) = \int_{\mathcal{M}} K(m;m')b(m')dm'. \tag{8}$$

Let

$$\Omega(m;m') = a(m')b(m), \tag{9}$$

i.e., Ω is the projection operator of K corresponding to ϱ, or

$$\int_{\mathcal{M}} \Omega(m;m^*)K(m^*;m')dm^* = \int_{\mathcal{M}} K(m;m^*)\Omega(m^*;m')dm^* = \varrho\Omega(m;m'). \tag{10}$$

When both $\kappa(m')$ and $s(m|m')$ are piecewise continuous functions, then the linear integral equations (7) and (8) can be viewed as the continuum limit of eigenvalue equations of the form

$$\sum_j M_{i,j}v_j = \varrho v_i,$$

where $(M_{i,j})$ is a matrix and (v_j) is a vector. When the dimension of (v_j) is n and of $(M_{i,j})$ is $n \times n$, there are n independent orthonormal eigenvectors and at most n eigenvalues. When a linear operator, which maps an element from one space into the same space, is applied in a compound way repeatedly, the behavior is similar to applying a projection operator and a scaling operator for many times, where the projection operator projects the element to the subspace spanned by the eigenvectors corresponding to the maximum eigenvalue, if this element is not orthonormal to these eigenvectors. When $n \to \infty$, there are at most a countable number of eigenvalues and the maximum eigenvalue is separated from others. In our case, the linear operator is K, the projective operator is Ω, and the scaling operator is ϱ. If the element is not orthonormal to the eigenelement (invariant direction) corresponding to the maximum eigenvalue, it will finally be projected into this eigendirection. It follows from (10) that Ω is a solution of the following fixed-point problem:

$$\text{Find } H\text{: } H(m,m') = \frac{1}{\varrho}\int_{\mathcal{M}} K(m;m^*)H(m^*,m')dm^*.$$

Thus, when $n \to \infty$,

$$\frac{K^{[n]}}{\varrho^n} \to \Omega \tag{11}$$

and

$$G_n(m) \to \varrho^n \int_{\mathcal{M}} \Omega(m;m')G_0(m')dm'. \tag{12}$$

From (12), we can see that if $\varrho < 1$, then $G_n \to 0$ when $n \to \infty$, and that if $\varrho > 1$, then $G_n \to \infty$ when $n \to \infty$. Here ϱ is called the *criticality* parameter because if $\varrho < 1$, the process is stable and otherwise the population explodes to infinity.

Equation (12) can be rewritten as

$$\begin{aligned} G_n(m) &\to \varrho^n \int_{\mathcal{M}} b(m)a(m')G_0(m')dm' \\ &= \varrho^n b(m)\int_{\mathcal{M}} a(m')G_0(m')dm' \\ &= \varrho^n b(m) \times \text{const}, \end{aligned} \tag{13}$$

implying that $b(m)$ is asymptotically proportional to the intensity of the population when $n \to \infty$. The eigenfunction $a(m')$ can be interpreted as the asymptotic ability in producing offspring, directly and indirectly, from an ancestor $\{m'\}$ because

$$\begin{aligned} &\lim_{n\to\infty}\sum_{i=n}^{\infty}\int_{\mathcal{M}} K^{[i]}(m;m')dm \\ &= \lim_{n\to\infty}\sum_{i=n}^{\infty}\varrho^i\int_{\mathcal{M}}\Omega(m;m')dm \\ &= \lim_{n\to\infty}\sum_{i=n}^{\infty}\varrho^i a(m')\int_{\mathcal{M}} b(m)dm \\ &= \lim_{n\to\infty}\frac{\varrho^n}{1-\varrho}a(m') \times \text{const}. \end{aligned} \tag{14}$$

Criticality of the ETAS model. For the ETAS model, whose magnitude density is separable and the background rate is constant, the eigenvalue equations are

$$\varrho a(m') = \kappa(m')\int_{\mathcal{M}} a(m)s(m)dm, \tag{15}$$

$$\varrho b(m) = s(m)\int_{\mathcal{M}} \kappa(m')b(m')dm', \tag{16}$$

where $\mathcal{M}$ is the magnitude range. We can see

$$a(m') = C_1\kappa(m') \tag{17}$$

and

$$b(m) = C_2 s(m). \tag{18}$$

Substituting $a(m')$ and $b(m)$ back into (15) and (16),

$$\varrho = \int_{\mathcal{M}} \kappa(m)s(m)dm. \tag{19}$$

For the ETAS model given in Eq. (2), substituting $\kappa(m) = A\,e^{\alpha(m-m_0)}$ and $s(m) = \beta\,e^{-\beta(m-m_0)}$ into (19), the criticality parameter is then

$$\varrho = \int_{m_0}^{\infty} s(m)\kappa(m)dm = \frac{A\beta}{\beta-\alpha},$$

where the last equality requires $\alpha < \beta$ unless the magnitude density is truncated or tapered (see, e.g., Eqs. (4) and (5) in Sornette and Werner [26]).

Details of the behavior of the ETAS model were discussed, for example, by Helmstetter and Sornette [15], Zhuang and Ogata [27], Saichev and Sornette [28], Lippiello *et al.* [18], and Vere-Jones and Zhuang [29].

IV. BRANCHING RATIO

The *branching ratio* is defined as the proportion of triggered events amongst all events. Suppose that the process is

stationary (which can be restricted only to the background process) and ergodic. Taking expectations over the time domain on both sides of (6),

$$\bar{\lambda} s_1(m) = \mathbf{E}\,[\lambda(t,m)] = \mu s_0(m) + \mathbf{E}\left[\sum_{t_i:t_i<t} \kappa(m_i) g(t-t_i) s(m\,|\,m_i)\right], \quad (20)$$

where $\mathbf{E}[\cdot]$ means the expectation over the time domain, $\bar{\lambda}$ is the total average rate, and $s_1(m)$ is the marginal magnitude density of all events. The summation on the right-hand side can be written as

$$\begin{aligned}&\mathbf{E}\left[\sum_{t_i:t_i<t} \kappa(m_i) g(t-t_i) s(m\,|\,m_i)\right]\\ &= \int_{\mathcal{M}} \int_{-\infty}^{t} \kappa(m^*) g(t-u) s(m\,|\,m^*) \times \bar{\lambda} s_1(m^*) du\, dm^*\\ &= \bar{\lambda} \int_{\mathcal{M}} \kappa(m^*) s(m\,|\,m^*) s_1(m^*) dm^* \quad \left(\because \int_0^{\infty} g(u) du = 1\right),\end{aligned} \quad (21)$$

i.e.,

$$\bar{\lambda} s_1(m) = \mu s_0(m) + \bar{\lambda} \int_{\mathcal{M}} \kappa(m^*) s(m\,|\,m^*) s_1(m^*) dm^*. \quad (22)$$

In the above, (21) can be obtained from martingale theories related to the properties of the conditional intensity (see, e.g., Zhuang [30] for justification).

Integrate both sides with respect to m, noting that s_1, s_0, and $s(\cdot\,|\,\cdot)$ are PDFs,

$$\bar{\lambda} = \mu + \bar{\lambda} \int_{\mathcal{M}} \kappa(m^*) s_1(m^*) dm^*. \quad (23)$$

The branching ratio is obtained by

$$\omega = 1 - \frac{\mu}{\bar{\lambda}} = \int_{\mathcal{M}} \kappa(m^*) s_1(m^*) dm^*, \quad (24)$$

which is also the average number of events that are triggered by an arbitrary event. This parameter is nonzero and less than 1 only when $\varrho < 1$.

The difference between criticality and the branching ratio is as follows: The criticality characterizes the average productivity of an arbitrary event after infinitely many generations, while the branching ratio characterizes the average productivity over all generations. For the ETAS model, the magnitude density is completely separable from the whole intensity, i.e., $s_1(m) = s_0(m) = s(m\,|\,m^*)$. By comparing (24) to (19), we have $\omega = \varrho$ (see also Ref. [31]).

V. THE BASS MODEL

The BASS model was developed by Turcotte, Holliday, and colleagues (see, e.g., Turcotte *et al.* [19] and Holliday *et al.* [20]), with the explicit consideration of the Båth law. The BASS model is based on the following assumptions:

(a) The magnitude density of the direct offspring is [Ref. [20], Eq. (15)]

$$s(m\,|\,m^*) = \beta_d\, e^{-\beta_d (m-m_c)}. \quad (25)$$

(b) The expected number of direct offspring from an event of magnitude m^* is [Ref. [20], Eq. (23)]

$$\kappa(m^*) = A_{\text{bass}}\, e^{\beta_d (m^*-m_c)}, \quad (26)$$

where A_{bass} is a constant.

(c) A_{bass} is restricted by the Båth law, i.e., [Ref. [20], Eq. (12)]

$$\text{No. (direct offspring} \geqslant m^* - \delta) = 1, \quad (27)$$

where δ is a constant magnitude difference between the main shock and the largest aftershock, according to the Båth law, usually taking a value of 1.2.

It is worthwhile to mention the following points: (1) The Båth law can be derived as an asymptotic property of the ETAS model and does not need to be specified explicitly. Here we refer to Vere-Jones and Zhuang [29] and Vere-Jones [32] for details. Helmstetter and Sornette [33] also showed that the Båth law could be recovered through simulations based on the ETAS model. (2) Turcotte and co-workers considered self-similarity of the branching process a desirable or even a necessary attribute of a model, and claim that the BASS model is self-similar while the ETAS model is not. Their notion of self-similarity appears to be that the model simply satisfies the four scaling relations: Gutenberger-Richter of magnitudes, Omori of time decay, modified Omori of spatial decay, and a modified Båth's law of the parent event (see the abstract in Ref. [17]). They assert in the abstract of Ref. [17] that this is not the case for the ETAS model. Their claim appears to focus on the required condition that $\alpha < \beta$ in the ETAS model (see Ref. [17], Sec. IV, and the second to last paragraph in Sec. V). We show in the following that the stochastic BASS model must be similarly constrained to enable the existence of a stable version. (3) Vere-Jones [16] defined a class of models using a tighter definition of self-similarity, and showed that this places quite different constraints on the magnitude. Vere-Jones' self-similar ETAS model does not impose hard (i.e., fixed) boundaries on the magnitude distribution, but uses a normalized stability factor to constrain the dependence between the magnitudes of the direct offspring and the parent. A fixed upper magnitude boundary, as in the truncated exponential distribution, is artificial and introduces other problems.

In early papers of the BASS model [19,20], a deterministic version is used, i.e., during simulations, the number of direct offspring is obtained by rounding $\kappa(m)$ in Eq. (26) to the nearest integer. To formulate the BASS model as a (stochastic) Poisson cluster process defined by the conditional intensity (6), we assume in the following that the number of the direct offspring is a Poisson distributed random variable. Moreover, we also note that (27), together with (26) and (25), constrains $A_{\text{bass}} = \exp(-\beta_d \delta)$, although this requirement is dropped in the discussion below.

In the original version of the BASS model proposed by Turcotte *et al.* [19] and Holliday *et al.* [20], the magnitude density of each triggered generation is the Gutenberg-Richter magnitude-frequency relation without any truncations. This may not cause any immediate issues when discussing or simulating a single cluster from a particular main shock. However, the average number of triggered shocks in the second generation from an arbitrary first-generation shock is infinite,

because

$$\int_{m_c} A_{\text{bass}}\, e^{\beta_d(m'-m_c)}\beta_d\, e^{-\beta_d(m'-m_c)}\, dm' = +\infty. \tag{28}$$

Therefore, to ensure that the average number of triggered events remains finite, the productivity exponent must either be smaller than the exponent of the magnitude density (e.g., $\alpha < \beta_d$, essentially recovering the standard ETAS model) or the magnitude density must be a truncated (Sornette and Werner [26]) or tapered exponential distribution (see, e.g., Kagan and Schoenberg [34]). A third solution, which is pursued in more detail below and apparently used in the simulations by Turcotte *et al.* [19] and Holliday *et al.* [20], requires that the magnitudes of triggered events are restricted to be less than the magnitude of their parent, or, similarly, the magnitude differences between the triggered events and the parent event do not exceed a certain level Δ.

When the BASS model is additionally endowed with a background rate and background magnitude density, further constraints are required. If the background magnitude density is the nontruncated Gutenberg-Richter relation with exponent β_0, then the productivity exponent β_d must be smaller than β_0 to ensure a finite mean number of direct offspring triggered by an arbitrary background event. This formulation of a stochastic Hawkes process version of the BASS model is investigated in detail in the following sections.

VI. LIPPIELLO *ET AL.*'S MODEL

Lippiello *et al.* [18,35,36] also proposed an alternative to the ETAS model and studied its properties, including criticality. The conditional intensity of this model is

$$\lambda(t,m) = \mu s_0(m) + \sum_{i:t_i<t} \Phi\left(\frac{t-t_i}{k\,10^{-b(m-m_i)}}\right), \tag{29}$$

where Φ is the response function. The above equation can be normalized as

$$\lambda(t,m) = \mu s_0(m) + \sum_{i:t_i<t} s(m)\kappa(m_i)g(t-t_i\,|\,m,m_i), \tag{30}$$

where

$$\kappa(m) = \frac{k10^{bm}}{b\ln 10}\int_0^\infty \Phi(u)du,$$
$$s(m) = b10^{-bm}\ln 10,$$

and

$$g(t\mid m,m') = \frac{\Phi\left(\frac{t}{k\,10^{-b(m-m')}}\right)}{k10^{-b(m-m')}\int_0^\infty \Phi(u)du},$$

providing that $\int_0^\infty \Phi(u)du$ is finite.

The model is similar to the stochastic version of the BASS model, except that the time distribution of triggered events also depends on the magnitudes of both the parent and offspring events. However, since the integral of $g(t\mid m,m')$ from 0 to ∞ is 1 for all the values of m and m', the results throughout this article also hold for this model. Lippiello *et al.* [18,35,36] also considered the case of dependence between occurrence times and locations. Without loss of generality, we can omit the spatial component.

VII. RESULTS FROM A CLUSTERING MODEL WITH TRUNCATED MAGNITUDE DISTRIBUTION

In this section, we consider a particular formulation of the class given by (6) motivated by the BASS model. We assume $\kappa(m) = A\,e^{\beta(m-m_c)}$, $s_0(m) = \beta_0\,e^{-\beta_0(m-m_c)}$,

$$s(m\,|\,m') = f(m)H(m'-m+\Delta)/F(m'+\Delta), \quad \Delta \geqslant 0,$$

where $f(m) = \beta\,e^{-\beta(m-m_c)}$, $F(m)$ is the cumulative probability distribution corresponding to $f(m)$, H is the Heaviside function, and A, β_0, β, and Δ are constants. In the following we take $m_c = 0$ and use β instead of β_d to shorten the notation. It can be seen that there are two differences between this model and the BASS model: (1) β_0 is introduced for the background seismicity and also constrained to be greater than β to ensure a finite expected rate of first generation events, and (2) A and Δ are free parameters, not restricted by (27).

A. Case $\Delta = 0$

Here we first discuss the left eigenequation and then the right eigenequation. The left eigenequation

$$\varrho a(m') = \frac{\kappa(m')}{F(m')}\int_0^{m'} f(m)a(m)dm \tag{31}$$

can be rewritten as

$$\varrho a(m)F(m) = \kappa(m)\int_0^m f(m')a(m')dm. \tag{32}$$

Taking derivatives of both sides and using $\kappa'(m) = \beta\kappa(m)$,

$$\varrho\frac{da(m)}{dm}F(m) + \varrho a(m)f(m) = \beta\varrho a(m)F(m) + \kappa(m)f(m)a(m),$$

i.e.,

$$\frac{a'(m)}{a(m)} = \beta + \frac{1}{\varrho}\frac{\kappa(m)f(m)}{F(m)} - \frac{f(m)}{F(m)}.$$

The above equation has a solution

$$\ln a(m) = \beta m + \frac{A\beta}{\varrho}\int\frac{dm}{1-e^{-\beta m}} - \ln F(m)$$

or

$$a(m) = C\exp\left[\beta\left(\frac{A}{\varrho}+1\right)m\right](1-e^{-\beta m})^{\frac{A}{\varrho}-1}.$$

A meaningful $a(m)$ requires that $a(0)$ be finite. This condition is equivalent to $\frac{A}{\varrho}-1 \geqslant 0$, i.e., $\varrho \leqslant A$.

The right eigenequation is

$$\varrho b(m) = \int_{\mathcal{M}} K(m,m')b(m')dm = \int_m^\infty \frac{A\beta e^{\beta(m'-m)}b(m')}{1-e^{-\beta m'}}dm'. \tag{33}$$

Take derivatives of both sides,

$$\varrho\frac{db(m)}{dm} = -\beta\varrho b(m) - K(m,m)b(m) = -\beta\varrho b(m) - \frac{A\beta b(m)}{1-e^{-\beta m}}, \tag{34}$$

and so

$$b(m) = C\, e^{-\beta m}[e^{\beta m} - 1]^{-\frac{A}{\varrho}}, \tag{35}$$

where C could be any constant. Since $b(m)$ is a PDF, we are only interested in the cases that have a finite integral over $[0,\infty]$. The required condition for

$$\int_0^\infty b(m)dm = C\int_0^\infty e^{-\beta m}[e^{\beta m} - 1]^{-\frac{A}{\varrho}}dm \tag{36}$$

to be finite is $\varrho > A$, since

$$\begin{aligned}\int_0^\infty e^{-\beta m}[e^{\beta m} - 1]^{-\frac{A}{\varrho}}dm &= \int_0^\infty e^{-(1+\frac{A}{\varrho})\beta m}(1 - e^{-\beta m})^{-\frac{A}{\varrho}}dm \\ &= \frac{1}{\beta}\int_0^1 u^{\frac{A}{\varrho}}(1-u)^{-\frac{A}{\varrho}}du \\ &= \frac{1}{\beta}\mathrm{B}\left(1+\frac{A}{\varrho},1-\frac{A}{\varrho}\right) \\ &= \frac{A\pi}{\varrho\beta\sin(A\pi/\varrho)}\end{aligned} \tag{37}$$

has a finite value only if $A/\varrho < 1$, i.e., $\varrho > A$, where B is the beta function.

Together with the conclusion that for the left eigenfunction $a(m)$ is a finite value at $a(0)$ only if $\varrho \leqslant A$, we can see that the meaningful criticality parameter ϱ is the value $\kappa(0) = A$ with eigenfunction $a(m) = C\, e^{2\beta m}$ when $\Delta = 0$.

The criticality parameter can be also obtained in the following way. We apply the limit operation to both sides of (31). When $m \to 0$, notice that $F(0) = 0$ and $f(m) = dF(m)/dm$. Therefore, by L'Hopital's rule,

$$\begin{aligned}\varrho a(0) &= \kappa(0)\lim_{m'\to 0}\frac{\int_0^{m'} f(m)a(m)dm}{F(m')} \\ &= \kappa(0)\lim_{m'\to 0}\frac{\frac{d}{dm'}\int_0^{m'} f(m)a(m)dm}{\frac{d}{dm'}F(m')} = \kappa(0)a(0).\end{aligned}$$

Thus, if $a(0)$ is nonzero and finite, the criticality parameter ϱ is

$$\varrho = \kappa(0) = A.$$

That is, the criticality of this model is determined by the average productivity of an earthquake at the threshold magnitude. In other words, the restriction on the magnitudes of the direct offspring stabilizes the model.

In this case, if $\kappa(m)$ is a monotonically increasing function of m, then the branching ratio given by (24) satisfies

$$\omega = \int_{\mathcal{M}}\kappa(m^*)s_1(m^*)dm^* \geqslant \int_{\mathcal{M}}\kappa(0)s_1(m^*)dm^* = \kappa(0) = \varrho,$$

since $\kappa(m) \geqslant \kappa(0)$ for all m and $\int_{\mathcal{M}} s_1(m)dm = 1$. In contrast to the ETAS model, the criticality parameter of this model is not identical to the branching ratio. We note that $\omega \to 1$ when $\varrho = 1$ since ω always lies between ϱ and 1 when the process is subcritical. We can also construct a model with $\varrho \geqslant \omega$, for example, by setting $\kappa(m)$ equal to a monotonically decreasing function.

Note that condition $\rho < 1$ is only necessary but not sufficient to ensure the stability of the model. For example, in this model, if we set $\kappa(0) = A < 0$ and $\beta \geqslant \beta_0$, then $\rho = \kappa(0) < 0$, but $\omega \to 1$, since the expected number of events in the first generation is infinity.

In the remainder of this section, we show that the marginal magnitude distribution of all events is no longer a Gutenberg-Richter (G-R) (exponential) distribution. From (22), the magnitude density for all the events is determined by the integral equation

$$s_1(m) = \frac{\mu}{\bar{\lambda}}s_0(m) + \int_{\mathcal{M}}\kappa(m^*)s(m\,|\,m^*)s_1(m^*)dm^*, \tag{38}$$

where

$$1 - \frac{\mu}{\bar{\lambda}} = \int_0^\infty \kappa(m)s_1(m)dm.$$

Substitute $s(m\,|\,m^*) = f(m)H(m^* - m)/F(m^*)$ into the above, we have

$$s_1(m) = \frac{\mu}{\bar{\lambda}}s_0(m) + f(m)\int_m^\infty \frac{\kappa(m^*)s_1(m^*)}{F(m^*)}dm^*. \tag{39}$$

It is easy to see that $s_1(m)$ has the form $Cs_0(m) + f(m)q(m)$ where $C = \mu/\bar{\lambda}$, and

$$q(m) = \int_m^\infty \frac{\kappa(m^*)s_1(m^*)}{F(m^*)}dm^*.$$

Substituting $s_1(m) = Cs_0(m) + f(m)q(m)$ into the above equation, we obtain

$$q(m) = \int_m^\infty \frac{\kappa(m^*)[Cs_0(m^*) + f(m^*)q(m^*)]}{F(m^*)}dm^*, \tag{40}$$

which is equivalent to the differential equation

$$q'(m) + \frac{f(m)\kappa(m)}{F(m)}q(m) = -\frac{Cs_0(m)\kappa(m)}{F(m)}. \tag{41}$$

Substituting $\kappa(m) = A\, e^{\beta m}$, $f(m) = \beta\, e^{-\beta m}$, and $F(m) = 1 - e^{-\beta m}$, the above equation has a solution

$$\begin{aligned}q(m) = &\frac{CA\beta_0}{\beta_0 - (1+A)\beta}e^{-(\beta_0-\beta)m}(1 - e^{-\beta m})^{-A} \\ &\times {}_2F_1\left(1 - A, \frac{\beta_0}{\beta} - (1+A), \frac{\beta_0}{\beta} - A, e^{-\beta m}\right) \\ &+ C_1\, e^{-A\beta m}[1 - e^{-\beta m}]^{-A},\end{aligned} \tag{42}$$

where C_1 is an arbitrary constant, and ${}_2F_1$ is the hypergeometric function defined by

$${}_2F_1(a,b;c;z) \equiv \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{n=0}^\infty \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)\Gamma(n+1)}z^n$$

$$\text{for} \quad |z| < 1, \quad \Gamma(x) \equiv \int_0^x u^x\, e^{-u}\, du$$

being the gamma function (e.g., Ref. [37]). From $s_1(m) = s_0(m)$ if $A = 0$, $C_1 = 0$ can be obtained. Therefore,

$$\begin{aligned}s_1(m) = Cs_0(m) + &\frac{CA\beta_0\beta}{\beta_0 - (1+A)\beta}e^{-\beta_0 m}(1 - e^{-\beta m})^{-A} \\ &\times {}_2F_1\left(1 - A, \frac{\beta_0}{\beta} - (1+A), \frac{\beta_0}{\beta} - A, e^{-\beta m}\right).\end{aligned} \tag{43}$$

The parameter C can be obtained by solving the equality $\int_0^\infty s_1(m) = 1$. By using MATHEMATICA, we can obtain

$$C = \left[1 + \frac{A\beta_0 \mathrm{B}(1-A,\beta_0/\beta)}{\beta_0 - (1+A)\beta}\, {}_3F_2\left(1-A, \frac{\beta_0}{\beta}, \frac{\beta_0}{\beta} - 1 - A; \frac{\beta_0}{\beta} - A, \frac{\beta_0}{\beta} + 1 - A; 1\right)\right]^{-1},$$

where

$$ {}_3F_2(a_1,a_2,a_3;b_1,b_2;z) \equiv \frac{\Gamma(b_1)\Gamma(b_2)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_2)} \sum_{n=0}^{\infty} \frac{\Gamma(a_1+n)\Gamma(a_2+n)\Gamma(a_3+n)}{\Gamma(b_1+n)\Gamma(b_2+n)\Gamma(n+1)} z^n \quad \text{for} \quad |z| < 1$$

is the generalized hypergeometric function and its value at 1 is defined by continuation. When $m \to 0$, by the Gaussian formula

$$ {}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \text{ when } \Re(c) > \Re(a+b),$$

hence

$$\lim_{m\to 0} {}_2F_1\left(1-A, \frac{\beta_0}{\beta} - (1+A), \frac{\beta_0}{\beta} - A, e^{-\beta m}\right) = \frac{\Gamma(\beta_0/\beta - A)\,\Gamma(A)}{\Gamma(\beta_0/\beta - 1)} \tag{44}$$

is a finite value. Therefore, when $m \to 0$,

$$s_1(m) \sim (1 - e^{-\beta m})^{-A} \sim m^{-A}. \tag{45}$$

It can be seen that $s_1(m)$ is singular at $m = 0$, but $s_1(m)$ still integrates to 1 over the whole magnitude range if the process is subcritical ($A < 1$). Also, when $m \to \infty$, it is obvious that $s_1(m) \sim e^{-\beta_0 m}$.

Solution (44) is valid only if $\beta_0/\beta > 1 + A$. When $\beta_0/\beta \leqslant 1 + A$, the magnitude distribution of the background events is overridden by the asymptotic magnitude distribution, i.e., by the solution to a reduced version of (38) when $\mu/\bar{\lambda} = 0$:

$$s_1(m) = \int_{\mathcal{M}} \kappa(m^*) s(m\,|\,m^*) s_1(m^*) dm^*. \tag{46}$$

Since the above equation is a special case of $\rho = 1$ of the right eigenequation in Eq. (33), the solution is

$$s_1(m) = C_1\, e^{-(1+A)\beta m}[1 - e^{-\beta m}]^{-A}, \tag{47}$$

where $C_1 = \frac{\beta \sin A\pi}{A\pi}$ is the normalizing constant. This solution has asymptotes of m^{-A} and $e^{-(1+A)\beta m}$ at 0 and ∞, respectively.

In summary, when background events have a constant occurrence rate and a magnitude distribution of the G-R law with parameter β_0, beside $\varrho = A < 1$, a further condition $\beta_0 > (1 + A)\beta$ is required to ensure that the whole process is stable. When $\beta_0 > (1 + A)\beta$ and $A < 1$, the process has a finite occurrence rate and the magnitude distribution for all events has an asymptote of an exponential law with parameter β_0; when $\beta_0 < (1 + A)\beta$ and $A < 1$, the process is not stable with an infinite occurrence rate and the magnitude distribution for all events has an asymptote of an exponential law with parameter $(1 + A)\beta$.

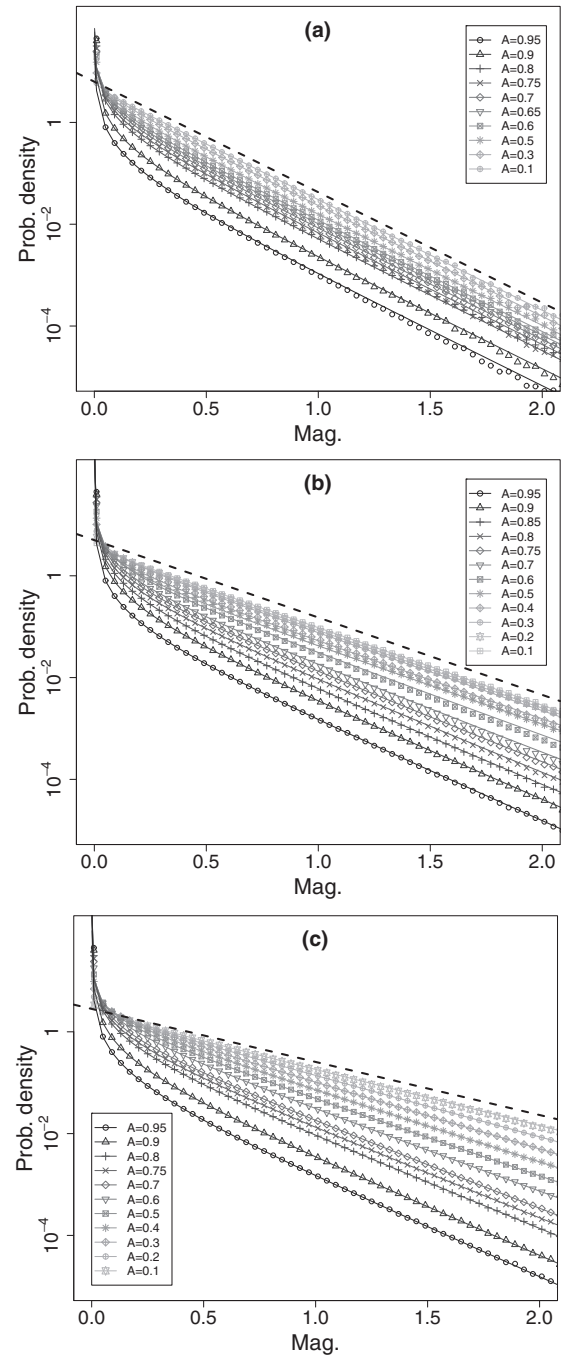

FIG. 1. Probability density functions for the magnitudes of all simulated events. β_0 is fixed at 5.0, 3.5, and 2.4 in (a)–(c), respectively, while β is fixed at 2.3 and A changes from 0.1 to 0.9. The scatter points represent the relative frequency of simulated magnitudes, the solid lines represent the corresponding theoretic results, and the dashed straight lines represent the slopes of β_0 in the density $s_0(m)$ of the background events.

We verify the above results through Monte Carlo simulation. We set the magnitude threshold $m_0 = 0$ and run three sets of simulations, all with model parameter $\beta = 2.3$ and the values of A changing between 0 and 1, but with different β_0 values. In the first set of simulations shown in Fig. 1(a), $\beta_0 = 5.0$, where the condition $\beta_0/\beta - 1 - A > 0$ is always satisfied for $0 < A < 1$, the higher end of the

magnitude density function is controlled by the background magnitude density, irrespective of the value of A. In the second set [Fig. 1(b)], $\beta_0 = 3.5$, the magnitude density functions have asymptotes given by $e^{-\beta(1+A)m}$ if $\beta_0/\beta < 1 + A$ (i.e., $A > 11/24$) and, otherwise, $e^{-\beta_0 m}$. In the third set [Fig. 1(c)] with $\beta_0 = 2.4$, where $\beta_0/\beta - 1 - A < 0$ holds for all selected A values, the high ends of the magnitude density functions have an asymptote as $e^{-\beta(1+A)m}$. In all three cases, there is a singularity point at zero, the magnitude threshold. In summary, the magnitude distribution for all events departs from the empirical Gutenberg-Richter magnitude-frequency relation, but in accordance with the theoretical results discussed above.

B. Case $\Delta > 0$

When $\Delta > 0$, the left eigenequation in Eq. (7) becomes the following functional equation,

$$\varrho a(m) = \frac{\kappa(m)}{F(m+\Delta)} \int_0^{m+\Delta} f(m')a(m')dm. \tag{48}$$

This equation suggests a possible way to solve it by iteration:

$$a^{(0)}(m) = 1,$$
$$a^{(n+1)}(m) = \frac{1}{\varrho}\frac{\kappa(m)}{F(m+\Delta)} \int_0^{m+\Delta} f(m')a^{(n)}(m')dm.$$

However, since $a(m)$ is a monotonic increasing function, the above iteration does not converge. We need to consider an alternative such as $c(m) = a(m)/\kappa(m)$ or $c(m) = a(m)/[\kappa(m)]^2$. It is easy to verify that $c(m) = a(m)/\kappa(m)$ does not decay quickly enough. Assuming that $c(m) = a(m)/[\kappa(m)]^2$, the left eigenequation (48) becomes

$$\varrho c(m)[\kappa(m)]^2 = \frac{\kappa(m)}{F(m+\Delta)} \int_0^{m+\Delta} f(m')c(m')[\kappa(m')]^2\, dm', \tag{49}$$

i.e.,

$$\varrho c(m) = \frac{\int_0^{m+\Delta} f(m')c(m')[\kappa(m')]^2\, dm'}{\kappa(m)F(m+\Delta)} = \frac{A\beta \int_0^{m+\Delta} c(m')\kappa(m')dm'}{\kappa(m)F(m+\Delta)}, \tag{50}$$

by noticing $f(m)\kappa(m) = A\beta$. Now consider the iterative scheme

$$c^{(0)}(m) = 1$$
$$c^{(n+1)}(m) = \frac{A\beta}{\varrho}\frac{\int_0^{m+\Delta} c^{(n)}(m')\kappa(m')dm'}{\kappa(m)F(m+\Delta)}, \quad n = 1,2,\ldots.$$

Then

$$c^{(1)}(m) = \frac{A\beta}{\varrho}\frac{\int_0^{m+\Delta} c^{(0)}(m')e^{\beta m'}\, dm'}{Ae^{\beta m}[1-e^{-\beta(m+\Delta)}]} = \frac{A\, e^{\beta\Delta}}{\varrho},$$
$$c^{(2)}(m) = \frac{A\beta}{\varrho}\frac{\int_0^{m+\Delta} c^{(1)}(m')e^{\beta m'}\, dm'}{A\, e^{\beta m}[1-e^{-\beta(m+\Delta)}]} = \frac{A^2\, e^{2\beta\Delta}}{\varrho^2},$$

and

$$c^{(n)}(m) = \frac{A^n\, e^{n\beta\Delta}}{\varrho^n}.$$

The convergence of $\{c^{(n)}(m)\}$ requires $\varrho \geqslant Ae^{\beta\Delta}$. Note that $\varrho > Ae^{\beta\Delta}$ corresponds to the trivial case of $c(m) = 0$ and $a(m) = 0$, and $\varrho = Ae^{\beta\Delta}$ corresponds to $c(m) = 1$ and $a(m) = [\kappa(m)]^2 = A^2\, e^{2\beta m}$. In summary, the left eigenequation has a nontrivial solution of $a(m) = [\kappa(m)]^2 = A^2\, e^{2\beta m}$ with corresponding eigenvalue $\varrho = A\, e^{\beta\Delta}$.

The right eigenequation takes the form of

$$\varrho b(m) = \int_{\mathcal{M}} K(m,m')b(m')dm = \int_{\max\{0,m-\Delta\}}^{\infty} \frac{A\beta e^{\beta(m'-m)}b(m')}{1-e^{-\beta(m'+\Delta)}}dm'. \tag{51}$$

Take derivatives of both sides,

$$\varrho\frac{db(m)}{dm} = -\beta\int_{\max\{0,m-\Delta\}}^{\infty} \frac{A\beta e^{\beta(m'-m)}b(m')}{1-e^{-\beta(m'+\Delta)}}dm' - K(m,\max\{0,m-\Delta\})b(\max\{0,m-\Delta\})\frac{d\max\{0,m-\Delta\}}{dm}$$
$$= -\beta\varrho b(m) - \begin{cases} 0, & \text{for} \quad m < \Delta, \\ K(m,m-\Delta)b(m-\Delta), & \text{for} \quad m \geqslant \Delta. \end{cases} \tag{52}$$

When $m < \Delta$, according to (52),

$$\varrho\frac{db(m)}{dm} = -\varrho\beta b(m') \Rightarrow b(m) = C_0 e^{-\beta m}, \tag{53}$$

where C_0 could be any positive constant, and for simplification, we take $C_0 = 1$. Substitute the above equation into (51), for $m \leqslant \Delta$,

$$\varrho\, e^{-\beta m} = A\beta\, e^{-\beta m}\int_{\max\{0,m-\Delta\}}^{\infty} \frac{e^{\beta m'}b(m')dm'}{1-e^{-\beta(m'+\Delta)}} = A\beta\, e^{-\beta m}\int_0^{\infty} \frac{e^{\beta m'}b(m')\, dm'}{1-e^{-\beta(m'+\Delta)}}, \tag{54}$$

which gives

$$\varrho = A\beta\int_0^{\infty} \frac{e^{\beta m'}b(m')dm'}{1-e^{-\beta(m'+\Delta)}}. \tag{55}$$

Now we consider the value of $b(m)$ when $m \geqslant 0$. Assume

$$b(m) = D_n(m-n\Delta)e^{-\beta m}, \quad \text{for} \quad n\Delta < m \leqslant (n+1)\Delta,$$
$$n = 0,1,2,\ldots,$$

where $D_{n+1}(0) = D_n(\Delta)$ is required for all non-negative integers n. We only need to find the values of $D_n(x)$ for $x \in (0,\Delta]$ to determine $b(m)$. Substitute the above form of

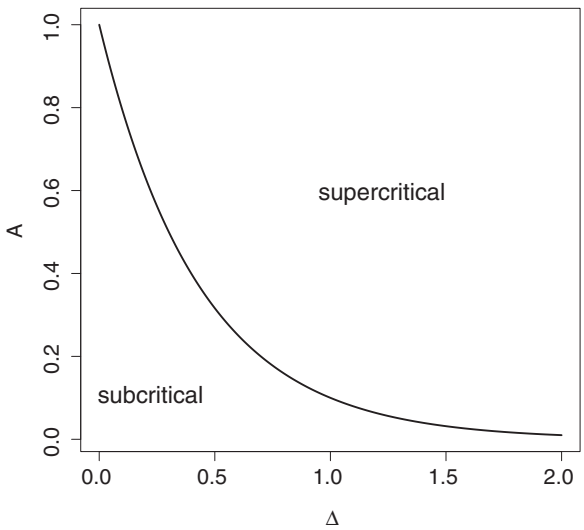

FIG. 2. Subcritical and supercritical regimes for the case of $\beta = \ln 10$ (i.e., $b = 1.0$).

$b(m)$ into (52), we have

$$\varrho \frac{dD_n(m - n\Delta)}{dm} = -\frac{A\beta D_{n-1}(m - n\Delta)}{1 - e^{-\beta m}}, \tag{56}$$

i.e.,

$$\varrho \frac{dD_n(x)}{dx} = -\frac{A\beta D_{n-1}(x)}{1 - e^{-\beta(x+n\Delta)}}, \quad 0 < x \leqslant \Delta, \tag{57}$$

for $n = 1,2,\ldots,$ with $D_0(x) = C_0 = 1$ and $D_n(0) = D_{n-1}(\Delta)$. The above equation can be solved iteratively,

$$D_0(x) = 1, \tag{58}$$

$$D_{n+1}(x) = D_n(\Delta) - \frac{A\beta}{\varrho} \int_0^x \frac{D_n(u)}{1 - e^{-\beta[u+(n+1)\Delta]}}\, du. \tag{59}$$

Considering $\varrho = A\, e^{\beta\Delta}$, the above equation can also be written as

$$D_{n+1}(x) = D_n(\Delta) - \beta\, e^{-\beta\Delta} \int_0^x \frac{D_n(u)}{1 - e^{-\beta[u+(n+1)\Delta]}}\, du. \tag{60}$$

Figure 2 shows the theoretical curve ($A = e^{-\beta\Delta}$) of the boundary between the subcritical and supercritical regimes when $\beta = \ln 10$. We can see from Fig. 2 that when $\Delta = 0$, the process is subcritical when $A = \kappa(0) < 1$, as derived above. Moreover, when Δ increases, A should be decreased to ensure that the process remains in the subcritical regime. The magnitude density function for the overall events in this case is more difficult to obtain than the case of $\Delta = 0$. However, we can still discuss its asymptotic properties in the following way: Consider the integral equation related to the branching ratio

$$\begin{aligned} s_1(m) &= \frac{\mu}{\bar{\lambda}} s_0(m) + \int_{\mathcal{M}} \kappa(m^*) s(m\,|\,m^*) s_1(m^*) dm^* \\ &= \frac{\mu}{\bar{\lambda}} s_0(m) + f(m) \int_{\max(0,m-\Delta)}^{\infty} \frac{\kappa(m^*) s_1(m^*)}{F(m^*)} dm^*, \end{aligned} \tag{61}$$

where $\mu/\bar{\lambda} = 1 - \int_{\mathcal{M}} \kappa(m) s_1(m)\, dm$. Suppose that $s_1(m) \sim D\, e^{-\gamma m}$, when m is sufficiently large, where D and γ are constants. When m is sufficiently large, $F(m) \approx 1$, so the above equation can be rewritten as

$$\begin{aligned} D\, e^{-\gamma m} &\approx \frac{\mu}{\bar{\lambda}} \beta_0\, e^{-\beta_0 m} + \beta\, e^{-\beta m} \int_{m-\Delta}^{\infty} A D\, e^{-(\gamma-\beta)m^*}\, dm^* \\ &= \frac{\mu}{\bar{\lambda}} \beta_0\, e^{-\beta_0 m} + \beta\, e^{-\beta m} \frac{A D\, e^{-(\gamma-\beta)(m-\Delta)}}{\gamma - \beta}. \end{aligned} \tag{62}$$

The following conclusions can be drawn from the above equation:

(1) $\gamma \geqslant \beta_0$. This is because $s(m\,|\,m_i)$ is a truncated probability density, and thus the tail of the magnitude density function does not increase when the direct offspring of any event are added into the population pool.

(2) When $0 < \mu/\bar{\lambda} \leqslant 1$, $\gamma = \beta_0$. It is clear that s_1 can be written in the form of a mixture of two probability densities:

$$s_1(m) = C s_0(m) + (1 - C) s_d(m),$$

where $C = \mu/\bar{\lambda}$ and $s_d(m)$ is a magnitude density function. If $s_1(m) \sim e^{-\gamma m}$, then so is $s_d(m)$. Also, if $\gamma < \beta_0$, then

$$\begin{aligned} 1 &= \lim_{m\to 0} \frac{s_1(m)}{s_1(m)} = C \lim_{m\to 0} \frac{s_0(m)}{s_1(m)} + (1 - C) \lim_{m\to 0} \frac{s_d(m)}{s_1(m)} \\ &= 1 - C, \end{aligned}$$

which conflicts with the assumption that $C > 0$.

(3) When $\mu/\bar{\lambda} = 0$, i.e., $\bar{\lambda} = \infty$ or the whole process has an infinitely large rate, by (62) the solution satisfies $\gamma = \beta[1 + A\, e^{(\gamma-\beta)\Delta}]$.

We use simulations to obtain the magnitude distribution for different cases of Δ and criticality parameters. Figures 3(a)–3(c) show the magnitude PDFs for the cases of $\beta_0 = 5.0$, 3.5, and 2.4, respectively, while Δ is fixed as 0.1, β is fixed as 2.3, and A changes between 0 and 1. Figures 3(d)–3(f) correspond to Figs. 3(a)–3(c), respectively, except that $\Delta = 0.1$ is replaced by $\Delta = 0.5$. When A is small enough, for example, $A = 0.1$ in Figs. 3(a)–3(e) or $A = 0.05$ in Fig. 3(e), the overall magnitude distribution can be well approximated by the G-R magnitude-frequency relation. However, in this case, what controls the global slope is β_0 but not β. When A is big and Δ is small, for example, $A = 0.6$–1.0 in Figs. 3(a)–3(c), the magnitude-frequency (magnitude probability density) curves can be divided into three parts: The tails have the asymptotic slopes as we discussed above; the middle (transitive) part corresponds to the transitive part in Fig. 4 from the singular point at $m = 0$ to the part with the asymptomatic slope; unlike the case of $\Delta = 0$, the starting part is not singular, but is replaced by a short and approximately straight line segment. When A is big and Δ is also big, e.g., $A = 0.6$–1.0 in Figs. 3(d)–3(f), the transitive part disappears and the starting short line segments are connected to the asymptotic-slope parts directly.

Based on the above analysis of Eq. (62) and the simulation results, we can summarize the asymptotic behavior of $s_1(m)$ as in Fig. 4. In this figure, the vertical dashed lines mark the cases for $\beta_0 = 2.4$, 3.5, and 5.0, the gray solid lines with unit slope mark the function $y = \gamma$, and the other gray solid curves mark the functions of $y = \beta[1 + A\, e^{(\gamma-\beta)\Delta}]$ with A varying between 0 and 1. The black lines with arrows in each panel show how the slope changes for different β_0 when A changes. For a more general discussion, we plot the cases of $\Delta = 0$, 0.1, 0.5 in Figs. 4(a)–4(c), respectively. We can draw the following conclusions:

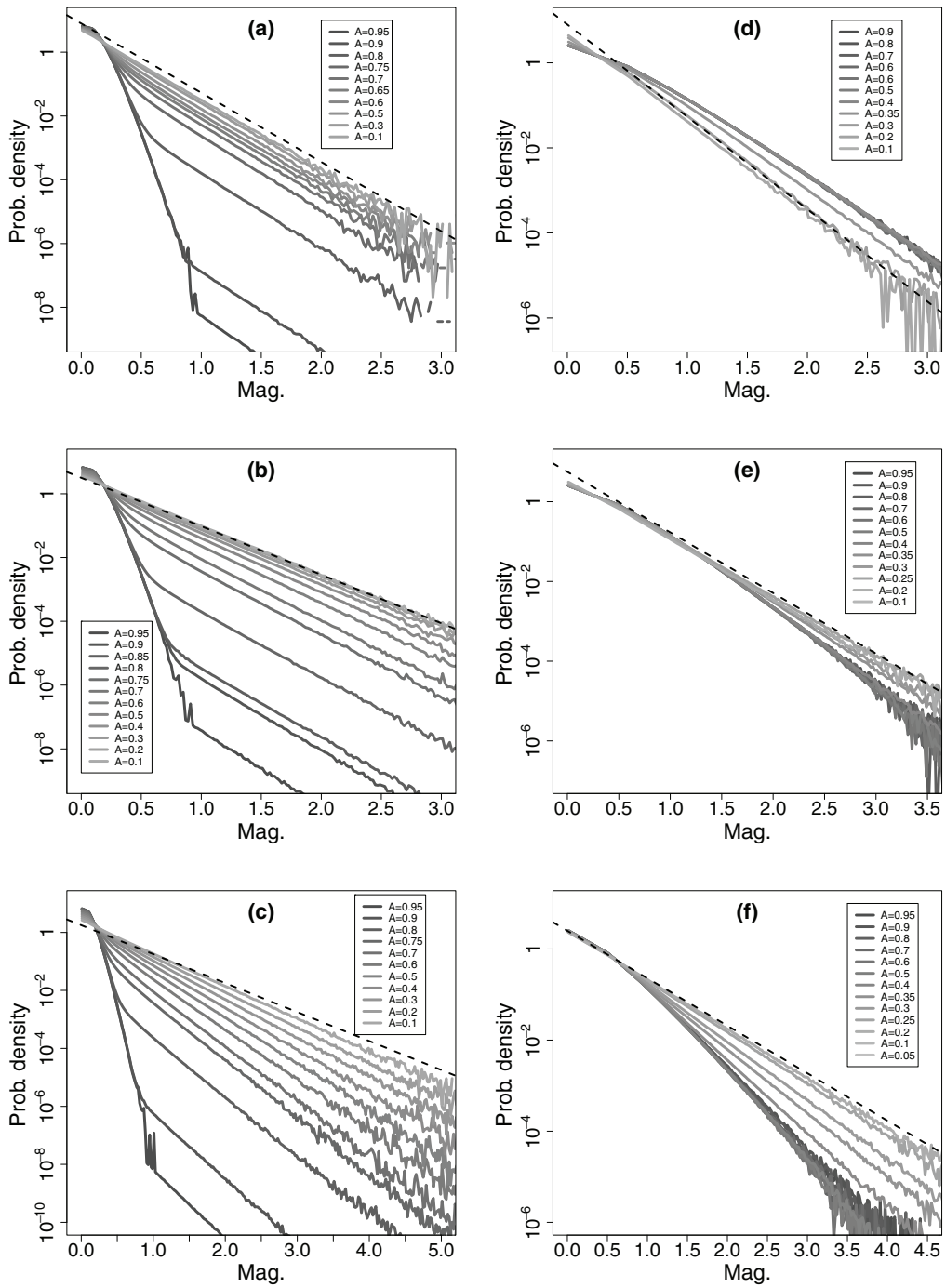

FIG. 3. Magnitude PDFs for all the events obtained by simulation for the BASS model with different parameters: (a) $\Delta = 0.1$ and $\beta_0 = 5.0$, (b) $\Delta = 0.1$ and $\beta_0 = 3.5$, (c) $\Delta = 0.1$ and $\beta_0 = 2.4$, (d) $\Delta = 0.5$ and $\beta_0 = 5.0$, (e) $\Delta = 0.5$ and $\beta_0 = 3.5$, and (f) $\Delta = 0.5$ and $\beta_0 = 2.4$, while β is fixed as 2.3, and A changes between 0 and 1, represented by curves of different gray scales. The dashed lines represent the slope β_0 in $s_0(m)$ of the background events.

(1) The asymptotic slope is determined by β_0 and the first intersection between $y = \gamma$ and $y = \beta(1 + A\,e^{(\gamma-\beta)\Delta})$, i.e., $\max\{\beta_0, \gamma_0\}$, where γ_0 is the smaller solution of $\gamma = \beta(1 + A\,e^{(\gamma-\beta)\Delta})$, if such solutions exist.

(2) When A is large enough such that a solution of $\gamma = \beta(1 + A\,e^{(\gamma-\beta)\Delta})$ does not exist, the asymptotic slope is determined by the point at which $y = \gamma$ cuts $\gamma = \beta(1 + A_0\,e^{(\gamma-\beta)\Delta})$ for a certain A_0, i.e., the asymptotic slope is the corresponding γ if it is greater than β_0, or β_0, otherwise. As shown in Fig. 4(c), the gray dashed curve is the one tangent to $y = \gamma$ among the set of curves $\{y = \beta[1 + A\,e^{(\gamma-\beta)\Delta}] : \ A \in [0, +\infty)\}$ and P is the cutting point. It can be shown that the solution at P is $\gamma = \beta + \frac{1}{\Delta}$.

(3) If β_0 is greater than all the values of γ_0, which correspond to the first intersection points between $y = \gamma$ and $y = \beta(1 + Ae^{(\gamma-\beta)\Delta})$ for possible values of A, then the slope

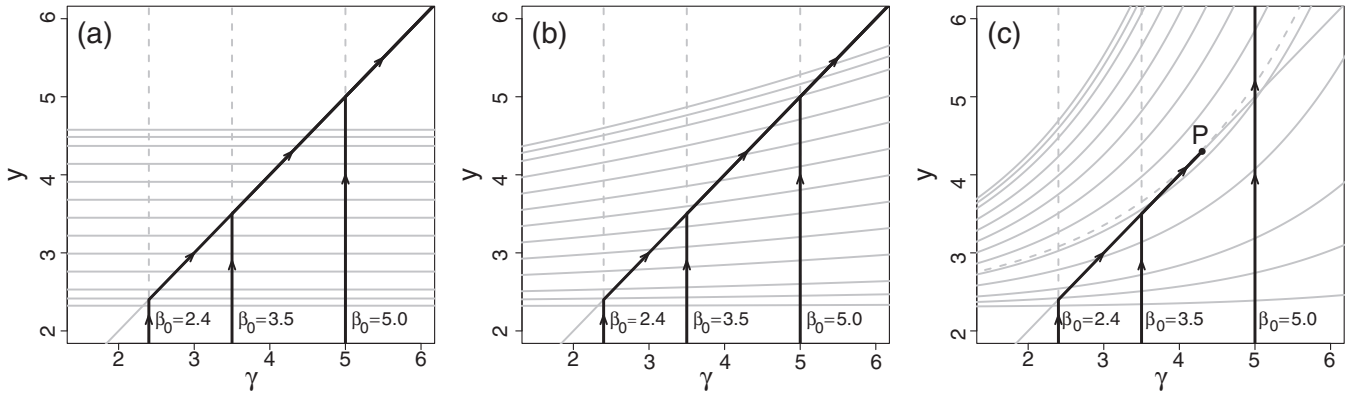

FIG. 4. Illustrations of obtaining the asymptotic slopes of the magnitude PDFs in Figs. 1 and 3: (a) $\Delta = 0$, (b) $\Delta = 0.1$, and (c) $\Delta = 0.5$. The vertical dashed lines mark the cases for $\beta_0 = 2.4$, 3.5, and 5.0, the gray solid lines with unit slope mark the function $y = \gamma$, and the other gray solid lines mark the functions of $y = \beta[1 + A\,e^{(\gamma-\beta)\Delta}]$ with A varying between 0 and 1. The gray dashed curve in (c) is the one which is tangent to $y = \gamma$ among the set of curves $\{y = \beta[1 + A\,e^{(\gamma-\beta)\Delta}] : A > 0\}$ at P. The black lines with arrows in each panel show the paths of how the slope changes for different β_0.

always remains β_0. For example, when $\Delta = 0.5$, $\beta_0 = 5.0$ and $\beta = 2.3$ in Fig. 4(c).

Moreover, there is a mode in the criticality regime where $\omega = 1$ and $\varrho < 1$. Here at this mode, each family line in the process eventually becomes extinct with probability 1 and a finite total family size, but the whole process does not have a finite occurrence rate. This shows again that $\varrho < 1$ is only a necessary condition but not a sufficient one. To distinguish different subcritical modes of the process, we say that the process is *essentially subcritical* when $\varrho < 1$ and $\omega < 1$, and *semicritical* or *pseudosubcritical* when $\varrho < 1$ and $\omega = 1$. From the above analysis, we can see that the process is semicritical when A takes values in (A_1, A_2) where $A_1 = (\beta_0/\beta - 1)e^{(\beta-\beta_0)\Delta}$ corresponds to the first turning point on the path of possible γ values in Fig. 4, and $A_2 = \min\{e^{-\beta\Delta}, \frac{1}{\beta\Delta e}\}$, with $A = e^{-\beta\Delta}$ corresponding to the case when the process is critical (i.e., $\varrho = A\,e^{\beta\Delta} = 1$), and $A = \frac{1}{\beta\Delta e}$ corresponding to the case when $y = \gamma$ cuts $y = \beta[1 + A\,e^{(\gamma-\beta)\Delta}]$ (i.e., $\gamma = \beta + 1/\Delta$). If $A_2 < A_1$, then (A_1, A_2) is the empty set.

C. Magnitude intensity of the first generation

Unlike the magnitude distribution $s_1(m)$ of all events, the magnitude distribution of the first generation can be derived in explicit form and can provide us with some insights into the overall magnitude distribution, especially in the subcritical case where the first generation comprises the majority of all the descendants. In this section, we will focus on the expected magnitude intensity of the first generation triggered by the background events, i.e.,

$$\begin{aligned} s_f(m) &= \int_{\mathcal{M}} \kappa(m')\, s(m|m') s_0(m')\, dm' \\ &= \int_{\max(0,m-\Delta)}^{\infty} A\, e^{\beta m'} \frac{\beta\, e^{-\beta m}}{1 - e^{-\beta(m'+\Delta)}} \beta_0 e^{-\beta_0 m'}\, dm' \\ &\quad (\text{let } u = e^{-\beta m'}) \\ &= A\beta_0 e^{-\beta m} \int_0^{e^{-\beta \max(0,m-\Delta)}} u^{\frac{\beta_0}{\beta}-2}(1 - u e^{-\beta\Delta})^{-1} du \qquad (63) \\ &= \frac{A\,\beta_0\,\beta}{\beta_0 - \beta}\, e^{-\beta \min(m,\Delta) - \beta_0 \max(0,m-\Delta)} \\ &\quad \times {}_2F_1\left(1, \frac{\beta_0}{\beta} - 1; \frac{\beta_0}{\beta}; e^{-\beta \max(m,\Delta)}\right). \qquad (64) \end{aligned}$$

A special case of the above equation is

$$s_f(m) = -\beta_0 e^{-\beta(m-\Delta)} \ln[1 - e^{-\beta \max(m,\Delta)}], \text{ when } \beta_0 = 2\beta, \qquad (65)$$

obtained by using the equality $\ln(1 - z) = -z\,{}_2F_1(1,1;2;z)$ for $|z| < 1$. To analyze the above equation, we consider the two cases of large and small values of m:

(1) Large m. To obtain the asymptotic behavior of $s_f(m)$ for large m, substitute the expansion $(1 - ue^{\beta\Delta})^{-1} \sim 1 + O(1)$ into (63), and then

$$s_f(m) \sim e^{-\beta_0 m}. \qquad (66)$$

This implies the tail of $s_f(m)$ is controlled by the magnitude structure of the background events.

(2) Small m. When $\Delta \neq 0$, it is easy to see $s_f(m) \sim e^{-\beta m}$. When $\Delta = 0$, using the following from page 63 of Andrews *et al.* [38],

$$\lim_{x\to 1-} {}_2F_1(a,b;a+b;x)/\ln[1/(1-x)] = \mathrm{B}(a,b),$$

where B is the beta function, we can obtain $s_f(m) \sim -\ln m$ by using the approximation $1 - e^{-\beta \max(m,\Delta)} \approx \beta m$. That is to say, in this case, when $\Delta = 0$, $s_f(m)$ has a singularity point at 0 with $-\ln m$ as the asymptote, and this singularity disappears with $\Delta > 0$ where the asymptote is replaced by $e^{-\beta m}$.

VIII. SUMMARY

In this article, we considered the stability conditions and some asymptotic properties of a general class of branching processes in which the magnitude distribution of offspring depends on the magnitude of the parent. This class includes the ETAS model as a special case in which offspring magnitudes are independent of parents, but the class also encompasses more general models, such as the BASS model and other recently proposed models.

For this general class, we derived the equations that determine the criticality parameter and the branching ratio, each of them describing a different property of the process. The criticality parameter characterizes the average number of offspring of an arbitrary parent after infinitely many generations, while the branching ratio describes the average number of offspring over all generations. For the special case of the ETAS model, where the magnitude component satisfies the Gutenberg-Richter law and can be separated from the whole process, these two quantities are identical, but they may differ in the more general case.

To explore these quantities, we reformulated the BASS model as a well-defined Poisson cluster process with a background rate, a Poisson-distributed number of offspring, and parent-dependent magnitude density of the offspring. We found that the model is asymptotically unstable unless exponents of the productivity law and the magnitude densities are suitably modified, or unless other conditions are imposed, such as a truncation of the magnitude density of the offspring.

We proposed one particular version of the BASS model that ensures that the model can be subcritical, in which the magnitude density of aftershocks is truncated from above by $m' + \Delta$, where m' is the parent's magnitude and Δ is a constant. We then explored the stability conditions of this model, including the asymptotic properties of the compound magnitude distribution. We analyzed these properties analytically by deriving expressions for the left and right eigenfunctions and the maximum eigenvalue of the equation that describes the branching. Based on these results and on numerical simulations of this modified BASS model, we derived stability conditions and showed that in the subcritical regime, there is a mode where the process is not stable with a finite occurrence rate but where all family trees still eventually become extinct. Furthermore, the compound magnitude distribution of all events may differ substantially from the Gutenberg-Richter distribution, unless the process is essentially subcritical or Δ is relatively large.

ACKNOWLEDGMENTS

J.Z. is partially supported by supported by Grants-in-Aid No. 22700299 for Young Scientists (B) and No. 2530052 for Scientific Research (C), both from the Japan Society for the Promotion of Science. M.J.W. was partially supported by the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement No. EAR-1033462 and USGS Cooperative Agreement No. G12AC20038. We thank David Vere-Jones for early discussions for stimulating this research. The authors thank Pathikrit Bhattacharya for stimulating discussions, and two anonymous reviewers for their insightful and constructive comments and encouragement.

[1] Y. Y. Kagan and L. Knopoff, J. Geophys. Res. **86**, 2853 (1981).
[2] F. Musmeci and D. Vere-Jones, Ann. Inst. Stat. Math. **44**, 1 (1992).
[3] Y. Y. Kagan, Geophys. J. Int. **106**, 135 (1991).
[4] Y. Ogata, Ann. Inst. Stat. Math. **50**, 379 (1998).
[5] Y. Ogata and J. Zhuang, Tectonophysics **413**, 13 (2006).
[6] Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988).
[7] Y. Y. Kagan and L. Knopoff, Science **236**, 1563 (1987).
[8] R. Console, M. Murru, and A. M. Lombardi, J. Geophys. Res. **108**, 2468 (2003).
[9] A. Helmstetter, Y. Y. Kagan, and D. D. Jackson, Bull. Seismol. Soc. Am. **96**, 90 (2006).
[10] J. Zhuang, A. Christophersen, M. K. Savage, D. Vere-Jones, Y. Ogata, and D. D. Jackson, J. Geophys. Res. **113**, B11302 (2008).
[11] J. Zhuang, Earth Planet Space **63**, 207 (2011).
[12] M. J. Werner, A. Helmstetter, D. D. Jackson, and Y. Y. Kagan, Bull. Seismol. Soc. Am. **101**, 1630 (2011).
[13] D. D. Daley and D. Vere-Jones, *An Introduction to Theory of Point Processes–Volume 1: Elementary Theory and Methods*, 2nd ed. (Springer, New York, 2003).
[14] A. Veen and F. P. Schoenberg, J. Am. Stat. Assoc. **103**, 614 (2008).
[15] A. Helmstetter and D. Sornette, J. Geophys. Res. **107**, 2237 (2002).
[16] D. Vere-Jones, Adv. Appl. Probab. **37**, 908 (2005).
[17] A. Saichev and D. Sornette, Phys. Rev. E **72**, 056122 (2005).
[18] E. Lippiello, C. Godano, and L. de Arcangelis, Phys. Rev. Lett. **98**, 098501 (2007).
[19] D. L. Turcotte, J. R. Holliday, and J. B. Rundle, Geophys. Res. Lett. **34**, L12303 (2007).
[20] J. R. Holliday, D. L. Turcotte, and J. B. Rundle, Physica A: Stat. Mech. Appl. **387**, 933 (2008).
[21] F. Omori, J. Coll. Sci., Imp. Univ. Tokyo **7**, 111 (1895).
[22] T. Utsu, Zisin J. Seismol. Soc. Jpn. **10**, 35 (1957).
[23] J. Zhuang, Y. Ogata, and D. Vere-Jones, J. Geophys. Res. **109**, B05301 (2004).
[24] E. Lippiello, C. Godano, and L. de Arcangelis, Geophys. Res. Lett. **39**, 051083 (2012).
[25] J. Zhuang, C.-P. Chang, Y. Ogata, and Y.-I. Chen, J. Geophys. Res. **110**, B05S13 (2005).
[26] D. Sornette and M. J. Werner, J. Geophys. Res. **110**, B08304 (2005).
[27] J. Zhuang and Y. Ogata, Phys. Rev. E **73**, 046134 (2006).
[28] A. Saichev and D. Sornette, Phys. Rev. E **71**, 056127 (2005).
[29] D. Vere-Jones and J. Zhuang, Phys. Rev. E **78**, 047102 (2008).
[30] J. Zhuang, J. R. Stat. Soc.: Ser. B **68**, 635 (2006).
[31] D. S. Harte, Geophys. J. Int. **192**, 390 (2013).
[32] D. Vere-Jones, Adv. Appl. Probab. **40**, 882 (2008).
[33] A. Helmstetter and D. Sornette, Geophys. Res. Lett. **30**, 2069 (2003).

[34] Y. Y. Kagan and F. Schoenberg, J. Appl. Probab. **38A**, 158 (2001).
[35] E. Lippiello, L. de Arcangelis, and C. Godano, Phys. Rev. Lett. **100**, 038501 (2008).
[36] E. Lippiello, L. de Arcangelis, and C. Godano, Int. J. Mod. Phys. B **23**, 5583 (2009).
[37] *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, edited by M. Abramowitz and I. A. Stegun, 1st ed. (Dover, New York, 1972), p. 1046.
[38] *Special Functions*, edited by G. E. Andrews, R. Askey, and R. Roy, 1st ed. (Cambridge University Press, Cambridge, UK, 1999), p. 664.