Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks

Jiancang Zhuang* and Yosihiko Ogata†
Institute of Statistical Mathematics, Research Organization of Information and Systems, 4-6-7 Minami Azabu, Minato-Ku, Tokyo 106-8659, Japan
(Received 21 November 2005; revised manuscript received 16 March 2006; published 27 April 2006)

The space-time epidemic-type aftershock sequence model is a stochastic branching process in which earthquake activity is classified into background and clustering components and each earthquake triggers other earthquakes independently according to certain rules. This paper gives the probability distributions associated with the largest event in a cluster and their properties for all three cases when the process is subcritical, critical, and supercritical. One of the direct uses of these probability distributions is to evaluate the probability of an earthquake to be a foreshock, and magnitude distributions of foreshocks and nonforeshock earthquakes. To verify these theoretical results, the Japan Meteorological Agency earthquake catalog is analyzed. The proportion of events that have 1 or more larger descendants in total events is found to be as high as about 15%. When the differences between background events and triggered event in the behavior of triggering children are considered, a background event has a probability about 8% to be a foreshock. This probability decreases when the magnitude of the background event increases. These results, obtained from a complicated clustering model, where the characteristics of background events and triggered events are different, are consistent with the results obtained in [Ogata *et al.*, Geophys. J. Int. **127**, 17 (1996)] by using the conventional single-linked cluster declustering method.

DOI: 10.1103/PhysRevE.73.046134 PACS number(s): 64.60.Ak, 02.50.Ey, 91.30.Dk

I. INTRODUCTION

The foreshock has been one of the most important topics in the research of seismicity pattern and earthquake prediction for decades (see, e.g., [1–22]). Some researchers believe that there is no difference between foreshocks and mainshocks except in magnitudes of following earthquakes, that is to say, foreshocks are mainshocks whose aftershocks happen to be bigger ([23–29]). This assertion leads to a proposition that prediction made based on foreshock studies cannot be improved from prediction based on clustering models.

To test the above assertion, we need a good model of earthquake clustering. Such a test can be carried out by finding whether there are distinguishing features between the patterns of foreshocks and mainshocks in triggering seismicity. Among them, it is important to know the probability of foreshocks under the assumptions of a simple clustering model.

There have been many studies on the modeling of earthquake clustering. The most primary model is to use the modified Omori formula for occurrence times of aftershocks, and the Gutenber-Richter law for their magnitude ([12,23,24]). Console ([30]) and Savage and dePolo ([31]) use this model to evaluate the probability that an earthquake has one or more aftershocks larger in magnitude.

A more advanced type of clustering models is branching type ([32–43]). Among them, the epidemic-type aftershock sequence (ETAS) models are generally adopted (see, e.g., [38,44]). These models assume that the seismicity can be divided into a background component and a clustering component, and that each event, no matter whether it is from the background or directly triggered by another event, triggers its own offspring according to some general rules.

In this paper, we first give a brief description of necessary concepts associated with the ETAS model in this study, then evaluate the probability that the magnitude of the largest descendant is less than a given magnitude. Based on this probability, we derive the probability density of the magnitude distribution of foreshocks. To verify our theory, the Japan Meterological Agency (JMA) catalog is analyzed using the stochastic reconstruction method in Sec. V.

II. DEFINITION OF THE ETAS MODEL

In the ETAS model, the time-varying seismicity rate (mathematically termed as conditional intensity; see, e.g., [45], Chap. 7) takes the form of

$$\lambda(t,x,y,m) = \mu(t,x,y,m) + \sum_{i:t_i<t} \xi(t,x,y,m;t_i,x_i,y_i,m_i), \tag{1}$$

where $\mu(t,x,y,m)$ represents the background seismicity rate and $\xi(t,x,y,m;t_i,x_i,y_i,m_i)$ is the contribution to seismicity rate by the ith event occurring previously. In this model, there is no difference in triggering seismicity between foreshocks, mainshocks, and aftershocks. In practice, it is usual to make the following assumptions:

(1) The whole process is magnitude separable, i.e.,

*Now at Department of Earth and Space Sciences, University of California, Los Angeles, CA 90095–1567. Electronic address: zhuangjc@ism.ac.jp, zhuang@moho.ess.ucla.edu
†Electronic address: ogata@ism.ac.jp

$$\lambda(t,x,y,m) = \lambda(t,x,y)s(m), \tag{2}$$

where

$$\lambda(t,x,y) = \mu(t,x,y) + \sum_{i:t_i<t} \xi(t,x,y;t_i,x_i,y_i,m_i), \tag{3}$$

and

$$s(m) = \beta e^{-\beta(m-m_c)}, \quad m \geqslant m_c \tag{4}$$

is the probability density function (pdf) for all the magnitudes (Gutenburg-Richter law), β being a constant and m_c being the magnitude threshold.

(2) The background rate is stationary, i.e.,

$$\mu(t,x,y) = \mu(x,y); \tag{5}$$

(3) The response function $\xi(t,x,y;t',x',y',m')$ is separable, in the form of

$$\xi(t,x,y;t',x',y',m') = \kappa(m')g(t-t')f(x-x',y-y',m'), \tag{6}$$

where

$$\kappa(m) = Ae^{\alpha(m-m_c)}, \quad m \geqslant m_c, \tag{7}$$

is the mean number of children (direct offspring) from an event sized m, and

$$g(t) = \frac{p-1}{c}\left(1+\frac{t}{c}\right)^{-p}, \quad t>0, \tag{8}$$

and

$$f(x,y;m) = \frac{q-1}{\pi D e^{\gamma(m-m_c)}}\left(1+\frac{x^2+y^2}{De^{\gamma(m-m_c)}}\right)^{-q} \tag{9}$$

are the pdf's for the occurrence times and locations of children (direct offspring), respectively. In Eqs. (7)–(9), A, α, c, p, D, q, and γ are constant parameters.

In this paper, we call $\kappa(m)$ the triggering ability from an event of magnitude m. Equation (8) is the probability density form of the Omori-Utsu formula (see [46], [47], or, [48] for a review). The form of (9) is based on the analysis in [43] and [44].

For a realization of the process, $\{(t_i,x_i,y_i,m_i): i=1,\ldots,N\}$, in a spatial region S and a time interval $[0,T]$, the logarithm-likelihood has the standard form

$$\ln L = \sum_i \ln \lambda(t_i,x_i,y_i,m_i) - \int_0^T \int\int_S \lambda(t,x,y)\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}y, + \sum_i \ln s(m_i), \tag{10}$$

where the subscript i runs over all the events in the region S and time interval $[0,T]$. The parameters in the model can be estimated by maximizing the logarithm-likelihood function. For the version of a model with nonhomogeneous background, the estimation procedure can be found in [41–43].

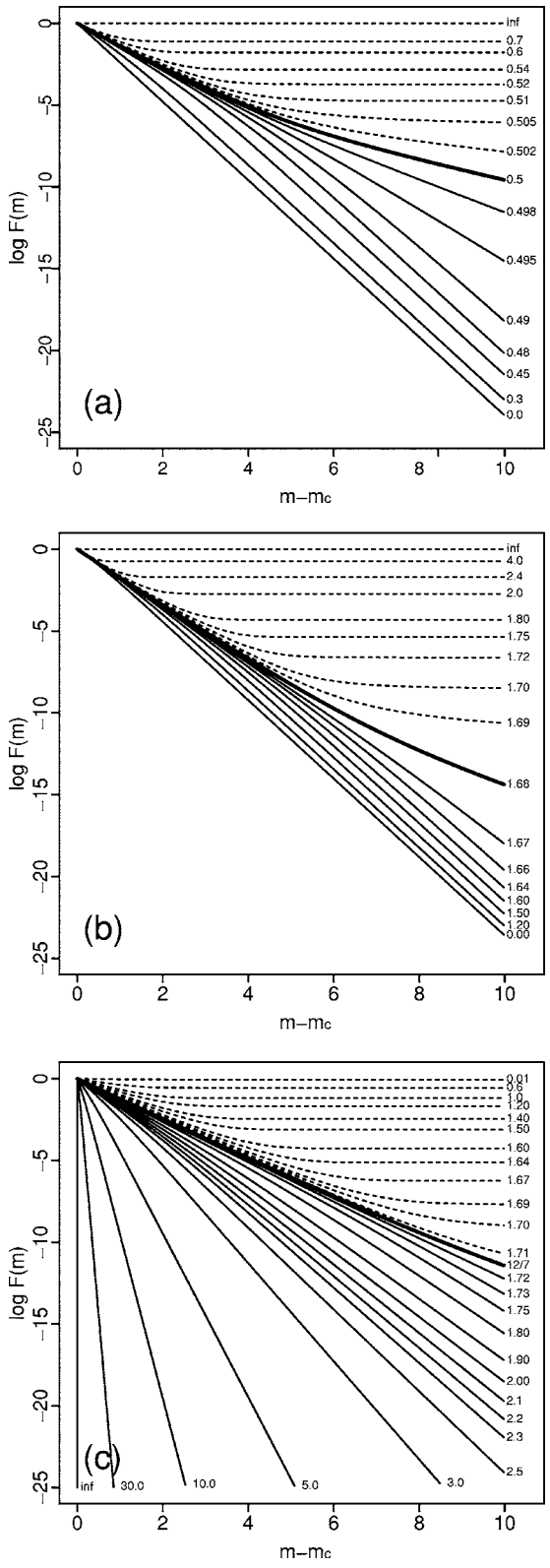

FIG. 1. Influence of A, α,and β to function $F(m)$. (a) Parameter A changes, but $\alpha=1.2$ and $\beta=2.4$ are fixed; (b) Parameter α changes, but $A=0.3$ and $\beta=2.4$ are fixed; (c) Parameter β changes, but $A=0.3$ and $\alpha=1.2$ are fixed. The thin solid, thick solid, and dashed curves represent the subcritical, critical, and supercritical regimes, respectively.

III. THEORETICAL DISTRIBUTIONS ASSOCIATED WITH THE LARGEST MAGNITUDE OF ALL THE DESCENDANTS FROM A GIVEN EVENT

A. Basic equations

Note the number of children from an event of magnitude m is a Poisson random variable with a mean of $\kappa(m)$. The probability that an event of magnitude m has no offspring greater than m' can be derived from the model analytically ([49]), i.e.,

$$\begin{aligned}\zeta(m,m') &= \Pr\{\text{an event sized } m \text{ has no offspring } > m'\} \\ &= \sum_{n=0}^{\infty} \Pr\{\text{each child is } \leqslant m' \text{ and has no offspring} \\ &\quad > m' | m \text{ has } n \text{ direct offspring}\} \\ &\quad \times \Pr\{m \text{ has } n \text{ children}\} \\ &= \sum_{n=0}^{\infty} \left[\int_{m_c}^{m'} s(u)\zeta(u,m')\mathrm{d}u \right]^n \frac{[\kappa(m)]^n}{n!} \mathrm{e}^{-\kappa(m)} \\ &= \exp\left\{ -\kappa(m) \left[1 - \int_{m_c}^{m'} s(m^*)\zeta(u,m')\mathrm{d}u \right] \right\}. \end{aligned} \quad (11)$$

It is evident that $\zeta(m,m')$ has the form

$$\zeta(m,m') = \exp[-\kappa(m)F(m')], \quad (12)$$

where

$$F(m') = 1 - \int_{m_c}^{m'} s(m^*)\exp[-\kappa(m^*)F(m')]\mathrm{d}m^* \quad (13)$$

represents the probability that the largest earthquake in an arbitrary cluster, including the initial event and all its descendants, is greater than m'.

Because $F(m)$ is a function determined by $\kappa(m)$ and $s(m)$, it is influenced only by the parameters A, α, and β in the model. Substituting Eqs. (4) and (7) into Eq. (13),

$$\begin{aligned}F(m') &= 1 - [\Gamma_{-\beta/\alpha}[AF(m')] - \Gamma_{-\beta/\alpha}(AF(m')\mathrm{e}^{\alpha(m'-m_c)})] \\ &\quad \times \frac{\beta}{\alpha}[AF(m')]^{\beta/\alpha}, \end{aligned} \quad (14)$$

where $\Gamma_a(x) = \int_x^{\infty} \mathrm{e}^{-u}u^a\mathrm{d}u$ is the complementary incomplete gamma function. Based on the analysis with moment generating functions, Saichev and Sornette [50] also gave similar forms of the above equations. But, they did not discuss the case when the process is supercritical. In this paper, we are going to discuss the properties of ζ under all the three cases: (1) subcritical case, where each family tree dies off finally and the whole process is stable and stationary; (2) critical case, where each family tree dies off with a long tail and the population of the whole process in unit time increases unboundedly; and, (3) supercritical case, where some of the family trees may never die off and the population of the whole process will be explosive.

According to Appendix A, the critical parameter ρ is determined by

$$\varrho = \int_{m_c}^{+\infty} \kappa(m)s(m)\mathrm{d}m. \quad (15)$$

Substituting (4) and (7) into the above equation, we have

$$\varrho = \frac{A\beta}{\beta - \alpha}. \quad (16)$$

It is evident that $\beta > \alpha$ is required. Also, when A increases, α increases, or β decreases, the critical parameter ϱ increases. As is shown later, the influences of these three parameters, A, α, and β, to ζ and F is mainly due to their influences on the critical parameter ϱ through the above equation.

Function $F(m)$ is closely related to the extinct probability of the family tree starting from an event of magnitude m, namely $P_c(m)$. $P_c(m)$ can be derived in the following way:

$$\begin{aligned}P_c(m) &= \Pr\{\text{The family tree from an event of magnitude } m \text{ extinguishes}\} \\ &= \Pr\{\text{an event of magnitude } m \text{ produces finite number of offspring}\} \\ &= \sum_{n=0}^{\infty} \Pr\{\text{each child produces finite number of offspring} | m \text{ has } n \text{ children}\} \Pr\{m \text{ has } n \text{ children}\} \\ &= \sum_{n=0}^{\infty} \left[\int_{m_c}^{+\infty} s(m^*)P_c(m^*)\mathrm{d}m^* \right]^n \frac{[\kappa(m)]^n}{n!} \mathrm{e}^{-\kappa(m)} = \exp\left[-\kappa(m)\left(1 - \int_{m_c}^{+\infty} s(m^*)P_c(m^*)\mathrm{d}m^* \right) \right]. \end{aligned} \quad (17)$$

Substitute $P_c(m) = \exp[-C\kappa(m)]$ into Eq. (17); we have

$$C = 1 - \int_{m_c}^{+\infty} s(m^*)\exp[-C\kappa(m^*)]\mathrm{d}m^*. \quad (18)$$

Substitute Eqs. (4) and (7) into Eq. (18),

$$C = 1 - \frac{\beta}{\alpha}C^{-\beta/\alpha}\Gamma\left(-\frac{\beta}{\alpha}, C \right). \quad (19)$$

Compare Eq. (18) to Eq. (13),

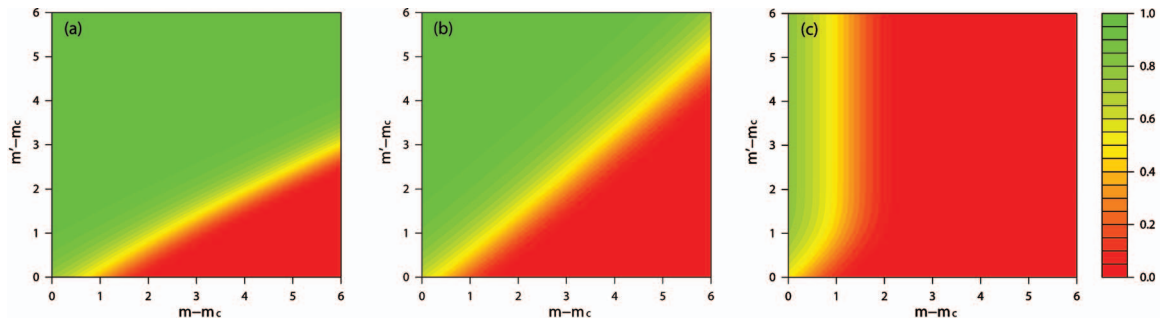

FIG. 2. (Color) Image of the function $\zeta(m,m')$: (a) Subcritical case with $A=0.3$, $\alpha=1.2$, and $\beta=2.4$; (b) Critical case with $A=0.5$, $\alpha=1.2$, and $\beta=2.4$; and (c) supercritical case with $A=0.7$, $\alpha=1.2$, and $\beta=2.4$.

$$\lim_{m\to+\infty} F(m) = C = -\frac{\ln P_c(m)}{\kappa(m)}. \tag{20}$$

It is easy to prove that (18) has one solution C in (0,1) if and only if the process is supercritical, i.e., $\varrho=\int_{m_c}^{\infty}\kappa(m)s(m)\mathrm{d}m>1$.

B. Criticality and $\zeta(m,m')$

For the subcritical case, which requires $\beta>\alpha$ and $\varrho=A\beta/(\beta-\alpha)<1$, it is easy to see that $\zeta(m,m')\to 1$ when $m'\to+\infty$ because $C=0$. That is to say, when the process is subcritical, the larger the event, the less chance that it has a larger descendant. To discuss how fast ζ tends to 1, or how fast $\ln\zeta$ tends to zero, it is useful to use the following approximation. If φ is not an integer,

$$\Gamma_\varphi(x_1)-\Gamma_\varphi(x_2)=\int_{x_1}^{x_2} u^{\varphi-1}\mathrm{e}^{-u}\,\mathrm{d}u=\sum_{n=0}^{+\infty}\frac{(-1)^n(x_2^{n+\varphi}-x_1^{n+\varphi})}{n!(n+\varphi)}; \tag{21}$$

if $\varphi=-k$ is a nonpositive integer, we can replace the kth item in the summation by $(-1)^k\ln(x_2/x_1)/k!$. Equations (14) and (21) give

$$F(m)=1-\frac{\beta}{\alpha}[AF(m)]^{\beta/\alpha}\sum_{n=0}^{+\infty}(-1)^n X_n(m), \tag{22}$$

where

$$X_n(m)=\begin{cases}\dfrac{[AF(m)\mathrm{e}^{\alpha(m-m_c)}]^{n-(\beta/\alpha)}-[AF(m)]^{n-(\beta/\alpha)}}{n!\left(n-\dfrac{\beta}{\alpha}\right)}, & n\neq\dfrac{\beta}{\alpha}\\[2ex] \dfrac{\alpha(m-m_c)}{n!}, & n=\dfrac{\beta}{\alpha}\end{cases} \tag{23}$$

i.e.,

$$F(m)=\mathrm{e}^{-\beta(m-m_c)}+\frac{A\beta}{\beta-\alpha}F(m)[1-e^{(\alpha-\beta)(m-m_c)}]+\cdots. \tag{24}$$

Because we are looking for the solution of $F(m)$ such that $F(m)\mathrm{e}^{\alpha(m-m_c)}\to 0$ when $m\to\infty$, Eq. (24) can approximate by keeping the first two terms, which gives

$$F(m)\approx\frac{\mathrm{e}^{-\beta(m-m_c)}}{1-\dfrac{A\beta}{\beta-\alpha}[1-\mathrm{e}^{(\alpha-\beta)(m-m_c)}]}=\frac{\mathrm{e}^{-\beta(m-m_c)}}{1-\varrho[1-\mathrm{e}^{(\alpha-\beta)(m-m_c)}]}. \tag{25}$$

The above equation implies that, when the process is subcritical,

$$\lim_{m\to+\infty}\frac{F(m)}{s(m)}=\frac{1}{\beta(1-\varrho)}. \tag{26}$$

Furthermore, in the subcritical case, according to Eqs. (12) and (26), for a fixed m, when m' is large enough,

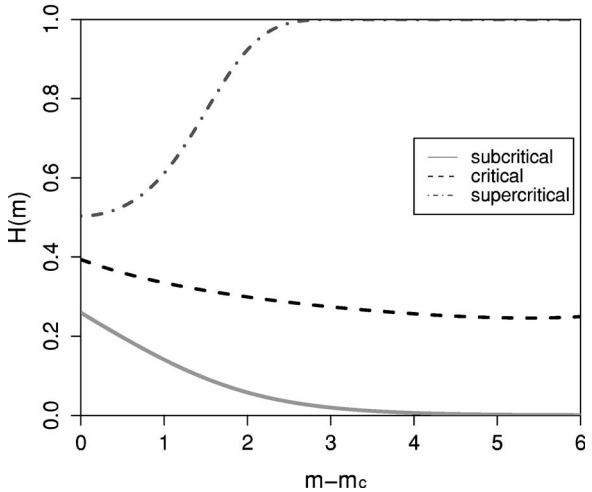

FIG. 3. Foreshock probabilities for background events of different magnitudes, i.e., $H(m)=1-\zeta(M,M)$, with parameters being ($A=0.3$, $\alpha=1.2$, $\beta=2.4$), ($A=0.5$, $\alpha=1.2$, $\beta=2.4$), and ($A=0.7$, $\alpha=1.2$, $\beta=2.4$) for the subcritical, critical, and supercritical, respectively.

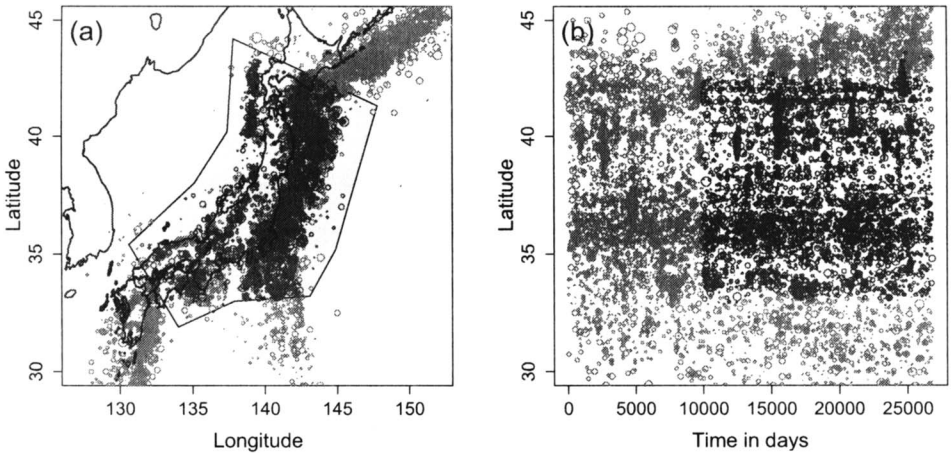

FIG. 4. Seismicity in the Japan region and nearby during 1926–1999 ($M_J \geqslant 4.2$). (a) Epicenter locations; (b) Latitudes of epicenter locations against occurrence times. The shaded region represents the studied space-time range.

$$\zeta(m,m') \approx \exp\left[-\frac{1}{\beta(1-\varrho)}s(m')\kappa(m)\right] = \exp\left[-\frac{A'}{\beta(1-\varrho)}e^{\alpha m-\beta m'}\right], \tag{27}$$

where $A'=Ae^{(\beta-\alpha)m_c}$, i.e., $\zeta(m,m')$ could be approximately regarded as a function of $\alpha m-\beta m'$.

The limit properties for the critical case can be obtained by letting $\varrho \to 1$ in Eq. (25),

$$F(m) \approx e^{-\alpha(m-m_c)} \propto [\kappa(m)]^{-1}. \tag{28}$$

In other words, when m is large enough and the process is critical,

$$\zeta(m,m') \approx \exp[-Ae^{\alpha(m-m')}] \tag{29}$$

is approximately a function of $m-m'$.

When the process is supercritical, $\varrho > 1$ and $C > 0$. Equation (20) yields

$$\lim_{m'\to\infty} \zeta(m,m') = \lim_{m'\to\infty} e^{-\kappa(m)F(m')} = e^{-C\kappa(m)} = P_c(m'), \tag{30}$$

implying that $\zeta(m,m')$ is only a function of m when m' is sufficiently large. That is, when m' is sufficiently large, descendants of any magnitudes may be produced, because the probability that the population of the family tree is infinite is greater than 0 in the supercritical case.

C. Numerical results

By using the iteration given in Eq. (13) or Eq. (14), $F(m)$ can be solved. In Fig. 1, we first fixed two of A, α, and β, and then changed the other to see how $F(m)$ changes. As we have discussed above, the influence of all the three parameters, α, β, and A, to $F(m)$ is through their influence to the criticality, i.e., the properties of $F(m)$ can be divided into three regimes: subcritical, critical, and supercritical, where $F(m) \sim s(m)$, $F(m) \sim 1/\kappa(m)$, and $F(m) \sim C$, respectively.

Once $F(m)$ is evaluated, we can plot the image of $\zeta(m,m')$ according to (12), as shown in Fig. 2. Basically, in the subcritical case, the contour line corresponding to $\zeta(m,m')=\text{const} \approx \exp\{-[A'/\beta(1-\rho)]e^{\alpha m-\beta m'}\}$ approximately parallel to a slope of α/β [see Eq. (27) for justification]. This

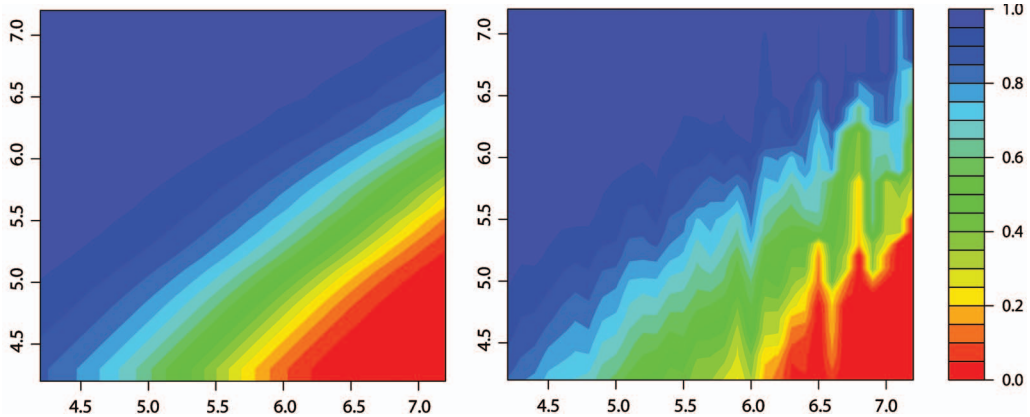

FIG. 5. (Color) Comparison between the theoretical $\xi(m,m')$ and reconstructed $\hat{\xi}(m,m')$ for the JMA catalog. Left panel: $\xi(m,m')$; right panel: $\hat{\xi}(m,m')$.

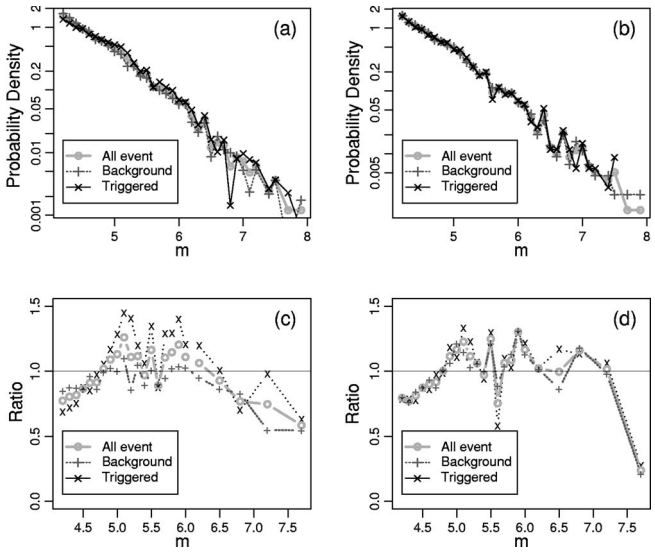

FIG. 6. Reconstruction of the magnitude distributions, $\hat{s}_b(m)$ of the background events, $\hat{s}_c(m)$ of the triggered events, and $s(m)$ of all the events (a) for the JMA catalog and (b) for the simulated catalog (from Zhuang *et al.* [42]). Corresponding ratios of the reconstructed densities to the theoretic density, $\hat{s}_b(m)/s(m)$, $\hat{s}_c(m)/s(m)$, $\hat{s}(m)/s(m)$ for (c) the JMA catalog and (d) the simulated catalog, where $s(m)$ is from the fitted ETAS model.

asymptotic slope becomes 1 when the process is critical [see Eq. (29)]. In the supercritical case, the contour lines are asymptotically parallel to the m' axis. Or explicitly, in the supercritical case, for a fixed m, when $m' \to \infty$ (20) and (30) give $F(m') \to C$ and $\zeta(m,m') \to \exp[-C\kappa(m)] = P_c(m)$, independent from m'. These results are consistent with our analytic discussions.

IV. DEFINITION AND MAGNITUDE DISTRIBUTIONS OF FORESHOCKS

As noted in the Introduction, in conventional studies foreshocks are defined as nonaftershock earthquakes that are followed by one or more larger earthquakes occurring nearby. Such a definition is only applicable to a de-aftershocked catalog. De-aftershocking is conventionally carried out by using a window-based or link-based declustering method (see Refs. [47,51,52], for window-based declustering method; and Refs. [53–55] for link-based declustering methods; or [56] for a review). However, the window size or the link distance in the conventional definitions of earthquake clusters are difficult to determine. Similarly, parameters of the maximum time lag and the maximum distance between the foreshock and the mainshock are hard to decide. To avoid the above difficulties, we define a foreshock by a background event that has at least one offspring, direct or indirect, with a larger magnitude. Here, we refer to [41,42] or Appendix B on how to determine whether an earthquake is a background event, a triggered event, or a foreshock by using the stochastic declustering method.

Recalling (11), the probability that an event of magnitude m is a foreshock is the same probability that an event of magnitude m produces at least one descendant greater than m, i.e.,

$$H(m) = 1 - \zeta(m,m). \tag{31}$$

Thus the probability density functions of the magnitude distributions of foreshocks and nonforeshock events are

$$s_f(m) = \frac{s(m)H(m)}{\int_{m_c}^{+\infty} s(m)H(m)\mathrm{d}m} \propto s(m)[1 - \zeta(m,m)], \tag{32}$$

and

$$s_n(m) = \frac{s(m)[1 - H(m)]}{\int_{m_c}^{+\infty} s(m)[1 - H(m)]\mathrm{d}m} \propto s(m)\zeta(m,m), \tag{33}$$

respectively. In other words, both magnitude distributions departure from the exponential distribution or the Gutenburg-Richter (GR) law for all the events.

$H(m)$ is shown in Fig. 3. When the process is supercritical, from Eqs. (30) and (31),

$$\lim_{m\to+\infty} H(m) = 1 - \lim_{m\to+\infty} \zeta(m,m) = 1 - \lim_{m\to+\infty} \mathrm{e}^{-C\kappa(m)} = 1, \tag{34}$$

indicating that a background event is always a foreshock if it is sufficiently large. When the process is critical, by Eq. (29),

$$\lim_{m\to+\infty} H(m) = 1 - \lim_{m\to+\infty} \zeta(m,m) = 1 - \mathrm{e}^{-A}, \tag{35}$$

indicating that, if the magnitude is sufficiently large, it has a constant probability between 0 and 1 to be a foreshock. For the subcritical case, by Eq. (27),

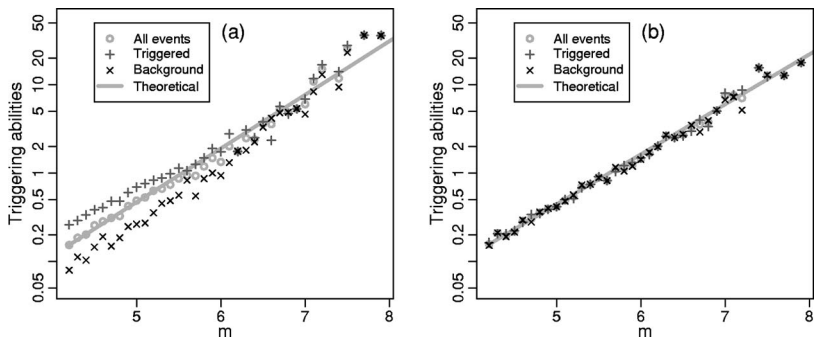

FIG. 7. Reconstruction of the triggering abilities $\kappa_b(m)$ of the background events and $\kappa_c(m)$ for the triggered events in (a) the JMA catalog and in (b) a simulated catalog (from Zhuang *et al.* [42]). For comparison, the empirical functions of the triggering abilities for all the events are plotted as gray circles, and the corresponding theoretical functions, $\kappa(m) = A\mathrm{e}^{\alpha m}$, are represented by the straight lines.

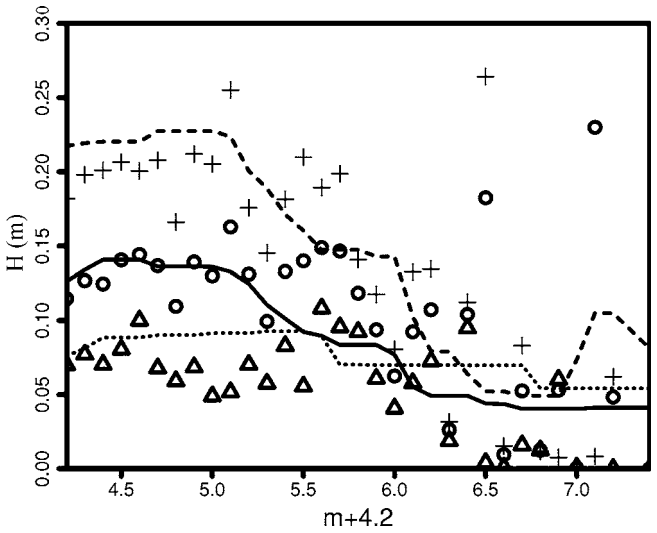

FIG. 8. Probabilities that events of different magnitudes have at least one large descendant in the JMA catalog. Circles, triangles, and crosses represent the reconstructed $\hat{H}(m)$, $\hat{H}_c(m)$, and $\hat{H}_b(m)$ for all the events, background events and triggered events, respectively. The solid, dotted, and dashed curves are theoretical results for all the events, background events, and triggered events, respectively, computed by using the empirical $\hat{\kappa}(m)$, $\hat{\kappa}_b(m)$, $\hat{\kappa}_c(m)$, $\hat{s}(m)$, $\hat{s}_b(m)$, and $\hat{s}_c(m)$.

$$\lim_{m\to+\infty} H(m) = 1 - \lim_{m\to+\infty} \zeta(m,m) = 1 - \lim_{m\to+\infty} \exp\left[-\frac{A e^{(\alpha-\beta)m}}{\beta(1-\rho)}\right] = 0,$$

indicating that the probability that a background event is a foreshock decreases to 0 when its magnitude increases to a sufficiently large value.

In practice, given an earthquake catalog, we can first fit the model to the catalog to obtain the parameters in the branching structure, then calculate $\zeta(m,m')$, $H(m)$, and $s_f(m)$ according to the formulas derived in Secs. III and IV, to find out the proportion of foreshocks and nonforeshock events in the catalog and their magnitude distributions.

V. DATA ANALYSIS AND RESULTS

The dataset used in this study is the Japan Meteorological Agency (JMA) catalog in a range of longitude 121°–155°E, latitude 21°–48°N, depth 0–100 km, time 1926/January/1–1999/December/31 and magnitude $\geq M_J 4.2$ (Fig. 4). For an earthquake catalog covering records of a long history, completeness and homogeneity are always problems causing trouble for statistical analysis. To tackle these problems, we choose a target space-time range, in which the seismicity seems to be relatively complete and homogeneous. The incompleteness of the early period and inhomogeneity in the JMA dataset can be easily seen from Fig. 4(b). We choose a polygon as shown in Fig. 4(a) as the target region and a time period of 10 000–26 814 days after 1926/Janary/1, with the same depth and magnitude ranges as the whole dataset. The events outside of this study space-time range are used as complementary events for calculating the boundary effect.

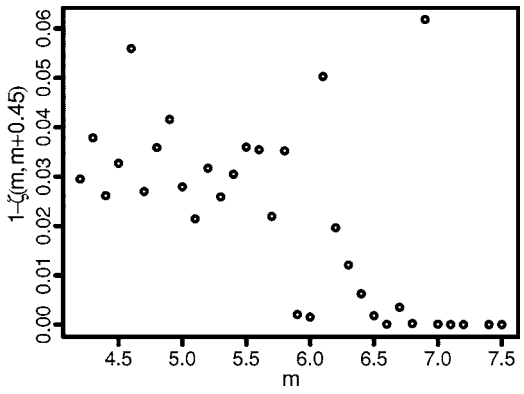

FIG. 9. Estimated probabilities that events of different magnitudes have at least one descendant of a magnitude 0.45 larger, i.e., $1-\hat{\zeta}(m,m+0.45)$, in the JMA catalog.

Using the technique described by Zhuang *et al.* [41,42], the model parameters are estimated by maximizing the likelihood function: $\hat{A}=0.1692$, $\hat{c}=0.0189$ day, $\hat{\alpha}=1.5014$, $\hat{p}=1.1181$, $\hat{D}^2=0.000\,864\,0$ deg^2, $\hat{q}=1.9107$, $\hat{\gamma}=1.0761$, and $\hat{\beta}=1.9585$.

There are two ways to evaluate $\zeta(m,m')$: one is based on solving Eq. (12) with the parameters from the fitting results; the other is based on the reconstruction technique given in Appendix B. We used the terms of theoretical results and reconstruction results to represent the results obtained from these two methods, respectively. In order to avoid the edge effect caused by missing observations of offspring of the events near the end of the time interval, we restrict the events within the time between 10 000–20 000 for reconstruction. The theoretical $\zeta(m,m')$ and the reconstructed $\hat{\zeta}(m,m')$ for the JMA are shown in Fig. 5. Roughly speaking, the theoretical $\zeta(m,m')$ is close to the model, while $\hat{\zeta}(m,m')$ is close to the data. In calculating the theoretical one in the left panel, we use the empirical magnitude distribution for the events falling in the study space-time range as $s(m)$. The overall impression of these two images is their similarity to each other, even though the theoretical one is more smooth than the reconstructed one. The contour lines in the left panel are not as straight because the magnitude distribution of the real catalog is slightly different from the GR law, which has an exponential distribution, while the real catalog has a pdf of magnitudes with a tail that decays more quickly than the pdf of the exponential distribution.

Plotting $H(m)$ and $\hat{H}(m)$ in Fig. 8, we can see that, even if there is no big discrepancy between these two functions, the probability for an event to produce at least one larger descendant is about 15%, which seems much higher than what has been estimated in Ref. [10].

According to Zhuang *et al.* [42], such high probabilities may be explained by the different characteristics of background events and their descendants, outlined as follows.

(1) The magnitude distributions of background events and triggered events are different. If we denote the magnitude pdf's by $s_b(m)$ and $s_c(m)$ for the background events and

triggered events, respectively, $s_b(m)$ has a higher Gutenburg-Richter b value (or $\beta=b\ln 10$ in Eq. (4), as shown in Fig. 6). For the JMA catalog, $\beta=2.1$ for background events and $\beta=1.8$ for triggered events. This seems contrary to our knowledge that mainshocks usually have a smaller b value than aftershocks. The background events are initial events of each cluster, but not the largest events (mainshocks), and thus have a lower mean magnitude than mainshocks and a higher b value. This low b value indicates that clusters tend to be initiated by small events.

(2) A triggered event triggers more direct children on average than a background event of the same magnitude (Fig. 7). This can be explained by the fact that the number of children events triggered by a parent event depends not only on the parent's magnitude, but also on the heterogeneity of the stress field. The more heterogeneous the stress field, the more children each parent triggers. At the beginning of an earthquake cluster, the stress field is at a relatively homogenous state. The stress field becomes more heterogeneous as the cluster evolves until it is adjusted by the occurrence of most events in the cluster. The stress field recovers to a state being relatively homogeneous at the final stage of the cluster. Here, we denote the productivity from a background event and a triggered event by $\kappa_b(m)$ and $\kappa_c(m)$, respectively.

To evaluate the probability that a triggered event of magnitude m has no descendant larger than m', we can just replace s and κ in Eq. (11) with s_c and κ_c,

$$\zeta_c(m,m') = \exp\left\{-\kappa_c(m)\left[1-\int_{m_c}^{m'} s_c(m^*)\zeta_c(m^*,m')\mathrm{d}m^*\right]\right\}. \tag{37}$$

For a background event of magnitude m, the corresponding probability becomes

$$\begin{aligned}\zeta_b(m,m') &= \Pr\{\text{a background event of magnitude } m \text{ has no offspring greater than } m'\}\\ &= \sum_{n=0}^{\infty} \Pr\{\text{each child is } \leqslant m' \text{ and has no offspring } > m' | m \text{ has } n \text{ children}\} \times \Pr\{m \text{ has } n \text{ direct offspring}\}\\ &= \sum_{n=0}^{\infty}\left[\int_{m_c}^{m'} s_c(m^*)\zeta_c(m^*,m')\mathrm{d}m^*\right]^n \frac{[\kappa_b(m)]^n}{n!}\mathrm{e}^{-\kappa_b(m)} = \exp\left\{-\kappa_b(m)\left[1-\int_{m_c}^{m'} s_c(m^*)\zeta_c(m^*,m')\mathrm{d}m^*\right]\right\}.\end{aligned} \tag{38}$$

Using the stochastic reconstruction techniques introduced in [42], we reconstruct κ_b, κ_c, and s_c, and then use them to calculate H_b and H_c. However, H_b and H_c can also be reconstructed directly (see Appendix B). As we can see in Fig. 8, theoretical curves of H_b and H_c are still close to the reconstructed ones, $\hat{H}_b$ and $\hat{H}_c$, respectively. Basically, the probability that a background earthquake of magnitude about 4.2 to 5 triggers a larger earthquake is around 8%.

It is worthwhile to mention the differences between our results and those obtained by Ogata *et al.* [10]. Ogata *et al.* [10] used the "magnitude-based clustering" (MBC) and the "single-link clustering" (SLC) methods to separate the JMA catalog into different clusters. They defined the largest event in a cluster to be mainshock. Their foreshocks were defined as shocks before the mainshock and 0.45 smaller than the mainshock in magnitude. They obtained the probability that a cluster has one or more foreshocks was 7.2% if the MBC method was used for declustering, and 3.7% if the SLC method was used. Their results implied that a good declustering method is essential to study foreshocks. From the definition of ζ, the probabilities that a background earthquake is a "foreshock defined by Ogata *et al.*" corresponds to $1-\zeta(m,m+0.45)$. We plot $1-\hat{\zeta}(m,m+0.45)$ for the analyzed JMA data in Fig. 9. Even though the values of $1-\hat{\zeta}(m,m+4.5)$ vary from 2% to 6% for lower magnitudes, their mean value is around 3–4%. Thus our results using the stochastic reconstruction methods are closer to the probability estimated by using the SLC method than by using the MBC method.

VI. DISCUSSION

We have used two methods to calculate the probability that an earthquake event of a given magnitude produces at least one descendent of another given magnitude: One is based on the explicit formulation of the clustering model, and the other is based on stochastic reconstruction with the probabilities that one event is the descendant of another previous event. For the second method, we need a large enough number of events to make the estimation. The advantage of this method is that it gives more information of the data than what can be obtained purely from the model.

Zhuang *et al.* [42] give a list of discrepancies between the ETAS model and the real seismicity and note that the ETAS model, even though it is proved to be the best model in describing earthquake clusters in practice at the current stage, is only just a "first-order" approximation of seismicity. Seismicity is much more complicated than in the ETAS model. The background seismicity and the triggering behaviors of earthquakes may change from time to time and place to place, because of the change of the complicated stress fields. An improved model, if these discrepancies and the complexity of seismicity are considered, definitely will benefit from our understanding of foreshocks and earthquake prediction.

As mentioned in the Introduction, the problem of whether foreshocks are mainshocks whose aftershocks happen to be bigger is of interest. From our analysis, the classification of background events and triggered events is more essential than the classification of foreshocks, mainshocks, and aftershocks. Foreshocks and mainshocks are not easily comparable in triggering other events, because a foreshock is always a background event while a mainshock may be a background event or an event triggered by a previous event. The events among the triggered events corresponding to foreshocks in the background events are the triggered events which have one or more larger descendants. They have higher ability in producing offspring, larger or not, than foreshocks, which is due to the main differences between triggered events and background events. Thus, the relation between foreshocks and mainshocks is that the foreshock and the background mainshock are the complementary parts in the background seismicity.

Another important problem is how to make use of the above results for the purpose of earthquake prediction. The results of this paper and other related ones prove these facts: (1) the occurrence of big earthquakes following the so-called "foreshocks" is not due to coincidence; (2) given an event, the probability that it is a foreshock is overestimated if it is estimated by using the ETAS model or the Omori-Utsu formula with the Gentenburg-Richter law for the magnitude; (3) prediction based on foreshocks can be implemented based on a clustering model more complicated than the ETAS model.

VII. CONCLUSIONS

In this paper, based on the space-time ETAS model, we obtained the key equation for the probability for the magnitude of the largest descendant from a given ancestor, as $\zeta(m,m')$ in Eq. (11). When $m'\to\infty$, $\zeta(m,m')\sim\exp[-\alpha e^{\alpha m-\beta m'}/\beta]$, $\zeta\sim\exp[-\alpha e^{m-m'}]$, and $\zeta(m,m')\sim P_c(m)$ for the cases where the process are subcritical, critical, and supercritical, respectively. We define foreshocks naturally by background events that have at least one larger descendant. The probability that a background event is a foreshock can obtained from Eq. (11).

We also analyze the JMA catalog to verify our theories. To obtain $\zeta(m,m')$, two methods are used: one is directly computed from the integral equations (12) and (13), and the other by using stochastic reconstruction [see (B3) in Appendix B]. Because of the different behaviors in triggering offspring between background events and triggered events, we evaluate the corresponding ζ function for both background events and triggered events. For a background event of a magnitude from 4.2 to 6.0, the probability that it is a foreshock is about 8%. Our outputs are very close to the results obtained by using a conventional foreshock definition and a conventional "single-linked cluster" (SLC) declustering method.

ACKNOWLEDGMENTS

This study is supported financially by Grant-in-Aid 14380128 and 17200021 for Scientific Research (B2) and (A), respectively, from Ministry of Education, Science, Sports and Culture of Japan and also by the 2004 and 2005 ISM project. This first author is also supported by a postdoctoral program P04039 from the Japan Society for Promotion of Science. Discussions with Annemarie Christophsen, Kazuyoshi Nanjo, Martha Savage, Euan Smith, Didier Sornetter, and David Vere-Jones were helpful and gratefully acknowledged.

APPENDIX A: CRITICALITY OF THE ETAS MODEL

In the process of the ETAS model, the background process is Poisson with an intensity function $\mu(t,x,y,m)$; once an event occurs, it produces an offspring process with an intensity as the aforementioned ξ. If we denote the background events as the generation G_0, with the intensity function $V_0(\mu,x,y,m)=\mu(t,x,y,m)$, then the intensity function for the first generation G_1 is

$$V_1(t,x,y,m)=\int\int\int\int_S \xi(t,x,y,m;t',x',y',m') \times V_0(t',x',y',m')\mathrm{d}t'\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}m', \quad \text{(A1)}$$

where S represents the whole time-space-magnitude range. For simplification, X represents the vector (t,x,y,m), X' is (t',x',y',m') and (A1) can be simplified as

$$V_1(X)=\int_S \xi(X;X')V_0(X')\mathrm{d}X'. \quad \text{(A2)}$$

Similarly, the intensity function for the nth generation is

$$V_n(X)=\int_S \xi(X;X')V_{n-1}(X')\mathrm{d}X'. \quad \text{(A3)}$$

Thus

$$V_n(X)=\int_S \xi^{[n]}(X;X')V_0(X')\mathrm{d}X', \quad \text{(A4)}$$

where

$$\xi^{[1]}(X;X')=\xi(X;X'), \quad \text{(A5)}$$

and

$$\xi^{[k]}(X;X')=\int_S \xi^{[k-1]}(X;X^*)\xi(X^*;X')\mathrm{d}X^*, \quad \text{(A6)}$$

where X^* is an abbreviation of (t^*,x^*,y^*,m^*). Suppose that $a(X')$ and $b(X)$ are the left and right eigenfunctions of ξ corresponding to the maximum eigenvalue ϱ, i.e.,

$$\varrho a(X')=\int_S a(X)\xi(X;X')\mathrm{d}X, \quad \text{(A7)}$$

and

$$\varrho b(X)=\int_S h(X;X')b(X')\mathrm{d}X', \quad \text{(A8)}$$

respectively, satisfying that

$$\int_S a(X)b(X)\mathrm{d}X = 1. \tag{A9}$$

Let

$$\Lambda(X;X') = a(X')b(X), \tag{A10}$$

i.e., Λ gives the projection operator corresponding to ϱ, or

$$\int_S \Lambda(X;X^*)\xi(X^*;X')\mathrm{d}X^* = \int_S h(X;X^*)\Lambda(X^*;X')\mathrm{d}X^* = \varrho\Lambda(X;X'). \tag{A11}$$

Reconsider Eq. (A4), when $n\to\infty$,

$$\frac{\xi^{[n]}}{\varrho^n} \to \Lambda, \tag{A12}$$

and thus

$$V_n(X) \to \varrho^n \int_S \Lambda(X;X')V_0(X')\mathrm{d}X'. \tag{A13}$$

From Eq. (A13), we can see that if $\varrho<1$, then $V_n\to 0$ when $n\to\infty$, and that if $\varrho>1$, then $V_n\to\infty$ when $n\to\infty$. Here, ϱ is called the *critical parameter* because if $\varrho<1$, the process is stable and otherwise the population is explosive to infinity.

Equation (A13) can be rewritten as

$$V_n(X) \to \varrho^n \int_S b(X)a(X')V_0(X')\mathrm{d}X' = \varrho^n b(X)\int_S a(X')V_0(X')\mathrm{d}X' = \varrho^n b(X)\times \text{const}, \tag{A14}$$

implying that $b(X)$ is the asymptotic intensity of the population in the nth generation, when $n\to\infty$.

The eigenfunction $a(X')$ can be interpreted as the asymptotic ability in producing offspring, directly and indirectly, from an ancestor $\{X'\}$ because

$$\lim_{n\to\infty}\sum_{i=n}^{\infty}\int_S \xi^{[k]}(X;X')\mathrm{d}X = \lim_{n\to\infty}\sum_{i=n}^{\infty}\varrho^i\int_S \Lambda(X;X')\mathrm{d}X = \lim_{n\to\infty}\sum_{i=n}^{\infty}\varrho^i a(X')\int_S b(X)\mathrm{d}X = \lim_{n\to\infty}\frac{\varrho^n}{1-\varrho}a(X')\times\text{const}. \tag{A15}$$

For an ETAS model with a magnitude-separable and temporal constant background rate and a fully separable uniform clustering response, i.e.,

$$\lambda(t,x,y,m) = \lambda(t,x,y)s(m) \tag{A16}$$

$$\lambda(t,x,y) = \mu(x,y) + \sum_{i:t_i<t}\kappa(m_i)g(t-t_i)f(x-x_i,y-y_i;m_i), \tag{A17}$$

the eigenvalue equations are

$$\varrho a(t',x',y',m') = \int\int\int\int_S a(t,x,y,m)\kappa(m')g(t-t') \times f(x-x',y-y';m')s(m)\mathrm{d}t\mathrm{d}x\mathrm{d}y\mathrm{d}m,$$

$$\varrho b(t,x,y,m) = \int\int\int\int_S \kappa(m')g(t-t')f(x-x',y-y';m') \times s(m)b(t',x',y',m')\mathrm{d}t'\,\mathrm{d}x'\,\mathrm{d}y'\,\mathrm{d}m'.$$

Only consider the functions of m' and m for the forms of a and b. The above eigenequations can be simplified as

$$\varrho a(m') = \kappa(m')\int_{\mathcal{M}} a(m)s(m)\mathrm{d}m, \tag{A18}$$

$$\varrho b(m) = s(m)\int_{\mathcal{M}} \kappa(m')b(m')\mathrm{d}m', \tag{A19}$$

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

$$a(m') = C_1\kappa(m'), \tag{A20}$$

and

$$b(m) = C_2 s(m). \tag{A21}$$

Substituting $a(m')$ and $b(m)$ back into Eqs. (A18) and (A19), we can get that

$$\varrho = \int_{\mathcal{M}} \kappa(m)s(m)\mathrm{d}m. \tag{A22}$$

For the model given in Sec. II, the criticality parameter

$$\varrho = \int_{m_c}^{\infty} s(m)\kappa(m)\mathrm{d}m = \frac{A\beta}{\beta-\alpha}.$$

APPENDIX B: $\zeta(m_1,m_2)$ BY RECONSTRUCTION

Sections III and IV give the theory and methods for calculating the probability that an event produces at least one larger descendant when the process of seismicity is controlled by the ETAS model. However, the true model is always unknown for seismicity. What we know is the observation, i.e., a catalog of earthquakes. The ETAS model is just a rough approximation of the seismicity. For example, Zhuang *et al.* [42] show that the triggering behaviors are different between background event and triggered events. To get more information on $\zeta(m,m')$ from the data than from the model, we need to introduce a special technique, which is called the stochastic reconstruction method, developed by Zhuang *et al.* ([41,42]), following up the initial idea of Kagan and Knopoff [7], to accomplish this task.

Once the conditional intensity function is estimated, it provides us a good estimate of the probability of how likely an event is a background event or triggered by others. The contribution of the background seismicity rate at the occurrence time and location of the ith event is

$$\varphi_i = \frac{\mu(x_i, y_i)}{\lambda(t_i, x_i, y_i)}. \tag{B1}$$

If we remove the ith event with probability $1-\varphi_i$ for all the events in the process, we can realize a process with the occurrence rate of $\mu(x,y)$ (see [57,58], for justification). Thus, it is natural to regard φ_i as the probability that the ith event is a background events. Similarly,

$$\rho_{ij} = \begin{cases} \dfrac{\kappa(m_i) g(t_j - t_i) f(x_j - x_i, y_j - y_i; m_i)}{\lambda(t_j, x_j, y_j)}, & \text{if } j > i, \\ 0, & \text{otherwise}, \end{cases} \tag{B2}$$

can be regarded as the probability that the jth event is directly produced by i. It is easy to verify that $\varphi_j + \Sigma_i \rho_{ij} = 1$.

The whole catalog can be separated into different clusters the following way: For each event j, generate a random variable u_j uniformly distributed on $[0, 1]$, then select the jth event as a background event if $U_j < \varphi_j$; otherwise, select the I_jth event as the parent of event j, where $I_j = \min\{k : \varphi_j + \Sigma_{i=1}^{k} \rho_{ij} \leqslant U_j, 1 \geqslant k < j\}$.

Such a separation can be repeated many times get many different versions of cluster classifications. Suppose that the total number of versions is $\tilde{N}$ and that, for each event i, $M_{i\ell}$ is the magnitude of the largest descendant from event i; then, $\zeta(m, m')$ can be estimated through

$$\hat{\zeta}(m, m') = \frac{\sum_{i,\ell} I(M_{i\ell} < m') I(m_i \approx m)}{\tilde{N} \sum_i I(m_i \approx m)}. \tag{B3}$$

Similarly, to estimate $\zeta(m, m')$ for a background event,

$$\hat{\zeta}_b(m, m') = \frac{\sum_{i,\ell} I(M_{i\ell} < m') \varphi_i I(m_i \approx m)}{\tilde{N} \sum_i I(m_i \approx m)}, \tag{B4}$$

and for a triggered event

$$\hat{\zeta}_c(m, m') = \frac{\sum_{i,\ell} I(M_{i\ell} < m')(1 - \varphi_i) I(m_i \approx m)}{\tilde{N} \sum_i I(m_i \approx m)}. \tag{B5}$$

More details on the uses of stochastic reconstruction techniques can be found in [42].

[1] I. L. Cifuentes, J. Geophys. Res. **94**, 665 (1989).
[2] W. J. Arabasz, W. D. Richins, and C. J. Langer, Bull. Seismol. Soc. Am. **71**, 803 (1981).
[3] T. Utsu, Zisin II (J. Seismol. Soc. Japan) **30**, 179 (1977) (in Japanese with English abstract).
[4] T. Utsu, Zisin II (J. Seismol. Soc. Japan) **31**, 129 (1978) (in Japanese with English abstract).
[5] J. Boatwright, Bull. Seismol. Soc. Am. **72**, 1049 (1982).
[6] R. E. Abercrombie and J. Mori, Bull. Seismol. Soc. Am. **84**, 725 (1994).
[7] Y. Y. Kagan and L. Knopoff, Geophys. J. R. Astron. Soc. **55**, 67 (1978).
[8] K. Maeda, Bull. Seismol. Soc. Am. **86**, 242–254 (1996).
[9] Y. Ogata, T. Utsu, and K. Katsura, Geophys. J. Int. **121**, 233 (1995).
[10] Y. Ogata, T. Utsu, and K. Katsura, Geophys. J. Int. **127**, 17 (1996).
[11] L. M. Jones, Bull. Seismol. Soc. Am. **74**, 1361 (1984).
[12] L. M. Jones, Bull. Seismol. Soc. Am. **75**, 1669 (1985).
[13] Z.-R. Liu, Bull. Seismol. Soc. Am. **74**, 255 (1984).
[14] L. Jones and P. Molnar, Nature (London) **262**, 677 (1976).
[15] L. M. Jones and P. Molnar, J. Geophys. Res. **84**, 3596 (1979).
[16] B. C. Papazachos, Pure Appl. Geophys. **112**, 627 (1974).
[17] E. Smith, N.Z. J. Geol. Geophys. **24**, 579 (1981).
[18] M. K. Savage and S. Rupp, N.Z. J. Geol. Geophys. **43**, 461 (2000).
[19] F. F. Evison and D. A. Rhoades, N.Z. J. Geol. Geophys. **36**, 51 (1993).
[20] F. F. Evison and D. A. Rhoades, N.Z. J. Geol. Geophys. **40**, 537 (1997).
[21] A. Helmstetter, S. Hergarten, and D. Sornette, Phys. Rev. E **70**, 046120 (2004).
[22] A. Merrifield, M. K. Savage, and D. Vere-Jones, N.Z. J. Geol. Geophys. **47**, 327 (2004).
[23] P. Reasenberg and L. M. Jones, Science **243**, 1173 (1994).
[24] P. Reasenberg and L. M. Jones, Science **265**, 1251 (1994).
[25] P. Reasenberg, J. Geophys. Res. **104**, 4755 (1999).
[26] A. Helmstetter, D. Sornette, and J.-R. Grasso, J. Geophys. Res. **108**, 2046 (2003).
[27] A. Helmstetter and D. Sornette, Geophys. Res. Lett. **30**, 2069 (2003).
[28] A. Helmstetter and D. Sornette, J. Geophys. Res. **108**, 2457 (2003).
[29] K. R. Felzer, R. E. Abercrombie, and G. Ekström, Bull. Seismol. Soc. Am. **94**, 88 (2004).
[30] R. Console, M. Murru, and B. Alessandrini, Bull. Seismol. Soc. Am. **83**, 1248 (1993).
[31] M. K. Savage and D. M. DePolo, Bull. Seismol. Soc. Am. **83**, 1910 (1992).
[32] R. Console and M. Murru, J. Geophys. Res. **106**, 8699 (2001).
[33] R. Console, M. Murru, and A. M. Lombardi, J. Geophys. Res. **108**, 2468 (2003).
[34] Y. Kagan, J. Geophys. Res. **106**, 135 (1991).
[35] F. Musmeci and D. Vere-Jones, Ann. Inst. Stat. Math. **44**, 1

(1992).
[36] Y. Ogata, Proc. Edinb. Math. Soc. **83**, 9 (1988).
[37] Y. Ogata, J. Geophys. Res. **97**, 19845 (1992).
[38] Y. Ogata, Ann. Inst. Stat. Math. **50**, 379 (1998).
[39] Y. Ogata, J. Geophys. Res. **109**, B03308 (2004).
[40] S. L. Rathbun, Bull. Internat. Statist. Inst. **55**, Book 2, 379 (1993).
[41] J. Zhuang, Y. Ogata, and D. Vere-Jones, J. Am. Stat. Assoc. **97**, 369 (2002).
[42] J. Zhuang, Y. Ogata, and D. Vere-Jones, J. Geophys. Res. **109**, B05301 (2004).
[43] J. Zhuang, C.-P. Chang, Y. Ogata, and Y.-I. Chen, J. Geophys. Res. **110**, B05S18 (2005).
[44] Y. Ogata and J. Zhuang, Tectonophysics **413**, 13 (2006).
[45] D. Daley and D. Vere-Jones, *An Introduction to Theory of Point Processes* (Springer, New York, 2003).
[46] F. Omori, J. Coll. Sci., Imp. Univ. Tokyo **7**, 111 (1894).
[47] T. Utsu, J. Fac. Sci., Hokkaido Univ., Ser. 7 **3**, 129 (1969).
[48] T. Utsu, Y. Ogata, and R. S. Matsu'ura, J. Phys. Earth **43**, 1 (1995).
[49] J. Zhuang and Y. Ogata, Eos Trans. AGU, 85(47), Fall Meet. Suppl., Abstract S23A-0300 (2004).
[50] A. Saichev and D. Sornette, Phys. Rev. E **71**, 056127 (2005).
[51] J. Gardner and L. Knopoff, Bull. Seismol. Soc. Am. **64**, 1363 (1974).
[52] V. I. Kellis-Borok, and V. I. Kossobokov, Comput. Seismol. **19**, 45 (1986).
[53] S. D. Davis and C. Frohlich, Geophys. J. Int. **104**, 289 (1991).
[54] C. Frohlich and S. D. Davis, Geophys. J. Int. **100**, 19 (1991).
[55] P. Resenberg, J. Geophys. Res. **90**, 5479 (1985).
[56] G. M. Molchan and O. E. Dmirtrieva, Geophys. J. Int. **190**, 501 (1992).
[57] Y. Ogata, IEEE Trans. Inf. Theory **IT-27**, 23 (1981).
[58] J. Zhuang, J. R. Stat. Soc. Ser. B. Methodol. (to be published).