Solution of the nonlinear theory and tests of earthquake recurrence times

D. Sornette,[1,2,*] S. Utkin,[3,†] and A. Saichev[1,3]
[1]*Department of Management, Technology and Economics, ETH Zurich, Kreuzplatz 5, CH-8032 Zurich, Switzerland*
[2]*Institute of Geophysics and Planetary Physics and Department of Earth and Space Sciences, University of California, Los Angeles, California 90095, USA*
[3]*Mathematical Department, Nizhny Novgorod State University, Gagarin Prospekt 23, Nizhny Novgorod, 603950, Russia*
(Received 15 December 2007; published 11 June 2008)

We develop an efficient numerical scheme to solve accurately the set of nonlinear integral equations derived previously in [A. Saichev and D. Sornette, J. Geophys. Res. **112**, B04313 (2007)], which describes the distribution of interevent times in the framework of a general model of earthquake clustering with long memory. Detailed comparisons between the linear and nonlinear versions of the theory and direct synthetic catalogs show that the nonlinear theory provides an excellent fit to the synthetic catalogs, while there are significant biases resulting from the use of the linear approximation. We then address the suggestions proposed by some authors to use the empirical distribution of interevent times to obtain a better determination of the so-called clustering parameter. Our theory and tests against synthetic and empirical catalogs find a rather dramatic lack of power for the distribution of interevent times to distinguish between quite different sets of parameters, casting doubt on the usefulness of this statistic for the specific purpose of identifying the clustering parameter.

DOI: 10.1103/PhysRevE.77.066109 PACS number(s): 89.75.Da

I. INTRODUCTION

Most complex systems of interest in the natural and social sciences exhibit intermittent bursts of activity interspersed within long times of reduced activity. A simple metric to characterize this property consists of the distribution of recurrence (also called "waiting" or "interevent") times between (suitably defined) events. Recently, the literature has undergone itself a burst of publication activity on this topic, motivated by the idea that distributions of recurrence times may be one of the most important complexity measures for both random fields and nonlinear dynamical systems [1]. The applications include recurrence time and anomalous transport [2], waiting times between earthquakes [3–7] and rock fractures [8], time intervals between consecutive e-mails [9], and between web browsing, library visits, and stock trading [10].

Much of the recent interest of the statistical physics community focused on applying scaling techniques, which are common tools in the study of critical phenomena, to the statistics of interearthquake recurrence times or waiting times [3,4,11–18]. Many of the claims made in these recent papers on recurrence statistics have either been challenged, refuted, or explained by previously known facts about earthquake statistics [6,7,19–22]. In particular, two of us [6,7] have developed a general theory of the statistics of interevent times in the framework of the general class of self-excited Hawkes conditional Poisson processes [23–25] adapted to modeling seismicity. The corresponding model is known as the epidemic-type aftershock sequence (ETAS) model, in which any earthquake may trigger other earthquakes, which in turn may trigger more, and so on. Introduced in slightly different forms by Kagan and Knopoff [26] and Ogata [27], the model describes statistically the spatiotemporal clustering of seismicity. Using three well-known statistical laws of statistical seismicity (the Gutenberg-Richter law, the Omori law, and the productivity law), the empirical observations on the distribution of earthquake recurrence times can be explained within this model without invoking additional mechanisms other than the well-known fact that earthquakes can trigger other earthquakes [6,7].

A recent development is the proposition that interevent time distributions may provide a new and more reliable way to measure of the so-called background earthquake activity [19,28]. This question arises as follows: if earthquakes trigger other earthquakes, how much of the observed seismicity is due to past seismicity (endogenous origin) and how much is resulting from an "external" driving source (exogenous origin) often referred to as "background" seismicity thought to reflect the driving tectonic forces at large scales. This question obviously generalizes to any system in which future events may be in part triggered by past events, such as in commercial sales [29] and web browsing activity [10,30]. Within the ETAS framework, the fraction of events in a given catalog which have been triggered by previous events can be shown [31] to be nothing but the so-called branching ratio n, defined mathematically as the average number of first-generation events triggered by a given preceding event [32]. Reciprocally, the fraction of background events is equal to $1-n$ (note that these models assume that the triggering branchinglike processes are subcritical: $n<1$). The degree to which the parameter n can be retrieved from the distribution of interevent times relies on a departure from universality pointed out by Hainzl *et al.* [19] and two of us [6,7]. In this respect, the ETAS model provides an excellent training ground. Using synthetic catalogs generated with the ETAS model, Hainzl *et al.* [19] found that the estimation of n using the distribution of interevent times is better than from the application of a standard declustering procedure [33].

*dsornette@ethz.ch
†sergei_utkin@mail.ru

More progress can be achieved by a better understanding of the sensitivity of the distribution of interevent times to the branching ratio n. In principle, the theoretical framework based on the technique of probability generating functions developed in Refs. [6,7] provides an ideal approach to this problem. However, this previous effort was limited on two accounts. First, while Saichev and Sornette derived the full exact nonlinear integral equations of the problem, they ended solving their linearized versions in order to derive the distribution of interevent times. The present paper keeps the full nonlinear integral equations and shows that the linear simplification leads to systematic biases in the estimations of the key parameters of the ETAS model, and in particular, of the branching ratio n, which has been the focus of recent interest in the seismological community [19,28]. Secondly, only preliminary sensitivity analysis was performed with respect to n. The present paper presents a detailed treatment of the full exact nonlinear equations providing the distribution of interevent times for the ETAS model and discusses how well n can be constrained.

The organization of the presentation is as follows. Section II describes the theoretical framework developed by Saichev and Sornette [6,7] and summarizes their main results, essentially based on a linear approximation to the full nonlinear equations that they derived. Section III focuses on these nonlinear equations and presents the numerical scheme that has been used to solve them. Detailed comparisons between the linear and nonlinear versions of the theory and direct ETAS simulated catalogs are presented. With improved adaptive mesh grids, it is shown that the nonlinear theory provides an excellent fit to the synthetic ETAS catalogs, while there are significant biases resulting from the use of the linear approximation. Section III concludes by a synthetic test demonstrating the possibility to use the nonlinear theory to invert for two of the unknown parameters, if constraints exist on the other three parameters of the model. Section IV applies these results to the empirical data set treated by Corral [3]. We find a rather dramatic lack of power for the distribution of interevent times to distinguish between quite different sets of parameters, casting doubt on the usefulness of this statistics for the specific purpose of identifying the clustering parameter n.

II. SUMMARY OF RESULTS OBTAINED BY SAICHEV AND SORNETTE [6,7]

A. ETAS model

The ETAS model views the flow of future seismicity as being triggered by past seismicity and by a few background events. Each earthquake is assumed to have the potential to trigger future earthquakes according to three laws capturing the nature of seismicity viewed as a marked point process. We restrict this study to the temporal domain only, summing over the whole spatial domain of interest. First, the magnitude of any earthquake, regardless of time, space, or magnitude of the mother shock, is drawn randomly from the exponential Gutenberg-Richter (GR) law. Its complementary cumulative probability distribution is expressed as

$$Q(m) = 10^{-b(m-m_0)}, \tag{1}$$

where the constant exponent b is typically close to one, and the cutoff m_0 serves to normalize the pdf. We do not consider the influence of an upper cutoff $m_{\max}$, usually estimated in the range 8–9.5 [34,35], because its impact is quite weak in the calculations.

Second, the model assumes that direct aftershocks are distributed in time according to the modified "direct" Omori law (see Ref. [36], and references therein). Denoting the usual Omori law exponent by $p=1+\theta$ and assuming $\theta>0$, the normalized pdf of the Omori law can be written as

$$\Phi(t) = \frac{\theta c^{\theta}}{(t+c)^{1+\theta}}, \tag{2}$$

where t is the time since the earthquake and c is a regularizing constant preventing the divergence of the rate at small times.

Third, the number of direct aftershocks of an event of magnitude m is assumed to follow the productivity law as follows:

$$\rho(m) = \kappa \times 10^{\alpha(m-m_0)}, \quad m_0 \le m, \tag{3}$$

where κ and α are constants. Note that the productivity law (3) is zero below the cutoff m_0, i.e., earthquakes smaller than m_0 do not trigger other earthquakes. The existence of the small-magnitude cutoff m_0 acts as an "ultraviolet" cutoff which is necessary to ensure the convergence of the models of triggered seismicity for $\alpha \le b$.

These laws are combined with the fundamental defining ETAS equation

$$\Lambda(t) = \omega + \sum_{i|t_i<t} \rho(m_i)\Phi(t-t_i), \tag{4}$$

giving the conditional Poisson intensity $\Lambda(t)$ for the occurrence of the next event, conditioned on the history of past events $\mathcal{H}(t)=\{\ldots(t_i,m_i),\ldots,(t_1,m_1)\}$. Here, $t_i<t$ (respectively, m_i) is the time of occurrence (respectively, magnitude) of the ith earthquake counted from the present time t. The term ω is the background contribution assumed to embody the effect of the large scale tectonic driving. Taking the expectation of Eq. (4) yields the average seismic rate

$$E[\Lambda(t)] = \frac{\omega}{1-n}, \tag{5}$$

where n is the key parameter of the ETAS model defined as the number of direct aftershocks per earthquake, averaged over all magnitudes as follows:

$$n \equiv \int_{m_0}^{+\infty} |dQ(m)/dm|\rho(m)dm = \frac{\kappa b}{(b-\alpha)}. \tag{6}$$

As recalled in the Introduction, the fraction of events in a given catalog which have been triggered by previous events can be shown [31] to be exactly given by this "branching ratio" n.

B. Mathematical formulation for the determination of the distribution of interevent times

Saichev and Sornette [6,7] used the formalism of probability generating functions to calculate from first principles

for the ETAS model the distribution $H(\tau,m)$ of waiting times between events of magnitudes larger than or equal to m in a region of seismicity rate $\lambda(m)$. In agreement with previous works [3–7], we express $H(\tau,m)$ as

$$H(\tau,m) \simeq \lambda(m) f(\lambda(m)\tau), \tag{7}$$

so that the dependence on the local seismicity rate is absorbed in the variable $\lambda(m)$ while the more general functional form is captured by the function $f(x)$. Saichev and Sornette first used the general relation

$$H(\tau,m) = \frac{1}{\lambda(m)} \frac{d^2 P(\tau,m)}{d\tau^2}, \tag{8}$$

where $P(\tau,m)$ is the probability of absence of events of magnitude larger than or equal to m within the interval $[t,t+\tau]$. The following expression was obtained:

$$P(\tau,m) = \exp\left(- \int_0^\infty N(t,\tau,m) dt - \int_0^\tau N_-(\tau',m) d\tau' \right), \tag{9}$$

with the auxiliary functions $N_-(\tau,m)$ and $N(t,\tau,m)$ given by the following nonlinear integral implicit equations:

$$N_-(\tau,m) = 1 - \Psi[\Phi(\tau) \otimes N_-(\tau,m)] + Q(m)\Psi[Q^{-1/\gamma}(m)\Phi(\tau) \otimes N_-(\tau,m)], \tag{10}$$

$$N(t,\tau,m) = 1 - \Psi[\Phi(t) \otimes N(t,\tau,m) + \Phi(t+\tau) \otimes N_-(\tau,m)], \tag{11}$$

where

$$\gamma \equiv \frac{b}{\alpha} \tag{12}$$

is assumed larger than 1 (but probably close to 1). $N_-(\tau,m)$ and $N(t,\tau,m)$ have the following probabilistic interpretation:

(1) $N_-(\tau,m)$ is the probability that either some background earthquake occurs in the time window $[t,t+\tau]$, which has a magnitude larger than m, or, if its magnitude is smaller than m, given that it occurred at time t, that it will generate at least one aftershock (or their subsequent daughters) of magnitude larger than m within the interval $[t,t+\tau]$.

(2) Analogously, $N(t,\tau,m)$ is the probability that some background earthquake of magnitude larger than m, occurring at instant $t=0$, will generate at least one aftershock of magnitude larger than m within the time interval $[t,t+\tau]$.

The symbol $\otimes$ stands for the convolution operation over the variable τ in Eq. (10) and in the second part of the argument of the function Ψ in Eq. (11), and over the variable t in the first part of the argument of the function Ψ in Eq. (11). The function $\Psi(z)$ is expressed through the incomplete Gamma function as follows:

$$\Psi(z) = \gamma(\kappa z)^\gamma \Gamma(-\gamma,\kappa z) \ldots . \tag{13}$$

For convenience, we use its expansion in powers of z as follows:

$$\Psi(z) = 1 - nz + \beta z^\gamma - \eta z^2 + \cdots, \tag{14}$$

where β and η are two numerical constants, which can be expressed in terms of κ and γ.

It is clear that the distribution $H(\tau,m)$ of interevent times τ depends on the magnitude cutoff m of events used to construct this distribution. A natural value for this cutoff is the magnitude m_d of so-called completeness of the considered catalog, above which all earthquakes are thought to be recorded by the existing seismic network. This detection threshold m_d has evolved over time together with the technology and density of the seismic networks. In our comparison with Corral's analysis presented below, we use the values m_d reported by him for each corresponding catalog.

The goal of this paper sequel is to calculate the full solution of Eqs. (10) and (11) leading to the expression of $H(\tau,m)$ given by Eq. (8) and to compare this prediction to the data analysis performed by Corral [3] in order to bracket the three key parameters of the ETAS model, θ, γ, and n. The first one describes the direct Omori law. The second one, given the well-known b value, provides a new estimate for the productivity exponent α. The third one n is directly associated with the fundamental question in seismicity of how much clustering occurs in recorded catalogs, as discussed in the Introduction.

C. Analytical solution using the linear approximation

The determination of the form of $f(x)$ defined in Eq. (7) can be analytically resolved only by reducing Eqs. (10) and (11) to their linear approximations, i.e., when only the two first summands of the expansion (14) are considered.

$$\Psi(z) \cong 1 - nz. \tag{15}$$

Using this approximation, Saichev and Sornette [6,7] introduced for convenience the auxiliary function $g(\tau,m_d)$ defined by

$$g(\tau,m_d) = 1 - \frac{N_-(\tau,m_d)}{N_-(m_d)}, \quad N_-(m_d) = \lim_{\tau\to\infty}(N_-(\tau,m_d)). \tag{16}$$

Here, m_d, which is usually larger than m_0, is considered in empirical data analysis to be the magnitude of completeness of seismic catalogs. With the linear approximation, Eq. (15), we have

$$N_-(m_d) = \frac{Q(m_d)}{1-\delta}, \quad \delta = n[1 - Q^{1-1/\gamma}(m_d)]. \tag{17}$$

With Eq. (15), the main remaining problem of solving Eq. (11) can be done by representing the first integral in Eq. (9) via the second one, so that one just needs to determine the function $g(\tau,m_d)$, from which one obtains

$$P(\tau) = \exp\left(- \frac{1-n}{1-\delta}\tau - \frac{1-n}{\delta}\Delta \int_0^\tau g(\tau',m_d) d\tau' \right), \tag{18}$$

where $\Delta = \frac{n}{1-n} - \frac{\delta}{1-\delta}$. Here, we have dropped the explicit dependence on the magnitude, except in the function g, which is written as dependent on the threshold magnitude m_d. Then, using a quasistatic approximation for $g(\tau,m_d)$, Saichev and Sornette obtained

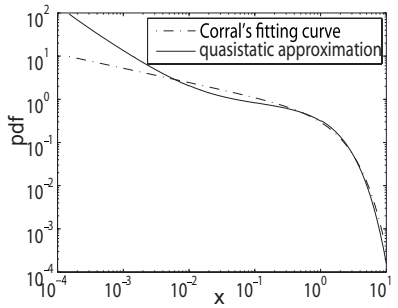

FIG. 1. Nondimensional probability density function $f(x)$ of interevent times for $m_d - m_0 = 2$, $\gamma = 1.2$, $n = 0.9$, $\theta = 0.03$ (continuous line) compared with Corral's phenomenological expression [3] (dashed-dotted line).

$$g(\tau, m_d) \cong \frac{\delta a(\tau)}{1 - \delta + \delta a(\tau)}, \tag{19}$$

where

$$a(\tau) = \int_{\tau}^{\infty} \Phi(t')dt' = \frac{c^{\theta}}{(\tau + c)^{\theta}}. \tag{20}$$

Using the dimensionless variable $x = \lambda\tau$, where $\lambda = \omega Q(m_d)/(1-\delta)$ and ω is the seismic rate of spontaneous seismic sources, expression (8) with Eq. (7) yields

$$f(x) = \frac{d^2\varphi(x, m_d)}{dx^2}, \quad \varphi(x, m_d) = P(\tau) = P\left(\frac{x}{\lambda}\right), \tag{21}$$

leading finally to the dimensionless distribution of interevent times as follows:

$$f(x) = \{\theta\nu(1-\delta)\epsilon^{-\theta}(x+\epsilon)^{\theta-1}g^2(x, m_d, \theta) + [\eta + \nu g(x, m_d, \theta)]^2\}\varphi(x, m_d). \tag{22}$$

Here, $\epsilon = \lambda c$, $\nu = (1-n)\Delta$, and $g(x, m_d, \theta) = g(\tau, m_d) = g(x/\lambda, m_d)$.

Figure 1 reproduces the comparison obtained in Refs. [6,7] between the function $f(x)$ given by Eq. (22) and Corral's phenomenological functional fit [3]. Refs. [6,7] found that expression (22) can fit rather well the empirical distributions of interevent times, so as to even improve on Corral's fit for short time scales, with $\alpha \approx 0.7-0.9$ and $n \approx 0.8-1.0$. These rather large intervals reflect a corresponding insensitivity of the quantitative shape of $f(x)$ with respect to α and n. An analysis of the impact of the first nonlinear term βz^{γ} in the expansion (14) of $\Psi(z)$ in the nonlinear equations (10) and (11) led Saichev and Sornette to expect "weak departures from the results obtained with the linear approximation." They added, "It thus appears that the statistics of recurrence times is not sensitive enough to reveal the importance of these nonlinear corrections, which describe the effect of cascades of generations of aftershocks." It turns out that this statement was premature, as shown by our full treatment of the nonlinear equations. In particular, we identify significant biases in the estimation of the parameter n when using the linear approximation. The reason lies in the fact that, for α close to 1 (specifically, $1 \le \gamma \le 1.2$) as found in Refs. [6,7], the first nonlinear correction βz^{γ} is only weakly nonlinear. We show below that the inclusion of the next term $\sim z^2$ changes somewhat the conclusions. In contrast, the higher-order terms beyond z^2 do not change the conclusions.

III. ANALYSIS OF THE FULL NONLINEAR EQS. (10) and (11)

A. Preparation of the equations and notations

Using the first four summands of the expansion (14), Eqs. (10) and (11) can be written in the following form:

$$N_-(\tau, m_d) = \frac{Q(m_d)}{N_-(m_d)} + \delta\Phi(\tau) \otimes N_-(\tau, m_d) + \sigma N_-(m_d)[\Phi(\tau) \otimes N_-(\tau, m_d)]^2, \tag{23}$$

$$N(t, \tau, m_d) = n[\Phi(t) \otimes N(t, \tau, m_d) + \Phi(t+\tau) \otimes N_-(\tau, m_d)] - \beta[\Phi(t) \otimes N(t, \tau, m_d) + \Phi(t+\tau) \otimes N_-(\tau, m_d)]^{\gamma} + \eta[\Phi(t) \otimes N(t, \tau, m_d) + \Phi(t+\tau) \otimes N_-(\tau, m_d)]^2, \tag{24}$$

where

$$\sigma = \eta[1 - Q^{1-2/\gamma}(m_d)]. \tag{25}$$

In the present case, $N_-(m_d)$ defined in Eq. (16) is not identical to $Q(m_d)/(1-\delta)$ as in the linear approximation. Instead, $N_-(m_d)$ is the root of a simple quadratic equation. However, the difference can be small: for instance, for $m_d - m_0 = 2$, $\theta = 0.03$, $\gamma = 1.2$, $n = 0.9$, we have $Q(m_d)/(1-\delta) \simeq 1.93 * 10^{-2}$ compared with $N_-(m_d) \simeq 1.90 * 10^{-2}$.

Defining the dimensionless variables $x = \lambda t$, $y = \lambda\tau$ with $\lambda = \omega N_-(m_d)$ and the functions

$$M_-(y) = N_-\left(\frac{y}{\lambda}, m_d\right), \quad M(x, y) = N\left(\frac{x}{\lambda}, \frac{y}{\lambda}, m_d\right), \tag{26}$$

$$\Phi_{\epsilon}(x) = \lambda\Phi\left(\frac{x}{\lambda}\right), \tag{27}$$

Eqs. (23) and (24) become

$$M_-(y) = \frac{Q(m_d)}{\lambda} + \delta\Phi_{\epsilon}(y) \otimes M_-(y) + \sigma\lambda[\Phi_{\epsilon}(y) \otimes M_-(y)]^2, \tag{28}$$

$$M(x, y) = n[\Phi_{\epsilon}(x) \otimes M(x, y) + \Phi_{\epsilon}(x+y) \otimes M_-(y)] - \beta[\Phi_{\epsilon}(x) \otimes M(x, y) + \Phi_{\epsilon}(x+y) \otimes M_-(y)]^{\gamma} + \eta[\Phi_{\epsilon}(x) \otimes M(x, y) + \Phi_{\epsilon}(x+y) \otimes M_-(y)]^2. \tag{29}$$

For the numerical calculations, we transform Eq. (28) into an equation for the new function $g(y) = 1 - M_-(y)/\lambda$ as follows:

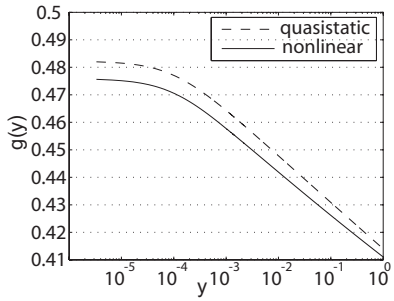

FIG. 2. Solution $g(y)$ of the nonlinear equation (30) (continuous line) and quasistatic approximation (19) (dashed line). The parameters of the ETAS model are $m_d-m_0=2$, $\gamma=1.2$, $n=0.9$, $\theta=0.03$.

$$g(y) = 1 - \delta - \frac{Q(m_d)}{\lambda} + \delta[a_\epsilon(y) + \Phi_\epsilon(y) \otimes g(y)] - \sigma\lambda[a_\epsilon(y) + \Phi_\epsilon(y) \otimes g(y) - 1]^2, \quad a_\epsilon(y) = \int_x^\infty \Phi_\varepsilon(x')dx'. \quad (30)$$

Solving for $g(y)$ instead of $M_-(y)$ is more efficient numerically because $g(y)$ is a monotonically decreasing function unlike $M_-(y)$. This ensures a faster numerical convergence and a weaker sensitivity to the finite mesh size of the discretization scheme. Equations (29) and (30) form the basis for our numerical calculations.

B. Numerical solution

The first step is to solve Eq. (30) for the function $g(y)$ that we reformulate as Eq. (A1) given in Appendix A. We use the method of successive approximations to obtain the value of the function $g(y_i)$ on a regular grid $y_i=y_0+idy$ with a small mesh dy. The performance of this method is discussed in Appendix A in the context of the linear approximation.

Figure 2 shows the difference between the quasistatic approximation (19) and the solution of the nonlinear equation (30). One can observe that the nonlinear solution lies under the quasistatic approximation, i.e., it gives a correction which is in the opposite direction compared with the linear solution (see Fig. 3).

The next step is to determine the function $M(x,y)$, obtained as the solution of Eq. (29). Note that the convolution operation involving $M_-(y)$ can now be expressed in terms of the known function $g(y)$ as follows:

$$\Phi_\epsilon(x+y) \otimes M_-(y) = a_\epsilon(t) - a_\epsilon(t+\tau) - \Phi_\epsilon(x+y) \otimes g(y). \quad (31)$$

In the nonlinear case, the function $M(x,y)$ cannot be represented analytically through the function $M_-(y)$ as in the linear case. This means that we have to calculate $M(x,y)$, a function of two variables (which significantly slows down the calculation speed). Equation (31) implies that the functions $\Phi_\epsilon(x+y)$ and $g(y)$ should be estimated on the same grid points $(x+y)_k$ and y_l. Therefore, the mesh sizes of x and of y should be identical: $dx=dy$.

In order to determine the probability $P(\tau)=\varphi(y,m_d)$, we must also estimate the integral

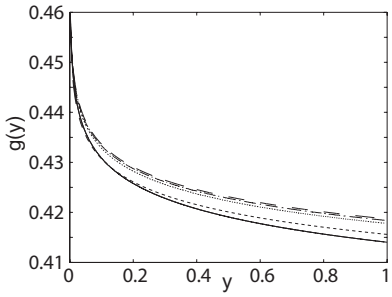

FIG. 3. Function $g(y)$ obtained with different schemes: (solid)—quasistatic approximation developed in [6,7]; numerical solution of Eq. (A1) for (dash-dotted) $dy=10^{-5}$ and (long dashed) -10^{-4}; numerical of solution of Eq. (28) for $M_-(y)$ yielding $g(y)=1-M_-(y)$: (dotted) $dy=10^{-5}$ and (dashed) 10^{-4}. The ETAS parameters are $m_d-m_0=2$, $\gamma=1.2$, $n=0.9$, $\theta=0.03$.

$$\int_0^\infty M(x,y)dy. \quad (32)$$

This requires one to span a large set of y values in order to approximate the theoretical one $[0;+\infty]$ which, together with the condition $dx=dy$, make the problem very demanding in memory capacity. For example, for $x\in[0;1]$ and $y\in[0;1]$ with $dy=dx\sim10^{-4}$, $M(x,y)$ is a matrix with 10^8 elements. To alleviate this burden on memory capacity, we divide the y interval into smaller intervals y_n, $n=1,2,\ldots,N$ and we determine the matrix $M(x,y_n)$ consecutively for each of these subintervals. Having determined the probability function $\varphi(y,m_d)$ on each such small interval y_n, we use a simple smoothing polynomial interpolating scheme in order to prevent jumps in its second-order derivative.

An example of the resulting probability $P(y)$ defined in Eq. (9) is shown in Fig. 4, which identifies a significant difference between the quasistatic approximation presented in Refs. [6,7] and the nonlinear solution.

C. Comparison between the linear and nonlinear versions of the theory and direct ETAS simulated catalogs

We present a comparison for the pdf of interevent times obtained with (i) the linear analytic quasistatic approxima-

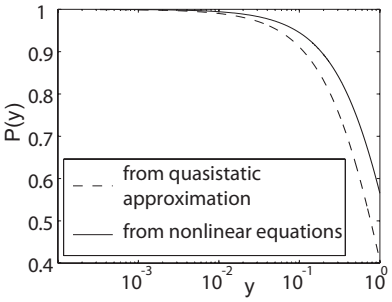

FIG. 4. Probability $P(y)$ defined in Eq. (9) obtained by the numerical solution of the nonlinear equations (29) and (30) (continuous line) compared with the quasistatic approximation reported in Refs. [6,7] (dashed line). The ETAS parameters are $m_d-m_0=2$, $\gamma=1.2$, $n=0.9$, $\theta=0.03$.

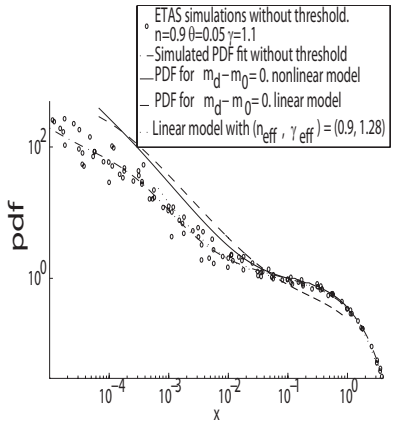

FIG. 5. Pdf of dimensionless interevent times for the ETAS parameters equal to $n=0.9$, $\theta=0.05$, $\gamma=1.1$, $m_d-m_0=0$. The pdf obtained with the nonlinear theory leading to the Eqs. (29) and (30) is shown with the solid line. The pdf obtained with the linear theory is shown as the dashed line. The exact pdf obtained by the simulation method described in Appendix B is shown in histogram form (circles) and with a smoothing fit (dashed-dotted) performed with a function defined as the ratio of a polynomial function of fifth order over another polynomial function of fourth order, in terms of the logarithm of the dimensionless interevent times. The dotted line shows the pdf obtained with the linear theory with effective parameters $n_{\text{eff}}=0.86$ and $\gamma_{\text{eff}}=1.28$, chosen to fit the exact pdf in the region $x \geq 4\times 10^{-2}$, where the nonlinear theory is performing well.

tion, (ii) the numerical solution of the nonlinear equations (29) and (30), and (iii) "exact" synthetic catalogs.

The two former solutions are obtained by taking the second-order derivative of the functions $P(y)$ $[\varphi(y,m)]$ shown in Fig. 4, according to Eq. (21). The synthetic catalog was obtained using the method described in Appendix B. The ETAS parameters used here are $n=0.9$, $\theta=0.05$, $\gamma=1.1$, $m_d-m_0=0$.

Figure 5 shows the three pdf's obtained by the three methods. For the "exact" pdf reconstructed from a synthetic ETAS catalog, we show both the histogram and a fit using a function constructed as the ratio between a polynomial function of order 5 divided by another polynomial function of order 4. These functions are expressed in terms of the logarithm of the dimensionless interevent time. This fit has no pretense of rigor, it only provides a useful guide to the eye.

Figure 5 shows that the linear theory is significantly in error while the nonlinear theory provides an excellent agreement with the exact pdf for values of the dimensionless time interval $x \geq 4\times 10^{-2}$. For smaller x's, the difference is due to numerical errors in the treatment of Eqs. (29) and (30), which can be removed by using an adaptive mesh size, as discussed shortly below. The discrepancy between the exact pdf and the one obtained using the linear approximation implies that a fit of empirical pdf's using the linear theory will likely provide spurious values for the significant parameters n and γ. Indeed, a good fit of the exact synthetic pdf shown in Fig. 5 is obtained with the linear theory using effective parameters $n_{\text{eff}}=0.86$ and $\gamma_{\text{eff}}=1.28$, showing here a systematic bias of 5% in the determination of n and over 16% in the determination of γ (and therefore of the productivity α).

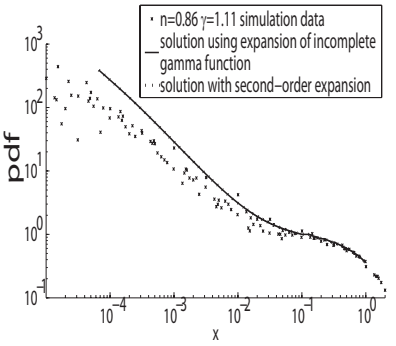

FIG. 6. Test showing that using the exact values of the incomplete gamma function (13) in the calculation of Eq. (29) is undistinguishable from the solution obtained with the expansion (14) which includes all four terms up to second order.

Let us now return to the discrepancy between the nonlinear theory and the exact pdf observed for $x<4\times 10^{-2}$ in Fig. 5. Two possible factors need to be discussed:

(1) impact of terms of order higher than z^2 in the expansion (14) of the function $\Psi(z)$ given by Eq. (13);

(2) lack of convergence of the numerical scheme to solve Eqs. (29) and (30), due to a too large mesh size.

Figure 6 rules out the first explanation, since the solution of the nonlinear equations obtained by using the full expression (13) in the calculation of Eq. (29) is undistinguishable from the solution obtained with the expansion (14). This check and other tests confirm that there is no need to complicate the computations by adding the calculation of the incomplete gamma function. This is important when using our theory for inverting the parameters from fits to empirical data, for instance.

With respect to the second factor, we improve the numerical precision by varying the mesh size dx so that dx/x remains approximately equal to 10^{-4} for $x<0.1$, while dx is fixed at 10^{-4} for $x \geq 0.1$. Thus, for $x \simeq 0.001$, we have chosen $dx \simeq 10^{-7}$, which is the limit that we have been able to handle due to limited numerical precision of the computer. Figure 7 shows for the example $n=0.86$, $\theta=0.05$, $\gamma=1.11$ that the problem previously noted in Fig. 5 disappears: there is a good agreement between the exact pdf obtained from the synthetic ETAS catalog and the nonlinear theory down to $x=0.001$.

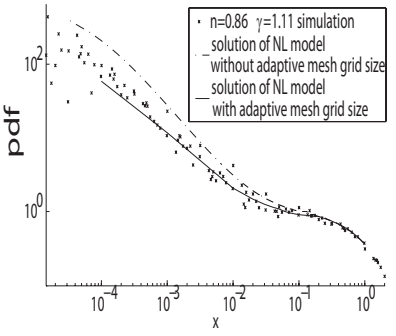

FIG. 7. Comparison between the exact pdf obtained from a synthetic ETAS catalog (crosses) with $n=0.86$, $\theta=0.05$, $\gamma=1.11$ and the nonlinear theory without (dashed line) and with (continuous line) adaptive mesh grid size as described in the text.

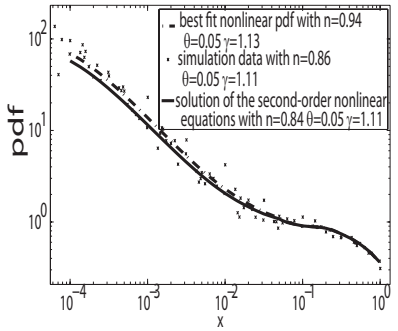

FIG. 8. Comparison between (i) the exact pdf of interevent times (crosses) generated with the ETAS model according to the method described in Appendix B with $n_{\text{cat}}=0.86$, $\theta_{\text{cat}}=0.05$, $\gamma_{\text{cat}}=1.11$, and $\{m_d-m_0\}_{\text{cat}}=0$, which mimics a real-life situation containing fluctuations, (ii) the best fit (dashed line) with the nonlinear theory, and (iii) the pdf obtained with the nonlinear theory with the true values of the parameters (continuous line). For simplicity, we impose the true values $\theta=\theta_{\text{cat}}=0.05$ and $m_d-m_0=\{m_d-m_0\}_{\text{cat}}=0$ in the best fit and invert for the two other parameters, which yields $n_{\text{best fit}}=0.94\pm0.02$ and $\gamma_{\text{best fit}}=1.13\pm0.02$.

D. Test of the inversion of the parameters n and γ using the nonlinear theory from a synthetic ETAS catalog

Consider a synthetic ETAS catalog of interevent times for some fixed values of the parameters n_{cat}, θ_{cat}, γ_{cat}, and $\{m_d-m_0\}_{\text{cat}}$. In this example, we take specifically $n_{\text{cat}}=0.86$, $\theta_{\text{cat}}=0.05$, $\gamma_{\text{cat}}=1.11$, and $\{m_d-m_0\}_{\text{cat}}=0$. Figure 8 shows the exact pdf of interevent times (crosses), which mimics a real-life situation with statistical fluctuations. In a real-life experiment, one would like to use the nonlinear theory to invert for the unknown parameters n,θ,γ,m_d-m_0. In this goal, using the nonlinear theory, we calculate the predicted pdf $f_{\text{NL}}(x)$ of interevent times for fixed values of the parameters n, θ, γ, and m_d-m_0. For a given set of these four parameters, we construct the mean-square error of the logarithm of the pdf over the N interevent times of the catalog as follows:

$$\mathcal{L}(n,\theta,\gamma,m_d-m_0)=\sum_{i=1}^{N}[\ln f_{\text{NL}}(x_i)-\ln f_{\text{cat}}(x_i)]^2, \quad (33)$$

where $f_{\text{NL}}(x_i)$ is the predicted pdf at the dimensionless interevent time x_i given by the nonlinear theory and $f_{\text{cat}}(x_i)$ is the corresponding empirical pdf (in the synthetic catalog). $\mathcal{L}(n,\theta,\gamma,m_d-m_0)$ quantifies how well the nonlinear prediction for the pdf of interevent times can describe the (synthetic) data. The unknown parameters n,θ,γ,m_d-m_0 are then obtained by finding the quadruplets which makes $\mathcal{L}(n,\theta,\gamma,m_d-m_0)$ minimum.

In practice, given the computational cost of the numerical solution of the nonlinear theory, we have found it unpractical to explore systematically the four-dimensional parameter space (with supercomputer resources, this is not excluded but the next section removes the motivation to explore further this option as we will see). For the sake of demonstration, we assume that we already know $\theta=\theta_{\text{cat}}=0.05$ and $m_d-m_0=\{m_d-m_0\}_{\text{cat}}=0$. We are then left with searching for the remaining parameters n and γ. For this, we form a grid in the (n,γ) plane over which we find the minimum of $\mathcal{L}(n,\gamma;\theta_{\text{cat}},\theta_{\text{cat}})$. The corresponding inverted values are $n_{\text{best fit}}=0.94\pm0.02$ and $\gamma_{\text{best fit}}=1.13\pm0.02$. The recovery of γ (and therefore of the productivity exponent α) is good, while there is 9% error on n. Figure 8 shows that this best pdf fits well the exact pdf obtained from the ETAS catalog and is not far from the pdf predicted by the nonlinear theory with the true parameters.

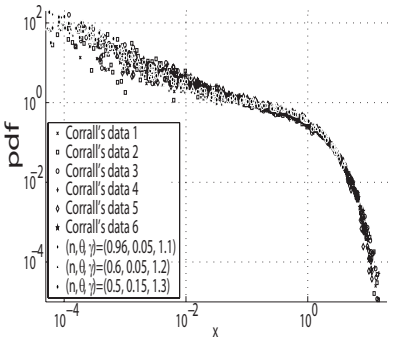

FIG. 9. Comparison between the pdf's of interevent times in synthetic ETAS catalogs with three different sets of parameters and the pdf's obtained by Corral in different regions of the world.

IV. ON THE LACK OF POWER OF THE PDF OF INTEREVENT TIMES TO INVERT FOR THE CLUSTERING PARAMETER n AND OTHER PARAMETERS

The title of this section is motivated by Fig. 9 comparing the pdf of interevent times in synthetic ETAS catalogs with three different sets of parameters and the pdf's obtained by Corral in different regions of the world [3].

First, one can observe that the three triplets ($n=0.96$, $\theta=0.05$, $\gamma=1.1$), ($n=0.6$, $\theta=0.05$, $\gamma=1.2$), and ($n=0.5$, $\theta=0.15$, $\gamma=1.1$) give almost the same pdf's over the whole range of dimensionless interevent times $10^{-4}\leq x\leq 15$. For $x>0.1$, the data collapse is almost perfect, while the scatter is larger for the smaller x values. This is bad news for the determination of the clustering parameter n in particular, since relatively small changes in the Omori law parameter θ and in the productivity law parameter γ can compensate for a quite significant change in the branching ratio n. This suggests that previous claims on the use of the pdf of interevent times to extract efficiently the clustering parameter have been overoptimistic [19,28]. While our simulations and those performed by Hainzl *et al.* [19] as well as the real data used in our respective studies are almost identical, we attribute the difference in our conclusion on (1) the distinct methods (we use the exact theory underlying ETAS) and (2) a matter of interpretation. While we fully endorse the conclusion of Hainzl *et al.* that the interevent times distribution is not universal, we demonstrate that it is almost insensitive to the background rate in the long-time limit, which is the regime quantified by the approach of Hainzl *et al.* [19], based on the gamma distribution [Eq. (1) in their paper]. Furthermore, we note that the correlation between the branching ratio n estimated from the gamma distribution obtained by fitting the

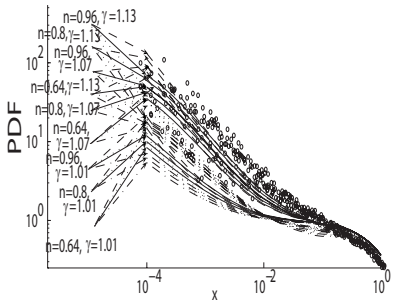

FIG. 10. Comparison between the pdf's obtained by Corral in different regions of the world and the 36 pdf's of interevent times calculated with the nonlinear theory for all combinations of the following sets (with a fixed $\theta=0.05$): $n=0.64,0.8,0.96$; (ii) $\gamma=1.01,1.07,1.13$; (iii) $m_d-m_0=0.1,1$; and (iv) $\epsilon(c)=10^{-4}$, 10^{-5}. Solid lines: $m_d-m_0=0.1$, $c=10^{-4}$; dashed lines: $m_d-m_0=0.1$, $c=10^{-5}$; dashed-dotted lines: $m_d-m_0=1$, $c=10^{-4}$; dotted lines: $m_d-m_0=1$, $c=10^{-5}$.

empirical (or ETAS generated) distributions of interevent times and the branching ratio estimated by a standard declustering algorithm (Reasenberg) is extremely weak: an n value of 0.8 obtained by one method can correspond easily to an n value as low as 0.1 by the other method (see Fig. 3 in Ref. [19]).

Second, Fig. 9 shows that the three chosen triplets of parameters are basically equally good at fitting Corral's data sets [3]. We note that, again for $x>0.1$, all empirical data and ETAS simulations present an almost perfect collapse on a quasiuniversal curve. Larger scatter characterizes smaller x values, which is the region to scrutinize in the hope of extracting some useful constraints on the parameter values.

Actually, the situation is even more involved since, in addition to the parameters n, θ, and γ, a genuine inversion needs also to determine m_d-m_0 (whose impact is significant as shown in Ref. [37]) as well as the regularizing constant c in the Omori law (2). Figure 10 presents the pdf's calculated with the nonlinear theory for different sets of four of these parameters (n,γ,m_d-m_0,c) with a fixed $\theta=0.05$, together with Corral's data. For the pdf's obtained from the nonlinear theory, we used all combinations between the three values $n=0.64,0.8,0.96$, the three values $\gamma=1.01,1.07,1.13$, two values $m_d-m_0=0.1,1$, and two values $\epsilon(c)\equiv\lambda c=10^{-4},10^{-5}$, corresponding to a total of 36 combinations. One can observe roughly two clusters among these 36 theoretical curves. All curves with $\gamma=1.01$ belong to the lower cluster, which is clearly not fitting the data. The upper cluster, which is in better agreement with the data, corresponds to the larger values $\gamma=1.07$ and 1.13. This suggests that the productivity parameter α is likely to be smaller than (instead of equal to) the b value of the Gutenberg-Richter law (recall that $\gamma=b/\alpha$). There is also a smaller impact of ϵ and of m_d-m_0: in general, higher values of these parameters displace the pdf downward.

The comparison between these 36 theoretical pdf's and Corral's data in Fig. 10 shows that there are large uncertainties in the inversion of the parameters. One could argue that the parameter θ should perhaps be modified to a value different from 0.05 in order to better describe the data for small x's. But this region is very sensitive to errors such as resulting from incompleteness [38,39] and its use is problematic.

V. CONCLUSION

We can conclude by the following rather conservative assessment. Recalling the definition of the dimensionless variable $x=\lambda\tau$, where λ is the average seismicity rate of a given region and τ is a realization of the random variable defined as the interevent time between two successive events in that region, we observe on the one hand that the range of dimensionless interevent times $x\geq 0.1$ is probably quite reliable from an empirical viewpoint but the corresponding pdf's are remarkably insensitive to the specific values of the clustering parameter, Omori law exponent, productivity exponent and completeness of the catalogs. On the other hand, the range of $x<0.1$, which would promise to give more sensitivity, is not only highly unreliable but also lacks significant power to obtain a good inversion due to the existence of many almost equally good fits with quite different sets of parameters. Our theoretical analysis and its comparison with Corral's data does not seem to support the proposition that interevent time distributions could provide a new and more reliable way to measure the so-called background earthquake activity as suggested in Refs. [19,28].

ACKNOWLEDGMENT

We are grateful to A. Corral for sharing his data with us.

APPENDIX A: NUMERICAL SOLUTION FOR THE LINEAR APPROXIMATION

This Appendix provides a validation step of the numerical discretization scheme that we have developed to solve the nonlinear equations (29) and (30). Here, we apply this scheme to the linear approximation and compare the result with those which are available analytically. In the linear case, Eq. (30) for $g(y)$ reduces to the following implicit linear integral equation:

$$g(y)=\delta[a_\epsilon(y)+\Phi_\epsilon(y)\otimes g(y)], \tag{A1}$$

where δ is defined in Eq. (17). To solve Eq. (A1), we use the method of successive approximations to obtain the value of the function $g(y_i)$ on a regular grid $y_i=y_0+idy$ with a small mesh dy. This method is adapted to the treatment of the convolution integral in the right-hand side of Eq. (A1). This simple method is fast and provides good convergence. For example, with $dy\sim 10^{-4}$, the calculation converges on the 15th iteration with a residual absolute error $\sim 10^{-15}$.

This is illustrated in Fig. 11, which shows the function $g(y)$ in the linear approximation, obtained directly from the numerical solution of Eq. (A1) for $dy=10^{-5}$ and 10^{-4}, and by using the equation for $M_-(y)$ for $dy=10^{-5}$ and 10^{-4}. As mentioned in the main text, the convergence is faster when using $g(y)$ compared with using $M_-(y)$. Eventually, as the mesh size goes to zero, both methods converge towards the same estimation. An illustration of this convergence is given with $dy=10^{-5}$, which shows much better agreement between the

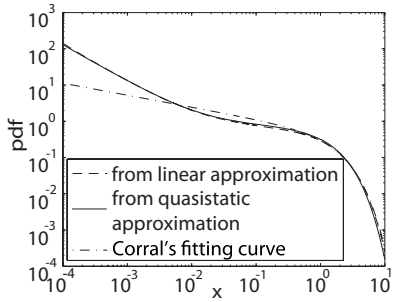

FIG. 11. Probability density functions (pdf) of interevent times obtained by using the function $g(y)$ solution of Eq. (A1) with $dy=10^{-5}$, with Eqs. (A2) and (21) (continuous line). The pdf obtained by Saichev and Sornette with the quasistatic approximation [6,7] is shown as the dashed line. Corral's fitting curve is the dotted-dashed line. The parameters of the ETAS model are $m_d-m_0=2$, $\gamma=1.2$, $n=0.9$, $\theta=0.03$.

two estimations compared with the results obtained for $dy=10^{-4}$. The function $g(y)$ obtained by solving Eq. (A1) is almost identical for $dy=10^{-5}$ and 10^{-4}, demonstrating the faster convergence of this scheme. Figure 11 also shows that the quasistatic approximation is not perfect, but exhibits a relative error of about no more than 1% in this example. Once $g(y)$ has been obtained, the distribution of interevent times is obtained from Eq. (18), which can be expressed here as

$$\varphi(y,m_d)=\exp\left(-\frac{1-n}{1-\delta}y-\frac{1-n}{\delta}\Delta\int_0^y g(x)dx\right), \quad \text{(A2)}$$

and with Eq. (21). Using the function $g(y)$ obtained by solving Eq. (A1) with $dy=10^{-5}$ gives the dimensionless pdf of interevent times shown in Fig. 11. For comparison is also shown the pdf obtained by Saichev and Sornette with the quasistatic approximation [6,7]. There is an excellent agreement between the two methods.

APPENDIX B: ETAS SIMULATIONS

This Appendix describes how we construct the pdf of interevent times in specific synthetic catalogs generated with the ETAS model. Actually, we do not generate synthetic catalogs. Instead, we use the analytical form of the cumulative distribution function (CDF) $F_{k+1}(\tau)$ of the waiting time τ between the kth and the $(k+1)$th event, knowing the times and magnitudes of the preceding k events, to draw the occurrence time of this $(k+1)$th event. Generating in this way 1000 or more interevent times, we use logarithmic bins to construct the histogram of these interevent times. This construction provides the "true" or exact numerical benchmark against which to compare our theory and the empirical data.

The main advantage of this algorithm compared with standard numerical codes that generate synthetic catalogs is that it runs significantly faster because it uses the specific feature of the ETAS model and its known conditional (CDF) of interevent times. Thus, parts of the calculations can be done analytically. We generate the occurrence times one by one by finding the CDF of the time of the next event based on the knowledge of the previous CDF and the time of the previous event. We use a standard Newton algorithm as well as a standard randomization algorithm.

The CDF $F_{k+1}(\tau)$ is obtained by recurrence as follows. For the first event, $F_1(\tau)$ is nothing but the cumulative probability of occurrence of a spontaneous (background) shock since the origin of time, given by definition by the Poisson law with rate ω as follows:

$$F_1(\tau)=1-e^{-\omega\tau}. \quad \text{(B1)}$$

The CDF $F_2(\tau)$ of the waiting time from the first to the second shock is made of two contributions: (i) the second shock may again be a background event or (ii) it may be triggered by the first shock. This yields

$$F_2(\tau)=1-e^{-\omega\tau}e^{-\rho_1[1-a(\tau)]}, \quad \text{(B2)}$$

where $\rho_1=\rho(m_1)$ is the productivity of the first shock obtained from expression 3 given its magnitude m_1, and $a(\tau)$ is defined in Eq. (20).

All following shocks are similarly either a background event or triggered by one of the preceding events. The CDF $F_3(\tau)$ of the waiting time between the second and the third shocks is thus given by

$$F_3(\tau)=1-e^{-\omega\tau}e^{-\rho_1[a(\tau_2)-a(\tau_2+\tau)]-\rho_2[1-a(\tau)]}, \quad \text{(B3)}$$

where τ_2 is the realized time interval between the first and the second shocks. Iterating, we obtain the CDF $F_k(\tau)$ for the waiting between the $(k-1)$th and kth shocks under the following form:

$$F_k(\tau)=1-e^{-\omega\tau}\exp\left\{-\sum_{i=1}^{k-1}\rho_{k-i}\left[a\left(\sum_{j=2}^{i}\tau_j\right)-a\left(\sum_{j=2}^{i}\tau_j+\tau\right)\right]\right\}, \quad \text{(B4)}$$

where τ_j is the waiting time between the $(j-1)$th event and the jth event, and $\rho_i=\rho(m_i)$ is the productivity of the ith shock obtained from expression (B4) given its magnitude m_i.

In order to generate the $(k+1)$th interevent time interval between the occurrence of the kth and $(k+1)$th shock, it is necessary to know the k previous interevent times between the k previous shocks and their k magnitudes. Since, in the ETAS model, the magnitudes are drawn independently according to the Gutenberg-Richter distribution (1), they can be generated once for all. In order to generate a catalog of N events, we thus draw N magnitudes from the law (1). In order to generate the corresponding N interevent times, we use the expression (B4) iteratively from $k=1$ to $k=N$ in a standard way: since any CDF $F(x)$ of a random variable x is by construction itself uniformly distributed in $[0,1]$, we obtain a given realization $x*$ of the random variable x by drawing a random number r uniformly in $[0,1]$ and by solving the equation $F(x*)=r$. In our case, we generate N-independent uniformly distributed random numbers $x_1,\ldots,x_N$ in $[0,1]$ and determine each τ_i successively as the solution of $F_i(\tau_i)=x_i$.

[1] J. Gao, Y. Cao, and J. Hu, *Recurrence Time Distribution, Renyi Entropy, and Pattern Discovery*, 2005 Conference on Information Sciences and Systems (The Johns Hopkins University, Baltimore, MD, 2005), p. 1618.
[2] G. M. Zaslavsky and M. K. Tippett, Phys. Rev. Lett. **67**, 3251 (1991).
[3] A. Corral, Phys. Rev. E **68**, 035102(R) (2003).
[4] A. Corral, Phys. Rev. Lett. **92**, 108501 (2004).
[5] V. Livina, S. Tuzov, S. Havlin, and A. Bunde, Physica A **348**, 591 (2005).
[6] A. Saichev and D. Sornette, Phys. Rev. Lett. **97**, 078501 (2006).
[7] A. Saichev and D. Sornette, J. Geophys. Res. **112**, B04313 (2007).
[8] J. Davidsen, S. Stanchits, and G. Dresen, Phys. Rev. Lett. **98**, 125502 (2007).
[9] A.-L. Barabási, Nature (London) **435**, 207 (2005).
[10] A. Vazquez, J. G. Oliveira, Z. Dezso, K. I. Goh, I. Kondor, and A. L. Barabasi, Phys. Rev. E **73**, 036127 (2006).
[11] P. Bak, K. Christensen, L. Danon, and T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002).
[12] A. Corral, Physica A **340**, 590 (2004).
[13] A. Corral, Nonlinear Processes Geophys. **12**, 89 (2005).
[14] A. Corral, Phys. Rev. Lett. **95**, 028501 (2005).
[15] A. Corral, Phys. Rev. Lett. **97**, 178501 (2006).
[16] A. Corral and K. Christensen, Phys. Rev. Lett. **96**, 109801 (2006).
[17] J. Davidsen and C. Goltz, Geophys. Res. Lett. **31**, 10.1029/2004GL020892 (2004).
[18] V. N. Livina, S. Havlin, and A. Bunde, Phys. Rev. Lett. **95**, 208501 (2005).
[19] S. Hainzl, F. Scherbaum, and C. Beauval, Bull. Seismol. Soc. Am. **96**, 313 (2006).
[20] M. Lindman, K. Jonsdottir, R. Roberts, B. Lund, and R. Bodvarsson, Phys. Rev. Lett. **94**, 108501 (2005).
[21] M. Lindman, K. Jonsdottir, R. Roberts, B. Lund, and R. Bodvarsson, Phys. Rev. Lett. **96**, 109802 (2006).
[22] G. Molchan, Pure Appl. Geophys. **162**, 1135 (2005).
[23] A. G. Hawkes, Biometrika **58**, 83 (1971a).
[24] A. G. Hawkes, J. R. Stat. Soc. Ser. B (Methodol.) **33**, 438 (1971).
[25] A. G. Hawkes and D. Oakes, J. Appl. Probab. **11**, 493 (1974).
[26] Y. Y. Kagan and L. Knopoff, J. Geophys. Res. **86**, 2853 (1981).
[27] Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988).
[28] J. Hardebeck, Background seismicity rates from inter-event-time statistics: Spatial patterns appear stationary through time, working paper, 2007 (unpublished).
[29] D. Sornette, F. Deschatres, T. Gilbert, and Y. Ageon, Phys. Rev. Lett. **93**, 228701 (2004).
[30] R. Crane and D. Sornette, Proc. Natl. Acad Sci USA (to be published) (reprint at http://arXiv.org/abs/0803.2189).
[31] A. Helmstetter and D. Sornette, Geophys. Res. Lett. **30**, 10.1029/2003GL017670 (2003).
[32] A. Helmstetter and D. Sornette, J. Geophys. Res. **107**, 2237 (2002).
[33] P. Reasenberg, J. Geophys. Res. **90**, 5479 (2005).
[34] Y. Y. Kagan, Pure Appl. Geophys. **155**, 537 (1999).
[35] V. F. Pisarenko, A. Sornette, D. Sornette, and M. V. Rodkin, Pure Appl. Geophys. (to be published); http://arxiv.org/abs/physics/0703010.
[36] T. Utsu, Y. Ogata, and R. S. Matsu'ura, J. Phys. Earth **43**, 1 (1995).
[37] D. Sornette and M. J. Werner, J. Geophys. Res. **110**, B09303 (2005).
[38] A. Helmstetter, Y. Y. Kagan, and D. D. Jackson, J. Geophys. Res. **110**, B05S08 (2005) .
[39] Y. Y. Kagan, Phys Earth Planet Inter. **135**, 173 (2003).