Eur. Phys. J. B **51**, 443–459 (2006)
DOI: 10.1140/epjb/e2006-00242-6

THE EUROPEAN PHYSICAL JOURNAL B

Renormalization of branching models of triggered seismicity from total to observable seismicity

A. Saichev[1,2] and D. Sornette[3,4,a]

[1] Mathematical Department, Nizhny Novgorod State University, Gagarin prosp. 23, Nizhny Novgorod, 603950, Russia
[2] Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095, USA
[3] Institute of Geophysics and Planetary Physics and Department of Earth and Space Sciences, University of California, Los Angeles, CA 90095, USA
[4] Laboratoire de Physique de la Matière Condensée, CNRS UMR 6622 and Université de Nice-Sophia Antipolis, 06108 Nice Cedex 2, France

Received 5 March 2006
Published online 21 June 2006 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2006

Abstract. Several recent works point out that the crowd of small unobservable earthquakes (with magnitudes below the detection threshold m_d) may play a significant and perhaps dominant role in triggering future seismicity. Using the ETAS branching model of triggered seismicity, we apply the formalism of generating probability functions to investigate how the statistical properties of observable earthquakes differ from the statistics of all events. The ETAS (epidemic-type aftershock sequence) model assumes that each earthquake can trigger other earthquakes ("aftershocks"). An aftershock sequence results in this model from the cascade of aftershocks of each past earthquake. The triggering efficiency of earthquakes is assumed to vanish below a lower magnitude limit m_0, in order to ensure the convergence of the theory and may reflect the physics of state-and-velocity frictional rupture. We show that, to a good approximation, the statistical distribution of seismic rates of events with magnitudes above m_d generated by an ETAS model with branching ratio n is the same as that of events generated by another ETAS model with effective parameter $n(m_d)$. Our present analysis thus confirms, for the full statistical (time-independent or large time-window approximation) properties, the results obtained previously by one of us and Werner, based solely on the average seismic rates (the first-order moment of the statistics). Our analysis also demonstrates that this correspondence is not exact, as there are small corrections which can be systematically calculated, in terms of additional contributions that can be mapped onto a different branching model. We also show that this approximate correspondence of the ETAS model onto itself obtained by changing m_0 into m_d, and n into $n(m_d)$ holds only with respect to its statistical properties and not for all its space-time properties.

PACS. 64.60.Ak Renormalization-group, fractal, and percolation studies of phase transitions – 91.30.Dk Seismicity – 02.50.Ey Stochastic processes

1 Introduction

In the last few years, physicists' interest for the space-time organization of seismicity in different regions of the world has spurred. This recent burst of attention is probably due to the introduction of new diagnostic tools applied to earthquake catalogs [1–19] and to improved insights from cartoon models of earthquakes [20–25]. The first class of papers in particular suggest to re-examine the standard statistical properties of earthquakes, usually documented under the following distinct power law and fractal properties: (i) the Gutenberg-Richter distribution $\sim 1/E^{1+\beta}$ (with $\beta \approx 2/3$) of earthquake energies E [26]; (ii) the Omori law $\sim 1/t^p$ (with $p \approx 1$ for large earthquakes) of the rate of aftershocks as a function of time t since a mainshock [27]; (iii) the productivity law $\sim E^a$ (with $a \approx 2/3$) giving the number of earthquakes triggered by an event of energy E [28]; (iv) the power law distribution $\sim 1/L^2$ of fault lengths L [29]; (v) the fractal structure of fault networks [30] and of the spatial organization of earthquake epicenters [31]; (vi) the distribution $1/s^{2+\delta}$ (with $\delta \geq 0$) of seismic stress sources s in earthquake focal zones due to past earthquakes [32]. Specifically, the statistical analysis based on (a) coarse-grained scaling ansatz [1,2,7–9,13] (b) entropic methods [4,15], and (c) network

[a] Now at D-MTEC, ETH Zurich, CH-8032 Zurich, Switzerland; e-mail: dsornette@ethz.ch

methods [6,10–12] suggest that the above standard seismological descriptions [26–32] are incomplete. It is not clear however what should be the correct physical model. Several papers have however questioned the novelty of the insights derived from these approaches [2,5,16].

The present authors are among those who have studied how the standard seismological laws [26–32] (in particular the laws (i)–(iii) mentioned above) could actually go a long way towards explaining most of the empirical phenomenology of seismicity, including the supposed anomalous or "novel" scaling laws proposed by the above quoted physicists (see for instance [33–38]). In this series of papers, we have developed a consistent statistical description of seismicity using models of triggered seismicity, which allows one to make quantitative predictions of observables that can be compared with empirical data. The simplest class of models of triggered seismicity combines the above mentioned Gutenberg-Richter (i), Omori (ii), and productivity laws (iii) which can be applied to a fractal spatial geometry of earthquake epicenters (v) [38]. The fundamental physical ingredient introduced in the construction of these models of triggered seismicity is that each earthquake can trigger other earthquakes and an earthquake sequence results in these models from the cascade of events triggered by past earthquakes. The usual notions of foreshocks, mainshocks and aftershocks lose their specificity as any earthquake can be triggered by previous earthquakes and may trigger itself subsequent earthquakes. Here, we continue our study of the general branching process, called the Epidemic-Type Aftershock Sequence (ETAS) model of triggered seismicity, introduced by Ogata in the present form [42] and by Kagan and Knopoff in a slightly different form [43] and whose main average statistical properties are reviewed in [44]. This model has been shown to constitute a powerful null hypothesis to test against other models [42]. The ETAS model belongs to a general class of branching processes [45,46]. It can be viewed as the monofractal approximation of the more general multifractal model of triggered seismicity introduced recently in [24,25], which derives from the physics of thermally activated rupture aided by stress.

The physical problem addressed here is the following. We start from the suggestion [28,47] that small earthquakes dominate or are at least equivalent collectively to large earthquakes in triggering other earthquakes. This conclusion relies on a combination of empirical evidence interpreted in the context of models of triggered seismicity such as the ETAS model. Combining the Gutenberg-Richter law (i) $\propto 10^{-bm}$ and the productivity law (iii) $\propto 10^{\alpha m}$, which are both empirically based, and using the conceptual framework of the ETAS model in which any earthquake can trigger or be triggered by other earthquakes, one obtains the typical number $\propto 10^{-(b-\alpha)m}$ of events triggered by earthquakes of magnitude between m and $m+1$. With the empirical estimates of $b \approx 1$ and $0.8 \leq \alpha \leq 1$ together with the observation that triggered events seem to have magnitudes with only weak or no relation with the magnitude of the triggering event (magnitude-independence law) [47] (i.e., large earthquakes can be triggered by small events), this implies the perhaps surprising conclusion that, according to the ETAS model, large earthquakes are triggered more by the swarm of small previous earthquakes than by preceding large earthquakes. This stems from (i) the observation that the number of small earthquakes increases faster as their magnitude decrease than their productivity decreases and (ii) the assumption that all earthquakes play the same role (which does not seem to be contradicted by observations). The conclusion that small earthquakes dominate triggering is thus intrinsically a collective effect. This picture, which emphasizes the collective organization of earthquakes or "many-body" view, can be contrasted with the "one-body" or few-body approach of Stein and co-workers [48,49] which focuses exclusively on how a few large earthquakes can promote subsequent shocks at some sites and inhibit them in others. If indeed the small earthquakes dominate in the triggering of future events, this begs to define how small "small" can be, since the smaller the earthquakes the larger their triggering influence. The evidence that small earthquakes should dominate triggering is based on the empirical statistics (i) and (iii) established for event magnitudes above magnitude 2 or 3 (depending on the completeness of the studied catalogs). The question of how small "small" amounts to asking how far in the small magnitude range can the productivity law and the magnitude-independence law be extrapolated. Because the Gutenberg-Richter law (i) has been observed at such small scales as individual dislocation motions, we know for sure that there must be a lower "ultra-violet" cut-off magnitude m_0 at which the productivity of events of magnitude smaller than m_0 tapers off or vanishes. Otherwise, the factor $\propto 10^{-(b-\alpha)m}$ would diverges as $m \to -\infty$ (energy goes to zero). Is the ultra-violet cut-off associated with an atomic scale for rupture? Or are other relevant scales? This question has been addressed in two recent papers by M. Werner and one of us [50,51] within the framework of the ETAS model. Consider a catalog complete for magnitudes above some observational threshold m_d, i.e., all earthquakes with magnitudes $m \geq m_d$ have been recorded but smaller earthquakes are not. Noting that the magnitude m_d of completeness of a seismic catalog is not in general the same as the magnitude m_0 of the smallest triggering earthquake, reference [50] showed that bounds for m_0 can be obtained from quantitative fits to observed aftershock sequences. In addition, reference [51] remarked that, in models of triggered seismicity and in their estimation from empirical data, the detection threshold m_d is commonly equated to the magnitude m_0 of the smallest triggering earthquake. This unjustified assumption neglects the possibility of shocks below the detection threshold triggering observable events, a process which should dominate according to our previous discussion. Reference [51] developed a mean field formalism within the ETAS model: by considering the branching structure of one complete cascade of triggered events, the catalog of observed events with magnitude above m_d was shown to be described by an effective "renormalized" ETAS model with its lower magnitude cut-off equal to m_d but with an apparent branching ratio n_a (which is the apparent fraction of aftershocks in a given catalog) and an apparent background source S_a, due to the presence of smaller undetected events capable of triggering larger events. This result is potentially very important since it implies that previous estimates of the clustering characteristics of seismicity may significantly underestimate the true values: for instance, an observed fraction of 55% of aftershocks is renormalized into a true value of 75% of triggered events.

The object of the present paper is to extend the previous mean field treatment to obtain the full earthquake statistics using the formalism of generating probability function (GPF) already developed in [35–38]. In a sense, the question addressed here is whether the ETAS model can be renormalized onto itself by moving m_0 to $m_d > m_0$ (which can be seen as a coarse-graining operation), that is, is there an effective ETAS model with minimum magnitude m_d and with renormalized parameters, which describes the observed catalogs? Beyond its interest and application to earthquakes, this problem is relevant to a general understanding of coarse-grained properties of marked branching processes, to which our formalism applies.

The focus of the paper is on the statistical properties of clusters in synthetic catalogs generated by the ETAS model. In its standard formulation, the ETAS model is defined in terms of the conditional Poisson intensity which controls the probability that a new earthquake occurs in the near future. This Poisson intensity is expressed as a sum of triggering contributions from all past earthquakes. It would thus seem that clusters cannot be clearly defined since the occurrence of future earthquakes depends on all past earthquakes and thus all earthquakes are in this sense linked or clustered together. In fact, it is easy to show that, due to the linear contribution of past earthquakes in the conditional intensity together with the exponential (Poisson) form of the probability distribution of waiting times for the next earthquake, the ETAS model is fundamentally a branching process (see for instance [51] where this is explicited). As a consequence, an equivalent formulation of the ETAS model amounts to generate earthquakes according to a specific branching hierarchy of ancestors and progenies. This remark is actually very useful to generate synthetic earthquake catalogs based on an algorithm [39] with is much faster than the original relying on a sum over all past earthquakes which needs to be re-estimated at each time step [42]. The notion that clusters of earthquakes have a genuine physical meaning in the ETAS model is further made explicit by the recently introduced "stochastic reconstruction" method developed by Zhuang et al. [40,41]. As this method shows, clusters have a precise statistical meaning.

The organization of the paper is the following. Section 2.1 introduces the general formulation of observable clusters of triggered events using the generating probability functions (GPF). It also presents a simple intuitive approximation which will be make rigorous in later sections. Section 2.2 derives general relations for the effective branching rates of observable events. Section 2.3 defines the ETAS model and recalls its main useful properties. Section 2.4 introduces the GPF for unobservable and observable aftershocks. Section 2.5 gives the main properties of the effective branching rates, which recover the previous analysis of [51] in a slightly different form. Section 2.6 explains that the present approach and that of [51] are equivalent physically but with a different mathematical formulation. The justification for introducing a physically equivalent but mathematically different formulation here is that it is more adapted to the calculations of the full statistics with the GPF formalism. Section 3 presents all our results on the statistics of observable events in the ETAS branching model. Section 3.1 derives the general equation governing the GPF of observable events. Section 3.2 uses the derivation of the previous section to give quantitative estimates for the fraction of observable events. Section 3.3 discusses the approximation of self-similarity, corresponding to a renormalization of the ETAS model onto itself by the change from m_0 to m_d, *for its statistical properties*. This self-similarity amounts to say that the statistics of observable events can be deduced entirely from the statistics of all events under a simple renormalization of the average branching ratio into an effective value. Section 3.4 derives the implications of the self-similar approximation for the distribution of the numbers of observable events. Section 3.5 discusses the deviations from self-similarity and identifies a correction in the form of a new branching model, which gives rather small corrections to the previously self-similar estimates. Section 3.6 shows that the effective parameter $n(m_d)$ can be interpreted as a normalized branching ratio only for the global statistical properties of seismic rates, which are time- and space-independent, but not for the full space and time-dependent properties. The last section concludes. Finally, an Appendix presents a brief pedagogical introduction to the general formalism of generating probability functions (GPF) and introduces the main formulas useful to follow the calculations presented below.

2 Definition and properties of effective rate of observable aftershocks

2.1 General formulation of observable clusters

In this section, we present the general formulation of generating probability function (GPF) (see the Appendix for notations and definitions) for marked branching processes with an observational constraint. Recall that, for general branching processes such as the ETAS model, the GPF formalism allows one to calculate the full statistical properties. Here, the mark associated with an event is its magnitude. The observational constraint is that only events with magnitude $m \geq m_d$, where m_d is the observation threshold, are known, while the process produces events which can have a lower magnitude, down to a lower triggering cut-off m_0.

To get a first feeling of how the observational constraint can be taken into account in the GPF formalism, consider the case of a large finite time window of size τ in which we count the number of events and let us use the approach developed in [36] for the statistics of the number of such windowed events. The time windows are considered large if their size is significantly larger than the typical life-time of the clusters, defined as the sequences of aftershocks, triggered by single background events (see [36] for a discussion). In this limit, the statistics of the total (observable and unobservable) number of events in a window of size τ is obtained from the generating probability function (GPF) $\Theta_w(z;\tau)$, which obeys the following equation

$$\Theta_w(z;\tau) = e^{\omega\tau[\Theta(z)-1]}. \tag{1}$$

$\Theta(z)$ is the GPF of the number of all the aftershocks triggered by a given source, including the source event itself and ω is the Poisson intensity of the background sources. We have shown [35] that $\Theta(z)$ has the structure

$$\Theta(z) = zG(z), \tag{2}$$

where the factor z to the left of $G(z)$ takes into account the contribution of the background source, while the GPF $G(z)$ describes the statistics of the number of all aftershocks within a given cluster. In view of (2), the GPF for finite time windows given by (1) is the natural generalization of the GPF obtained for the Poissonian background events:

$$\Theta_b(z;\tau) = e^{\omega\tau(z-1)}. \tag{3}$$

Now, the statistics of observable events requires to replace the GPF $\Theta(z)$ in (1) by the GPF $\Theta(z;m_d)$ of the number of aftershocks (and their sources) whose magnitudes m are larger than the detection threshold m_d to obtain

$$\Theta_w(z;m_d,\tau) = e^{\omega\tau[\Theta(z;m_d)-1]}. \tag{4}$$

Note that some clusters might have no observable events at all. This means that there is a non-zero probability

$$p(m_d) = \Theta(z=0;m_d) \neq 0 \tag{5}$$

that the cluster is completely unobservable. The complementary probability

$$q(m_d) = 1 - p(m_d) = 1 - \Theta(0;m_d) \tag{6}$$

is the probability that there is at least one observable event (source or some aftershock) in the cluster under inspection. In what follows, we refer to a cluster as "observable," if it contains at least one observable event (with magnitude $m \geq m_d$). Accordingly, $q(m_d)$ defined in (6) is the probability that a cluster is observable; it is also the fraction of observable clusters.

It is convenient to express the GPF $\Theta(z;m_d)$ in the form

$$\Theta(z;m_d) = q(m_d)\tilde{\Theta}(z;m_d) + 1 - q(m_d), \tag{7}$$

where

$$\tilde{\Theta}(z;m_d) = \frac{1}{q(m_d)}\left[\Theta(z;m_d) - \Theta(0;m_d)\right] \tag{8}$$

is nothing but the conditional GPF of the number of observable events within observable clusters. This definition implies that it has the same structure

$$\tilde{\Theta}(z:m_d) = z\tilde{G}(z;m_d), \tag{9}$$

as that given by (2) of the GPF $\Theta(z)$ of the total number of events belonging to some cluster. Expression (9) implies that one can treat the observable event which comes first in time as the "observable source," and then interpret $\tilde{G}(z;m_d)$ as the GPF of its observable aftershocks.

Substituting (7) into (4) yields the following representation for the GPF of the number of observable windowed events

$$\Theta_w(z;m_d,\tau) = e^{\omega(m_d)\tau[\tilde{\Theta}(z;m_d)-1]}, \tag{10}$$

where

$$\omega(m_d) = \omega q(m_d) \tag{11}$$

is the renormalized intensity of "observable sources," which is the same as the intensity of observable clusters by definition.

There is a physically transparent way to estimate the probability $q(m_d)$ that a cluster is observable. It is indeed always possible to represent $q(m_d)$ in the form

$$q(m_d) = q^+(m_d)Q(m_d) + q^-(m_d)[1 - Q(m_d)]. \tag{12}$$

Here, $q^+(m_d)$ (respectively $q^-(m_d)$) is the probability that the cluster is observable under the condition that its generating source is also observable (respectively unobservable). In addition, $Q(m_d)$ is the probability that the source is observable. Obviously, we have

$$q^+(m_d) \equiv 1 , \qquad q^-(m_d) = 1 - p^-(m_d), \tag{13}$$

where $p^-(m_d)$ is the probability that all aftershocks triggered by an unobservable event are unobservable. To estimate $p^-(m_d)$, we make the assumption that each unobservable event either triggers only one first-generation aftershock, with probability $\nu(m_d)$, or does not trigger any aftershocks with the probability $1-\nu(m_d)$. This approximation is quite reasonable, as can be seen from the application of the productivity law $\sim E^a \sim 10^{\alpha m}$ (with $a \approx 2/3$, $\alpha \approx 1$): if an event of magnitude $m = 7$ produces about 10^5 observable events on average, an event of magnitude 2 triggers about 0.1 events on average. In this example, $\nu(m_d) \approx 0.1$ and $1-\nu(m_d) \approx 0.9$ and the error in neglecting the possibility for this event to trigger two aftershocks is of order 0.01. This error becomes even smaller for smaller unobservable sources.

Suppose additionally that the magnitudes of the triggered aftershocks are statistically independent of each other. Then, the probability that all aftershocks, triggered by an unobservable event, are unobservable, is given by

$$p^-(m_d) \simeq \sum_{k=0}^{\infty}[1-\nu(m_d)]\nu^k(m_d)[1-Q(m_d)]^k, \tag{14}$$

where $(1-\nu)\nu^k$ is the geometrical probability that an unobservable event triggers k aftershocks, while $(1-Q)^k$

is the probability that they are all unobservable. After summation, we obtain

$$p^-(m_d) \simeq \frac{1-\nu(m_d)}{1-\nu(m_d)[1-Q(m_d)]}. \quad (15)$$

Substituting this expression into (13) and then (13) into (12), we obtain the probability that a cluster is observable under the form

$$q(m_d) \simeq \frac{Q(m_d)}{1-\nu(m_d)[1-Q(m_d)]}. \quad (16)$$

In the following, we obtain with the framework of the ETAS model, a physically transparent relation for the probability $\nu(m_d)$, which will allow us to obtain an accurate estimation of the probability $q(m_d)$ given by (16) and the corresponding renormalized intensity $\omega(m_d)$ of "observable sources" given by (11). Specifically, the role of $\nu(m_d)$ is derived in expression (37) below for the GPF of first-generation aftershocks triggered by unobservable event.

2.2 Effective observable aftershocks rate

Before turning to the specifics of the ETAS model and its statistics of observable events, it is useful to discuss the properties of the average rate of general branching processes. The results obtained in this section recover those obtained in [51] within a slightly different interpretation, that we present to be self-contained and to connect with the subsequent derivation of the full number statistics in following sections.

It is well-known that the key parameter controlling the properties of cascades of triggered events is the branching rate n, which is nothing but the average number of first generation aftershocks, where the average is performed over all possible triggering event of arbitrary magnitude. The cases $n<1$, $n=1$ and $n>1$ correspond respectively to the sub-critical, critical and super-critical regimes, with the first-two giving stationary time series in the presence of a stationary immigration of sources and the later giving explosive time series with a positive probability [44–46].

In branching processes (of which the ETAS model is an example), we can use the representation that each shock triggers independently its own aftershocks sequence (see [51] for a discussion on the two interpretations in terms of decoupled branches used here or of collective triggering; the two views are equivalent due to the linear sum over past events and the conditional Poisson process formulation). The independence between different branches allows us to obtain the average $\langle R\rangle$ of the total number of events (mainshock itself and all its offsprings over all generations) triggered by an arbitrary mainshock as [52]

$$\langle R\rangle = 1+n+n^2+\cdots = \frac{1}{1-n}, \quad (17)$$

where n^k is the contribution of the k-th generation of aftershocks. Thus, if the average number $\langle R\rangle$ of events per cluster is known, the aftershocks rate can then be obtained as

$$n = \frac{\langle R\rangle - 1}{\langle R\rangle}. \quad (18)$$

This simple remark will be useful in the following to derive an apparent or renormalized branching ratio $n(m_d)$ and test its usefulness to describe the full number statistics.

The average number of observable events, which are triggered by some arbitrary source, is simply $\langle R\rangle$ given by (17) multiplied by the probability $Q(m_d)$ that an event is observable:

$$Q(m_d) = \int_{m_d}^{\infty} P(m)dm, \quad (19)$$

where $P(m)$ is the probability density function (PDF) of their random magnitudes (assumed to be the same for sources and all aftershocks). This gives

$$\langle R\rangle(m_d) = \langle R\rangle Q(m_d) = \frac{Q(m_d)}{1-n}. \quad (20)$$

Consider now the conditional average $\langle\tilde{R}\rangle(m_d)$ of the number of observable events within some observable cluster. It is simply given by

$$\langle\tilde{R}\rangle(m_d) = \frac{\langle R\rangle(m_d)}{q(m_d)} = \frac{Q(m_d)}{(1-n)q(m_d)}, \quad (21)$$

as a result of the law of conditional probabilities, using the fact that $q(m_d)$ is the probability that the cluster is observable, as given by (16). This quantity $\langle\tilde{R}\rangle(m_d)$ can not be derived as straightforwardly as the average number $\langle R\rangle$ over all events obtained with (17). Indeed, an observable cluster may result from an effective "observable source" which might belong, for instance, to the 3th or even 7th generation of the total aftershock sequence. Moreover, it seems impossible to classify uniquely observable events of observable clusters as belonging to observable aftershocks of first, second or kth generations. Therefore, it is not possible to use for $\langle\tilde{R}\rangle(m_d)$ the reasonings underlying relation (17). Notwithstanding this limitation, we can introduce an effective branching ratio for observed clusters, based on a natural extension of relation (18). Let us thus define the effective branching ratio of observable clusters as

$$n(m_d) = \frac{\langle\tilde{R}\rangle(m_d) - 1}{\langle\tilde{R}\rangle(m_d)}. \quad (22)$$

With (21), this gives

$$n(m_d) = \frac{Q(m_d) - (1-n)q(m_d)}{Q(m_d)}. \quad (23)$$

Substituting in (23) the r.h.s. of equality (16) yields

$$n(m_d) \simeq \frac{n-\nu(m_d)[1-Q(m_d)]}{1-\nu(m_d)[1-Q(m_d)]}, \quad (24)$$

expressing the effective average aftershock rate via the probability $Q(m_d)$ that a background event is observable and the probability $\nu(m_d)$ that an unobservable background event triggers some first-generation aftershock. In the critical case $n = 1$, expression (24) predicts that the effective branching ratio $n(m_d)$ does not depend on m_d and is equal to $n(m_d) \equiv 1$ as well. This is the consequence of the fact that, for $n = 1$, $\langle \tilde{R} \rangle = \infty$ and, due to equation (22), we have $n(m_d) = 1$. On the other hand, note that for n close to 1 but not exactly equal to 1, for instance $n = 0.95$ shown in Figure 2, the effective branching ratio $n(m_d)$ still strongly depends on m_d.

We stress that the introduction of the effective branching ratio $n(m_d)$ in (23) is solely based on average statistical properties of aftershock numbers. It is not obvious that it also describes the space-time properties of aftershock cascades. As we show in section 3.6, this turns out not to be the case, in particular as the time-dependent Omori law decay of observable events is still controlled by the true branching ratio n and not by the effective one $n(m_d)$.

2.3 Basic properties of the ETAS model

To make further progress and in particular to calculate the probability $q(m_d)$ given by (16) that a cluster is observable and to obtain the effective branching rate $n(m_d)$ given by (24), we need to specify the properties of ETAS branching model. The ETAS model is defined by the following rules. Each event of magnitude m triggers a Poissonian sequence of aftershocks characterized by the Poissonian GPF [35]

$$G_1(x; m|\kappa) = e^{\kappa\mu(m)(z-1)}, \tag{25}$$

where $\kappa\mu(m)$ is the average number of first generation aftershocks triggered by a mainshock of magnitude m, κ is a numerical constant and $\mu(m)$ describes the so-called productivity law, i.e., the dependence of the number of first generation aftershocks number on the mainshock magnitude m. Previous empirical studies have established that the productivity law is approximately exponential [28,47]:

$$\mu(m) = 10^{\alpha(m-m_0)}. \tag{26}$$

with an exponent α in the range 0.8–1. Here, m_0 is the lower magnitude threshold below which events are supposed not to be able to trigger any aftershock. The ETAS model also uses the well-known Gutenberg-Richter (GR) law for the PDF of earthquake magnitudes

$$P(m) = b\ln(10)10^{-b(m-m_0)}\ , \quad \text{with} \quad b \approx 1, \tag{27}$$

which are assumed to be independently drawn at each event occurrence. Averaging the GPF defined by (25) over all possible random source magnitudes m weighted by the GR distribution (27), we obtain the GPF of first generation aftershock numbers triggered by an arbitrary source:

$$G_1(z|\kappa) = F[\kappa(1-z)], \tag{28}$$

where

$$F(x) = \gamma \int_1^\infty e^{-\kappa xy} \frac{dy}{y^{\gamma+1}} = \gamma y^\gamma \Gamma(-\gamma, y)\ , \qquad \gamma = \frac{b}{\alpha}, \tag{29}$$

and $\Gamma(-\gamma, y)$ is the incomplete Gamma function. In the sequel, we shall use the following power law expansion of the function $F(x)$

$$F(x) \simeq 1 - \frac{\gamma}{\gamma-1}\ x + \beta\ x^\gamma,$$
$$\beta = \gamma\Gamma(-\gamma) = \frac{\Gamma(2-\gamma)}{\gamma-1}. \tag{30}$$

Thus, for $\gamma \to 1+$, both coefficients of x and x^γ grow together.

Our previous calculations have shown that this expansion is very accurate for $\gamma \leq 1.25$, which is the relevant range [35–38].

This expansion (30) allows us to express the main properties of the statistics of the number of aftershocks. For this, let us substitute (30) into (28) to obtain the corresponding approximate expression for the GPF of the number of first generation aftershocks

$$G_1(z|\kappa) \simeq 1 - n(1-z) + \beta\kappa^\gamma(1-z)^\gamma, \tag{31}$$

where

$$n = \langle R_1 \rangle = \frac{\gamma\kappa}{\gamma-1} \tag{32}$$

is the average aftershock branching ratio, i.e., the average $\langle R_1 \rangle$ of the total number of first generation aftershocks triggered by an arbitrary source. Recall that the last term $\beta\kappa^\gamma(1-z)^\gamma$ in the r.h.s. of expression (31) expresses the property that the distribution $\mathcal{P}_1(r|\kappa)$ of the total number of first generation aftershocks triggered by an arbitrary source has a power law tail

$$\mathcal{P}_1(r|\kappa) \simeq \frac{\gamma\kappa^\gamma}{r^{1+\gamma}}. \tag{33}$$

Expression (33) is the leading asymptotic of the exact expression

$$\mathcal{P}_1(r|\kappa) = \frac{1}{r!}\frac{d^r G_1(z|\kappa)}{dz^r}\Big|_{z=0} = \gamma\frac{\kappa^\gamma}{r!}\Gamma(r-\gamma,\kappa), \tag{34}$$

corresponding to the exact GPF (28) of the number of first generation aftershocks. Reference [35] has shown that the power law (33) leads to a PDF of the total number of aftershocks of all generations which are triggered by an arbitrary source, which has a fatter tail $\sim 1/r^{1+(1/\gamma)}$, close to criticality $n \approx 1$.

2.4 Observable and unobservable aftershocks

Let us now consider a different averaged GPF (25) obtained by using a truncated GR law constrained to unobservable earthquakes (with magnitudes m between m_0 and m_d):

$$P^-(m|m_d) = \frac{b\ln(10)10^{-b(m-m_0)}}{1-Q(m_d)} H(m_d - m)H(m - m_0), \tag{35}$$

where $H(x)$ is the unit step (Heaviside) function and $Q(m_d) = 10^{-b(m_d - m_0)}$ according to (19) and (27) is the probability for an earthquake to be observable. Averaging expression (25) over all magnitudes weighted by $P^-(m|m_d)$ given by (35) yields the GPF of the number of first-generation aftershocks triggered by an unobservable event:

$$G_1^-(z|\kappa, m_d) = \frac{F[\kappa(1-z)] - Q(m_d)F[\kappa\mu(m_d)(1-z)]}{1 - Q(m_d)}. \tag{36}$$

Substituting the expansion (30) in (36) yields finally

$$G_1^-(z|\kappa, m_d) \simeq 1 - \nu(m_d)(1-z) + \mathcal{O}[(1-z)^2], \tag{37}$$

where the coefficient $\nu(m_d)$ appears here from its definition as the probability that an unobservable background event triggers some aftershock. The expansion (37) at this linear order for the GPF of first-generation aftershocks triggered by unobservable event is equivalent to saying that an unobservable event can trigger at most a single aftershock, in agreement with the approximation used to obtain (15) and (16).

Expressions (30) and (36) thus yield

$$\nu(m_d) = n\frac{1 - \rho(m_d)}{1 - Q(m_d)}, \tag{38}$$

where

$$\rho(m_d) = Q(m_d)\mu(m_d) = [\mu(m_d)]^{1-\gamma} \tag{39}$$

describes the competition between the GR and productivity laws at the observational magnitude threshold m_d. Multiplying (38) by the fraction $1 - Q(m_d)$ of unobservable sources yields the average number $\langle R_1^- \rangle$ of first generation aftershocks triggered by an unobservable source. $\langle R_1^- \rangle$ can be interpreted as the branching rate $n^-(m_d)$ of first-generation aftershocks triggered by unobservable sources:

$$\langle R_1^- \rangle(m_d) = n^-(m_d) = n[1 - \rho(m_d)]. \tag{40}$$

Note that the GPF (37) does not contain a term of the form $\sim(1-z)^\gamma$ as in (31), which was responsible for the power law tail (33) of the PDF of the total number of first generation aftershocks. As a consequence, the tail of the PDF of first-generation aftershocks triggered by unobservable sources is thinner than a power law. The power law tail (33) is simply due to the interplay between the productivity law (26) and the GR law (27) for the sources. Now, constraining the source magnitudes to be smaller than m_d truncates the GR law and thus the PDF of the number of first-generation events.

Let us now turn to the statistics of first-generation aftershocks triggered by observable sources. The corresponding GPF is obtained by averaging (25) over all magnitudes weighted by the following modified GR law:

$$P^+(m; m_d) = \frac{P(m)}{Q(m_d)}H(m - m_d) = b\ln(10)10^{-b(m-m_d)}H(m - m_d). \tag{41}$$

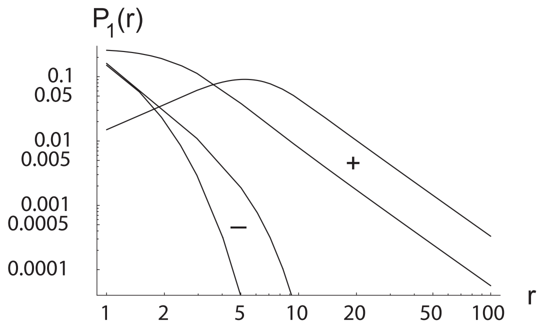

Fig. 1. Dependence of the PDF's $\mathcal{P}_1^+(r|\kappa, m_d)$ and $\mathcal{P}_1^-(r|\kappa, m_d)$ as a function of number r for $\gamma = 1.1$, $n = 0.9$ and for $\mu(m_d) = 10$ and 50, illustrating the presence of a power law tail $\sim r^{-\gamma-1}$ for first-generation aftershocks triggered by observable sources and of fast decaying tails for first-generation aftershocks triggered by unobservable sources (of magnitude less than m_d).

This leads to

$$G_1^+(z|\kappa, m_d) = F[\kappa\mu(m_d)(1-z)]. \tag{42}$$

Note that expression (42) differs from the GPF (28) of the total number of first-generation events only through the renormalization

$$\kappa \to \kappa(m_d) = \kappa\mu(m_d). \tag{43}$$

This allows us to interpret the average number of first-generation aftershocks triggered by an arbitrary observable source as an effective branching rate $n^+(m_d)$ equal to

$$\langle R_1^+ \rangle = n^+(m_d) = n\kappa(m_d)Q(m_d) = n\rho(m_d), \tag{44}$$

where $\rho(m_d)$ is defined by (39). Not surprisingly, the PDF of the number of first-generation aftershocks triggered by observable background events has a power law tail,

$$\mathcal{P}_1^+(r|\kappa, m_d) = \frac{1}{r!}\frac{d^r G_1^+(z|\kappa, m_d)}{dz^r}\Big|_{z=0} \simeq \gamma\frac{\kappa^\gamma}{r^{1+\gamma}}, \tag{45}$$

analogous to (33).

These different results are summarized in Figure 1 which shows the PDF's $\mathcal{P}_1^+(r|\kappa, m_d)$ and $\mathcal{P}_1^-(r|\kappa, m_d)$ as a function of the number r of events obtained from the exact relation (34), for two different values of $\mu(m_d)$.

2.5 Properties of effective aftershock rates

We are now armed to discuss in the framework of the ETAS model the properties of the probability $q(m_d)$ given by (16) for a cluster to be observable and the corresponding expression (24) for the effective branching rate $n(m_d)$ of observable clusters. Note again that the expansion (37) at this linear order writes that an unobservable event can trigger at most a single aftershock, in agreement with the approximation used to obtain (15) and (16). This entitles us to substitute (38) into (16) and (24) to obtain

$$q(m_d) = \frac{Q(m_d)}{1 - n[1 - \rho(m_d)]}, \quad (46)$$

and

$$n(m_d) = \frac{n\rho(m_d)}{1 - n[1 - \rho(m_d)]}, \quad (47)$$

where $\rho(m_d)$ is defined by (39). In the following sections, these two relations (46) and (47) will be derived from the exact equations obeyed by the GPF's of the number of aftershocks over all generations. In the mean time, let us discuss their properties and seismological implications.

Using the notations (40) and (44), we can rewrite the effective rate (47) of observable clusters in the form

$$n(m_d) = \frac{n^+(m_d)}{1 - n^-(m_d)}, \quad (48)$$

and interpret it as the rate $n^+(m_d)$ of aftershocks triggered by observable sources, amplified by the impact of aftershocks triggered by unobservable sources since the denominator in (46) and (47) describes the influence of aftershocks triggered by unobservable sources.

First, notice that, in the critical case $n = 1$, we have $n(m_d) \equiv 1$. Thus, the critical regime for all events is also critical for observable events. In this case, the probability $q(m_d)$ that a cluster is observable is given by

$$q(m_d) = \mu^{-1}(m_d) \qquad (n = 1), \quad (49)$$

and decreases as the observation threshold m_d increases, which parallels the intensity of effective observable sources given by (11).

Two cases are worth discussing. For

$$n^-(m_d) = n[1 - \rho(m_d)] \ll 1 , \quad (50)$$

which corresponds to a negligible productivity of unobservable events, then the impact of unobserved sources is small and

$$q(m_d) \simeq Q(m_d), \qquad n(m_d) \simeq n^+(m_d) = n\rho(m_d), \quad (51)$$

as if all aftershocks, which are triggered by observable sources, were observable.

In contrast, for

$$\rho(m_d) \ll 1, \quad (52)$$

we have

$$q(m_d) \simeq \frac{Q(m_d)}{1 - n}, \qquad n(m_d) \simeq \frac{n\rho(m_d)}{1 - n}, \quad (53)$$

where the factor $1/(1 - n)$, quantifying the impact of clusters triggered by unobservable background events, becomes predominant.

Expression (47) can be rewritten as

$$n(m_d) = \frac{1}{1 + \frac{1-n}{n}\,[\mu(m_d)]^{\gamma - 1}}. \quad (54)$$

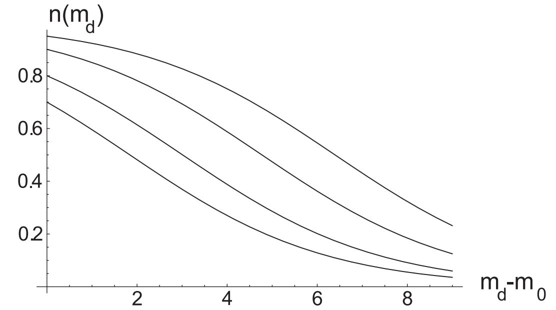

Fig. 2. Dependence of the effective rate $n(m_d)$ given by (47) and (54) for $\alpha = 0.8$ and $b = 1$ ($\gamma = 1.25$) for different value of n: $n = 0.7; 0.8; 0.9; 0.95$ from bottom to top.

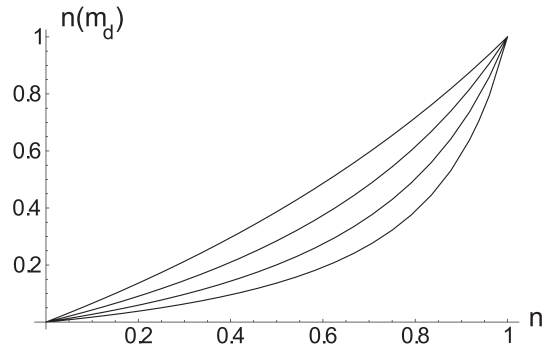

Fig. 3. Dependence of effective rate $n(m_d)$ given by (47) and (54) as a function of the branching ratio n of all first-generation events, for $\alpha = 0.8$ and $b = 1$ and several values of $m_d - m_0$: $m_d - m_0 = 1; 2; 3; 4$ from bottom to top.

Thus, for

$$m_d - m_0 \ll \Delta^* \equiv \frac{1}{b - \alpha} \log_{10}\left(\frac{n}{1 - n}\right), \quad (55)$$

the effective aftershocks rate is critical: $n(m_d) \simeq n$. For example, if $n = 0.9$, $b = 1$ and $\alpha = 0.8$ we have $\Delta^* \simeq 4.77$. Figure 2 (respectively 3) shows the dependence of the effective rate $n(m_d)$ as a function of $m_d - m_0$ (respectively n) for various n (respectively $m_d - m_0$).

2.6 Correspondence between the present formalism and Sornette-Werner representation [51]

At this point, the astute reader will have noticed that the expression (47) for the effective rate of observable events of first-generation is not the same as expression (10) of [51], which also gives an apparent branching ratio denoted n_a for observable aftershocks of the first generation. Our present form (47) for $n(m_d)$ departs from expression (10) of [51] for n_a via the denominator, that is, by the fact that $n^-(m_d)$ defined in (50) is not zero. The two approaches are actually equivalent as we now explain. Expression (10) of [51] defines an apparent branching rate n_a as only due to observable sources while $n(m_d)$ given by (47) takes also into account the unobservable sources on observable aftershocks. In other words, $n(m_d)$ given by (47) counts the effect of unobserved sources in the production of first generation events and thus describes the average number of first-generation daughters from unobservable aftershocks which are themselves "sources" for the future generations. In contrast, Sornette and Werner construct a representation in which the introduction of the observational cut-off $m_d > m_0$ not only renormalizes n into n_a given by their equation (10) but also introduces a renormalization of the spontaneous source rate [51]: for each real observable spontaneous sources, there are many apparent spontaneous sources which result from the fact that an event triggered by an unobservable previous aftershock is considered a spontaneous source since one can not track its ancestor. The two approaches can thus be summarized as follows:

– Sornette–Werner:

$$\{n, 1 \text{ spontaneous source}\} \rightarrow \{n_a = n\rho(m_d),\ S_a = N_{\text{obs}}(n - n_a) \text{ spontaneous sources}\}\ , \quad (56)$$

where N_{obs} is the total number of observed aftershocks;

– present work:

$$\{n, 1 \text{ spontaneous source}\} \rightarrow \{n(m_d) = \frac{n\rho(m_d)}{1 - n[1 - \rho(m_d)]},\ 1 \text{ spontaneous source}\}. \quad (57)$$

Note that the fraction $f_a = S_a/N_{\text{obs}}$ of apparent sources among all observed events given by expression (23) of [51] can be written

$$f_a = S_a/N_{\text{obs}} = n - n_a = n - n\rho(m_d) = n^-(m_d), \quad (58)$$

where the last equality results from definition (50). This provides a physically intuitive interpretation of $n^-(m_d)$. Expression (47) can thus be written

$$n(m_d) = \frac{n_a}{1 - f_a} = n_a(1 + f_a + f_a^2 + ...). \quad (59)$$

This formula clarifies completely the relationship between Sornette-Werner's formation and the present one: the first term n_a in the r.h.s. of (59) corresponds to the average number of daughters of first-generation due to an observable initial source; The second term $n_a f_a$ corresponds to the average number of daughters of first-generation which are due to an apparent observable source which is triggered from a first-generation unobservable aftershock of the initial spontaneous source. The third term $n_a f_a^2$ corresponds to the average number of daughters of first-generation which are due to an apparent source which was itself triggered by an apparent source of a first-generation unobserved aftershock of the initial spontaneous source; and so on... This reasoning demonstrates that the two formulations are physically equivalent, even though they have been obtained by different physical arguments.

3 Statistical description of observable events

Until now, we have explored some properties of the fraction $q(m_d)$ of observable clusters and its corresponding effective observable aftershock rate $n(m_d)$, using a physically transparent but non-rigorous approach based on the properties of first-generation aftershocks triggered by observable and unobservable sources. In the following, we study the full statistical properties of observable events in large time window using the GPF's of the number of events of all generations. It is again important to stress that our results apply only to the statistics of clusters and not to the time-dependent properties of aftershock sequences, as shown in Section 3.6.

3.1 Derivation of the GPF of observable events

We start by the remark that the GPF of a single source of magnitude m, which takes into account the observability of the source, is equal to

$$z^{H(m-m_d)} = 1 + (z-1)H(m - m_d) = \begin{cases} 1\ , & m < m_d \\ z\ , & m > m_d. \end{cases} \quad (60)$$

Let us define $G(z; m, m_d)$ as the GPF of the number of observable aftershocks of all generations which are triggered by a source (which can be observable or unobservable). Multiplying $G(z; m, m_d)$ by (60) yields the GPF $\Theta(z; m, m_d)$ of the number of observable events triggered by a source of given magnitude m:

$$\Theta(z; m, m_d) = z^{H(m-m_d)} G(z; m, m_d). \quad (61)$$

Averaging this expression over all possible source magnitudes weighted by the GR law (27) gives the GPF

$$\Theta(z; m_d) = \int_{m_0}^{\infty} \Theta(z; m, m_d) P(m) dm \quad (62)$$

of the total number of all observable events, which include all observable sources and all their observable aftershocks of all generations. $\Theta(z; m_d)$ can be expressed as

$$\Theta(z; m_d) = G(z; m_d|m_0) + (z-1)G(z; m_d|m_d) \quad (63)$$

where

$$G(z; m_d|x) = \int_x^{\infty} G(z; m, m_d) P(m) dm. \quad (64)$$

Thus, determining the GPF $\Theta(z; m_d)$ requires to calculate the GPF $G(z; m, m_d)$ of the number of all observable aftershocks of all generations belonging to the same cluster. The later can be obtained by using the statistical independence of sources and aftershocks magnitudes, which leads to replacing z within the exponential of the r.h.s. of (25) by $\Theta(z; m_d)$, which yields

$$G(z; m, m_d) = e^{\kappa\mu(m)[\Theta(z; m_d) - 1]}. \quad (65)$$

Substituting (65) and (27) into (64) yields

$$G(z; m_d|x) = Q(x) F(\kappa\mu(x)[1 - \Theta(z; m_d)]). \quad (66)$$

Using this expression (66) to express the terms in the r.h.s. of (63) leads to the following equation determining the sought GPF $\Theta(z;m_d)$:

$$\Theta(z;m_d) = F(\kappa[1-\Theta(z;m_d)]) \\ + (z-1)Q(m_d)F(\kappa\mu(m_d)[1-\Theta(z;m_d)]). \quad (67)$$

For $m_d = m_0$ ($Q = \mu = 1$) such that all events can be observed, this equation reduces to the standard functional equation

$$\Theta(z|\kappa) = zF(\kappa[1-\Theta(z|\kappa)]) = zG_1(\Theta(z|\kappa)|\kappa), \quad (68)$$

where $G_1(z|\kappa)$ is the GPF given by (28) of the number of first-generation aftershocks while $\Theta(z|\kappa) = \Theta(z;m_0)$ is the GPF of the total number of event in a cluster. We make explicit the dependence on the parameter κ because it is going to play a crucial role in the following discussion.

There is a physically natural partition of the GPF $\Theta(z;m_d)$ given by (67) according to

$$\Theta(z;m_d) = G^-(z;m_d) + zG^+(z;m_d), \quad (69)$$

where

$$G^-(z;m_d) = F(\kappa[1-\Theta(z;m_d)]) \\ - Q(m_d)F(\kappa\mu(m_d)[1-\Theta(z;m_d)]) \quad (70)$$

describes the statistics of observable aftershocks triggered by an unobservable event, while

$$G^+(z;m_d) = Q(m_d)F(\kappa\mu(m_d)[1-\Theta(z;m_d)]) \quad (71)$$

describes the statistics of observable aftershocks triggered by an observable event.

There are a few exact consequences of relations (67)–(71) which can now be obtained. Consider the average number of events over all generations of a given cluster, given by definition by

$$\langle R\rangle(m_d) = \frac{d\Theta(z;m_d)}{dz}\Big|_{z=1}. \quad (72)$$

Using equation (67), it is easy to show that it satisfies the equation

$$\langle R\rangle(m_d) = n\langle R\rangle(m_d) + Q(m_d), \quad (73)$$

whose solution (20) was already obtained from a direct probabilistic argument. By construction, $\langle R\rangle(m_d)$ given by (20) is equal to the sum

$$\langle R\rangle(m_d) = Q(m_d) + \langle R\rangle^+(m_d) + \langle R\rangle^-(m_d) \quad (74)$$

of the contributions of observable events and aftershocks, which are triggered by both observable and unobservable events, with

$$\langle R\rangle^\pm(m_d) = \frac{d\Theta^\pm(z;m_d)}{dz}\Big|_{z=1}. \quad (75)$$

It follows from (70), (71) and (75) that

$$\langle R\rangle^+(m_d) = n\rho(m_d)\langle R\rangle(m_d) = n^+(m_d)\langle R\rangle(m_d), \quad (76)$$

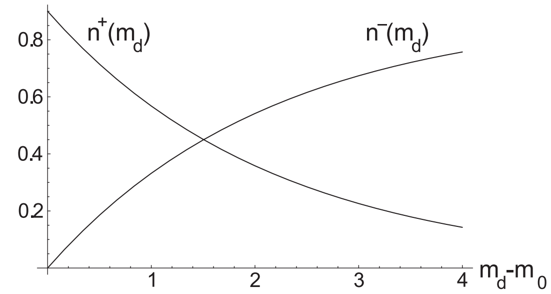

Fig. 4. Dependence of the rates $n^\pm(m_d)$ quantifying the relative impact of aftershocks triggered by observable ($^+$) versus unobservable ($^-$) events, as a function of $m_d - m_0$, for $\alpha = 0.8$, $b = 1$, and $n = 0.9$.

and

$$\langle R\rangle^-(m_d) = n[1-\rho(m_d)]\langle R\rangle(m_d) = n^-(m_d)\langle R\rangle(m_d). \quad (77)$$

These two relations confirm the physical meaning of the rates $n^\pm(m_d)$ defined in (40) and (44), which quantify the relative impact of aftershocks triggered by observable versus unobservable events. Expressions (74), (76) and (77) show that the rates $n^\pm(m_d)$ are the fractions of aftershocks of all generations which are triggered by observable ($^+$) versus unobservable ($^-$) events.

Figure 4 plots these two rates $n^\pm(m_d)$ as a function of $m_d - m_0$ for $\alpha = 0.8$, $b = 1$, $n = 0.9$, showing that the impact of unobserved events may easily dominate for quite reasonable values of the model parameters.

3.2 Fraction of observable clusters

One of the key parameters governing the statistics of windowed observable events is the fraction $q(m_d)$ defined by (6) of observable clusters. Equation (67) allows us to calculate it exactly. Indeed, it is easy to show that expression (67) implies that $q(m_d)$ is solution of the equation

$$q(m_d) = 1 - F[\kappa q(m_d)] + Q(m_d)F[\kappa\mu(m_d)q(m_d)]. \quad (78)$$

Noticing that $Q(m_d) = [\mu(m_d)]^{-\gamma} \equiv \mu^{-\gamma}$, we can rewrite (78) in the form

$$\Psi[q(m_d)] = 0, \quad (79)$$

where

$$\Psi(x) = 1 - x - F(\kappa x) + \mu^{-\gamma}F(\kappa\mu x). \quad (80)$$

A good approximate solution of (79) can be obtained by substituting the polynomial approximation (30) for F to obtain

$$\Psi(x) \simeq \mu^{-\gamma} - x[1 - n(1-\mu^{1-\gamma})]. \quad (81)$$

The corresponding solution of (79) then reads

$$q(m_d) = \frac{1}{n\mu + (1-n)\mu^\gamma}, \quad (82)$$

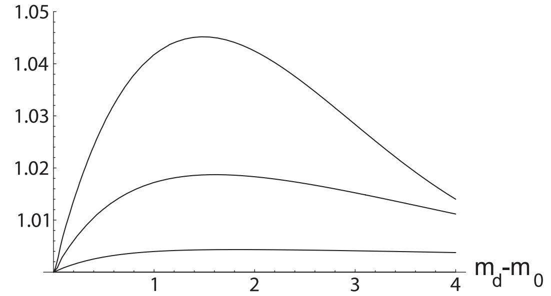

Fig. 5. Dependence of the ratio of the approximation (82) divided by the numerical solution of the exact equation (78), as a function of $m_d - m_0$, demonstrating the good accuracy of the approximate expression (82), for $n = 0.9$, $b = 1$: $\alpha = 0.7; 0.8; 0.9$ from top to bottom.

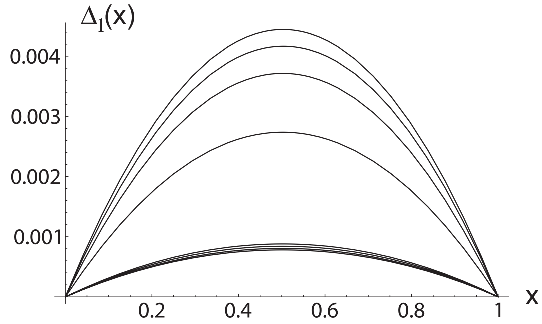

Fig. 6. Dependence of the difference $\Delta_1(x; m_d)$ given by (89) as a function of the variable x, for $n = 0.9, \gamma = 1.25$ and several values of $m_d - m_0 = 1; 2; 3; 4$ (four upper curves from top to bottom). The group of almost undistinguishable curves at the bottom of the graph corresponds to $n = 0.9, m_d - m_0 = 1; 2; 3; 4$ and $\gamma = 1.1$.

which is equivalent to expression (46) derived above using an intuitive nonrigorous reasoning. In contrast, expression (82) and thus (46) is now obtained as an approximate solution of the exact equation (78). The accuracy of this approximation (82) (or (46)) can thus be checked by comparing it with the numerical solution of the exact equation (78). Correlatively, this also directly check the quality of expression (47). Figure 5 shows the ratio of the approximation (82) divided by the numerical solution of the exact equation (78), as a function of $m_d - m_0$ for $n = 0.9, b = 1$ and three values of $\alpha = 0.7, 0.8, 0.9$. One can observe that the quality of the approximation (82) improves as α gets closer to 1.

3.3 Self-similarity of the statistics of observable events

We now have the tools to calculate the conditional GPF $\tilde{\Theta}(z; m_d)$ defined by (8) of the total number of observable events within an observable cluster. Substituting (7) into (67) yields the equation for $\tilde{\Theta}(z; m_d)$:

$$\varphi(\tilde{\Theta}; m_d) = zF(\kappa(m_d)(1 - \tilde{\Theta})), \tag{83}$$

where

$$\kappa(m_d) = \kappa\mu(m_d)q(m_d), \tag{84}$$

and

$$\varphi(x; m_d) = \frac{\Psi(q(m_d)(1 - x))}{Q(m_d)}. \tag{85}$$

Definition (80) and equation (79) imply that the following identities are true

$$\varphi(0; m_d) \equiv 0 \ , \qquad \varphi(1; m_d) \equiv 1. \tag{86}$$

Using (80) and the approximate expression (30), we obtain the linear approximation

$$\varphi(x; m_d) \simeq \varphi_1(x) \ , \qquad \varphi_1(x) = x, \tag{87}$$

which is consistent with (86). Then, substituting (87) into (83) yields an approximate equation for the GPF $\tilde{\Theta}(z; m_d)$:

$$\tilde{\Theta}(z; m_d) \simeq zF(\kappa(m_d)[1 - \tilde{\Theta}(z; m_d)]). \tag{88}$$

We can check the accuracy of the linear approximation (87), and thus its consequence for $\tilde{\Theta}(z; m_d)$ given by (88) by comparing the linear function $\varphi_1(x) = x$ of (87) with the exact one given by (83). Figure 6 shows the difference

$$\Delta_1(x; m_d) = \varphi(x; m_d) - x, \tag{89}$$

where $\varphi(x; m_d)-$ is given by (83), as a function of the variable x, for $n = 0.9, \gamma = 1.25$ and $\gamma = 1.1$ and several values of $m_d - m_0$. This figure confirms the good accuracy of the linear approximation. The two following subsections will extract the consequence of this formulation for the distribution of aftershock numbers and will quantify the impact of the next order correction to the linear approximation (87).

Note that equation (88) coincides, after the application of the renormalization (43) where $\kappa(m_d)$ is now given by expression (84), with the equation (68) for the GPF $\Theta(z|\kappa)$ of the total number of events within an arbitrary cluster. This has an important consequence for the physical understanding of seismicity according to the ETAS model: as long as the linear approximation (87) is applicable, the statistics of the number of observable events within observable clusters is identical, up to the renormalization (43), to the statistics of the total number of events within an arbitrary cluster in which all events can be observed.

This can be restated as the following self-similar property for the statistical properties of observable clusters:

$$\tilde{\Theta}(z; m_d) \simeq \Theta(z|\kappa(m_d)). \tag{90}$$

This self-similarity property means that the statistics of observable windowed events within large time windows is identical after the correspondence

$$\omega \to \omega(m_d) = \omega q(m_d),$$
$$n \to n(m_d) = \frac{\gamma\kappa(m_d)}{\gamma - 1}, \tag{91}$$

to the statistics of the total number of windowed events. The effective branching rate $n(m_d)$ defined in (91) coincides, using the expression (84) for $\kappa(m_d)$ and (82) for

$q(m_d)$, with expression (47) that we have previously obtained for the effective rate of observable aftershocks. This shows again that the intuitive probabilistic reasoning of Section 2.5 on effective aftershock rates is equivalent to the linear approximation (87) for the GPF. As we are going to probe in greater depth, this suggests that the self-similar property (90) is a resilient and general feature of the ETAS model.

3.4 Distribution of the number of observable events

We now derive the consequences of the above results for the distribution of the numbers of observable events.

Let us denote by $\tilde{\mathcal{P}}(r; m_d)$ the probability corresponding to the GPF $\tilde{\Theta}(z; m_d)$ defined by (8), that there are r observable events in a given observable cluster. Similarly, we denote $\mathcal{P}(r; m_d)$ the distribution of the numbers of observable events within an arbitrary cluster corresponding to the GPF $\Theta(z; m_d)$. The two GPF $\tilde{\Theta}(z; m_d)$ and $\Theta(z; m_d)$ are linked through equation (7). It follows from (7) and (88) that, within the domain of application of the linear approximation (87), these two probabilities can be expressed in terms of the probability $\mathcal{P}(r|\kappa)$ of the total number of events of an arbitrary cluster via the following self-similar relations

$$\begin{aligned} &\tilde{\mathcal{P}}(r; m_d) \simeq \mathcal{P}(r|\kappa(m_d)), \\ &\mathcal{P}(r; m_d) \simeq q(m_d)\mathcal{P}(r|\kappa(m_d)), \quad r \geqslant 1. \end{aligned} \tag{92}$$

Thus, the statistical properties of observable events are known from those of all events via the scaling relations (92) (within the linear approximation (87) of the GPF). The self-similar properties (90) and (92) mean that, for the statistical properties, the ETAS model is renormalized onto itself under the transformation $m_0 \to m_d$, with just a renormalization from κ to $\kappa(m_d)$ and, as a consequence, a renormalization of the branching ratio from n to $n(m_d)$. Our present analysis thus confirms for the full statistical properties the results obtained previously, based solely on the average seismic rates [51].

The statistics properties of the total number of events in individual aftershock clusters (without the constraint of observability) has been derived in our previous paper [35,35]. Therefore, we just need to recall briefly some of its key properties which are useful for understanding the statistics of observable events.

Recall that the probability density $\mathcal{P}(r|\kappa)$ is given by the Cauchy integral [38]

$$\mathcal{P}(r|\kappa) = \frac{1}{2\pi i r} \oint_{\mathcal{C}} \frac{d\Theta(z|\kappa)}{z^r}, \tag{93}$$

where $\mathcal{C}$ is sufficiently small contour enveloping the origin $z = 0$, and $\Theta(z|\kappa)$ is solution of the functional equation (68). The main difficulty in the calculation of the integral (93) is that the GPF $\Theta(z|\kappa)$ is defined only implicitly, via the solution of equation (68). To overcome this obstacle, we perform a change of variable and use the new integration variable $y = \Theta(z|\kappa)$. It follows from (68) that

$$z = \frac{y}{G_1(y|\kappa)}, \tag{94}$$

which yields the explicit integral for $\mathcal{P}(r|\kappa)$:

$$\mathcal{P}(r|\kappa) = \frac{1}{2\pi i r} \oint_{\mathcal{C}'} G_1^r(y|\kappa) \frac{dy}{y^r}, \tag{95}$$

where $\mathcal{C}'$ is some small contour in the complex plane y enveloping the origin $y = 0$.

It is interesting to point out that relation (95) has an intuitive probabilistic interpretation, as it can be expressed as

$$\mathcal{P}(r|\kappa) = \frac{1}{r} \Pr(Y_r = r - 1), \tag{96}$$

where

$$Y_s = \sum_{k=1}^{s} U_k, \tag{97}$$

and $\{U_1, U_2, \dots\}$ are mutually independent random integers with GPF $G_1(z|\kappa)$ given by (28) with (31). In the relevant regime for earthquakes for which, probably, $1 < \gamma < 2$, and for $r \gg 1$, the PDF of the sum (97) tends asymptotically to

$$\begin{aligned} \Pr(Y_r = s) &= \frac{1}{(\nu r)^{1/\gamma}} \ell_\gamma\left(\frac{s - nr}{(\nu r)^{1/\gamma}}\right), \\ \nu &= -\kappa^\gamma \Gamma(1-\gamma), \end{aligned} \tag{98}$$

where $\ell_\gamma(x)$ is the stable Lévy distribution such that its two-sided Laplace transform is equal to

$$\int_{-\infty}^{\infty} \ell_\gamma(x) e^{-ux} dx = e^{u^\gamma}. \tag{99}$$

This Lévy distribution has the following properties

$$\ell_\gamma(x) \sim \frac{x^{-\gamma-1}}{\Gamma(-\gamma)} \quad (x \to \infty), \qquad \ell_\gamma(0) = \frac{1}{\gamma \Gamma(1 - 1/\gamma)}. \tag{100}$$

One can calculate $\ell_\gamma(x)$ numerically for any value $1 < \gamma < 2$ using, for instance, the following integral representation

$$\begin{aligned} \ell_\gamma(x) = \frac{1}{\pi} \int_0^\infty &\exp\left[-u^\gamma + ux\cos\left(\frac{\pi}{\gamma}\right)\right] \\ &\times \sin\left[ux\sin\left(\frac{\pi}{\gamma}\right) + \frac{\pi}{\gamma}\right] du. \end{aligned} \tag{101}$$

The asymptotic expression for the probability (96) corresponding to (98) is

$$\mathcal{P}(r|\kappa) \simeq \frac{1}{r(\nu r)^{1/\gamma}} \ell_\gamma\left(\frac{r(1-n) - 1}{(\nu r)^{1/\gamma}}\right) \qquad (r \gg 1). \tag{102}$$

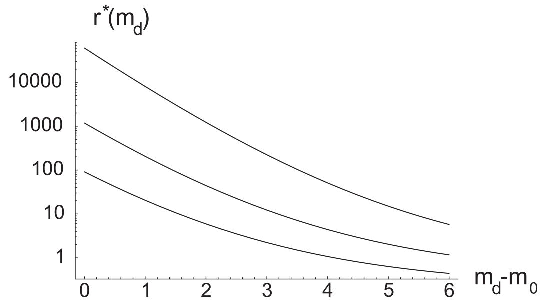

Fig. 7. Dependence of the cross-over value $r^*(m_d)$ separating the two power laws (103) and (104) for the statistics of the number of observable events, for $\gamma = 1.25$ and $n = 0.9; 0.8; 0.7$ (top to bottom).

When the average branching ratio n defined by (18) is close to 1, equation (102) with (100) predict the existence of two characteristic power laws for the probability $\mathcal{P}(r|\kappa)$:

$$\mathcal{P}(r|\kappa) \sim r^{-1-1/\gamma} \qquad (r \ll r^*), \tag{103}$$

and

$$\mathcal{P}(r|\kappa) \sim r^{-1-\gamma} \qquad (r \gg r^*), \tag{104}$$

where

$$r^* = \nu^{1/(\gamma-1)} \left(\frac{1}{1-n}\right)^{\gamma/(\gamma-1)}. \tag{105}$$

The power law (104) reflects the intrinsic distribution of the number of first-generation aftershocks given by relation (33), while the heavier power law tail (103) reflects the effects of cascades over many generations in the branching aftershocks triggering process [35]. See Figure 2 in reference [35] for a visualization of the two power laws (103) and (104) and their cross-over.

Then, substituting in (105) the effective branching rate $n(m_d)$ for observable clusters, we obtain the dependence of the cross-over value $r^*(m_d)$ separating the two power laws (103) and (104) for the statistics of the number of observable events, as a function of the threshold magnitude $m_d - m_0$. Figure 7 shows $r^*(m_d)$ as a function of $m_d - m_0$ for $\gamma = 1.25$ and several values of n. One can observe a fast decrease of $r^*(m_d)$ with $m_d - m_0$, which implies that increasing the observation magnitude threshold m_d amounts to deviate more and more from criticality, as shown also directly in Figure 2.

3.5 Deviations from self-similarity

All the results above on the self-similarity of the statistics of observable events expressed by relations (90) and (92) can be viewed as the consequence of the linear approximation (87). It is thus important to explore how strong can be the deviations from self-similarity resulting from the properties of the exact equation (83) for the GPF $\tilde{\Theta}$ of the number of observable events within an observable cluster. In this goal, we rewrite equation (83) in the form

$$\tilde{\Theta} = z\tilde{G}_1(\tilde{\Theta}; m_d), \tag{106}$$

where

$$\tilde{G}_1(z; m_d) = G_1(z|\kappa(m_d)g(z; m_d), \tag{107}$$

and

$$g(z; m_d) = \frac{z}{\varphi(z; m_d)}. \tag{108}$$

One can interpret (106) and (107) as describing some branching process such that the GPF of the number of first-generation aftershocks is equal to $\tilde{G}_1(z; m_d)$. In other words, the random number $R_1(m_d)$ of first-generation aftershocks in this new branching model is equal to the sum of the two statistically independent random integers

$$R_1(m_d) = U(m_d) + V(m_d), \tag{109}$$

where $U(m_d)$ has the self-similar GPF $G_1(z|\kappa(m_d)$, while $V(m_d)$ has the GPF $g(z; m_d)$.

Using (95), (96) and (107), we immediately obtain the exact integral representation of the PDF of the number of observable events in an observable cluster:

$$\mathcal{P}(r; m_d) = \frac{1}{2\pi i r} \oint_{\mathcal{C}'} G_1^r[z|\kappa(m_d)] g^r(y; m_d) \frac{dy}{y^r}. \tag{110}$$

Its corresponding probabilistic representation reads

$$\tilde{\mathcal{P}}(r; m_d) = \frac{1}{r} \Pr(Y_r = r-1), \tag{111}$$

where

$$Y_s = \sum_{k=1}^{s} (U_k + V_k). \tag{112}$$

The random integers $\{U_1, U_2, \dots\}$ have the GPF $G_1[z|\kappa(m_d)]$, while the random numbers $\{V_1, V_2, \dots\}$ are random integers which are mutually statistically independent from each other (and from U's) with the GPF $g(y; m_d)$ given by (108). As the transformation $m_0 \to m_d$ is equivalent to a decimation step in the language of the renormalization group (see [54] and [53] as well as Chap. 11 of [55] for pedagogical introductions), the random variables $\{V_1, V_2, \dots\}$ with their GPF $g(y; m_d)$ correspond to a new relevant direction (branching process different from the time-independent ETAS) in the space of branching processes.

To obtain the statistical properties of the random integers V, consider the quadratic approximation of the function $\varphi(x; m_d)$ defined by (85):

$$\varphi(x; m_d) \simeq \varphi_2(x; m_d),$$
$$\varphi_2(x; m_d) = x + \eta(m_d)x(1-x). \tag{113}$$

This second-order approximation is again consistent with the identities (86). Here,

$$\eta(m_d) = 4\Delta_1(1/2; m_d), \tag{114}$$

where Δ_1 is defined in (89). Figure 8 shows the difference

$$\Delta_2(x; m_d) = \varphi(x; m_d) - \varphi_2(x; m_d), \tag{115}$$

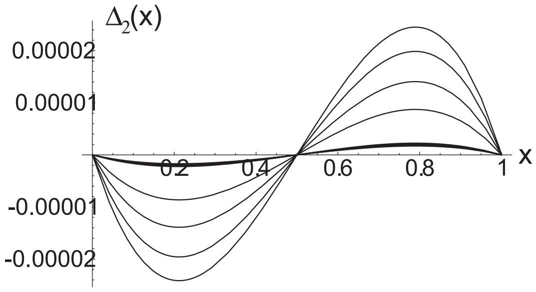

Fig. 8. Dependence of the difference $\Delta_2(x; m_d)$ defined in (115) as a function of x for the same parameters as in Figure 6, demonstrating the high accuracy of the quadratic approximation (113).

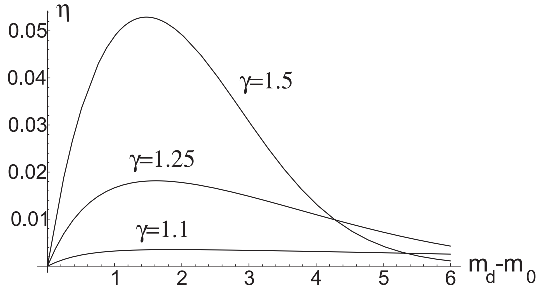

Fig. 9. Dependence of the shift $\eta(m_d)$ to the effective branching rate for observable events as a function of $m_d - m_0$ for different values of γ and for $n = 0.9$.

as a function of x for different values of $m_d - m_0$ for the same parameters as for Figure 6, and for γ 1.1 and 1.25. Comparison between Figures 6 and 8 demonstrate the large improvement from the linear to the quadratic approximation (113).

Substituting (113) into (108) yields the approximate GPF of the auxiliary random integers V as

$$g(z; m_d) = \frac{1}{1 + \eta(m_d)(1-z)}. \tag{116}$$

This GPF means that V has a geometric distribution with the following average and variance

$$\langle V \rangle = \eta \;, \qquad \sigma_V^2 = \eta(\eta + 1). \tag{117}$$

Thus, if $\eta \ll 1$, the random variable V has a small impact on the statistics of the number of observable events. In this case, we obtain the leading asymptotical contribution of the variable V to the statistics of observable events by using a power law expansion for the GPF $\tilde{G}_1(z; m_d)$, similar to (31):

$$\tilde{G}_1(z; m_d) \simeq 1 + \tilde{n}(m_d)(1-z) + \beta\kappa^\gamma(m_d)(1-z)^\gamma, \tag{118}$$

where

$$\tilde{n}(m_d) = n(m_d) + \eta(m_d). \tag{119}$$

Expression (118) shows that the main contribution of the random integer V resulting from the first-order correction to the linear approximation (87) is to introduce a small shift (for $\eta \ll 1$) equal to $\eta(m_d)$ to the effective branching rate $n(m_d)$ obtained within the linear approximation (87). Figure 9 shows the dependence of this shift $\eta(m_d)$ as a function of $m_d - m_0$ for different values of γ. Since $n(m_d)$ is typically in the range 0.5–1, this shows that the corrections are no more than about 10% in the value of the effective branching rate for observable events.

3.6 Un-renormalized branching ratio of the time-dependent properties

We have stressed several times that the renormalization of the branching ratio n into an effective value $n(m_d)$ given by (91) has been demonstrated only for the global statistical properties of the ETAS model. It turns out that the time-dependent properties of a catalog of observable events with magnitudes $m > m_d$ are not described by $n(m_d)$ but by the un-renormalized true branching ratio n. Roughly speaking, because the magnitude of an earthquake is randomly chosen in the ETAS model, observed and unobserved events have the same distribution in time and space. Therefore, the branching parameter n in the ETAS model controls the time decay of the aftershock rate after a mainshock, and the effect of changing the local magnitude cut-off from m_0 to m_d is only to rescale the seismicity rate by a factor $Q(m_d)$ defined in (19), without changing the decay of aftershocks with time. Thus, while the ETAS branching model is renormalized onto itself by changing m_0 into m_d and n into $n(m_d)$ with respect to its statistical properties, the time-dependent properties of observable events with $m > m_d$ are controlled by the un-renormalized branching ratio n, as we now demonstrate explicitly.

We start from the exact equation for the generating probability function (GPF) $\Theta(z; m, m_d, t)$ of the total number until time t of observable aftershocks triggered after a main earthquake that occurred at time 0 which magnitude m. This equation has the form of equation (13) in [56] extended in the time domain:

$$\Theta(z; m, m_d, t) = \exp\Big[\int_{m_0}^{\infty} dm' \int_0^t dt' \lambda(m, m', t')(\Theta(z, m', m_d, t-t') - 1) + (z-1)\int_{m_d}^{\infty} dm' \int_0^t dt' \lambda(m, m', t')\Theta(z, m', m_d, t-t')\Big]. \tag{120}$$

$\lambda(m, m', t')dt'$ is the number of first-generation events of magnitude m' triggered between time t' and $t' + dt'$ by an event of magnitude m. The first (resp. second) double integral in the exponential accounts for unobservable (resp. observable) aftershocks which trigger observable events. The corresponding average number of aftershocks satisfies the linear integral equation

$$\langle R\rangle(m,m_d,t) = \int_{m_0}^{\infty} dm' \int_0^t dt' \lambda(m,m',t') \langle R\rangle(m',m_d,t-t') + \langle R_1\rangle(m,m_d,t) \ , \quad (121)$$

where

$$\langle R_1\rangle(m,m_d,t) = \int_{m_d}^{\infty} dm' \int_0^t dt' \lambda(m,m',t') \quad (122)$$

is the average number of observable aftershocks of first generation triggered till time t. Note that the above equations are very general and hold for any branching process.

The ETAS model is defined by

$$\lambda(m,m',t) = \kappa\mu(m)p(m')\Phi(t), \quad (123)$$

where $\mu(m)$ is the productivity law (26), $p(m)$ is the Gutenberg-Richter (GR) law for the PDF of earthquake magnitudes defined by (27) and $\Phi(t)$ is the bare Omori law (whose expression is not needed here). Using (123) in (121) leads to

$$\langle R\rangle(m,m_d,t) = \kappa\mu(m) \times \int_{m_0}^{\infty} dm' \int_0^t dt' p(m')\Phi(t') \langle R\rangle(m',m_d,t-t') + \langle R_1\rangle(m,m_d,t) \quad (124)$$

and

$$\langle R_1\rangle(m,m_d,t) = \kappa\mu(m)Q(m_d)F(t), \quad (125)$$

where

$$F(t) = \int_0^t \Phi(t')dt' \quad \text{and} \quad Q(m_d) = \int_{m_d}^{\infty} p(m')dm'. \quad (126)$$

The lower limit m_0 in the integrals of equations (121) and (124) means that we take into account the contribution of observable aftershocks which are triggered by unobservable progenitors. For $m < m_d$, $\langle R\rangle(m,m_d,t)$ describes the average number of observable aftershocks, which are triggered by an unobservable mainshock.

To solve equation (124), we apply the Laplace transform to it and obtain the following equation

$$\hat{R}(m,m_d,s) = \kappa\mu(m)\hat{\Phi}(s) \int_{m_0}^{\infty} p(m')\hat{R}(m',m_d,s)dm' + \kappa\mu(m)Q(m_d)\frac{1}{s}\hat{\Phi}(s) \ , \quad (127)$$

where

$$\hat{R}(m,m_d,s) = \int_0^{\infty} e^{-st} \langle R\rangle(m,m_d,t)dt, \qquad \hat{\Phi}(s) = \int_0^{\infty} e^{-st}\Phi(t)dt. \quad (128)$$

Let us introduce the auxiliary function

$$\hat{W}(m_d,s) = \int_{m_0}^{\infty} p(m')\hat{R}(m',m_d,s)dm', \quad (129)$$

which is nothing but the average number of observable aftershocks, which are triggered by a mainshock of random magnitude m. It follows from (127) that $\hat{W}(m_d,s)$ is solution of the algebraic equation

$$\hat{W}(m_d,s) = n\ \hat{\Phi}(s)\hat{W}(m_d,s) + nQ(m_d)\frac{1}{s}\hat{\Phi}(s), \quad (130)$$

where

$$n = \int_0^{\infty} \mu(m)p(m)dm = \frac{\kappa\gamma}{\gamma-1} \quad (131)$$

is the branching ratio previously introduced in (18). The solution of equation (130) is

$$\hat{W}(m_d,s) = Q(m_d)\frac{1}{s}\hat{\Phi}(s)\frac{n}{1-n\hat{\Phi}(s)}. \quad (132)$$

Substituting it into (127) yields

$$\hat{R}(m,m_d,s) = Q(m_d)\frac{1}{s}\hat{\Phi}(s)\frac{\kappa\mu(m)}{1-n\hat{\Phi}(s)}. \quad (133)$$

Thus, $\hat{R}(m,m_d,s)$ differs from $\hat{R}(m,m_0,s)$, which takes into account all aftershocks, only by the factor $Q(m_d)$ defined in (126). In other words, the following relation is true

$$\langle R\rangle(m,m_d,t) \equiv Q(m_d)\langle R\rangle(m,t), \quad (134)$$

where $\langle R\rangle(m,t)$ is the average number of aftershocks with magnitude above m_0, which are triggered by a mainshock of magnitude m. Thus, the time-dependence of $\langle R\rangle(m,m_d,t)$ is the same as that of $\langle R\rangle(m,t)$. Since the time-dependence of $\langle R\rangle(m,t)$ is controlled by the characteristic time scale $t^* \sim 1/(1-n)^{1/\theta}$ (for $\Phi(t) \sim 1/t^{1+\theta}$) [44], we conclude that the un-renormalized true branching ratio n (and not $n(m_d)$) controls the time-dependent properties of the ETAS model.

4 Conclusion

We have shown that, to a good approximation, the statistical properties of catalogs generated by the ETAS model which are restricted to earthquakes above an observable magnitude threshold $m_d > m_0$ are the same as those of catalogs generated with another ETAS model with a renormalized effective branching ratio $n(m_d)$ given in (91). Our present analysis thus confirms, for the full statistical properties, the results obtained previously in reference [51], based solely on the average seismic rates (the first-order moment of the statistics). However, our analysis also demonstrates that this renormalization is not exact, as there are small corrections which can be systematically calculated, in terms of additional contributions that can be mapped onto a different branching model (a new relevant direction in the language of the renormalization group). However, for practical applications, due to the strong stochasticity of the ETAS branching model, these deviations from exact self-similarity will be difficult to observe. In addition, we have shown that the effective parameter $n(m_d)$ cannot be interpreted as a genuine renormalized branching ratio, because observed and unobserved events have the same distribution in time and space, and thus have the same branching parameter n which controls the time decay of aftershock rate after a mainshock. Thus, the ETAS model is not renormalized onto itself by changing m_0 into m_d, and n into $n(m_d)$ with respect to all its space-time properties.

We thank M. Werner for stimulation discussions. This work was partially supported by NSF-EAR02-30429, and by the Southern California Earthquake Center (SCEC) SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 928.

Appendix on the formalism of generating probability functions (GPF)

In this Appendix, we recall the definition of the GPF of some non-negative random integer R (it may be, for instance, the number of earthquakes within some space-time window) and illustrate possible applications of the GPF formalism to explore the statistical properties of branching processes. Let us denote by $P(r)$ the probability that the random number R is equal to some r. Then, by definition, the GPF of the random integer R is equal to the series

$$G(z) = P(0) + P(1)z + P(2)z^2 + \cdots = \sum_{r=0}^{\infty} P(r)z^r. \quad (135)$$

As a first illustration, consider the case where the random integer R is distributed according to Poisson statistics with a mean value $\langle R\rangle = \nu$, such that

$$P(r) = \frac{\nu^r}{r!}e^{-\nu}. \quad (136)$$

Then, the summation of the series in (135) gives

$$G(z) = e^{\nu(z-1)}. \quad (137)$$

One interest of the GPF formalism is that, if one knows the GPF $G(z)$ of some random integer R, one can then calculate the corresponding probabilities $P(r)$ using

$$P(r) = \frac{1}{r!}\frac{d^rG(z)}{dz^r}\bigg|_{z=0}. \quad (138)$$

In some cases, especially for numerical calculations, it is more convenient to use the Cauchy integral formula (which is equivalent to (138))

$$P(r) = \frac{1}{2\pi i}\oint_{\mathcal{C}} G(z)\frac{dz}{z^{r+1}}, \quad (139)$$

where $\mathcal{C}$ is an arbitrary contour which lies inside the circle $|z| \leqslant 1$ in the complex plane z and envelops the origin $z = 0$. From the knowledge of the GPF $G(z)$, one can also obtain easily all the statistical moments of the random integer R. For instance, the average of the random integer R is given by

$$\langle R\rangle \equiv \sum_{r=1}^{\infty} rP(r) = \frac{dG(z)}{dz}\bigg|_{z=1}. \quad (140)$$

The GPF formalism is particular useful to study the statistical properties of branching processes, because it uses optimally the independence between the different branches. Consider the following branching process in which some event triggers a random number R_1 of other (first-generation) events, where the random number R_1 is described by a probability function associated with the GPF $G_1(z)$. Let in turn each first-generation event trigger independently random second-generation events, whose number per first-generation event is also characterized by the same GPF $G_1(z)$. Then, due to the independence of the random numbers of events of first and second generations, the GPF of the total number of events of both first- and second-generation events is equal to

$$G_2(z) = G_1[zG_1(z)]. \quad (141)$$

Let in turn each second-generation event trigger independently random events, whose numbers are again characterized by the same GPF $G_1(z)$, and so on over the infinite range of all possible generations. Then, the GPF $G(z)$ of the total number of events over all generations satisfies the functional equation

$$G(z) = G_1(zG(z)). \quad (142)$$

As an illustration, let us find the average of the total number of triggered events over all generations. Using relation (140) and the normalizing condition $G(z = 1) = 1$, differentiating equation (142) with respect to z, we obtain

$$\langle R\rangle = n(1 + \langle R\rangle) \to \langle R\rangle = \frac{n}{1-n}, \quad (143)$$

where

$$n = \frac{dG_1(z)}{dz}\bigg|_{z=1} \quad (144)$$

is the average number of first generation events. If $n < 1$, then one call the branching process subcritical. If $n = 1$, it is critical and supercritical (explosive) for $n > 1$.

References

1. P. Bak, K. Christensen, L. Danon, T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002)
2. M. Lindman, K. Jonsdottir, R. Roberts, B. Lund, R. Bdvarsson, Phys. Rev. Lett. **94**, 108501 (2005)
3. K.F. Tiampo, J.B. Rundle, W. Klein, J.S.S. Martins, C.D. Ferguson, Phys. Rev. Lett. **91**, 238501 (2003)

4. M.S. Mega, P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda, S. Vinciguerra, Phys. Rev. Lett. **90**, 18850 (2003)
5. A. Helmstetter, D. Sornette, Phys. Rev. Lett. **92**, 129801 (2004)
6. S. Abe, N. Suzuki, Europhys. Lett. **65** (4), 581 (2004)
7. A. Corral, Phys. Rev. E. **6803**(3 Part 2), 5102 (2003)
8. A. Corral, Phys. Rev. Lett. **92**, 108501 (2004)
9. A. Corral, Phys. Rev. E **71**, 017101 (2005)
10. M. Baiesi, M. Paczuski, Phys. Rev. E **69**, 066106 (2004)
11. M. Baiesi, M. Paczuski, Nonlinear Processes in Geophysics, **12**, 1–11 (2005)
12. M. Baiesi, Physica A **359**, 775 (2006)
13. J. Davidsen, M. Paczuski, Phys. Rev. Lett. **94**, 048501 (2005)
14. A. Helmstetter, G. Ouillon, D. Sornette, J. Geophys. Res. **108**, 2483 (2003)
15. D. Marsan, C.J. Bean, Phys. Rev. E **69**, 063101 (2004)
16. A. Helmstetter, D. Sornette, Phys. Rev. E **69**, 063102 (2004)
17. N. Scafetta, B.J. West, Phys. Rev. Lett. **92**, 138501 (2004)
18. X. Yang, S. Du, J. Ma, Phys. Rev. Lett. **92**, 228501 (2004)
19. R. Woodard, D.E. Newman, R. Snchez, B.A. Carreras, Phys. Rev. Lett. **93**, 249801 (2004)
20. P.B. Rundle, J.B. Rundle, K.F. Tiampo, J.S. Sa Martins, S. McGinnis, W. Klein, Phys. Rev. Lett. **87**, 148501 (2001)
21. S. Hergarten, H.J. Neugebauer, Phys. Rev. Lett. **88**, 238501 (2002)
22. A. Helmstetter, D. Sornette, Phys. Rev. E. **66**, 061104 (2002)
23. A. Helmstetter, S. Hergarten, D. Sornette, Phys. Rev. E **70**, 046120 (2004)
24. D. Sornette, G. Ouillon, Phys. Rev. Lett. **94**, 038501 (2005)
25. G. Ouillon, D. Sornette, J. Geophys. Res. **110**, B04306 (2005)
26. L. Knopoff, Y.Y. Kagan, R. Knopoff, Bull. Seism. Soc. Am. **72**, 1663 (1982)
27. T. Utsu, Y. Ogata, S. Matsu'ura, J. Phys. Earth **43**, 1 (1995)
28. A. Helmstetter, Phys. Rev. Lett. **91**, 058501 (2003)
29. D. Sornette, P. Davy, Geophys. Res. Lett. **18**, 1079 (1991)
30. P. Davy, A. Sornette, D. Sornette, Nature **348**, 56 (1990)
31. Y.Y. Kagan, L. Knopoff, Geophys. J. Roy. Astr. Soc. **62**, 303 (1980)
32. Y.Y. Kagan, Nonlin. Proc. Geophys. **1**, 171 (1994)
33. A. Helmstetter, D. Sornette, J. Geophys. Res. (Solid Earth) **108** (B10), 2457 (2003)
34. A. Helmstetter, D. Sornette, Geophys. Res. Lett. **30**, 2069 (2003)
35. A. Saichev, A. Helmstetter, D. Sornette, Pure and Applied Geophysics **162**, 1113 (2005)
36. A. Saichev, D. Sornette, Phys. Rev. E **70**, 046123 (2004)
37. A. Saichev, D. Sornette, Phys. Rev. E **71**, 056127 (2005)
38. A. Saichev, D. Sornette, Eur. J. Phys. B **49**, 377401 (2006)
39. K.R. Felzer, T.W. Becker, R.E. Abercrombie, G. Ekström, J.R. Rice, J. Geophys. Res. **107**, B9, 2190 (2002)
40. J. Zhuang, Y. Ogata, D.Vere-Jones, J. Am. Stat. Assoc. **97**, 369 (2002)
41. J. Zhuang, Y. Ogata, D. Vere-Jones, J. Geophys. Res. **109**, No. B5, B05301 (2004)
42. Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988)
43. Y.Y. Kagan, L. Knopoff, J. Geophys. Res. **86**, 2853 (1981)
44. A. Helmstetter, D. Sornette, J. Geophys. Res. **107** (B10), 2237 (2002)
45. *Classical and modern branching processes*, edited by K.B. Athreya, P. Jagers (Springer, New York, 1997)
46. G. Sankaranarayanan, *Branching processes and its estimation theory* (Wiley, New York, 1989)
47. A. Helmstetter, Y.Y. Kagan, D.D. Jackson, J. Geophys. Res. **110**, B05S08 (2005)
48. R.S. Stein, Scientific American **288**, 72 (2003)
49. S. Toda, R.S. Stein, K. Richards-Dinger, S. Bozkurt, J. Geophys. Res., in press (2005)
50. D. Sornette, M.J. Werner, J. Geophys. Res. **110**, B8, B08304 (2005)
51. D. Sornette, M.J. Werner, J. Geophys. Res. **110**, B9, B09303, 10.1029/2005JB003621 (2005)
52. A. Helmstetter, D. Sornette, Geophys. Res. Lett. **30** (11), (2003)
53. K.G. Wilson, Scientific American **241**, August 158 (1979)
54. K.G. Wilson, J. Kogut, Phys. Rep. **12**, 75 (1974)
55. D. Sornette, *Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, 2nd edn. (Springer Series in Synergetics, Heidelberg, 2004)
56. A. Saichev, D. Sornette, Phys. Rev. E **72**, 056122 (2005)