Eur. Phys. J. B **49**, 377–401 (2006)
DOI: 10.1140/epjb/e2006-00075-3

THE EUROPEAN PHYSICAL JOURNAL B

Power law distribution of seismic rates: theory and data analysis

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 28 September 2005 / Received in final form 1 January 2006
Published online 10 March 2006 –

Abstract. We report an empirical determination of the probability density functions $P_{\rm data}(r)$ (and its cumulative version) of the number r of earthquakes in finite space-time windows for the California catalog, over fixed spatial boxes 5×5 km^2, 20×20 km^2 and 50×50 km^2 and time intervals $\tau = 10,\ 100$ and 1000 days. The data can be represented by asymptotic power law tails together with several cross-overs. These observations are explained by a simple stochastic branching process previously studied by many authors, the ETAS (epidemic-type aftershock sequence) model which 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. We develop the full theory in terms of generating functions for describing the space-time organization of earthquake sequences and develop several approximations to solve the equations. The calibration of the theory to the empirical observations shows that it is essential to augment the ETAS model by taking account of the pre-existing frozen heterogeneity of spontaneous earthquake sources. This seems natural in view of the complex multi-scale nature of fault networks, on which earthquakes nucleate. Our extended theory is able to account for the empirical observation but some discrepancies, especially for the shorter time windows, point to limits of both our theoretical approach and of the ETAS model.

PACS. 64.60.Ak Renormalization-group, fractal, and percolation studies of phase transitions (see also 61.43.Hv Fractals; macroscopic aggregates) – 02.50.Ey Stochastic processes – 91.30.Dk Seismicity (see also 91.45.gd–in geophysics appendix)

1 Introduction

Many papers purport to characterize the space-time organization of seismicity in different regions of the world. Recent claims of universal laws for the distribution of waiting times and seismic rates between earthquakes have derived from the analyses of space-time windows [1,2]. The flurry of interest from physicists comes from their fascination with the self-similar properties exhibited by seismicity (Gutenberg-Richter power law of earthquake seismic moments, Omori decay law of aftershock rates, fractal and multifractal space-time organization of earthquakes and faults) together with the development of novel concepts and techniques that may provide new insights [3–7].

The interest is no less vivid among seismologists and geophysicists in characterizing the space-time properties of seismicity, because it allows them to understand the dynamics of plate motion (at large scales), to constrain the interaction between faults, as well as to develop better hazard assessment. Recently, an additional incentive is provided by the development of forecasting models of seismicity, for instance within the RELM (Regional Earthquake Likelihood Models: `www.relm.org`) project in Southern California. In the RELM project, a forecast is expressed as a vector of earthquake rates specified for each multi-dimensional bin [8], where a bin is defined by an interval of location, time, magnitude and focal mechanism and the resolution of a model corresponds to the bin sizes. Then, expectations and likelihoods can be estimated and used for the comparison between different forecasting methods.

A fundamental issue in testing models' prediction is to take into account so-called aftershock clustering. In one way or another, many models use some form of declustering approach to remove the effect of aftershocks

[a] e-mail: `sornette@moho.ess.ucla.edu`

which otherwise dominate and obscure the desired information about the model's performance [8]. Then, with such declustered catalogs, the likelihood of forecasts are estimated using Poissonian probabilities. But, if the catalog is only partially declustered (which it will most probably be as there are no agreed upon fully efficient method of declustering), then our contribution in this paper is to show that the distribution of event numbers should present a tail much more heavy than predicted by the Poissonian statistics and to propose a theoretical explanation for it. Pisarenko and Golubeva [11] introduced a model to decluster catalogs by so-to-say Poisson "with random parameter," which resulted in a law with slowly decreasing probabilities (actually a stable Lévy law) for the distribution of rates. We note however that several short term forecasts have been developed recently to forecast seismicity rate, which do not remove aftershocks from the catalog in order to compare the model with the data and estimate the model performance [9,10]. Our criticism does not apply to them.

We improve on preceding results on several points. Our first contribution is to show that the heavy tail nature of the distribution of seismic rate is intrinsic to a class of generic models of triggered seismicity. Specifically, our theory is based on a simple model of earthquake triggering, in which future seismicity is a conditional Poisson process, with average rates (or Poisson intensity) conditioned on past seismicity. We show that the exponential Poisson rate is renormalized into a power law tail by the mechanism involving a cascade of earthquake triggering. Our theory thus provides a prediction for the distribution of seismic rates in space-time bins in the form of a power law tail distribution. Our second contribution is to show that our prediction is verified by empirical seismic rates in Southern California over more than two decades. In addition, our theory accounts well for the evolution of the distribution of seismic rates as a function of the time window size from 1 day to 1000 days.

This implies that spontaneous fluctuations of the number of triggered earthquakes in space-time bins may be simply due to the cascades of triggering processes, which lead to dramatic departures from the Poisson model used as one of the building block of standard testing procedures. Accounting for the intrinsic heavy tail nature of the distribution of seismic rate may explain, we believe, many of the contradictions and rejections of models assessed on the basis of Poisson statistics of so-called declustered catalogs. This also suggests the need for fundamentally different earthquake prediction models and testing methods. Our results also offer a simple alternative explanation to so-called universal laws [1,2] in terms of cascades of triggered earthquakes: our proposed framework explains the observed power law distributions of seismic rates from the fundamentals of seismicity characterized by a few exponents.

The organization of the paper is the following. Section 2 presents the epidemic-type aftershock sequence (ETAS) branching model. Section 3 applies the formalism of generating probability functions (GPF) to the problem of calculating the distribution of seismic rates in finite space-time windows. Appendix A offers a general tutorial of the GPF formalism. Section 4 calculates the averages and the distributions of aftershock numbers in finite space-time windows within the ETAS model. Section 5 presents the large time window approximation to obtain explicit solutions of the implicit GPF equations. Section 6 describes the empirical analysis of the seismic rates in Southern California and compares with the theoretical predictions. Section 7 presents additional statistical tests of the theory using statistics conditioned on generation numbers. Section 8 concludes. A list of symbols is offered and two additional Appendices B and C clarify some technicalities associated with derivations in the body of the text.

2 The Epidemic-Type Aftershock Sequence (ETAS) branching model of earthquakes with long memory

We study the general branching process, called the Epidemic-Type Aftershock Sequence (ETAS) model of triggered seismicity, introduced by Ogata in the present form [12] and by Kagan and Knopoff in a slightly different form [13] and whose main statistical properties are reviewed in [14]. For completeness and in order to fix notations, we recall its definition and ingredients used in our analysis that follows. In this model, all earthquakes are treated on the same footing and there is no distinction between foreshocks, mainshocks and aftershocks, other than from retrospective human-made classification. The advantage of the ETAS model is its conceptual simplicity based on three independent well-found empirical laws and its power of explanation of other empirical observations (see for instance [15] and references therein).

The ETAS model belongs to a general class of branching processes [18,19], and has in addition the property that the variance of the number of earthquake progenies triggered in direct lineage from a given mother earthquake is mathematically infinite. Moreover, a long-time (power law) memory of the impact of a mother on her first-generation daughters describes the empirical Omori law for aftershocks. These two ingredients together with the mechanism of cascades of branching have been shown to give rise to subdiffusion [21,22] and to non mean-field behavior in the distribution of the total number of aftershocks per mainshock, in the distribution of the total number of generations before extinctions [23] and in the distribution of the total duration of an aftershock sequence before extinction [24].

In the ETAS model, each earthquake is a potential progenitor or mother, characterized by its conditional average number

$$N_m \equiv \kappa\mu(m) \tag{1}$$

of children (triggered events or aftershocks of first generation), where

$$\mu(m) = 10^{\alpha(m-m_0)} \ , \tag{2}$$

is a mark associated with an earthquake of magnitude $m \geqslant m_0$ (in the language of "marked point processes"),

κ is a constant factor and m_0 is the minimum magnitude of earthquakes capable of triggering other earthquakes. The meaning of the term "conditional average" for N_m is the following: for a given earthquake of magnitude m and therefore of mark $\mu(m)$, the number r of its daughters of first generation are drawn at random according to the Poissonian statistics

$$p_\mu(r) = \frac{N_m^r}{r!}\, e^{-N_m} = \frac{(\kappa\mu)^r}{r!}\, e^{-\kappa\mu}\,. \tag{3}$$

N_m is the expectation of the number of daughters of first generation, conditioned on a fixed magnitude m and mark $\mu(m)$. The expression (2) for $\mu(m)$ is chosen in such a way that it reproduces the empirical dependence of the average number of aftershocks triggered directly by an earthquake of magnitude m (see [25] and references therein). Expression (1) with (2) gives the so-called productivity law of a given mother as a function of its magnitude. The challenge of our present analysis is to understand how the exponential distribution (3) is changed by taking into account all earthquake triggering paths simultaneously and at all possible generations.

The ETAS model is complemented by the Gutenberg-Richter (GR) density distribution of earthquake magnitudes

$$p(m) = b\ \ln(10)\ 10^{-b(m-m_0)}\ , \quad m \geqslant m_0\ , \tag{4}$$

such that $\int_m^\infty p(x)dx$ gives the probability that an earthquake has a magnitude equal to or larger than m. This magnitude distribution $p(m)$ is assumed to be independent of the magnitude of the triggering earthquake, i.e., a large earthquake can be triggered by a smaller one [15,25].

Combining equations (4) and (2) shows that the earthquake marks μ and therefore the conditional average number N_m of daughters of first generation are distributed according to the normalized power law

$$p_\mu(\mu) = \frac{\gamma}{\mu^{1+\gamma}}\ , \quad 1 \leq \mu < +\infty, \qquad \gamma = b/\alpha\ . \tag{5}$$

For earthquakes, $b \approx 1$ and $0.5 < \alpha < 1$ giving $1 < \gamma < 2$ (see [26] for a review of values quoted in the literature and their implications). This range $1 < \gamma < 2$ implies that the mathematical expectation of μ and therefore of N_m (performed over all possible magnitudes) is finite but its variance is infinite (the marginal case $\alpha = 1$ leading to $\gamma = 1$ requires the existence of an upper magnitude cut-off [26]).

For a fixed γ, the coefficient κ then controls the value of the average number n (or branching ratio) of children of first generation per mother:

$$n = \langle N_m\rangle = \kappa\langle\mu\rangle = \kappa\frac{\gamma}{\gamma-1}\ , \tag{6}$$

where the average $\langle N_m\rangle$ is taken over all mothers' magnitudes drawn from the GR law. Recall that the values $n < 1$, $n = 1$ and $n > 1$ correspond respectively to the sub-critical, critical and super-critical regimes.

The next ingredient of the ETAS model consists in the specification of the space-time rate function $N_m\Phi(\boldsymbol{r} - \boldsymbol{r}_i, t-t_i)$ giving the average rate of first generation daughters at time t and position $\boldsymbol{r}$ created by a mother of magnitude $m \geqslant m_0$ occurring at time t_i and position $\boldsymbol{r}_i$:

$$\Phi(\boldsymbol{x}, t) = \Phi(t)\,\phi(\boldsymbol{x})\,. \tag{7}$$

The time propagator $\Phi(t)$ has the Omori law form

$$\Phi(t) = \frac{\theta c^\theta}{(c+t)^{1+\theta}}\ H(t) \tag{8}$$

where $H(t)$ is the Heaviside function, $0 < \theta < 1$, c is a regularizing time scale that ensures that the seismicity rate remains finite close to the mainshock. The time decay rate (8) is called the "direct Omori law" [14,27]. Due to the process of cascades of triggering by which a mother triggers daughters which then trigger their own daughters and so on, the direct Omori law (8) is renormalized into a "dressed" or "renormalized" Omori law [14,27], which is the one observed empirically. The analysis below will retrieve and extend this result.

The space propagator is given by

$$\phi(\boldsymbol{x}) = \frac{\eta\ d^\eta}{2\pi(x^2+d^2)^{(\eta+2)/2}}\,. \tag{9}$$

For our comparison with the empirical data, we shall consider the epicenter position of earthquakes, that is, the 2D-projection on the earth surface of the real 3D-distribution of earthquake hypocenters. Numerical implementations of the theory developed below will thus be done in 2D but it is easy to generalize to 3D if/when the empirical data will be of sufficient quality to warrant it. For the calculations performed below, we assume that $\phi(\boldsymbol{x})$ is decoupled from the distribution of magnitudes, which is not entirely correct since, in reality, d is proportional to the mainshock rupture length $L \sim 10^{m/2}$. As a consequence, the density of aftershocks over a fault area of size less than d does not depend on m in real data. Our theory developed below provides a way to approximately account for this fact. Technically, this decoupling is convenient for the calculations but disappears eventually when we introduce the effective parameter p defined in equations (64), (65), which is the cost to pay for providing an approximate analytical solution (through the large time-window approximation developed in Sect. 5 and the factorization procedure of Sect. 5.1). Our final formulas (113), (115) are expressed in term of this semi-phenomenological parameter p (and not in terms of Eq. (9) nor of the parameter d), which we have used to fit real data and theoretical curves. In a nutshell, this parameter is the fraction (among all aftershocks) of aftershocks, triggered by some inner source and occurring inside the window. Thus, the parameter p takes into account the spatial aftershocks distribution around of mainshock. Its adjustment in our comparison with the data gives eventually (see Sect. 6.2) $p = 0.45; 0.85; 0.92$ for $L = 5; 20; 50$ km, respectively. These values take into account in an effective way the mentioned law $d \sim L \sim 10^{m/2}$.

In the following, we will make use of the Laplace transform of the Omori law

$$\hat{\Phi}(u) = \int_0^\infty \Phi(t)\, e^{-ut}\, dt = \theta\, (cu)^\theta\, e^{cu}\, \Gamma(-\theta, cu) \quad (10)$$

and of its asymptotic behavior

$$\hat{\Phi}^{-1}(u) \sim 1 + \Gamma(1-\theta)(cu)^\theta\,, \qquad cu \ll 1\,. \quad (11)$$

The Fourier transform of the space propagator (9) will also be useful:

$$\tilde{\phi}(\boldsymbol{q}) = \iint_{-\infty}^{\infty} \phi(\boldsymbol{x}) e^{i(\boldsymbol{q}\cdot\boldsymbol{x})}\, d\boldsymbol{x} = 2\left(\frac{dq}{2}\right)^{\eta/2} \frac{K_{\eta/2}(dq)}{\Gamma(\eta/2)}. \quad (12)$$

In particular

$$\tilde{\phi}(\boldsymbol{q}) = e^{-dq} \quad (\eta = 1)\,, \qquad \tilde{\phi}(\boldsymbol{q}) = e^{-dq}(1+dq) \quad (\eta = 3)\,. \quad (13)$$

The last ingredient of the ETAS model is to assume that plate tectonic motion induces spontaneous mother earthquakes, which are not triggered by previous earthquakes, according to a Poissonian point process, such that the average number of spontaneous mother earthquakes per unit time and per unit surface is ϱ. In the ETAS branching model, each such spontaneous mother earthquake then triggers independently its own space-time aftershocks branching process.

It is a well-established facts that, at large scale, earthquakes are preferentially clustered near the plate boundaries while, at smaller scales, earthquakes are found mostly along faults and close to nodes between several faults [28]. It is natural to extend the ETAS model to allow for the heterogeneity of the spontaneous earthquake sources ϱ reflecting the influence of pre-existing fault structures, some rheological heterogeneity and complex spatial stress distributions. We get some guidelines from the distribution of the stress field in heterogeneous media and due to earthquakes [29] which should be close to a Cauchy distribution ($\frac{1}{x^2+\pi^2}$) or probably more generally to a power law [30,31] (see also Chap. 17 of [32]).

The simplest prescription is thus to assume that ϱ is itself random and distributed according to

$$\frac{1}{\langle\varrho\rangle}\, f\left(\frac{\varrho}{\langle\varrho\rangle}\right)\,, \quad (14)$$

where $\langle\varrho\rangle$ is the statistical average of the random space-time Poissonian source intensity ϱ. In the numerical applications below, we shall use the form

$$f_\delta(x) = \frac{\delta+1}{\delta}\left(1+\frac{x}{\delta}\right)^{-2-\delta} \qquad (\delta > 0). \quad (15)$$

Note that the distribution of spontaneous source $f(\rho)$ is not exactly the same as a spatial distribution of background seismicity. The former is a probability density function (PDF) and the later is a field. A field is described by its PDF (corresponding to a the one-point statistics) and its higher-order moments which capture the spatial dependence structure. By using $f(\rho)$, we account for the one-point statistics but not for the higher-order moments of the spatial structure. In other words, we sample over statistics to account for the sample over different boxes with different source efficiencies. Using $f(\rho)$ is a simple way to take into account the fractal space structure of event sources, when one is interested in the statistics of event numbers. If we wanted to calculate the spatial correlation properties of seismicity rates, using $f(\rho)$ would be insufficient. But for a one-point statistics of event numbers in finite space-time windows, a one-point statistics of spontaneous sources seems sufficient.

The value $\delta = 0$ gives the same tail as the Cauchy distribution advocated in [29] for the stress field. We have considered other functions, such as half-Gaussian, exponential, half-Cauchy but none of them give satisfactory fits to the data (see below). The parametrization (15) with $\delta > 0$ allows us to have only a single scale $\langle\varrho\rangle$ controlling the typical fluctuation of the random sources. We have found that only slightly positive values of δ (corresponding to tails a little faster than the Cauchy law) gives reasonable fits to the data (see below). It is interesting to observe that the data on the distribution of seismic rates thus seems to constrain significantly the fractal distribution of seismic sources.

3 Generating Probability Function (GPF) of earthquakes branching process

In this section, we describe the statistical properties of earthquake branching processes using the technology of generating functions. For the theory of GPF, we refer to [16–19]. These books are devoted mostly to the analysis of the statistics of the number of events per generation, and explore the asymptotics of these kinds of statistics for the Markovian case. Our contribution is to extend this theory to the more complex situation of continuous space-time propagators and power law distributions of marks. Appendix A recalls the definition of the GPF of arbitrary non-negative random integer R and illustrate possible applications of the GPF formalism to explore the statistical properties of branching processes.

First, let us recall the GPF of the total number R_1 of the first-generation aftershocks of a mother event of magnitude m can be easily obtained as (see App. A)

$$e^{\mu\kappa(z-1)}\,, \quad (16)$$

using the fact that the rate of first-generation aftershocks is Poissonian according to equation (3). In this expression (16), κ and μ are given by their definition in equations (1) and (2).

Averaging equation (16) over the random parameter μ gives the GPF of the number R_1 of first generation aftershocks triggered by a mother aftershock of arbitrary magnitude

$$G(z) = \gamma\kappa^\gamma(1-z)^\gamma\, \Gamma(-\gamma, \kappa(1-z))\,. \quad (17)$$

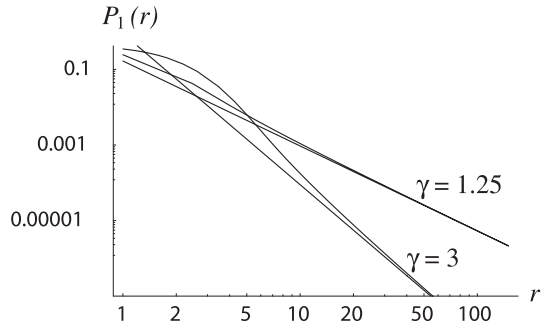

Fig. 1. Plot of the probabilities (19) and their power law asymptotics (20) for the infinite variance case $\gamma = 1.25$ and for a finite variance case $\gamma = 3$.

Note that $\langle R_1 \rangle$ is nothing but the branching ratio n defined above in equation (6). Knowing the expression (17) for the GPF $G(z)$, one finds the corresponding probabilities of the random numbers R_1 by using expression (145) of Appendix A together with the property of incomplete Gamma function (formula (11) of Chap. 9, Sect. 2 of the book [20])

$$\frac{d^n}{dx^n}\left[x^{-\alpha}\Gamma(\alpha,x)\right] = (-1)^n x^{-\alpha-n}\Gamma(\alpha+n,x)\ . \qquad (18)$$

In our case $\alpha = -\gamma$, $n = r$, $x = \kappa(1-z)$ and we calculate the r-th derivation at $z = 0$ ($x = \kappa$) to obtain

$$P_1(r) = \Pr\{R_1 = r\} = \gamma\frac{k^\gamma}{r!}\Gamma(r-\gamma,\kappa)\ , \qquad (19)$$

which have the following asymptotics

$$P_1(r) \simeq \frac{\gamma\kappa^\gamma}{r^{\gamma+1}} = n^\gamma\,\gamma^{1-\gamma}\,(\gamma-1)^\gamma\,r^{-\gamma-1} \qquad (r \gg 1)\,. \qquad (20)$$

Expression (20) implies that, for $1 < \gamma < 2$, the variance of the random number R_1 is infinite. For $\gamma > 2$, the variance is finite and is equal to

$$\sigma_1^2 = \frac{n^2}{\gamma(\gamma-2)} + n\,. \qquad (21)$$

Figure 1 shows the probabilities (19) and their power law asymptotics for $n = 1$, $\gamma = 1.25$ and $\gamma = 3$.

Let us now consider the set of independent space-time aftershock branching processes, triggered by spontaneously arising mother earthquakes. Due to the independence between each sequence triggered by each spontaneous event, it is easy to show that the GPF of the number of events (including mother earthquakes and all their aftershocks of all generations), falling into the space-time window $\{[t, t+\tau] \times \mathcal{S}\}$ is equal to

$$\Theta_{\rm sp}(z,\tau,\mathcal{S}) = e^{-\varrho\,L(z,\tau,\mathcal{S})} \qquad (22)$$

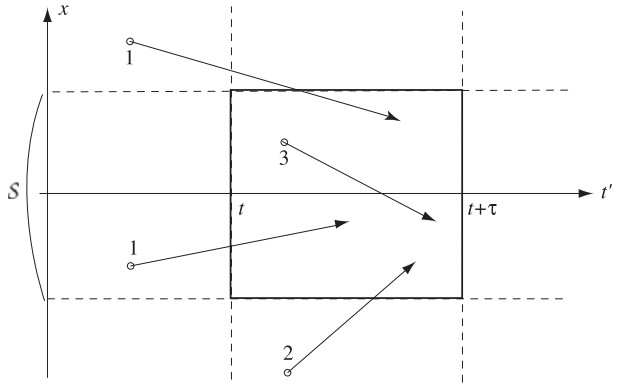

Fig. 2. Illustration of the three different sets of space-time locations for mother earthquakes contributing to the three terms in the r.h.s. of expression (23).

where

$$L(z,\tau,\mathcal{S}) = \int_0^\infty dt \iint\limits_{-\infty}^{\infty} d\boldsymbol{x}\,[1-\Theta(z,t,\tau,\mathcal{S};\boldsymbol{x})]$$
$$+ \int_0^\tau dt \iint\limits_{-\infty}^{\infty} d\boldsymbol{x}\,[1-\Theta(z,t,\mathcal{S};\boldsymbol{x})]\,[1-I_{\mathcal{S}}(\boldsymbol{x})] \qquad (23)$$
$$+ \int_0^\tau dt \iint\limits_{\mathcal{S}} d\boldsymbol{x}\,[1-z\Theta(z,t,\mathcal{S};\boldsymbol{x})]\ .$$

The three above summands have the following transparent geometrical meaning.

- The first summand describes the contribution to the GPF $\Theta_{\rm sp}$ from aftershocks triggered by mother earthquakes that occurred before the time window (i.e. at instants t' such that $t' < t$) (positions 1 in Fig. 2). The corresponding GPF $\Theta(z, t-t', \tau, \mathcal{S}; \boldsymbol{x})$ of the number of aftershocks triggered inside the space-time window $\{[t, t+\tau] \times \mathcal{S}\}$ by some mother event that occurred at time t' satisfies the relation

$$\Theta(z,t,\tau,\mathcal{S};\boldsymbol{x}) = G[1-\Psi(z,t,\tau,\mathcal{S};\boldsymbol{x})]\,, \qquad (24)$$

where the auxiliary function $\Psi(z,t,\tau,\mathcal{S};\boldsymbol{x})$, describing the space-time dissemination of aftershocks triggering by some mother event, is equal to

$$\Psi(z,t,\tau,\mathcal{S};\boldsymbol{x}) =$$
$$\Phi(\boldsymbol{x},t) \otimes [1-\Theta(z,t,\tau,\mathcal{S};\boldsymbol{x})]$$
$$+ \Phi(\boldsymbol{x},t+\tau) \otimes [1-\Theta(z,\tau,\mathcal{S};\boldsymbol{x})]$$
$$+ (1-z)\Phi(\boldsymbol{x},t+\tau) \otimes I_{\mathcal{S}}(\boldsymbol{x})\Theta(z,\tau,\mathcal{S};\boldsymbol{x}). \qquad (25)$$

The symbol $\otimes$ represents the convolution operator. $\Phi(\boldsymbol{x}-\boldsymbol{x}', t')$, which has been defined in equation (7), is the probability density function (PDF) of the position $\boldsymbol{x}'$ and instant t' of some first generation aftershock, triggered by the mother event, arising at the instant $t = 0$ and at the point $\boldsymbol{x}$. The function $I_{\mathcal{S}}(\boldsymbol{x})$ in equations (23) and (25) is the indicator of the space window $\mathcal{S}$ and $G(z)$ in equation (24) is the GPF of the

number R_1 of first generation aftershocks, triggered by some mother earthquake. $G(z)$ given in equation (17) is common to all mother earthquakes and to all aftershocks.

- The last two terms in expression (23) (positions 2 and 3 in Fig. 2) describe the contribution of aftershocks triggered by earthquakes, occurring inside the time window (i.e., $t' \in [t, t+\tau]$). The second term (position 2 in Fig. 2) corresponds to the subset spatially outside the domain $\mathcal{S}$. The third term (position 3 in Fig. 2) corresponds to the subset spatially inside the domain $\mathcal{S}$. These last two terms in expression (23) depend on the GPF

$$\Theta(z,\tau,\mathcal{S};\boldsymbol{x}) = \Theta(z,t=0,\tau,\mathcal{S};\boldsymbol{x}) \tag{26}$$

of the numbers of aftershocks triggered till time τ inside the space window $\mathcal{S}$ by some mother event arising at the instant $t=0$ and at the point $\boldsymbol{x}$. It follows from equations (24) and (25) that it satisfies the relations

$$\Theta(z,\tau,\mathcal{S};\boldsymbol{x}) = G[1-\Psi(z,\tau,\mathcal{S};\boldsymbol{x})] \tag{27}$$

and

$$\Psi(z,\tau,\mathcal{S};\boldsymbol{x}) = \Phi(\boldsymbol{x},\tau) \otimes [1-\Theta(z,\tau,\mathcal{S};\boldsymbol{x})] + (1-z)\Phi(\boldsymbol{x},\tau) \otimes I_{\mathcal{S}}(\boldsymbol{x})\Theta(z,\tau,\mathcal{S};\boldsymbol{x})\,. \tag{28}$$

In addition, we shall need the GPF

$$\Theta(z,\mathcal{S};\boldsymbol{x}) = \Theta(z,\tau=\infty,\mathcal{S};\boldsymbol{x}) \tag{29}$$

of the total numbers of aftershocks triggered by some mother earthquake inside the area $\mathcal{S}$. As seen from equations (27) and (28), it satisfies the relations

$$\Theta(z,\mathcal{S};\boldsymbol{x}) = G[1-\Psi(z,\mathcal{S};\boldsymbol{x})] \tag{30}$$

and

$$\Psi(z,\mathcal{S};\boldsymbol{x}) = 1-\phi(\boldsymbol{x}) \otimes \Theta(z,\mathcal{S};\boldsymbol{x}) + (1-z)\phi(\boldsymbol{x}) \otimes I_{\mathcal{S}}(\boldsymbol{x})\Theta(z,\mathcal{S};\boldsymbol{x})\,. \tag{31}$$

Taking into account the distribution of the source intensities ϱ amounts to averaging equation (22) over ϱ weighted with the statistics (14). This gives

$$\Theta_{\rm sp}(z,\tau;\mathcal{S}) = \hat{f}[\langle\varrho\rangle\, L(z,\tau,\mathcal{S})]\,, \tag{32}$$

where $\hat{f}(u)$ is the Laplace transform of the PDF $f(x)$. For the specification (15), expression (32) becomes

$$\hat{f}_\delta(u) = (1+\delta)(\delta u)^{1+\delta}\, e^{\delta u}\, \Gamma(-1-\delta,\delta u)\,. \tag{33}$$

4 Averages and rates of aftershocks within the space-time window $\{[t,t+\tau]\times\mathcal{S}\}$

Before discussing the properties of the distributions of aftershocks, we consider their simplest statistical characteristics, namely the averages and rates of different kinds of aftershocks. This introduces the relevant characteristic scales in the time and in the space domains, which are found inherent to the space-time branching processes. This also suggests the natural "large time window approximation" used and tested below within the more general probabilistic treatment.

4.1 Average of the total number of events in the space-time window $\{[t,t+\tau]\times\mathcal{S}\}$

Let us first calculate the average of the total number of events inside the space-time window given by

$$\langle R_{\rm sp}(\tau,\mathcal{S})\rangle = \left.\frac{\partial\Theta_{\rm sp}(z,\tau,\mathcal{S})}{\partial z}\right|_{z=1}\,. \tag{34}$$

It follows from equations (32) and (23) that it is equal to

$$\langle R_{\rm sp}(\tau,\mathcal{S})\rangle = \langle R_{\rm out}(\tau,\mathcal{S})\rangle + \langle R(\tau,\mathcal{S})\rangle + \langle\varrho\rangle S\tau\,, \tag{35}$$

where $\langle R_{\rm out}(\tau,\mathcal{S})\rangle$ is the average number of aftershocks triggered by spontaneous "mother" earthquake sources that occurred before time t (positions 1 in Fig. 2), $\langle R(\tau,\mathcal{S})\rangle$ is the average of number of aftershocks triggering by spontaneous earthquake sources that occur within the time interval $[t,t+\tau]$ (positions 2 and 3 in Fig. 2) and $\langle\varrho\rangle S\tau$ is the average of number of spontaneous earthquakes inside the space-time window. Here and everywhere in the following, S is the area of the spatial domain $\mathcal{S}$ and thus $S\tau$ is the space-time volume associated with the space-time window.

In what follows, it will be useful to introduce the rate of events

$$N_{\rm sp}(\tau,\mathcal{S}) = \frac{d\langle R_{\rm sp}(\tau,\mathcal{S})\rangle}{d\tau}\,, \tag{36}$$

with

$$N_{\rm sp}(\tau,\mathcal{S}) = N_{\rm out}(\tau,\mathcal{S}) + N(\tau,\mathcal{S}) + \langle\varrho\rangle S\,. \tag{37}$$

Using equations (24), (25) and equations (27), (28), Appendix B shows that

$$N(\tau,\mathcal{S}) = \frac{n}{1-n}\langle\varrho\rangle\, S - N_{\rm out}(\tau,\mathcal{S}) \tag{38}$$

where $N_{\rm out}(\tau,\mathcal{S})$ satisfies the equation

$$N_{\rm out}(\tau,\mathcal{S}) - n\, N_{\rm out}(\tau,\mathcal{S}) \otimes \Phi(\tau) = \frac{n\,\langle\varrho\rangle S}{1-n} a(\tau)\,, \tag{39}$$

and n is the branching ratio defined in equation (6). Here and below, the following notation is used

$$a(\tau) = \int_\tau^\infty \Phi(t)\, dt = \left(\frac{c}{\tau+c}\right)^\theta\,. \tag{40}$$

Substituting equations (38) into (37) gives the obvious equality

$$N_{\rm sp}(\tau,\mathcal{S}) = \frac{\langle\varrho\rangle S}{1-n}\,, \tag{41}$$

which implies that, due to the cascade of earthquake triggering processes, the average of the total number $\langle R_{\rm sp}(\tau, \mathcal{S})\rangle$ of events is amplified by the factor $1/(1-n)$ compared with the average number $\langle\varrho\rangle\tau S$ of earthquake sources. This factor $1/(1-n)$ has a simple intuitive meaning [33]: one event gives on average n daughters in direct lineage; each of these first-generation daughters give n grand-daughters, the average number of grand-daughters is thus n^2, and the reasoning continues over all generations. Summing over all generations, the total number of events triggered by a given source plus the source itself is $1 + n + n^2 + n^3 + ...$, which sums to $1/(1-n)$. It is clear from this reasoning that formula (41) only holds for $n < 1$. Otherwise, the regime $n > 1$ corresponds to the super-critical case for which there is a non-zero probability $P(r = \infty) > 0$ that number of aftershocks is infinite.

4.2 Impact of mother earthquake sources occurring before the time window

Here, we use previous and other related relations in the goal of estimating the contribution of the different terms in the r.h.s. of equation (23) to the GPF $\Theta_{\rm sp}(z, \tau, \mathcal{S})$. Our goal is to prepare and check for approximations that will be used below.

For instance, we shall assume that the contribution of the first term in equation (23), which is responsible for aftershocks triggered by earthquakes occurring before time t (i.e., outside the time window $[t, t+\tau]$), is negligible if the corresponding relative events rate obeys the following condition

$$\mathcal{N}_{\rm out}(\tau) = \frac{N_{\rm out}(\tau, \mathcal{S})}{N_{\rm sp}(\tau, \mathcal{S})} \ll 1\,. \tag{42}$$

To check when this condition holds, notice that, due to equations (39), (41), $\mathcal{N}_{\rm out}(\tau)$ is solution of

$$\mathcal{N}_{\rm out}(\tau) - n\mathcal{N}_{\rm out}(\tau) \otimes \Phi(\tau) = na(\tau)\,. \tag{43}$$

Applying the Laplace transform to both sides of this equation gives

$$\hat{\mathcal{N}}_{\rm out}(u) = n\frac{1-\hat{\Phi}(u)}{u[1-n\hat{\Phi}(u)]}\,. \tag{44}$$

Using the asymptotic formula (11), we obtain

$$\hat{\mathcal{N}}_{\rm out}(u) \simeq \frac{n\,(c_1 u)^{\theta-1}}{1+(c_1 u)^{\theta}}\,, \tag{45}$$

where

$$c_1 = \left(\frac{\Gamma(1-\theta)}{1-n}\right)^{1/\theta} c \tag{46}$$

is a characteristic time-scale of aftershock branching processes separating a $1/t^{1-\theta}$ law at $t < c_1$ from a $1/t^{1+\theta}$ law at $t > c_1$ for the decay with time of the average aftershock rate triggered by a single mother earthquake [27,14]. The estimation (46), and the more accurate estimation (130) related to it which is developed below, are highly sensitive to the values of θ and n, especially when the former (resp. later) approaches 0 (resp. 1). As a result, changing θ and n "insignificantly" may lead to very different values c_1. For instance, for $c = 10$ s, $n = 0.9$ (resp. $n = 0.96$), and $\theta = 0.2$, then $c_1 \simeq 20$ days (resp. $c_1 \simeq 7$ years). But, for $\theta = 0.1$ [44], we obtain $c_1 \simeq 6000$ years (resp. $c_1 \simeq 56$ million years) for $n = 0.9$ (resp. for $n = 0.96$). For $\theta \geq 0.2$ and $n \leq 0.9$, c_1 is of the order of days or less and the large window approximation used below holds.

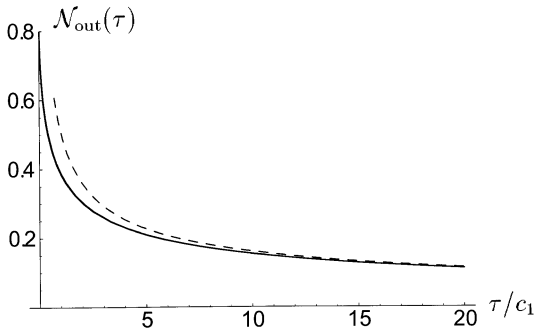

Fig. 3. Plots of the exact rate $\mathcal{N}_{\rm out}(\tau)$, its fractional approximation (47) (which actually coincides with the exact value) and its asymptotic approximation (51) obtained from (49) (dashed line), for $n = 0.9$ and $\theta = 1/2$.

Taking the inverse Laplace transform of equation (45) gives

$$\mathcal{N}_{\rm out}(\tau) = n\,E_\theta\left[-\left(\frac{\tau}{c_1}\right)^{\theta}\right] \tag{47}$$

where $E_\theta(x)$ is the Mittag-Leffler function defined by

$$E_\theta(-x) = \frac{x}{\pi}\sin\pi\theta\int_0^\infty \frac{y^{\theta-1}e^{-y}\,dy}{y^{2\theta}+x^2+2xy^\theta\cos\pi\theta} \qquad (x>0,\ 0<\theta<1)\,. \tag{48}$$

The following asymptotic property holds:

$$E_\theta(-x) \sim \frac{1}{x\ \Gamma(1-\theta)} \qquad (x\to\infty). \tag{49}$$

In addition,

$$E_{1/2}(-x) = e^{x^2}\,\mathrm{erfc}\,x\,. \tag{50}$$

Figure 3 plots the exact rate $\mathcal{N}_{\rm out}(\tau)$, its fractional approximation (47) and corresponding asymptotics derived from equation (49):

$$\mathcal{N}_{\rm out}(\tau) \simeq \frac{n}{\Gamma(1-\theta)}\left(\frac{c_1}{\tau}\right)^{\theta}\,, \tag{51}$$

for $n = 0.9$ and $\theta = 1/2$. One can observe that the asymptotic result (51) is rather precise even if τ is close to c_1. Equation (51) means that, if

$$\tau \gg c_1\ , \tag{52}$$

then one can neglect the contribution of aftershocks triggered by the spontaneous earthquake sources occurring before the time window $[t, t+\tau]$. The remark will be used

in our following investigation and the condition (52) will be refered to as the "large-time window approximation." As discussed before, this approximation will hold typically for $\theta \geq 0.2$ and $n \leq 0.9$, for which c_1 is of the order of days or less. Note that the errors on the knowledge of the parameters θ and n, while difficult to quantify accurately, are probably of the order of 0.1 or more, in absolute value. This is the range over which c_1 changes from smaller than a day to astronomical. It thus seems premature to conclude on the precise value of c_1. Other information is necessary to constrain c_1. In this vein, let us mention that we have also compared our theory with real data, not only for 10–1000 days presented here, but for 1 day and even 0.1 day. The discrepancy between theory and real data (not shown) for the smaller time windows might serve as some indirect argument that $\theta < 0.3$ while $\theta > 0.1$. Moreover, the reasonable agreement of the predictions of our analysis obtained with the help of the large window approximation with our data analysis for time windows of 10 days and larger suggests that c_1 is in the range of a few days.

4.3 Impact of mother earthquake sources occurring inside the space-time window $\{[t, t+\tau] \times \mathcal{S}\}$

Let us calculate the contribution to the rate (37) of events corresponding to the last term of equation (23) (position 3 in Fig. 2). This term describes all the aftershocks triggered by mother earthquakes occurring within the space-time window $\{[t, t+\tau] \times \mathcal{S}\}$. It is easy to show that the corresponding seismic rate, which can be compared with the contribution (41), is equal to

$$\mathcal{N}_{\rm in}(\tau;\mathcal{S}) = 1 - n + \frac{1-n}{S} \iint_{\mathcal{S}} \langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle d\boldsymbol{x}\,, \qquad (53)$$

where

$$\langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle = \left.\frac{\partial \Theta(z,\tau,\mathcal{S})}{\partial z}\right|_{z=1} \qquad (54)$$

is the average number of aftershocks triggered within the space window $\mathcal{S}$ till instant τ by some mother earthquake occurring at position $\boldsymbol{x}$ and at the time $t = 0$. Using relations (27)–(31), one can show that

$$\langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle = \langle R(\mathcal{S};\boldsymbol{x})\rangle - \langle R_+(\tau,\mathcal{S};\boldsymbol{x})\rangle\,, \qquad (55)$$

where $\langle R(\mathcal{S};\boldsymbol{x})\rangle$ is the total number of aftershocks falling inside the space domain $\mathcal{S}$ which are triggered by a earthquake source occurring at position $\boldsymbol{x}$ and at the time $t = 0$. $\langle R_+(\tau,\mathcal{S};\boldsymbol{x})\rangle$ is the corresponding number of aftershocks falling with the space domain $\mathcal{S}$ after the instant τ.

It is easy to show that the Laplace (with respect to τ) and the Fourier (with respect to $\boldsymbol{x}$) transform of $\langle R_+(\tau,\mathcal{S};\boldsymbol{x})\rangle$ is equal to

$$\langle \hat{R}\rangle_+(u,\mathcal{S};\boldsymbol{q}) = n\,\tilde{\phi}_{\mathcal{S}}(\boldsymbol{q})\frac{1}{u}\left(\frac{1}{1-n\tilde{\phi}(\boldsymbol{q})} - \frac{\tilde{\Phi}(u)}{1-n\tilde{\phi}(\boldsymbol{q})\tilde{\Phi}(u)}\right). \qquad (56)$$

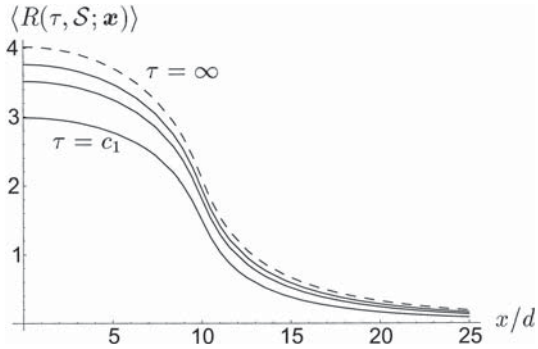

Fig. 4. Plots of $\langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle$ for $\eta = 1$, $\ell = 10\,d$, $\theta = 1/2$ and for $\tau = c_1$; $5\,c_1$; $20\,c_1$. The dashed line is the plot of the average of the total number of aftershocks, falling into the circle $\mathcal{S}$.

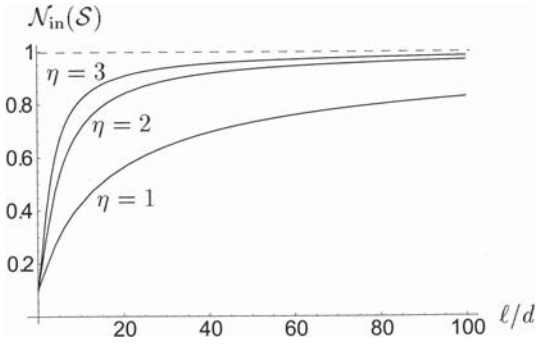

Fig. 5. Plots of the relative rate $\mathcal{N}_{\rm in}(\mathcal{S})$ given by (59) as a function of the dimensionless size ℓ/d of the space domain, for different exponents $\eta = 1$; 2; 3 of the space propagator.

Appendix C shows how to go from equation (56) to

$$\langle \tilde{R}\rangle(\tau,\mathcal{S};\boldsymbol{q}) \simeq \langle \tilde{R}\rangle(\mathcal{S};\boldsymbol{q})\left[1 - E_\theta\left(-\frac{1-n\tilde{\phi}(\boldsymbol{q})}{1-n}\left(\frac{\tau}{c_1}\right)^\theta\right)\right]. \qquad (57)$$

Using expression (57), we construct Figure 4 which plots $\langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle$ for different values of τ/c_1, in order to illustrate its convergence to the average $\langle R(\mathcal{S};\boldsymbol{x})\rangle$ of the total number of aftershocks falling inside the area $\mathcal{S}$.

As can be seen from Figures 3 and 4, it follows from equation (57) and from the properties of Mittag-Leffler functions that, if the large time window approximation (52) holds, one may use the approximate equality

$$\langle R(\tau,\mathcal{S};\boldsymbol{x})\rangle \simeq \langle R(\mathcal{S};\boldsymbol{x})\rangle. \qquad (58)$$

In this large time window approximation, the relative rate (53) is transformed into

$$\mathcal{N}_{\rm in}(\mathcal{S}) \simeq 1 - n + \frac{n(1-n)}{2\pi^2\ell^2}\int_0^\infty \frac{\tilde{\phi}(q)}{1-n\phi(q)}\tilde{I}^2_{\mathcal{S}}(q)\,dq\,. \qquad (59)$$

As the space domain $\mathcal{S}$ increases in size, $\mathcal{N}_{\rm in}(\mathcal{S})$ increases towards 1. Figure 5 plots $\mathcal{N}_{\rm in}(\mathcal{S})$ using the space propagator $\phi(\boldsymbol{x})$ given by equation (9) for different values of the exponent η.

5 Large time window approximation

The analysis of the previous section gives us the possibility to explore the probabilistic properties of the number of events in given space-time windows, in the regime where the large time window approximation (52) holds. If the time duration τ of the space-time window is sufficiently large, the previous section has shown that the statistical averages and the seismic rates become independent of τ. It seems reasonable to conjecture that the GPF $\Theta(z,\tau,\mathcal{S};\boldsymbol{x})$ of the total number of aftershocks triggered by some earthquake source inside the space domain $\mathcal{S}$ until time τ coincides approximately with the saturated GPF $\Theta(z,\mathcal{S};\boldsymbol{x})$ of the total number of aftershocks triggered by some earthquake source inside the space domain $\mathcal{S}$. Within this approximation of large time windows, the effect of aftershocks triggered by earthquake sources occurring till the beginning t of the time window is negligible. Section 7 below will explore in more details the applicability of this conjecture.

Within this large time window approximation, one may ignore the first term in the r.h.s. of equation (23) and replace $\Theta(z,t,\mathcal{S};\boldsymbol{x})$ by $\Theta(z,\mathcal{S};\boldsymbol{x})$ in the remaining terms. As a result, equation (23) takes the following approximate form

$$L(z,\tau,\mathcal{S}) \simeq \tau\iint\limits_{-\infty}^{\infty}[1-\Theta(z,\mathcal{S};\boldsymbol{x})][1-I_{\mathcal{S}}(\boldsymbol{x})]d\boldsymbol{x}+\tau\iint\limits_{\mathcal{S}}[1-z\Theta(z,\mathcal{S};\boldsymbol{x})]d\boldsymbol{x}\,, \tag{60}$$

where $\Theta(z,\mathcal{S};\boldsymbol{x})$ is the solution of equations (30) with (31) or, equivalently, is the solution of

$$\Theta = G\left[\Theta\otimes\phi-(1-z)I_{\mathcal{S}}\Theta\otimes\phi\right]\,, \tag{61}$$

where the function G is defined in equation (17).

5.1 Factorization procedure

To find a reasonable approximate expression for the sought GPF $\Theta(z,\mathcal{S};\boldsymbol{x})$, notice that if $\ell \gg d$ (or if n is close to 1) then the characteristic spatial scale associated with the GPF $\Theta(z,\mathcal{S};\boldsymbol{x})$ becomes greater than d. Therefore, without essential error, one may replace $\Theta\otimes\phi$ by Θ in equation (61). In addition, we take into account the finiteness of the domain $\mathcal{S}$ by using the factorization procedure of replacing the last term of the argument of the function G in equation (61) as follows:

$$I_{\mathcal{S}}(\boldsymbol{x})\Theta(z,\mathcal{S};\boldsymbol{x})\otimes\phi(\boldsymbol{x}) \simeq \Theta(z,\mathcal{S};\boldsymbol{x})\;p_{\mathcal{S}}(\boldsymbol{x})\,, \tag{62}$$

where $p_{\mathcal{S}}(\boldsymbol{x})$ remains to be specified. We will show below that $p_{\mathcal{S}}(\boldsymbol{x})$ may be interpreted as the overall fraction of aftershocks, triggered by a mother earthquake at position $\boldsymbol{x}$, which fall within the domain $\mathcal{S}$. The factorization procedure amounts to replacing a convolution integral by an algebraic term. This factorization approximation is a crucial step of our analysis and will be justified further below. As a result of its use, the nonlinear integral equation (61) transforms into the functional equation

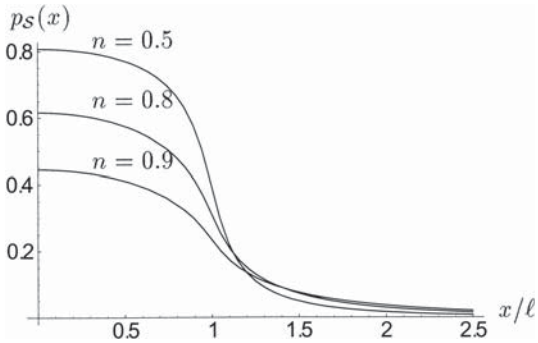

Fig. 6. Plots of the self-consistent factor $p_{\mathcal{S}}(\boldsymbol{x})$ defined by (65) for $\ell = 10\,d$ and for $n = 0.5;\ 0.8;\ 0.9$.

$$\Theta = G[(1+(z-1)p_{\mathcal{S}}(\boldsymbol{x}))\Theta]\,. \tag{63}$$

It is easy to show that, if relation (63) holds, then the average of the number of aftershocks corresponding to it is equal to

$$\langle R\rangle = \frac{n}{1-n}\,p_{\mathcal{S}}\,. \tag{64}$$

In the next subsection, we shall clarify what is the probabilistic sense of the parameter $p_{\mathcal{S}}$. Here, it is sufficient to remark that one can determine it from a consistency condition: choose $p_{\mathcal{S}}(\boldsymbol{x})$ such that the r.h.s. of equation (64) is equal to the true $\langle R(\mathcal{S};\boldsymbol{x})\rangle$. This gives

$$p_{\mathcal{S}}(\boldsymbol{x}) = \frac{1-n}{n}\langle R(\mathcal{S};\boldsymbol{x})\rangle, \quad \tilde{p}_{\mathcal{S}}(\boldsymbol{q}) = \tilde{I}_{\mathcal{S}}(\boldsymbol{q})\tilde{\phi}(\boldsymbol{q})\frac{1-n}{1-n\tilde{\phi}(\boldsymbol{q})}\,. \tag{65}$$

Figure 6 plots $p_{\mathcal{S}}(\boldsymbol{x})$ defined by expression (65) as a function of dimensionless distance x/ℓ.

One can observe that, for $\ell \gg d$, the factor $p_{\mathcal{S}}(\boldsymbol{x})$ approaches a rectangular function. We can use this observation to help determine the statistics of the number of events in a finite space-time window, using the approximation $p_{\mathcal{S}}(\boldsymbol{x}) \simeq$ const. $= p$ for $\boldsymbol{x}\in\mathcal{S}$. We define the parameter p as the space average of $p_{\mathcal{S}}(\boldsymbol{x})$ over the window's area $\mathcal{S}$:

$$p \simeq \frac{1}{S}\iint\limits_{\mathcal{S}} p_{\mathcal{S}}(\boldsymbol{x})d\boldsymbol{x}, \tag{66}$$

p is thus the average over all possible spatial positions of mother earthquakes of the fraction of aftershocks which fall within the space-time window $\mathcal{S}$. The approximation $p_{\mathcal{S}}(\boldsymbol{x}) \simeq \text{const} = p$ for $\boldsymbol{x}\in\mathcal{S}$ allows us to simplify the last term of equation (60) as follows:

$$\tau\iint\limits_{\mathcal{S}}[1-z\Theta(z,\mathcal{S};\boldsymbol{x})]d\boldsymbol{x} \simeq \tau S[1-z\Theta(z;p)]\,, \tag{67}$$

where $\Theta(z,p)$ is the solution of

$$\Theta(z;p) = G[(1+(z-1)p)\Theta(z;p)]\,, \tag{68}$$

which is derived from equation (63).

Complementarily, as can be seen from Figure 6, $p_{\mathcal{S}}(\boldsymbol{x})$ is small outside the window space domain $\mathcal{S}$. It implies that, outside $\mathcal{S}$, one may replace equation (63) by its linearized version. As a result, we get

$$1 - \Theta(z, \mathcal{S}; \boldsymbol{x}) \simeq \frac{n}{1-n}(1-z) p_{\mathcal{S}}(x) \,. \tag{69}$$

Therefore, the first term in the r.h.s. of equation (60) transforms into

$$\tau \iint_{-\infty}^{\infty} [1 - \Theta(z, \mathcal{S}; \boldsymbol{x})][1 - I_{\mathcal{S}}(\boldsymbol{x})] d\boldsymbol{x} \simeq q\tau S \, \frac{n}{1-n}(1-z) \;, \tag{70}$$

where

$$q = \frac{1}{S} \iint_{-\infty}^{\infty} p_{\mathcal{S}}(\boldsymbol{x})[1 - I_{\mathcal{S}}(\boldsymbol{x})] d\boldsymbol{x} \,. \tag{71}$$

Taking into account that, due to equation (65),

$$\iint_{-\infty}^{\infty} p_{\mathcal{S}}(\boldsymbol{x}) d\boldsymbol{x} = S \,, \tag{72}$$

we obtain

$$q \simeq 1 - p \;. \tag{73}$$

Putting all these approximations together allows us to rewrite equation (60) in the form

$$L(z, \tau, \mathcal{S}) \simeq \tau S \left[\frac{n}{1-n}(1-p)(1-z) + 1 - z\Theta(z; p) \right] . \tag{74}$$

In what follows, we shall select a value of the parameter p which takes into account the finiteness of the window's spatial domain $\mathcal{S}$, to better fit empirical data on the statistics of seismic rates in finite space-time bins.

5.2 Probabilistic meaning of the factorization approximation (62) leading to (63) and (65)

The factorization approximation (62) has the following implication. Calling $\Theta(z)$ the GPF of the total number of aftershocks triggered over the whole space by some earthquake source, one can then determine the corresponding GPF $\Theta(z, \mathcal{S}; \boldsymbol{x})$, taking into account approximately the finiteness of the space domain $\mathcal{S}$ and satisfying the functional equation (63) obtained from the factorization procedure, by using the relation

$$\Theta(z, \mathcal{S}; \boldsymbol{x}) = \Theta(q_{\mathcal{S}} + z p_{\mathcal{S}}) \,, \qquad q_{\mathcal{S}}(\boldsymbol{x}) = 1 - p_{\mathcal{S}}(\boldsymbol{x}) \,. \tag{75}$$

This expression (75) has the following interpretation. Let the above mentioned earthquake source triggers r aftershocks. Then, the number of those aftershocks which fall into the space domain $\mathcal{S}$, is equal to

$$R_m(\mathcal{S}; \boldsymbol{x}|r) = X_1 + X_2 + \cdots + X_r \,, \tag{76}$$

where $\{X_1, \ldots, X_r\}$ are mutually independent random variables equal to 1 with probability $p_{\mathcal{S}}$ and 0 with probability $q_{\mathcal{S}} = 1 - p_{\mathcal{S}}$. Thus, $p_{\mathcal{S}}$ is the fraction of the aftershocks which fall into the domain $\mathcal{S}$.

The corresponding expression (75) can be interpreted as follows. It gives the exact solution for the GPF of some specific space-time branching process, such that the PDF $f(\boldsymbol{y}; \boldsymbol{x})$ of the space positions $\boldsymbol{y}$ of each aftershock is the same for all aftershocks and depends only on the position $\boldsymbol{x}$ of the earthquake source. For this problem, we have

$$p_{\mathcal{S}}(\boldsymbol{x}) = \iint_{-\infty}^{\infty} f(\boldsymbol{y}; \boldsymbol{x}) I_{\mathcal{S}}(\boldsymbol{y}) d\boldsymbol{y} \,. \tag{77}$$

In the general case, the relation (75) offers the possibility, at least semi-quantitatively, to describe the characteristic features of the space-time branching processes, by using the probability $p_{\mathcal{S}}$ as an effective independent parameter of the theory.

Let us mention a few useful consequences of the relation (75). It implies that the probability that r aftershocks fall into the spatial domain $\mathcal{S}$ is equal to

$$P(r, \mathcal{S}; \boldsymbol{x}) = \sum_{k=r}^{\infty} P(k) \; B(k, r, \mathcal{S}; \boldsymbol{x}) \,, \tag{78}$$

where $P(k)$ is the probability that some earthquake source triggers k aftershocks and

$$B(k, r, \mathcal{S}; \boldsymbol{x}) = \binom{k}{r} p_{\mathcal{S}}^r \, q_{\mathcal{S}}^{k-r} \,. \tag{79}$$

This binomial probability $B(k, r, \mathcal{S}; \boldsymbol{x})$ is nothing but the conditional probability that, if the mother earthquake triggers $k \geqslant r$ aftershocks then, r of them will fall into the spatial domain $\mathcal{S}$. If $r \gg 1$, expression (79) can be approximated by its well-known Gaussian asymptotics

$$B(k, r, \mathcal{S}; \boldsymbol{x}) = \frac{1}{\sqrt{2\pi k p_{\mathcal{S}} \, q_{\mathcal{S}}}} \exp\left[-\frac{(r - k p_{\mathcal{S}})^2}{2k \, p_{\mathcal{S}} \, q_{\mathcal{S}}} \right] . \tag{80}$$

If, in addition, $P(k)$ decays slowly, for instance if it has a power asymptotic for $k \to \infty$, then expressions (78) and (80) imply the asymptotic relation

$$P(r, \mathcal{S}; \boldsymbol{x}) \sim \frac{1}{p_{\mathcal{S}}} \, P\left(\frac{r}{p_{\mathcal{S}}} \right) \qquad (r \to \infty) \,. \tag{81}$$

In view of this last relation (81), it seems reasonable to assume that the asymptotic behavior of the probabilities of the number of events for $r \gg 1$ are the same for the case of a finite $\mathcal{S}$ ($p_{\mathcal{S}} < 1$) and for an unbounded one ($p_{\mathcal{S}} = 1$).

5.3 Taking into account a magnitude threshold of completeness $m_d > m_0$

Recall that the ETAS model contains an "ultra-violet cutoff m_0, which is the minimum magnitude of earthquakes

capable of triggering other earthquakes (see definition (2) of the productivity). This ultra-violet cut-off is a necessary regularization of any model of triggered seismicity which combines self-similarity and factorization of the Gutenberg-Richter (i.e., the magnitude of triggered events is independent of the magnitude of the mother event). Vere-Jones has recently introduced a new class of models generalizing the ETAS model in which the magnitudes are no more independent, which allows one to get rid of the need of an ultra-violet cut-off through the existence of self-similar dependence relations [34]. Specifically, exact self-similarity (without any magnitude cut-off) requires that the magnitude of aftershocks are dependent upon the magnitude of the triggering earthquake, a property which is not commonly observed in the known phenomenology of seismicity but is not completely excluded and will require new elaborate empirical tests to falsify (see [35] for a theoretical study of Vere-Jones' model and a comparison with ETAS model).

In addition to the physically motivated ultra-violet cut-off m_0 of the ETAS model, seismic catalogs introduce another characteristic magnitude m_d, which is the magnitude of so-called completeness, i.e., roughly speaking, only earthquakes of magnitude larger than m_d can be observed. Since the value of m_d is controlled by observational and instrumental constraints which change with time as the technology and density of seismic stations evolve, m_d has no reason to be the same as the ultra-violet cut-off m_0. The magnitude threshold m_d is approximately 1.5 for the studied catalog (see below Sect. 6). For most modern catalogs, $m_d \approx 2$ and is larger than m_0 (see [26] for estimations of m_0). Until now, our theory has not considered this fact, which amounts to assuming that one can observe all earthquakes of any magnitude above m_0. But since we will compare our predictions with empirical catalogs, it is necessary to calculate the impact of the fact that empirical data only reports seismic rates for event magnitudes $m \geq m_d > m_0$. Sornette and Werner has shown that the impact of $m_d > m_0$ in the ETAS model is to renormalize the branching ratio n and the rate ϱ of spontaneous seismic sources into apparent values n_a and ϱ_a [36]. This result holds with a good approximation for the full statistical properties of aftershock numbers [37].

We now derive the corresponding consequences for our problem. It turns out that a procedure analog to the factorization procedure presented above in Section 5.1, which accounts approximately for the finiteness of space-time window, allows us to calculate exactly the effect of observing only earthquakes above the threshold $m_d > m_0$. Because, in the ETAS model, magnitudes of different events are statistically independent, then the random number of observable events in a given space-time window can be written similarly to (76) as a sum of independent summands, each summand equal to 1 if the corresponding magnitude is larger than m_d, and equal to 0 otherwise. Each event generates an independently random number of observable events, possessing the following GPD

$$D(z,Q) = 1 - Q(m_d)(1-z) \ , \tag{82}$$

where $Q(m_d)$ is the complementary CDF of random magnitudes. With the Gutenberg-Richter law (4), it is given by

$$Q(m_d) = 10^{-b(m_d - m_0)} \ . \tag{83}$$

Note that $Q(m_d) \to 1$ for $m_d \to m_0$ and $D(z,Q) \to z$. Replacing in the GPF $\Theta_{\rm sp}(z,\tau;\mathcal{S})$ given by (32) the argument z by $D(z)$, we obtain the GPF of only observable events:

$$\Theta_{\rm sp}^{\rm obs}(z,\tau;\mathcal{S}) = \hat{f}[\langle\varrho\rangle L(D(z,Q),\tau,\mathcal{S})] \ . \tag{84}$$

This replacement procedure $z \to D(z,Q)$ is based on the property that different branches of triggered earthquakes are independent. Applying this procedure to the previously derived expression (68) and taking into account the equality

$$D(D(z,Q),p) = D(z,Qp) \ , \tag{85}$$

expressing the statistical independence of the two conditions that an event falls into the time-space window and be in the observable range $m > m_d$, leads to

$$L(D(z,Q),\tau,\mathcal{S}) \simeq$$
$$\tau S\left[\frac{n}{1-n}Q(1-p)(1-z)+1-(1-Q(1-z))\Theta(z,Qp)\right]. \tag{86}$$

This expression (86) replaces (74) to account for the constraint that only events of magnitudes m larger than $m_d > m_0$ are observed. Expression (86) reduces to (74) as it should for $Q(m_d) \to 1$ corresponding to $m_d \to m_0$.

Note that the general approach leading to (84) can be used more generally to determine many other properties deriving from event statistics. For instance, for $m_d \simeq 5-6$, then expression (84) describes the statistics of "dangerous" events. Generalizing, if Q in (84) represents the probability that some event falls within the magnitude interval $M \in (m_1, m_2)$, then the r.h.s. of equation (84) describes the statistics of the number of events whose magnitudes belong to this interval. Generalizing further, consider that each event of magnitude m can cause some damage quantified by the function $X(m)$, with characteristic function

$$\varphi(u|m) = \langle e^{iuX(m)}\rangle \ . \tag{87}$$

If we assume that the total damage Y is additive (equal to the sum of the independent damage of each separate event), then the characteristic function of the whole damage

$$C(u,\rho) = \langle e^{iuY}\rangle \tag{88}$$

is equal to

$$C(u) = \hat{f}[\langle\varrho\rangle L(\varphi(u),\tau,\mathcal{S})] \ , \tag{89}$$

where

$$\varphi(u) = \int_{m_0}^{\infty} \varphi(u|m)p(m)dm. \tag{90}$$

5.4 Large space-time windows

For the discussion of this section, we omit for a while the impact of $m_d > m_0$ of the previous section, in order to not burden too much the notations.

In order to get more insight into the properties of the statistics of seismic rates in finite space-time windows, it is useful to study the statistics of seismic rates in the limit where the space and time windows are large. In this case, $\mathcal{N}_{\rm in}$ in (59) and p in (66) are both close to 1 and one may replace equation (74) by

$$L(z,\tau,\mathcal{S}) \simeq \tau S[1 - z\Theta(z)]\,, \tag{91}$$

where $\Theta(z)$ is solution of the functional equation

$$\Theta(z) = G[z\Theta(z)]\,. \tag{92}$$

Accordingly, the GPF of the number of events in a (large) space-time window as given by (32) takes the form

$$\Theta_{\rm sp}(z,\rho) = \hat{f}(\rho\,[1 - z\Theta(z)])\,. \tag{93}$$

Here and everywhere below,

$$\rho = \langle\varrho\rangle\,\tau S\ . \tag{94}$$

Knowing the GPF $\Theta_{\rm sp}(z,\tau;\mathcal{S})$, the probability $P_{\rm sp}(r;\rho)$ of event numbers r is obtained from the formula

$$P_{\rm sp}(r;\rho) = \frac{1}{r!}\,\frac{\partial^r \hat{f}(\rho\,[1 - z\Theta(z)])}{\partial z^r}\bigg|_{z=0}\ . \tag{95}$$

Equivalently, the integral representation of (95) reads

$$P_{\rm sp}(r;\rho) = \frac{1}{2\pi i}\oint_{\mathcal{C}} \hat{f}(\rho\,[1 - z\Theta(z)])\frac{dz}{z^{r+1}}\,, \tag{96}$$

where $\mathcal{C}$ is a sufficiently small contour in the complex plane z around the origin $z = 0$.

The main difficulty in calculating $P_{\rm sp}(r;\rho)$ comes from the fact that the GPF $\Theta(z)$ is defined only implicitly by equation (92). To overcome this difficulty, we rewrite the integral in (96) in the following equivalent form (see App. A for more details on the method)

$$P_{\rm sp}(r;\rho) = \frac{1}{2\pi i r}\oint_{\mathcal{C}} \frac{d\hat{f}(\rho\,[1 - z\Theta(z)])}{z^r} \qquad (r > 0) \tag{97}$$

and use the new integration variable $y = z\Theta(z)$. Expression (92) shows that the inverse function of y is $z = y/G(y)$. As a result, the equation (97) transforms into

$$P_{\rm sp}(r;\rho) = \frac{\rho}{2\pi i r}\oint_{\mathcal{C}'} G^r(y)Q(y;\rho)\frac{dy}{y^r}\,, \tag{98}$$

where

$$Q(z;\rho) = \frac{1}{\rho}\frac{d\hat{f}[\rho\,(1-z)]}{dz} \tag{99}$$

and $\mathcal{C}'$ is a contour enveloping the origin $y = 0$ in the complex plane y. One may interpret $Q(z;\rho)$ in (99) as the GPF of some random integer R_ρ such that $\langle R_\rho\rangle = \rho$.

Notice that equation (98) has a simple probabilistic interpretation. Indeed, it follows from (98) that

$$P_{\rm sp}(r;\rho) = \frac{\rho}{r}\Pr\{R_\rho + R(r) = r - 1\}\,, \tag{100}$$

where R_ρ is a random integer with GPF $Q(z;\rho)$ given by (99) while

$$R(r) = R_1 + R_2 + \cdots + R_r\ , \tag{101}$$

where $\{R_1, R_2, \ldots, R_r, \ldots\}$ are mutually independent random integers with GPF $G(z)$. This implies that the probabilities of each such random variable R_i, $i = 1, ..., r$, has the power law asymptotics (20).

This remark provides a simple analysis of the asymptotic behavior of the probabilities $P_{\rm sp}(r;\rho)$ for $r \gg 1$, by using expression (100). For $1 < \gamma < 2$, the asymptotics of the probability $P(k;r)$ that the sum (101) is equal to k goes to, for large r,

$$P(k|r) \simeq \frac{1}{(\epsilon r)^{1/\gamma}}\psi_\gamma\left(\frac{k - nr}{(\epsilon r)^{1/\gamma}}\right)\,, \tag{102}$$

where

$$\epsilon = -\left(n\frac{\gamma - 1}{\gamma}\right)^\gamma \Gamma(1-\gamma) \tag{103}$$

and $\psi_\gamma(x)$ is the stable Lévy distribution with the two-sides Laplace transform

$$\int_{-\infty}^{\infty}\psi_\delta(x)e^{-sx}\,dx = e^{s^\delta}\,. \tag{104}$$

It is known that

$$\psi_\delta(x) \sim \frac{x^{-\delta-1}}{\Gamma(-\delta)} \qquad (x \to \infty)\,, \qquad \psi_\delta(0) = \frac{1}{\delta\Gamma(1 - \frac{1}{\delta})}\,. \tag{105}$$

One can calculate $\psi_\delta(x)$ for any $1 < \delta < 2$, using, for instance, the following integral representation

$$\psi_\delta(x) = \frac{1}{\pi}\int_0^\infty \exp\left[-u^\delta + ux\cos\left(\frac{\pi}{\delta}\right)\right]\sin\left[ux\sin\left(\frac{\pi}{\delta}\right) + \frac{\pi}{\delta}\right]\,du\,. \tag{106}$$

For some numerical illustrations, we will use the case $\delta = 3/2$ for which the following analytical expression is available

$$\psi_{3/2}(x) = \frac{1}{\pi\sqrt{3}}\left[\Gamma\left(\frac{2}{3}\right){}_1F_1\left(\frac{5}{6},\frac{2}{3},\frac{4x^3}{27}\right) - x\Gamma\left(\tfrac{4}{3}\right){}_1F_1\left(\frac{7}{6},\frac{4}{3},\frac{4x^3}{27}\right)\right]. \tag{107}$$

For $r \gg \rho$, one can neglect the random integer R_ρ in the r.h.s. of equation (100) and obtain from equations (100) and (102) the following asymptotic formula

$$P_{\rm sp}(r;\rho) \simeq \frac{\rho}{r(\epsilon r)^{1/\gamma}} \psi_\gamma \left(\frac{(1-n)r-1}{(\epsilon r)^{1/\gamma}} \right) \qquad (r \gg \rho) \,. \tag{108}$$

If $1-n \ll 1$ (the branching is close to but not exactly critical), equation (108) predicts the existence of two characteristic power laws in the dependence of the probabilities $P_{\rm sp}(r;\rho)$ with r, a result already derived in [23].

1. For

$$1 \ll r \ll r_* \,, \qquad \text{with} \quad r_* = \left(\frac{1}{1-n} \right)^{\gamma/(\gamma-1)} \epsilon^{1/(\gamma-1)} \tag{109}$$

we have

$$P_{\rm sp}(r;\rho) \sim r^{-1-1/\gamma} \,. \tag{110}$$

2. For

$$r \gg r_* \,, \tag{111}$$

we recover the original power law (20) of the number of first generation aftershocks

$$P_{\rm sp}(r;\rho) \sim r^{-1-\gamma} \,. \tag{112}$$

For values of the parameters $\gamma = b/\alpha$ and n which are typical of real seismicity modeled by aftershock triggering processes, the cross-over number r_* separating the two power laws can be very large. For instance, for $\gamma = 1.25$ and $n = 0.9$, we obtain $r_* \simeq 10^4$.

We should mention that the asymptotics (110) and (112) are universal in the sense that they are inherent not only to the statistics of windowed events, but also to statistics of the total number of aftershocks triggered by some mainshocks. These asymptotics were obtained in our previous paper [23] (by using the Lagrange series technique), in the calculation of the total number of aftershocks triggered by some mainshock. The technique based on Cauchy integrals presented here is however more general and offers an efficient method for exploring the probabilistic properties of events statistics in general branching models.

5.5 Prediction of the distribution of event numbers for large time windows

We start with the general expression (32) of the GPF $\Theta_{\rm sp}(z,\tau;\mathcal{S})$ with the approximation (74) for $L(z,\tau,\mathcal{S})$ and (86) for its extension including the effect of the threshold m_d derived in Section 5.3. We use the relationship between the probability $P_{\rm sp}(r;\rho,p)$ and its GPF similar to expression (95) and its integral representation similar to (96). Putting all this together, we obtain the following expression valid in the limit of large time windows and which includes the effect of the threshold m_d:

$$P_{\rm sp}(r;\rho,p) = \frac{1}{2\pi i} \times \oint_{\mathcal{C}} \hat{f}\Big(\rho \Big[\frac{n}{1-n} Q(1-p)(1-z) + 1 - (1-Q(1-z))\Theta(z;Qp) \Big] \Big) \frac{dz}{z^{r+1}} \tag{113}$$

where $\Theta(z;p)$ is the solution of the functional equation (68). Similarly to the change of variable used to go from (97) to (98), we introduce the new integration variable

$$y = (1+(z-1)Qp)\Theta(z;Qp) \iff z = Z(y) = \frac{1}{Qp}\left(\frac{y}{G(y)} + Qp - 1 \right) . \tag{114}$$

By construction of y, $\Theta(z;Qp) = G(y)$ which allows us to obtain the following explicit expression

$$P_{\rm sp}(r;\rho,p) = \frac{1}{2\pi i} \oint_{\mathcal{C}'} \frac{dZ(y)}{dy} \frac{dy}{Z^{r+1}(y)} \times \hat{f}\Big(\rho \Big[\frac{n}{1-n} Q(1-p)(1-Z(y)) + 1 - (1-Q(1-Z(y)))G(y) \Big] \Big) . \tag{115}$$

This expression (115) allows us to make a precise quantitative prediction for the dependence of the distribution $P_{\rm sp}(r;\rho,p)$ of the number r of earthquakes per space-time window as a function of r, once the model parameters $n, \gamma, p, \delta, \rho$ and Q are given. We note that Pisarenko and Golubeva have shown that the distribution of numbers has the same tail as the distribution of seismic rates [11]. Thus, for the tails, our results can be interpreted either as statements on the distribution of realized numbers of earthquakes or on the distribution of average seismic rates.

We now turn to a brief description of the data analysis and of their fits with (115).

6 Empirical analysis and comparison with theory

6.1 Description of the data and procedure

We use the Southern Californian earthquakes catalog with revised magnitudes (available from the Southern California Earthquake Center) as it is among the best one in terms of quality and time span. Magnitudes M_L are given with a resolution of 0.1 from 1932 to 2003, in a region from approximately 32° to 37°N in latitude, and from −114° to −122° in longitude. In order to maximize the size and quality of the data used for the analysis (to improve the statistical significance), we consider the sub-catalog spanning the time interval 1994–2003 for $M_L > 1.5$, which contains a total of 86 228 earthquakes. The completeness of this sub-catalog has been verified in the standard way in [38] by computing the complementary cumulative magnitude distribution for each year from 1994 to 2003 included. The stability of the linear relationship of the logarithm of the number as a function of magnitude M_L for $M_L > 1.5$ is taken as a diagnostic of completeness.

The spatial domain is covered by square boxes of $L \times L$, where we have varied L from 5 km to 50 km. Different sizes for the time window have been considered from

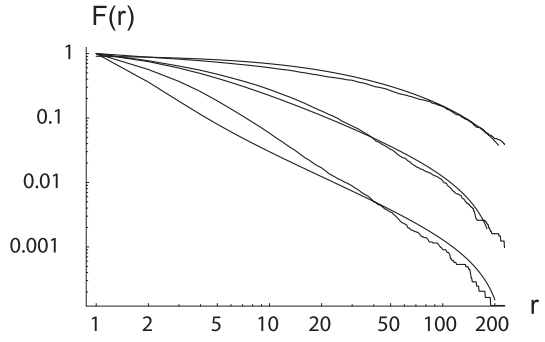

Fig. 7. Empirical complementary cumulative distributions $F_{\rm data}(r) \equiv \int_r^\infty P_{\rm data}(r')dr'$ of the number r of earthquakes in the space-time bins for $L = 20$ km and $\tau = 10$ day, 100 days and 1000 days (bottom to top). The smoothed curves are obtained by numerical integration of expression (115) for the set of parameters (116).

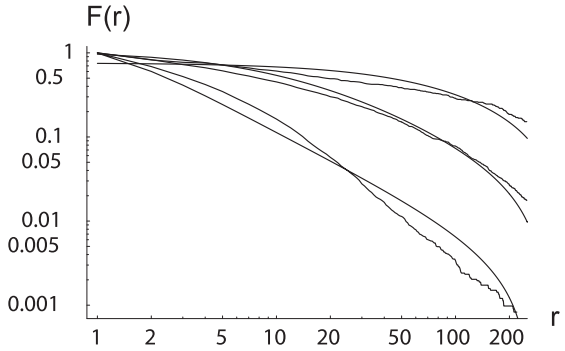

Fig. 8. Same as Figure 8 for $L = 50$ km.

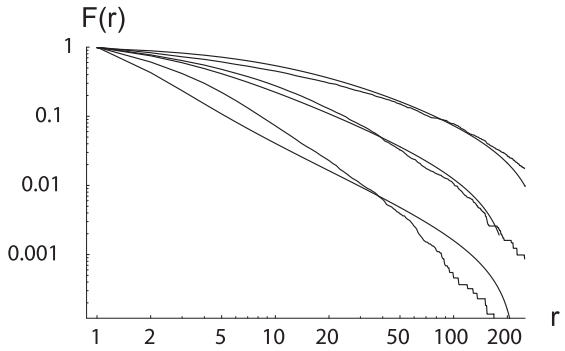

Fig. 9. Same as Figure 8 but with fixed $\tau = 100$ days and for $L = 5$ km, $L = 20$ km and $L = 50$ km (bottom to top).

$\tau = 0.1$ day to $\tau = 1000$ days. Combining the space and time windows leads to space-time windows or bins of size $L^2 \times \tau$. In each space-time bin, we count the number of events and then construct the empirical distribution of binned numbers (which are coarse-grained proxies for the seismic rates).

6.2 Results

For a quantitative comparison between our theory and the experimental data, we take $L = 20$ km and 50 km with $\tau = 10$, 100 and 1000 days as they correspond approximately to the domain of validity of our theory which uses the "large window" approximation developed in Section 5. Figure 7 (respectively Fig. 8) plots the empirical complementary cumulative distributions $\int_r^\infty P_{\rm data}(r')dr'$ of the number r of earthquakes in the space-time bins described above for $L = 20$ km (respectively $L = 50$ km) and $\tau = 10$ day, 100 days and 1000 days. Figure 9 is the same but with fixed $\tau = 100$ days and for $L = 5$ km, $L = 20$ km and $L = 50$ km. These (complementary) cumulative distributions minimize the statistical fluctuations that dominate the statistics and provide a clear comparison with theoretical results.

The theoretical curves shown in Figures 7–9 are obtained by numerical integration of expression (115) for the set of parameters

$$n \in [0.7; 1], \quad \gamma = 1.1, \quad \delta = 0.15, \quad Q = 0.04, \quad \rho = 0.96 \times S \times \tau \text{days} \quad (116)$$

where $S = 1$ for the 20×20 km^2 domain. These parameters are determined by fitting the distribution obtained for $L = 20$ km and $\tau = 100$ days and are then held fixed for all other distributions obtained for the other space-time windows. There is however one remaining parameter which is adjusted as a function of the size of each window: $p = 0.45; 0.85; 0.92$ for $L = 5; 20; 50$ km, respectively. Indeed, recall that p introduced in (66) quantifies the average fraction of the total set of aftershocks which fall within the space-time window $\mathcal{S}$. This free parameter has been introduced as a device in the factorization procedure of Section 5.1. The slow variation of p from $p = 0.85$ for the 20×20 km^2 area to $p = 0.92$ for the 50×50 km^2 area is related physically to the slow power law decay of the propagator $\phi(x)$ defined by (9).

In summary, the distribution for $L = 20$ km and $\tau = 100$ days allows us to determine the parameters given by (116) which are kept fixed for the other distributions. Then, for each space size L, we further determine their corresponding parameter p by a fit of the distribution for $\tau = 100$ days. The p value thus obtained is kept fixed in the theoretical calculation of the distributions for the other time windows $\tau = 10$ day and $\tau = 1000$ days with the same L.

6.3 Discussion of the obtained parameter values

Let us comment on the obtained values (116) of the parameters.

- The value of the branching ratio $n \in [0.7; 1]$ is poorly constrained but seems to be compatible with values previously reported in the literature which range from 0.4 to 0.8 and more. The reason for this lack of sensibility to n has already been alluded to in our theoretical development. Consider the critical value $n = 1$ for which the average number of aftershocks is equal to $\langle R \rangle = \infty$, while the corresponding PDF $P(r)$ of event

numbers only insignificantly differs from the analogous PDF in subcritical regimes. Recall that they differ significantly when $r \gtrsim r_*$, where r_* is given by relation (109). The estimations given below relation (112) show that, for $\gamma = 1.25$ and $n = 0.9$ for instance, we have $r_* \simeq 10^4$, which is much larger than the maximal aftershock number $r_{\max} \sim 200$ which we use to compare the theory and real data. The solution of the inverse problem (i.e., looking for the values n for which $r_* \simeq r_{\max} = 200$, and for which the PDF $P(r)$ is indistinguishable from the critical one) gives that, if $\gamma = 1.25$, $n \gtrsim 0.8$, while if $\gamma = 1.1$, then $n \gtrsim 0.7$. Therefore, we have the same theoretical results relevant to the data range for any n in the interval $n \in (0.7, 1)$. We have indeed checked that the theoretical curves for windows of 100 and 1000 days and the parameters listed in (116) (specifically the values $n = 0.7; 0.8; 0.9$ and $n = 0.96$) gives practically undistinguishable results.

It is worth noting an additional source of uncertainty in the value of the branching ratio n as discussed previously [36,37]: the effect of not accounting for non-observable seismicity (i.e., earthquakes of magnitude $m_0 \leq m < m_d$), which may neverthess trigger observable earthquakes, is to renormalize the true n into a significantly smaller apparent value n_a.

- The possibility that the branching ratio is large (above 0.7 and perhaps close to 1) suggests that real earthquake sequences are close to criticality. This is reminiscent of the view that earthquakes are similar to the avalanches of a self-organized critical system [39,40] functioning at a dynamical critical point. In our context, it is noteworthy that mean-field models of self-organized criticality can be formulated in terms of branching models [41,42].
- The value $\gamma = 1.1$, corresponding to $\alpha \approx 0.9$ (through $\gamma = b/\alpha$), is in reasonable agreement with previous estimations of its value: $\alpha \approx 1$ [43,44], $0.5 \leq \alpha < 1$ [25,45,46]. The quality of the fits does not deteriorate significantly for γ up to 1.3. Therefore, our inversion of the distribution of seismic rates does not provide a rigid constraint on this parameter.
- We obtain practically the same quality of fits for δ in the ranges $0.1 \leq \delta < 0.2$. The choice δ close to zero is consistent with the choice of the Cauchy distribution as a proxy for the heterogeneity of the spatial distribution of spontaneous earthquake sources inferred for the stress field and deduced from previous theoretical [30,31] and empirical analysis of earthquake sources [29]. We have also used functional forms for $f_\delta(x)$ other than (15) to describe the pre-existing heterogeneity of spontaneous earthquake sources, such as half-Gaussian, exponential as well as different variants of power laws. Overall, we find that we need $f_\delta(x)$ to have a power law tail close to the Cauchy distribution in order to get reasonable fits. This shows that the ETAS model as well as any other model of this class of triggered seismicity need to be generalized to account for a pre-existing heterogeneity of the crust, which controls the occurrence of the spontaneous earthquake sources. The strong sensitivity of the quality of the fits with respect to the fractal structure of the spontaneous sources quantified by the exponent δ is a surprising but positive bonus of this work. We did not expect a priori that the distribution of seismic rates would teach so much about the heterogeneity of the seismic active regions.
- The value $Q = 0.04$ leads to $m_0 \approx 0$ with expression (83), using $m_d = 1.5$ for the Southern Californian earthquakes catalog and $b = 1$. This value for the minimum triggered earthquake magnitude is compatible with the estimations obtained independently in [26].
- The value $\rho = 0.96 \times \tau$days for the 20×20 km^2 domain translates into the following average number of events inside that window: $\langle R_{\rm sp}(\tau, \mathcal{S})\rangle \simeq \frac{\rho}{1-n} \simeq 10 \times \tau$ (in units of days), taking $n = 0.9$. This would predict about 10 events on average per day in a space window of 20×20 km^2. But of course, we only observe the fraction $Q = 0.04$, leading the observation of about 0.5 event per day in a space window of 20×20 km^2, which is reasonable.
- We cannot exclude that these results may be dependent on the temporal window of our study (1994–2003), which was chosen for its quality, if earthquake processes are not fully stationary over this time scale.
- In the above analysis, we have assumed that the catalog of seismicity for Southern California is complete for $m > 1.5$. This is unfortunately not the case after a large earthquake. The completeness magnitude 1 h after a $m = 7$ earthquake is roughly $m_0 = 4$ [44,47]. Therefore assuming $m_0 = 1.5$ underestimates the number of events within the first hour by a factor about $10^{b(4-1.5)} \sim 300$. Can this "transient time-magnitude" incompleteness of the catalog change the distribution of event numbers?

  It is difficult to conclude with certainty. First, in contrast with [44,47] which explored the conditional windowed statistics, under the condition that there is some large mainshock at the beginning of the time series, we are here exploring the unconditional windowed statistics, where the statistical weight of large events is not so much dominating (especially if $\alpha < b$). We can also expect the "transient time-magnitude" incompleteness to become negligible for sufficiently large time windows. We have checked that this is the case by generating synthetic catalogs with the ETAS model and by mimicking the "transient time-magnitude" incompleteness by pruning the catalog to remove the events with too small magnitudes and too close to previous events according to the empirical laws reported in [44,47]. We find indeed that the distortion is negligible for the largest time windows but may be significant for the smallest time window of 10 days. This "transient time-magnitude" incompleteness may thus provide an additional source of discrepancy with our theoretical predictions. Accounting for such distortion is difficult as it requires a full space-time description to generalize the "transient time-magnitude" incompleteness

into a "transient space-time-magnitude" incompleteness, which prevents us from applying the correction procedure of [44].

To address this issue, we have performed the following test.

1. for each event, we define a spatial neighborhood (as usual proportional to the length L of the rupture) and a temporal neighborhood (using the formula of [44]). For the temporal neighborhood, the magnitude threshold is taken equal to 1.5 (which is equal to the magnitude detection threshold in the catalog).
2. After each event, we attribute a tag equal to 1 to all events which belong to the space-time neighborhood of the event, as previously defined. All other events have the tag equal to 0.
3. We then construct the histograms of the numbers of events in the space-time boxes, over all boxes which have only earthquakes with tag equal to 0. All other boxes are excluded from the statistics. By so doing , we exclude from the statistics the events which may be affected by the temporal "shadow" of one or more preceding events.

This procedure is rather conservative as an event falling in the "shadow" of a $M = 2$ event, for instance, will be excluded, but it is useful to provide an insight on the impact of temporal incompleteness on our results. We find that the difference between our theoretical predictions and the obtained histograms of the statistics of event numbers is not significantly better or worse than the difference shown in Figures 7–9, suggesting that the temporal incompleteness of the catalog is not the main explanation for the discrepancies between theory and empirical data.

6.4 Limits of the theoretical approach

Overall, the theory provides a reasonable representation of the behavior of the complementary cumulative distributions as L and τ are varied. However, it is clear that there are discrepancies, and all the more so, the smaller τ is. This is not really surprising as the following reasons can be invoked.

1. The empirical data may be incomplete for different reasons, in particular as the "ergodic" time over which intermittent changes of seismic activity should be averaged out is of the order of the repeat time of the largest earthquakes in California, which is of the order of 150 years.
2. In order to obtain the explicit formula (115), we have been forced to use an approximation which requires relatively large L and τ, as explained in section 5. In particular, it follows from the condition (52) and from estimations of the time scale c_1 given by (46) for different values of the Omori power exponent θ and of the branching ratio n that the large time-window approximation works for the time windows investigated here only if $\theta \geq 0.2$ and $n \leq 0.9$ typically. The reasonable agreement with our data analysis for time windows of 10 days and larger suggests the validity of this range of values for θ and n.
3. The experimental distribution $P_{\rm data}(r)$ exhibits an approximate power law tail, as reported in Figures 7–9. In contrast, while the numerical evaluation of formula (115) is reasonably in agreement with the data for $r \lesssim 100$, the asymptotic tail of expression (115) seems to decay faster. This is curious given that the tail of the theoretical distribution should be given asymptotically by the power laws (110) and (112) for unwindowed events. We believe the discrepancy may result from the combination of the windowing effect (the formulas (110) and (112) are actually derived strictly for unbounded domains), the threshold magnitude and the value of *gamma* close to one.
4. Another possible source of error lies in the approximate relation (69). From a probabilistic point of view, this approximation amounts to neglect the possibility that more than one event in the space-time window is triggered by a spontaneous source positioned outside the space window. In addition, the factorization procedure of Section 5.1 (see also Sect. 5.2) accounts only semi-quantitatively for the finiteness of the space domain. A detailed discussion of the applicability of the factorization procedure can be found in Section 7.

Using the same technique in terms of GPF's shows that the distribution of the total number of aftershocks has two power law regimes $\sim 1/r^{1+\frac{1}{\gamma}}$ for $r < r_* \simeq 1/(1-n)^{\gamma/(\gamma-1)}$ and $\sim 1/r^{1+\gamma}$ for $r > r_*$ [23]. We expect the same laws to describe the intermediate asymptotics and asymptotics of windowed distributions of event numbers. Pisarenko and Golubeva [11], with a different approach applied to much larger spatial box sizes in California, Japan and Pamir-Tien Shan, have fitted the distribution of event numbers by a power law

$$P_{\rm data}(r) \sim 1/r^{1+\zeta} \tag{117}$$

with an exponent $\zeta < 1$ which could perhaps be associated with the intermediate asymptotics characterized by the exponent $1/\gamma < 1$, found in our analysis [23]. By using data collapse with varying spatial box sizes on a California catalog, Corral finds that the distribution of seismic rates exhibits a double power-law behavior with $\zeta \approx 0$ for small rates and $\zeta \approx 1.2$ for large rates [2]. The first regime might be associated with the non universal bulk par of the distribution found in our analysis. The second regime is compatible with the prediction for the asymptotic exponent $\zeta = \gamma$. Concerning the collapse of the distributions for different time windows proposed by Corral, first we would like to remark that this collapse involves a rescaling of the different distributions as a function of the size of the space-time windows, which can be predicted from our theoretical analysis. Secondly, a collapse in log-log plot is only an indication and certainly not a proof of existence. In other terms, any collapse contains always some noise and cannot be taken as a proof "by eye." Complementary to the empirical data collapse proposed by Corral [2], we have provided a physical mechanism and a theory to

predict the variation of the distributions with the size of the space-time windows.

Note that the derivation of section 5.4 using the large time window approximation is a priori sensitive to the value of n, and in particular the approximation should not hold for $n \to 1$. In contrast, the fitting procedure leading to the set of parameters (116) is quite robust to changes in n and in particular it works also for $n \simeq 1$. The resolution of this paradox is obtained as follows. We indeed offered a first estimation of the domain of validity of the large time window approximation in terms of condition $\tau \gg c_1$ given by equations (52) with (46), showing that c_1 is highly sensitive to the closeness n to 1. In particular, $c_1 = \infty$ for $n = 1$. It turns out that this condition (52) becomes irrelevant for $n \to 1$ because it is based on the analysis of the average seismicity rate, while the relevant quantity should be the full distribution of rates, which is well-behaved as $n \to 1$, as explained in the preceding paragraph. Using the properties of the statistics of seismic rates conditioned on generation number discussed in the next section 7, we will derive an improved condition (130) for the validity of the large window approximation, which reads $t_* \simeq c\,(k_*/\omega)^{1/\theta}$, where ω is some tolerance level (say $\omega = 0.1$) while k_* is the number of generations for which most earthquakes have been triggered in the space-time window. Figure 11 below shows that the large time window approximation indeed remains valid for n close to 1 and is rather insensitive to the value of n.

Let us also mention other related works. Neglecting the space-time differentiation of events within a cascade, simple branching models of earthquakes have been formulated in the energy space and give also a power law distribution of numbers with exponent determined from the rate of energy cascade and its heterogeneity [51–53].

7 Theoretical tests of the theory using statistics conditioned on generation number

In our theoretical development to obtain the prediction (115) that could be compared with empirical data, we have been obliged to make two main approximations: (1) assuming that the duration τ of the time window $[t, t+\tau]$ is sufficiently large (i.e., the inequality (52) holds), we have replaced the GPF $\Theta(z, \tau, \mathcal{S}; \boldsymbol{x})$ by its asymptotics $\Theta(z, \mathcal{S}; \boldsymbol{x})$; (2) we have used a factorization procedure to take into account quantitatively the finiteness of the spatial window $\mathcal{S}$.

In this section, we attempt to clarify further the domain of application of these two approximations by testing them on other event statistics conditioned on fixed number of generations. Numerical calculations of the exact PDF are compared with the approximations.

7.1 Large spatial windows

Let us consider the statistics of aftershocks triggered over the whole space during the first k generations by some mother event. The corresponding GPF of the number of aftershocks triggered in the course of k generations is defined by the following iterative recurrence equation

$$\Theta_k(z) = G[z\Theta_{k-1}(z)]\,, \qquad \Theta_1(z) = G(z)\,. \tag{118}$$

One can calculate the corresponding probabilities of aftershock numbers by using the Cauchy integral

$$P_k(r) = \frac{1}{2\pi} \oint\limits_{\mathcal{C}} \Theta_k(z) \frac{dz}{z^{r+1}}\,. \tag{119}$$

Furthermore, we can make use of the knowledge that, as $k \to \infty$, the GPF $\Theta_k(z)$ converges to the asymptotic GPF $\Theta(z)$ which is the solution of equation (92). It is easy to show that one can calculate the corresponding probabilities using an equality analogous to (98):

$$P(r) = \frac{1}{2\pi i(r+1)} \oint\limits_{\mathcal{C}'} G^{r+1}(y) \frac{dy}{y^{r+1}}\,. \tag{120}$$

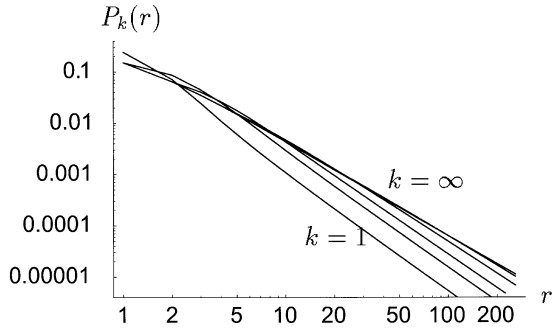

Fig. 10. Distributions $P_k(r)$ given by (119) for $\gamma = 1.25$, $n = 1$, for different numbers of generations $k = 1; 2; 3; 5; 8$. The upper curve corresponds to asymptotic distribution $P(r)$ given (120) corresponding to $k \to +\infty$, while lower one corresponds to the distribution $P_1(r)$ given by (19) of the number of aftershocks of first generation.

Figure 10 shows the distribution $P_k(r)$ of aftershocks numbers, obtained by a numerical calculation of the integral (119) for different values $k = 1, 2, 3, 5, 8$ and for $\gamma = 1.25$ and in the critical case $n = 1$. It also shows the corresponding asymptotic distribution for $k \to +\infty$ obtained by integration of (120). Note that, even for in this critical case, the distribution for $k = 8$ generations is already almost undistinguishable from the asymptotic distribution including an infinite number of generations, at least for $r \leqslant 250$. Figure 11 clarifies further the convergence rate by plotting the ratio

$$p_k(r) = \frac{P_k(r)}{P(r)} \tag{121}$$

for different values k.

The information on the number of generations necessary to reach the asymptotic regime gives us the possibility of estimating the corresponding characteristic time beyond which the asymptotic distribution $P(r)$ becomes a

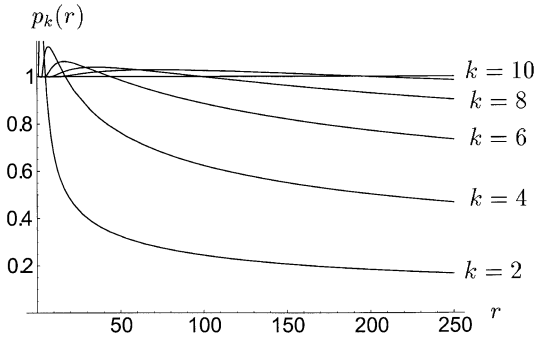

Fig. 11. Ratios (121) as a function of event numbers r for different values of the generation number k, demonstrating the convergence of the distributions $P_k(r)$ to the asymptotic distribution $P(r)$. Bottom to top: $k = 2$; 4; 6; 8; 10.

good approximation of $P_k(r)$. Let $T(k)$ denote the random time at which a k-th generation aftershock is triggered. It is equal to

$$T(k) = \tau_1 + \tau_2 + \cdots + \tau_k \,, \tag{122}$$

where $\{\tau_1, \tau_2, \ldots \tau_k, \ldots\}$ are mutually independent random waiting times between the occurrence of a mother earthquake and one of its first-generation aftershock. We define the ω-th waiting time quantile $t(\omega, k)$ of generation k by

$$\mathcal{Q}[t(\omega,k),k] = \Pr\{T(k) > t(\omega,k)\} = \omega \ . \tag{123}$$

Thus, $1-\omega$ is the probability that the duration of any chain of k successive generations of triggered aftershocks is smaller than $t(\omega, k)$. Choosing some confidence level (for example $1-\omega = 0.9$), one may assert that, during the time $t(\omega, k)$, all aftershocks of the k-th generation have already been triggered. Let $k = k_*$ be such that the corresponding probability $P_{k_*}(r)$ is close to the asymptotic $P(r)$. Then, one may interpret

$$t_* = t(\omega, k_*) \tag{124}$$

as an estimation of the characteristic time for the validity of the asymptotic distribution $P(r)$.

The asymptotic expression for the probability $\mathcal{Q}(t, k)$ defined by (123) for $k \gg 1$ can be determined by using the fact that the terms τ_k of the sum (122) are determined by Omori's law (8) with $0 < \theta < 1$. For $k \gg 1$, $\mathcal{Q}(t, k)$ is asymptotically close to

$$\mathcal{Q}(t,k) = F_\theta\left(\frac{t}{c[k\Gamma(1-\theta)]^{1/\theta}}\right) \ , \tag{125}$$

where

$$F_\theta(x) = \int_x^\infty \varphi_\theta(y)\, dy \tag{126}$$

and $\varphi_\theta(x)$ is the one-sided Lévy stable distribution defined by the Laplace transform

$$\hat{\varphi}_\theta(u) = \int_0^\infty \varphi_\theta(x) e^{-ux}\, dx = e^{-u^\theta} \,. \tag{127}$$

In particular

$$F_{1/2}(x) = \mathrm{erf}\left(\frac{1}{2\sqrt{x}}\right) \,. \tag{128}$$

The following asymptotic behavior holds

$$F_\theta(x) \simeq \frac{x^{-\theta}}{\Gamma(1-\theta)} \qquad (x \gg 1) \,. \tag{129}$$

Substituting (125) and (129) into (123), we obtain the following estimation for t^* defined by (124)

$$t_* \simeq c\left(\frac{k_*}{\omega}\right)^{1/\theta} \ . \tag{130}$$

For instance, for $\theta = 1/2$, $k_* = 8$, $\omega = 0.1$ and $c = 2$ min, then $t_* \simeq 9$ days. Note that t_* is highly sensitive to the value of θ. Indeed, $\theta = 1/3$ (resp. $2/3$) with all other parameters being the same gives $t_* \simeq 700$ days (resp. $t_* \simeq 1$ day).

Expression (130) is related to condition (52) (and actually improves on it) as follows. The condition (52) with (51) can be expressed by introducing, similarly to the reasoning leading to (130), some small threshold $\omega \ll 1$ such that (52) translates into $\mathcal{N}_{\mathrm{out}}(\tau) \lesssim \omega$. Correspondingly, we can introduce τ^* such that, if $\tau > \tau^*$, then the duration of the window is large at the ω level:

$$\omega \simeq \frac{n}{\Gamma(1-\theta)}\left(\frac{c_1}{\tau^*}\right)^\theta \ . \tag{131}$$

Using (46), we get

$$\tau^* \simeq c\left(\frac{n}{\omega(1-n)}\right)^{1/\theta} \ . \tag{132}$$

The two expressions (132) and (130) have a similar structure. The only difference is that the characteristic generation number k_* is replaced by the factor $n/(1-n)$. For n not too close to 1, $n/(1-n)$ gives a not unreasonable estimation of k_*. For n close to 1, expression (130) should be preferred as it provides an improvement to (52) based on the calculation of quantiles rather than on the mean rate behavior.

7.2 Finite spatial windows

A natural generalization of the iterative procedure (118) for finite spatial window $\mathcal{S}$ allows to estimate the corresponding aftershock statistics. Consider an earthquake occurring at point $\boldsymbol{x}$. Then, the PDF of the space positions $\boldsymbol{y}$ of an aftershock of the kth generation is given by $\phi_k(\boldsymbol{y} - \boldsymbol{x})$, where

$$\phi_k(\boldsymbol{y}) = \phi(\boldsymbol{y}) \underbrace{\otimes \cdots \otimes}_{k} \phi(\boldsymbol{y}) \tag{133}$$

is the k-times convolution of the space propagator $\phi(\boldsymbol{y})$ (one example is given by (9)). Correspondingly, the probability for an aftershock of the k-th generation to fall into the space window $\mathcal{S}$ is equal to

$$p_k(\mathcal{S};\boldsymbol{x}) = \iint_{\mathcal{S}} \phi_k(\boldsymbol{y}-\boldsymbol{x})d\boldsymbol{y}\,. \tag{134}$$

Provided these probabilities are known, one can determine the GPF $\Theta_k(z,\mathcal{S};\boldsymbol{x})$ of the number of aftershocks of the k generation occurring in the spatial domain $\mathcal{S}$ by using the following iteration

$$\begin{aligned}
\Theta_1(z,\mathcal{S};\boldsymbol{x}) =& G\left(p_1(z-1)+1\right)\,,\\
\Theta_2(z,\mathcal{S};\boldsymbol{x}) =& G\left[(p_1(z-1)+1)\,G\left(p_2(z-1)+1\right)\right]\\
\Theta_2(z,\mathcal{S};\boldsymbol{x}) =& G\left[(p_1(z-1)+1)\,G\left[(p_2(z-1)\right.\right.\\
&\left.\left.+1)\,G\left[(p_3(z-1)+1)\right]\right]\right]\,,
\end{aligned} \tag{135}$$

and so on up to the order k. Then, the distribution of the number of aftershocks of the k generation is given by

$$P_k(r,\mathcal{S};\boldsymbol{x}) = \frac{1}{2\pi}\oint_{\mathcal{C}} \Theta_k(z,\mathcal{S};\boldsymbol{x})\frac{dz}{z^{r+1}}\,. \tag{136}$$

These expressions are general and hold for any space propagator $\phi(\boldsymbol{y})$. Let us now specialize to the form (9) for the spatial propagator $\phi(\boldsymbol{y})$, with $\eta = 1$,

$$\phi_k(\boldsymbol{x}) = \frac{kd}{2\pi(x^2+k^2d^2)^{3/2}} \tag{137}$$

corresponding to an asymptotic $1/|\mathbf{x}|^3$ decay law often argued on the basis of the shape of the elastic Green function in a three dimensional space. From (134), we then have

$$p_k(\mathcal{S};\boldsymbol{x}) = \mathcal{P}\left(\frac{\ell}{kd},\frac{x}{kd}\right) \tag{138}$$

where

$$\begin{aligned}
\mathcal{P}(u,v) = \frac{2}{\pi}\int_0^u E\left(\frac{4vs}{1+(v+s)^2}\right)\\
\times\frac{sds}{[1+(v-s)^2]\sqrt{1+(v+s)^2}}\,.
\end{aligned} \tag{139}$$

Here, $E(m)$ is the complete elliptic integral

$$E(m) = \int_0^{\pi/2}\sqrt{1-m\sin^2\epsilon}\,d\epsilon\,. \tag{140}$$

Figure 12 shows the distributions (138) for $\ell = 10d$ (recall that ℓ is the radius of the assumed circular domain $\mathcal{S}$ centered on the origin, which has been used in (180)) and for different positions $\boldsymbol{x}$ of the mother earthquake, given by $x/\ell = 0$; 0.4; 0.6; 0.8; 1; 1.2; 1.4; 1.6 from top to bottom. The separation by the curve for $x/\ell = 1$ into two families has a simple explanation. For $x/\ell \le 1$, the mother event lies within the spatial domain $\mathcal{S}$ of interest and it is thus counted as generation 0. Its immediate aftershocks are most probably adjacent to it and thus have a large probability to also fall within $\mathcal{S}$. As the number k of generation increases, aftershocks diffuse away and are less and less likely to fall within $\mathcal{S}$. In contrast, for $x/\ell > 1$, the mother earthquake falls outside $\mathcal{S}$. Therefore, there is not event at the zeroth generation in $\mathcal{S}$, hence the curves start from zero. The first generations of aftershocks which are most likely to be nearby the mother earthquake fall rarely within $\mathcal{S}$. Only as aftershocks of higher generation levels develop and diffuse away from the mother earthquake, can they invade $\mathcal{S}$. Of course, at large generation numbers, the aftershocks diffuse away from any finite spatial domain, explaining the decay of $P_k(\mathcal{S};\boldsymbol{x})$ to zero for large k's.

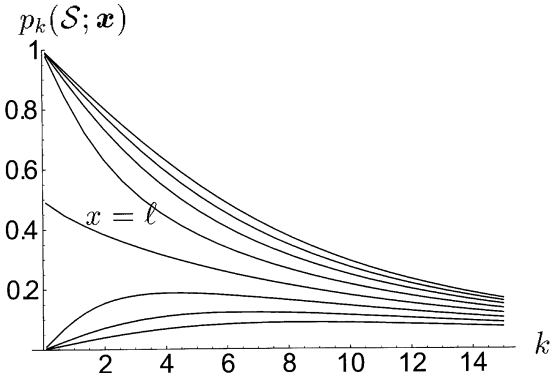

Fig. 12. Plots of the probabilities $P_k(r,\mathcal{S};\boldsymbol{x})$ given by (138) for $\gamma = 1.25$, $n = 0.99$, $\ell = 10d$ and for different positions of the mother earthquake: $x/\ell = 0$; 0.4; 0.6; 0.8; 1; 1.2; 1.4; 1.6. Recall that ℓ is the radius of the assumed circular domain $\mathcal{S}$ centered on the origin. The two families of curves separated by the central one for $x/\ell = 1$ are explained in the text.

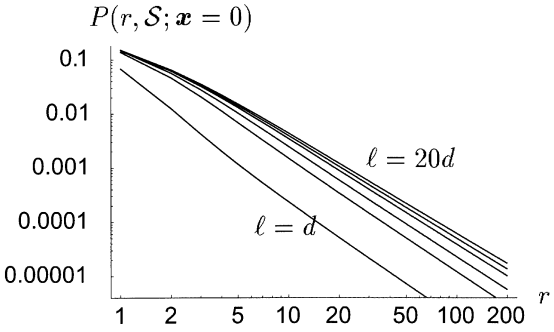

Fig. 13. Plots of the distribution $P(r,\mathcal{S};\boldsymbol{x})$ of the total number r of aftershocks falling within the disk $\mathcal{S}$ for a mother earthquake at the origin $\boldsymbol{x} = 0$ and for different values of the circle radius ℓ. Bottom to top: $\ell/d = 1$; 3; 5; 10; 20.

Figure 13 plots the asymptotic distribution $P(r,\mathcal{S};\boldsymbol{x})$ as a function of the number r of aftershocks for $\gamma = 1.25$, $n = 0.99$, for different values of the radius ℓ of the disk $\mathcal{S}$. The mother earthquake is assumed to occur at the origin, that is, at the center of the disk $\mathcal{S}$. $P(r,\mathcal{S};\boldsymbol{x})$ is obtained by using (136) for $k = 25$ generations, which is certainly a very good approximation to $P(r,\mathcal{S};\boldsymbol{x}) = P_{k\to+\infty}(r,\mathcal{S};\boldsymbol{x})$. Figure 14 plots $P(r,\mathcal{S};\boldsymbol{x})$ as a function of the number r

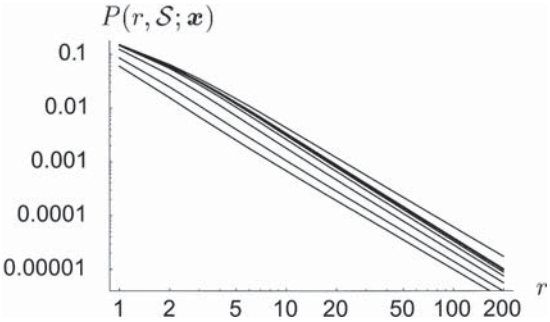

Fig. 14. Plots of the distribution $P(r,\mathcal{S};\boldsymbol{x})$ of the total number r of aftershocks falling within the disk $\mathcal{S}$ for different positions $\boldsymbol{x}$ of the mother earthquake at fixed disk radius $\ell/d = 10$. Bottom to top: $x/\ell = 1.4;\ 1.2;\ 1;\ 0.8;\ 0.6;\ 0.4;\ 0.2;\ 0$. The upper curve thus corresponds to an infinite disk $\ell = +\infty$.

of aftershocks, at fixed $\ell/d = 10$ for various positions of the mother earthquake, for the same parameters $\gamma = 1.25$, $n = 0.99$. One can observe that, when the mother earthquake is inside the disk $\mathcal{S}$ ($x = 0.8\ell;\ 0.6\ell;\ 0.4\ell;\ 0.2\ell;\ 0$), the corresponding distributions are close to each other as predicted in Section 5.1. When the mother is outside the disk $\mathcal{S}$ ($x = 1.4\ell;\ 1.2\ell;\ 1$), the distributions differ from the previous case and are significantly smaller. This gives additional support in favor of the linear approximation (69), which we used in Section 5.1. More precisely, these properties result directly from the analysis of Section 5.1, in which we notice below equation (65) that, for $\ell \gg d$, the factor $p_{\mathcal{S}}(\boldsymbol{x})$ approaches a rectangular function, which leads to the approximation $p_{\mathcal{S}}(\boldsymbol{x}) \simeq \text{const.} = p$ for $\boldsymbol{x} \in \mathcal{S}$. This leads to the natural assumption that the GPF $\Theta(z,\mathcal{S};\boldsymbol{x})$ is almost the same for all interior source positions $\boldsymbol{x} \in \mathcal{S}$. This means in turn that the corresponding distribution $P(r,\mathcal{S};\boldsymbol{x})$ should be almost the same for all $\boldsymbol{x} \in \mathcal{S}$. This remarkable fact is illustrated in Figure 14 in which the curves for the interior source probabilities merge.

7.3 Testing the factorization approach

We can now test the factorization approximation developed in Section 5.1 to take into account the finiteness of the space window $\mathcal{S}$ by comparing it with the approach in term of the statistics over successive generations of the previous section. We thus compare the asymptotic distribution $P(r,\mathcal{S},\boldsymbol{x} = 0)$, obtained by calculating the integral (136) for a large enough generation number k ($k = 25$ is found to be sufficient), with the factorization approximation

$$P(r,p) = \frac{p^r}{2\pi r} \oint_{\mathcal{C}'} \frac{dG(y)}{dy} \frac{G^r(y)\,dy}{[y-(1-p)G(y)]^r}\,, \qquad (141)$$

with an appropriate value of the parameter $p = p_{\mathcal{S}}$, defined as the fraction of the aftershocks which fall into the domain $\mathcal{S}$. Figure 15 shows this comparison for four different values of $\ell/d = 10;\ 7.5;\ 5;\ 2.5$. The upper curve in each panel is the distribution (120) for an infinite domain $\ell/d = +\infty$, as a reference. There is some discrepancy between $P(r,\mathcal{S},\boldsymbol{x} = 0)$ and $P(r,p)$ given by (141). The main difference is that the true distribution $P(r,\mathcal{S},\boldsymbol{x} = 0)$ decays faster than the factorization approximation $P(r,p)$ for large r. We recall that, due to (81), the asymptotic behavior of the distribution $P(r,p)$ obtained under the factorization approximation is the same as for an infinite domain given by (120). Notice that the crossing between theoretical and empirical data curves in Figure 7 for $L = 20$ km and in Figure 9 for $L = 5$ and 20 km are consistent with the crossing observed in Figure 15 between the factorization approximation and the true distribution. Furthermore, the faster rate of decay of the PDF's for finite areas may explain in part that the observation that the exponent $q \approx 1.6$ in (117) may be larger than $1 + \gamma$. Nevertheless, Figure 15 shows that an appropriate choice of the parameter $p = p_{\mathcal{S}}$ allows us, at least semi-quantitatively, to take into account the finiteness of the area $\mathcal{S}$.

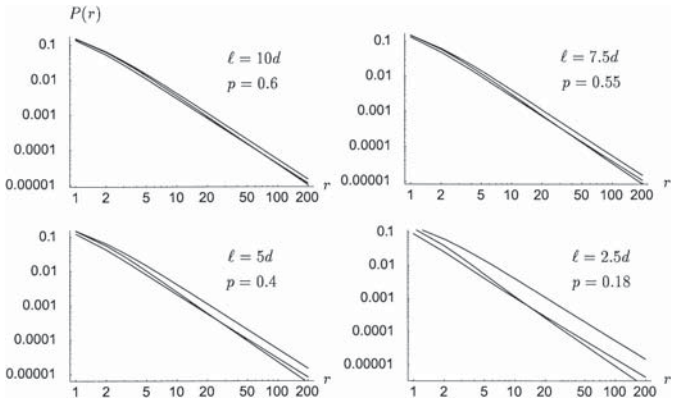

Fig. 15. Comparison of the asymptotic distribution $P(r,\mathcal{S},\boldsymbol{x} = 0)$, obtained by calculating the integral (136) for a large enough generation number k ($k = 25$ is found to be sufficient), with the factorization approximation $P(r,p)$ given by (141), where $p = p_{\mathcal{S}}$ is defined as the fraction of the aftershocks which fall into the domain $\mathcal{S}$. The upper curve in each panel is the distribution (120) for an infinite domain $\ell/d = +\infty$, as a reference. The different panels correspond to $\ell/d = 10;\ 7.5;\ 5;\ 2.5$, with $\gamma = 1.25$, $n = 0.99$, $\boldsymbol{x} = 0$.

8 Discussion

We have presented a general formulation in terms of generating functions of the space-time organization of earthquake sequences, in the framework of general branching processes. We have applied this approach to the ETAS (Epidemic-Type Aftershock Sequence) model of triggered seismicity. In view of the formidable difficulty in obtaining exact solutions to the nonlinear integral equations involving the generating functions, we have developed several approximation schemes which have been tested by comparison with exact numerical calculations. We have used the corresponding predictions to fit the distribution of

seismic rates in four finite space-time windows in a California seismic catalog. The space-time windows differ by their time interval going from $\tau = 10$ day to $\tau = 1000$ days. The fits have been found to account satisfactorily for the empirical observation. In particular, we have adjusted the parameters of the theory on the time window $\tau = 100$ days and have then used these frozen parameter values in the theory to calculate the distribution for the other time windows. This tests the rigidity of the theory to account simultaneously for the distributions at different time scales. In this process, we have found it necessary to augment the ETAS model by taking account of the pre-existing frozen heterogeneity of spontaneous earthquake sources. We have discussed the physical justification of this generalization in terms of pre-existing stress and fault networks, which constrain the form of the pre-existing heterogeneity. Our findings have also important implications to assess the quality of models developed to forecast future seismicity, and suggest to re-examine current procedures which assume Poisson statistics in the construction of likelihood scores.

We thank warmly G. Ouillon for help in the analysis of the data and in the preparation of the corresponding figures. This work is 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 861.

List of symbols

$\otimes$: convolution operator;
α: productivity exponent defined in (2);
$a(t)$: complementary cumulative Omori law defined in (40);
b: b-value of the Gutenberg-Richter distribution (empirically close to 1);
c: regularizing time scale that ensures that the seismicity rate remains finite close to the mainshock
c_1: characteristic time-scale of aftershock branching processes defined in (46);
d: characteristic spatial scale of the spatial propagator $\phi(\boldsymbol{x})$;
δ: exponent of the power law distribution $f_\delta(x)$ of heterogenous sources;
η: exponent of the spatial propagator $\phi(\boldsymbol{x})$;
$E(m)$: complete elliptic integral (140);
$E_\theta(x)$: Mittag-Leffler function;
$f_\delta(x)$: probability density distribution of spontaneous sources defined by (15);
$\Phi(t)$: time propagator quantifying the rate of daughter triggering of first generation at time t after a mother that occurred at time 0;
$\hat{\Phi}(u)$: Laplace transform of $\Phi(t)$;
$\hat{f}_\delta(u)$: Laplace transform of $f_\delta(x)$;
$\phi(\boldsymbol{x})$: spatial propagator quantifying the probability for a daughter to be triggered at a distance $\boldsymbol{x}$ from its mother;
$\tilde{\phi}(\boldsymbol{q})$: Fourier transform of $\phi(\boldsymbol{x})$;
$\tilde{\phi}_{\mathcal{S}}(\boldsymbol{q})$: Fourier transform of $\phi(\boldsymbol{x})$ restricted in the space domain $\mathcal{S}$;
$\varphi_\theta(x)$: one-sided Lévy stable distribution;
$\boldsymbol{\Phi}(\boldsymbol{x}-\boldsymbol{x}', t')$: PDF of the position $\boldsymbol{x}'$ and instant t' of some first generation aftershock, triggered by a mother event, arising at the instant $t = 0$ and at the point $\boldsymbol{x}$, defined in (7);
GPF: generating probability function;
$G(z)$: GPF of the number R_1 of first generation aftershocks triggered by a mother aftershock of arbitrary magnitude;
γ: reduced exponent defined by $\gamma = b/\alpha$;
$\Gamma(x)$: complete Gamma function;
$\Gamma(\theta, x)$: incomplete Gamma function;
κ: constant factor controlling the value (6) of the branching ratio;
$K_{\eta/2}(x)$: modified Bessel function of the second kind;
ℓ: radius of the circular domain $\mathcal{S}$;
$L(z, \tau, \mathcal{S})$: function occurring in the definition (23) of $\Theta_{\rm sp}(z, \tau, \mathcal{S})$;
m: magnitude of an earthquake;
m_0: minimum magnitude of earthquakes capable of triggering other earthquakes (ultraviolet cut-off of the theory);
m_d, M_L: magnitude of completeness of seismic catalogs;
μ: mark associated with an earthquake of magnitude m according to (2);
n: branching ratio equal to the average number of aftershocks of first generation per mother;
N_m: average number of children (triggered events or aftershocks) of first generation) given by (1);
PDF: probability density function;
$N_{\rm sp}(\tau, \mathcal{S})$: rate of events defined in (36);
p: average over all possible spatial positions of mother earthquakes of the fraction of aftershocks which fall within the space-time window $\mathcal{S}$;
$P_1(r)$: PDF of R_1;
$P_{\rm data}(r)$: probability density function of the number r of events;
$p(m)$: Gutenberg-Richter distribution (4);
$p_\mu(r)$: Poisson statistics;
$p_{\mathcal{S}}(\boldsymbol{x})$: overall fraction of aftershocks, triggered by a mother earthquake at position $\boldsymbol{x}$, which fall within the domain $\mathcal{S}$;
$\psi_\delta(x)$: two-sided Levy distribution defined by (104)–(106);
$\Psi(z, t, \tau, \mathcal{S}; \boldsymbol{x})$: auxiliary function describing the space-time dissemination of aftershocks triggering by some mother event;
$Q = Q(m_d)$: fraction of observable events;
$\Theta(z, t, \tau, \mathcal{S}; \boldsymbol{x})$: GPF of aftershocks triggering by some mother event inside the space-time bin;
r: number of events;
R_1: number of first generation aftershocks triggered by a mother of arbitrary magnitude;
$\langle R_{\rm sp}(\tau, \mathcal{S})\rangle$: average of the total number of events inside the space-time window defined in (34);
σ_1: standard deviation of R_1;
ρ: average number of spontaneous sources in the space-

time bin τS defined in (94);
ϱ: average number of spontaneous mother earthquakes per unittime and per unit surface;
$\mathcal{S}$: spatial domain (or bin);
τ: time interval (in unit of days) used to define the space-time bins;
θ: exponent of the time propagator $\Phi(t)$;
$\Theta_{\rm sp}(z,\tau,\mathcal{S})$: GPF of the number of events (including mother earthquakes and all their aftershocks of all generations), falling into the space-time window $\{[t,t+\tau]\times\mathcal{S}\}$;
$\Theta(z,t-t',\tau,\mathcal{S};\boldsymbol{x})$: GPF of the number of aftershocks triggered inside the space-time window $\{[t,t+\tau]\times\mathcal{S}\}$ by some mother event that occurred at time t';
$\Theta(z,\tau,\mathcal{S};\boldsymbol{x})$: GPF of the numbers of aftershocks triggered till time τ inside the space window $\mathcal{S}$ by some mother event arising at the instant $t=0$ and at the point $\boldsymbol{x}$ (defined in (26));
z: running variable of generating probability functions;
ζ: exponent of the power law distribution (117) of earthquake numbers in space-time bins.

Appendix A: 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

$$G(z)=P(0)+P(1)z+P(2)z^2+\cdots=\sum_{r=0}^{\infty}P(r)z^r\ . \quad (142)$$

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

$$P(r)=\frac{\nu^r}{r!}e^{-\nu}\ . \quad (143)$$

Then, the summation of the series in (142) gives

$$G(z)=e^{\nu(z-1)}\ . \quad (144)$$

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

$$P(r)=\frac{1}{r!}\frac{d^rG(z)}{dz^r}\bigg|_{z=0}\ . \quad (145)$$

In some cases, especially for numerical calculations, it is more convenient to use the Cauchy integral formula (which is equivalent to (145))

$$P(r)=\frac{1}{2\pi i}\oint_{\mathcal{C}}G(z)\frac{dz}{z^{r+1}}\ , \quad (146)$$

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

$$\langle R\rangle\equiv\sum_{r=1}^{\infty}rP(r)=\frac{dG(z)}{dz}\bigg|_{z=1}\ . \quad (147)$$

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

$$G_2(z)=G_1[zG_1(z)]\ . \quad (148)$$

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

$$G(z)=G_1(zG(z))\ . \quad (149)$$

As an illustration, let us find the average of the total number of triggered events over all generations. Using relation (147) and the normalizing condition $G(z=1)=1$, differentiating equation (149) with respect to z, we obtain

$$\langle R\rangle=n(1+\langle R\rangle)\qquad\rightarrow\qquad\langle R\rangle=\frac{n}{1-n}\ , \quad (150)$$

where

$$n=\frac{dG_1(z)}{dz}\bigg|_{z=1} \quad (151)$$

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 $n>1$.

We end this brief tutorial by mentioning a calculation technique based on the Cauchy integral formula (146), that we use in this paper to calculate the probabilities $P(r)$ of the total number of events for the branching process under study, for instance to obtain equation (98). The problem is that, in the general case, there is no known explicit solution of the functional equation (149). To overcome this difficulty, let us introduce the new integration variable $y=zG(z)$, such that equation (149)

reads $G(z) = G_1(y)$ and thus z can be expressed as function of y as

$$z = \frac{y}{G_1(y)} \ . \tag{152}$$

By integration by part, the integral in (146) can be written

$$P(r) = \frac{1}{2\pi i r} \oint\limits_{\mathcal{C}} \frac{dG(z)}{z^r} \ . \tag{153}$$

Since $G(z) = G_1(y)$, $dG(z) = dG_1(y)$. Using (152) and again integrating by part, we obtain

$$P(r) = \frac{1}{2\pi i (r+1)} \oint\limits_{\mathcal{C}'} G_1^{r+1}(y) \frac{dy}{y^{r+1}} \ , \tag{154}$$

where $\mathcal{C}'$ is some contour inside the circle $|y| \leqslant 1$ enveloping the origin $y = 0$ in the complex plane y. One can, in particular, replace $\mathcal{C}'$ by the circle $|y| = 1$ and rewrite (154) in the more convenient form for numerical calculations

$$P(r) = \frac{1}{2\pi (r+1)} \int_{-\pi}^{\pi} G_1^{r+1}(e^{is}) e^{-irs} ds \ . \tag{155}$$

Appendix B: Derivation of expressions (38) and (41)

The structure of the first term in the r.h.s. of expression (38) is common to all expressions for the average number of events in arbitrary branching processes. It is however not trivial to derive it in the present context.

First, let us recall that the sought average $\langle R_{\rm sp}(\tau, \mathcal{S})\rangle$ is given by (34). To obtain the explicit value $\langle R_{\rm sp}(\tau, \mathcal{S})\rangle$, we thus need to differentiate the GPF $\Theta_{\rm sp}(z, \tau, \mathcal{S})$ given by (22) with respect to z. Using the normalization condition $\Theta(z = 1, \dots) \equiv 1$ and the randomness of the parameter ϱ, we obtain

$$\langle R_{\rm sp}(\tau, \mathcal{S})\rangle = \langle \varrho \rangle \left[M_{\rm out}(\tau, \mathcal{S}) + M(\tau, \mathcal{S}) + S\tau \right] \ , \tag{156}$$

where

$$M_{\rm out}(\tau, \mathcal{S}) = \int_0^{\infty} dt \iint\limits_{-\infty}^{\infty} d\boldsymbol{x} \langle R(t, \tau, \mathcal{S}; \boldsymbol{x})\rangle \tag{157}$$

and

$$M(\tau, \mathcal{S}) = \int_0^{\tau} dt \iint\limits_{-\infty}^{\infty} d\boldsymbol{x} \langle R(t, \mathcal{S}; \boldsymbol{x})\rangle \ . \tag{158}$$

Recall that

$$\langle R(t, \tau, \mathcal{S}; \boldsymbol{x})\rangle = \left. \frac{d\Theta(z, t, \tau, \mathcal{S}; \boldsymbol{x})}{dz} \right|_{z=1} \ ,$$
$$\langle R(\tau, \mathcal{S}; \boldsymbol{x})\rangle = \left. \frac{d\Theta(z, \tau, \mathcal{S}; \boldsymbol{x})}{dz} \right|_{z=1} \ . \tag{159}$$

Using these equalities and equations (24), (25) yields

$$\langle R(t, \tau, \mathcal{S}; \boldsymbol{x})\rangle = n\Phi(\boldsymbol{x}, t) \otimes \langle R(t, \tau, \mathcal{S}; \boldsymbol{x})\rangle + n\Phi(\boldsymbol{x}, t+\tau) \otimes \langle R(\tau, \mathcal{S}; \boldsymbol{x})\rangle + n\Phi(\boldsymbol{x}, t+\tau) \otimes I_{\mathcal{S}}(\boldsymbol{x}) \ . \tag{160}$$

Analogously, from equations (27), (28), we obtain

$$\langle R(\tau, \mathcal{S}; \boldsymbol{x})\rangle = n\Phi(\boldsymbol{x}, \tau) \otimes \langle R(\tau, \mathcal{S}; \boldsymbol{x})\rangle + n\Phi(\boldsymbol{x}, \tau) \otimes I_{\mathcal{S}}(\boldsymbol{x}) \ . \tag{161}$$

To obtain the equation for $M_{\rm out}(\tau, \mathcal{S})$ defined in (157), let us integrate equation (160) with respect to $\boldsymbol{x}$ and t. This gives

$$M_{\rm out}(\tau, \mathcal{S}) - nM_{\rm out}(\tau, \mathcal{S}) = na(\tau) \otimes \langle R(\tau, \mathcal{S})\rangle + nS \int_0^{\tau} dt \ a(t) \ , \tag{162}$$

with the following notations:

$$\iint\limits_{-\infty}^{\infty} \Phi(\boldsymbol{x}, t) d\boldsymbol{x} = \Phi(t) \ , \quad a(\tau) = \int_{\tau}^{\infty} \Phi(t) dt,$$
$$b(\tau) = \int_0^{\tau} \Phi(t) dt,$$
$$a(\tau) + b(\tau) = 1 \ . \tag{163}$$

Analogously, we obtain from (161)

$$M(\tau, \mathcal{S}) - n\Phi(\tau) \otimes M(\tau, \mathcal{S}) = nS \int_0^{\tau} b(t) dt \ . \tag{164}$$

Note that the following identity is true

$$\langle R(\tau, \mathcal{S})\rangle \otimes a(\tau) \equiv M(\tau, \mathcal{S}) - \langle R(\tau, \mathcal{S})\rangle \otimes b(\tau) \tag{165}$$

which is just a consequence of the identity $a(\tau) + b(\tau) \equiv 1$. We also have

$$\langle R(\tau, \mathcal{S})\rangle \otimes b(\tau) = M(\tau, \mathcal{S}) \otimes \Phi(\tau) \ , \tag{166}$$

and

$$\langle R(\tau, \mathcal{S})\rangle \otimes a(\tau) \equiv M(\tau, \mathcal{S}) - M(\tau, \mathcal{S}) \otimes \Phi(\tau) \ . \tag{167}$$

Using this identity and equation (164), we can rewrite expression (162) in the form

$$M_{\rm out}(\tau, \mathcal{S}) = \frac{n}{1-n} S\tau - M(\tau, \mathcal{S}) \ , \tag{168}$$

which is nothing but expression (38) written in a different form. Substituting it into (156) yields

$$\langle R_{\rm sp}(\tau, \mathcal{S})\rangle = \frac{\langle \varrho \rangle \tau}{1-n} \tag{169}$$

which is equivalent to (41).

Appendix C: Derivation of expression (57) starting from (56)

Let us first consider the last term in the r.h.s. of (56)

$$\frac{\tilde{\Phi}(u)}{1-n\tilde{\phi}(\boldsymbol{q})\tilde{\Phi}(u)}\ . \tag{170}$$

Let us rewrite it in the equivalent form

$$\frac{1}{\tilde{\Phi}^{-1}(u)-n\tilde{\phi}(\boldsymbol{q})} \tag{171}$$

and substitute in this expression the asymptotic relation (11). This gives

$$\frac{\tilde{\Phi}(u)}{1-n\tilde{\phi}(\boldsymbol{q})\tilde{\Phi}(u)} \simeq \frac{1}{\Gamma(1-\theta)(cu)^{\theta}+1-n\tilde{\phi}(q)}\ . \tag{172}$$

Using in (56) the expression (172), we obtain after some simple algebraic transformations

$$\langle\hat{R}\rangle_{+}(u,\mathcal{S};\boldsymbol{q}) \simeq \langle\tilde{R}\rangle(\mathcal{S};\boldsymbol{q})\,\frac{\Gamma(1-\theta)c^{\theta}\,u^{\theta-1}}{\Gamma(1-\theta)(cu)^{\theta}+1-n\tilde{\phi}(\boldsymbol{q})}\,, \tag{173}$$

where $\langle\tilde{R}\rangle(\mathcal{S};\boldsymbol{q})$ is the Fourier transform of the average $\langle R(\mathcal{S};\boldsymbol{x})\rangle$ of the total number of aftershocks falling inside the space domain $\mathcal{S}$ and triggered by an earthquake occurring at position $\boldsymbol{x}$.

This average $\langle R(\mathcal{S};\boldsymbol{x})\rangle$ is equal to

$$\langle R(\mathcal{S};\boldsymbol{x})\rangle = \left.\frac{d\Theta(z,\mathcal{S};\boldsymbol{x})}{dz}\right|_{z=1}\ , \tag{174}$$

where the GPF $\Theta(z,\mathcal{S};\boldsymbol{x})$ satisfies equation (30). Differentiating equation (30) with respect to z and taking into account that

$$\left.\frac{dG(z)}{dz}\right|_{z=1} = n \tag{175}$$

yields

$$\langle R(\mathcal{S};\boldsymbol{x})\rangle = -n\left.\frac{d\Psi(z,\mathcal{S};\boldsymbol{x})}{dz}\right|_{z=1}\ . \tag{176}$$

Differentiating, in turn, expression (31) with respect to z yields

$$\left.\frac{d\Psi(z,\mathcal{S};\boldsymbol{x})}{dz}\right|_{z=1} = -\phi(\boldsymbol{x})\otimes[\langle R(\mathcal{S};\boldsymbol{x})\rangle + I_{\mathcal{S}}(\boldsymbol{x})]\ . \tag{177}$$

We have used here the normalizing condition $\Theta(z=1,\mathcal{S};\boldsymbol{x})\equiv 1$. Combining the two last relations gives

$$\langle R(\mathcal{S};\boldsymbol{x})\rangle - n\phi(\boldsymbol{x})\otimes\langle R(\mathcal{S};\boldsymbol{x})\rangle = \phi(\boldsymbol{x})\otimes I_{\mathcal{S}}(\boldsymbol{x})\ . \tag{178}$$

Applying the Fourier transform with respect to $\boldsymbol{x}$ to both sides of (178) gives, after simple algebraic manipulations,

$$\langle\tilde{R}\rangle(\mathcal{S};\boldsymbol{q}) = \frac{n\,\tilde{\phi}(\boldsymbol{q})}{1-n\tilde{\phi}(\boldsymbol{q})}\,\tilde{I}_{\mathcal{S}}(\boldsymbol{q})\,, \tag{179}$$

where $\tilde{I}_{\mathcal{S}}(\boldsymbol{q})$ is the Fourier transform of the indicator function of the space window $\mathcal{S}$. In this paper, we assume that $\mathcal{S}$ is the circular domain of radius ℓ centered at the origin of the plane $\boldsymbol{x}$. Then

$$\tilde{I}_{\mathcal{S}}(\boldsymbol{q}) = 2\pi\,\frac{\ell}{q}\,J_1(\ell q)\,. \tag{180}$$

To obtain (57), we use the theory of Mittag-Leffler functions, in particular the expression of the Laplace transform of the function $E_{\theta}(-\tau^{\theta})$

$$\int_0^{\infty} e^{-u\tau}E_{\theta}(-\tau^{\theta})d\tau = \frac{u^{\theta-1}}{1+u^{\theta}}\ . \tag{181}$$

Reciprocally, if the Laplace transform of a function is equal to the r.h.s. of (181), then this function is necessarily $E_{\theta}(-\tau^{\theta})$. We briefly indicate how (181) is obtained. In [20] (Vol. 3, Chap. 18, Sect. 1), the elegant formula (18) reads

$$\int_0^{\infty} e^{-t}E_{\alpha}(t^{\alpha}z)dt = \frac{1}{1-z}\ . \tag{182}$$

Introducing the new integration variable τ such that $t=\tau u$, we obtain

$$u\int_0^{\infty} e^{-u\tau}E_{\theta}(\tau^{\alpha}zu^{\alpha})d\tau = \frac{1}{1-z}\ . \tag{183}$$

The next step is to choose $z=-u^{-\alpha}$ which gives

$$u\int_0^{\infty} e^{-u\tau}E_{\theta}(-\tau^{\alpha})d\tau = \frac{1}{1+u^{-\alpha}}\ . \tag{184}$$

It is easy to see that, after replacing α by θ and after simple algebraic manipulations, we obtain (181).

Coming back to equation (173), it implies that $\langle R_{+}(\tau,\mathcal{S};\boldsymbol{x})\rangle$ is given by

$$\langle R_{+}(\tau,\mathcal{S};\boldsymbol{x})\rangle = \langle R(\mathcal{S};\boldsymbol{x})\rangle\otimes\mathcal{H}(\tau;\boldsymbol{x})\,, \tag{185}$$

where the Fourier transform of the function $\mathcal{H}(\tau;\boldsymbol{x})$ is equal to

$$\tilde{\mathcal{H}}(\tau;\boldsymbol{q}) = E_{\theta}\left(-\frac{1-n\tilde{\phi}(\boldsymbol{q})}{1-n}\left(\frac{\tau}{c_1}\right)^{\theta}\right)\ . \tag{186}$$

Thus, the Fourier transform (with respect to $\boldsymbol{x}$) of the sought average given by (55) is given by expression (57).

References

1. P. Bak, K. Christensen, L. Danon, T. Scanlon, Phys. Rev. Lett. **88**, 178501 (2002)
2. A. Corral, Phys. Rev. E. **6803**(3 Part 2), 5102 (2003)
3. M.S. Mega, P. Allegrini, P. Grigolini, V. Latora, L. Palatella, A. Rapisarda, S. Vinciguerra, Phys. Rev. Lett. **90**, 18850 (2003)
4. S. Abe, N. Suzuki, Europhys. Lett. **65** (4), 581 (2004)
5. M. Baiesi, M. Paczuski, Phys. Rev. E **69**, 066106 (2004)

6. M. Baiesi, M. Paczuski, Nonlinear Processes in Geophysics, **12**, 1–11 (2005)
7. M. Baiesi, Physica A **359**, 775–783 (2006)
8. D. Schorlemmer, M. Gerstenberger, S. Wiemer, D. Jackson, Earthquake likelihood model testing, preprint, 2005.
9. M.C. Gerstenberger, S. Wiemer, L.M. Jones, P.A. Reasenberg, Nature **435**, 328 (2005)
10. A. Helmstetter, Y. Kagan, D. Jackson, *Comparison of short-term and long-term earthquake forecast models for Southern California*, in press in Bull. Seism. Soc. Am. (2005)
11. V.F. Pisarenko, T.V. Golubeva, Computational Seismology and Geodynamics **4**, 127 (1996)
12. Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988)
13. Y.Y. Kagan, L. Knopoff, J. Geophys. Res. **86**, 2853, (1981)
14. A. Helmstetter, D. Sornette, J. Geophys. Res. **107**, (B10) 2237 (2002)
15. A. Helmstetter, D. Sornette, J. Geophys. Res. **108**, (B10) 2457 (2003)
16. B.A. Sevastyanov, *Branching Processes* (Moscow, Nauka, 1971)
17. T.E. Harris, *The theory of branching processes* (Prentice-Hall, Englewood Cliffs, New York, 1964)
18. *Classical and modern branching processes*, edited by K.B. Athreya, P. Jagers (Springer, New York, 1997)
19. *Branching processes and its estimation theory*, edited by G. Sankaranarayanan (Wiley, New York, 1989)
20. H. Bateman, A. Erdélyi, *Higher Transcendental Functions*, Vol. 2 (Mac Graw-Hill, 1953)
21. A. Helmstetter, D. Sornette, Phys. Rev. E. **6606**, 061104 (2002)
22. A. Helmstetter, G. Ouillon, D. Sornette, J. Geophys. Res. **108**, 2483 (2003)
23. A. Saichev, A. Helmstetter, D. Sornette, *Pure and Applied Geophysics* **162**, 1113 (2005)
24. A. Saichev, D. Sornette, Phys. Rev. E **70**, 046123 (2004)
25. A. Helmstetter, Phys. Rev. Lett. **91**, 058501 (2003)
26. D. Sornette, M.J. Werner, J. Geophys. Res. **110**, No. B8, B08304, (2005)
27. A. Sornette, D. Sornette, Geophys. Res. Lett. **6**, 1981 (1999)
28. A. Gorshkov, V. Kossobokov, A. Soloviev, in *Nonlinear dynamics of the lithosphere and earthquake prediction*, edited by V.I. Keilis-Borok and A.A. Soloviev (Springer, Heidelberg) pp. 239–310
29. Y.Y. Kagan, Nonlinear Processes in Geophysics **1**, 171 (1994)
30. V.M. Zolotarev, B.M. Strunin, Soviet Phys. Solid State **13**, 481 (1971)
31. V.M. Zolotarev, *One-dimensional Stable Distributions*, Amer. Math. Soc. Providence R.I. (1986)
32. D. Sornette, *Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, 2nd edn. (Springer Series in Synergetics, Heidelberg, 2004)
33. A. Helmstetter, D. Sornette, Geophys. Res. Lett. **30** (11) (2003)
34. D. Vere-Jones, *A class of self-similar random measures*, in press in Adv. Appl. Probab. (2005)
35. A. Saichev, D. Sornette, Phys. Rev. E **72**, 056122 (2005)
36. D. Sornette, M.J. Werner, J. Geophys. Res. **110**, No. B9, B09303 (2005)
37. A. Saichev, D. Sornette, *Renormalization of the ETAS branching model of triggered seismicity from total to observable seismicity*, preprint at `http://arxiv.org/abs/physics/0507024`
38. G. Ouillon, D. Sornette, J. Geophys. Res. **110**, B04306 (2005) `http://arXiv.org/abs/cond-mat/0407208`
39. A. Sornette, D. Sornette, Europhys. Lett. **9**, 197 (1989)
40. P. Bak, C. Tang, J. Geophys. Res. **94**, 15635 (1989)
41. P. Alstrom, Phys. Rev. A **38**, 4905 (1988)
42. P. Alstrom, Phys. Rev. A **41**, 7049 (1990)
43. K.R. Felzer, T.W. Becker, R.E. Abercrombie, G. Ekstroem, J.R. Rice, J. Geophys. Res. **107** (B9), 2190 (2002)
44. A. Helmstetter, Y.Y. Kagan, D.D. Jackson, J. Geophys. Res. **110**, B05S08 (2005)
45. R. Console, M. Murru, A.M. Lombardi, J. Geophys. Res. **108** (B10), 2468, (2002)
46. J. Zhuang, Y. Ogata, D. Vere-Jones, J. Geophys. Res. **109**, B05301 (2004)
47. Y.Y. Kagan, Bull. Seism. Soc. Am. **94** (4), 1207 (2004)
48. J.K. Gardner, L. Knopoff, Bull. Seismol. Soc. Amer. **64**, 1363 (1974)
49. P. Reasenberg, J. Geophys. Res. **90**, 5479 (1985)
50. S.D. Davis, C. Frohlich, J. Geophys. Res. **96** (B4), 63356350 (2001)
51. V.N. Gaisky, *Statistical investigation of earthquake processes* (Moscow, Nauka, 1970)
52. S. Shlien, M.N. Toksoz, Bull. Seismol. Soc. Amer. **60**, No. 6, (1970)
53. Y.Y. Kagan, *Statistical methods in the study of seismic processes*, in *Proceedings of the 39th session*, Bulletin of the International Statistical Institute, Vienna, Vol. XLV, Book 3, pp. 437–453 (1973)