Diffusion of Earthquake Aftershock Epicenters, Omori's Law and Generalized Continuous-Time Random Walk Models

A. Helmstetter, D. Sornette

Diffusion of epicenters of earthquake aftershocks, Omori's law, and generalized continuous-time random walk models

A. Helmstetter[1,*] and D. Sornette[2,3,†]

[1]*Laboratoire de Géophysique Interne et Tectonophysique, Observatoire de Grenoble, Université Joseph Fourier, Boîte Postale 53X, 38041 Grenoble Cedex, France*

[2]*Laboratoire de Physique de la Matière Condensée, CNRS UMR6622 and Université de Nice–Sophia Antipolis, Boîte Postale 71, Parc Valrose, 06108 Nice Cedex 2, France*

[3]*Institute of Geophysics and Planetary Physics and Department of Earth and Space Science, University of California, Los Angeles, California 90095*

(Received 25 March 2002; published 30 December 2002)

The epidemic-type aftershock sequence (ETAS) model is a simple stochastic process modeling seismicity, based on the two best-established empirical laws, the Omori law (power-law decay $\sim 1/t^{1+\theta}$ of seismicity after an earthquake) and Gutenberg-Richter law (power-law distribution of earthquake energies). In order to describe also the space distribution of seismicity, we use in addition a power-law distribution $\sim 1/r^{1+\mu}$ of distances between triggered and triggering earthquakes. The ETAS model has been studied for the last two decades to model real seismicity catalogs and to obtain short-term probabilistic forecasts. Here, we present a mapping between the ETAS model and a class of CTRW (continuous time random walk) models, based on the identification of their corresponding master equations. This mapping allows us to use the wealth of results previously obtained on anomalous diffusion of CTRW. After translating into the relevant variable for the ETAS model, we provide a classification of the different regimes of diffusion of seismic activity triggered by a mainshock. Specifically, we derive the relation between the average distance between aftershocks and the mainshock as a function of the time from the mainshock and of the joint probability distribution of the times and locations of the aftershocks. The different regimes are fully characterized by the two exponents θ and μ. Our predictions are checked by careful numerical simulations. We stress the distinction between the ''bare'' Omori law describing the seismic rate activated directly by a mainshock and the ''renormalized'' Omori law taking into account all possible cascades from mainshocks to aftershocks of aftershock of aftershock, and so on. In particular, we predict that seismic diffusion or subdiffusion occurs and should be observable only when the observed Omori exponent is less than 1, because this signals the operation of the renormalization of the bare Omori law, also at the origin of seismic diffusion in the ETAS model. We present predictions and insights provided by the ETAS to CTRW mapping which suggest different ways for studying seismic catalogs. Finally, we discuss the present evidence for our predicted subdiffusion of seismicity triggered by a main shock, stressing the caveats and limitations of previous empirical works.

DOI: 10.1103/PhysRevE.66.061104 PACS number(s): 05.40.Fb, 91.30.Px

I. INTRODUCTION

The spatiotemporal complexity of earthquakes is often invoked as an illustration of the phenomenon of critical self-organization with scale-invariant properties [1–5]. This concept points to the importance of developing a system approach in which large scale properties can emerge from the repeating interactions occurring at smaller scales. Such ideas are implemented in models proposing links between the physics of earthquakes and concepts of statistical physics, such as critical points, self-organized criticality, spinodal decomposition, critical depinning, etc., in order to explain the most solidly established facts in the phenomenology of earthquakes, of which we cite the three most important.

(i) Law 1. The Gutenberg-Richer law [6] states that the cumulative distribution of earthquake magnitudes m sampled over broad regions and large time intervals is proportional to 10^{-bm}, with a b value $b \approx 1$. Translating into energies E with the correspondence $m = (2/3)\log_{10} E + \text{const}$ leads to a power law $\sim 1/E^B$ with $B \approx 2/3$.

(ii) Law 2. Omori's law for aftershocks [7] states that the rate of earthquakes triggered by a mainshock decays with time according to an inverse power $1/t^p$ of time with an exponent $p \approx 1$.

(iii) Law 3. The earthquakes are clustered in space along hierarchical fault structures [8] and their spatial distribution over long times can be approximately described by a fractal dimension close to 2.2 (in three dimensions) [9].

There are many other empirical ''laws'' but these three characterize the very fundamentals of seismicity in size, time, and space.

We should immediately point out that these three laws come with significant caveats.

(1) There have been ongoing controversies on the universality of the exponent B or b value of the Gutenberg-Richter law [10,11].

(2) The exponent p of Omori's law exhibits a large variability from one aftershock sequence to another aftershock sequence and is found typically in the range from 0.3 to 2. We note, however, that not all these values, especially the extreme ones, automatically reflect a bona-fide power-law

*Email address: ahelmste@obs.ujf-grenoble.fr

†Email address: sornette@ess.ucla.edu

decay and one should exert caution in attributing too much confidence to them.

(3) The view that geological faults and earthquake hypocenters are fractal objects is now recognized to be a naive description of a much more complex reality in which a hierarchy of scales occur with possibly different organizations at different scales [8].

In addition, a major difficulty for making progress in modeling and predicting earthquakes is that these three and other laws may be ''explained'' by a large variety of models, with many different mechanisms. For instance, with respect to the first two laws, we observe the following.

(a) There are many mechanisms that create a power-law distribution of earthquake sizes (see, for instance, the list of mechanisms described in chapter 14 of Ref. [12]).

(b) Omori's law is essentially a slowly decaying ''propagator'' describing a long time memory of past events impacting on the future seismic activity. Such slow power-law time decay of the Omori propagator may result from several and not necessarily exclusive mechanisms (see [13] and references therein): pore-pressure changes due to pore-fluid flows coupled with stress variations, slow redistribution of stress by aseismic creep, rate-and-state dependent friction within faults, coupling between the viscoelastic lower crust and the brittle upper crust, stress-assisted microcrack corrosion [14,15], slow tectonic driving of a hierarchical geometry with avalanche relaxation dynamics [16], etc.

The zeroth-order description of earthquakes is to consider a single isolated homogeneous fault on which earthquakes are recurrent to accommodate the long-term slow tectonic loading. But faults are not isolated and the most conspicuous observation is that earthquakes interact and influence each other on complex fault structures. Understanding these interactions is essential for understanding earthquakes and fault self-organization. However, the full impact of interactions between earthquakes is still far from being well understood. The simplest and clearest observation of earthquake interaction is provided by aftershocks whose phenomenology is captured by Omori's law (Law 2). Indeed, aftershocks are the most obvious and striking signature of the clustering of the seismicity in time and space, and are observed after all large shallow earthquakes. Most aftershocks are triggered a few hours or days after the mainshock. However, due to the very slow power-law decay of the rate of aftershocks, known as the Omori law [7], aftershocks can be triggered up to a hundred years after the mainshock [17]. Aftershocks often occur near the rupture zone of the mainshock with a variety of focal mechanisms suggesting that they are actually on separate structures [18,19]. They are also sometimes triggered at very large distances from the mainshock [20–24]. As an example, Hill *et al.* [20] observed aftershocks of the Landers earthquake as far as 1250 km from the epicenter. Similarly to the temporal distribution of aftershocks, a power-law distribution seems to describe well the distribution of distances between pairs of events [22]. Since a power law decays slowly, it describes a slow decay of the probability of observing aftershocks at large distances to the mainshock.

Thus, Omori's law can be considered as the simplest and best-established description of earthquake interactions of a certain kind. The question we investigate is whether it can be used fruitfully to explain a larger variety of earthquake interactions beyond the class of observations that were used to establish it. In a series of papers [25–27], we find that Omori's law for aftershocks plus the constrain that aftershocks are distributed according to the Gutenberg-Richter power law for earthquake size distribution *independently* of the magnitude of their progenitor is enough to derive many of the other empirical ''laws,'' as well the variability of the p exponent. Here, we test the potential of this approach to account for and to quantify observations on aftershock diffusion.

Aftershock diffusion refers to the phenomenon of expansion or migration of aftershock zone with time [28–36]. Immediately after the mainshock occurrence, most aftershocks are located close to the rupture plane of the mainshock, then aftershocks seem to migrate away from the mainshock, at velocities ranging from 1 km/h to 1 km/year [36,37]. Note that this expansion is not universally observed, but is more important in some areas than in others [31].

The diffusion of aftershocks is usually interpreted as a diffusion of the stress induced by the mainshock, either by a viscous relaxation process [37], or due to fluid transfer in the crust [38,39,35]. Another interpretation of the expansion of aftershocks is given by Dieterich [40], who reproduces the Omori law decay of aftershocks and the expansion of the aftershock zone with time, using a rate and state friction law and assuming that the rate of aftershocks is proportional to the stress rate. In his model, the expansion of aftershock zone arises from the nonuniform stress induced by the mainshock. Another alternative explanation is that the diffusion of aftershocks is mainly due to the occurrence of large aftershocks, and to the localization of secondary aftershock close to the largest aftershocks, as observed by Ouchi [34]. The apparent diffusion of the seismicity may thus result from a cascade process; the mainshock triggers aftershocks that in turn trigger their own aftershocks, and thus lead to an expansion of the aftershock zone.

In the present paper, we investigate the epidemic-time aftershock sequence (ETAS) model, and show that the cascade of secondary aftershocks can indeed explain the reported diffusion of aftershocks. The ETAS model was introduced by Kagan and Knopoff [41] (in a slightly different form than used here) and Ogata [42] to describe the temporal and spatial clustering of seismicity. This model provides a tool for understanding the clustering of the seismic activity, without arbitrary distinction between aftershocks, foreshocks, and mainshocks. In this model, all earthquakes are assumed to be simultaneously mainshocks, aftershocks, and possibly foreshocks. Each earthquake generates aftershocks that decay with time according to Omori's law, which will in turn generate their own aftershocks. The seismicity rate at any given time and location is given by the superposition of aftershock sequences of all events impacting that region at that time according to space-time ''propagators.'' The additional ingredient in the version of the ETAS model that we study is that the number of aftershocks per earthquake increases exponentially $\propto 10^{\alpha m}$ with the magnitude m of the mainshock (i.e., as a power law $\propto E^{2\alpha/3}$ of the energy released by the main-

shock), in agreement with the observations [43,44]. Since the energy of an earthquake is a power law of its rupture length, this law expresses the very reasonable idea that the number of events related to a given earthquake is proportional to a power of its volume of influence. The value of the exponent α controls the nature of the seismic activity, that is, the relative role of small compared to large earthquakes. Few studies have measured α in seismicity data [43,45,46]. This parameter α is often found close to b [43] or fixed arbitrarily equal to b [41,47]. In the case where α is close to the Gutenberg-Richter b value, this law also reproduces [47] the self-similar empirical Bath' s law [48], which states that the average difference $m_M - m_A$ in magnitude between a mainshock and its largest aftershock is approximately 1.2 units, regardless of the mainshock magnitude: $m_A = m_M - 1.2$. If $\alpha < b$, small earthquakes, taken together, trigger more aftershocks than larger earthquakes. In contrast, large earthquakes dominate earthquake triggering if $\alpha \geqslant b$. This case $\alpha \geqslant b$ has been studied analytically in the framework of the ETAS model by Ref. [27] and has been shown to eventually lead to a finite time singularity of the seismicity rate. This explosive regime cannot, however, describe a stationary seismic activity.

A natural way to tame this singular behavior is to introduce an upper cutoff for the magnitude distribution at large magnitudes, mirroring the cutoff m_0 used for the low-magnitude range. The physical argument for introducing this cutoff is based on the finiteness of the maximum earthquake that the earth is capable of carrying. The specific way of introducing such a cutoff (abrupt or smooth with a transition to a power law with larger exponent or to an exponential taper) is not very important qualitatively because all these laws will regularize the singular behavior and make the average branching ratio defined below finite. Such regularization with a maximum upper magnitude then allows $\alpha \geqslant b$. The special case $\alpha = b$ required for Bath's law to hold exactly cannot therefore be excluded.

However, based on a recent reanalysis of seismic catalogs using the powerful collapse technique, one of us [46] has presented strong evidence that α is strictly smaller than b. In this paper, we will therefore consider only the case $\alpha < b$ and take $\alpha = 0.5$ specifically in our numerical simulations. In this regime $\alpha < b$, Bath's law cannot be reproduced because the average difference in size between a mainshock and its largest aftershock increases with the mainshock magnitude. For $\alpha < b$, it is easy to show that Bath's law is replaced by $m_A = (\alpha/b) m_M - \text{const}$, where m_M and m_A are the magnitudes of the mainshock and of the largest aftershock. Tests of this prediction will be reported in a future publication but we expect that distinguishing this modified Bath's law from Bath's law will be a difficult task due to the limited range of the studied magnitudes as well as the dependence of the distribution of $m_M - m_A$ on the magnitude thresholds chosen for the mainshocks and for the aftershocks [49].

We assume that the distribution of all earthquakes follow the Gutenberg-Richter distribution and take this distribution of aftershock sizes to be independent of the magnitude of the mainshock. Therefore, an earthquake can trigger a larger earthquake, albeit with a small probability. This model can thus describe *a priori* both aftershock and foreshock sequences. The ETAS model has been calibrated to real seismicity catalogs to retrieve its parameters [42,50–54,45,47] and to give short-term probabilistic forecast of seismic activity by extrapolating past seismicity into the future via the use of its space-time propagator [41,55,56].

The ETAS model is a branching model that exhibits different regimes [26] depending upon the value of the branching ratio n, defined by the average number of primary aftershocks per earthquake. The critical case $n = 1$ corresponds to exactly one primary aftershock per earthquake, when averaging over all mainshock magnitudes larger than a threshold m_0. Let us stress that n is an average quantity that does not reflect adequately the large variability of the number of aftershocks per main shock, as a function of its magnitude. Indeed, the number of aftershocks per mainshock increases exponentially fast as a function of the mainshock magnitude, so that large mainshocks will have significantly more than n aftershocks. For $\alpha = 0.5$, a magnitude-7 earthquake gives typically ten times more direct aftershocks than a magnitude-5, and 100 times more direct aftershocks than a magnitude-3 earthquake. The increase in triggered seismic activity with the magnitude of the mainshock is obviously stronger for a larger value of α. Note that these numbers refer to aftershocks of the first generation; the total number of triggered events is larger by the factor $1/(1-n) \sim 10$ (for $n \approx 0.9$ which is typical), due to the cascades of secondary aftershocks. Notwithstanding this large variability, the average number n of primary aftershocks per earthquake controls the global regime. For n exactly equal to 1, seismicity is at the border between death and growth. In the subcritical regime $n < 1$, since each earthquake triggers on average less that one aftershock, starting from a large event, the seismicity will decrease with time and finally die out. The supercritical regime $n > 1$ corresponds to more that one primary aftershock per earthquake on average. Starting from a large earthquake, after a transient regime, the average seismicity will finally increase exponentially with time [26], but there is still a finite probability for aftershock sequences to die out.

The numerical simulations reported below have been performed with $\alpha = 0.5$. It is probable that a good fit to seismic data is obtained by using a value of $\alpha \approx 0.8$ larger that the value 0.5, as reviewed and documented recently by one of us [46]. We have checked that results similar to those presented below hold true qualitatively for larger values $0.5 < \alpha < 1$. Such larger values of α lead, however, to stronger fluctuations that are more difficult to handle numerically because the variance of the number $\rho(m)$ of direct triggered aftershocks defined below in Eq. (3) becomes undefined for $\alpha > 0.5$. A full understanding of this regime requires a special treatment that will be reported elsewhere.

Sornette and Sornette [25] studied analytically a particular case of this model, without magnitude and spatial dependence, and they considered only the subcritical regime $n < 1$. Starting with one event at time $t = 0$ and considering that each earthquake generates an aftershock sequence with a ''local'' Omori exponent $p = 1 + \theta$, where $\theta > 0$, they studied the decay law of the ''global'' aftershock sequence, composed of all secondary aftershock sequences, i.e., by taking into account that the primary aftershocks can create second-

ary aftershocks which themselves may trigger tertiary aftershocks and so on. They found that the global aftershock rate decays according to an Omori law with an exponent $p=1-\theta<1$, up to a characteristic time [25,26]

$$t^*=c\left(\frac{n\Gamma(1-\theta)}{|1-n|}\right)^{1/\theta}, \tag{1}$$

and then recovers the local Omori exponent $p=1+\theta$ for time larger than t^*. Helmstetter and Sornette [26] extended their analysis to the general ETAS model with magnitude dependence, and considered both the subcritical and the supercritical regime, but still restricted the analysis to the temporal distribution of the seismicity, without spatial dependence. In the subcritical regime, they recovered the crossover found by Sornette and Sornette [25]. In addition, Helmstetter and Sornette [26] give the explicit mathematical formula for the gradual transition between the Omori law with exponent $p=1-\theta$ for $t\ll t^*$ to the Omori law with exponent $p=1+\theta$ for $t\gg t^*$. This smooth transition can be observed in Fig. 2 on the line calculated for $t^*=10^9$ days with $n<1$. t^* can thus be viewed as the time where the apparent exponent p of the Omori law is approximately in between the two asymptotic values $1-\theta$ and $1+\theta$. A more rigorous mathematical definition [26] is that t^* is the characteristic time scale such that βt^* is the dimensionless variable of the Laplace transform (with variable β) of the seismicity rate.

In the supercritical regime, Helmstetter and Sornette [26] found a novel transition between a power-law decay with exponent $p=1-\theta$ at early times, similar to the subcritical regime, to an exponential increase of the seismicity at large times. The regime where $\alpha>b$ or equivalently $2\alpha/3>B$ has been found to lead to a new kind of critical stochastic finite-time-singularity [27], relying on the interplay between long-memory and extreme fluctuations. Recall that the number of aftershocks per earthquake increases as a power law $\propto E^{2\alpha/3}$ of the energy released by the mainshock whereas the number of earthquakes of energy E decreases as the Gutenberg-Richter law $\propto 1/E^{1+B}$. Intuitively, when $2\alpha/3>B$, the increase in the rate of creation of aftershocks with the mainshock energy more than compensates the decrease of the probability to get a large mainshock when the mainshock energy increases. This theory based solely on the ETAS model has been found to account for the main observations (power-law acceleration and discrete scale invariant structure) of critical rupture of heterogeneous materials, of the largest sequence of starquakes ever attributed to a neutron star as well as of some earthquake sequences [27].

In the sequel, we extend the analytical study of the temporal ETAS model [25–27] to the spatio-temporal domain. To model the spatial distribution of aftershocks, we assume that the distance between a mainshock and each of its direct aftershock is drawn from a given distribution, independently of the magnitude of the mainshock and of the delay between the mainshock and its aftershocks. For illustration, but without loss of generality, for the mapping to the continuous time random walk (CTRW) model discussed later, we shall take a power-law distribution of distances between earthquakes. We take the simplest and most parsimonious hypothesis that space, time, and magnitude are decoupled in the earthquake propagator. Our first result is to establish a correspondence between the ETAS model and the CTRW model, first introduced by Montroll and Weiss [57] and used to model many physical processes. We then build on this analogy to derive the joint probability distribution of the times and locations of aftershocks. We show analytically that, for sufficiently short times $t<t^*$, the average distance between a mainshock and its aftershock increases subdiffusively as $R\sim t^H$, where the exponent H depends on the local Omori exponent $1+\theta$ and on the distribution of the distances between an earthquake and its aftershocks. We also demonstrate that the local Omori law is not universal, but varies as a function of the distance from the mainshock. Due to the diffusion of aftershocks with time, the decay of aftershock is faster close to the mainshock than at large distances. These nontrivial space-time couplings occur notwithstanding the decoupling between space, time, and magnitude in the ''bare'' propagator, and are due to the existence of cascades of aftershocks.

A recent work of Krishnamurthy *et al.* [58] substantiates the general modeling strategy used here of representing the space-time dynamics of earthquakes by an effective stochastic process (the ETAS model) entirely defined by two exponents [corresponding to our μ and $H(\theta,\mu)$ defined below], where μ is the exponent of the power-law distribution of jumps between successive active sites and H is the (sub-) diffusion exponent. Indeed, Krishnamurthy *et al.* [58] show that the Bak and Sneppen model and the Sneppen model of extremal dynamics (corresponding to a certain class of self-organized critical behavior [12]) can be completely characterized by a suitable stochastic process called ''linear fractional stable motion.'' Beyond recovering the scaling exponents of this model, the stochastic process strategy predicts the conditional probabilities of successive activations at different sites and thus offers important insights. We note that this approach with the linear fractional stable motion is extremely close in spirit as well as in form to our approach mapping the ETAS model to the CTRW model. The ETAS model can thus be taken to represent an effective stochastic process of the complex self-organization of seismicity.

II. THE ETAS MODEL

A. Definitions and specific parametrization of the ETAS model

We assume that a given event (the ''mother'') of magnitude m_i occurring at time t_i and position $\vec{r}_i$ gives birth to other events (''daughters'') of any possible magnitude chosen with some independent Gutenberg-Richter distribution at a later time between t and $t+dt$ and at point $\vec{r}\pm\vec{dr}$ to within $\vec{dr}$ at the rate

$$\phi_{m_i}(t-t_i,\vec{r}-\vec{r}_i)=\rho(m_i)\Psi(t-t_i)\Phi(\vec{r}-\vec{r}_i). \tag{2}$$

We will refer to $\phi_{m_i}(t-t_i,\vec{r}-\vec{r}_i)$ both as the seismic rate induced by a single mother or as the ''bare propagator.'' It is the product of three independent contributions.

(1) $\rho(m_i)$ gives the number of daughters born from a mother with magnitude m_i. This term will, in general, be

chosen to account for the fact that large earthquakes have many more triggered events than have small earthquakes. Specifically, we take

$$\rho(m_i) = K 10^{\alpha(m_i - m_0)}, \tag{3}$$

which, as we said earlier, is justified by the power-law dependence of the volume of stress perturbation as a function of the earthquake size. α quantifies how fast the average number of daughters per mother increases with the magnitude of the mother.

(2) $\Psi(t-t_i)$ is a normalized waiting time distribution giving the rate of daughters born at time $t-t_i$ after the mother. The normalization condition reads $\int_0^{+\infty} dt \Psi(t) = 1$. $\Psi(t-t_i)dt$ can thus be interpreted as the probability for a daughter to be born between t and $t+dt$ from the mother that was born at time t_i. $\Psi(t-t_i)$ embodies Omori's law: it is the ''bare'' or ''direct'' Omori law,

$$\Psi(t) = \frac{\theta c^{\theta}}{(t+c)^{1+\theta}} H(t), \tag{4}$$

where $\theta > 0$ and $H(t)$ is the Heaviside function.

(3) $\Phi(\vec{r} - \vec{r}_i)$ is a normalized spatial ''jump'' distribution from the mother to each of her daughters, quantifying the probability for a daughter to be triggered at a distance $|\vec{r} - \vec{r}_i|$ from the mother. Specifically, we take

$$\Phi(\vec{r}) = \frac{\mu}{d\left(\frac{|\vec{r}|}{d} + 1\right)^{1+\mu}}, \tag{5}$$

which has the form of an (isotropic) elastic Green function dependence describing the stress transfer in an elastic upper crust. The exponent μ is left adjustable to account for heterogeneity and the possible complex modes of stress transfers. The normalization condition reads $\int d\vec{r} \Phi(\vec{r}) = 1$, where the integral is carried out over the whole space.

The physical justification for this decoupled model (2) in which $\phi_{m_i}(t-t_i, \vec{r} - \vec{r}_i)$ is the product of three independent distributions is that elastic waves propagate at kilometers per second and thus almost instantaneously reset the stress field after a large main shock. In other words, there is a well-defined separation of time scales between the time of propagation of seismic waves (seconds to minutes) which control the convergence to a new mechanical equilibrium after the main shock and the time scales involved in aftershock sequences (hours, days, months, or many years). The spatial dependence in Eq. (2) reflects the stress redistribution. This new stress field then relaxes slowly and more or less independently from point to point leading to the local Omori law $\Psi(t-t_i)$. Notwithstanding this argument, the decoupling in Eq. (2) between the local responses in magnitudes, space, and time is mostly performed because of its simplicity. It constitutes an approximation that should be checked and relaxed in future studies.

We assume a distribution $P(m)$ of earthquake sizes expressed in magnitudes m which follows the Gutenberg-Richter distribution

$$P(m) = b \ln(10) 10^{-b(m-m_0)}, \tag{6}$$

with a b value usually close to 1. m_0 is a lower bound magnitude below which no daughter is triggered.

B. The branching ratio n

A key parameter of the ETAS is the average number n of daughter earthquakes created per mother event, summed over all possible magnitudes. As we shall see, it is also natural to call it the ''branching ratio.'' To see this, consider the integral of the seismic rate $\phi_{m_i}(t-t_i, \vec{r} - \vec{r}_i)$ induced by one earthquake over all times after t_i, over all spatial positions and over all magnitudes $m_i \geqslant m_0$, which must give, by definition, the average number n of direct (or primary) daughter earthquakes created per mother event independently of its magnitude. For $\alpha < b$, and using Eqs. (2), (3), and (6), it is exactly given by

$$n \equiv \int d\vec{r} \int_{t_i}^{+\infty} dt \int_{m_0}^{+\infty} dm_i P(m_i) \phi_{m_i}(t-t_i, \vec{r} - \vec{r}_i)$$
$$= \int_{m_0}^{+\infty} dm_i P(m_i) \rho(m_i) = \frac{Kb}{b-\alpha}, \tag{7}$$

since the two integrals over time and space contribute each a factor 1 by the normalization of Ψ and Φ. This result (7) is identical to that found in absence of spatial dependence of $\phi_{m_i}(t-t_i)$ with respect to $\vec{r} - \vec{r}_i$ due to the factorization of the rate ρ, time Ψ, and space Φ dependences [26]. The branching ratio has also been evaluated in the case where the magnitude distribution follows a gamma distribution [54].

We stress again that n is an *average* quantity that does not reflect the large fluctuations in the number of aftershocks from event to event. Indeed, large events with magnitudes M produce, in general, many more aftershocks than small events with magnitude $m < M$, simply because $\rho(M) \gg \rho(m)$ if $M > m$ [see the exponential dependence (3) of $\rho(m)$ on the magnitude m].

C. Numerical simulation of the spatial ETAS model

The ETAS model has been simulated numerically using the algorithm described in Refs. [59,52]. Starting with a large event of magnitude M at time $t=0$, events are then simulated sequentially. At any given time t, we calculate the conditional seismic rate $\lambda(t)$ defined by

$$\lambda(t) = \sum_{t_i \leqslant t} K 10^{\alpha(m_i - m_0)} \frac{\theta c^{\theta}}{(t - t_i + c)^{1+\theta}}, \tag{8}$$

where $K = n(b-\alpha)/b$, and t_i and m_i are the times and magnitudes of all preceding events that occurred at time $t_i \leqslant t$. Note that we use the bare propagator because the sum in Eq. (8) is performed exhaustively on the complete catalog of past events. The time of the following event is then determined

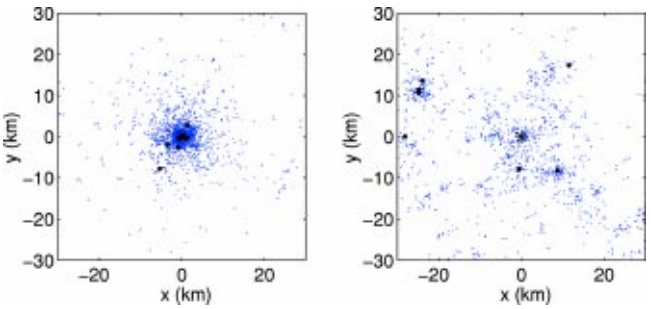

FIG. 1. Maps of seismicity generated by the ETAS model with parameters $b=1$, $\theta=0.2$, $\mu=1$, $d=1$ km, $\alpha=0.5$, $c=0.001$ day, and a branching ratio $n=1$. The mainshock occurs at the origin of space with magnitude $M=7$. The minimum magnitude is fixed at $m_0=0$. The distances between mainshock and aftershocks follow a power law with parameter $\mu=1$ and the local (or bare) Omori's law is $\propto 1/t^{1+\theta}$. According to the theory developed in the text, the average distance between the first mainshock and the aftershocks is thus expected to grow as $R\sim t^H$ with $H=0.2$ [Eq. (58)]. The two plots are for different time periods of the same numerical simulation, such that the same number of earthquakes $N=3000$ is obtained for each graph. (a) Time between 0 and 0.3 days; (b) time between 30 and 70 yr. Real aftershock sequences are indeed observed to last decades up to a century. Large black dots indicate large aftershocks around which other secondary aftershocks cluster. The mainshock is shown by a black star. At early times, aftershocks are localized close to the mainshock, and then diffuse and cluster around the largest aftershocks.

according to the nonstationary Poisson process of conditional intensity $\lambda(t)$, and its magnitude is chosen according to the Gutenberg-Richter distribution with parameter b. To determine the position in space of this new event, we first choose its mother randomly among all preceding events with a probability proportional to their rate of aftershocks $\phi_{m_i}(t-t_i)$ evaluated at the time of the new event. Once the mother has been chosen, we generate the distance r between the new earthquake and its mother according to the power-law distribution $\Phi(\vec{r})$ given by Eq. (5). The location of the new event is determined by assuming an isotropic distribution of aftershocks. By this rule, it is clear that new events tend to be close, in general, to the last large earthquakes, leading to space clustering.

Note that this two-steps procedure is equivalent to but more convenient for a numerical implementation than the one-step method, consisting of calculating at each point on a fine space-covering grid the seismic rate, equal to the sum over all preceding mothers weighted by the bare space $\Phi(\vec{r})$ and time $\Psi(t)$ propagators given by Eqs. (5) and (4); after normalizing, these rates then provide to each grid point a probability for the event to occur on that point. The equivalence between our two-step procedure and the direct calculation of the seismic rates is based on the law of conditional probabilities: [probability of next event (A)]=[probability of next event conditioned on its mother (event B)$\times$[probability of choosing the mother], i.e., $P(A,B)=P(A|B)P(B)$.

Figure 1 shows the result of a numerical simulation of the ETAS model which exhibits a diffusion of the seismic activity. We simulate a sequence of aftershocks and secondary

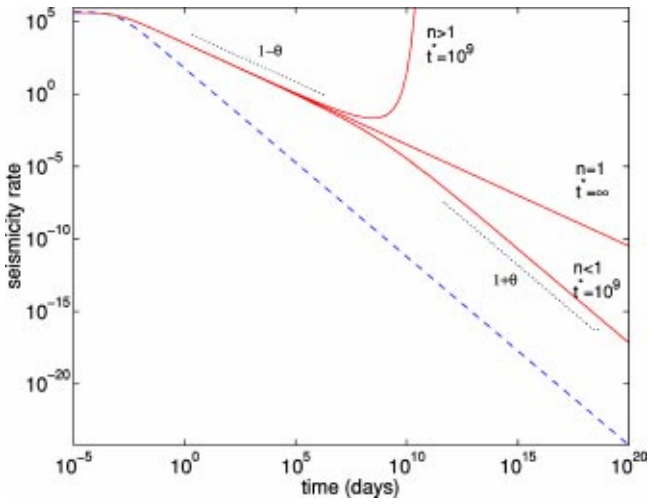

FIG. 2. Seismicity rate $N(t)$ for the temporal ETAS model calculated for $\theta=0.3$ and $c=0.001$ day. The local law $\phi(t)\propto 1/t^{1+\theta}$, which gives the probability distribution of times between an event and its (first-generation) aftershocks is shown as a dashed line. The global law $N(t)$, which includes all secondary and successive aftershocks generated by all the aftershocks of the first event, is shown as a solid line for the three regimes, $n<1$, $n=1$, and $n>1$. In the critical regime $n=1$, the seismicity rate follows a renormalized or dressed Omori law $\propto 1/t^p$ for $t>c$ with an exponent $p=1-\theta$, smaller than the exponent of the local law $1+\theta$. In the subcritical regime $(n<1)$, there is a crossover from an Omori law $1/t^{1-\theta}$ for $t<t^*$ to $1/t^{1+\theta}$ for $t>t^*$. In the supercritical regime $(n>1)$, there is a crossover from an Omori law $1/t^{1-\theta}$ for $t<t^*$ to an exponential increase $N(t)\sim\exp(t/t^*)$ for $t>t^*$. We have chosen on purpose values of $n=0.9997<1$ and $n=1.0003>1$ very close to 1 such that the crossover time $t^*=10^9$ days given by Eq. (1) is very large. In real data, such large t^* would be undistinguishable from an infinite value corresponding to the critical regime $n=1$. This representation is chosen for pedagogical purpose to make clear the different regimes occurring at times smaller and larger than t^*. In reality, we can expect n to be significantly smaller or larger than 1, such that t^* becomes maybe of the order of months, years, or decades and the observed Omori law will thus lie in the crossover regime, given an apparent Omori exponent anywhere from $1-\theta$ to $1+\theta$.

aftershocks starting from a mainshock of magnitude $M=7$, with the following parameters: $\theta=0.2$, $b=1$, $\alpha=0.5$, $n=1$, and $\mu=1$. At early times, aftershocks are localized close to the mainshock, and then diffuse and cluster close to the largest aftershocks. This (sub-)diffusion is extremely slow, as we shall quantify in the sequel. Our purpose is to provide a theory for this process based on the ETAS model. This theory will be tested by numerical simulations.

The different regimes are illustrated in Fig. 2, which shows the seismicity rate $N(t)$ for the temporal ETAS model studied in Refs. [25,26] obtained by summing the seismic activity over all space for the three cases $n<1$ (subcritical), $n=1$ (critical), and $n>1$ (supercritical). The subcritical regime is characterized by the existence of the time scale t^* given by Eq. (1). There is no difference between the critical case $n=1$ and the subcritical case for $t<t^*$ (see Fig. 2). Indeed, the difference between the subcritical regime and the critical regime can be observed only for $t>t^*$. A simple way

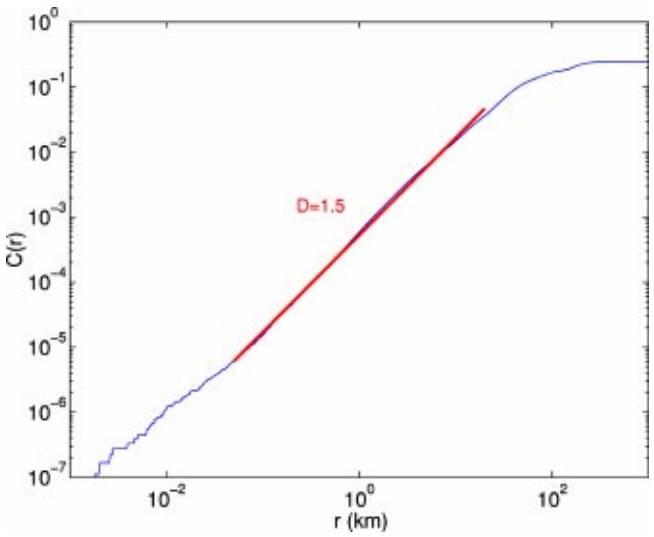

FIG. 3. Plot of the correlation function of the 3 000 epicenters generated in the time interval [30, 70] yrs and shown in the right panel of Fig. 1, calculated following Grassberger-Procaccia's algorithm, as a function of scale r, in double-logarithmic scales.

to see this is to realize that the critical regime $n=1$ gives $t^*=+\infty$, meaning that, in the critical regime, one is always in the situation $t<t^*$.

It is interesting to note that the spatial distribution of epicenters shown in the right panel of Fig. 1 has the visual appearance of a fractal set of points. This is confirmed by the calculation of the correlation dimension of this set of $N=3000$ points generated in the time interval [30, 70] yrs, which is found approximately equal to $D_2=1.5\pm0.05$ over more than two decades in spatial scales, as shown in Fig. 3. If we use instead all 30 000 events of the simulation performed up to time $t=70$ yr, we find $D_2=1.85\pm0.05$ while the correlation dimension of the geometrical set made of the epicenters of the 10 000 last events (time interval [7, 70] yr) is $D=1.7\pm0.05$, also over more than two decades in scale. These values are similar to those reported for two-dimensional maps of active fault systems [60–62,8], and are in good agreement with D_2 values in the range [1.65,1.95] measured for aftershocks epicenters [63]. The fractal clustering of the earthquake epicenters, according to the ETAS model, occurs because of a self-similar process taking place on many different scales. However, the description of this multiscale process solely in terms of a single fractal dimension fails to fully embody the complex spatial superposition of local ''singularities'' associated with each aftershock on the one hand and finite-size effects (stemming from the finite lifetime of each aftershock sequence) on the other hand. Each event indeed creates its cloud of direct aftershocks which can be characterized by its singular exponent $1-\mu$ for $\mu\leqslant1$ and 0 for $\mu>1$, defined by the scaling $\propto\int_0^R rdr/r^{1+\mu}\propto R^{1-\mu}$ of the ''mass'' of the cloud with its radius R. Finite-size effects and randomness have been documented to generate realistic but sometimes spurious fractal signatures [64–67]. This problem requires a special study which is left for another work.

D. Relationship with the space-independent ETAS model

The spatial ETAS model reduces to the space-independent ETAS model solved in Ref. [26] by integrating the dressed propagator obtained below over all space. In the Fourier representation [see expression (26)] this corresponds to putting the wave number k to zero. Indeed, for $k=0$, the Fourier transform amounts to performing a simple integration over all space. Since $\hat{\Phi}(\vec{k}=\vec{0})=1$, expression (26) derived below reduces to the form studied at length in Ref. [26]. Therefore, all results reported previously hold also for the version of the space-dependent ETAS model studied here, when averaging over the whole space. This is an important property that all the solutions discussed below must obey.

III. MAPPING OF THE ETAS MODEL ON THE CTRW MODEL

In order to study the space-time properties of the ETAS model, it is very useful to use a correspondence between the ETAS model and the CTRW that we establish here. In this way, we can adapt and use the wealth of results previously derived for the CTRW. But first, let us demonstrate the correspondence between the ETAS and CTRW models. For this, our strategy is to derive the master equations for both models and show that they are identical.

A. The master equation of the ETAS model

The ETAS model can be rephrased by defining the rate $\phi_{m_i\to m}(t-t_i,\vec{r}-\vec{r}_i)$ at which a given event (the ''mother'') of magnitude $m_i\geqslant m_0$ occurring at time t_i and position $\vec{r}_i$ gives birth to other events (''daughters'') of specified magnitude m at a later time between t and $t+dt$ and at point $\vec{r}$ to within an infinitesimal volume $|d\vec{r}|$. Note that the only difference with respect to the previous definition (2) is that we now specify also the magnitude m of the daughter. $\phi_{m_i\to m}(t-t_i,\vec{r}-\vec{r}_i)$ is given by

$$\phi_{m_i\to m}(t-t_i,\vec{r}-\vec{r}_i)=\rho(m_i\to m)\Psi(t-t_i)\Phi(\vec{r}-\vec{r}_i), \quad (9)$$

where $\Psi(t-t_i)$ and $\Phi(\vec{r}-\vec{r}_i)$ are the same as previously while

$$\rho(m_i\to m)=P(m)\rho(m_i). \quad (10)$$

With the parametrizations (3) and (6), this reads

$$\rho(m_i\to m)=n\ln(10)(b-\alpha)10^{\alpha(m_i-m_0)}10^{-b(m-m_0)}. \quad (11)$$

Let us consider the case where there is an origin of time $t=0$ at which we start recording the rate of earthquakes, assuming that a large earthquake has just occurred at $t=0$ and somehow reset the clock. In the following calculation, we will forget about the effect of events at times prior to $t=0$ and count all aftershocks that are created only by this main shock.

Let us call $N_m(t,\vec{r})dtdmd\vec{r}$ the number of earthquakes occurring between t and $t+dt$ of magnitude between m and $m+dm$ inside a box of volume $|d\vec{r}|$ centered at point $\vec{r}$. $N_m(t,\vec{r})$ is the solution of a self-consistency equation that formalizes mathematically the following process: an earthquake may trigger aftershocks; these aftershocks may trigger their own aftershocks, and so on. The rate of seismicity at a given time t and position $\vec{r}$ is the result of this cascade process. The self-consistency equation that sums up this cascade reads

$$N_m(t,\vec{r}) = S(t,\vec{r},m) + \int \vec{d}r' \int_{m_0}^{\infty} dm' \times \int_0^t d\tau \phi_{m'\to m}(t-\tau,\vec{r}-\vec{r}') N_{m'}(\tau,\vec{r}'). \quad (12)$$

The rate $N_m(t,\vec{r})$ at time t and position $\vec{r}$ is the sum over all induced rates from all earthquakes of all possible magnitudes that occurred at all previous times and locations propagated to the present time t and to the position $\vec{r}$ of observation by the corresponding bare propagator. The induced rate of events per earthquake that occurred at an earlier time τ and position $\vec{r}'$ is equal to $\phi_{m'\to m}(t-\tau,\vec{r}-\vec{r}')$. The source term $S(t,\vec{r})$ is the main shock plus the background seismicity, if any. In absence of background seismicity, a main earthquake that occurs at the origin of time $t=0$ at position $\vec{r}=\vec{0}$ with magnitude M gives

$$S(t,\vec{r},m) = \delta(t)\delta(m-M)\delta(\vec{r}), \quad (13)$$

where δ is the Dirac distribution. Other arbitrary source functions can be chosen.

The source term corresponding to a single mainshock is indeed the δ function (13) rather than the direct Omori law created by this mainshock in direct lineage. To see this, notice that the direct Omori law is recovered from Eq. (12) by replacing $N_{m'}(\tau,\vec{r}')$ in the integral by $S(t,\vec{r},m)$ given by Eq. (13). This shows that the difference between the renormalized and the direct Omori laws comes from taking into account the secondary, tertiary, etc., cascades of aftershocks.

As we have seen, a key assumption of the ETAS model is that the daughters born from a given mother have their magnitude drawn independently of the magnitude of the mother and of the process that give them birth, with a probability given by the Gutenberg-Richter distribution (6). The consequences of relaxing this hypothesis will be reported elsewhere. Keeping this assumption, it can be shown [68] that, for $\alpha \leqslant b/2$, an ensemble of realizations will obey

$$N_m(t,\vec{r}) = P(m)N(t,\vec{r}) \quad \text{for} \quad t>0, \quad (14)$$

which makes explicit the separation of the magnitude from the time and space variables. $N(t,\vec{r})$ is the number of events at position $\vec{r}$ at time t of any possible magnitude. Expression (14) means that the Gutenberg-Richter distribution is preserved at all times. That Eq. (14) holds for the ETAS model stems from the fact that the waiting time $\Psi(t)$ distribution (4) and jump size $\Phi(\vec{r})$ distribution (5) are independent of the magnitudes and that fluctuations in the seismicity rate are not too wild for $\alpha \leqslant b/2$. Note that, in a more complex model in which time, space, and magnitudes are interdependent, expression (14) would become a mean-field approximation, in which the fluctuations of the rates induced by the fluctuations of the realized magnitudes of the daughters factorize from the process.

Putting Eq. (14) in Eq. (12), for $t>0$ when the source term $S(t,\vec{r},m)$ is identically zero, one can simplify by $P(m)$ and obtain

$$N(t,\vec{r}) = \int d\vec{r}' \int_0^t d\tau \phi(t-\tau,\vec{r}-\vec{r}')N(\tau,\vec{r}'), \quad t>0, \quad (15)$$

where

$$\phi(t-\tau,\vec{r}-\vec{r}') = \int_{m_0}^{\infty} dm' P(m') \phi_{m'}(t-\tau,\vec{r}-\vec{r}'). \quad (16)$$

Equation (15) is nothing but the expectation (or statistical average, i.e., average over an ensemble of realizations) of expression (8), with the definition $N(t,\vec{r}) \equiv E[\lambda(t)\Phi(\vec{r})]$. Therefore, the master equation obtained here gives us only the first moment of the space-time dynamics of seismicity. It is not difficult to derive the equations for the variance and covariance of the seismic rate as well as higher moments.

The value of the source term at $t=0$ that should be incorporated in Eq. (15) requires more care. Indeed, a naive treatment would give a source term $\delta(t)\delta(m-M)\delta(\vec{r})/P(M)$ obtained by simply dividing by $P(m)$, expressed at $m=M$ due to the Dirac distribution $\delta(m-M)$. However, this source term still depends on m via the Dirac distribution $\delta(m-M)$ and is thus unsuitable as a source term of Eq. (15) which is independent of m. In order to circumvent this difficulty, one has to get rid of the Dirac distribution $\delta(m-M)$. The corresponding procedure has been described in details in Ref. [26] and consists in applying the integral operator $\int_{m_0}^{\infty} dm \hat{\phi}(\beta,\vec{r})$ to Eq. (12), where $\hat{\phi}(\beta,\vec{r})$ is the Laplace transform with respect to the time variable of $\phi(t,\vec{r})$. In this way, the Dirac distribution $\delta(m-M)$ is regularized. Identifying with the results of Ref. [26], we obtain that $N(t,\vec{r})$ is the solution of Eq. (15) with a source term

$$S_M(t,\vec{r}) = \delta(r)\delta(t)\rho(M)/n, \quad (17)$$

where $\rho(M)$ is defined in Eq. (3) and n is given by Eq. (7). Thus, the complete master equation for the number $N(t,\vec{r})$ of events at position $\vec{r}$ at time t of any possible magnitude is solution of

$$N(t,\vec{r}) = S_M(t,\vec{r}) + \int d\vec{r}' \int_0^t d\tau \phi(t-\tau,\vec{r}-\vec{r}')N(\tau,\vec{r}'), \quad t>0, \quad (18)$$

$N(t,\vec{r})$ is the ''dressed'' or ''renormalized'' propagator, obtained by summing the bare Omori propagator over all possible aftershock cascades. $N(t,\vec{r})$ can also be called the renormalized Omori law [25].

The essential assumption used to derive Eq. (12) is that the fluctuations of the earthquake magnitudes in a given sequence can be considered to be decoupled from those of the seismic rate. This approximation can be shown to be valid for $\alpha \leqslant b/2$ [68], for which the random variable $\rho(m_i)$ has a finite variance. In this case, any coupling between the fluctuations of the earthquake energies and the instantaneous seismic rate provides only subdominant corrections to Eq. (12). For $\alpha > b/2$, the variance of $\rho(m_i)$ is mathematically infinite or undefined as $\rho(m_i)$ is distributed according to a power law with exponent $b/\alpha < 2$. In this case, the master equation (12) is not completely correct as an additional term must be included to account for the effect of the dependence between the fluctuations of earthquake magnitudes and the instantaneous seismic rate. Our results are presented below for $\alpha = 0.5$, which belongs to the first regime $\alpha \leqslant b/2$. For $\alpha > b/2$, Ref. [68] has shown that the renormalization of the bare propagator into the dressed propagator is weaker than for $\alpha \leqslant b/2$, all the more so as $\alpha \to b$. Preliminary numerical simulations for $\alpha > b/2$ shows that our results presented below hold qualitatively but with a reduction of the observed spatial diffusion exponent compared to the value predicted from the master equation approach developed here. This regime $\alpha > b/2$ is probably relevant to the real seismicity [43,45,46], even if a precise estimation of α is very difficult.

B. A master equation of the CTRW model

We now demonstrate that the self-consistent mean field equation (18) is identical to the master equation of a CTRW. Random walks underlie many physical processes and are often the basis of first-order description of natural processes. The CTRW model, which is a generalization of the naive model of a random walker that jumps by ± 1 spatial step on a discrete lattice at each time step, was introduced by Ref. [57] and investigated by many other workers [69–73]. The CTRW considers a continuous distribution of spatial steps as well as time steps (which can be seen either as waiting times between steps or as durations of the steps). The CTRW model is thus based on the idea that the length of a given jump, as well as the waiting time $\tau_i = t_i - t_{i-1}$ elapsing between two successive jumps are drawn from a joint probability density function (PDF) $\phi(\vec{r},t)$, which is usually referred to as the jump PDF. From a mathematical point of view, a CTRW is a process subordinated to random walks under the operational time defined by the process $\{t_i\}$.

From $\phi(\vec{r},t)$, the jump length PDF $\Phi(\vec{r}) = \int_0^{+\infty} dt\, \phi(\vec{r},t)$ and the waiting time PDF $\Psi(t) = \int d\vec{r}\, \phi(\vec{r},t)$ can be deduced. Thus, $\Phi(\vec{r})d\vec{r}$ produces the probability for a jump length in the interval $(\vec{r},\vec{r}+d\vec{r})$ and $\Psi(t)dt$ the probability for a waiting time in the interval $(t,t+dt)$. When the jump length and waiting time are independent random variables, this corresponds to the decoupled form $\phi(\vec{r},t) = \Psi(t)\Phi(\vec{r})$. If both are coupled, a jump of a certain length involves a time cost or, vice versa in a given time span the walker can only travel a maximum distance. With these definitions, a CTRW process can be described through a master equation (see Refs. [73–75] for a review and references therein) which turns out to be given by an equation that is identical to Eq. (18).

This connection between the ETAS model of earthquakes and a model of random walks provides an important advance for the understanding of spatiotemporal earthquake processes, as it allows one to borrow the deep knowledge accumulated in past decades on random walks. In the same spirit, polymer physics acquired its status as a fundamental physical problem from its previous status of an applied field of research in chemistry when Flory, Edwards, de Gennes, des Cloizeaux, and others showed how to formulate problems in polymer physics in the language of random walks and how to extract novel results. In the sequel of this paper, we use this analogy to provide a wealth of predictions as well as important questions for earthquake aftershocks.

In the context of the CTRW, we have the following correspondence.

(a) $N(t,\vec{r})$ is the PDF for the random walker to just arrive at position $\vec{r}$ at time t.

(b) The source term $S_M(t,\vec{r})$ given by Eq. (17) denotes the initial condition of the random walk, here chosen to be at the origin of space at time $t=0$. The constant $\rho(M)/n$ adds the possibility via the parameter M to have more than one initial walker at the origin.

(c) In the CTRW context, the master equation (18) states that the PDF $N(t,\vec{r})$ of just having arrived at position $\vec{r}$ at time t comes from all possible paths in number $N(\tau,\vec{r}')$ having crossed a position $\vec{r}'$ at an earlier time τ, weighted by a transfer or propagator function $\phi(t-\tau,\vec{r}-\vec{r}')$ describing all the possible steps of the random walker from $(\tau,\vec{r}')$ to $(t,\vec{r})$.

It is important to stress that $N(t,\vec{r})$ defined above is different from the standard quantity $W(t,\vec{r})$ usually studied in random walk problems, defined as the probability to find the random walk at position $\vec{r}$ at time t. The relationship between $N(t,\vec{r})$ and $W(t,\vec{r})$ is

$$W(t,\vec{r}) = \int_0^t dt' \left[1 - \int_0^{t-t'} dt'' \Psi(t'') \right] N(t',\vec{r}). \qquad (19)$$

The term $1 - \int_0^{t-t'} dt'' \Psi(t'')$ in bracket is the probability for the walker not to jump in the time interval $[t',t]$ and the integral in the right-hand side of Eq. (19) means that the probability $W(t,\vec{r})$ for the random walker to be at position $\vec{r}$ at time t is the sum over all possible scenarios in which the walker just arrives at $\vec{r}$ at an earlier time t' and then does not jump until time t. In the context of earthquake aftershocks, $W(t,\vec{r})$ is the probability that an event at $\vec{r}$ has occurred at a time $t' \leqslant t$ and that the whole system has remained quiescent from t' to t.

In the Fourier-Laplace domain (see below), expression (19) reads

TABLE I. Correspondence between the ETAS (epidemic-type aftershock sequence) and CTRW (continuous-time random walk) models. ''PDF'' stands for probability density function.

	ETAS	CTRW
$\Psi(t)$	PDF for a ''daughter'' to be born at time t from the mother that was born at time 0	PDF of waiting times
$\Phi(\vec{r})$	PDF for a daughter to be triggered at a distance $\vec{r}$ from its mother	PDF of jump sizes
m	Earthquake magnitude	Tag associated with each jump
$\rho(m)$	Number of daughters per mother of magnitude m	Local branching ratio
n	Average number of daughters created per mother summed over all possible magnitudes	Control parameter of the random walk survival (branching ratio)
$n<1$	Subcritical aftershock regime	Subcritical ''birth and death''
$n=1$	Critical aftershock regime	The standard CTRW
$n>1$	Supercritical exponentially growing regime	Explosive regime of the ''birth and death'' CTRW
$N(t,\vec{r})$	Number of events of any possible magnitude at $\vec{r}$ at time t	PDF of just having arrived at $\vec{r}$ at time t
$W(t,\vec{r})$	PDF that an event at $\vec{r}$ has occurred at a time $t' \leqslant t$ and that no event occurred anywhere from t' to t	PDF of being at $\vec{r}$ at time t

$$\hat{W}(\beta,\vec{k}) = \frac{1-\hat{\Psi}(\beta)}{\beta}\hat{N}(\beta,\vec{k}). \tag{20}$$

In general, the CTRW model models transport phenomena in any heterogeneous media. It has, for instance, been used successfully for describing the behavior of chemical species as they migrate through porous media [76,77]. In insight, it is rather natural that it can be applied to the ''transport of stress'' through the heterogeneous crust and thus to the description of the anomalous diffusion of seismic activity.

Table I synthesizes the correspondence between the ETAS and CTRW models and then draws its consequences.

C. Experimental verifications of the crossover between the two power-law Omori decays in photoconductivity in amorphous semiconductors and in fractal stream chemistry using the correspondence between the ETAS and CTRW models

The crossover from an Omori law $1/t^{1-\theta}$ for $t<t^*$ to $1/t^{1+\theta}$ for $t>t^*$ found in Refs. [25,26] with t^* given by Eq. (1) has actually a counterpart in the CTRW. This behavior was first studied by Scher and Montroll [70] in a CTRW with absorbing boundary condition to model photoconductivity in amorphous semiconductors As_2Se_3 and an organic compound finding $\theta \approx 0.5$ and $\theta = 0.8$, respectively. In a semiconductor experiment, electric holes are injected near a positive electrode and then transported to a negative electrode where they are absorbed. The transient current follows exactly the transition $1/t^{1-\theta}$ for $t<t^*$ to $1/t^{1+\theta}$ for $t>t^*$ found for Omori's law for earthquake aftershocks in the ETAS model. In the semiconductor context, the finiteness of t^* results from the existence of a force applied to the holes while in the ETAS model it results from a finite distance $1-n$ to the critical point $n=1$ in the subcritical regime. When the force goes to zero or $n \to 1$, $t^* \to +\infty$.

A similar transition has been recently proposed to model long-term time series measurements of chloride, a natural passive tracer, in rainfall and runoff in catchments [79]. The quantity analogous to the dressed Omori propagator is the effective travel time distribution $h(t)$ which governs the global lag time between injection of the tracer through rainfall and outflow to the stream. $h(t)$ has been shown to have a power-law form $h(t) \sim 1/t^{1-m}$ with m between -0.3 and 0.2 for different time series [80]. This variability may be due to the transition between an exponent $1-\theta$ at short times to $1+\theta$ at long times [79], where θ is the exponent of the bare distribution of individual transition times.

D. General and formal solution of the spatial ETAS model

Let us solve Eq. (18) for the number $N(t,\vec{r})$ of events at position $\vec{r}$ at time t of any possible magnitude. Recall that $N(t,\vec{r})$ can also be interpreted as the dressed Omori propagator. Extending Ref. [26] to the spatial domain and also in analogy with the standard approach to solve the CTRW, the Laplace-in-time Fourier-in-space transform $\hat{N}(\beta,\vec{k})$ of $N(t,\vec{r})$ is given by

$$\hat{N}(\beta,\vec{k})=\frac{\hat{S}_M(\beta,\vec{k})}{1-n\hat{\Psi}(\beta)\hat{\Phi}(\vec{k})}, \tag{21}$$

where $\hat{S}_M(\beta,\vec{k})$ is the Laplace-Fourier transform of the source $S_M(t,\vec{r})$ given by Eq. (17) and $\hat{\Psi}(\beta)[\hat{\Phi}(\vec{k})]$ is the Laplace (Fourier) transforms of $\Psi(t)$ $[\Phi(\vec{r})]$. For a mainshock of magnitude M occurring at time $t=0$ and position $\vec{r}=0$, the source term is thus $\hat{S}_M(\beta,\vec{k})=\rho(M)/n$. The only difference between expression (21) and the Laplace-Fourier transform of the PDF of the CTRW of just having arrived at $\vec{r}$ at time t occurs when the branching ratio n is different from 1. In general, solutions of CTRW models are expressed for $n=1$ and for the variable $W(t,\vec{r})$ which is simply related to $N(t,\vec{r})$ according to Eq. (19). Using Eqs. (19) and (21) leads to

$$\hat{W}(\beta,\vec{k})=\frac{1-\hat{\Psi}(\beta)}{\beta}\frac{\hat{S}_M(\beta,\vec{k})}{1-n\hat{\Psi}(\beta)\hat{\Phi}(\vec{k})}. \tag{22}$$

In the following, we exploit Eq. (22) to obtain analytical solutions of the spatial ETAS model in different regimes, that provide specific predictions on the conditions necessary for observing aftershock diffusion. In addition, we provide specific predictions on the exponent H of the diffusion law $R\sim t^H$ that are tested by numerical simulations.

IV. CRITICAL REGIME $n=1$

A. Classification of the different regimes

Numerous works on the CTRW have investigated many possible forms for $\Psi(t)$ and $\Phi(\vec{r})$ and have provided the asymptotic long time and large scale dependence of $W(t,\vec{r})$ (see Refs. [73–75,77] and references therein). Here, we restrict our discussion to the cases where both $\Psi(t)$ and $\Phi(\vec{r})$ have power-law tails as given by Eqs. (4) and (5). The long-time and large scale behavior of the ETAS and CTRW are controlled by the behavior of the Laplace-Fourier transforms for small β and small $|\vec{k}|$.

Two cases must be distinguished depending on the exponent μ controlling the weight of the tail of $\Phi(\vec{r})$.

For $\mu>2$, the variance $\langle(\vec{r})^2\rangle=\sigma^2$ of the jump size distribution exists. To leading order in $k=|\vec{k}|$, $\hat{\Phi}(\vec{k})$ can be expanded as

$$\hat{\Phi}(\vec{k})=1-\sigma^2k^2+O(k^o) \quad \text{with} \quad o>2. \tag{23}$$

For $\mu\leqslant 2$, the variance $\langle(\vec{r})^2\rangle$ is infinite. This regime of ''long jumps'' leads to so-called Lévy flights. In this case, to leading order in $k=|\vec{k}|$, $\hat{\Phi}(\vec{k})$ can be expanded as

$$\hat{\Phi}(\vec{k})=1-\sigma^\mu k^\mu+\mathcal{O}(k^o), \quad \text{where} \quad 0<\mu\leqslant 2 \quad \text{with} \quad o>\mu, \tag{24}$$

where σ is a characteristic distance defined by

$$\sigma=\begin{cases} d[\Gamma(1-\mu)]^{1/\mu}, & 0<\mu<1 \\ \dfrac{d\pi}{\mu\Gamma(\mu-1)\sin(\pi\mu/2)}, & 1<\mu<2. \end{cases} \tag{25}$$

For a distribution $\Psi(t)$ of waiting times of the form of a local Omori law (4) with exponent $\theta<1$, $\hat{\Psi}(\beta)$ can be expanded for small β as

$$\hat{\Psi}(\beta)=1-(\beta c')^\theta+\mathcal{O}(\beta^\omega) \quad \text{with} \quad \omega\geqslant 1, \tag{26}$$

where c' is proportional to c up to a numerical constant $c'=c(\Gamma(1-\theta))^{1/\theta}$ in the case $\theta<1$.

Putting the leading terms of the expansions of $\hat{\Phi}(\vec{k})$ for small $|\vec{k}|$ and of $\hat{\Psi}(\beta)$ for small β in Eq. (21) gives

$$\hat{N}(\beta,\vec{k})=\frac{\hat{S}_M(\beta,\vec{k})}{1-n+n(\beta c')^\theta+n\sigma^\mu k^\mu}. \tag{27}$$

The corresponding $\hat{W}(\beta,\vec{k})$ is obtained from Eq. (22) by

$$\hat{W}(\beta,\vec{k})=\hat{S}_M(\beta,\vec{k})\frac{(\beta)^{\theta-1}c'^{\,\theta}}{1-n+n(\beta c')^\theta+n\sigma^\mu k^\mu}. \tag{28}$$

The critical regime $n=1$ gets rid of the constant term $1-n$ in the denominator of Eqs. (27) and (28). This case is analyzed in details below.

The regime $n\neq 1$ introduces a characteristic time t^* given by Eq. (1). In the subcritical regime, Eq. (27) can be rewritten as

$$\hat{N}(\beta,\vec{k})=\frac{\hat{S}_M(\beta,\vec{k})}{1-n}\frac{1}{1+(\beta t^*)^\theta+(kr^*)^\mu}, \tag{29}$$

where r^* is defined by

$$r^*=\sigma\left(\frac{n}{1-n}\right)^{1/\mu}. \tag{30}$$

For $t<t^*$ and $r<r^*$, the dressed propagator is given by the same expression as for the critical case and all our results below hold. For large times $t>t^*$ and large distances $r>r^*$, we can factorize Eq. (29) as a product of a function of time and a function of space,

$$\hat{N}(\beta,\vec{k})\simeq\frac{\hat{S}_M(\beta,\vec{k})}{1-n}\frac{1}{1+(\beta t^*)^\theta}\frac{1}{(1+(kr^*)^\mu}. \tag{31}$$

Thus, there is no diffusion in the subcritical regime for $t>t^*$ and $r>r^*$. We shall not analyze further this trivial regime $n<1$ and $t>t^*$ and will only analyze the case $t<t^*$. If there is the need, the crossover can be calculated explicitly using Eq. (27).

In order to get the leading behavior of $N(t,\vec{r})$ from that of $W(t,\vec{r})$, we see from Eqs. (21) and (22) that $\hat{N}(\beta,\vec{k})=\{\beta/[1-\hat{\Psi}(\beta)]\}\hat{W}(\beta,\vec{k})\approx\beta^{1-\theta}c'^{\,-\theta}\hat{W}(\beta,\vec{k})$. The inverse Laplace transform of $1/\beta^\theta$ is $1/[\Gamma(\theta)t^{1-\theta}]$. Using the fact

that the Laplace transform of df/dt is β times the Laplace transform of $f(t)$ minus $f(0)$, we get $N(t,\vec{r})$ as the derivative of a convolution

$$N(t,\vec{r}) = \frac{c'^{-\theta}}{\Gamma(\theta)} \frac{d}{dt} \int_0^t dt' \frac{W(t',\vec{r})}{(t-t')^{1-\theta}} = c'^{-\theta}_0 D_t^{1-\theta} W(t,\vec{r}). \tag{32}$$

In Eq. (32), we have dropped the Dirac function coming from the inverse Laplace transform of the constant term $f(0)$, which provides a contribution only at the origin of time $t=0$. Note that the operator $[1/\Gamma(\theta)] \times (d/dt) \int_0^t dt' [W(t',\vec{r})/(t-t')^{1-\theta}]$ is nothing but the so-called fractional Riemann-Liouville derivative operator of order $1-\theta$ applied to the function $W(t,\vec{r})$ of time t and is usually denoted ${}_0D_t^{1-\theta} W(t,\vec{r})$.

B. The standard diffusion case $\theta > 1$ and $\mu > 2$

The standard diffusion process is recovered for $\theta \geqslant 1$ (for which the average waiting time is finite) and for $\mu \geqslant 2$ (for which the variance of the jump length is finite). In this case, $\hat{N}(\beta,\vec{k}) = \hat{S}_M(\beta,\vec{k})/(\beta c' + \sigma^2 k^2)$. For an impulsive source leading to $\hat{S}_M(\beta,\vec{k}) = \text{const}$, this is the Laplace-Fourier transform of the standard diffusion propagator

$$N(t,\vec{r}) \propto \frac{1}{(Dt)^{d/2}} \exp[-(\vec{r})^2/Dt] \quad \text{where} \quad D = \sigma^2/c', \tag{33}$$

where d is here the space dimension. This solution is valid for $|\vec{r}|/\sqrt{Dt}$ not too large. For larger values, large deviations lead to corrections with the power-law tail of the input jump distribution $\Phi(\vec{r}) \sim 1/|\vec{r}|^{1+\mu}$ defined in Eq. (5), along the lines presented, for instance, in Ref. [12] (Sec. 3.5). This regime is not relevant to the aftershock problem for which usually $0 < \theta < 1$.

C. Long waiting times ($\theta < 1$) and finite variance of the jump sizes ($\mu > 2$)

Putting the leading terms of the expansions of $\hat{\Phi}(\vec{k})$ (23) and of $\hat{\Psi}(\beta)$ (26) in Eq. (21) gives

$$\hat{N}(\beta,\vec{k}) = \frac{1}{(\beta c')^\theta + (\sigma k)^2}. \tag{34}$$

The expression (34) can be inverted with respect to the Fourier transform, and then inverted with respect to the Laplace transform using Fox functions [75,81]. The solution for $W(t,\vec{r})$ in one dimension is given, for instance, in Ref. [75] in terms of an infinite sum

$$W(t,\vec{r}) = \frac{1}{2D} \frac{1}{t^{\theta/2}} \sum_{k=0}^{\infty} \frac{(-1)^k z^{-k}}{k! \Gamma[1-\theta(k+1)/2]}, \tag{35}$$

where

$$z = \frac{Dt^{\theta/2}}{|\vec{r}|} \tag{36}$$

and $D = \sigma/c'^{\theta/2}$.

Expression (35) and many others below involve the Γ function of negative arguments. We recall that the function $\Gamma(u)$ can be analytically continued to the whole complex plane, except for the simple poles $u = 0, -1, -2, -3, \ldots$. Thus, $\Gamma(u)$ is defined everywhere but at these poles. In order to get the expression of the Γ function for negative arguments, one can use two formulas: $\Gamma(1-u) \times \Gamma(u) = \pi/\sin(\pi u)$ and $\Gamma(1+u) = u\Gamma(u)$. Both these formulas are valid for all points with the possible exception of the arguments at poles $0, -1, -2, \ldots$. For instance, $\Gamma(-\theta) = \Gamma(1-\theta)/(-\theta) = -[\pi/\theta \sin(\pi\theta)]/\Gamma(\theta)$, for $0 < \theta < 1$.

Expression (35) can be rewritten as a Fox function [82],

$$W(t,z) = \frac{1}{2D} \frac{1}{t^{\theta/2}} H_{1,1}^{1,0}\left[\frac{1}{z} \middle| \begin{matrix} (1-\theta/2, \theta/2) \\ (0,1) \end{matrix} \right], \tag{37}$$

whose asymptotic dependence for large z, obtained from a standard theorem of the Fox function [Eq. (1.6.3) of Ref. [82]], is

$$W(t,z) \sim \frac{1}{Dt^{\theta/2}} \frac{1}{z^{(1-\theta)/(2-\theta)}} \times \exp\left[-\left(1 - \frac{\theta}{2}\right) \left(\frac{\theta}{2}\right)^{\theta/(2-\theta)} z^{2/(2-\theta)} \right] \tag{38}$$

is in agreement with the result of Roman and Alemany [83] and Barkai *et al.* [81] for a space dimension $d_f = 1$, including the dependence in the power law prefactor to the exponential. The exponential dependence $W(t,r) \sim \exp[-\text{const}(r/Dt^{\theta/2})^{2/(2-\theta)}]$ in Eq. (38) holds in arbitrary dimensions d_f, the only modification occurring in the prefactor whose power of z change with the space dimension d_f as [83,81]

$$W_{d_f}(t,z) \sim \frac{1}{Dt^{\theta/2}} \frac{1}{z^{d_f(1-\theta)/(2-\theta)}} \times \exp\left[-\left(1 - \frac{\theta}{2}\right) \left(\frac{\theta}{2}\right)^{\theta/(2-\theta)} z^{2/(2-\theta)} \right). \tag{39}$$

The expression of $N(t,\vec{r})$ can be obtained from $W(t,\vec{r})$ using the fractional Riemann-Liouville derivation (32) of order $[1-\theta]$. Inserting expression (35) in Eq. (32) and using the expression of the fractional Riemann-Liouville derivative operator ${}_0D_t^\alpha$ applied to an arbitrary power t^μ, i.e., ${}_0D_t^\alpha t^\mu = [\Gamma(1+\mu)/\Gamma(1+\mu-\alpha)] t^{\mu-\alpha}$, we obtain

$$N(t,\vec{r}) = \frac{c'^{-\theta}}{2Dt^{1-(\theta/2)}} \sum_{k=0}^{\infty} \frac{(-1)^k z^k}{k! \Gamma[(1-k)\theta/2]}. \tag{40}$$

Expression (40) can be used to evaluate $N(t,\vec{r})$ for small z, but the numerical evaluation of Eq. (40) is impossible for

large z. In order to obtain the asymptotic behavior of $N(t,\vec{r})$, expression (40) can be rewritten as a Fox function [82],

$$N(t,\vec{r}) = \frac{c'^{-\theta}}{2Dt^{1-(\theta/2)}} H^{1,0}_{1,1}\left[\frac{1}{z}\middle|\begin{matrix}(\theta/2,\theta/2)\\(0,1)\end{matrix}\right]. \tag{41}$$

Employing again the standard theorem of the Fox function [Eq. (1.6.3) of Ref. [82]), the asymptotic behavior of $N(t,r)$ for large distances r such that $r > Dt^{\theta/2}$ is given by

$$N(t,r) \sim \frac{c'^{-\theta}}{Dt^{1-(\theta/2)}} \left(\frac{|\vec{r}|}{Dt^{\theta/2}}\right)^{(1-\theta)/(2-\theta)} \times \exp\left[-\left(1-\frac{\theta}{2}\right)\left(\frac{\theta}{2}\right)^{\theta/(2-\theta)}\left(\frac{|\vec{r}|}{Dt^{\theta/2}}\right)^{2/(2-\theta)}\right]. \tag{42}$$

The exponential dependence $N(t,r) \sim \exp[-\text{const}(r/Dt^{\theta/2})^{2/2-\theta}]$ in Eq. (42) holds in arbitrary dimensions.

This expression becomes incorrect for very large distances because it would predict an exponential or slightly superexponential decay with r. This cannot be true as the global law cannot decay faster than the local law (5). The reason for Eq. (42) to become incorrect at large distances is that the expansion of $\hat{N}(\beta,\vec{k})$ for small $|\vec{k}|$ (large distances) given by Eq. (34) has been truncated at the order k^2. There is, however, a subdominant term $\propto k^{\mu}$ that describes the power-law tail of the local law (5) and also of the global law asymptotically. A similar situation occurs in the application of the central limit theorem for sums of N random variables with power-law distributions with exponents $\mu > 2$ [12]: the distribution of the sum S is a Gaussian in its bulk for $|S| < \sqrt{N \ln N}$ and crosses over to a power law with tail exponent μ for larger S. In a similar way, the crossover of $N(t,r)$ to the asymptotic local power law (5) can be recovered by an analysis including the subleading correction $\propto k^{\mu}$ to the expansion (34).

Expression (40) shows that the global rate of seismicity cannot be factorized as a product of a distribution of times and a distribution of distances. This space-time coupling implies that the seismic activity diffuses with time, and that the decay of the rate of aftershocks depends on the distance from the first mainshock. This coupling of space and time stems from the cascade of aftershocks, from the primary aftershocks to the secondary aftershocks to the tertiary aftershocks, and so on.

Figure 4 presents the decay of the seismic activity $N(r,t)$ obtained using expression (40) for small z and expression (42) for large z, as a function of the time from the mainshock and as a function of the distances r. Close to the mainshock epicenter, expression (40) predicts that the global seismicity rate decays with time as the renormalized Omori law

$$N(t,0) \sim \frac{1}{t^{1-\theta/2}}. \tag{43}$$

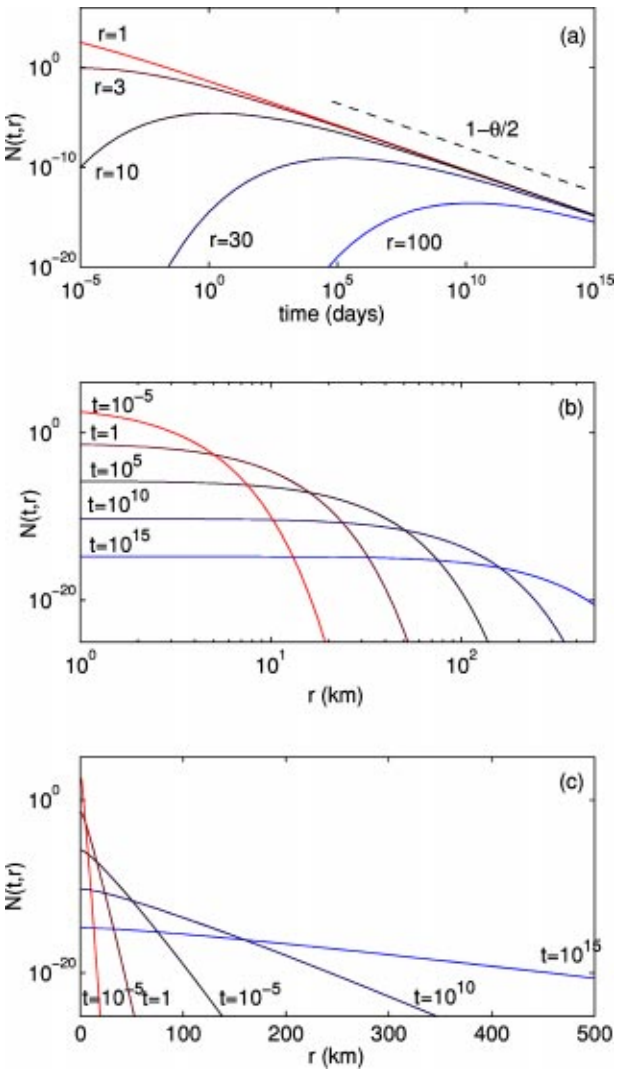

FIG. 4. Rate of seismicity $N(t,r)$ in the critical regime $n=1$ for $\theta=0.2$, $\mu>2$, $c'=1$ day, and $\sigma=1$ km, evaluated from expressions (40) and (42), plotted as a function of the time (a) for different values of the distance r between the mainshock and its aftershocks, and (b),(c) as a function of r [logarithmic scale for r in (b) and linear scale for r in (c)] for different values of the time between the mainshock and its aftershocks. The temporal decay of seismicity with time is characterized by a power-law decay $N(r,t) \sim 1/t^{1-\theta/2}$ close to the mainshock epicenter or at large times for $r \ll Dt^{\theta/2}$. For large distances $r \gg Dt^{\theta/2}$, there is a truncation of the power-law decay at early times $t^{\theta/2} \ll r/D$, because the seismicity has not yet diffused up to the distance r. Although the distribution of distances between a mainshock and its direct aftershocks $\Phi(r)$ follows a power-law distribution with exponent $1+\mu$, the log-linear graph (c) shows that the global rate of aftershocks $N(\vec{r},t)$ decreases approximately exponentially as a function of the distance from the mainshock, with a characteristic distance that increases with time. This is in agreement with expression (42) that predicts $N(t,r) \sim \exp[(|\vec{r}|/Dt^{\theta/2})^{2/(2-\theta)}]$, i.e., $N(t,r) \sim \exp[C(t)|\vec{r}|^q]$ with an exponent $q=2/(2-\theta)$ close to 1 within the exponential. The same remark as for Fig. 2 applies: the representation of our predictions for very large times is made for pedagogical purpose to illustrate clearly the different regimes.

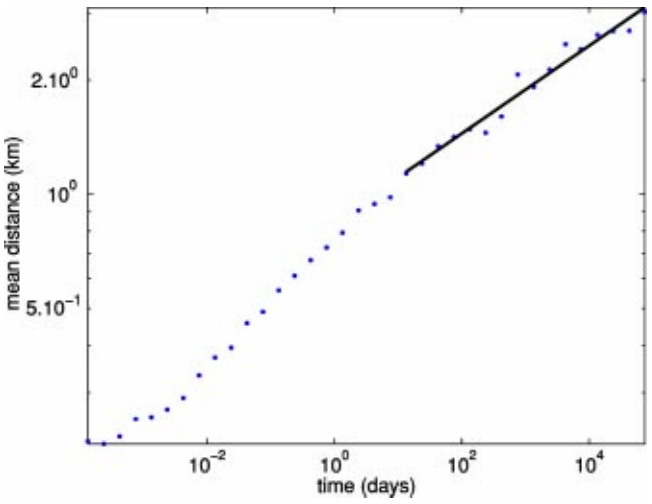

FIG. 5. Average distance between the first mainshock and its aftershocks as a function of the time from the mainshock, for numerical simulations of the ETAS model in the critical regime $n=1$, generated with the parameters $\theta=0.2$, $d=1$ km, $\mu=3$, and $c=10^{-3}$ day. The theoretical prediction for the diffusion exponent is thus $H=\theta/2=0.1$. We observe a crossover from a larger exponent at early times when the mean distance is close to the characteristic scale $d=1$ km of the distribution of distances between an aftershock and its progenitor, to a subdiffusion with an exponent close to the theoretical prediction at large times. The solid line is a fit of the numerical data for times $t>10$ days, which gives an exponent $H=0.12$ slightly larger than the predicted value $H=0.1$.

The same decay is found at any fixed point $\vec{r}$ for times $t>(|\vec{r}|/D)^{2/\theta}$. At all times, the same decay $1/t^{1-\theta/2}$ is also obtained by measuring the aftershock seismicity in a local box at a distance from the main shock origin increasing with time as $r\sim t^{\theta/2}$ [this is nothing but putting $z=$const in Eq. (40)]. At large distances $r>Dt^{\theta/2}$, the global decay law is different from a power-law decay. Figure 4 shows that the rate of aftershocks presents a truncation at early times, which increases as the distance r increases. At large times, the rate of aftershocks recovers the $1/t^{1-\theta/2}$ power-law decay (43). We stress that a fit of the global law $N(r,t)$ over the whole time interval by an Omori law would yield an apparent exponent $p<1-\theta/2$ that decreases with r.

Integrating Eq. (40) over the whole one-dimensional space, we recover the global Omori law,

$$N(t)=\int drN(t,r)\sim\frac{1}{t^{1-\theta}} \tag{44}$$

found in Refs. [25,26]. Thus, we have found an additional source of variability of the exponent p of the Omori law: if measured over the whole catalog, we should measure $p=1-\theta$ in the critical regime $n=1$ while $p=1-\theta/2$ is slightly larger when measured in certain time and space-windows, as described above. Thus, in this regime, pruning of catalogs may lead to continuous change from the value $1-\theta$ to $1-\theta/2$. In addition, as we have mentioned, the crossover in time may lead to still smaller apparent exponents, thus enhancing the impression of variability of the exponent p. In reality, this range of p values are seen to result from the complex spatiotemporal organization of the aftershock seismicity of the ETAS model. These results should lead us to be cautious when analyzing real catalogs with respect to the conditions and regimes under which the analysis is performed.

There is another observable that characterizes how an aftershock sequence invades space as a function of time. Expression (40) indeed predicts a subdiffusion process quantified by

$$\langle|\vec{r}|^2\rangle\sim t^{2H}, \tag{45}$$

with $H=\theta/2$ since the natural variable is z given by Eq. (36). Indeed, expression (40) tells us that, up to a global rescaling function of time, the rate of aftershocks is identical for a fixed value of z. Thus, any aftershock structure diffuses according to Eq. (45).

This prediction is checked in Fig. 5 by numerical simulations. 1000 synthetic catalogs have been generated with $\mu=3$, $\theta=0.2$, and $n=1$. The average distance between the first mainshock and its aftershocks as a function of the time from the mainshock has been averaged over these 1000 simulations. The theoretical diffusion exponent is $H=\theta/2=0.1$, in good agreement with the asymptotic behavior observed in the numerical simulation. In practice, in order to minimize the effect of fluctuations and optimize the speed of convergence, we estimate numerically $\exp[\langle\ln|\vec{r}|\rangle]$ which is also expected to scale as $\exp[\langle\ln|\vec{r}|\rangle]\sim t^{\theta/2}$ due to the simple scaling form of Eq. (41).

This problem has also been solved exactly in Ref. [84] in the context of the so-called fractional Fokker-Planck equation, which amounts to replacing the distribution $\Phi(\vec{r})$ of jumps (5) by a Gaussian function. This fractional Fokker-Planck equation allows one to introduce the possibility of bias or drift in the CTRW and therefore in the aftershock sequence.

D. Exponential waiting time distribution and long jump size Lévy distribution ($\mu<2$)

This case with an exponential distribution

$$\Psi(t)=\lambda e^{-\lambda t} \tag{46}$$

of waiting times with a Lévy distribution $\Phi(\vec{r})=L_\mu(|\vec{r}|)$ of jump sizes with tail exponent $\mu<2$ has been investigated by Budde *et al.* [85]. One finds

$$\langle|\vec{r}|^2\rangle^{1/2}\sim t^{1/\mu}, \tag{47}$$

corresponding to a superdiffusion regime with Hurst exponent $H=1/\mu>1/2$. The full distribution function $W(t,\vec{r})$ corresponding to the critical regime $n=1$ is known for $\lambda t\gg 1$,

$$W(t,\vec{r})\propto\frac{1}{(\lambda t)^{1/\mu}}L_\mu\left(\frac{|\vec{r}|}{(\lambda t)^{1/\mu}}\right). \tag{48}$$

The corresponding $N(t,\vec{r})$ is obtained from Eq. (20). The Laplace transform of the exponential distribution (46) is $\hat{\Psi}(\beta)=\lambda/(\beta+\lambda)$. We thus get

$$\hat{N}(\beta,\vec{k})=(\beta+\lambda)\hat{W}(\beta,\vec{k}), \tag{49}$$

and thus

$$N(t,\vec{r})=\frac{\partial W(t,\vec{r})}{\partial t}+\lambda W(t,\vec{r}). \tag{50}$$

Expression (50) together with Eq. (48) predicts a diffusion law $r\sim t^H$ with $H=1/\mu$ which is in good agreement with our simulations. At large times $|\vec{r}|\ll(\lambda t)^{1/\mu}$, $N(t,\vec{r})\approx\lambda W(t,\vec{r})\sim 1/t^{1/\mu}$, giving an apparent local Omori exponent $\theta=1-1/\mu$. This offers a different mechanism for generating Omori's law for aftershocks from purely exponential local relaxation but with a heavy distribution of jump sizes. This power-law decay should be observed only at a fixed distance r or over a limited domain from the mainshock in the regime of large times.

Integrating over the whole space, $\int d\vec{r}W(t,\vec{r})=1$, which gives $N(t)=\delta(t)+\lambda$ equal to a constant seismic rate. This results from an initial mainshock at $t=0$ leading to the cascade of aftershocks adjusting delicately to this constant rate for the critical value $n=1$ of the branching parameter. In the subcritical regime $n<1$, the Omori law integrated over space gives instead $N(t)\propto\exp[-(1-n)\lambda t]$, showing that the characteristic decay time $1/(1-n)\lambda$ of the dressed Omori propagator $N(t)$ becomes much larger (much longer memory) that the decay time $1/\lambda$ of the bare Omori propagator.

For $\mu>2$, we recover the standard diffusion corresponding to $\theta>1$ and $\mu>2$ discussed in Sec. IV B.

E. Long waiting times ($\theta<1$) and long jump sizes (Lévy flight regime for $\mu\leqslant 2$)

Putting the leading terms of the expansions of $\hat{\Phi}(\vec{k})$ and of $\hat{\Psi}(\beta)$ in Eq. (21) gives

$$\hat{N}(\beta,\vec{k})=\hat{S}_M(\beta,\vec{k})\frac{1}{(\beta c')^{\theta}+(\sigma k)^{\mu}}. \tag{51}$$

The corresponding $\hat{W}(\beta,\vec{k})$ is given by

$$\hat{W}(\beta,\vec{k})=\hat{S}_M(\beta,\vec{k})\frac{(\beta)^{\theta-1}c'^{\theta}}{(\beta c')^{\theta}+(\sigma k)^{\mu}}. \tag{52}$$

Equation (52) has been studied extensively in the context of the CTRW model as a long wavelength $|\vec{k}|\to 0$ and long time $\beta\to 0$ approximation to investigate the long time behavior of the CTRW. Kotulski [86] has developed a rigorous approach, based on limit theorems, to classify the asymptotic behaviors of different type of CTRWs and justifies the approximation (52) for the long time behavior. Barkai [87] has studied the quality of the long wavelength $|\vec{k}|\to 0$ and long time $\beta\to 0$ approximation (52) by solving the exact CTRW problem for the case when the waiting time distribution $\Psi(t)$ is a one-sided stable Lévy law of index θ with the same tail as Eq. (4) and the distribution $\Phi(\vec{r})$ of jumps is a symmetric stable Lévy of index μ with the same tail as Eq. (5). Their Laplace and Fourier transforms that appear in the denominator of Eq. (22) are, respectively, $\hat{\Psi}(\beta)=\exp[-\beta^{\theta}]$ and $\hat{\Phi}(\vec{k})=\exp[-|\vec{k}|^{\mu}/2]$. Note that the long wavelength $|\vec{k}|\to 0$ and long time $\beta\to 0$ approximation gives $1-\exp[-(c'\beta)^{\theta}]\exp[-|\sigma\vec{k}|^{\mu}]=(c'\beta)^{\theta}+|\sigma\vec{k}|^{\mu}$, which recovers Eq. (51). By comparing the exact solution of Eq. (21) for $\Psi(t)$ and $\Phi(\vec{r})$ of the above Lévy form with that of the long wavelength $|\vec{k}|\to 0$ and long time $\beta\to 0$ approximation (52), Barkai [87] finds that certain solutions of Eq. (52) diverge on the origin, a behavior not found for the corresponding solutions of Eq. (21). In addition, certain solutions of the full equation (21) converge only very slowly for $\mu<1$ to the solutions of the long-time approximation (52). These results validate our use of the asymptotic long time behavior with respect to the scaling laws but provide a note of caution if one needs more precise nonasymptotic information. In this case, such information can be obtained by a suitable analysis of the full equation (21).

Using power counting, expression (52) predicts a diffusion process (45) with exponent

$$H=\frac{\theta}{\mu}. \tag{53}$$

This prediction is checked by numerical simulation of the ETAS model in the critical regime $n=1$, with $\theta=0.2$, $\mu=0.9$, shown in Fig. 6. The average distance between the first mainshock and its aftershocks as a function of the time from the mainshock indeed increases according to Eq. (45) with an exponent H in very good agreement with the prediction $H=\theta/\mu=0.2$. As the form of the denominator in Eq. (52) is independent of the space dimension, the prediction (53) is valid in any space dimension.

The natural variable for the expansions given below allowing to compute $N(t,\vec{r})$ is

$$z=\frac{Dt^{\theta/\mu}}{|\vec{r}|}, \tag{54}$$

where $D=\sigma/c'^{\theta/\mu}$ and $c'=c[\Gamma(1-\theta)]^{1/\theta}$.

1. z expansion of the solution

$W(t,\vec{r})$ can be obtained as the following sum [Eq. (5.10) of Ref. [88]]

$$W(t,\vec{r})=\frac{1}{\pi|\vec{r}|}\sum_{m=0}^{+\infty}(-1)^m z^{m\mu}\frac{\Gamma(m\mu+1)}{\Gamma(m\theta+1)}\cos\left[\frac{\pi}{2}(m\mu+1)\right]. \tag{55}$$

Applying Eq. (32) to Eq. (55) term by term in the sum, we get

$$N(t,\vec{r}) = \frac{c'^{-\theta}}{D\pi t^{1-\theta+\theta/\mu}} \sum_{m=0}^{+\infty} (-1)^m z^{1+m\mu} \frac{\Gamma(m\mu+1)}{\Gamma((m+1)\theta)} \times \cos\left[\frac{\pi}{2}(m\mu+1)\right]. \quad (56)$$

The asymptotics

$$\frac{\Gamma(m\mu+\mu+1)\Gamma(m\theta+1)}{\Gamma(m\theta+\theta+1)\Gamma(m\mu+1)} \sim \frac{\Gamma(m\mu+\mu+1)\Gamma((m+1)\theta)}{\Gamma((m+2)\theta)\Gamma(m\mu+1)} \sim m^{\mu-\theta} \quad (57)$$

show that the series (55) and (56) exist only for $\mu<\theta$. It can be shown that these series exist for all z in this case. This series converges very slowly for large z but the Padé summation method [89] can be used to improve the convergence of Eq. (56) in the case $\mu<\theta$, and can also be used to estimate Eq. (56) in the case $\mu>\theta$ for which the series diverges.

The space integral $\int dr N(t,r)$ over the whole one-dimensional volume V, with $N(t,r)$ given by Eq. (56), recovers the global Omori law

$$\int_V dr N(t,r) \sim \frac{1}{t^{1-\theta}}. \quad (58)$$

Note the nontrivial phenomenon in which the superposition of all aftershock activities transforms the local Omori law or "bare propagator" (4) $\Psi(t)\sim 1/t^{1+\theta}$ into the global Omori law or "dressed propagator" $1/t^{1-\theta}$. This effect was predicted in Refs. [25,26] in the version of the ETAS model without space dependence. These results are consistent with the claim of Sec. II D, according to which all results reported previously for the version of the ETAS model without space dependence hold also for the version of the space-dependent ETAS model studied here, when averaging over the whole space.

The asymptotic behavior for $|\vec{r}| \gg D t^{\theta/\mu}$ (i.e., $z \ll 1$) and $\mu<\theta$ is obtained by keeping only the first nonzero term ($m=1$) in Eq. (56) which is convergent for all z in the case $\mu<\theta$,

$$N(t,\vec{r}) = \frac{\sin\left(\frac{\pi\mu}{2}\right)}{\sigma c' \pi} \frac{\Gamma(1+\mu)}{\Gamma(2\theta)} \left(\frac{c'}{t}\right)^{1-2\theta} \left(\frac{\sigma}{|\vec{r}|}\right)^{1+\mu} \quad \text{for} \quad |\vec{r}| \gg D t^{\theta/\mu}. \quad (59)$$

At fixed large $|\vec{r}|$ and for $t<|\vec{r}/D|^{\mu/\theta}$, this predicts a local Omori law with exponent $p=1-2\theta$.

2. 1/z expansion of the solution

We use the theory of Fox functions [82] to obtain $N(t,\vec{r})$ as an infinite series in $1/z$. For this, we first rewrite expression (56) as a Fox function [82],

$$N(t,\vec{r}) = \frac{c'^{-\theta}}{D\mu\pi t^{1-\theta+\theta/\mu}} \times R\left(H_{2,2}^{1,2}\left[z e^{i\pi/2} \middle| \begin{matrix} (1/\mu,1/\mu),(1,1) \\ (1/\mu,1/\mu),(\theta/\mu-\theta+1,\theta/\mu) \end{matrix} \right] \right), \quad (60)$$

where $R(z)$ indicates the real part of z.

The $1/z$ expansion of $N(t,\vec{r})$ can be obtained using the dual expansion of the Fox function (60) [expression (3.7.2) of Ref. [82]]

$$N(t,\vec{r}) = \frac{c^{-\theta}}{D\pi\mu t^{1-\theta+\theta/\mu}} \sum_{m=0}^{+\infty} (-1)^m \times \left[\mu z^{1-\mu-m\mu} \frac{\Gamma(1-(m+1)\mu)\sin((m+1)\mu\pi/2)}{\Gamma(-m\theta)} + \frac{z^{-m}}{m!} \frac{\pi\cos(m\pi/2)}{\sin[(m+1)\pi/\mu]\Gamma(\theta-(m+1)\theta/\mu)} \right]. \quad (61)$$

This expansion exists only for $\mu>\theta$ [conditions of p. 71 below Eq. (3.7.2) of Ref. [82]]. This is easily checked by the behavior of an asymptotics similar to Eq. (57). Note that the series (61) is not defined in the special case $\mu=1$ due to the presence of the ill-defined ratio $\Gamma(0)/\Gamma(0)$ and a different approach is required, such as the integral representation of $W(t,\vec{r})$ developed in Ref. [88]. The global Omori law obtained by integrating over the whole space (61) is again $N(t)\sim 1/t^{1-\theta}$, as expected from the analysis of the ETAS model without space dependence [26].

Keeping only the largest term of Eq. (61) for large z, we obtain the asymptotic behavior for small distances $r<Dt^{\theta/\mu}$,

$$N(t,r) \simeq \frac{\Gamma(1-2\mu)\sin(\pi\mu)\sin(\pi\theta)}{c'\sigma\pi^2} \frac{\Gamma(1+\theta)}{(r/\sigma)^{1-2\mu}} \frac{1}{(t/c')^{1+\theta}} \quad \text{for} \quad \mu<0.5,$$

$$N(t,r) \simeq \frac{c'^{-\theta}}{c'\sigma\mu\Gamma(\theta-\theta/\mu)\sin(\pi/\mu)} \frac{1}{(t/c')^{1-\theta+\theta/\mu}} \quad \text{for} \quad 0.5<\mu<2. \quad (62)$$

Note that for $r<Dt^{\theta/\mu}$ and $0.5<\mu<2$, the leading behavior of $N(t,r)$ is independent of r.

Equation (62) thus predicts an apparent exponent

$$p=1+\theta \quad \text{for} \quad \mu<0.5,$$

$$p=1-\theta+\theta/\mu \quad \text{for} \quad 0.5<\mu<2, \quad (63)$$

for small distances $r<Dt^{\theta/\mu}$. This prediction is valid only in the case $\mu>\theta$ for which the series (61) is convergent. However, the same asymptotic results are also obtained by differ-

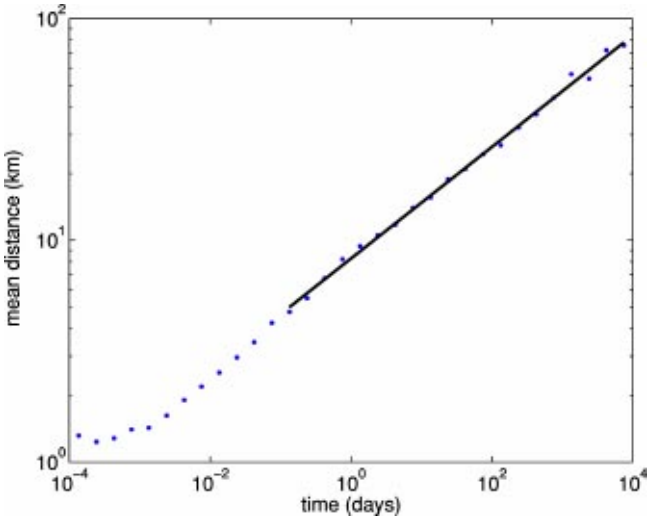

FIG. 6. Average distance between the first mainshock and its aftershocks as a function of the time from the mainshock, for a numerical simulation of the ETAS model in the critical regime $n=1$, with $\theta=0.2$, $\mu=0.9$, $c'=1$ day, and $d=1$ km. The solid line is a fit of the data which gives an exponent $H=0.25$ in good agreement with the predicted value $H=0.22$.

ent methods in the case $\mu<\theta$, for instance, expression (63) is recovered for all $\mu<2$ using the integral representation of Ref. [88]. The numerical evaluation of Eq. (56), which converges for $\mu<\theta$, also recovers the asymptotic results (62). The two regimes $\mu<0.5$ and $0.5<\mu<2$ are illustrated in Figs. 7 and 8, respectively. The seismicity rate $N(t,\vec{r})$ is evaluated from expression (56) for small z and from expression (61) for large z.

We also performed numerical simulations of the ETAS and CTRW models and the results are in good agreement with expressions (56) and (61) for $N(\vec{r},t)$ for $t\gg c$ and $r\gg d$. For very small times $t\ll c$, or for very small distances $r\ll d$, expressions (56) and (61) are not valid because they are based on a long wavelength $|\vec{k}|\to 0$ and long time $\beta\to 0$ approximation. Numerical simulations of the ETAS model in the case $\theta=0.2$ and $\mu=0.9$ are presented in Fig. 9, and are in good agreement with the analytical solutions (56) and (61) shown in Fig. 8 for the same parameters, except from the truncation of $N(t,r)$ for times $t\ll c$ and distances $r\ll d$ that are not reproduced by the analytical solution.

F. A simple nonseparable joint distribution of waiting times and jump sizes: coupled spatial diffusion and long waiting time distribution

Consider the choice for $\phi_{m_i}(t-t_i,\vec{r}-\vec{r}_i)$ replacing Eq. (2) by

$$\phi_{m_i}(t-t_i,\vec{r}-\vec{r}_i)=\rho(m_i)\Psi(t-t_i)\Phi(|\vec{r}-\vec{r}_i|/\sqrt{Dt}), \tag{64}$$

where $\rho(m_i)$ and $\Psi(t)$ are again given by Eqs. (3) and (4) while Eq. (5) is changed into

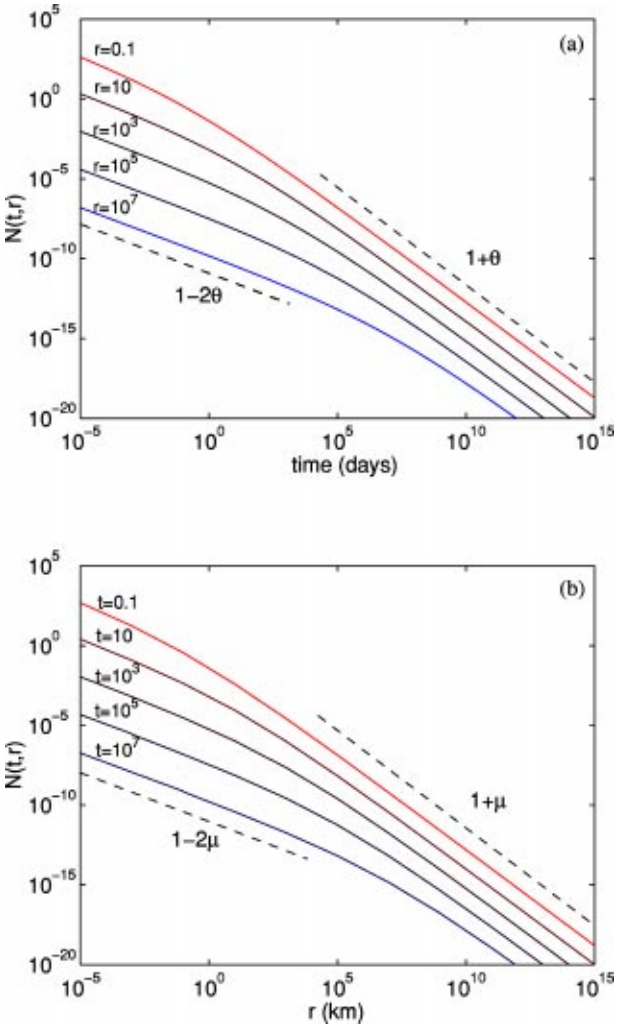

FIG. 7. Rate of seismicity $N(t,r)$ for $\theta=0.2$, $\mu=0.2$, $c'=1$ day, and $\sigma=1$ km, evaluated from expressions (56) and (62), plotted as a function of the time (a) for different values of the distance r between the mainshock and its aftershocks, and (b) as a function of r for different values of the time between the mainshock and its aftershocks. We stress again that the time scales shown here do not necessarily correspond to real observable time scales but are presented to demonstrate clearly the existence of the two regimes. The dashed lines give the predicted asymptotic dependence in each regime.

$$\Phi(|\vec{r}-\vec{r}_i|/\sqrt{Dt})=\frac{1}{\sqrt{2Dt}}\exp(-|\vec{r}-\vec{r}_i|^2/Dt). \tag{65}$$

The spatial diffusion of seismic activity is now coupled to the waiting time distribution. Expression (65) captures the effect that, in order for aftershocks to spread over large distances by the underlying physical process, they need time. In fact, returning to the discussion in the Introduction on the various proposed mechanisms for aftershocks, expression (65) embodies a microscopic diffusion process.

In this case, Eq. (21) must be replaced by

$$\hat{N}(\beta,\vec{k})=\frac{\hat{S}_M(\beta,\vec{k})}{1-n\hat{\phi}(\beta,\vec{k})}, \tag{66}$$

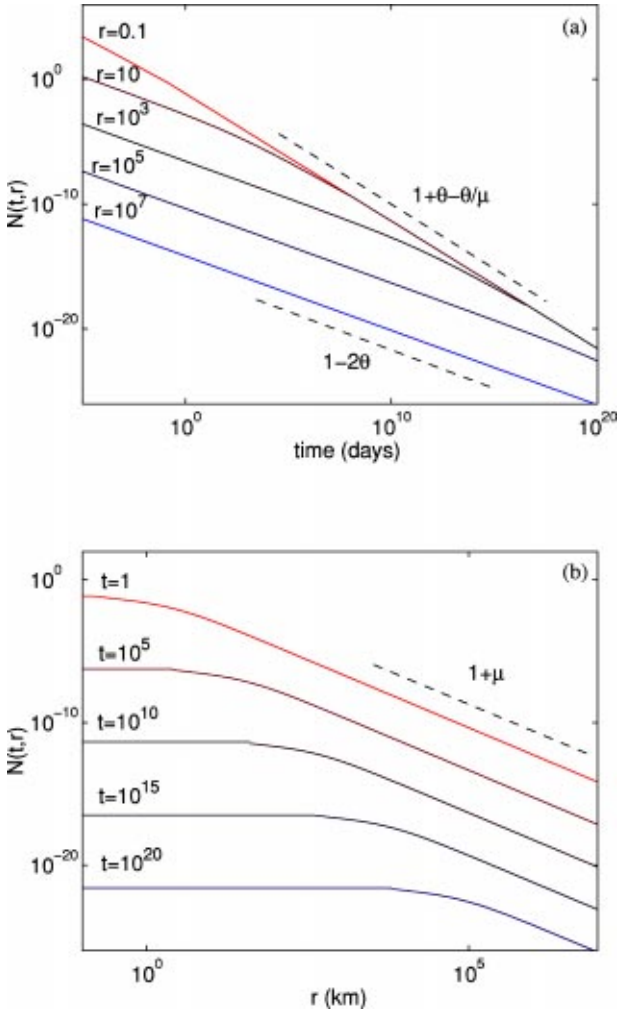

FIG. 8. Rate of seismicity $N(t,r)$ for $\theta=0.2$, $\mu=0.9$, $c'=1$ day, and $\sigma=1$ km, evaluated from expressions (56) and (62), plotted as a function of the time (a) for different values of the distance r between the mainshock and its aftershocks, and (b) as a function of r for different values of the time between the mainshock and its aftershocks. The dashed lines give the predicted asymptotic dependence in each regime.

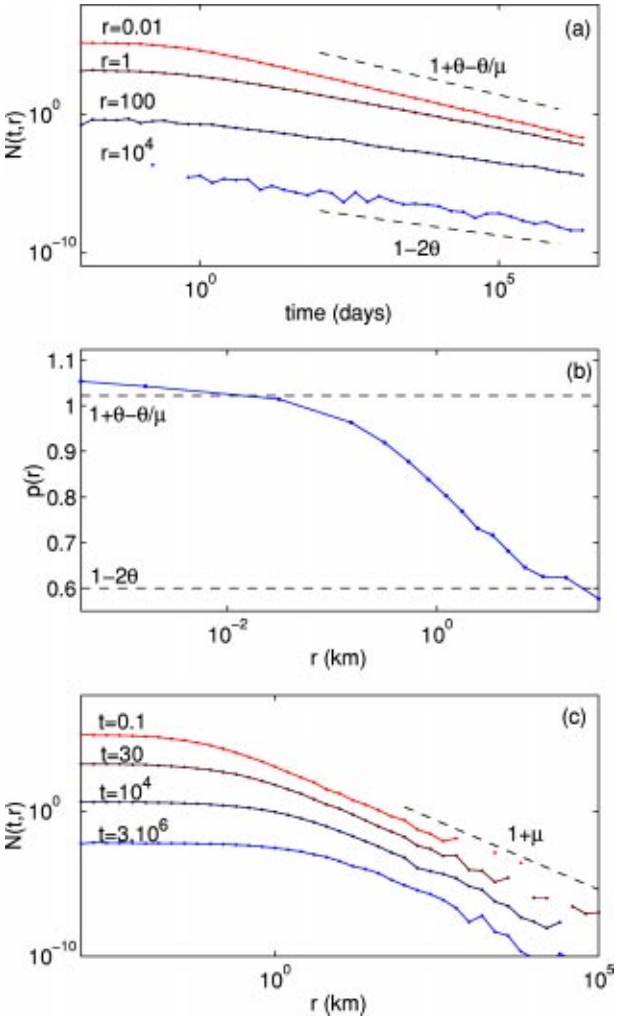

FIG. 9. Rate of seismicity $N(t,r)$ obtained from numerical simulations of the ETAS model generated with the same parameters as in Fig. 8 ($\theta=0.2$, $\mu=0.9$, $c'=1$ day, and $d=1$ km). $N(r,t)$ is computed by averaging over 500 numerical realizations of the ETAS model. (a) Aftershock rate as a function of the time from the mainshock for several distances $|\vec{r}|$ ranging from 0.01 to 10^4 km. (b) Apparent Omori exponent measured for times $t>10$ as a function of the distance from the mainshock. The aftershock decay rate (with time) is larger close to the mainshock epicenter than at large distances from the mainshock. The asymptotic values for small and large distances are in agreement with the predictions (63) for $r \ll Dt^{\theta/\mu}$ and (59) for $r \gg Dt^{\theta/\mu}$, which are shown as the horizontal dashed lines. (c) Rate of seismicity $N(t,r)$ as a function of the distance between aftershocks and mainshock for various times. The theoretical prediction for large distances is shown as the dashed line with slope $-(1+\mu)$.

where $\hat{\phi}(\beta,\vec{k})$ is the Laplace-Fourier transform of the product $\Psi(t)\Phi(|\vec{r}|/\sqrt{Dt})$. For large times and long distances for which the first terms in the expansion in β and k are sufficient, and for $n=1$, we obtain

$$\hat{\phi}(\beta,\vec{k}) \propto \frac{\hat{S}_M(\beta,\vec{k})}{(\beta+Dk^2)^{\theta}}. \tag{67}$$

The inverse Laplace-Fourier transform of Eq. (66) is

$$N(t,\vec{r}) \sim \frac{1}{t^{1-\theta}} \frac{1}{\sqrt{2\pi Dt}} \exp(-|\vec{r}|^2/Dt). \tag{68}$$

As expected, expression (68) recovers the dressed Omori propagator in the case of absence of space dependence [26]. At finite r and long times, the dressed Omori law also decay as $1/t^{1-\theta}$. The diffusion of aftershocks is normal with the standard diffusion exponent $H=1/2$.

V. NEW QUESTIONS ON AFTERSHOCKS DERIVED FROM THE CTRW ANALOGY

We list a series of comments and questions suggested from the analogy between the ETAS model and the CTRW model. In particular, we discuss the possibility of defining

new observables for earthquake aftershocks, which could be worthwhile to investigate in future empirical studies of earthquake aftershocks.

A. Recurrence of aftershock activity in the proximity of the main shock

A quantity often investigated in studies of random walks is the probability $W(t,\vec{0})$ to find the random walker at its starting point (the origin) at time t. In the earthquake framework, this is the seismic aftershock rate close to the main shock.

B. First-passage times

The first-passage time of a random walk is the first arrival time of the random walk at a given point $\vec{r}$. In the earthquake context, this translates into the study of the waiting time for a given region to have its own first aftershock after the main shock occurs. The distribution of such first passage waiting times gives the distribution of times with no nearby seismic activity. See, for instance, Ref. [84] in the case of a power-law distribution of waiting times and Gaussian distribution of jump sizes. Margolin and Berkowitz [76] give the distribution of first-passage times in the case where the jump distribution is narrow and the waiting distribution is long tailed $\sim 1/t^{1+\theta}$. They analyze the three different regimes $\theta<1$, $1<\theta<2$, and $\theta \geqslant 2$.

C. Occupation time of seismic activity

Weiss and Calabrese [90] have studied the total amount of time spent by a lattice CTRW on a subset of points. In the seismic language, this amounts to studying the probability distribution of the durations of aftershock sequences that are localized in a specific subset of the space. In other words, how probable are aftershock sequences that are found only within a given spatial subset over a certain duration?

D. Transience and recurrence of seismic activity

Another question that has been studied in some details in the CTRW framework is whether random walks are transient or recurrent. A transient random walk visits any point $\vec{r}$ at most a finite number of times before escaping to infinity. For earthquakes, the transient regime corresponds to the activation of at most a finite number of aftershocks in any given point $\vec{r}$. In contrast, a recurrent random walk may return a growing number of times to all or a subset of points at time increases. In the aftershock language, this means that these points will have a never-ending (decaying) aftershock activity. We stress here the difference between the global Omori law giving a never-ending power-law decay of the aftershock activity (in the subcritical regime $n<1$) and its spatial dependence which must exhibit important variations. In particular, in the recurrent regime, an Omori law can be documented by counting aftershocks in those limited regions of space which are activated again and again.

E. Probability for the cumulative number of aftershocks

Let us define a basic quantity in the CTRW formalism, namely the probability $\chi_m(t)$ to make exactly m steps up to time t. In the earthquake context, $\chi_m(t)$ is the probability to have exactly m aftershocks after the main shock. In the case in which the spatial transition probability $\Phi(\vec{r})$ between different positions is independent of the waiting times [corresponding to factorizing $\phi_{m_i}(t-t_i,\vec{r}-\vec{r}_i)$ as in Eq. (2)], the probability density $W(t,\vec{r})$ to find the walker at position $\vec{r}$ at time t can be written

$$W(t,\vec{r})=\sum_{m=0}^{+\infty} W_m(\vec{r})\chi_m(t), \tag{69}$$

where $W_m(\vec{r})$ is the probability to reach $\vec{r}$ from $\vec{0}$ in m steps. In the earthquake context, $W_m(\vec{r})$ is the probability that there has been exactly m events in the time interval $[0,t]$ and that the last one occurred at $\vec{r}$. Equation (69) states that the CTRW is a random process subordinated to simple random walks described by $W_m(\vec{r})$ under the operational time given by the $\chi_m(t)$ distribution [91,92].

F. Random walk models with birth and death and background seismicity from localized sources

Bender *et al.* [93] have studied models of random walks in which walkers are born in proportion to the population at one specific site (for instance, the origin) with probability $a-1$ (with $a>1$) and die at all other sites with probability $1-n$ (with $n\leqslant 1$). In the earthquake context, this consists in assuming that the aftershock activity is fed by a localized region in space, which is itself activated by the aftershocks returning to this region, furthering the overall activity. This may be considered to describe the seismic activity close to a plate boundary, in which the plate boundary is the constant self-consistent source of a seismic activity which may spread over a significant region away from the boundary. The excursion of the random walkers quantify the spread of the seismic activity away from the main fault structure. The rate of death of the walkers correspond exactly to the distance $1-n$ from the critical value $n=1$. Bender *et al.* [93] find a phase diagram in the $(a-1,1-n)$ parameter space in which a boundary separates two possible asymptotic regimes.

(1) For small $a-1$ and large $1-n$, the seismic activity at the origin and everywhere eventually dies off.

(2) For large $a-1$ and small $1-n$, the average seismic activity at the origin approaches a positive constant at long times. In this regime, there is a transition as $a-1$ is decreased or as $1-n$ is increased, between a case where the global seismic activity outside the origin goes to zero and a case where it diverges at long times. On the boundary between these two regimes in the $(a-1,1-n)$ parameter space, the distribution of seismic activity approaches a steady state at long times. There is a critical point (for space dimensions different from 2) at a certain value $(a_c-1,1-n_c)$, for which the long-time seismic activity away from the source is given by $\sim(a-a_c)^{\nu}$ where ν is a critical ex-

ponent equal to 2 in three dimensions.

Note that the results of Ref. [93] are obtained for random walks on a lattice. This can easily be converted into a CTRW by the fact that a CTRW is nothing by a process subordinated to discrete random walks under the operational time defined by the process $\{t_i\}$ of the time of just arrival to a given site, as given by Eq. (69).

VI. DISCUSSION

Using the analogy between the ETAS model and the CTRW model established here, we have derived the relation between the average distance between aftershocks and the mainshock as a function of the time from the mainshock, and the joint probability distribution of the times and locations of aftershocks.

We have assumed that each earthquake triggers aftershocks at a distance r and time t according to the bare propagator $\phi(r,t)$, which can be factorized as $\Psi(t)\Phi(r)$. This means that the distribution $\Phi(r)$ of the distances between an event and its direct aftershocks is decoupled from the distribution $\Psi(t)$ of waiting time. Hence, the direct aftershocks triggered by a single mainshock do not diffuse in space with time. Notwithstanding this decoupling in space and time of the bare propagator $\phi(r,t)$, we have shown that the global law or dressed propagator $N(t,\vec{r})$ defined as the global rate of events at time t and at position $\vec{r}$, cannot be factorized into two distributions of waiting times and space jumps. This joint distribution of waiting times and positions of the whole sequence of aftershocks cascading from a mainshock is different from the product of the bare time and space propagators.

The mean distance between the mainshock and its aftershocks, including secondary aftershocks, increases with the time from the mainshock, due to the cascade process of aftershocks triggering aftershocks triggering aftershocks, and so on. In the critical case $n=1$, this diffusion takes the form of a power-law relation $R\sim t^H$ of the average distance R between aftershocks and the mainshock, as a function of the time t from the mainshock. If the local Omori law is characterized by an exponent $0<\theta<1$, and if the space jumps follow a power law $\Phi(r)\sim 1/(r+d)^{1+\mu}$, the diffusion exponent is given by $H=\theta/\mu$ in the case $\mu<2$ and $H=\theta/2$ in the case $\mu>2$. Depending on the θ and μ values, we can thus observe either subdiffusion ($H<1/2$) or superdiffusion ($H>1/2$), as summarized in Fig. 10. In the subcritical ($n<1$) and supercritical ($n>1$) regimes, this relation is still valid up to the characteristic time t^* given by Eq. (1) and for distances smaller than $r^*\propto Dt^{*H}$ given by Eq. (30). For $t>t^*$ and $r>r^*$ in the subcritical regime, the global distributions of times and distances between the mainshock and its aftershocks are decoupled and there is therefore no diffusion. In the supercritical regime, the aftershock rate increases exponentially for $t>t^*$ and the aftershocks diffuses more rapidly than before t^*.

In the critical regime, the cascade of secondary aftershocks introduces a variation of the apparent Omori exponent as a function of the distance from the mainshock. The asymptotic values of the Omori exponent in the different regimes are summarized in Table II. In the regime $\mu<2$, we observe a transition from an Omori law decay with an exponent $p=1-2\theta$ at early times $t^H \ll r/D$ to a larger exponent at large times. This provides another mechanism to explain the observed variability of the Omori exponent. In the regime $\mu>2$, a power-law decay of the seismicity with time is observed only at large times $t^H \gg r/D$. At early times, or at large distances $r \gg Dt^H$, the seismicity rate is very small, because the seismicity has not yet diffused up to the distance r.

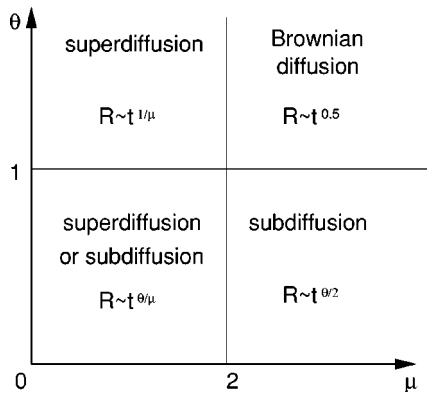

FIG. 10. Classification of the different regime of the diffusion of aftershocks in space as a function of time from the main shock. The bare Omori law for aftershocks decay with time as $1/t^{1+\theta}$. The jump size distribution between the earthquake ''mother'' and its ''daughters'' is proportional to $1/r^{1+\mu}$. $R(t)$ is the average distance between all aftershocks triggered up to time t after the mainshock.

We should emphasize that our theoretical analysis of aftershock diffusion predicts the behavior of the ensemble average of aftershock sequences. Individual sequences may depart from this ensemble average, especially for sequences with few earthquakes and limited durations. For long sequences (20 000 events say), we have verified that the exponent H measured on individual sequences does not deviate from the ensemble average value by more than about 20%. As already discussed, the impact of fluctuations becomes, however, more effective as the parameter α increases above $b/2$.

TABLE II. Asymptotic values of the (renormalized) Omori exponent (of the dressed propagator) in the different regimes for $z \ll 1$ and $z \gg 1$, where $z \equiv Dt^H/r$.

	Large z ($r \ll Dt^H$)	Small z ($r \gg Dt^H$)
$\mu<0.5$	$p=1+\theta$	$p=1-2\theta$
$0.5\leqslant\mu<2$	$p=1-\theta+\theta/\mu$	$p=1-2\theta$
$2\leqslant\mu$	$p=1-\theta/2$	Not defined [a]

[a]The Omori exponent is not defined in this case because the dependence of $N(t,\vec{r})$ with respect to time given by expression (42) and represented in Fig. 4 has a contribution from the exponential asymptotics which is different from a power law for large distances $r \gg Dt^H$.

The diffusion of the seismicity also renormalizes the spatial distribution of the seismicity, which is very different from the local distribution $\Phi(r)$ of distances between a triggering event and its direct aftershocks. In the regime $\mu>2$, the global seismicity rate $N(t,\vec{r})$ decays exponentially with the distance from the mainshock, whereas the local distribution of distances $\Phi(r)$ is a power-law distribution. In the regime $\mu<2$, the local law $\Phi(r)\sim r^{-1-\mu}$ is recovered at large distances, but a slower decay for $0.5<\mu<2$ or a constant rate for $\mu<0.5$ is observed at small distances $r \ll Dt^H$. These predictions on the decrease of the Omori exponent with r have not yet been observed in earthquake catalogs, but an expansion of the aftershock zone has been reported in many studies [28–36]. However, very few studies have quantified the diffusion law. Noir *et al.* [35] show that the earthquake Dobi sequence (central Afar, August 1989) composed of 22 $M>4.6$ earthquakes presented a migration that was in agreement with a diffusion process due to fluid transfer in the crust, characterized by a normal diffusion process with exponent $H=0.5$. Tajima and Kanamori [31,32] studied several aftershock sequences in subduction zone and observed a much slower logarithmic diffusion, which is compatible with a low exponent H close to 0.1. In some cases, the aftershock sequence displays no expansion with time. For instance, Shaw [94] studied several aftershock sequences in California and concluded that the distribution of distances between the mainshock and its aftershocks is independent of time. This can be explained by the fact that the Omori exponent measured in Ref. [94] is very close to 1, thus θ is very small and our prediction is that the exponent H should be very small.

In fact, the ETAS model predicts that diffusion should be observed only for aftershock sequences with a measured Omori exponent p significantly smaller than 1, which can only occur according to our model when the bare Omori propagator with exponent $1+\theta$ is renormalized into the dressed propagator with global exponent $1-\theta$. We have shown that this renormalization of the exponent only occurs at times less than t^*, while for longer times in the subcritical regime $n<1$ the dressed Omori propagator recovers the value of the bare exponent $1+\theta>1$ (see Fig. 2). Therefore, identifying an empirical observation of $p<1$ with our prediction $p=1-\theta$ indicates that the aftershock sequence falls in the ''good'' time window $t<t^*$ in which the renormalization operates. We have also shown that the dressed propagator gives a diffusion only for $t<t^*$. We can thus conclude that, according to the ETAS model, the observation of an empirical Omori exponent larger than 1 is indicative of the large time $t>t^*$ behavior in the subcritical regime $n<1$, for which there is no diffusion. This provides a possible explanation for why many sequences studied in Refs. [31,32,94] do not show a diffusion of the aftershock epicenters. Reciprocally, a prerequisite for observing diffusion in a given aftershock sequence is that the empirical p value be less than 1 in order to qualify the regime $t<t^*$.

An alternative model has been discussed by Dieterich [40] who showed that the spatial variability of the stress induced by a mainshock, coupled with a rate and state friction law, results in an expansion of the aftershock zone with time. This expansion does not take the form of a diffusion law as observed in the ETAS model, the relation between the characteristic size of the aftershock zone does not grow as a power law of the time from the mainshock [Eq. (22) and Fig. 6 of Ref. [40]].

Marsan *et al.* [95,96] and Marsan and Bean [97] studied several catalogs at different scales, from the scale of a deep mine to the world-wide seismicity, and observed that the average distance between two earthquakes increases as a power law of the time between them, with an exponent often close to 0.2, indicative of a subdiffusion process. They interpreted their results as a mechanism of stress diffusion, that may be due to fluid transfer with heterogeneous permeability leading to subdiffusion. Their analysis is quite different from those used in other studies, because they consider all pairs of events, without distinction between aftershocks and mainshocks. This analysis can, however, lead to spurious diffusion, and in some cases this method does not detect diffusion in a synthetic data set with genuine diffusion. We have tested their analysis on a synthetic catalog generated by superposing a background seismicity with uniform spatial and temporal distribution, and ten mainshocks with Poissonian distribution in time and space, and with a power-law distribution of energies. Each of these mainshocks generates only *direct* aftershocks, without secondary cascades of aftershocks, and the number of aftershocks increases exponentially with the magnitude of the mainshock. This way, we generate a synthetic catalog without any physical process of diffusion, and which includes all the other well-established characteristics of real seismicity: clustering in space and time superposed to a seismicity background. Applying the analysis of Refs. [95–97] to this synthetic data set leads to an apparent diffusion process with a well-defined exponent $H=0.5$. However, this apparent diffusion does not reflect a genuine diffusion but simply describes the crossover from the characteristic size of an aftershock zone at early times to the larger average distance between uncorrelated events at large times. In plain words, the apparent power law $R\propto t^H$ is nothing but a crossover and is not real. Furthermore, applying this analysis to a synthetic catalog generated using the ETAS model, without seismicity background, and with a theoretical diffusion exponent $H=0.2$, the method yields $H=0.01$ if we use all the events of the catalog. If we select only events up to a maximum distance r_{max} to apply the same procedure as in Refs. [95–97], we obtain larger values of H which are more in agreement with the theoretical exponent $H=0.2$ but with large fluctuations that are function of r_{max}. Therefore, it is probable that the diffusion reported in Refs. [95–97] is not real and results from a crossover between two characteristic scales of the spatial earthquake distribution. It may be attributed to the analyzing methodology which mixes up uncorrelated events. We are thus reluctant to compare the results of Marsan *et al.* [95–97] with the predictions obtained with the ETAS model.

One can similarly question the results on anomalous diffusion of seismicity obtained by Sotolongo-Costa *et al.* [98], who considered 7500 microearthquakes recorded by a local Spanish network from 1985 to 1995. They interpret the se-

quence of earthquakes as a random walk process, in which the walker jumps from an earthquake epicenter to the next in a sequential order. The time between two successive events is seen as a waiting time between two jumps and the distance between these events is taken to correspond to the jump size. Since the distributions of time intervals and of distances between successive earthquakes are both heavy tailed (approximately power laws), their model is a CTRW. We cannot stress enough that their CTRW model of seismicity has nothing to do with our results on the mapping of the ETAS model onto a CTRW. Their procedure is ad hoc and their results depend obviously strongly on the space domain of the analysis since distant earthquakes that are completely unrelated can be almost simultaneous. We also stress that our mapping of the ETAS model onto the CTRW model does not correspond to identifying an earthquake sequence as a *single* realization of a CTRW, as assumed arbitrarily by Sotolongo-Costa *et al.* [98].

Our predictions obtained here are thus difficult to test on seismicity data, due to the small number of events available and the restricted time periods and distance ranges, and because the seismicity background can strongly affect the results. New methods should hence be developed to investigate if there is a real physical process of diffusion in seismic activity and to compare the observations of real seismicity with the quantitative predictions of the ETAS model. Preliminary study of aftershock sequences in California leads to the conclusion that most aftershock sequences are characterized by an Omori exponent $p>1$, indicative of the subcritical regime with $t>t^*$. As expected from our predictions in this regime, we do not observe an expansion of the aftershock zone. However, a few sequences give a value $p<1$ and also exhibit an increase of the average distance between the mainshock and its aftershocks consistent with our predictions. A detailed report of this analysis will be reported elsewhere.

VII. CONCLUSION

We have studied analytically and numerically the ETAS model, which is a simple stochastic process modeling seismicity, based on the two best-established empirical laws for earthquakes, the power-law decay of seismicity after an earthquake and a power-law distribution of earthquake energies. This model assumes that each earthquake can trigger aftershocks, with a rate increasing with its magnitude. In this model, the seismicity rate is the result of the whole cascade of direct and secondary aftershocks.

We have first established an exact correspondence between the ETAS model and the CTRW model. We have then used this analogy to derive the joint probability of times and distances of the seismicity following a large earthquake and we have characterized the different regimes of diffusion.

We have shown that the diffusion of the seismicity should be observed only for times $t<t^*$, where t^* is a characteristic time depending on the model parameters, corresponding to an observed Omori exponent smaller than one. Most aftershock sequences have an observed Omori exponent larger than one, corresponding to the subcritical regime of the ETAS model, for which there is no diffusion. The diffusion of the seismicity produces a decrease of the Omori exponent as a function of the distance from the mainshock, the decay of aftershocks being faster close to the mainshock than at large distances. The spatial distribution of seismicity is also renormalized by the cascade process, so that the observed distribution of distances between the mainshock and its aftershocks can be fundamentally different from the bare propagator $\Phi(r)$ which gives the distribution of the distances between triggered and triggering earthquakes. We have also noted that the ETAS model generates apparent but realistic fractal spatial patterns.

Assuming that the distances between triggering and triggered events are independent of the time between them, this model generates a diffusion of the whole sequence of aftershocks with the time from the mainshock, which is induced by the cascade of aftershocks triggering aftershocks, and so on. Our results thus provides a simple explanation of the diffusion of aftershock sequences reported by several studies, which was often interpreted as a mechanism of anomalous stress diffusion. We see that no such ''anomalous stress diffusion'' is needed and our theory provides a parsimonious account of aftershock diffusion resulting from the minimum physical ingredients of the ETAS model. As Einstein once said: ''A theory is more impressive the greater the simplicity of its premises, the more different the kinds of things it relates and the more extended its range of applicability.''

ACKNOWLEDGMENTS

We are very grateful to B. Berkowitz, S. Gluzman, J.-R. Grasso, Y. Klafter, L. Margerin, A. Saichev, and G. Zaslavsky for useful suggestions and discussions.

[1] D. Sornette, *Spontaneous Formation of Space-Time Structures and Criticality*, Vol. 39 of *NATO Advanced Study Institute, Series B: Physics*, edited by T. Riste and D. Sherrington (Kluwer, Dordrecht, 1991), pp. 57–106.
[2] J. B. Rundle and W. Klein, Rev. Geophys. **33**, 283 (1995).
[3] I. Main, Rev. Geophys. **34**, 433 (1996).
[4] D. Sornette, Phys. Rep. **313**, 238 (1999).
[5] D. L. Turcotte, Rep. Prog. Phys. **62**, 1377 (1999).
[6] B. Gutenberg and C. F. Richter, Bull. Seismol. Soc. Am. **34**, 185 (1944).
[7] F. Omori, On the aftershocks of earthquakes, J. Coll. Sci. Imp. Univ. Tokyo **7**, 111 (1894).
[8] G. Ouillon, C. Castaing, and D. Sornette, J. Geophys. Res., [Solid Earth] **101** B3, 5477 (1996).
[9] Y. Y. Kagan and L. Knopoff, Geophys. J. R. Astron. Soc. **62**, 303 (1980).
[10] V. F. Pisarenko and D. Sornette, e-print cond-mat/0011168, Pure Appl. Geophys. (to be published).
[11] Y. Y. Kagan, Pure Appl. Geophys. **155**, 537 (1998).
[12] D. Sornette, *Critical Phenomena in Natural Sciences*, Springer

Series in Synergetics (Springer, Heidelberg, 2000).
[13] R. A. Harris, in International Handbook of Earthquake and Engineering Seismology, edited by W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger (unpublished), Chap. 73.
[14] T. Yamashita and L. Knopoff, Geophys. J. R. Astron. Soc. **91**, 13 (1987).
[15] M. W. Lee and D. Sornette, Eur. Phys. J. B **15**, 193 (2000).
[16] Y. Huang, H. Saleur, C. G. Sammis, and D. Sornette, Europhys. Lett. **41**, 43 (1998).
[17] T. Utsu, Y. Ogata, and S. Matsuúra, J. Phys. Earth **43**, 1 (1995).
[18] M. Bath and C. F. Richter, Bull. Seismol. Soc. Am. **48**, 133 (1958).
[19] G. C. Beroza and M. D. Zoback, Science **259**, 210 (1993).
[20] D. P. Hill *et al.*, Science **260**, 1617 (1993).
[21] D. W. Steeples and D. D. Steeples, Bull. Seismol. Soc. Am. **86**, 921 (1996).
[22] Y. Y. Kagan and D. D. Jackson, J. Geophys. Res., [Solid Earth] **103**, 24 453 (1998).
[23] A. J. Meltzner and D. J. Wald, Bull. Seismol. Soc. Am. **89**, 1109 (1999).
[24] D. Dreger and B. Savage, Bull. Seismol. Soc. Am. **89**, 1094 (1999).
[25] A. Sornette and D. Sornette, Geophys. Res. Lett. **26**, 1981 (1999).
[26] A. Helmstetter and D. Sornette, e-print http://arXiv.org/abs/cond-mat/0109318, J. Geophys. Res. (to be published).
[27] D. Sornette and A. Helmstetter, Phys. Rev. Lett. **89**, 158501 (2002).
[28] K. Mogi, J. Phys. Earth **16**, 30 (1968).
[29] M. Imoto, On migration phenomena of aftershocks following large thrust earthquakes in subduction zones, Report of the National Research Center for Disaster Prevention, No. 25, 1981, pp. 29–71.
[30] J. L. Chatelain, R. K. Cardwell, and B. L. Isacks, Geophys. Res. Lett. **10**, 385 (1983).
[31] F. Tajima and H. Kanamori, Phys. Earth Planet. Inter. **40**, 77 (1985).
[32] F. Tajima and H. Kanamori, Geophys. Res. Lett. **12**, 345 (1985).
[33] R.L. Wesson, Tectonophysics **144**, 214 (1987).
[34] T. Ouchi and T. Uekawa, Phys. Earth Planet. Inter. **44**, 211 (1986).
[35] J. Noir, E. Jacques, S. Bekri, P. M. Adler, P. Tapponnier, and G. C. P. King, Geophys. Res. Lett. **24**, 2335 (1997).
[36] E. Jacques, J. C. Ruegg, J. C. Lepine, P. Tapponnier, G. C. P. King, and A. Omar, Geophys. J. Int. **138**, 447 (1999).
[37] P. A. Rydelek and I. S. Sacks, Geophys. Res. Lett. **28**, 3079 (2001).
[38] A. Nur and J. R. Booker, Science **175**, 885 (1972).
[39] K. W. Hudnut, L. Seeber, and J. Pacheco, Geophys. Res. Lett. **16**, 199 (1989).
[40] J. Dieterich, J. Geophys. Res., [Solid Earth] **99**, 2601 (1994).
[41] Y. Y. Kagan and L. Knopoff, Science **236**, 1563 (1987).
[42] Y. Ogata, Am. Stat. **83**, 9 (1988).
[43] Y. Yamanaka and K. Shimazaki, J. Phys. Earth **38**, 305 (1990).
[44] G. Drakatos and J. Latoussakis, Journal of Seismology **5**, 137 (2001).
[45] Z. Guo, and Y. Ogata, J. Geophys. Res., [Solid Earth] **102**, 2857 (1997).
[46] A. Helmstetter, Geophys. Res. Lett. (to be published).
[47] K. R. Felzer, T. W. Becker, R. E. Abercrombie, G. Ekström, and J. R. Rice, J. Geophys. Res., [Solid Earth] (to be published).
[48] M. Bath, Tectonophysics **2**, 483 (1965).
[49] R. Console, A. M. Lombardi, M. Murru, and D. Rhoades, J. Geophys. Res. (to be published).
[50] Y. Ogata, Tectonophysics **169**, 159 (1989).
[51] Y. Ogata, J. Geophys. Res., [Solid Earth] **97**, 19 845 (1992).
[52] Y. Ogata, Pure Appl. Geophys. **155**, 471 (1999).
[53] Y. Ogata, J. Geophys. Res., [Solid Earth] **106**, 8729 (2001).
[54] Y. Y. Kagan, Geophys. J. Int. **106**, 135 (1991).
[55] Y. Y. Kagan and D. D. Jackson, Geophys. J. Int. **143**, 438 (2000).
[56] R. Console and M. Murru, J. Geophys. Res., [Solid Earth] **166**, 8699 (2001).
[57] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
[58] S. Krishnamurthy, A. Tanguy, P. Abry, and S. Roux, Europhys. Lett. **51**, 1 (2000).
[59] Y. Ogata, Ann. Inst. Stat. Math. **50**, 379 (1998).
[60] C. H. Scholz and B. B. Mandelbrot, *Fractals in Geophysics* (Birkhäuser Verlag, Boston, 1989).
[61] A. Sornette, P. Davy, and D. Sornette, J. Geophys. Res. [Solid Earth] **98**, 12 111 (1993).
[62] C. C. Barton and P. R. La Pointe, *Fractals in the Earth Sciences* (Plenum Press, New York, 1995).
[63] K. Nanjo and H. Nagahama, Pure Appl. Geophys. **157**, 575 (2000).
[64] G. Ouillon and D. Sornette, Geophys. Res. Lett. **23**, 3409 (1996).
[65] D. Hamburger, O. Biham, and D. Avnir, Phys. Rev. E **53**, 3342 (1996).
[66] M. Eneva, Geophys. J. Int. **124**, 773 (1996).
[67] O. Malcai, D. A. Lidar, O. Biham, and D. Avnir, Phys. Rev. E **56**, 2817 (1997).
[68] A. Helmstetter, D. Sornette, and J.-R. Grasso, e-print http://arXiv.org/abs/ cond-mat/0205499, J. Geophys. Res. (to be published).
[69] E. W. Montroll and H. Scher, J. Stat. Phys. **9**, 101 (1973).
[70] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
[71] V.M. Kenkre, E. W. Montroll, and M. F. Shlesinger, J. Stat. Phys. **9**, 45 (1973).
[72] M. F. Shlesinger, J. Stat. Phys. **10**, 421 (1974).
[73] G. H. Weiss, *Aspects and Applications of Random Walk* (North-Holland, Amsterdam, 1994).
[74] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, New York, 1995).
[75] R. Metzler and J. Klafter, Phys. Rep. **339**, 1 (2000).
[76] G. Margolin and B. Berkowitz, J. Phys. Chem. **104**, 3942 (2000).
[77] B. Berkowitz and H. Scher, Transp. Porous Media **42**, 241 (2001).
[78] S. Redner, Am. J. Phys. **58**, 267 (1990).
[79] H. Scher, H. G. Margolin, R. Metzler, J. Klafter, and B. Berkowitz, Geophys. Res. Lett. **29**, 5 (2002).
[80] J. W. Kirchner, X. Feng, and C. Neal, Nature (London) **403**, 524 (2000).

[81] E. Barkai, R. Metzler, and J. Klafter, Phys. Rev. E **61**, 132 (2000).
[82] A. M. Mathai and R. K. Saxena, *The H-function with Applications in Statistics and Other Disciplines* (Wiley Eastern Limited, New Delhi, 1978).
[83] H. E. Roman and P. A. Alemany, J. Phys. A **27**, 3407 (1994).
[84] E. Barkai, Phys. Rev. E **63**, 046118 (2001).
[85] C. Budde, D. Prato, and M. Re, Phys. Lett. A **283**, 309 (2001).
[86] M. Kotulski, J. Stat. Phys. **81**, 777 (1995).
[87] E. Barkai, e-print cond-mat/0108024.
[88] A. I. Saichev and G. M. Zaslavsky, Chaos **7**, 753 (1997).
[89] C. M. Bender and S. Orzag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).
[90] G. H. Weiss and P. P. Calabrese, Physica A **234**, 443 (1996).
[91] I. M. Sokolov, A. Blumen, and J. Klafter, e-print cond-mat/0107632.
[92] I. M. Sokolov, Phys. Rev. E **63**, 011104 (2001).
[93] C. M. Bender, S. Boettcher, and P. N. Meisinger, Phys. Rev. Lett. **75**, 3210 (1995).
[94] B. E. Shaw, Geophys. Res. Lett. **20**, 907 (1993).
[95] D. Marsan, C. J. Bean, S. Steacy, and J. McCloskey, J. Geophys. Res. [Solid Earth] **105**, 28 (2000).
[96] D. Marsan, C. J. Bean, S. Steacy, and J. McCloskey, Geophys. Res. Lett. **26**, 3697 (1999).
[97] D. Marsan and C. J. Bean, Geophys. J. Int. (to be published).
[98] O. Sotolongo-Costa, J. C. Antoranz, A. Posadas, F. Vidal, and A. Vásquez, Geophys. Res. Lett. **27**, 1965 (2000).