Nonuniversal Power Law Distribution of Intensities of the Self-Excited Hawkes Process: A Field-Theoretical Approach

Kiyoshi Kanazawa[1] and Didier Sornette[2,3,4]
[1]Faculty of Engineering, Information and Systems, The University of Tsukuba, Tennodai, Tsukuba, Ibaraki 305-8573, Japan
[2]ETH Zurich, Department of Management, Technology and Economics, Zurich 8092, Switzerland
[3]Tokyo Tech World Research Hub Initiative, Institute of Innovative Research, Tokyo Institute of Technology, Tokyo 152-8550, Japan
[4]Institute of Risk Analysis, Prediction and Management, Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen 518055, China

(Received 6 January 2020; revised 29 June 2020; accepted 25 August 2020; published 21 September 2020)

The Hawkes self-excited point process provides an efficient representation of the bursty intermittent dynamics of many physical, biological, geological, and economic systems. By expressing the probability for the next event per unit time (called "intensity"), say of an earthquake, as a sum over all past events of (possibly) long-memory kernels, the Hawkes model is non-Markovian. By mapping the Hawkes model onto stochastic partial differential equations that are Markovian, we develop a field theoretical approach in terms of probability density functionals. Solving the steady-state equations, we predict a power law scaling of the probability density function of the intensities close to the critical point $n = 1$ of the Hawkes process, with a nonuniversal exponent, function of the background intensity ν_0 of the Hawkes intensity, the average timescale of the memory kernel and the branching ratio n. Our theoretical predictions are confirmed by numerical simulations.

DOI: 10.1103/PhysRevLett.125.138301

The self-excited conditional Poisson process introduced by Hawkes [1–3] is the simplest point process modeling epidemic dynamics, in which the whole past history influences future activity. It captures the ubiquitous phenomenon of time (and space) intermittency and clustering due to endogenous interactions. The Hawkes process is enjoying an explosion of interest in many complex systems, including in physics, biology, geology, and seismology and in financial and economic markets. For instance, the Hawkes model remains the standard reference in statistical seismology [4–7] and is now used to model a variety of phenomena in finance, from microstructure dynamics to default risks [8,9]. The Hawkes process is also fashionable to model social dynamics [10].

A key ingredient of the Hawkes model is the memory kernel $h(t)$, which quantifies how much a past event influences the triggering of a future event. When $h(t)$ is a pure exponential function, the Hawkes model can be represented as a Markovian process by adding an auxiliary variable. But most systems exhibit longer memories, with $h(t)$ containing multiple time scales and often describing power law decaying impacts, which makes the Hawkes model non-Markovian in general. Here, we present a general field master equation, which represents the self-excited Hawkes process as being equivalent to a Markovian stochastic partial differential equation (SPDE). This novel representation allows us to use the mathematical apparatus to solve master equations and derive a new result on the distribution of activity rates, which is found to take the form of a nonuniversal power law.

The Hawkes process is defined via its intensity $\hat{\nu}$, which is the frequency of events per unit time. An event can be a burst of electrons in an amorphous semiconductor or an organic compound associated with photoconductivity [11], a rainfall or runoff in catchments [12], an earthquake [4], an epidemic [13], an epileptic seizure [14], a firm's bankruptcy or credit default [15,16], a financial volatility burst [17,18], a transaction in foreign exchange markets [19], a book sale [20], a view of a YouTube video [21], or any other occurrence in social dynamics [22]. Such an event (or shock) occurs during $[t, t + dt)$ with the probability of $\hat{\nu}dt$, with

$$\hat{\nu}(t) = \nu_0 + n \sum_{i=1}^{\hat{N}(t)} h(t - \hat{t}_i), \tag{1}$$

where ν_0 is the background intensity, $\hat{t}_i$ represents the time series of events, n is a positive number called branching ratio, $h(t)$ is a normalized nonnegative function satisfying $\int_0^\infty h(t)dt = 1$, and $\hat{N}(t)$ is the number of events during the interval $[0, t)$ (called "counting process"), as shown in Fig. 1(a) for a schematic. By convention, we denote stochastic variables with a hat symbol, such as $\hat{A}$, to distinguish them from the nonstochastic real numbers A, corresponding for instance to a specific realization of the random variable. The memory kernel $h(t)$ represents the

usually non-Markovian influence of a given event. The branching ratio n is the average number of events of first generation ("daughters") triggered by a given event [4,23] and is also the fraction of events that are endogenous, i.e., that have been triggered by previous events [24]. The Hawkes process has three different regimes: (i) $n < 1$: subcritical; (ii) $n = 1$: critical and (iii) $n > 1$: supercritical or explosive (with a finite probability). The Hawkes process is a model for out-of-equilibrium systems without detailed balance [25] and does not satisfy the fluctuation-dissipation relation [26].

Let us decompose the memory kernel as a continuous superposition of exponential kernels,

$$h(t) = \frac{1}{n}\int_0^\infty \frac{n(x)}{x} e^{-t/x} dx, \qquad n = \int_0^\infty dx n(x), \quad (2)$$

satisfying the normalization $\int_0^\infty h(t)dt = 1$ with the set of continuous timescale $x \in \boldsymbol{R}_+ := (0, \infty)$. This decomposition is equivalent to applying the Laplace transform, a standard method even for non-Markovian Langevin equations [26–30]. Here $n(x)$ quantifies the contribution of the xth exponential with memory length x to the branching ratio and $n(x)/n$ is the normalized distribution of timescales present in the memory kernel. In this Letter, we require the existence of its first moment

$$\frac{\alpha}{n} := \langle \tau \rangle := \int_0^\infty x \frac{n(x)}{n} dx < \infty. \quad (3)$$

This condition (3) means that $n(x)$ should decay faster than $1/x^2$ at large x and thus $h(t)$ decays at large times faster than $1/t^2$. In addition, we restrict our analysis to the subcritical case $n < 1$.

The starting point of our approach is to express $\hat{\nu}(t)$ (1) as the continuous sum

$$\hat{\nu}(t) = \nu_0 + \int_0^\infty dx \hat{z}(t,x), \quad (4)$$

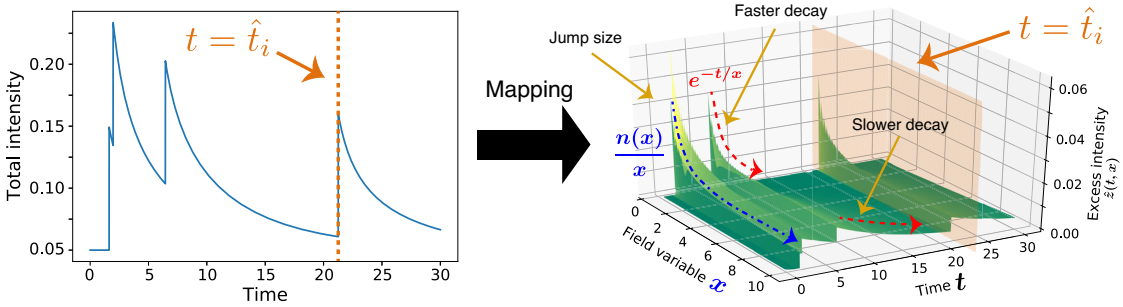
where each excess intensity $\hat{z}(t,x)$ is the solution of a simple time-derivative equation

$$\frac{\partial \hat{z}(t,x)}{\partial t} = -\frac{\hat{z}(t,x)}{x} + \frac{n(x)}{x}\hat{\xi}^P_{\hat{\nu}}(t), \qquad \forall\ x \in \boldsymbol{R}_+, \quad (5)$$

and the same state-dependent Poisson noise $\hat{\xi}^P_{\hat{\nu}}(t)$, defined by

$$\hat{\xi}^P_{\hat{\nu}}(t) = \sum_{i=1}^{\hat{N}(t)} \delta(t - \hat{t}_i), \quad (6)$$

acts on the Langevin Eq. (5) for each excess intensity $\hat{z}(t,x)$ [see Fig. 1(b)]. The excess intensity $\{\hat{z}(t,x)\}_{x\in\boldsymbol{R}_+}$ can be viewed as a one-dimensional field variable distributed on the x axis; correspondingly, Eq. (5) should be considered as an SPDE describing the classical stochastic dynamics of the field. This interpretation has the advantage of allowing us to apply functional methods available for SPDEs [25]. The introduction of the $\hat{z}(t,x)$ is called Markovian embedding, a technique to transform a non-Markovian dynamics onto a Markovian one by adding a sufficient number of variables (see [31–33] for the cases of non-Markovian Langevin equations). Markovian embedding is related to the trick proposed in [34] for an efficient estimation of the maximum likelihood of the Hawkes process. Each SPDE (5) describes a Markovian relaxation of the field variable $\hat{z}(t,x)$, hit by intermittent simultaneous shocks $\hat{\xi}^P_{\hat{\nu}}(t)$ with x-dependent sizes $n(x)/x$, whose influence decays exponentially with the characteristic time x. Equation (5) together with (6) implies that $\hat{\nu}(t)$ given by (4) recovers the standard Hawkes definition (1).



(a) Non-Markovian dynamics of total intensity (b) Markovian field dynamics of excess intensities

FIG. 1. Mapping from (a) a non-Markovian description of the total intensity $\hat{\nu}(t)$, obeying the original Hawkes process (1), to (b) a Markovian description of the excess intensity $\{\hat{z}(t,x)\}_x$ on the auxiliary field variable $x \in \boldsymbol{R}_+$, which obeys an SPDE (5). Notice the shocks (6) impact all excess intensity $\{\hat{z}(t,x)\}_x$ for all different variables x simultaneously (orange plane at $t = \hat{t}_i$). One can observe the dependence $n(x)/x$ of the jump size given in Eq. (5) along the field variable x (blue chain arrow). Along the time axis, the exponential decay $\sim e^{-t/x}$ is shown to be faster (slower) for smaller (larger) x (red broken arrow). The sample trajectory was generated with the branching ratio density $n(x) = c_{\text{ini}}(x - x_{\text{ini}}) + n_{\text{ini}}$ if $x \in [x_{\text{ini}}, x_{\text{fin}}]$ [$n(x) = 0$ otherwise], with $\nu_0 = 0.05$, $x_{\text{ini}} = 0.5$, $x_{\text{fin}} = 10.0$, $c_{\text{ini}} = 0.0044$, $n_{\text{ini}} = 0.021$, $\Delta x = (x_{\text{fin}} - x_{\text{ini}})/200$, and $\Delta t = 0.006$.

We have thus transformed a non-Markovian point process into a Markovian SPDE, which allows us to derive the corresponding master equation for the probability density functional (PDF) $P[\{\hat{z}(t,x)=z(x)\}_{x\in\boldsymbol{R}_+}]=P_t[z]$ for any field configuration $\{z(x)\}_{x\in\boldsymbol{R}_+}$, such that $P_t[z]\mathcal{D}z$ is the probability that the system is in the state specified by $\{z(x)\}_{x\in\boldsymbol{R}_+}$ at time t, with the functional integral volume element $\mathcal{D}z$. The corresponding master equation for the PDF $P_t[z]$ reads

$$\frac{\partial P_t[z]}{\partial t}=\int dx\frac{\delta}{\delta z}\left(\frac{z}{x}P_t[z]\right)+\left\{\nu_0+\int dx\left(z-\frac{n}{x}\right)\right\}P_t\left[z-\frac{n}{x}\right]-\left\{\nu_0+\int dxz\right\}P_t[z], \tag{7}$$

with the condition $P_t[z]=0$ holding over the boundary of the function space $z\in\partial\boldsymbol{R}_+^\infty:=\{z|z(x)=0$ for some $x\in(0,\infty)\}$. This can be derived by performing an ensemble average in a weak integral sense, namely considering an arbitrary functional $f[\{\hat{z}(t,x)\}_x]$ and averaging it over all possible realizations of $\hat{z}(t,x)$ weighted by their PDF (see Ref. [35] for details). The functional description (7) is interpreted as a formal continuous limit of a discrete formulation according to the convention [25] (see Ref. [35] for technical details).

It is convenient to transform (7) using the functional Laplace transformation $\mathcal{L}_{\text{path}}$ of an arbitrary functional $f[z]$ defined by the functional integration (or path integral) $\mathcal{L}_{\text{path}}(f[z];s):=\int\mathcal{D}ze^{-\int dxs(x)z(x)}f[z]$. Then, the Laplace representation of the PDF is $\tilde{P}_t[s]:=\mathcal{L}_{\text{path}}(P_t[z];s)$ for an arbitrary nonnegative function $\{s(x)\}_{x\in\boldsymbol{R}_+}$. The resulting Laplace transformed master Eq. (7) takes the following simple first-order functional differential equation in the steady state ($\partial P_t[z]/\partial t=0$):

$$\int dx\mathcal{H}[s;x]\frac{\delta\Phi[s]}{\delta s(x)}=-\nu_0\mathcal{K}[s], \tag{8}$$

where $\Phi[s]:=\log\tilde{P}_{\text{ss}}[s]:=\lim_{t\to\infty}\log\tilde{P}_t[s]$ is the steady state cumulant functional, $\mathcal{H}[s;x]:=e^{-\int dx's(x')n(x')/x'}-1+s(x)/x$, and $\mathcal{K}[s]:=e^{-\int dx's(x')n(x')/x'}-1$. This hyperbolic equation can be solved by the method of characteristics; the corresponding Lagrange-Charpit (LC) equations are the following partial-integro equations,

$$\frac{\partial s(l;x)}{\partial l}=-\mathcal{H}[s;x],\qquad\frac{\partial\Phi(l)}{\partial l}=\nu_0\mathcal{K}[s], \tag{9}$$

with the curvilinear parameter l indexing the position along a characteristic curve. The tail of the distribution of intensities $\hat{\nu}$ corresponds to the neighborhood of $s=0$ in the Laplace transform domain (i.e., $\tilde{P}_{\text{ss}}(s)\sim|s|^\gamma$ for $s\to0\Leftrightarrow P_{\text{ss}}(\nu)\sim\nu^{-\gamma-1}$ for $\nu\to\infty$ [36]). We first study the subcritical case $n<1$ and then the critical regime $n=1$ via a stability analysis of (9) for small s.

Remarkably, the LC equations can be interpreted as a dynamical system where l plays the role of time. This mapping allows us to use the standard stability analysis for bifurcations of dynamical systems, particularly for asymptotic analyses near criticality. Indeed, the stability analysis for $s\to0$ corresponds to the long time limit $l\to\infty$ and the critical condition of the original Hawkes process (1) corresponds to the transcritical bifurcation condition for the dynamical system described by Eq. (9).

Subcritical case $n<1$.—Linearizing the LC Eq. (9) yields

$$\frac{\partial s(l;x)}{\partial l}=-\int dx'H(x,x')s(x'), \tag{10a}$$

$$\frac{\partial\Phi(l)}{\partial l}=\nu_0\int dx'K(x')s(x'), \tag{10b}$$

with $x'H(x,x'):=\delta(x-x')-n(x')$ and $K(x'):=n(x')/x'$. Introducing the eigenvalues $\lambda\geq\lambda_{\min}$ and eigenfunctions $e(x;\lambda)$ of the operator $H(x,x')$, satisfying the relation

$$\int dx'H(x,x')e(x';\lambda)=\lambda e(x;\lambda), \tag{11}$$

we verify that all eigenvalues are real and the inverse matrix of $H(x,x')$, denoted by $H^{-1}(x,x')$, exists and has a singularity at $n=1$ (see Ref. [35] for the proof), recovering the critical condition of this Hawkes process.

We now introduce a set of variables to obtain a new representation based on the eigenfunctions,

$$s(x)=\sum_\lambda e(x;\lambda)X(\lambda)\Leftrightarrow X(\lambda)=\int dxe^{-1}(\lambda;x)s(x). \tag{12}$$

Here the inverse matrix $e^{-1}(\lambda;x)$ is introduced, satisfying $\int dxe^{-1}(\lambda;x)e(x;\lambda')=\delta_{\lambda,\lambda'}$. The existence of the inverse matrix is equivalent to the assumption that the set of all eigenfunctions is complete, and thus $H(x,x')$ can be diagonalized: $\int dxdx'e^{-1}(\lambda;x)H(x,x')e(x';\lambda')=\lambda\delta_{\lambda,\lambda'}$. In this representation, the linearized LC equations read

$$\frac{\partial X(l;\lambda)}{\partial l}=-\lambda X(l;\lambda). \tag{13}$$

For subcriticality, all the eigenvalues are positive, indicating that the fixed point $\{X(\lambda)=0\}_\lambda$ (i.e., $\{s(x)=0\}_x$) is the stable attractor in the functional space. Using straightforward calculations (see Ref. [35]), we obtain

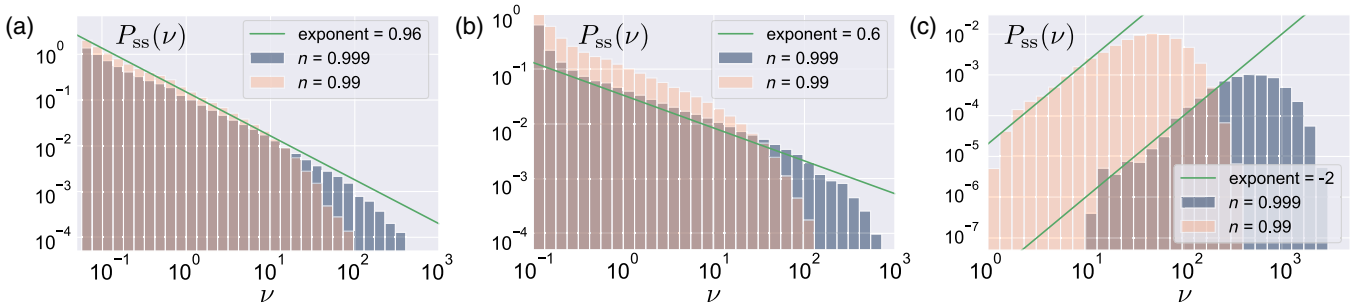


FIG. 2. Numerical steady state PDFs of the Hawkes intensity $\hat{\nu}$ for the double exponential case with $(\tau_1, \tau_2) = (1, 3)$, $(n_1, n_2) = (0.5, 0.499)$ or $(n_1, n_2) = (0.5, 0.49)$, near the critical point. (a) Background intensity $\nu_0 = 0.01$, leading to the power law exponent 0.96. (b) $\nu_0 = 0.1$, leading to the power law exponent 0.6. (c) $\nu_0 = 0.75$, leading to the negative (i.e., growing PDF) power law exponent −2.0. Here the sampling time interval and total sampling time are $dt = 0.001$ and $T_{\rm tot} = 10000$ from the initial condition $\hat{z}(0) = 0$. The initial 10% of the sample was discarded from the statistics for initialization.

$$\Phi[s] \simeq -\nu_0 \int dx \int dx' K(x) H^{-1}(x, x') s(x'), \quad (14)$$

from which we find, for small s,

$$\log \tilde{P}_{\rm ss}(s) := \log \tilde{P}_{\rm ss}[s\mathbf{1}(x)] = \Phi[s\mathbf{1}(x)] \simeq \frac{-\nu_0}{1-n} s, \quad (15)$$

where $\mathbf{1}(x)$ is the constant function equal to 1 for any x. The mean intensity thus converges at long times to $\langle \hat{\nu}(t) \rangle \to \nu_0/(1-n)$, which is a well-known result [4,23].

Critical case $n = 1$.—At criticality, the smallest eigenvalue vanishes, $\lambda_{\rm min} = 0$, which is associated to the zero eigenfunction $e(x; \lambda = 0) = x$, as verified by direct substitution: $\int dx H(x, x') e(x'; \lambda = 0) = 1 - n = 0$. From the linear LC equation (13), it is clear that the dominant contribution comes from the component $X(\lambda = 0)$, associated with the zero eigenfunction $e(x; \lambda = 0)$. The explicit representation of $X(\lambda = 0)$ is given by $X(\lambda = 0) = \int_0^\infty dx n(x) s(x)/\alpha$, where α is defined by Eq. (3).

We also obtain the LC equations for each component to leading order for $\lambda' \neq 0$,

$$\frac{\partial X(l; 0)}{\partial l} \simeq -\frac{X^2(l; 0)}{2\alpha}, \qquad \frac{\partial X(l; \lambda')}{\partial l} = -\lambda' X(l; \lambda'). \quad (16)$$

This is the normal form of transcritical bifurcations, leading to a log-type singularity in the cumulant for small s. Indeed, after straightforward calculations, we obtain

$$\log \tilde{P}_{\rm ss}(s) := \log \tilde{P}_{\rm ss}[s\mathbf{1}(x)] \simeq \nu_0 s - 2\nu_0 \alpha \log |s|, \quad (17)$$

which by inverse Laplace transform yields

$$P_{\rm ss}(\nu) \sim \nu^{-1+2\nu_0\langle\tau\rangle}, \quad (18)$$

using definition (3). The exponent $1 - 2\langle\tau\rangle\nu_0$ of the PDF is nonuniversal and a function of the background intensity ν_0 of the Hawkes intensity and of the average time scale of the memory kernel $\langle\tau\rangle$. As the tail exponent is smaller than 1, the steady-state PDF $P_{\rm ss}(\nu)$ would be not normalizable in absence of some cutoff [37], coming either from finite-time effects or nonexact criticality ($n \to 1^-$). This means that this power-law scaling (18) actually corresponds to an intermediate asymptotics of the PDF, according to the classification of Barenblatt [38], which, for n close to 1, can be observed over many orders of magnitude of the intensity for near-critical systems, as shown in figure 2. The intermediate power law asymptotic (18) is our main novel quantitative result. See Ref. [35] for details.

Example 1.—The above general derivation of (18) is rather involved and one can develop more intuition by studying simplest cases where the memory function $h(t)$ is a single exponential or the sum of two exponentials. In the former case $h(t) = (1/\tau)e^{-t/\tau}$, all functions become single variables and functional derivatives and integrations become standard derivative and integration operators. Then, the general master Eq. (7) reduces to

$$\frac{\partial P_t}{\partial t} = \frac{1}{\tau}\frac{\partial}{\partial z} z P_t + \left(\nu_0 + z - \frac{n}{\tau}\right) P_t\left(z - \frac{n}{\tau}\right) - (\nu_0 + z) P_t, \quad (19)$$

for the PDF $P_t := P_t(z)$ under the boundary condition $P_t(z)|_{z=0} = 0$. Its Laplace transform of the steady-state PDF $\tilde{P}_{\rm ss}(s) := \int_0^\infty d\nu e^{-s\nu} P_{\rm ss}(z)$ reads

$$\mathcal{H}(s) \frac{d\Phi(s)}{ds} = -\nu_0 \mathcal{K}(s), \quad (20)$$

by introducing the cumulant function $\Phi(s) := \log \tilde{P}_{\rm ss}(s)$, $\mathcal{H}(s) := e^{-ns/\tau} - 1 + s/\tau$, and $\mathcal{K}(s) := e^{-ns/\tau} - 1$. It can be directly solved exactly below the critical point $n < 1$, leading to

$$P_{\rm ss}(\nu) \propto \nu^{-1+2n\nu_0\tau} e^{-2\tau(1-n)\nu} \quad \text{(for large } \nu\text{)} \quad (21)$$

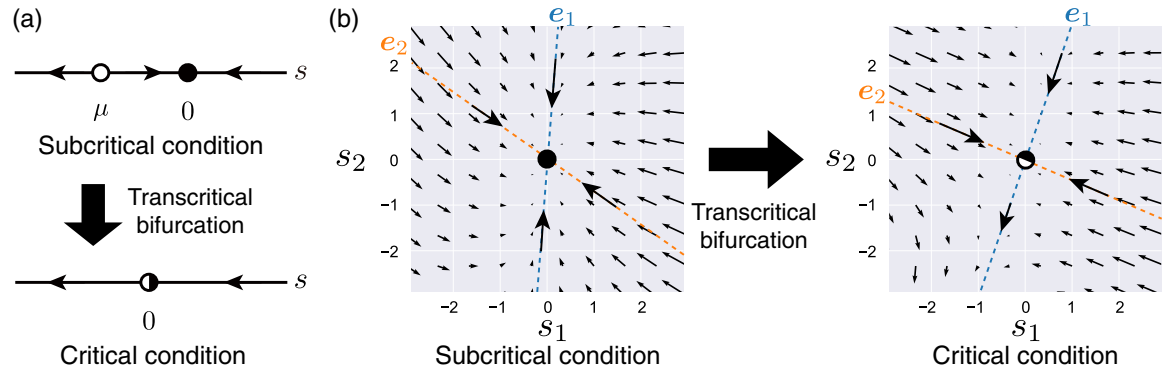


FIG. 3. Phase space description of the dynamical system following the LC equations. (a) The one-dimensional velocity field is defined by $V(s) \coloneqq ds/dl = -\mathcal{H}(s)$. For subcriticality ($n < 1$, top), there are two fixed points at $s = 0$ (attractor) and $s = \mu < 0$ (repeller). The repeller merges with the attractor at criticality ($n = 1$, bottom), which is a consistent picture of transcritical bifurcations. (b) The two-dimensional velocity field is defined by $\boldsymbol{V}(\boldsymbol{s}) \coloneqq d\boldsymbol{s}/dl$ with $\boldsymbol{s} \coloneqq (s_1, s_2)$. By linearization $\boldsymbol{V}(\boldsymbol{s}) \simeq -\boldsymbol{H}\boldsymbol{s}$, the eigenvectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are introduced by $\boldsymbol{H}\boldsymbol{e}_i = \lambda_i \boldsymbol{e}_i$ with $0 \le \lambda_1 < \lambda_2$. For subcriticality ($n < 1$, left), the origin $\boldsymbol{s} = \boldsymbol{0}$ is the stable attractor. At criticality ($n = 1$, right), the origin becomes marginal in terms of the linear stability ($\lambda_1 = 0$): a repeller merges with the attractor along the $\boldsymbol{e}_1$ (i.e., a transcritical bifurcation).

near criticality $1 - n \ll 1$. Remarkably, the LC equation $ds/dl = -\mathcal{H}(s)$ reduces to the normal form of transcritical bifurcations [see Fig. 3(a)]:

$$\frac{ds}{dl'} = \mu s - s^2 + O(s^3) \quad \text{(for small } s\text{)}, \tag{22}$$

with $l' \coloneqq n^2 l/(2\tau^2)$ and $\mu \coloneqq -2(1-n)/n^2$.

Example 2.—For two exponentials, the memory kernel is given by $h_t = \sum_{k=1}^{2} [n_k/(n\tau_k)] e^{-t/\tau_k}$, where each coefficient n_k quantifies the contribution of the kth exponential with memory length τ_k to the branching ratio $n = n_1 + n_2$. In calculations paralleling those for the general and one exponential cases, we can derive the master equation for the two-exponential case and its Laplace representation. Finally, the corresponding LC equations read

$$\frac{ds_i}{dl} = -\mathcal{H}_i(s_1, s_2), \qquad \frac{d\Phi}{dl} = \nu_0 \mathcal{K}(s_1, s_2) \tag{23}$$

with $i = 1, 2$, $\mathcal{H}_i(s_1, s_2) \coloneqq e^{-\sum_{k=1}^{2} n_k s_k/\tau_k} - 1 + s_i/\tau_i$, and $\mathcal{K}(s_1, s_2) \coloneqq e^{-\sum_{k=1}^{2} n_k s_k/\tau_k} - 1$. Following the same approach as for the general case (2), but now dealing with operators that are 2×2 matrices, we recover (18) with $\langle \tau \rangle = (n_1\tau_1 + n_2\tau_2)/n$ (see [35] for details). Our theoretical prediction is confirmed numerically for a memory kernel with two exponentials, as shown in Fig. 2. We note that the LC Eq. (23) exhibits the transcritical bifurcation as illustrated in Fig. 3(b).

While the Hawkes process was believed to be unable to produce power-law fluctuations [34], our finding demonstrates that it does produce them in the form of intermediate asymptotics, thus filling an important gap for applications to real systems. Our methodology can be readily generalized to various nonlinear Hawkes processes, highlighting broader power-law asymptotics [39]. Furthermore, our main result fills a gap in the study of the Hawkes and other point process, by focusing on the distribution of the number νdt of events in the limit of infinitely small time windows $[t, t + dt]$. This limit is in contrast to the other previously studied limit of infinitely large and finite but very large time windows, as standard results of branching processes (of which the Hawkes model is a special case) give the total number of events generated by a given triggering event (see Ref. [40] for a detailed derivation and [41] for the case of large time windows $[t, t + T]$, i.e., in the limit of large T). The corresponding PDFs are totally different from (18) which corresponds to the other limit $T \to 0$. There are also deep relationship between quantum field theories and ours. Indeed, Eq. (7) can be formally regarded as a Schrödinger equation for a non-Hermitian quantum field theory (see Ref. [35]).

This work was supported by the Japan Society for the Promotion of Science KAKENHI (Grants No. 16K16016 and No. 20H05526) and Intramural Research Promotion Program in the University of Tsukuba.

[1] A. Hawkes, J. R. Stat. Soc. Ser. B **33**, 438 (1971).
[2] A. Hawkes, Biometrika **58**, 83 (1971).
[3] A. Hawkes and D. Oakes, J. Appl. Probab. **11**, 493 (1974).
[4] A. Helmstetter and D. Sornette, J. Geophys. Res. **107**, 2237 (2002).
[5] Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988).
[6] Y. Ogata, Pure Appl. Geophys. **155**, 471 (1999).
[7] S. Nandan, G. Ouillon, D. Sornette, and S. Wiemer, Seismol. Res. Lett. **90**, 1650 (2019).
[8] V. Filimonov and D. Sornette, Phys. Rev. E **85**, 056108 (2012).
[9] A. G. Hawkes, Quant. Finance **18**, 193 (2018).
[10] Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, SEISMIC: A self-exciting point process model

for predicting tweet popularity, in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2015), p. 1513.
[11] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
[12] H. Scher, H. G. Margolin, R. Metzler, J. Klafter, and B. Berkowitz, Geophys. Res. Lett. **29**, 5 (2002).
[13] M. Feng, S.-M. Cai, M. Tang, and Y.-C. Lai, Nat. Commun. **10**, 3748 (2019).
[14] D. Sornette and I. Osorio, Prediction, in *Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Physics and Engineering*, edited by I. Osorio *et al.* (CRC Press, Taylor & Francis Group, Boca Raton, 2010), pp. 203–237.
[15] G. D'Amico, S. Dharmaraja, R. Manca, and P. Pasricha, Rep. Econ. Fin. **5**, 15 (2019).
[16] E. Errais, K. Giesecke, and L. R. Goldberg, SIAM J. Fin. Math. **1**, 642 (2010).
[17] A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel, Quant. Finance **11**, 991 (2011).
[18] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, and D. Sornette, Rep. Prog. Phys. **82**, 125901 (2019).
[19] M. Takayasua and H. Takayasu, Physica (Amsterdam) **324A**, 101 (2003).
[20] D. Sornette, F. Deschatres, T. Gilbert, and Y. Ageon, Phys. Rev. Lett. **93**, 228701 (2004).
[21] R. Crane and D. Sornette, Proc. Natl. Acad. Sci. U.S.A. **105**, 15649 (2008).
[22] D. Sornette, in *Extreme Events in Nature and Society*, edited by S. Albeverio, V. Jentsch, and H. Kantz, The Frontiers Collection (Springer, Heidelberg, 2005), pp 95–119.
[23] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes* (Springer, Heidelberg, 2003), Vol. I.
[24] A. Helmstetter and D. Sornette, Geophys. Res. Lett. **30**, 1576 (2003).
[25] C. W. Gardiner, *Handbook of Stochastic Methods*, 4th ed. (Springer, Berlin, 2009).
[26] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II*, 2nd ed. (Springer-Verlag, Berlin, 1991).
[27] H. Mori, Prog. Theor. Phys. **33**, 423 (1965); **34**, 399 (1965).
[28] M. H. Lee, Phys. Rev. B **26**, 2547 (1982).
[29] R. Morgado, F. A. Oliveira, G. G. Batrouni, and A. Hansen, Phys. Rev. Lett. **89**, 100601 (2002).
[30] J.-D. Bao, Y.-Z. Zhuo, F. A. Oliveira, and P. Hänggi, Phys. Rev. E **74**, 061111 (2006).
[31] I. Goychuk, Phys. Rev. E **80**, 046125 (2009).
[32] R. Kupferman, J. Stat. Phys. **114**, 291 (2004).
[33] F. Marchesoni and P. Grigolini, J. Chem. Phys. **78**, 6287 (1983); M. Ferrario and P. Grigolini, J. Math. Phys. (N.Y.) **20**, 2567 (1979).
[34] J.-P. Bouchaud, J. Bonart, J. Donier, and M. Gould, *Trades, Quotes and Prices* (Cambridge University Press, Cambridge, England, 2018), Sect. 9.3.4.
[35] K. Kanazawa and D. Sornette, companion paper, Phys. Rev. Research **2**, 033442 (2020).
[36] J. Klafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, 2011).
[37] The cutoff tail is typically characterized by an exponential [e.g., see Eq. (21) and Ref. [34] for the exponential and power-law memory kernel cases, respectively].
[38] G. I. Barenblatt, *Scaling, Self-Similarity, and Intermediate Asymptotics* (Cambridge University Press, Cambridge, England, 1996).
[39] K. Kanazawa and D. Sornette (to be published).
[40] A. Saichev, A. Helmstetter, and D. Sornette, Pure Appl. Geophys. **162**, 1113 (2005).
[41] A. I. Saichev and D. Sornette, Eur. Phys. J. B **49**, 377 (2006).