Bursting transition in a linear self-exciting point process

Tomokatsu Onaga[*] and Shigeru Shinomoto[†]

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

(Received 20 January 2014; published 29 April 2014)

Self-exciting point processes describe the manner in which every event facilitates the occurrence of succeeding events, as in the case of epidemics or human activity. By increasing excitability, the event occurrences start to exhibit bursts even in the absence of external stimuli. We revealed that the transition is uniquely determined by the average number of events added by a single event, $1 - 1/\sqrt{2} \approx 0.2929$, independently of the temporal excitation profile. We further extended the theory to multidimensional processes, to be able to incite or inhibit bursting in networks of agents by altering their connections.

PACS number(s): 89.75.Hc, 05.40.−a

I. INTRODUCTION

Irregular occurrences of events are modeled by the Poisson process such that point events are independently drawn in time at a given rate. Event occurrences that are not mutually independent may be modeled by adding a supplementary probability for event occurrence after every event [1]. This simple model called the Hawkes process has been widely applied to the analysis of earthquakes [2], genome sequences [3], urban crime [4], human activity [5–7], and neuronal activity [8,9].

The process is called self-exciting, if each event is associated with a positive supplementary probability for succeeding events, which we call "excitability." Given large excitability, the system may exhibit unstable bursts of events leading to nonstationary occurrence rate, even in the absence of external stimuli (Fig. 1). Contrariwise, under small excitability, the fluctuation in the occurrence rate may become undetectable from a single sparse series of irregular events.

We decide the nonstationarity of a single series of events based on whether principled rate estimators indicate a fluctuating rate or a constant rate [10,11]. Herein we obtain the critical condition for the excitability at which the estimated rate changes between constant and fluctuating. Based on the second order transition, the criticality condition is obtained in a universal formula applicable to a wide range of self-exciting processes associated with various temporal excitation profiles.

We then extend our analysis to the multidimensional Hawkes process, in which multiple agents mutually influence each other. This process is exemplified by a social system in which people influence activity through events such as emails [12,13]. It is known that people may exhibit autonomous bursts of activity without exogenous stimuli [14–16]. Knowing the conditions under which bursts occur, we can control the occurrence of burst activity by reconnecting people or agents.

II. CORRELATION IN THE HAWKES PROCESS

In the self-exciting process (or Hawkes process), the rate of event occurrence $\lambda(t)$ is modulated by past events as

$$\lambda(t) = \rho + \alpha \sum_k f(t - t_k), \qquad (1)$$

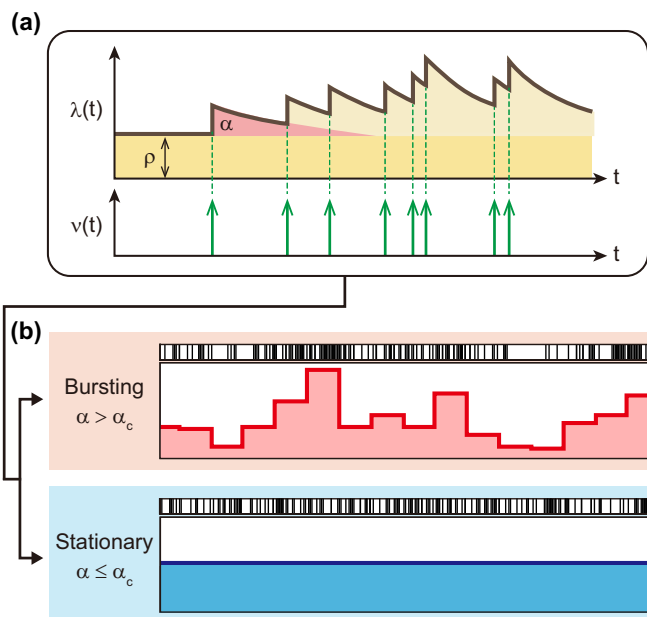*onaga@scphys.kyoto-u.ac.jp
†shinomoto@scphys.kyoto-u.ac.jp

FIG. 1. (Color online) Self-exciting processes. (a) The rate of event occurrence $\lambda(t)$ is modulated according to generated events $\nu(t) = \sum_k \delta(t - t_k)$. (b) Event sequences depicted in rasters may exhibit nonstationary bursts or remain stationary, depending on whether the excitability α is larger or smaller than a critical value α_c, respectively. The bin size of the histogram shown under each raster was selected using the method of minimizing the mean squared error between the histogram and the underlying rate.

where ρ is the base rate and t_k is the occurrence time of the kth event. The kernel function $f(t)$, representing the time course of the supplementary probability, satisfies two conditions: the causality, $f(t) = 0$ for $t < 0$, and the normalization, $\int_0^\infty f(t)\,dt = 1$. Accordingly, the coefficient α represents the excitability or the supplementary probability added after each event.

As a step to calculate the condition for the bursting transition, we estimate the correlation function of the event occurrence rate according to Hawkes [17], extending the originally proposed range of validity. By representing a series of event occurrences as a sum of Dirac delta functions, $\nu(t) = \sum_k \delta(t - t_k)$, the Hawkes process (1) may be represented as

$$\lambda(t) = \rho + \alpha \int_{-\infty}^{\infty} f(t - u)\nu(u)\,du. \qquad (2)$$

Because the ensemble average of the event occurrence $\langle \nu \rangle$ equals that of the rate $\langle \lambda \rangle$, the average rate is obtained as $\langle \lambda \rangle = \rho/(1 - \alpha)$. The excitability α should be less than unity to avoid explosion of event occurrences, in which $\lambda(t)$ diverges.

The complete covariance density $\phi^{(c)}(s) \equiv \langle \nu(t + s)\nu(t) \rangle - \langle \nu \rangle^2$ has a singularity at $s = 0$ with $\langle \nu \rangle \delta(s)$. Because $\langle \nu(t + s)\nu(t) \rangle = \langle \lambda(t + s)\nu(t) \rangle$ for $s > 0$, the covariance density satisfies the relation,

$$\phi^{(c)}(s) = \alpha \int_{-\infty}^{\infty} f(s - u)\phi^{(c)}(u)\,du, \qquad (3)$$

for $s > 0$. The correlation function of the rate fluctuation $\delta\lambda(t) \equiv \lambda(t) - \langle \lambda \rangle$ is given by removing the singularity from the complete covariance density, or $\phi(s) \equiv \langle \delta\lambda(t + s)\delta\lambda(t) \rangle = \phi^{(c)}(s) - \langle \lambda \rangle \delta(s)$. Inserting this relation into the integral equation (3), we obtain an integral equation for the correlation function,

$$\phi(s) = \alpha \langle \lambda \rangle f(s) + \alpha \int_{-\infty}^{\infty} f(s - u)\phi(u)\,du, \qquad (4)$$

which holds for $s > 0$. Define a function,

$$g(s) \equiv \alpha \langle \lambda \rangle f(s) + \alpha \int_{-\infty}^{\infty} f(s - u)\phi(u)\,du - \phi(s), \quad (5)$$

which satisfies $g(s) = 0$ for $s > 0$. Taking the Fourier transformation $\tilde{g}_\omega \equiv \int_{-\infty}^{\infty} g(t)\exp(-i\omega t)\,dt$ is given as

$$\tilde{g}_\omega = \alpha \langle \lambda \rangle \tilde{f}_\omega + \alpha \tilde{f}_\omega \tilde{\phi}_\omega - \tilde{\phi}_\omega. \qquad (6)$$

Considering the time reversal symmetry of the correlation function, $\phi(s) = \phi(-s)$ or $\tilde{\phi}_\omega = \tilde{\phi}_{-\omega}$, we obtain

$$\frac{\alpha \langle \lambda \rangle \tilde{f}_\omega - \tilde{g}_\omega}{1 - \alpha \tilde{f}_\omega} = \frac{\alpha \langle \lambda \rangle \tilde{f}_{-\omega} - \tilde{g}_{-\omega}}{1 - \alpha \tilde{f}_{-\omega}}, \qquad (7)$$

or equivalently,

$$\alpha \langle \lambda \rangle \tilde{f}_\omega + (1 - \alpha \tilde{f}_\omega)\tilde{g}_{-\omega} = \alpha \langle \lambda \rangle \tilde{f}_{-\omega} + (1 - \alpha \tilde{f}_{-\omega})\tilde{g}_\omega. \quad (8)$$

Because $f(t) = 0$ for $t < 0$ and $g(t) = 0$ for $t > 0$, their Fourier images $\tilde{f}_\omega$ and $\tilde{g}_\omega$ converge to 0 in the limit of $|\omega| \to \infty$ in half planes of $\mathrm{Im}(\omega) < 0$ and $\mathrm{Im}(\omega) > 0$, respectively. Because the left-hand side and right-hand side of Eq. (8) are regular in the lower and upper half imaginary planes, they vanish in respective half planes. While Hawkes derived the relation assuming an exponentially decaying kernel function, we may permit long-tailed kernels, such as power law functions $f(t) \propto (1 + t)^{-b}$ ($b > 1$), by trimming the range of the functional regularity to the adjoining half planes of $\mathrm{Im}(\omega) \geqslant 0$ and $\mathrm{Im}(\omega) \leqslant 0$.

Inserting the identity relation into Eq. (6), we obtain the relation

$$\tilde{\phi}_\omega = \frac{\alpha \tilde{f}_\omega + \alpha \tilde{f}_{-\omega} - \alpha^2 \tilde{f}_\omega \tilde{f}_{-\omega}}{(1 - \alpha \tilde{f}_\omega)(1 - \alpha \tilde{f}_{-\omega})} \langle \lambda \rangle, \qquad (9)$$

by which the correlation function $\phi(t)$ is obtained for a given excitation kernel $f(t)$.

III. DETECTABILITY OF THE RATE FLUCTUATION

Though the self-exciting process is a stationary process whose statistical properties are invariant with time as an ensemble, individual processes may significantly fluctuate in

time, causing bursts of events. The nonstationarity of a single series of event occurrences can be determined by whether principled rate estimators indicate fluctuating rate or constant rate. Herein we construct an optimal histogram in which the bin size is selected to minimize the mean integrated squared error (MISE) between the histogram and the underlying rate, and derive the condition under which the optimal bin size diverges, or equivalently an optimal histogram indicates constant rate [10].

The bin size Δ is selected to minimize MISE between the underlying rate $\lambda(t)$ and the histogram $\hat{\lambda}_\Delta(t)$. The MISE is a function of the bin size

$$S(\Delta) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \langle (\lambda(t) - \hat{\lambda}_\Delta(t))^2 \rangle dt, \qquad (10)$$

where T is the entire observation interval. Replacing the long time average with the average over the bin size, $\hat{\lambda}_\Delta(t)$ can be treated as a single rectangle whose height is the number of events K divided by the bin size Δ. Thus the MISE is given as

$$S(\Delta) = \left\langle \frac{1}{\Delta} \int_0^\Delta \left[\lambda^2(t) - \frac{2K}{\Delta}\lambda(t) \right] dt + \frac{K^2}{\Delta^2} \right\rangle. \quad (11)$$

The expected number of events in each interval is given by integrating the underlying rate: $E[K] = \int_0^\Delta \lambda(t)\,dt$. Because events are independently drawn, the Poisson relation holds: $E[K^2] = E[K]^2 + E[K]$. Inserting these relations into Eq. (11), we have

$$S(\Delta) = \phi(0) + \frac{\langle \lambda \rangle}{\Delta} - \frac{1}{\Delta^2} \int_0^\Delta dt \int_{-t}^t \phi(s)\,ds, \qquad (12)$$

where $\phi(s) = \langle \lambda(t + s)\lambda(t) \rangle - \langle \lambda \rangle^2 = \langle \delta\lambda(t + s)\delta\lambda(t) \rangle$.

If a series of events is derived from a constant rate process, the MISE is a monotonically decreasing function, and therefore, the optimal bin size diverges. By contrast, the MISE of inhomogeneous point processes may have a minimum at some finite Δ, provided that

$$\left. \frac{dS}{d(1/\Delta)} \right|_{\Delta = \infty} < 0. \qquad (13)$$

This can be summed up as

$$\frac{1}{\langle \lambda \rangle} \int_{-\infty}^{\infty} \langle \delta\lambda(t + s)\delta\lambda(t) \rangle ds > 1, \qquad (14)$$

on condition that $\int_0^\infty s\phi(s)\,ds$ is finite. Note that this instability condition derived from the histogram optimization is identical to the instability condition derived from the marginal likelihood maximization of the Bayesian rate estimator [11].

IV. TRANSITION IN THE SELF-EXCITING PROCESS

Applying the above-mentioned consideration to the self-exciting process, we obtain the condition for the detectable-undetectable criticality. The integral of the correlation function is given by the Fourier zero mode, or, $\int_{-\infty}^{\infty} \langle \delta\lambda(t + s)\delta\lambda(t) \rangle ds = \int_{-\infty}^{\infty} \phi(s)\,ds = \tilde{\phi}_0$. For the self-exciting point process, the critical condition is obtained from Eq. (9) and $\tilde{f}_0 = \int_{-\infty}^{\infty} f(t)\,dt = 1$ as

$$\frac{1}{\langle \lambda \rangle} \int_{-\infty}^{\infty} \langle \delta\lambda(t + s)\delta\lambda(t) \rangle ds = \frac{2\alpha - \alpha^2}{(\alpha - 1)^2} = 1. \quad (15)$$
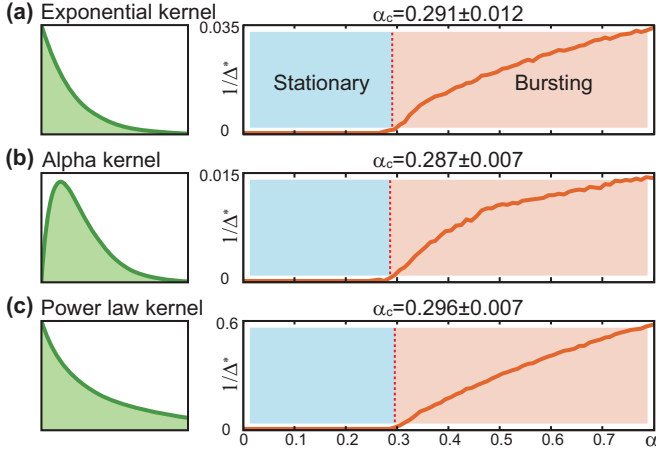
FIG. 2. (Color online) Transitions in self-exciting processes of different kernels. Inverse optimal bin size $1/\Delta^*$ is plotted with the excitability α. Event sequences were generated with the exponential kernel, $f(t) = 0.1\exp(-0.1t)$ (a), the alpha kernel, $f(t) = 0.01t\exp(-0.1t)$ (b), and the power law kernel, $f(t) = 2(1+t)^{-3}$ (c). The bin size Δ^* was selected by applying the method of optimizing a histogram [18] to every sequence of 500 000 events generated by the Hawkes process. The ranges of critical excitabilities α_c are 95% confidence intervals estimated with regression analysis.

Thus the rate fluctuation in the self-exciting point process is detectable or undetectable, respectively, if the excitability is larger or smaller than the critical value of

$$\alpha_c = 1 - 1/\sqrt{2} \approx 0.2929. \qquad (16)$$

Note that this bursting transition occurs with the excitability much smaller than $\alpha = 1$, at which the explosion of events occurs. Sample series of events generated with the excitability larger and smaller than the critical value are demonstrated in Fig. 1, from which we may observe burst of event occurrences and the apparent absence of rate fluctuation, respectively.

The detectability of rate fluctuation can be quantitatively examined using principled rate estimation methods. We first generated event sequences of self-exciting processes with three kinds of kernels; the exponential function $f(t) = \tau^{-1}e^{-t/\tau}$, the alpha function $\tau^{-2}te^{-t/\tau}$, and the power law function $2(1+t)^{-3}$. For each series of events obtained under given excitability α, we determined the optimal bin size Δ^* by using a method that enables to minimize the expected MISE even

without knowing the underlying rate [18]. Figure 2 shows how the inverse of the optimal bin size varies with the excitability. We observe that $1/\Delta^*$ begins to deviate from 0 when the excitability α exceeds some critical value. The departure from 0 can be approximated as linear in $\alpha - \alpha_c$, because the bifurcation is transcritical [19]. The critical excitabilities α_c estimated by the linear regression analysis applied to an interval $\alpha \in [0.31, 0.36]$ are consistent with the theoretical value.

V. MULTIDIMENSIONAL HAWKES PROCESS

Finally, we extend the theory to multidimensional self-exciting processes to discuss the criticality in networks of agents, such as people communicating with emails. Let $\lambda_i(t)$ and $v_i(t)$ represent the occurrence rate and a series of events of the ith agent ($i = 1, 2, \ldots, N$). The multidimensional process is given by

$$\lambda_i(t) = \rho_i + \sum_{j=1}^{N} \alpha_{ij} \int_{-\infty}^{\infty} f_{ij}(t-s) v_j(s)\, ds, \qquad (17)$$

where ρ_i is the base rate and α_{ij} and f_{ij} represent the interagent excitability or the supplementary probability and the kernel function for ith agent caused by an event of jth agent, respectively.

Given an excitability matrix $\boldsymbol{\alpha} = \{\alpha_{ij}\}$, the average firing rate $\langle \boldsymbol{\lambda} \rangle = \{\langle \lambda \rangle_i\}$ is obtained from $\boldsymbol{\rho} = \{\rho_i\}$ as

$$\langle \boldsymbol{\lambda} \rangle = \boldsymbol{C} \cdot \boldsymbol{\rho}, \qquad (18)$$

where $\boldsymbol{C}$ represents effective connections [9]:

$$\boldsymbol{C} \equiv \sum_{n=0}^{\infty} \boldsymbol{\alpha}^n = (\boldsymbol{I} - \boldsymbol{\alpha})^{-1}. \qquad (19)$$

Thus all eigenvalues of $\boldsymbol{\alpha}$ should be smaller than 1 to avoid explosion of event occurrences.

We represent the correlation functions of the rate fluctuation of agents or event sources by a matrix $\boldsymbol{\phi}(s)$. Similarly to the one-dimensional process (9), we may obtain the Fourier image of the correlation matrix $\tilde{\boldsymbol{\phi}}_\omega$ [17]. In particular, we may obtain the Fourier zero mode as $\tilde{\boldsymbol{\phi}}_0 = \boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^T - \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\langle \boldsymbol{\lambda} \rangle)$. From this, we can obtain the condition under which a network of agents exhibits fluctuating rate. That is

$$\max_i (\boldsymbol{C}\boldsymbol{\Lambda}\boldsymbol{C}^T\boldsymbol{\Lambda}^{-1})_{ii} > 2. \qquad (20)$$
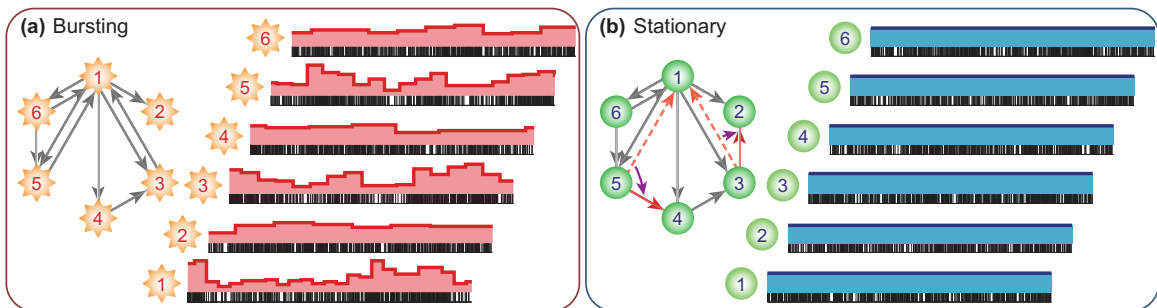

FIG. 3. (Color online) Multidimensional self-exciting processes. Networks of $N = 6$ agents linked with 10 connections of the excitability $\alpha_{ij} = 0.4$ with the kernel $f(t) = 0.2\exp(-0.2t)$. (a) A network exhibits bursts of event occurrences. (b) Bursting is inhibited by reconnecting agents (from dashed lines to solid lines).

Note that this is the condition that at least one agent exhibits bursting.

It is possible to incite or inhibit the bursting by changing the connections among agents by taking account of the above condition. We demonstrated this by simulating a network of $N = 6$ agents linked with 10 connections. The network satisfying the condition (20) exhibits bursts of event occurrences, whereas we are able to eliminate the bursting by reconnecting the agents (Fig. 3).

VI. DISCUSSION

In this study, we have shown that a linear self-exciting point process undergoes a transition at which the rate fluctuation changes from "invisible" to "visible". It should be noted that this transition does not belong to typical critical phenomena: Because the event generation process is linear, the correlation function varies smoothly with the excitability. The criticality is derived from observers in such a way that the rate fluctuation becomes "visible" in the sense of L^2 measure or maximizing the marginal likelihood. In deriving this criticality condition, we assumed a second-order phase transition in which $1/\Delta^*$ continuously deviates from 0. It is possible that the occurrence rate destabilizes from a finite time scale under the smaller excitability, but in practice this occurs only if the kernel possesses a very strong oscillation component of a finite time scale. We further extended the analysis to multi-dimensional processes and derived the necessary condition for bursting. Knowing the criticality condition, Eq. (20), we can control bursts of events in agent networks by reorganizing the connectivity among the agents.

[1] A. G. Hawkes, Biometrika **58**, 83 (1971).
[2] Y. Ogata, J. Am. Stat. Assoc. **83**, 9 (1988).
[3] P. Reynaud-Bouret and S. Schbath, Ann. Statist. **38**, 2781 (2010).
[4] G. O. Mohler *et al.*, J. Am. Stat. Assoc. **106**, 100 (2011).
[5] C. Blundell, J. Beck, and K. A. Heller, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., New York, 2012), p. 2609.
[6] L. Mitchell and M. E. Cates, J. Phys. A **43**, 045101 (2010).
[7] N. Masuda *et al.*, in *Temporal Networks*, edited by P. Holme and J. Saramäki (Springer, New York, 2013), p. 245.
[8] M. Krumin *et al.*, Front. Comput. Neurosci. **4**, 147 (2010).
[9] V. Pernice and S. Rotter, J. Stat. Mech.: Theor. Exp. (2013) P03008.
[10] S. Koyama and S. Shinomoto, J. Phys. A Math. Gen. **37**, 7255 (2004).

[11] S. Koyama *et al.*, J. Phys. A Math. Theor. **40**, F383 (2007).
[12] M. A. de Menezes and A. L. Barabasi, Phys. Rev. Lett. **92**, 028701 (2004).
[13] D. Rybski *et al.*, Sci. Rep. **2**, 560 (2012).
[14] A. L. Barabasi, Nature (London) **435**, 207 (2005).
[15] D. Sornette, F. Deschatres, T. Gilbert, and Y. Ageon, Phys. Rev. Lett. **93**, 228701 (2004).
[16] R. Crane and D. Sornette, Proc. Natl. Acad. Sci. USA **105**, 15649 (2008).
[17] A. G. Hawkes, J. R. Stat. Soc. Ser. B (Methodol.) **33**, 438 (1971).
[18] H. Shimazaki and S. Shinomoto, Neural Comput. **19**, 1503 (2007).
[19] T. Shintani and S. Shinomoto, Phys. Rev. E **85**, 041139 (2012).