Asymptotic solutions to nonlinear Hawkes processes: A systematic classification of the steady-state solutions

Kiyoshi Kanazawa[1] and Didier Sornette[2]

[1]*Faculty of Engineering, Information and Systems, University of Tsukuba, Tennodai, Tsukuba, Ibaraki 305-8573, Japan*
[2]*Institute of Risk Analysis, Prediction and Management, Academy for Advanced Interdisciplinary Studies,*
Southern University of Science and Technology, Shenzhen, Guangdong Province 518055, China

The linear Hawkes point process is a first-order non-Markovian stochastic model of intermittent bursty dynamics. While its nonlinear extensions, called nonlinear Hawkes processes, are expected to be more powerful in describing the coexistence of excitatory and inhibitory effects (or negative feedback) as occurs, for instance, in seismic and neural systems, such nonlinear Hawkes processes have been found hitherto to be analytically intractable due to the interplay between their non-Markovian and nonlinear characteristics, with no analytical solutions available. Here we systematically classify the solutions of the nonlinear Hawkes processes and then present their various exact/asymptotic solutions using the field master equation approach introduced previously by us. We report explicit power-law formulas for the steady-state intensity distributions $P_{ss}(\lambda) \propto \lambda^{-1-a}$, where the tail exponent a is expressed analytically as a function of parameters of the nonlinear Hawkes models. We introduce the basic analytical tools for advanced Hawkes modeling, particularly for model calibration to time-series data in various complex systems.

I. INTRODUCTION

Intermittent bursts are ubiquitously observed with temporal and spatial clustering characters in physical [1,2], seismic [3–6], epidemic [7], financial [8–10], and social systems [11,12]. Such bursty dynamics can be well described by the Hawkes process [13–15], a non-Markovian self-excited point process capturing both long-memory effects and critical bursts such that past events keep their potential influence to trigger future bursty events for a long time, potentially leading to critical bursts. However, the essential non-Markovian nature of this model has been an obstacle preventing the development of a unified analytical theory because the established framework of Markovian stochastic processes is not applicable.

Recently, however, a new theoretical scheme was developed to address such non-Markovian stochastic processes directly, in particular for the Hawkes process [16,17]. This scheme is based on a mapping from the non-Markovian Hawkes model to an equivalent stochastic partial differential equation (SPDE). The SPDE is then mapped to an equivalent field master equation (ME), i.e., a functional-differential equation for the probability density functional (PDF) of the intensity. The solutions of this equation can be obtained analytically in their asymptotic form, in particular near criticality. This theoretical framework predicted a novel nonuniversal power-law relation for the intensity as an intermediate asymptotics [18]. It has the potential for further explorations of the theoretical properties of more general Hawkes processes.

Since the basic linear Hawkes (LH) process is analytically solved in this framework, it is natural to seek further generalization of the framework, such as for nonlinear Hawkes (NLH) processes [19,20]. Nonlinear Hawkes processes are particularly important to account for the presence of inhibitory effects: In addition to positive feedback, many systems are characterized by coexisting negative feedback. In the context of point processes, while the standard Hawkes process describes only excitatory processes, many systems are kept in balance by the additional occurrence of inhibitory processes. For instance, inhibitory effects naturally appear in seismicity [21,22] as any earthquake creates a tensorial stress perturbation within the earth's visco-plasto-elastic crust with the presence of "stress shadows" in certain regions around the ruptured fault where future earthquakes are less likely [23], while other regions are brought closer to rupture by an increase in the local relevant stress component. Similarly, neurobiological brains are kept in balance by the interplay between excitatory and inhibitory neurotransmitters, with the resulting cascades of excitations exhibiting power-law statistics [24–26].

The long-standing problem of combining inhibitory and excitatory effects in point processes requires considering nonlinear extensions of the Hawkes processes in order to fulfill the condition that the intensity (a probability per unit time) remains non-negative. Our recent Letter [27] presented a step toward a general theory of NLH processes by applying the framework of the field ME [16,17]. In that Letter we discovered the existence of an asymptotic ubiquitous power-law distribution of the intensity for NLH processes in the case of

mark distributions with nonpositive mean. Since NLH processes may have a huge variety of forms, and thus of control parameters, for instance, in the tension-intensity map defined below and in the mark distribution, it would be useful to further study various NLH processes by systematically classifying their solutions according to the asymptotic analyses of the field MEs.

The present article supplements our Letter [27] by providing a systematic classification of various NLH processes, together with various explicit exact and asymptotic solutions. In this paper we present a general formulation for the NLH processes and provide their explicit solutions for various cases. In particular, we report three interesting asymptotic features which are valid for a wide class of memory kernels. (i) In the absence of inhibitory effects, i.e., when events all have positive marks, we find a nonuniversal power-law relation for the intensity distribution at criticality, with an exponent a that can take any value, i.e., corresponding to a genuine power law ($a > 0$) or to an intermediate power-law asymptotics ($a \leqslant 0$). This is in contrast to the LH model, where only a negative exponent $a < 0$ exists [16,17]. (ii) In the presence of inhibitory effects, i.e., both positive and negative marks coexist, in the case where the mark distribution has zero mean corresponding to a balance between inhibitory and excitatory effects, a wide class of NLH processes exhibit Zipf's law ($a \approx 1$) for their intensity distributions. (iii) For negative-mean marks, we derive the asymptotic formula for the intensity probability density function (PDF), whose tail becomes thinner than in the zero-mean mark case. This provides a different mechanism for the ubiquity of power laws, including Zipf's law, in the form of a universal property of the NLH family composed of intensity maps growing sufficiently fast as a function of the tension (to be defined below) and with balanced marks.

This article is organized as follows. We present the detailed mathematical formulation of the NLH processes in Sec. II. In Sec. III the NLH processes are mapped onto Markovian SPDEs, whose time evolution are described by MEs. We also develop a mathematical scheme to analyze the MEs, such as the functional Kramers-Moyal expansion and system size expansion for the diffusive limit. We refer to Sec. III G for a presentation of the flow chart of the methodology and our main results. In Sec. IV we study the exact solutions to NLH processes with an exponential-memory kernel without inhibitory effect, i.e., only the positive feedback effects are taken into account. In Secs. V and VI we study the exact solutions of NLH processes with an exponential-memory kernel and in the presence of inhibitory effects, i.e., when both positive and negative feedback are considered. In Secs. VII and VIII we present the asymptotic solutions of the NLH models with an arbitrary memory in the absence and presence of inhibitory effects, respectively. Section IX discusses future possible extensions and progress that can be derived from our present work. Section X summarizes and is followed by nine Appendixes presenting detailed derivations omitted from the main text for the sake of conciseness.

Readers interested only in the overview of our results should go to Sec. II and Tables I and II. Indeed, all our results are summarized in Tables I and II,. which map the inputs of the model, i.e., setups, to the outputs, i.e., the resultant asymptotic PDFs.

II. SETUP

We first introduce the mathematical notation used to define the NLH model. We then review the NLH processes and their applications for real data analysis of complex systems, to highlight their utility and importance in various contexts.

A. Mathematical notation

We denote any stochastic variable $\hat{A}$ with a circumflex to distinguish it from a nonstochastic real number A. The ensemble average of any stochastic variable $\hat{A}$ is written as $\langle \hat{A} \rangle$. The probability density function is denoted by $P_t(A) := \langle \delta(A - \hat{A}(t)) \rangle$, which characterizes the probability that $\hat{A}(t) \in [A, A + dA]$ as $P_t(A)dA$. Using the probability density function, the ensemble average can be rewritten as

$$\langle \hat{A}(t) \rangle := \int A P_t(A) dA. \tag{1}$$

We define the real number space by $\mathcal{R}$. Its nonnegative part is defined as $\mathcal{R}^+ := \{x \mid x \geqslant 0, \ x \in \mathcal{R}\}$. The K-dimensional real number space is defined as $\mathcal{R}_K := \{(x_1, \ldots, x_K) \mid x_k \in \mathcal{R}, \ k = 1, 2, \ldots, K\}$ and its nonnegative part is written as $\mathcal{R}_K^+ := \{(x_1, \ldots, x_K) \mid x_k \in \mathcal{R}, \ x_k \geqslant 0 \text{ for } k = 1, 2, \ldots, K\}$. We also define the functional space by $\mathcal{S}_F$. For example, a function f defined on $\mathcal{R}^+$ is in the function space $\mathcal{S}_F$ such that $\{f(x)\}_{x \in \mathcal{R}^+} \in \mathcal{S}_F$.

In this paper functionals, i.e., maps from a function space $\mathcal{S}_F$ to a real number space $\mathcal{R}$, appear to characterize the path probability density. For any $\{z(x)\}_{x \in \mathcal{R}_+} \in \mathcal{S}_F$, a functional f is defined as $f[z] := f[\{z(x)\}_{x \in \mathcal{R}^+}]$. Here the square brackets emphasize that f is a functional, i.e., its argument is a function, but not an ordinary function.

For a stochastic variable $\{\hat{z}(t, x)\}_{x \in \mathcal{R}^+}$ defined on a field $x \in \mathcal{R}^+$, the PDF is written as $P_t[z] := \langle \delta[z - \hat{z}] \rangle = P_t[\{z(x)\}_{x \in \mathcal{R}^+}]$, with the δ functional $\delta[z - \hat{z}] := \prod_{x \in \mathcal{R}^+} \delta(z(x) - \hat{z}(t, x))$. Here the PDF is defined over paths so that probability weighted quantities involve path integrals. For instance, the ensemble average is defined by

$$\langle \hat{A}(t) \rangle = \int A(t) P_t[z] \mathcal{D}z, \quad \mathcal{D}z := \prod_{x \in \mathcal{R}^+} dz(x), \tag{2}$$

where $\mathcal{D}z$ is the path-integral volume element.

B. Model

Let us now formulate the marked NLH process studied in this paper. Let us consider an internal variable $\hat{v}(t)$ that represents the total tension of the system and which obeys a non-Markovian stochastic differential equation (SDE)

$$\hat{v}(t) = \sum_{i=1}^{\hat{N}(t)} \hat{y}_i h(t - \hat{t}_i), \tag{3a}$$

where $\{\hat{y}_i\}_i$ is an independent and identically distributed random sequence of random numbers (jumps) obeying a distribution $\rho(y)$, $h(t)$ is a non-negative memory kernel, $\{\hat{t}_i\}_i$ is a Poisson process conditional on a time-dependent intensity $\hat{\lambda}(t)$, and $\hat{N}(t)$ is the total number of events during $[0, t)$ (called counting process). The jump size $\hat{y}_i$ is called a mark in the point process literature. Here the intensity $\hat{\lambda}(t)$ is assumed
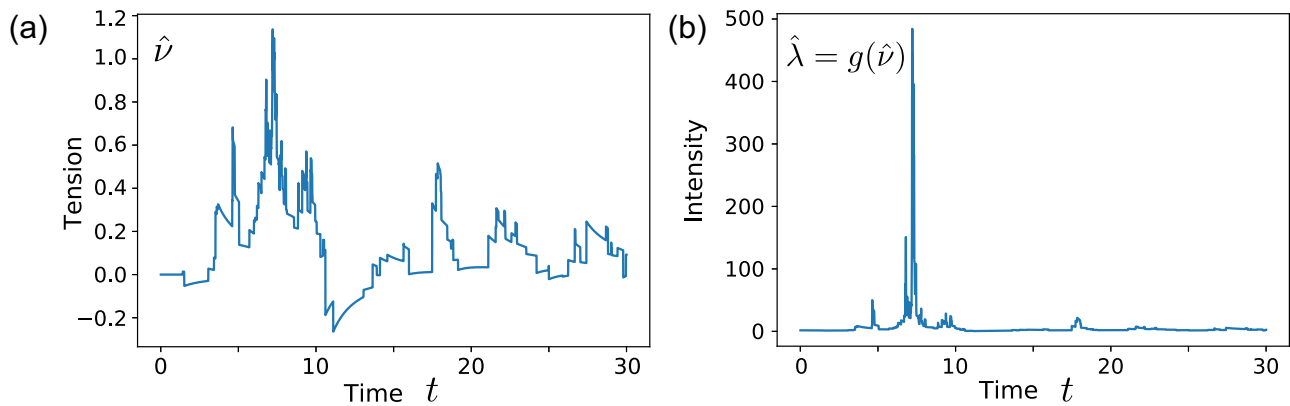
FIG. 1. Sample trajectory of (a) the tension $\hat{v}$ and (b) the intensity $\hat{\lambda} = g(\hat{v})$ in the NLH process (3). The functions and model parameters are $\rho(y) = e^{-x^2/2}/\sqrt{2\pi}$, $g(v) = e^{\beta v + v_0}$, $\beta = 5$, and $v_0 = 0.5$. The memory kernel $h(t)$ is defined by the expression (14) with $\tilde{h}(x) = \{\eta_{\text{ini}} + c_{\text{ini}}(x - x_{\text{ini}})\}/x$ for $x \in [x_{\text{ini}}, x_{\text{fin}}]$ and $\tilde{h}(x) = 0$ for $x \notin [x_{\text{ini}}, x_{\text{fin}}]$, with $x_{\text{ini}} = 0.5$, $x_{\text{fin}} = 10$, $\eta_{\text{ini}} = 0.021$, and $c_{\text{ini}} = 0.0044$, and we use the discrete step size $dx = 0.0475$.

to be stochastic and is a non-negative nonlinear function of the total tension $\hat{v}(t)$, defined as

$$\hat{\lambda}(t) = g(\hat{v}(t)) > 0. \tag{3b}$$

In this paper we call $g(\hat{v})$ the tension-intensity map or intensity function. The intensity is the probability per unit time for an event to be triggered: Assuming $\hat{N}(t) = k$, $\hat{\lambda}(t)dt$ gives the probability that $\hat{t}_{k+1} \in [t, t + dt)$ for an infinitely small time interval $dt \to 0$. We can rewrite Eqs. (3a) and (3b) as

$$\hat{\lambda}(t) = g\left(\sum_{i=1}^{\hat{N}(t)} \hat{y}_i h(t - \hat{t}_i)\right). \tag{3c}$$

This is the fundamental dynamical equation governing the NLH processes. See Fig. 1 for a schematic trajectory.

In this article we particularly focus on power-law forms of the steady PDF of the intensity for large λ as

$$P_{\text{ss}}(\lambda) = \lim_{t \to \infty} \langle \delta(\lambda - \hat{\lambda}(t)) \rangle \propto \lambda^{-1-a}, \tag{4}$$

where a is the exponent of the complementary cumulative distribution function.[1]

Remark. The model (3) is a natural nonlinear generalization of the conventional (linear) Hawkes process. Indeed, the LH process is recovered by choosing a linear intensity function

$$\lambda = g(v) = v + v_0, \tag{5}$$

assuming both $h(t)$ and v are non-negative. In contrast to the conventional Hawkes process, we do not assume non-negativity of $\hat{v}$ and $\hat{y}_i$ for the case of the general non-negative nonlinear intensity function $g(\hat{v})$.

For the LH process, the integral of the memory kernel

$$\eta := \int_0^\infty h(t)dt \tag{6}$$

is an important parameter (called the branching ratio) since it controls the fertility of events to trigger descendants (triggered events). Indeed, the LH process is subcritical for $\eta < 1$, critical at $\eta = 1$, and supercritical for $\eta > 1$.

C. Motivation and literature review

We now present a brief self-contained review of the existing literature on NLH processes for statistical physics readers who may be unfamiliar with this topic. Readers interested only in our main results may skip this section.

Nonlinear Hawkes processes were first introduced by Brémaud and Massoulié [19], who were concerned with general conditions for the existence of the processes. Since then, there have been a few applications to seismic, financial, and neural modeling, in particular for empirical comparisons. However, beyond the derivation of general conditions for existence, obtaining analytical solutions of these models is very difficult due to the complex interplay between their nonlinear and non-Markovian structures. Only a few studies exist, such as the analysis of the stability of these processes (conditions for nonexplosiveness) [19], a special solution for the Zumbach Hawkes processes with an exponential memory in the diffusive limit [28], and an asymptotic analysis for high-baseline intensity using the functional central limit theorem [29].

There are several motivations for introducing NLH processes. Here we focus on two interesting properties: (i) inhibitory effects and (ii) physical underpinning of the nonlinear tension-intensity maps. Indeed, one of the motivations for introducing NLH processes is to describe inhibitory effects [20] such that previous events can produce negative feedback effects on the total tension $\hat{v}$. For simplicity, let us consider the case where the tension-intensity map $\lambda = g(v)$ is an increasing function. For this setup, an event with positive mark $\hat{y}_i > 0$ is likely to induce future events and, inversely, an event with negative mark $\hat{y}_i < 0$ is likely to inhibit future events. This means that negative marks $\hat{y}_i < 0$ represent inhibitory effects, while positive marks $\hat{y}_i > 0$ represent excitatory effects.

---

[1] We used the PDF exponent $a_{\text{PDF}} = 1 + a_{\text{CCDF}}$ for the description of the power-law relations in Refs. [16,17], where CCDF stands for the complementary cumulative distribution function.

To implement such inhibitory effects, nonlinearity in the tension-intensity map is essential because the LH process cannot accommodate inhibitory effects. Indeed, if we assume an affine tension-intensity map $g(\nu) = \nu_0 + \nu$ with non-negative constant ν_0, ν must take a value larger than $-\nu_0$ in order for the tension-intensity map to remain non-negative. This condition requires that the mark distribution must be one sided toward the positive direction, i.e., $\rho(y) = 0$ for $y < 0$; otherwise, ν takes a value smaller than $-\nu_0$ with nonzero probability and the model assumption is violated. In fact, the model cannot be defined as a negative intensity or probability density cannot be given mathematical sense.

The second nice property of NLH processes is that the nonlinearity of the tension-intensity maps captures in a natural way the real mechanisms occurring in the modeled systems. Let us illustrate this point by reviewing several versions of the NLH processes studied in the literature.

1. Example 1: Seismic modeling

One of the most illustrative cases is found in the modeling of statistical seismicity. Let us regard the tension $\hat{\nu}(t)$ as the total stress component along the fault best oriented for rupture at a given point $\vec{r}$ in the earth's crust. Let $\hat{t}_i$ be the time of occurrence of the ith earthquake. This earthquake creates a tensorial stress field that adds to the preexisting stress field. Again, for our discussion, we simplify the picture by taking this stress perturbation as being a scalar, for instance, the Mohr-Coulomb stress amplitude along the fault best oriented for rupture at point $\vec{r}$. Furthermore, we take into account the viscoelastic property of the crust, which means that a stress perturbation is progressively relaxed via a memory kernel $h(t - \hat{t}_i)$ that tends to 0 at long times. Then the total stress at $\vec{r}$ is obtained as the sum of the stress perturbations created by all past earthquakes

$$\hat{\nu}(t) = \sum_{i=1}^{N(t)} \hat{y}_i h(t - \hat{t}_i). \tag{7}$$

Note that the marks $\hat{y}_i$ can be positive (negative), corresponding to the ith earthquake bringing the point $\vec{r}$ closer to (further away from) failure. The former case is the most intuitive and represents the stress load on $\vec{r}$ due to the redistribution of forces by the earthquake fault slip in its neighborhood, especially close to its fault tips and in its stress lobes of positive influence. The latter case is known as stress shadow [23] and is associated with the tensorial nature of the stress disturbances induced by an earthquake. Given the stochasticity in the distribution of earthquake sizes, in their positions and orientations, the marks $\hat{y}_i$ are stochastic variables. Given the total stress (tension) (7), the next ingredient is to recognize that mechanical rupture and earthquakes are thermally activated with an effective inverse temperature β that is renormalized via the quenched heterogeneity of the medium [30–32]. Then the probability for the next earthquake to occur is given by the Arrhenius formula, thus formulating the intensity λ as a decreasing exponential function $e^{-\beta \Delta E(t)}$ of the energy barrier $\Delta E(t)$ for nucleation. The key point is to approximate the energy barrier as a decreasing affine function of the stress field $\Delta E(t) = E - \hat{\nu}(t)$, where E is a constant. Altogether,

this yields

$$\lambda(t) = \lambda_0 e^{\beta \hat{\nu}(t)}. \tag{8}$$

We finally obtain the NLH with an exponential intensity

$$\hat{\lambda}(t) = g\left(\sum_{i=1}^{N(t)} \hat{y}_i h(t - \hat{t}_i)\right), \quad g(\hat{\nu}) := \lambda_0 e^{\beta \hat{\nu}}. \tag{9}$$

In addition, given that the prediction of earthquake magnitudes is empirically very difficult (while the short-term prediction of their rates is rather possible [33]), it is a plausible assumption that the marks are drawn independently of the current tension $\hat{\nu}(t)$.

In this simplified presentation, we have restricted our attention to the temporal version of the general formulation, which is known as the multifractal stress activation (MSA) model [21,22] and involves space in addition to time in the formulation of the tension and intensity. It is remarkable that both inhibitory effects and nonlinear intensity function appear naturally for this system, as the result of the random stress perturbations induced by earthquakes and from the Arrhenius law (renormalized by quenched disorder), respectively.

It should noted that Refs. [21,22] offered only an approximate scaling theory to derive magnitude-dependent Omori law exponents and that no analytical results exist for the MSA model or for its temporal-only version (9).

2. Example 2: Financial modeling

Reference [34] is one of the very first uses in finance of the LH process (in its bivariate form) in order to model the joint dynamics of trades and midprice changes of the NYSE. Reference [35] provided the first quantitative framework using the LH process to study and quantify the level of endogeneity (or reflexivity) of market fluctuations. The basic idea is that trades and price changes are analyzed by investors (humans or machines) as one of the useful information channels to improve trading decisions, on the basis (or belief) that past actions reveal intentions and that there is a persistence in price trends, volume, volatility, and more generally of trading activity. In this sense, the self-exciting Hawkes process is a natural candidate to model the point processes of discrete trades and midprice changes [36].

As an improved model, a nonlinear version of the Hawkes process was introduced by Blanc *et al.* [28], where the intensity dynamics is given by a quadratic extension to the standard Hawkes process

$$\hat{\lambda}(t) = \lambda_0 + \int_{-\infty}^{t} L(t-s) \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s) ds$$

$$+ \int_{-\infty}^{t} ds \int_{-\infty}^{t} du\, K(t-s, t-u)$$

$$\times \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s) \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(u)}(u), \tag{10}$$

where $\rho(y) = \frac{1}{2}[\delta(y-1) + \delta(y+1)]$ and the term $\hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s)$ is the compound Poisson process with intensity $\hat{\lambda}(t)$ and jump size distribution $\rho(y)$ as defined below by the expression (14c). This model is called the quadratic Hawkes processes and has been analyzed theoretically in Ref. [28].

Since this model is nonlinear and non-Markovian, its systematic analysis is difficult and only limited results are available. However, by assuming $K(t, s) = h(t)h(s)$ and $L(t) = 0$, this model reduces to a simpler NLH process

$$\hat{\lambda}(t) = g\left(\int_{-\infty}^{t} h(t-s)\hat{\xi}_{\rho(y);\hat{\lambda}(s)}(s)ds\right) = g\left(\sum_{i=1}^{\hat{N}(t)} \hat{y}_i h(t - \hat{t}_i)\right),$$

$$g(\nu) := \lambda_0 + \nu^2, \tag{11}$$

where we have used $\hat{\xi}_{\rho(y);\hat{\lambda}(s)}(s) = \sum_{i=1}^{\hat{N}(s)} \hat{y}_i\delta(s - \hat{t}_i)$. This NLH process is a special case of the Zumbach Hawkes process, without the Hawkes feedback. While the Zumbach Hawkes process is simpler than the quadratic Hawkes process, it is still difficult to solve analytically. Therefore, the analysis in Ref. [28] focused on the special case of an exponential memory $h(t) = (\eta/\tau)e^{-t/\tau}$ and considered the diffusive limit.[2] For this special case, the steady-state PDF of the intensity obeys a power law with nonuniversal exponent

$$P_{ss}(\lambda) \propto \lambda^{-1-a}, \quad a = \frac{1}{2} + \frac{1}{2n_Z}, \tag{12}$$

with a constant n_Z called the Zumbach norm (see Ref. [28] for details). It is remarkable that a power-law relation (12) appears even for short-memory kernels without introducing any power-law distributions. This special solution was the only available analytical solution for a NLH process before our work [27].

One of the main claims in Ref. [28] is that the power-law relation (12) provides a validation step supporting the relevance of the quadratic Hawkes process for financial data analyses, because it matches the empirical power-law price-change distribution, which is a well-known stylized fact in market microstructure. From this viewpoint, the authors of Ref. [28] claim that the quadratic Hawkes process is a minimal generalization beyond the LH process that is essentially needed to account for empirical facts.

3. Example 3: Self-excited multifractal model

It is also useful to mention the self-excited multifractal model [37], which is not *per se* a point process but got its inspiration from self-excited point processes, the concept of reflexivity [38], the multifractal random-walk model [39], and its generalizations [40,41]. Reminiscent of a NLH model with a much stronger nonlinearity than quadratic, the self-excited multifractal model is defined such that the amplitudes of the increments of the process are expressed as exponentials of a long memory of past increments

$$dX(t) = \sigma \exp\left(-\frac{1}{\sigma}\int_{-\infty}^{t} h(t - t')dX(t')\right)dW(t), \tag{13}$$

———————

[2]Blanc *et al.* call their analysis the low-frequency asymptotics, taking the long-time limit and a constant endogeneity rescaling. This asymptotic method is essentially equivalent to the diffusive limit in the framework of the system size expansion, a traditional asymptotic analysis developed for statistical physics, which is formulated in Sec. III E.

where $dW(t)$ is the increment of the regular Wiener noise process, $h(t)$ is a memory kernel function, and σ controls the amplitude of the noise as well as the dimension and scale of $X(t)$. Interpreting $dX(t)$ as a log-return of a financial price, the self-excited multifractal process recovers all the standard stylized facts documented in empirical financial time series. The exploration of the links between the self-excited multifractal model and the exponential NLH process is left for the future.

4. Goal of this study: Solutions for various nonlinear Hawkes processes

The above summaries highlight the fact that analytical solutions for NLH processes have not been obtained yet, except for special cases (such as the Zumbach Hawkes case with exponential memory in the diffusive limit). In this context, our goal is to systematically classify NLH processes and then provide analytical (both exact and asymptotic) solutions for various NLH processes, in particular for the steady-state intensity PDF $P_{ss}(\lambda)$. All our results are summarized as Tables I and II, with the mapping between the inputs of the model, i.e., setups, to the outputs, i.e., the resultant asymptotic PDFs.

III. MASTER EQUATION FORMULATION

In this section we introduce an analytical framework for the general NLH process based on the field MEs. We first provide a Markovian mapping from the original non-Markovian NLH process to a Markovian SPDE. We then derive the corresponding field ME for any memory kernel, which is shown to simplify for the special case of an exponential-memory kernel. We next develop two useful tools that have a long tradition in the history of physical stochastic processes: the Kramers-Moyal (KM) expansion and the system size expansion (SSE) for the diffusive limit. The field ME is then shown to reduce to the functional Fokker-Planck equations (FPEs) for a special case.

A. Mapping to Markovian SPDEs

Following Ref. [16], let us present the mapping from the original non-Markovian stochastic process (3) to Markovian SPDEs. Let us decompose the total tension $\hat{v}(t)$ and the memory kernel $h(t)$ as continuous sums

$$h(t) = \int_0^\infty dx\, \tilde{h}(x)e^{-t/x}, \quad \hat{v}(t) = \int_0^\infty dx\, \hat{z}(t, x). \tag{14a}$$

The intuition behind this decomposition is that the memory kernel is decomposed into a continuous sum of exponential terms with amplitude $\tilde{h}(x)$. This then suggests to use x as an auxiliary field $x \in (0, \infty)$ and then to decompose the tension as a continuous sum over the "excess tensions" $\hat{z}(t, x)$. The excess tensions are assumed to satisfy the SPDEs

$$\frac{\partial \hat{z}(t, x)}{\partial t} = -\frac{\hat{z}(t, x)}{x} + \tilde{h}(x)\hat{\xi}_{\rho(y);\hat{\lambda}(t)}^{CP},$$

$$\hat{\lambda}(t) = G[\hat{z}] := g\left(\int_0^\infty dx\, \hat{z}(t, x)\right) \tag{14b}$$

TABLE I. Summary of the results obtained in the present work for both one-sided and two-sided mark distributions. The obtained steady-state intensity distributions of intensities are systematically classified for various NLH processes. Here FAI and MSA stand for fast-accelerating intensity $[g(v) \gg v^2]$ and multifractal stress activation model $[g(v) \propto e^{\beta v}, \beta > 0]$. In this paper we assume that the moment-generating function $\Phi(x) := \int_{-\infty}^{\infty} \rho(y)(e^{xy} - 1)dy$ exists and c^* is the positive root of $\Phi(c^*) = 0$ for $m < 0$ or is equal to zero for $m = 0$. In addition, we define $m := \int_{-\infty}^{\infty} y\rho(y)dy$, $p_+ := \int_0^{\infty} \rho(y)dy$, and $p_- := \int_{-\infty}^0 \rho(y)dy$.

Model	Mark PDF $\rho(y)$	Tension-intensity map $g(v)$	Critical?	PDF $P_{\rm ss}(\lambda)$	Exponent a	Section for $h(t) = (\eta/\tau)e^{-t/\tau}$	Section for general $h(t)$
linear	one sided: $\rho(y) = 0$ for $y < 0$	$v_0 + v_1$, $v_0 > 0$	yes ($\eta \uparrow 1$)	$\propto \lambda^{-1-a} e^{-\lambda/\lambda_{\rm cut}}$ ($\lim_{\eta\uparrow 1} \lambda_{\rm cut} = \infty$)	$a < 0$ (intermediate asymptotics, nonuniversal)	Sec. IV	Sec. VII
ramp		$\max\{v_0, v - v_1\}$, $v_0 > 0$			any real number a (nonuniversal)		
ramp	two-sided, symmetric: $\rho(y) = \rho(-y)$	$\max\{v_0, v - v_1\}$, $v_0 > 0$	no	$\propto e^{-\lambda/\lambda_{\rm cut}}$ ($\lambda_{\rm cut} < \infty$)	absent	Sec. V	not treated
quadratic		$\propto v^2$		$\propto \lambda^{-1-a}$	$a > 1/2$ (nonuniversal)		
polynomial of order $n > 2$ (FAI)		$\propto v^n$, $n > 2$			$a = 1 - 1/n$ (universal)		Sec. VIII
exponential (FAI, MSA)		$\propto e^{\beta v}$			$a = 1$ (universal, Zipf)		
Gaussian (FAI)		$\propto e^{\beta v^2}$					
FAI		$\gg v^2$		$\propto \lambda^{-1}[\frac{dg(v)}{dv}]^{-1}|_{v=g^{-1}(\lambda)}$			
ramp	two-sided, nonpositive mean: $p_+ > 0$, $p_- > 0$, $m \leqslant 0$	$\max\{v_0, v - v_1\}$, $v_0 > 0$	no	$\propto e^{-\lambda/\lambda_{\rm cut}}$ ($\lambda_{\rm cut} < \infty$)	absent	Sec. VI	not treated
exponential (FAI, MSA)		$\propto e^{\beta v}$		$\propto \lambda^{-1-a}$	$a \geqslant 1$ (nonuniversal)		Sec. VIII
FAI		$\gg v^2$		$\propto \lambda^{-1}[e^{-uv}|\frac{dg(v)}{dv}|^{-1}]_{v=g^{-1}(\lambda)}$ $u := c^*/h(0)$			

TABLE II. Second summary table of our main results, highlighting the differences from previous articles. In this paper, 22 different solutions (14 exact and 8 asymptotic solutions) are reported.

Memory kernel $h(t)$	Mark distribution $g(\nu)$	Tension-intensity map $g(\nu)$	Main results	Exact or asymptotic?	Related work	Section
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	one-sided exponential, $\frac{1}{y^*}e^{-y/y^*}\Theta(y)$	general $g(\nu)$	(59)	exact	present result	IV
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	one-sided exponential, $\frac{1}{y^*}e^{-y/y^*}\Theta(y)$	linear, $\nu+\nu_0$	(67)	exact	Refs. [16,17,50]	IV
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	one-sided exponential, $\frac{1}{y^*}e^{-y/y^*}\Theta(y)$	ramp, $\max\{\nu_0,\nu-\nu_1\}$	(70)	exact	present results	IV
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	one-sided exponential, $\frac{1}{y^*}e^{-y/y^*}\Theta(y)$	ramp, $\max\{\nu_0,\nu-\nu_1\}$	(76)	asymptotic	present results	IV
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	two-sided exponential, $\frac{1}{2y^*}e^{-|y|/y^*}$	general $g(\nu)$	(90)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	two-sided exponential, $\frac{1}{2y^*}e^{-|y|/y^*}$	ramp, $\max\{\nu_0,|\nu|\}$	(92), (96)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	two-sided exponential, $\frac{1}{2y^*}e^{-|y|/y^*}$	quadratic, $k\nu^2+\lambda_0$	(98), (100)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	two-sided exponential, $\frac{1}{2y^*}e^{-|y|/y^*}$	exponential, $\lambda_0|\nu|e^{\beta\nu}$	(104), (105)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	general $g(\nu)$	(108)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	ramp, $\max\{\nu_0,|\nu-\nu_1|\}$	(110)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	quadratic, $k\nu^2+\lambda_0$	(112)	exact	Ref. [28]	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	polynomial, $k|\nu|^n+\lambda_0$	(114), (116)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	exponential, $\lambda_0e^{\beta\nu}$	(122), (123)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	diffusive limit	FAI, $g(\nu)\gg\nu^2$	(125), (126)	exact	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	symmetric (general), $\rho(y)=\rho(-y)$	ramp, $g(\nu)\simeq\nu+\nu_0$	(127)	asymptotic	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	symmetric (general), $\rho(y)=\rho(-y)$	quadratic, $k\nu^2+\lambda_0$	(132)	asymptotic	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	symmetric (general), $\rho(y)=\rho(-y)$	exponential, $\lambda_0e^{\beta\nu}$	(138)	asymptotic	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	symmetric (general), $\rho(y)=\rho(-y)$	FAI, $g(\nu)\gg\nu^2$	(146)	asymptotic	present results	V
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	asymmetric two-sided exponential[a]	general $g(\nu)$	(162)	exact	present results	VI
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	asymmetric two-sided exponential[a]	ramp, $\max\{\nu_0,|\nu|\}$	(165)	exact	present results	VI
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	asymmetric two-sided exponential[a]	exponential, $\lambda_0|\nu|e^{\beta\nu}$	(169), (171)	exact	present results	VI
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	asymmetric (general), $\rho(y)\neq\rho(-y)$	ramp, $\max\{\nu_0,\nu-\nu_1\}$	(172c)	asymptotic	present results	VI
exponential, $\frac{\eta}{\tau}e^{-t/\tau}$	asymmetric (general), $\rho(y)\neq\rho(-y)$	FAI, $g(\nu)\gg\nu^2$	(175)	asymptotic	present results	VI
general $h(t)$	general $\rho(y)$	ramp, $\max\{\nu_0,\nu-\nu_1\}$	(198), (232)	asymptotic	present result[b]	VII
general $h(t)$	general $\rho(y)$	exponential, $\lambda_0e^{\beta\nu}$	(235), (256)	asymptotic	Ref. [27]	VIII
general $h(t)$	general $\rho(y)$	FAI, $g(\nu)\gg\nu^2$	(259)	asymptotic	Ref. [27]	VIII

[a] $p(y) = \begin{cases} \frac{p_+}{y_+^*}e^{-y/y_+^*}, & y\geqslant 0 \\ \frac{p_-}{y_-^*}e^{+y/y_+^*}, & y<0. \end{cases}$

[b] A special case was studied in Refs. [16,17].

under the initial condition $\hat{z}(t=0,x)=0$. The term $\hat{\xi}^{\text{CP}}_{\rho(y);\hat{\lambda}(t)}$ is the compound Poisson process with intensity $\hat{\lambda}(t)$ and jump size distribution $\rho(y)$,

$$\hat{\xi}^{\text{CP}}_{\rho(y);\hat{\lambda}(t)} = \sum_{i=1}^{\hat{N}(t)} \hat{y}_i\delta(t-\hat{t}_i), \qquad (14c)$$
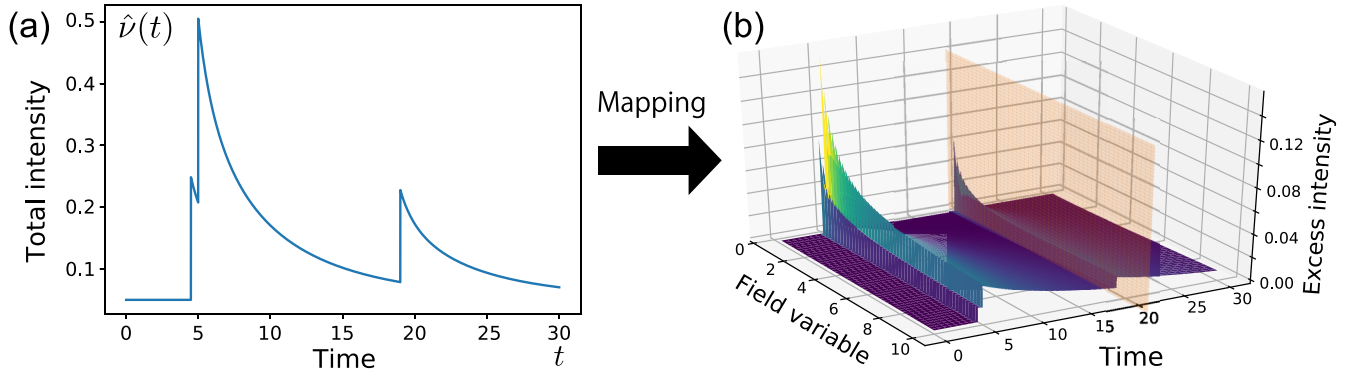
FIG. 2. Schematics of the Markovian embedding. (a) The original non-Markovian one-dimensional dynamics [described by the SDE (3)] is mapped onto (b) the Markovian field dynamics [described by the SPDE (14)].

which means that the random marks y_i obey the distribution $\rho(y)$.

This mapping can be schematically illustrated as shown in Fig. 2: The original dynamics is one dimensional, governed by the SDE (3). In this low-dimensional representation, the dynamics is non-Markovian. However, by applying the Markovian embedding, we can construct an infinite-dimensional Markovian dynamics governed by the SPDE (14) by adding sufficiently many auxiliary variables $\hat{z}(t, x)$.

Proof of equivalence

The SPDE (14b) together with the decomposition formula (14a) is equivalent to the original marked NLH process (3). Indeed, the formal solution of Eq. (14b) is given by

$$
\begin{aligned}
\hat{z}(t, x) &= \tilde{h}(x) \int_0^t dt' e^{-(t-t')/x} \hat{\xi}^{\mathrm{CP}}_{\rho(y);\hat{\lambda}(t)} \\
&= \sum_{i=1}^{\hat{N}(t)} \hat{y}_i \tilde{h}(x) \int_0^t dt' e^{-(t-t')/x} \delta(t' - \hat{t}_i) \\
&= \sum_{i=1}^{\hat{N}(t)} \hat{y}_i \tilde{h}(x) e^{-(t-\hat{t}_i)/x},
\end{aligned}
\tag{15}
$$

leading to

$$
\hat{v}(t) = \sum_{i=1}^{\hat{N}(t)} \hat{y}_i \int_0^\infty dx \tilde{h}(x) e^{-(t-\hat{t}_i)/x} = \sum_{i=1}^{\hat{N}(t)} \hat{y}_i h(t - \hat{t}_i). \tag{16}
$$

It is noteworthy that this derivation does not make explicit reference to the definition of $\hat{\lambda}(t) = g(\hat{v}(t))$ and is independent of the specific function $g(v)$.

B. Field master equation

In this section we study the functional ME corresponding to the SPDE (14b). The field ME of the PDF $P_t[z]$ is given by

$$
\frac{\partial P_t[z]}{\partial t} = (\mathcal{L}_{\mathrm{adv}} + \mathcal{L}_{\mathrm{jump}}) P_t[z], \tag{17a}
$$

with advective and jump Liouville operators

$$
\mathcal{L}_{\mathrm{adv}} P_t[z] := \int_0^\infty dx \frac{\delta}{\delta z(x)} \left(\frac{z(x)}{x} P_t[z] \right), \tag{17b}
$$

$$
\mathcal{L}_{\mathrm{jump}} P_t[z] := \int_{-\infty}^\infty dy \, \rho(y) G[z - y\tilde{h}] P_t[z - y\tilde{h}] - G[z] P_t[z]. \tag{17c}
$$

In this paper we provide various analytical exact or asymptotic solutions of (17).

1. Derivation

It is useful to provide a derivation of the field ME (17) via a discrete approach, which gives a sound mathematical interpretation and control of the functional derivatives [42]. Let us consider the case of the memory kernel composed of a discrete sum of K exponentials (which we refer to as K exponentials)

$$
h(t) = \sum_{k=1}^K \tilde{h}_k e^{-t/\tau_k}. \tag{18a}
$$

The NLH process (3) together with the K exponentials (18a) can be mapped onto a Markovian equation by introducing $\hat{z} := (\hat{z}_1, \ldots, \hat{z}_K)$,

$$
\frac{d\hat{z}_k(t)}{dt} = -\frac{\hat{z}_k(t)}{\tau_k} + \tilde{h}_k \hat{\xi}^{\mathrm{CP}}_{\rho(y);\hat{\lambda}(t)},
$$

$$
\hat{\lambda}(t) = g(\hat{v}(t)), \quad \hat{v}(t) := \sum_{k=1}^K \hat{z}_k(t), \tag{18b}
$$

which is parallel to the Markovian embedding procedure for Eqs. (14). We introduce the following function G, which will be convenient for future developments:

$$
G(\hat{z}) := g(\hat{v}) = g\left(\sum_{k=1}^K \hat{z}_k \right). \tag{18c}
$$

The ME for the SDE (18) is derived as follows. Let us consider an arbitrary function $f(\hat{z})$ and its time evolution $df(\hat{z}(t)) := f(\hat{z}(t+dt)) - f(\hat{z}(t))$ during $[t, t+dt)$,

$$df(\hat{z}) = \begin{cases} -\sum_{k=1}^{K} \frac{\hat{z}_k}{\tau_k} \frac{\partial f(\hat{z})}{\partial \hat{z}_k} dt & \text{(no jump during } [t, t+dt)\text{: probability } = 1 - \hat{\lambda}(t)dt) \\ f(\hat{z} + \hat{y}\tilde{\boldsymbol{h}}) - f(\hat{z}) & \text{(jump in } [t, t+dt) \text{ with } \hat{y} \in [y, y+dy)\text{: probability } = \hat{\lambda}(t)\rho(y)dtdy), \end{cases} \tag{19}$$

with $\tilde{\boldsymbol{h}} := (\tilde{h}_1, \dots, \tilde{h}_K)$. By taking the ensemble average of both sides over realizations of the excess tensions $\hat{z} := (\hat{z}_1, \dots, \hat{z}_K)$, we obtain

$$\langle df(\hat{z}) \rangle = \left\langle -\sum_{k=1}^{K} \frac{\hat{z}_k}{\tau_k} \frac{\partial f(\hat{z})}{\partial \hat{z}_k} dt \right\rangle + \int_{-\infty}^{\infty} dy\, \rho(y) \langle G(\hat{z})[f(\hat{z} + \hat{y}\tilde{\boldsymbol{h}}) - f(\hat{z})]dt \rangle, \tag{20}$$

which is equivalent to

$$\int_{-\infty}^{\infty} d\boldsymbol{z}\, f(\boldsymbol{z}) \frac{\partial P_t(\boldsymbol{z})}{\partial t} = \int_{-\infty}^{\infty} d\boldsymbol{z}\, P_t(\boldsymbol{z}) \left(-\sum_{k=1}^{K} \frac{z_k}{\tau_k} \frac{\partial f(\boldsymbol{z})}{\partial z_k} + \int_{-\infty}^{\infty} dy\, \rho(y) G(\boldsymbol{z})[f(\boldsymbol{z} + y\tilde{\boldsymbol{h}}) - f(\boldsymbol{z})] \right)$$

$$= \int_{-\infty}^{\infty} d\boldsymbol{z}\, f(\boldsymbol{z}) \left(\sum_{k=1}^{K} \frac{\partial}{\partial z_k} \frac{z_k}{\tau_k} P_t(\boldsymbol{z}) + \int_{-\infty}^{\infty} dy\, \rho(y)[G(\boldsymbol{z} - y\tilde{\boldsymbol{h}})P(\boldsymbol{z} - y\tilde{\boldsymbol{h}}) - G(\boldsymbol{z})P(\boldsymbol{z})] \right), \tag{21}$$

where we have performed an integration by parts and used the variable transformation $\boldsymbol{z} + y\tilde{\boldsymbol{h}} \to \boldsymbol{z}$ and the relation

$$\langle df(\hat{z}) \rangle = \langle f(\hat{z}(t+dt)) - f(\hat{z}(t)) \rangle = \int_{-\infty}^{\infty} d\boldsymbol{z}\, f(\boldsymbol{z}) P_{t+dt}(\boldsymbol{z}) - \int_{-\infty}^{\infty} d\boldsymbol{z}\, f(\boldsymbol{z}) P_t(\boldsymbol{z}) = dt \int_{-\infty}^{\infty} d\boldsymbol{z} f(\boldsymbol{z}) \frac{\partial P_t(\boldsymbol{z})}{\partial t} + O(dt^2). \tag{22}$$

Since Eq. (21) is an identity holding for any $f(\boldsymbol{z})$, we obtain the corresponding ME

$$\frac{\partial P_t(\boldsymbol{z})}{\partial t} = \sum_{k=1}^{K} \frac{\partial}{\partial z_k} \frac{z_k}{\tau_k} P_t(\boldsymbol{z}) + \int_{-\infty}^{\infty} dy\, \rho(y)[G(\boldsymbol{z} - y\tilde{\boldsymbol{h}})P(\boldsymbol{z} - y\tilde{\boldsymbol{h}}) - G(\boldsymbol{z})P(\boldsymbol{z})]. \tag{23}$$

We then proceed with the continuous limit for the memory kernel. We first rewrite

$$h(t) = \sum_{k=1}^{K} \tilde{h}_k e^{-t/\tau_k} \to \sum_{k=1}^{K} dx\, \tilde{h}(x_k) e^{-t/x_k}, \quad \hat{v}(t) = g\left(\sum_{k=1}^{K} \hat{z}_k(t) \right) \to g\left(\sum_{k=1}^{K} dx\, \hat{z}(t, x_k) \right) \tag{24}$$

for the formal replacement

$$\tau_k \to x_k, \quad \tilde{h}_k \to \tilde{h}(x_k)dx, \quad \hat{z}_k \to \hat{z}(t, x_k)dx \tag{25}$$

obtained by introducing the lattice interval dx and $x_k := kdx$. By introducing the formal functional derivative and integration for the limits $K \to \infty$ and $dx \to 0$,

$$\frac{\delta}{\delta z(x_k)}[\cdots] := \lim_{dx \to 0} \frac{1}{dx} \frac{\partial}{\partial z(x_k)}[\cdots], \quad \int_0^{\infty} dx[\cdots] := \lim_{dx \to 0} \sum_{k=1}^{K} dx[\cdots], \tag{26}$$

we obtain

$$\frac{\partial P_t[z]}{\partial t} = \int_0^{\infty} dx \frac{\delta}{\delta z(x)} \frac{z(x)}{x} P_t[z] + \int_{-\infty}^{\infty} dy\, \rho(y) G[z - y\tilde{h}]P_t[z - y\tilde{h}] - G[z]P_t[z] \tag{27}$$

and

$$h(t) = \int_0^{\infty} dx\, \tilde{h}(x) e^{-t/x}, \quad \hat{v}(t) = g\left(\int_0^{\infty} dx\, \hat{z}(t, x) \right), \tag{28}$$

which is equivalent to Eq. (17) (see Appendix A for the definition of the Dirac δ function and the functional derivative). See Appendix B for another derivation based on direct manipulation of functional derivatives.

2. Mathematical remark

Master (or Fokker-Planck) equations based on functional derivatives often appear in the description of SPDEs, such as for stochastic chemical reactions [42]. While this continuous description is a useful tool for formal calculations, unfortunately, its mathematical foundation has not been established yet. Indeed, one can easily observe that there is the potential problem of encountering a divergence, such as $[\delta/\delta z(x)]z(x)P[z] = \delta(0)P[z] + z(x)\delta P[z]/\delta z(x)$. This problem might be serious for nonlinear SPDEs even for physical observables (see the divergence problem of nonlinear stochastic chemical reaction, e.g., Chap. 13.3.3 in [42]), while it

might not be for linear SPDEs. One can find that this divergence is not serious for the LH process and the generalized Langevin equation [16] at least in understanding physical observables. Remarkably, for the generalized Langevin equation, this divergence problem is essentially the same as the one encountered in quantum field theory and can be renormalized in the same manner with which the divergence problem of the zero-point energy is solved in quantum electrodynamics. We note that, in the case of the NLH process, the SPDE (14) itself is fortunately linear, while the intensity function $g(\hat{v})$ is nonlinear.

To avoid such mathematically delicate issues, our strategy is to follow a safer interpretation that follows Ref. [42]: We regard the field ME (17) [or the FPE (48)] as a formal limit of the discrete ME (23). If we encounter a potential problem of divergence in Eq. (17), we return to the discrete ME (23) to proceed with the calculations and then return to its formal limit (17). We confirm that our main results hold for the general discrete cases (18) and we then generalize them to the continuous limit.

C. Special case: Exponential-memory kernel

Let us focus here on the simplest case of the single exponential-memory kernel

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \tag{29}$$

or equivalently

$$\tilde{h}(x) = \frac{\eta}{\tau} \delta(x - \tau), \tag{30}$$

with positive real numbers η and τ. Consistent with the definition (6), the parameter η is the branching ratio. This special case is easier to analyze analytically, since the functional ME (17) reduces to the ME for a PDF of the total tension v,

$$\frac{\partial P_t(v)}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial v} [v P_t(v)] + \int dy\, \rho(y) g\left(v - \frac{\eta y}{\tau}\right)$$
$$\times P_t\left(v - \frac{\eta y}{\tau}\right) - g(v) P_t(v). \tag{31}$$

D. Functional Kramers-Moyal expansion

One of the standard analytical prescriptions to analyze MEs is the KM expansion. The KM expansion was historically introduced for a formal validation of the FP description from MEs. This formal expansion was criticized by Van Kampen due to its ambiguous validity as an asymptotic series. Later, Van Kampen developed a mathematically sophisticated formulation in the form of the SSE [43]. We present a sound formulation of the KM functional expansion for the field ME, which will be utilized for a further generalization of the SSE in Sec. III E.

1. Exponential-memory case

To first present the key idea, let us focus on the exponential-memory case (29). By considering the expansion

$$\int dy\, \rho(y) g\left(v - \frac{\eta y}{\tau}\right) P_t\left(v - \frac{\eta y}{\tau}\right)$$

$$= \sum_{k=0}^{\infty} \frac{\alpha_k}{k!} \left(-\frac{\eta}{\tau}\right)^k \frac{\partial^k}{\partial v^k} g(v) P_t(v),$$
$$\alpha_k := \int_{-\infty}^{\infty} dy\, \rho(y) y^k, \tag{32}$$

the ME (31) can be rewritten as

$$\frac{\partial P_t(v)}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial v} [v P_t(v)] + \sum_{k=1}^{\infty} \frac{\alpha_k}{k!} \left(-\frac{\eta}{\tau}\right)^k \frac{\partial^k}{\partial v^k} g(v) P_t(v). \tag{33}$$

This is the KM expansion for the ME (31) for this special case. We have assumed that all the KM coefficients $\{\alpha_k\}_{k \geqslant 1}$ are finite, which excludes some singular classes of mark distributions, e.g., power-law mark distributions.

2. General cases

The above formulation can be generalized by considering the functional Taylor expansion (see Appendix A)

$$\int_{-\infty}^{\infty} dy\, \rho(y) G[z - y\tilde{h}] P_t[z - y\tilde{h}]$$

$$= \sum_{k=0}^{\infty} \frac{\alpha_k}{k!} \left(-\int_0^{\infty} dx\, y\tilde{h}(x) \frac{\delta}{\delta z(x)}\right)^k G[z] P_t[z] \tag{34}$$

with the KM coefficients defined by

$$\alpha_k := \int_{-\infty}^{\infty} dy\, \rho(y) y^k, \tag{35}$$

and assuming that all the KM coefficients $\{\alpha_k\}_{k \geqslant 1}$ are finite. Using this relation, the field ME can be rewritten as

$$\frac{\partial P_t[z]}{\partial t} = \int_0^{\infty} dx \frac{\delta}{\delta z(x)} \left(\frac{z(x)}{x} P_t[z]\right)$$
$$+ \sum_{k=1}^{\infty} \frac{\alpha_k}{k!} \left(-\int_0^{\infty} dx\, \tilde{h}(x) \frac{\delta}{\delta z(x)}\right)^k G[z] P_t[z]. \tag{36}$$

E. Diffusive limit: System size expansion

We next consider the diffusive limit for the mark distribution according to the SSE, by assuming that (i) the mark distribution is symmetric,

$$\rho(y) = \rho(-y), \tag{37}$$

i.e., this is the case where inhibitory effects are as prevalent as excitatory effects. This situation will be further studied in detail in Sec. V. This model is essentially different from the positive mark cases, i.e., $\rho(y) = 0$ for $y \leqslant 0$, because both positive and negative feedback effects occur with the same probability. For instance, such an assumption is natural for seismic models as the stress perturbations induced by earthquakes indeed present this symmetry (which has a complex tensorial spatial rendering; see, for instance, [44]). With this symmetry condition, all the odd-order KM coefficients are zero: $\alpha_{2k+1} = 0$ for non-negative integer k.

As the second assumption (ii), let us introduce a small parameter $\varepsilon > 0$ scaling the jump size in the original Hawkes
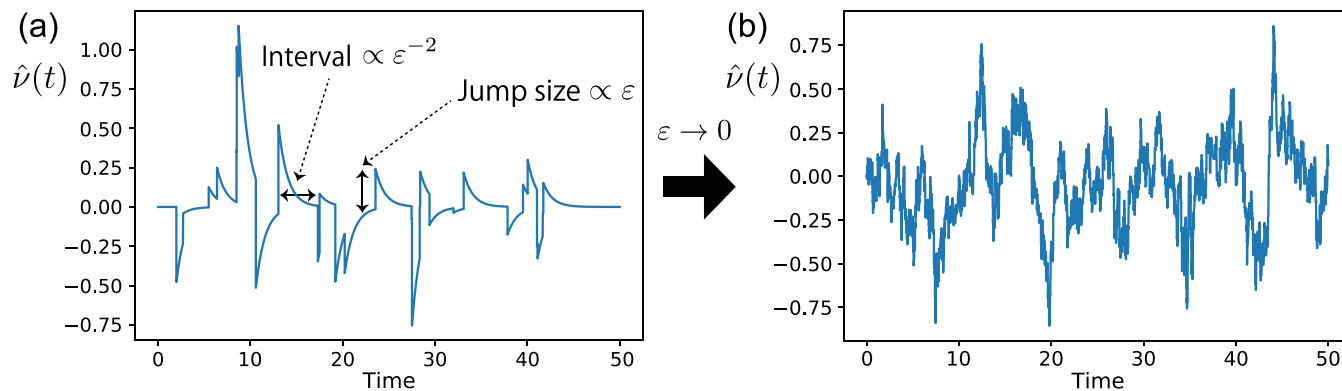
FIG. 3. Schematic trajectory of the NLH process in the diffusive limit. While (a) the trajectory is composed of sparse jumps for large ε ($\varepsilon = 1.0$), (b) the trajectory is composed of many small jumps for small ε and becomes approximately continuous ($\varepsilon = 0.1$). The trajectories were generated by assuming $h(t) = (\eta/\tau)e^{-t/\tau}$, $\rho_\varepsilon(y) = e^{-y^2/2\varepsilon^2}/\sqrt{2\pi\varepsilon^2}$, and $g(\nu) = \lambda_0 + k\nu^2$, with $\eta = 0.5$, $\tau = 1$, $k = 1$, and $\lambda_0 = 0.5$. The discrete time step is $\Delta t = 10^{-4}$.

process,

$$\hat{y}_i := \varepsilon \hat{Y}_i, \tag{38}$$

or equivalently

$$\hat{\lambda}(t) = g\left(\varepsilon \sum_{i=1}^{\hat{N}(t)} \hat{Y}_i h(t - \hat{t}_i) \right). \tag{39}$$

In other words, each jump size $\hat{y}_i$ is assumed proportional to a small parameter ε and thus the rescaled jump size $\hat{Y}_i$ appears as the renormalized jump size independent of ε [see Fig. 3(a)]. For explicit clarification of the ε dependence, we denote below the original mark distribution $\rho(y)$ by $\rho_\varepsilon(y)$. This assumption can be interpreted as a weak-coupling limit between the system and the noise term. Considering the Jacobian relation, i.e., preservation of probability,

$$\rho_\varepsilon(y)dy = \tilde{\rho}(Y)dY \tag{40}$$

with the scaled jump-size distribution $\tilde{\rho}(Y)$, the above scaling assumption on the trajectory level is equivalent to that for the mark distribution

$$\rho_\varepsilon(y) := \frac{1}{\varepsilon} \tilde{\rho}\left(\frac{y}{\varepsilon} \right). \tag{41}$$

We note that this scaling assumption is equivalent to the system size expansion (often called the Ω expansion), which was originally introduced by Van Kampen for a systematic derivation of the Langevin equation within this kinetic theory (see the textbook by Van Kampen [43] and a review [45] including recent extended SSEs [46,47]). With this assumption, the KM coefficients α_k have the scaling

$$\alpha_k = \begin{cases} \varepsilon^k \tilde{\alpha}_k & \text{for even } k \\ 0 & \text{for odd } k \end{cases} \tag{42}$$

with the ε-independent KM coefficient $\tilde{\alpha}_k := \int_{-\infty}^{\infty} Y^k \tilde{\rho}(Y)dY$.

In the weak-coupling limit $\varepsilon \to 0$, each jump size is very small and thus the noise term becomes irrelevant if the intensity is constant. To keep the effect of the noise minimally relevant, let us take the diffusive limit by increasing the intensity as a function of ε, i.e., $g(\nu)$ is a function of ε. As the

third assumption (iii), therefore, we assume that the intensity function satisfies the diffusive scaling

$$g(\nu) = \frac{1}{\varepsilon^2} \tilde{g}(\nu) \tag{43}$$

with the ε-independent intensity function $\tilde{g}(\nu)$ [see Fig. 3(b)]. In other words, the model is explicitly written in the following form:

$$\hat{\lambda}(t) = \frac{1}{\varepsilon^2} \tilde{g}\left(\varepsilon \sum_{i=1}^{\hat{N}(t)} \hat{Y}_i h(t - \hat{t}_i) \right). \tag{44}$$

These three assumptions enable us to rewrite the field ME exactly in terms of the functional FPE in the diffusive limit $\varepsilon \to 0$ [see Fig. 3(b)] as we will elaborate in the following.

It is interesting to mention a report by Gao and Zhu [29], where a similar but still different form of asymptotics is studied by assuming a one-sided mark distribution $\rho(y) = \delta(y - 1)$ and a scaling for the tension-intensity map $g(\nu) = (1/\varepsilon)\tilde{g}(\nu)$ for a nonlinear version of the high-baseline-intensity regime for the LH processes [48]. For this setup, the trajectory fluctuates around a deterministic trajectory and thus shows quite different phenomenology.

1. Exponential-memory case

To understand the main ingredients of our calculations, let us first focus on the exponential-memory case (29). The KM expansion can be rewritten as

$$\frac{\partial P_t(\nu)}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial \nu} [\nu P_t(\nu)] + \sum_{k=1}^{\infty} \varepsilon^{2k-2} \frac{\tilde{\alpha}_{2k}}{(2k)!}$$
$$\times \left(-\frac{\eta}{\tau} \right)^{2k} \frac{\partial^{2k}}{\partial \nu^{2k}} \tilde{g}(\nu) P_t(\nu). \tag{45}$$

By taking the diffusive limit $\varepsilon \to 0$ [Fig. 3(b)], we obtain the exact FPE

$$\frac{\partial P_t(\nu)}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial \nu} [\nu P_t(\nu)] + D \frac{\partial^2}{\partial \nu^2} \tilde{g}(\nu) P_t(\nu), \quad D := \frac{\tilde{\alpha}_2 \eta^2}{2\tau^2}. \tag{46}$$

We note that this FPE is equivalent to an Itô process described by

$$\frac{d\hat{v}}{dt} = -\frac{\hat{v}}{\tau} + \sqrt{2D\tilde{g}(\hat{v})} \cdot \hat{\xi}^{\text{G}}, \tag{47}$$

with the Itô product denoted by a centered dot and the standard white Gaussian noise $\hat{\xi}^{\text{G}}$, satisfying $\langle \hat{\xi}^{\text{G}} \rangle = 0$ and $\langle \hat{\xi}^{\text{G}}(t)\hat{\xi}^{\text{G}}(t') \rangle = \delta(t - t')$.

2. General cases

The above formulation can be extended for the field ME. Indeed, we obtain the exact functional FPE

$$\frac{\partial P_t[z]}{\partial t} = \int_0^\infty dx \frac{\delta}{\delta z(x)}\left(\frac{z(x)}{x}P_t[z]\right)$$
$$+ \int_0^\infty dx \int_0^\infty dx' D(x, x')$$
$$\times \frac{\delta^2}{\delta z(x)\delta z(x')}\tilde{G}[z]P_t[z], \tag{48}$$

with the coefficient

$$D(x, x') := \frac{\alpha_2}{2}\tilde{h}(x)\tilde{h}(x'). \tag{49}$$

The functional FPE (48) implies that the stochastic dynamics finally reduces to

$$\frac{\partial \hat{z}(t, x)}{\partial t} = -\frac{\hat{z}(t, x)}{x} + \sqrt{2\tilde{G}[\hat{z}]} \cdot \hat{\xi}^{\text{G}}(t; x) \tag{50}$$

for the diffusive limit [Fig. 3(b)], with the white Gaussian noise satisfying

$$\langle \hat{\xi}^{\text{G}}(t; x) \rangle = 0, \quad \langle \hat{\xi}^{\text{G}}(t; x)\hat{\xi}^{\text{G}}(t'; x') \rangle = D(x, x')\delta(t - t'). \tag{51}$$

F. Laplace transformation

Here we introduce the relevant notation for the Laplace transformation. We first define the K-dimensional Laplace transformation as

$$\mathcal{L}_K[f(\mathbf{z}); \mathbf{s}] := \int_0^\infty d\mathbf{z}\, e^{-\mathbf{s}\cdot\mathbf{z}} f(\mathbf{z}), \quad \mathbf{s} \in \mathcal{R}_K^+. \tag{52}$$

In a parallel manner, the Laplace transformation in the function space can be defined as a straightforward generalization as follows:

$$\mathcal{L}_{\text{path}}[f[z]; s] := \int_0^\infty \mathcal{D}z \exp\left(-\int_0^\infty dx\, s(x)z(x)\right)f[z],$$
$$s \in \mathcal{S}_F. \tag{53}$$

We note that this Laplace transformation is a kind of path integral.

G. How to apply the field master equation as the fundamental equation for analytical calculations

We have formulated the field master equation (17) as the basis for our analytical calculations. In conventional Markovian stochastic processes, the ME is one of the most fundamental equations for analytical calculations. Indeed, while stochastic differential equations are generally nonlinear equations, the corresponding master equations are always

linear; thus, analytical calculations reduce to eigenvalue problems in linear algebra. If the solutions to the eigenvalue problems are obtained, one can access various statistical quantities, such as ensemble averages, correlation functions, and probability density functions [42].

Thus, for our field ME (17) approach, such a linear-algebra method is formally applicable. In other words, a general analysis of the field ME (17) reduces to the eigenvalue problem

$$\mathcal{L}\phi_\lambda[z] = -\lambda\phi_\lambda[z], \tag{54}$$

with the total linear operator $\mathcal{L} := \mathcal{L}_{\text{adv}} + \mathcal{L}_{\text{jump}}$, the eigenvalue λ, and the corresponding eigenfunctional $\phi_\lambda[z]$. The time evolution of the PDF $P_t[z]$ is given by

$$P_t[z] = \sum_\lambda c_{\text{init}}(\lambda)e^{-\lambda t}\phi_\lambda[z], \tag{55}$$

with coefficients $c_{\text{init}}(\lambda)$ fixed by the initial condition $P_0[z] = \sum_\lambda c_{\text{init}}(\lambda)\phi_\lambda[z]$. Typically, due to the Perron-Frobenius theorem, the real part of the eigenvalues is non-negative $\text{Re}(\lambda) \geqslant 0$ and the zero eigenvalue exists $\lambda = 0$ as the most important eigenvalue, characterizing the steady state.

If we can solve the eigenvalue problem (54), we can straightforwardly calculate many useful quantities. For example, the ensemble average of an arbitrary observable $A[\hat{z}_t]$ as a functional of $\hat{z}_t$ is formally given by

$$\langle A[\hat{z}_t] \rangle := \int A[z]P_t[z]\mathcal{D}z = \sum_\lambda c_{\text{init}}(\lambda)e^{-\lambda t}\int A[z]\phi_\lambda[z]\mathcal{D}z. \tag{56}$$

In addition, the steady-state intensity PDF $P_{\text{ss}}(\lambda) := \lim_{t\to\infty}\langle \delta(\lambda - G[\hat{z}_t]) \rangle$ is written as

$$P_{\text{ss}}(\lambda) \propto \int \delta(\lambda - G[z])\phi_0[z]\mathcal{D}z. \tag{57}$$

Thus, only the zeroth eigenfunctional $\phi_0[z]$ is necessary for deriving the steady-state intensity PDF $P_{\text{ss}}(\lambda)$. In this sense, the study of the non-Markovian dynamics (3) is exactly mapped onto the linear-algebra problem (54) in principle. The field ME (17) provides a clear view regarding analytical solutions to the NLH processes. See Fig. 4(a) for the flow chart for our analytical calculations. In this paper we focus on the steady-state intensity PDF $P_{\text{ss}}(\lambda)$, requiring only the zeroth eigenfunctional based on the formula (57). We analytically obtain the zeroth eigenfunctional $\phi_0[z]$ corresponding to various setups. Since the NLH is uniquely identified by the memory kernel $h(t)$, mark distribution $\rho(y)$, and tension-intensity map $g(\nu)$, we enumerate the corresponding solutions by systematically specifying the setup parameters $\mathcal{P} := (h, \rho, g)$ in the following sections to be read like a dictionary. The flow chart for our setup-parameter classification is presented in Fig. 4(b).

IV. SOLUTION 1: EXPONENTIAL-MEMORY KERNEL WITH ONE-SIDED MARK DISTRIBUTION AND LINEAR AND RAMP INTENSITY MAPS

In this section we focus on exact solutions for the simplest case with the exponential-memory kernel (29), whose dynamics is characterized by a simple ME (31). In particular, we
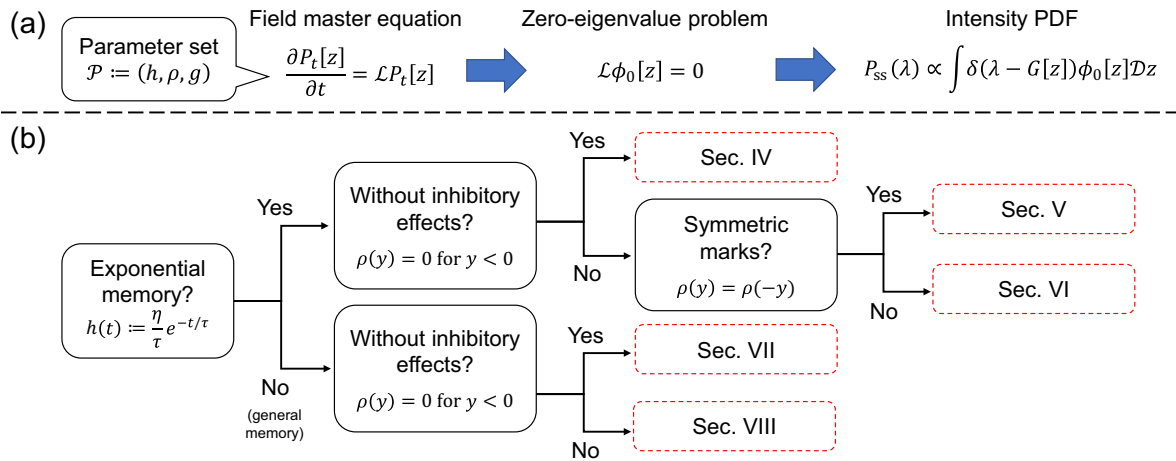
FIG. 4. Logical structure of this work. (a) Flow chart of our analytical calculations based on the field master equation (17). The field ME (17) is uniquely identified by the parameters $\mathcal{P} := (h, \rho, g)$. By solving the zero-eigenvalue problem (54), we obtain the steady-state intensity PDF $P_{ss}(\lambda)$ via formula (57). (b) Flow chart for our systematic classification of the possible intensity PDFs over the parameter set $\mathcal{P}$. We first categorize the memory kernel $h(t)$ according to whether it is a pure exponential relaxation or instead represents a general memory process and then classify the mark distribution $\rho(y)$ in terms of the inhibitory effects and symmetry.

assume here that all the marks are positive $\hat{y} > 0$, implying the absence of inhibitory effects.

A. Exact solutions for one-sided exponential jump

Let us consider the case with the exponential-memory kernel (29) and with the one-sided exponential jump size[3] (see Fig. 5)

$$\rho(y) := \frac{1}{y^*} e^{-y/y^*} \Theta(y), \qquad (58)$$

whose ME is known to be exactly tractable due to its special form [49]. We assume $y^* = 1$ without losing generality because the scale can be absorbed into the branching ratio η. Since both the memory kernel and jump size are non-negative, the inhibitory effects are absent in this model. Interestingly, even this simple model can exhibit nontrivial steady-state

_____

[3]Here $\Theta(y)$ is the Heaviside function defined by $\Theta(y) = 1$ for $y > 0$, $\Theta(0) = 1/2$, and $\Theta(y) = 0$ for $y < 0$.

FIG. 5. Schematic of the one-sided exponential mark distribution (58) with $y^* = 1$.

distribution functions of intensities resulting from the nonlinearity of the tension-intensity map $g(\hat{v})$. This case is special because the exact steady-state solution to the ME (31) is available. In the steady state, the exact steady-state solution is given by

$$P_{ss}(v) = \frac{v^{-1}}{Z} \exp\left(-cv + \tau \int \frac{g(v)}{v} dv\right), \qquad (59)$$

with

$$c := \frac{\tau}{\eta} \qquad (60)$$

and with a normalization constant given by

$$Z := \int_0^\infty dv \, v^{-1} \exp\left(-cv + \tau \int \frac{g(v)}{v} dv\right). \qquad (61)$$

Derivation

By utilizing the identity (see Appendix C for the technical derivation)

$$\left(1 + \frac{1}{c}\frac{\partial}{\partial v}\right) \int_0^\infty dy \, e^{-y} g\left(v - \frac{\eta y}{\tau}\right) P_t\left(v - \frac{\eta y}{\tau}\right) = g(v) P_t(v), \qquad (62)$$

we can rewrite the ME as

$$\frac{\partial P_t(v)}{\partial t} = \frac{1}{\tau}\partial_v[v P_t(v)] - \frac{\partial_v/c}{1 + \partial_v/c} g(v) P_t(v), \qquad (63)$$

with the differential operator $\partial_v := \partial/\partial v$. We note that a similar calculation technique can be found in Ref. [49]. This ME can be rewritten as

$$\frac{\partial P_t(v)}{\partial t} = -\frac{\partial}{\partial v} J_t(v),$$

$$J_t(v) := -\frac{1}{\tau} v P_t(v) + \frac{1/c}{1 + \partial_v/c} g(v) P_t(v). \qquad (64)$$

Here we assume the natural boundary condition [42]

$$\lim_{v\to\infty} J_t(v) = 0, \qquad (65)$$
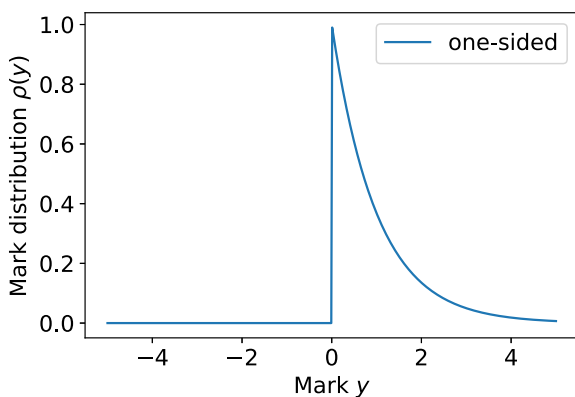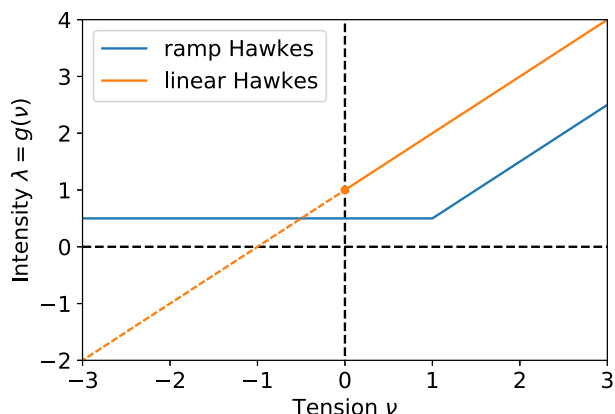
FIG. 6. Schematic of the tension-intensity maps for the linear, i.e., $g(\nu) = \nu_0 + \nu$ with $\nu_0 = 1$, and ramp, i.e., $g(\nu) = \max\{\nu_0, \nu - \nu_1\}$ with $\nu_0 = 1/2$ and $\nu_1 = 1/2$, Hawkes processes. While ν_0 must be non-negative due to the non-negativity of the probability, ν_1 can be either positive or nonpositive.

ensuring that the mean probability velocity $J_t(\nu)/P_t(\nu)$ is zero at infinity. We reject the possibility of periodic boundary conditions which are nonphysical. In the steady state, we thus obtain the exact steady-state solution (59).

B. Example 1: Linear Hawkes process

For the linear intensity function

$$g(\nu) = \nu + \nu_0 \tag{66}$$

with base intensity $\nu_0 > 0$, the model recovers the conventional LH process. In the subcritical case $\eta < 1$, the exact steady-state solution is given by the Γ distribution as analyzed in Ref. [50],

$$P_{\text{ss}}(\nu) = \frac{1}{Z}\nu^{-1-a}e^{-\nu/\nu_{\text{cut}}}, \quad Z = \nu_{\text{cut}}^{\tau\nu_0}\Gamma(\tau\nu_0), \quad a := -\tau\nu_0, \tag{67}$$

with the Gamma function $\Gamma(x) := \int_0^\infty dt\, t^{x-1}e^{-t}$. The characteristic tension for the exponential cutoff is defined as

$$\nu_{\text{cut}} := \frac{\eta}{1-\eta}\frac{1}{\tau}. \tag{68}$$

The PDF $P_{\text{ss}}(\nu)$ and thus the PDF of λ is a power law with a nonuniversal negative exponent a up to the cutoff tension ν_{cut}. Since the cutoff tension diverges near criticality, the power-law tail described by $\nu^{-1+\tau\nu_0}$ and $\lambda^{-1+\tau\nu_0}$ corresponds to an intermediate asymptotics [18], as reported in Ref. [16].

C. Example 2: Ramp tension-intensity map

Let us consider the ramp tension-intensity map (also called a rectified linear unit in the context of recent works in machine learning)

$$g(\nu) = \max\{\nu_0, \nu - \nu_1\} \tag{69}$$

for positive ν_0 and any real number ν_1 (see Fig. 6). In this paper the NLH process with the ramp tension-intensity map (69) is called the ramp Hawkes process. While the ramp Hawkes process is quite similar to the LH process, its minimal

nonlinearity leads to a genuine asymptotic power-law tail, thus very different from the LH process. In the subcritical regime $\eta < 1$, the exact steady-state solution is given by

$$P_{\text{ss}}(\nu) = \begin{cases} \frac{1}{Z}\nu^{-1-\tau\nu_1}e^{-\nu/\nu_{\text{cut}}}, & \nu > \nu_0 + \nu_1 \\ \frac{(\nu_0+\nu_1)^{-\tau(\nu_0+\nu_1)}}{Z}\nu^{-1+\tau\nu_0}e^{\tau(\nu_0+\nu_1)-c\nu}, & \nu \leqslant \nu_0 + \nu_1, \end{cases} \tag{70}$$

with the exponential cutoff ν_{cut} given by the expression (68), the constant c given by (60), and the normalization constant Z given by (61).

Interestingly, for $\nu_1 > 0$ and at criticality $\eta = 1$, for $\nu > \nu_0 + \nu_1$, $P_{\text{ss}}(\nu)$ becomes a pure power law

$$P_{\text{ss}}(\nu) \propto \nu^{-1-a}, \quad a := \tau\nu_1, \tag{71}$$

which is normalizable without truncation. Given the asymptotic linear relationship between ν and λ, the same power-law behavior holds for the PDF of λ. This power law is different from the intermediate asymptotic power-law distribution (67) for the LH process. In this sense, the ramp Hawkes process can reproduce any power-law relationship (including both true and intermediate asymptotics) at criticality, which may be useful to account for power-law distributions observed empirically in various systems. It is remarkable that such a slight change from the affine structure (66) to the rectified linear (69) structure creates this large difference in the asymptotic intensity distribution. Note also that, since $\max\{\nu_0, \nu - \nu_1\} < \nu + \nu_0 \,\forall\, \nu_1 > 0$, the ramp tension-intensity map has a smaller intensity than that of the LH process, which explains the thinner tail (71) compared with (67) (the latter becoming so heavy tailed close to criticality so as to become non-normalizable). Intuitively, the base tension ν_0 in the ramp tension-intensity map (69) acts as a replenishing engine that ensures a minimum activity, which can become the source of bursts. This structure of the ramp tension-intensity map is somewhat reminiscent of the Kesten process [51–53], which is well known to produce power-law distributions with the tail exponent depending on the distribution of the multiplicative factors. It is interesting that the exponent $a = \tau\nu_1$ is independent of the resourcing term.

D. Existence of steady-state solutions

The exact solution (59) is useful in understanding the condition for the existence of a steady-state solution. For example, let us consider the case of the exponential tension-intensity map

$$g(\nu) = \lambda_0 e^{\beta\nu}, \quad \beta > 0, \tag{72}$$

which has been used in the statistical calibration of neural spike time series in neural science [54]. The exact solution (59) predicts that this NLH process has no steady-state solution. Indeed,

$$P_{\text{ss}}(\nu) \propto \nu^{-1}\exp\left(-c\nu + \lambda_0\tau\int\nu^{-1}e^{\beta\nu}d\nu\right)$$

$$\propto \nu^{-1}\exp[-c\nu + \lambda_0\tau\,\text{Ei}(\beta\nu)]$$

$$\simeq \nu^{-1}\exp\left(-c\nu + \lambda_0\tau\frac{e^{\beta\nu}}{\beta\nu}\right) \tag{73}$$

for large ν with the exponential integral $\text{Ei}(x) := -\int_{-x}^{\infty} t^{-1} e^{-t} dt$. This PDF is not normalizable, implying that this NLH process is always unstable independently of the model parameters.

To avoid this problem, one of the easiest solutions is to introduce an upper bound in the intensity function

$$g(\nu) = \min\{\lambda_0 e^{\beta \nu}, \lambda_{\max}\}, \tag{74}$$

with the finite upper boundary parameter $\lambda_{\max} > 0$. In Ref. [54] Gerhard *et al.* introduced a similar regularization to guarantee the stability of their model. However, it is remarkable that this NLH process is always unstable in the absence of the upper bound and thus simulation results sensitively depend on the specific value of the cutoff $\lambda_{\max}$.

In general, if the tension-intensity map diverges faster than the linear (or ramp) function, there is no stationary solution. Indeed, for $g(\nu) \simeq \lambda_0 \nu^n$ with $n > 1$, we obtain

$$P_{\text{ss}}(\nu) \simeq \frac{\nu^{-1}}{Z} \exp\left(-c\nu + \lambda_0 \tau \int \nu^{n-1} d\nu\right)$$

$$\propto \nu^{-1} \exp\left(-c\nu + \frac{\lambda_0 \tau}{n} \nu^n\right), \tag{75}$$

the steady-state intensity distribution $P_{\text{ss}}(\nu)$ is given by the nonuniversal power-law relation

$$P_{\text{ss}}(\nu) \propto \nu^{-1-a}, \quad a := \frac{2\tau\nu_1}{\alpha_2}. \tag{76}$$

Given the asymptotic linear relationship between ν and λ, the same power-law behavior holds for the PDF of λ.

We stress that ν_1 can take any real value, either positive, negative, or zero. If negative or zero, the derivation does not extend all the way to the limit $\eta = 1$ and the power law (76) is truncated as in (67) by an exponential cutoff. This result implies a true power-law tail for positive ν_1, i.e., normalizable without a cutoff tail even at criticality, or intermediate asymptotic power-law tail for nonpositive ν_1, i.e., not normalizable without a cutoff tail near criticality. Notably, this recovers Eq. (71) for the one-sided exponential mark distribution (58) for which $\alpha_2 = 2$.

Derivation

Since we are interested only in the tail of the intensity PDF, let us focus on the asymptotic properties of the ME (31) for large ν. The ME (31) has the asymptotic expression

$$\frac{1}{\tau} \frac{\partial}{\partial \nu}[\nu P_{\text{ss}}(\nu)] + \int dy\, \rho(y)\left(\nu - \nu_1 - \frac{\eta y}{\tau}\right) P_{\text{ss}}\left(\nu - \frac{\eta y}{\tau}\right)$$

$$- (\nu - \nu_1) P_{\text{ss}}(\nu) \simeq 0 \quad \text{for large } \nu \tag{77}$$

in its steady state, obtained by replacing $g(\nu)$ by $\nu - \nu_1$ asymptotically. Applying the Laplace transform

$$\tilde{P}_{\text{ss}}(s) := \mathcal{L}_1[P_{\text{ss}}(\nu); s] = \int_0^{\infty} d\nu\, e^{-s\nu} P_{\text{ss}}(\nu) \tag{78}$$

which is not normalizable. In this sense, the ramp Hawkes process is the boundary between the stationary and nonstationary Hawkes processes under the assumption of an exponential memory (29) and one-sided exponential marks (58).

Thus, an NLH process with one-sided positive marks is not so flexible, if we require its stationarity. However, this situation drastically changes if we allow for the coexistence of excitatory and inhibitory effects, i.e., marks can take both positive and negative values. Indeed, as will be shown in Sec. V, NLH processes with two-sided marks are flexible enough to accommodate various nonlinearities without losing their stationarity.

E. Robust asymptotic results

The previous presentation of exact solutions for the ramp tension-intensity map (69) for the special case of (a) an exponential memory and (b) an exponential jump-size distribution allowed us to highlight the appearance of power-law tails for the distribution of tensions near and at criticality. Here we show that such a power-law behavior is asymptotically robust for general jump-size distributions, assuming that the memory is exponential. With the notation

$$\alpha_k := \int_0^{\infty} y^k \rho(y) dy < \infty \quad \text{for any } k \geqslant 1, \quad \alpha_1 = 1,$$

to Eq. (77) yields

$$-\frac{s}{\tau} \frac{d}{ds} \tilde{P}_{\text{ss}}(s) - \Phi(s)\left(\frac{d}{ds}\tilde{P}_{\text{ss}}(s) + \nu_1 \tilde{P}_{\text{ss}}(s)\right) \simeq 0,$$

$$\Phi(s) := \int_0^{\infty} dy (e^{-s\eta y/\tau} - 1) \rho(y). \tag{79}$$

Its solution is given by

$$\ln \tilde{P}_{\text{ss}}(s) \simeq -\nu_1 s + \int_0^s \frac{\nu_1 s' ds'}{s' + \tau \Phi(s')}, \tag{80}$$

with the normalization condition $\tilde{P}_{\text{ss}}(s = 0) = 1$. Considering the expansion

$$\Phi(s) = -\frac{\eta}{\tau} s + \frac{\eta^2 \alpha_2}{2\tau^2} s^2 + \cdots, \tag{81}$$

$\ln \tilde{P}_{\text{ss}}(s)$ has the asymptotic form for small s near criticality $1 - \eta \ll 1$,

$$\ln \tilde{P}_{\text{ss}}(s) \simeq \frac{2\tau\nu_1}{\alpha_2} \ln s \quad \text{for small } s, \tag{82}$$

implying, by inverse Laplace transform, a power-law asymptotics for the steady intensity PDF:

$$P_{\text{ss}}(\nu) \propto \nu^{-1-2\tau\nu_1/\alpha_2}. \tag{83}$$

For nonpositive ν_1, this PDF is not normalizable and thus requires a cutoff tail, such as the exponential given by Eq. (67).

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad g(\nu) \simeq \nu - \nu_1 + o(\nu^0) \quad \text{for large } \nu,$$
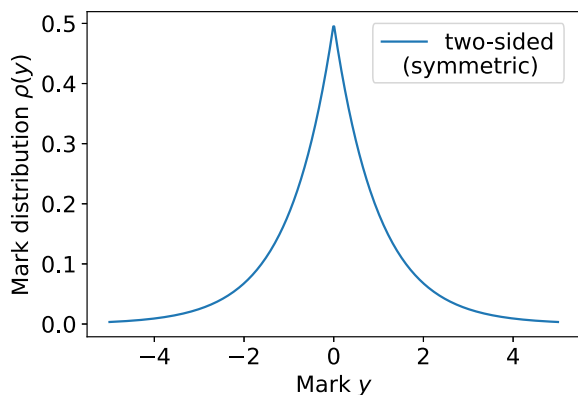
FIG. 7. Schematic of the two-sided symmetric exponential mark distribution (84) with $y^* = 1$.

V. SOLUTION 2: EXPONENTIAL-MEMORY KERNEL WITH TWO-SIDED SYMMETRIC MARK DISTRIBUTION FOR LINEAR TO FAST-ACCELERATING INTENSITY MAPS

In this section we study both exact and asymptotic solutions of the ME (31) for the PDF of the total tension ν valid for an exponential-memory kernel (29) and in the presence of inhibitory effects (i.e., marks $\hat{y}$ can be both positive and negative). The inhibitory effects imply that events can sometimes suppress or decrease the amplitude of bursts, which can lead to phenomena essentially different from those in the preceding section.

A. Exact solutions to two-sided symmetric exponential mark distribution

Let us focus on the case with the two-sided symmetric exponential mark distribution (see Fig. 7)

$$\rho(y) = \frac{1}{2y^*}e^{-|y|/y^*}, \tag{84}$$

which corresponds to the existence of symmetric positive ($y > 0$) and negative ($y < 0$) feedback effects with zero mean. We again assume $y^* = 1$, without loss of generality. This negative feedback effect is called the inhibitory effect in Ref. [20] and is known to be difficult to deal with in analytical approaches. We present the exact solution of Eq. (31) with (84) for some specific forms of $\lambda = g(\nu)$.

Let us recall the identity

$$\left(1 - \frac{1}{c}\frac{\partial}{\partial \nu}\right)\int_{-\infty}^{0}dy\,e^{-|y|}g\left(\nu - \frac{\eta y}{\tau}\right)P_t\left(\nu - \frac{\eta y}{\tau}\right)$$
$$= g(\nu)P_t(\nu), \tag{85}$$

where $c = \tau/\eta$ has been defined in (60) (see Appendix C for the derivation). This identity together with the other identity (62) implies a third identity useful to solve the ME (31):

$$\int_{-\infty}^{+\infty}dy\,\frac{e^{-|y|}}{2}g\left(\nu - \frac{\eta y}{\tau}\right)P_t\left(\nu - \frac{\eta y}{\tau}\right)$$
$$= \frac{1}{2}\left(\frac{1}{1 + \partial_\nu/c} + \frac{1}{1 - \partial_\nu/c}\right)g(\nu)P_t(\nu)$$
$$= \frac{1}{1 - \partial_\nu^2/c^2}g(\nu)P_t(\nu). \tag{86}$$

We thus obtain a simple representation of the ME (31):

$$\frac{\partial P_t(\nu)}{\partial t} = \frac{1}{\tau}\partial_\nu[\nu P_t(\nu)] + \frac{\partial_\nu^2/c^2}{1 - \partial_\nu^2/c^2}g(\nu)P_t(\nu). \tag{87}$$

This ME can be written in the more familiar form

$$\frac{\partial P_t(\nu)}{\partial t} = -\frac{\partial J_t(\nu)}{\partial \nu}, \tag{88}$$

where the probability current is defined by

$$J_t(\nu) := -\left(\frac{\nu}{\tau} + \frac{\partial_\nu/c^2}{1 - \partial_\nu^2/c^2}g(\nu)\right)P_t(\nu). \tag{89}$$

This formulation makes more transparent the meaning of the boundary condition $\lim_{\nu \to \infty}J_t(\nu) = 0$ ensuring that the mean probability velocity $J_t(\nu)/P_t(\nu)$ is zero at infinity. We reject the possibility of periodic boundary conditions which are nonphysical.

Then the steady-state solution satisfies the second-order differential equation

$$\frac{d^2}{d\nu^2}[\nu P_{ss}(\nu)] - \tau\frac{d}{d\nu}[g(\nu)P_{ss}(\nu)] - c^2\nu P_{ss}(\nu) = 0. \tag{90}$$

This is obtained by setting $\frac{\partial P_t(\nu)}{\partial t} = 0$ in (88) and using $\lim_{\nu \to \infty}J_t(\nu) = 0$, which leads to $J_t(\nu) = 0 \,\forall\, \nu$, from which Eq. (90) is derived.

1. Example 1: Ramp tension-intensity map

For the ramp tension-intensity map

$$\lambda = g(\nu) = \max\{\nu_0, |\nu|\}, \tag{91}$$

which corresponds to setting $\nu_1 = 0$ in Eq. (69) and adding the absolute value, the solution of (90) is a truncated-Lévy-type intensity asymptotic tail

$$P_{ss}(\lambda) \propto \lambda^{-1}e^{-\lambda/\lambda_{cut}}, \quad \frac{1}{\lambda_{cut}} := \frac{2c^2}{\tau + \sqrt{4c^2 + \tau^2}} \quad \text{for large } \lambda, \tag{92}$$

where $c = \tau/\eta$ has been defined in (60). For $\nu_1 \neq 0$ in Eq. (69), the exact form of the intensity distribution is also available.

Remarkably, this model has no critical point: The process is always stationary for all $\eta < \infty$, due to the stabilization effect of the inhibitory component of the process, and thus λ_{cut} is always finite. This is in contrast to the ramp Hawkes process with one-sided exponential jumps (without inhibitory effect) and thus highlights the fact that the inhibitory effects can be crucial in understanding even the qualitative behavior of the NLH processes in general.

Derivation of the exact solution. The ME (90) reduces to the following set of modified Bessel differential and constant-coefficient second-order differential equations:

$$\nu^2\frac{d^2\phi(\nu)}{d\nu^2} + \nu\frac{d\phi(\nu)}{d\nu} - (c^2\nu^2 + \gamma^2)\phi(\nu) = 0,$$
$$\gamma := \frac{1 - \tau\nu_0}{2}, \quad \phi(\nu) := \nu^\gamma P_{ss}(\nu) \quad (0 \leqslant \nu \leqslant \nu_0), \tag{93}$$

$$\frac{d^2}{d\nu^2}\psi(\nu) - \tau\frac{d}{d\nu}\psi(\nu) - c^2\psi(\nu) = 0,$$

$$\psi(\nu) := \nu P_{ss}(\nu) \quad (\nu_0 \leqslant \nu). \tag{94}$$

The exact solution is then given by

$$P_{ss}(\nu) = \begin{cases} \nu^{-\gamma}[C^{[1]}I_\gamma(c|\nu|) + C^{[2]}K_\gamma(c|\nu|)], & |\nu| \leqslant \nu_0 \\ C^{[3]}|\nu|^{-1}e^{-|\nu|/\nu_{cut}}, & |\nu| > \nu_0, \end{cases}$$

$$\frac{1}{\nu_{cut}} := \frac{2c^2}{\tau + \sqrt{4c^2 + \tau^2}}, \quad c := \frac{\tau}{\eta}, \tag{95}$$

with integral constants $C^{[1]}$, $C^{[2]}$, and $C^{[3]}$ and modified Bessel functions of the first and second kinds [denoted by $I_\gamma(x)$ and $K_\gamma(x)$, respectively; see Appendix D 1]. The integral constants are determined by the normalization and continuity conditions $\int_{-\infty}^{\infty} d\nu\, P_{ss}(\nu) = 1$ and $\lim_{\nu\uparrow\nu_0} P_{ss}(\nu) = \lim_{\nu\downarrow\nu_0} P_{ss}(\nu)$. We thus obtain that the intensity distribution is given by the sum of a δ function centered on ν_0 and the truncated Lévy distribution

$$P_{ss}(\lambda) = [1 - C^{[3]}\Gamma(0, \nu_0/\nu_{cut})]\delta(\nu - \nu_0)$$

$$+ C^{[3]}\lambda^{-1}e^{-\lambda/\nu_{cut}}\Theta(\nu - \nu_0). \tag{96}$$

where the incomplete Gamma function is $\Gamma(a, x) := \int_x^\infty dt\, t^{a-1}e^{-t}$.

2. Example 2: Quadratic tension-intensity map

For the quadratic tension-intensity map corresponding to the Zumbach Hawkes process (see Sec. II C 2)

$$\lambda = g(\nu) = k\nu^2 + \lambda_0, \tag{97}$$

the solution of (90) is a power-law steady-state distribution[4]

$$P_{ss}(\lambda) \propto |\lambda|^{-1-a}, \quad a := \frac{1}{2} + \frac{c^2}{2k\tau}, \quad c := \frac{\tau}{\eta}, \tag{98}$$

with $\lambda_0 > 0$ and the power-law exponent $a > 1/2$. The exact form of the intensity distribution is also available. We note that this nonuniversal power-law scaling is consistent with Eq. (12), which was reported for the diffusive limit of the Zumbach Hawkes process in Ref. [28].

Derivation of the exact solution. By the variable transformation $x = k\tau\nu^2/2$, the ME (90) for $\nu > 0$ reduces

to

$$x\frac{d^2\phi(x)}{dx^2} + \left(\frac{3}{2} - \frac{\tau\lambda_0}{2} - x\right)\frac{d\phi(x)}{dx} - \left(1 + \frac{c^2}{2k\tau}\right)\phi(x) = 0,$$

$$\phi(x) := P(\nu(x)), \quad \nu(x) := \sqrt{\frac{2x}{k\tau}}. \tag{99}$$

This is the confluent hypergeometric differential equation and thus its exact solution is given by

$$\phi(x) = C^{[1]}{}_1F_1\left(1 + \frac{c^2}{2k\tau}, \frac{3}{2} - \frac{\tau\lambda_0}{2}; x\right)$$

$$+ C^{[2]}{}_1U_1\left(1 + \frac{c^2}{2k\tau}, \frac{3}{2} - \frac{\tau\lambda_0}{2}; x\right), \tag{100}$$

with integral constants $C^{[1]}$ and $C^{[2]}$ and the confluent hypergeometric functions of the first ${}_1F_1$ and the second kind ${}_1U_1$ (see Appendix D 2). The integral constants are determined by the normalization condition $\int_{-\infty}^{\infty} d\nu\, P_{ss}(\nu) = 1$. Interestingly, this solution has the following asymptotic form for large x [see Eq. (D4)]:

$$\phi(x) \propto x^{-1-c^2/2k\tau}. \tag{101}$$

The steady-state distribution of the intensity λ is then given by

$$P_{ss}(\lambda) = \left|\frac{d\nu}{d\lambda}\right|P_{ss}(\nu) \propto \lambda^{-1-a}, \quad a := \frac{1}{2} + \frac{c^2}{2k\tau}, \quad c := \frac{\tau}{\eta} \tag{102}$$

for the tail $\lambda \to \infty$. This is a power-law asymptotic distribution with a nonuniversal exponent a without truncation.

3. Example 3: Exponential intensity map

For the exponential tension-intensity map

$$\lambda = g(\nu) = \lambda_0|\nu|e^{\beta\nu}, \tag{103}$$

the solution of (90) is Zipf's law for the intensity distribution

$$P_{ss}(\lambda) \propto \lambda^{-2} \tag{104}$$

up to a logarithmic factor $\ln\lambda$, for large λ with positive constants λ_0 and β. This intensity map is inspired by the MSA model [21,22], where the dominant contribution comes from the exponential factor originating from the Arrhenius law (see Sec. V B 4 for more details).

Derivation of the exact solution. The exact steady-state solution of (90) is given by

$$P_{ss}(\nu) = \begin{cases} \frac{C^{[1]}}{\nu}e^{c\nu}{}_1U_1\left(1 + \frac{c}{\beta}, 1 + \frac{2c}{\beta}; \frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right) + \frac{C^{[2]}}{\nu}e^{c\nu}L_{-1-c/\beta}^{2c/\beta}\left(\frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right), & \nu \geqslant 0 \\ \frac{C^{[3]}}{\nu}\exp\left(c\nu - \frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right){}_1U_1\left(\frac{c}{\beta}, 1 + \frac{2c}{\beta}; \frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right) + \frac{C^{[4]}}{\nu}\exp\left(c\nu - \frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right)L_{-c/\beta}^{2c/\beta}\left(\frac{\lambda_0\tau}{\beta}e^{\beta\nu}\right), & \nu < 0, \end{cases} \tag{105}$$

with integral coefficients $C^{[1]}$, $C^{[2]}$, $C^{[3]}$, and $C^{[4]}$ and the generalized Laguerre function $L_a^b(x)$ (see Appendix D 3). Considering the asymptotic formulas (D4) and (D6), $C^{[2]}$ must

be zero since $P_{ss}(\nu) \to 0$ for $\nu \to \infty$. This means that the asymptotic tail is given by

$$P_{ss}(\nu) \propto \nu^{-1}e^{-\beta\nu}, \tag{106}$$

which leads to the Zipf law (104) for the steady intensity distribution, by using the Jacobian relation $d\nu P_{ss}(\nu) = d\lambda P_{ss}(\lambda) \iff P_{ss}(\lambda) = |d\nu/d\lambda|P_{ss}(\nu)$. As shown

---

[4]If λ_0 is zero, the steady-state distribution is singular at $\nu = 0$ as $P_{ss}(\lambda) \propto \nu^{-1}$ and thus is not normalizable.

in Secs. V B 4 and V C 3, this asymptotic Zipf law (104) is robust for exponential-type tension-intensity maps under general symmetric mark distribution (or more generally when the mark average is zero), on the condition that the memory kernel is exponential and the mark distribution has its moments at all orders being finite.

B. Exact solutions in the diffusive limit

Let us now consider the diffusive limit formulated in Sec. III E and assume an exponential-memory kernel (29):

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad \rho_\varepsilon(y) = \frac{1}{\varepsilon} \tilde{\rho}\left(\frac{y}{\varepsilon}\right), \quad g(\nu) = \frac{1}{\varepsilon^2} \tilde{g}(\nu). \tag{107}$$

For this case, by solving the FPE (46) in the steady state, we obtain the explicit solution

$$P_{\rm ss}(\nu) \underset{\varepsilon \to 0}{\propto} \frac{1}{\tilde{g}(\nu)} \exp\left(-\frac{1}{\tau D} \int \frac{\nu d\nu}{\tilde{g}(\nu)}\right), \quad D := \frac{\tilde{\alpha}_2 \eta^2}{2\tau^2} \tag{108}$$

for any $\tilde{g}(\nu)$, assuming that $\tilde{\rho}(y)$ and $\tilde{g}(\nu)$ are independent of ε and that all the integrals appropriately converge.

1. Example 1: Ramp tension-intensity map

Let us first consider the example of the ramp tension-intensity map

$$g(\nu) = \max\{\nu_0, |\nu - \nu_1|\} \tag{109}$$

with positive real number $\nu_0 > 0$ and arbitrary real number[5] ν_1. For $\nu > \nu_0 + \nu_1$, the solution of the FPE (46) is the truncated Lévy distribution for the tension (and thus for the intensity)

$$P_{\rm ss}(\nu) = \frac{1}{Z} \frac{e^{-\nu/\nu_{\rm cut}}}{(\nu - \nu_1)^{1+a}}, \quad \nu_{\rm cut} := D\tau, \quad a := \frac{\nu_1}{D\tau}, \tag{110}$$

with an exponential tail tapering the intermediate power-law tail. We note that this model has no critical point due to the inhibitory effects leading to the characteristic intensity for the exponential cutoff $\nu_{\rm cut}$ to be always finite.

2. Example 2: Quadratic tension-intensity map

We next study the exact solution of the quadratic intensity (97), corresponding to the Zumbach Hawkes process (see Sec. II C 2) in the presence of inhibitory effects. Using the formula (108), the exact solution of the FPE (46) in the steady-state regime is given by

$$P_{\rm ss}(\nu) = \frac{1}{Z}(k\nu^2 + \lambda_0)^{-1-1/2kD\tau}, \tag{111}$$

which is equivalent to

$$P_{\rm ss}(\lambda) = \frac{1}{Z'} \frac{\lambda^{-1-1/2kD\tau}}{\sqrt{\lambda - \lambda_0}} \propto \lambda^{-1-a} \quad (\lambda \to \infty),$$
$$a := \frac{1}{2} + \frac{1}{2kD\tau}, \tag{112}$$

---

[5] Here ν_1 can be either positive, zero, or negative.

with $Z' := 2\sqrt{k}Z$. This is a power-law distribution without truncation and with a nonuniversal exponent a. We note that this nonuniversal power-law scaling is essentially identical to Eq. (12) for the diffusive limit of the Zumbach Hawkes process reported in Ref. [28].

3. Example 3: Polynomial tension-intensity map

Let us consider the case of the polynomial intensity given by

$$\tilde{g}(\nu) = k|\nu|^n + \nu_0, \quad n > 2 \tag{113}$$

with positive constant $\nu_0 > 0$. Using the formula (108), we obtain the exact steady-state distribution, the solution of the FPE (46),

$$P_{\rm ss}(\nu)$$
$$\propto \frac{1}{k|\nu|^n + \nu_0} \exp\left[-\frac{\nu^2}{2\tau D\nu_0} {}_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -\frac{k|\nu|^n}{\nu_0}\right)\right] \tag{114}$$

with the hypergeometric function ${}_2F_1$ (see Appendix D 4). By considering the asymptotic expansion

$$\nu^2 {}_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -\frac{k\nu^n}{\nu_0}\right)$$
$$= \begin{cases} \underbrace{\frac{\frac{2\pi}{n}}{\sin\frac{2\pi}{n}}\left(\frac{\nu_0}{k}\right)^{2/n}}_{\rm const} + o(\nu^0), & n > 2 \\ \frac{\nu_0}{k}\ln\left(\frac{k\nu^2}{\nu_0}\right) + o(\nu^0), & n = 2 \end{cases} \tag{115}$$

for large ν, we obtain the asymptotic form of the steady PDF for $|\nu| \to \infty \implies \lambda \to \infty$ as

$$P_{\rm ss}(\nu) \propto \frac{1}{k|\nu|^n + \nu_0} \implies P_{\rm ss}(\lambda) \propto \lambda^{-1-a}, \quad a := 1 - \frac{1}{n}. \tag{116}$$

Note that the limit $n \to +\infty$ recovers Zipf's law. The manner with which the exact solution (114) recovers Eq. (112) for $n \to 2$ is now elaborated.

Crossover between $n = 2$ and $n > 2$. Remarkably, the solution (112) for the quadratic Hawkes (i.e., $n = 2$) and the one (114) for the polynomial Hawkes with $n > 2$ are slightly different. This qualitative difference can be seen from the analytical singularity of the hypergeometric function ${}_2F_1$ at $n = 2$ and suggests a crossover between two power-law regimes. Here we explicitly estimate the crossover point.

Let us introduce a small positive parameter ϵ as

$$n := \frac{2}{1 - \epsilon} > 2 \tag{117}$$

and consider the limit $\epsilon \downarrow 0$. We focus on the discontinuous switching in Eq. (115) between $n > 2$ and $n = 2$. To estimate the crossover point, it is necessary to evaluate their higher-order asymptotic behavior for large ν with nonzero ϵ as given by Eq. (D12). As summarized in Appendix D 4 b, the threshold intensity is estimated to be

$$\lambda^* := \nu_0 e^{2/\epsilon}, \tag{118}$$

which characterizes the crossover between the two regimes. We thus obtain the explicit crossover formula as

$$P_{ss}(\lambda) \propto \begin{cases} \lambda^{-1-a_1} & \text{for } \lambda \ll \lambda^*, \ a_1 := \frac{1}{2} + \frac{1}{2kD\tau} \\ \lambda^{-1-a_2} & \text{for } \lambda \gg \lambda^*, \ a_2 := \frac{1}{2}. \end{cases} \quad (119)$$

The existence of this crossover point can be intuitively understood as follows. Let us go back to the SDE representation (47). Remarkably, the cases $n = 2$ and $n > 2$ are critically different in the sense that the relaxation term $-\nu/\tau$ is the same order as the fluctuation $\sqrt{2D\tilde{g}(\nu)}\xi^G$ for $n = 2$, whereas it is negligible for $n > 2$

$$\left| -\frac{\nu}{\tau} \right| \ll \sqrt{2D\tilde{g}(\nu)}\xi^G \quad (120)$$

for a sufficiently large $\nu \gg \nu^*$. Such a crossover point can be roughly estimated by the relationship $\nu^*/\tau = \sqrt{2D\tilde{g}(\nu^*)}$, suggesting $\nu^* = C^{1/\epsilon}$ with some constant C. We therefore obtain $\ln \lambda^* \propto \epsilon^{-1}$ consistently with Eq. (118).

4. Example 4: Multifractal stress activation model

An interesting example is the MSA model for earthquake triggering proposed in Refs. [21,22] and summarized for our purposes in Sec. II C 1, which corresponds to

$$\tilde{g}(\nu) = \lambda_0 \exp(\beta\nu), \quad (121)$$

with the base intensity $\lambda_0 > 0$ and effective inverse temperature $\beta > 0$. From the steady-state solution (108) of the FPE (46), we obtain the steady-state solution $P_{ss}(\nu)$ for the tension ν and $P_{ss}(\lambda)$ of the intensity $\lambda = \lambda_0 e^{\beta\nu}$:

$$P_{ss}(\nu) = \frac{1}{Z} \frac{e^{-\beta\nu}}{\lambda_0} \exp\left(\frac{e^{-\beta\nu}(1 + \beta\nu)}{\lambda_0 \beta^2 \tau D} \right)$$

$$\Longrightarrow P_{ss}(\lambda) = \frac{1}{\beta Z} \lambda^{-2} \exp\left(\frac{\lambda^{-1}[1 + \ln(\lambda/\lambda_0)]}{\beta^2 \tau D} \right). \quad (122)$$

The derivation of $P_{ss}(\lambda)$ from $P_{ss}(\nu)$ uses the Jacobian relation $d\nu P_{ss}(\nu) = d\lambda P_{ss}(\lambda) \Longleftrightarrow P_{ss}(\nu) = \beta\lambda P_{ss}(\lambda)$. This steady-state intensity distribution exhibits Zipf's law similarly to the aforementioned result (104):

$$P_{ss}(\lambda) \propto \lambda^{-2} \quad \text{for large } \lambda. \quad (123)$$

5. Example 5: Fast-accelerating intensity

Let us focus on a large class of intensity map $g(\nu)$ satisfying

$$\tilde{g}(\nu) \gg \nu^2 \quad \text{for large } \nu, \quad (124)$$

which we refer to as a fast-accelerating intensity (FAI) map. For example, the polynomial intensity (113) beyond second

order and the MSA intensity (121) belong to this class. Fast-accelerating intensity maps are special in the sense that the asymptotic PDF of ν, which is solution of the FPE (46), is given by

$$P_{ss}(\nu) \propto \frac{1}{\tilde{g}(\nu)} \exp\left(-\frac{1}{\tau D} \int \frac{\nu \, d\nu}{\tilde{g}(\nu)} \right)$$

$$= \exp[-\ln \tilde{g}(\nu) - o(\nu^{-c})], \quad (125)$$

with some positive constant $c > 0$. This expression is derived from Eq. (108), considering that $\nu/\tilde{g}(\nu) = o(\nu^{-1})$. We thus obtain a general asymptotic form

$$P_{ss}(\lambda) \propto \lambda^{-1} \left| \frac{d\tilde{g}(\nu)}{d\nu} \right|_{\nu=\tilde{g}^{-1}(\lambda)}^{-1}. \quad (126)$$

C. Robust asymptotic solutions

1. Robust exponential tail for the ramp intensity

Here we show that the exponential tail (92) and (96) for the ramp intensity (91) of the steady-state solution (108) of the FPE (46) remains valid for general symmetric mark distributions, assuming appropriate convergence of the moment-generating function and with an exponential-memory function

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad \rho(y) = \rho(-y), \quad g(\nu) \simeq \nu + \nu_0$$

for large $\nu \Longrightarrow P_{ss}(\lambda) \propto e^{-\lambda/\lambda_{cut}} \quad \text{for large } \lambda, \quad (127)$

up to a subleading contribution in the form of a truncated power law. The parameter λ_{cut} is given by the self-consistent relation

$$\frac{1}{\tau\lambda_{cut}} = \Phi\left(\frac{\eta}{\tau\lambda_{cut}} \right),$$

$$\Phi(x) := \int_{-\infty}^{\infty} dy \, \rho(y)(e^{xy} - 1) = \int_{-\infty}^{\infty} dy \, \rho(y)(\cosh yx - 1), \quad (128)$$

where $\Phi(x)$ is the moment-generating function. The equation for λ_{cut} has a single positive solution (see Appendix F).

Derivation. The solution (127) can be derived by direct substitution into the ME (31) as follows. Let us make an ansatz that the solution is given by

$$P_{ss}(\nu) \propto e^{-\nu/\lambda_{cut}} \quad \text{for large } \nu. \quad (129)$$

By considering the relations for large ν,

$$\frac{1}{\tau} \frac{d}{d\nu}(\nu e^{-\nu/\lambda_{cut}}) + \int_{-\infty}^{\infty} dy \, \rho(y)\left(\nu + \nu_0 - \frac{\eta y}{\tau} \right) e^{-\nu/\lambda_{cut} + \eta y/\tau\lambda_{cut}} - (\nu + \nu_0)e^{-\nu/\lambda_{cut}}$$

$$= \nu e^{-\nu/\lambda_{cut}} \left(-\frac{1}{\tau\lambda_{cut}} + \int_{-\infty}^{\infty} dy \, \rho(y)e^{\eta y/\tau\lambda_{cut}} - 1 \right) + o(\nu e^{-\nu/\lambda_{cut}})$$

$$= \nu e^{-\nu/\lambda_{cut}} \left[-\frac{1}{\tau\lambda_{cut}} + \Phi\left(\frac{\eta}{\tau\lambda_{cut}} \right) \right] + o(\nu e^{-\nu/\lambda_{cut}}), \quad (130)$$

the ME (31) in the steady state reads

$$\nu e^{-\nu/\lambda_{\text{cut}}}\left[-\frac{1}{\tau\lambda_{\text{cut}}}+\Phi\left(\frac{\eta}{\tau\lambda_{\text{cut}}}\right)\right]+o(\nu e^{-\nu/\lambda_{\text{cut}}})=0 \quad (131)$$

for large ν. This relation is equivalent to the self-consistent relation (128).

2. Robust power-law tail for quadratic intensity

We show that the power-law tail for the quadratic intensity, such as Eqs. (102) and (112), of the steady-state solution (108) of the FPE (46) is generally valid for general symmetric mark size distributions

$$h(t)=\frac{\eta}{\tau}e^{-t/\tau}, \quad \rho(y)=\rho(-y),$$

$$g(\nu)\simeq k\nu^2+\lambda_0 \quad \text{for large } \nu \Longrightarrow P_{\text{ss}}(\lambda)\propto\lambda^{-1-a} \quad (132)$$

for large λ with $a>1/2$, assuming appropriate convergence of the KM coefficients.

Note that the authors of Ref. [28] conjectured that the PDF of the intensity of various Zumbach Hawkes processes

should be a power law with a nonuniversal exponent. Our results confirm this conjecture, as least for an exponential-memory kernel, in the sense that the power-law asymptotics with a nonuniversal exponent $a>1/2$ is a robust property of Zumbach Hawkes processes, independently of the shape of the mark distribution, as long as it is symmetric with finite moments.

Derivation. Let us return to the ME in the steady state,

$$\frac{1}{\tau}\frac{\partial}{\partial\nu}[\nu P_{\text{ss}}(\nu)]$$

$$+\int_{-\infty}^{\infty}dy\,\rho(y)\left[k\left(\nu-\frac{\eta y}{\tau}\right)^2+\lambda_0\right]$$

$$\times P_{\text{ss}}\left(\nu-\frac{\eta y}{\tau}\right)-(k\nu^2+\lambda_0)P_{\text{ss}}(\nu)=0. \quad (133)$$

We make the ansatz that the asymptotic solution is given by

$$P_{\text{ss}}(\nu)\simeq C\nu^{-\kappa}+o(\nu^{-\kappa}) \quad \text{for large } \nu, \quad (134)$$

with a positive κ and a certain constant C. This implies that

$$C\left\{\frac{1-\kappa}{\tau}\nu^{-\kappa}+\int_{-\infty}^{\infty}dy\,\rho(y)\left[k\left(\nu-\frac{\eta y}{\tau}\right)^2+\lambda_0\right]\left(\nu-\frac{\eta y}{\tau}\right)^{-\kappa}-(k\nu^2+\lambda_0)\nu^{-\kappa}+o(\nu^{-\kappa})\right\}$$

$$=C\left\{\frac{1-\kappa}{\tau}\nu^{-\kappa}+\int_{-\infty}^{\infty}dy\,\rho(y)\left[k\nu^{-\kappa+2}\left(1-\frac{\eta y}{\tau\nu}\right)^{-\kappa+2}+\lambda_0\nu^{-\kappa}\left(1-\frac{\eta y}{\tau\nu}\right)^{-\kappa}\right]-(k\nu^2+\lambda_0)\nu^{-\kappa}+o(\nu^{-\kappa})\right\}$$

$$=C\left\{\frac{1-\kappa}{\tau}\nu^{-\kappa}+\int_{-\infty}^{\infty}dy\,\rho(y)\left[k\nu^{-\kappa+2}\left(1+\frac{(\kappa-2)\eta y}{\tau\nu}+\frac{\eta^2 y^2(\kappa-1)(\kappa-2)}{2\tau^2\nu^2}\right)+\lambda_0\nu^{-\kappa}\right]-(k\nu^2+\lambda_0)\nu^{-\kappa}+o(\nu^{-\kappa})\right\}$$

$$=C\left(\frac{1-\kappa}{\tau}\nu^{-\kappa}+\frac{k\alpha_2\eta^2}{2\tau^2}(\kappa-1)(\kappa-2)\nu^{-\kappa}+o(\nu^{-\kappa})\right)\simeq 0, \quad (135)$$

which leads to the self-consistent relation

$$\kappa=2+\frac{1}{kD\tau}, \quad D:=\frac{\alpha_2\eta^2}{2\tau^2}. \quad (136)$$

We thus obtain the power-law tail of the intensity distribution

$$P_{\text{ss}}(\lambda)\propto\lambda^{-1-a}, \quad a:=\frac{1}{2}+\frac{1}{2kD\tau}. \quad (137)$$

3. Robust Zipf's law for the multifractal stress activation model

We have shown that the exact steady-state solution of the FPE (46) exhibits Zipf's law for the MSA model (121) with exponential-memory kernel in the diffusive limit. Here we show that Zipf's law universally and robustly appears for the MSA model with any general symmetric mark distribution, on the condition that the memory is exponential and the appropriate integrals converge, i.e.,

$$h(t)=\frac{\eta}{\tau}e^{-t/\tau}, \quad g(\nu)=\lambda_0 e^{\beta\nu}, \quad \rho(y)=\rho(-y)$$

$$\Longrightarrow P_{\text{ss}}(\lambda)\propto\lambda^{-2} \quad \text{for large } \lambda. \quad (138)$$

a. Derivation. By defining $\phi(\nu):=g(\nu)P_{\text{ss}}(\nu)$, the steady-state ME is given by

$$\frac{1}{\tau}\frac{\partial}{\partial\nu}[\nu e^{-\beta\nu}\phi(\nu)]+\int_{-\infty}^{\infty}dy\,\rho(y)\phi\left(\nu-\frac{\eta y}{\tau}\right)-\phi(\nu)=0. \quad (139)$$

For large ν, the first term on the left-hand side is negligible due to the exponential factor $e^{-\beta\nu}$, implying

$$\int_{-\infty}^{\infty}dy\,\rho(y)\phi\left(\nu-\frac{\eta y}{\tau}\right)-\phi(\nu)\simeq 0 \quad \text{for large } \nu. \quad (140)$$

Assuming that $\phi(\nu)$ is non-negative, this integral equation has a general solution

$$\phi(\nu)=C_0+C_1\nu, \quad (141)$$

with constants C_0 and C_1 (see Appendix F). By imposing the natural boundary condition, C_1 must be zero, as shown later, and therefore the general solution is given by $\phi(\nu)=C_0$. This implies the following asymptotic form of the steady-state PDF:

$$P_{\text{ss}}(\nu)=\frac{\phi(\nu)}{g(\nu)}\propto e^{-\beta\nu}. \quad (142)$$

We thus obtain Zipf's law for the intensity PDF, from the Jacobian relation $P_{\text{ss}}(\lambda)=P_{\text{ss}}(\nu)|d\nu/d\lambda|$.

b. Natural boundary condition. Here we impose the natural boundary condition to remove C_1. Let us use the KM expansion (33) to define the probability current as

$$\frac{\partial P_t(v)}{\partial t} = -\frac{\partial}{\partial v} J_t(v),$$

$$J_t(v) := -\frac{1}{\tau}[v P_t(v)] - \sum_{k=1}^{\infty} \frac{\alpha_{2k}}{(2k)!} \frac{\eta^{2k}}{\tau^{2k}} \frac{\partial^{2k-1}}{\partial v^{2k-1}} g(v) P_t(v).$$

(143)

For the steady-state distribution, let us ignore the first term in $J_t(v)$ for large v to obtain

$$J_{ss}(v) \simeq -\sum_{k=1}^{\infty} \frac{\alpha_{2k}}{(2k)!} \frac{\eta^{2k}}{\tau^{2k}} \frac{\partial^{2k-1}}{\partial v^{2k-1}} g(v) P_{ss}(v) \quad \text{for large } v.$$

(144)

By direct substitution of the general solution $g(v)P_{ss}(v) = \phi(v) = C_0 + C_1 v$, we obtain

$$J_{ss}(v) \simeq -\sum_{k=1}^{\infty} \frac{\alpha_{2k}}{(2k)!} \frac{\eta^{2k}}{\tau^{2k}} \frac{\partial^{2k-1}}{\partial v^{2k-1}} (C_0 + C_1 v) = -\frac{\eta^2 \alpha_2}{2\tau^2} C_1.$$

(145)

Since the natural boundary condition implies $\lim_{v \to \infty} J_t(v) = 0$ for any t, we obtain $C_1 = 0$.

4. Robust asymptotic form for fast-accelerating intensity maps

We now show that the asymptotic form (126) of the steady-state solution (108) of the FPE (46) is robust even for general mark distribution for any FAI map:

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad g(v) \gg v^2 \text{ (for large } v), \quad \rho(y) = \rho(-y)$$

$$\implies P_{ss}(\lambda) \propto \lambda^{-1} \left| \frac{dg(v)}{dv} \right|_{v=g^{-1}(\lambda)}^{-1} \quad \text{for large } \lambda.$$

(146)

a. Derivation. By defining $\phi(v) := g(v)P_{ss}(v)$, the steady-state ME is given by

$$\frac{1}{\tau} \frac{\partial}{\partial v} \left(\frac{v}{g(v)} \phi(v) \right) + \int_{-\infty}^{\infty} dy \, \rho(y) \phi \left(v - \frac{\eta y}{\tau} \right) - \phi(v) = 0.$$

(147)

For large v, the first term on the left-hand side is negligible because $g(v)$ is an FAI. The self-consistency of this assumption will be confirmed later. This implies

$$\int_{-\infty}^{\infty} dy \, \rho(y) \phi \left(v - \frac{\eta y}{\tau} \right) - \phi(v) \simeq 0 \quad \text{for large } v.$$

(148)

Assuming the non-negativity of $\phi(v)$ and the natural boundary condition, this integral equation has a single solution $\phi(v) = C_0$ with a constant C_0, using the same logic as that in Sec. V C 3. Finally, this implies the following asymptotic form of the steady-state PDF of v:

$$P_{ss}(v) \propto \frac{1}{g(v)}.$$

(149)

The formula (146) for the intensity PDF then derives from the Jacobian relation $P_{ss}(\lambda) = P_{ss}(v)|dv/d\lambda|$.

b. Self-consistency of the assumption. Let us check whether this solution is consistent with the assumption that the first term in Eq. (147) is irrelevant for large v. For simplicity, we focus on the case of $g(v) = v^n$ with integer $n \geqslant 2$. We first assume the expansion of the solution

$$\phi(v) = \phi_0(v) + \phi_1(v) + \cdots, \quad \phi_0(v) = C^{[0]},$$

$$|\phi_0(v)| \gg |\phi_1(v)|$$

(150)

for large v with a constant $C^{[0]}$. By assuming that the first term on the left-hand side of Eq. (147) is subleading, we obtain

$$\frac{(1-n)C^{[0]}}{\tau} v^{-n} + \int_{-\infty}^{\infty} dy \, \rho(y) \phi_1 \left(v - \frac{\eta y}{\tau} \right) - \phi_1(y) \simeq 0.$$

(151)

We solve this nonhomogeneous integral equation by assuming a solution ansatz

$$\phi_1(v) \simeq C^{[1]} v^{-\kappa}, \quad \kappa > 0.$$

(152)

By using $\int_{-\infty}^{\infty} y \rho(y) dy = 0$, we obtain

$$\int_{-\infty}^{\infty} dy \, \rho(y) \phi_1 \left(v - \frac{\eta y}{\tau} \right)$$

$$= \int_{-\infty}^{\infty} dy \, \rho(y) C^{[1]} \left(v - \frac{\eta y}{\tau} \right)^{-\kappa}$$

$$= \int_{-\infty}^{\infty} dy \, \rho(y) C^{[1]} v^{-\kappa} \left(1 - \frac{\eta y}{\tau} v^{-1} \right)^{-\kappa}$$

$$= \int_{-\infty}^{\infty} dy \, \rho(y) C^{[1]} v^{-\kappa} \left(1 + \frac{\kappa \eta y}{\tau} v^{-1} \right.$$

$$\left. + \frac{\eta^2 \kappa(\kappa+1) y^2}{2\tau^2} v^{-2} + O(v^{-3}) \right)$$

$$= C^{[1]} v^{-\kappa} [1 + C^{[2]} v^{-2} + O(v^{-3})],$$

(153)

with the constant $C^{[2]}$ defined by

$$C^{[2]} := \frac{\eta^2 \kappa(\kappa+1)}{2\tau^2} \int_{-\infty}^{\infty} y^2 \rho(y) dy.$$

(154)

Equation (151) is thus equivalent to

$$C^{[1]} C^{[2]} v^{-\kappa-2} \simeq \frac{(n-1)C^{[0]}}{\tau} v^{-n}.$$

(155)

This implies that the exponent κ must satisfy

$$\kappa = n - 2.$$

(156)

This means that the subleading term $\phi_1(v)$ is actually negligible when $\kappa > 0 \iff n > 2$. We thus confirm that the first term in Eq. (147) can be dropped for FAI maps with dependence on v faster than v^2 for large v.

VI. SOLUTION 3: EXPONENTIAL-MEMORY KERNEL WITH TWO-SIDED ASYMMETRIC MARK DISTRIBUTION FOR LINEAR TO FAST-ACCELERATING INTENSITY MAPS

We study here both exact and asymptotic results for the case with the exponential-memory kernel and the two-sided
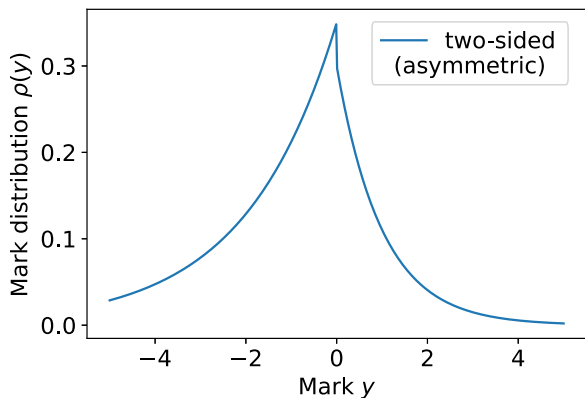
FIG. 8. Schematic of the two-sided asymmetric exponential mark distribution (158) with $(p_+, p_-, y_+^*, y_-^*) = (0.3, 0.7, 1, 2)$.

asymmetric mark distribution with negative-mean mark

$$h(t) = \frac{\eta}{\tau}e^{-t/\tau}, \quad p_+ := \int_0^\infty dy\, \rho(y) > 0,$$

$$p_- := \int_{-\infty}^0 dy\, \rho(y) > 0, \quad m := \int_{-\infty}^\infty y\rho(y)dy < 0 \quad (157)$$

and consider FAI maps $g(\nu) \gg \nu^2$.

A. Exact solution for two-sided asymmetric exponential mark distribution

Let us focus on the case with the two-sided asymmetric exponential mark distribution (see Fig. 8):

$$\rho(y) = \begin{cases} \frac{p_+}{y_+^*}e^{-y/y_+^*}, & y \geq 0 \\ \frac{p_-}{y_-^*}e^{y/y_-^*}, & y < 0, \end{cases} \quad (158)$$

where $p_+ + p_- = 1$, $y_+^* > 0$, and $y_-^* > 0$. The mean mark is given by

$$m := y_+^* p_+ - y_-^* p_- < 0. \quad (159)$$

By using the identities (62) and (85), the ME reads

$$\frac{\partial P_t(\nu)}{\partial t} = \frac{1}{\tau}\partial_\nu[\nu P_t(\nu)] + \partial_\nu \frac{\left(\frac{p_-}{c_-} - \frac{p_+}{c_+}\right) + \frac{\partial_\nu}{c_+c_-}}{\left(1 + \frac{\partial_\nu}{c_+}\right)\left(1 - \frac{\partial_\nu}{c_-}\right)}g(\nu)P_t(\nu),$$

$$c_\pm := \frac{\tau}{\eta y_\pm^*}. \quad (160)$$

This means that the ME expresses the condition of probability conservation

$$\frac{\partial P_t(\nu)}{\partial t} = -\frac{\partial J_t(\nu)}{\partial \nu},$$

$$J_t(\nu) := -\frac{\nu}{\tau}P_t(\nu) - \frac{\left(\frac{p_-}{c_-} - \frac{p_+}{c_+}\right) + \frac{\partial_\nu}{c_+c_-}}{\left(1 + \frac{\partial_\nu}{c_+}\right)\left(1 - \frac{\partial_\nu}{c_-}\right)}g(\nu)P_t(\nu). \quad (161)$$

By requiring the natural boundary condition $\lim_{\nu\to\infty} J_t(\nu) = 0$, we obtain the second-order partial differential equation that

the steady-state PDF $P_{ss}(\nu)$ satisfies

$$\tau\left(A + B\frac{d}{d\nu}\right)\phi(\nu) + \left(1 + C\frac{d}{d\nu} - B\frac{d^2}{d\nu^2}\right)\left(\frac{\nu}{g(\nu)}\phi(\nu)\right) = 0,$$

$$\phi(\nu) := g(\nu)P_{ss}(\nu), \quad (162)$$

with

$$A := \frac{p_-}{c_-} - \frac{p_+}{c_+} := -\frac{\eta}{\tau}m > 0, \quad B := \frac{1}{c_+c_-} > 0,$$

$$C := \frac{1}{c_+} - \frac{1}{c_-}. \quad (163)$$

We note that the coefficients are simplified for the symmetric mark distribution $y_+^* = y_-^*$ and $p_+ = p_- = 1/2$ such that $A = C = 0$.

1. Example 1: Ramp tension-intensity map

For the ramp tension-intensity map

$$\lambda = g(\nu) = \max\{\nu_0, |\nu|\}, \quad (164)$$

we obtain the exact solution for $\lambda > |\nu_0|$,

$$P_{ss}(\lambda) \propto \lambda^{-1}e^{-\lambda/\lambda_{cut}},$$

$$\lambda_{cut} := \frac{C + \tau B + \sqrt{(C + \tau B)^2 + 4B(1 + \tau A)}}{2(1 + \tau A)} > 0 \quad (165)$$

under the natural boundary condition. This implies that an exponential tail is observed for this model.

2. Example 2: Exponential tension-intensity map

For the exponential tension-intensity map

$$\lambda = g(\nu) = \lambda_0|\nu|e^{\beta\nu}, \quad (166)$$

we obtain the second-order partial differential equation for $\nu > 0$ as

$$y\frac{d^2\psi(y)}{dy^2} + \left(2\gamma - 1 - \frac{C}{B\beta} - y\right)\frac{d\psi(y)}{dy}$$

$$+ \left(-\gamma - \frac{A}{B\beta} + \frac{B\beta(\gamma-1)^2 - C(\gamma-1) - \beta^{-1}}{B\beta y}\right) = 0 \quad (167)$$

after a variable transformation

$$y(\nu) = \frac{\lambda_0\tau e^{\beta\nu}}{\beta}, \quad \phi(\nu) = y^\gamma\psi(y). \quad (168)$$

Here γ is either one of the roots of $B\beta(\gamma-1)^2 - C(\gamma-1) - \beta^{-1} = 0$ given by

$$\gamma_1 := 1 + \frac{C - \sqrt{4B + C^2}}{2\beta B}, \quad \gamma_2 := 1 + \frac{C + \sqrt{4B + C^2}}{2\beta B}. \quad (169a)$$

Equation (167) is equivalent to the confluent hypergeometric differential equation (D3). We thus obtain the exact solution for $\nu > 0$ as

$$\phi(\nu) = C^{[1]}y^{\gamma_1}{}_1U(q_1, r_1; y) + C^{[2]}y^{\gamma_2}{}_1U(q_2, r_2; y), \quad (169b)$$

where

$$q_1 := 1 + \frac{C + 2A - \sqrt{4B + C^2}}{2\beta B}, \quad r_1 := 1 - \frac{\sqrt{4B + C^2}}{\beta B},$$
$$(169c)$$

$$q_2 := 1 + \frac{C + 2A + \sqrt{4B + C^2}}{2\beta B}, \quad r_2 := 1 + \frac{\sqrt{4B + C^2}}{\beta B},$$
$$(169d)$$

with two constants of integration $C^{[1]}$ and $C^{[2]}$. For large $\nu \to \infty$, assuming the boundary condition $\lim_{\nu \to \infty} P_{ss}(\nu) = 0$, we obtain the asymptotic formula

$$\phi(\nu) \propto y^{-A/\beta B}, \qquad (170)$$

leading to the PDF tail of the intensity:

$$P_{ss}(\lambda) \propto \lambda^{-2 - \beta^{-1} u}, \quad u = \frac{A}{B} = -\frac{\tau m}{\eta y_+^* y_-^*}. \qquad (171)$$

This result implies that the power-law scaling deviates from Zipf's law in proportion to the amplitude m of the asymmetry of the mark distribution.

B. Robust asymptotic solutions

In this section we generalize the above exact results in the form of robust asymptotic results under a wide range of two-sided mark distributions with negative mean, in the presence of an exponential-memory kernel.

1. Robust exponential tail for the ramp intensity

Let us assume that the ramp intensity is asymptotically

$$g(\nu) \simeq \nu - \nu_1 \quad \text{for large } \nu, \qquad (172a)$$

with an exponential memory with two-sided mark distribution of negative mean:

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad p_+ := \int_0^\infty \rho(y) dy > 0,$$
$$p_- := \int_{-\infty}^0 \rho(y) dy > 0, \quad m := \int_{-\infty}^\infty y \rho(y) dy < 0. \qquad (172b)$$

Here ν_1 is an arbitrary real number, either positive or nonpositive in contrast to the LH process. Under this assumption, we obtain

$$P_{ss}(\lambda) \propto e^{-\lambda/\lambda_{cut}} \quad \text{for large } \lambda. \qquad (172c)$$

The parameter λ_{cut} is given by the self-consistent relation

$$\frac{1}{\tau \lambda_{cut}} = \Phi\left(\frac{\eta}{\tau \lambda_{cut}}\right), \quad \Phi(x) := \int_{-\infty}^\infty dy \, \rho(y)(e^{xy} - 1), \qquad (173)$$

where $\Phi(x)$ is the moment-generating function. The equation for λ_{cut} has a single positive solution (see Appendix F). This relation can be derived by a straightforward generalization of the derivation in Sec. V C 1.

2. Robust power-law tail for fast-accelerating intensity maps

We show that, under the general assumptions

$$h(t) = \frac{\eta}{\tau} e^{-t/\tau}, \quad g(\nu) \gg \nu^2 \quad \text{for large } \nu,$$
$$p_+ := \int_0^\infty \rho(y) dy > 0, \quad p_- := \int_{-\infty}^0 \rho(y) dy > 0,$$
$$m := \int_{-\infty}^\infty y \rho(y) dy < 0, \qquad (174)$$

we obtain the robust asymptotic relationship

$$P_{ss}(\lambda) \propto \lambda^{-1} \left(\left| \frac{dg(\nu)}{d\nu} \right|^{-1} e^{-u\nu} \right)_{\nu = g^{-1}(\lambda)}, \quad u := \frac{\tau c^*}{\eta}, \quad (175)$$

where c^* is the unique positive root of $\Phi(c^*) = 0$, with the moment-generating function is defined by $\Phi(x) := \int_{-\infty}^\infty dy \, \rho(y)(e^{xy} - 1)$.

a. Examples. From this formula, we readily deduce the power-law PDF for the exponential intensity

$$g(\nu) \simeq \lambda_0 e^\beta \implies P_{ss}(\lambda) \propto \lambda^{-2 - \beta^{-1} u}. \qquad (176)$$

We note that this result is consistent with the aforementioned exact result (171) by considering Eq. (F19) in Appendix F for the case with the exponential intensity and the two-sided asymmetric exponential mark distribution.

In addition, we obtain the truncated power-law PDF for the polynomial intensity

$$g(\nu) \simeq \lambda_0 \nu^n, \quad n > 2 \implies P_{ss}(\lambda) \propto \lambda^{-2 + 1/n} e^{-u(\lambda/\lambda_0)^{1/n}}, \qquad (177)$$

where the cutoff length appears due to the asymmetry of the mark distribution. For the zero-mean mark limit $m \uparrow 0$, the cutoff disappears as $u \downarrow 0$.

b. Derivation. By defining $\phi(\nu) = g(\nu) P_{ss}(\nu)$, the ME is given by

$$\frac{1}{\tau} \frac{\partial}{\partial \nu} \left(\frac{\nu}{g(\nu)} \phi(\nu) \right) + \int_{-\infty}^\infty dy \, \rho(y) \phi\left(\nu - \frac{\eta y}{\tau}\right) - \phi(y) = 0. \qquad (178)$$

As an asymptotic assumption for the solution, let us first neglect the first term of Eq. (178) to obtain

$$\int_{-\infty}^\infty dy \, \rho(y) \phi\left(\nu - \frac{\eta y}{\tau}\right) - \phi(y) \simeq 0 \quad \text{for large } \nu \qquad (179)$$

for the case (174). The self-consistency of this assumption will be confirmed later. According to Appendix E, the general solution is given by the superposition of exponentials

$$\phi(\nu) \simeq \sum_i C_i e^{-(\tau c_i/\eta)\nu}, \qquad (180)$$

where the c_i are the roots of the moment-generating function $\Phi(x) = 0$. The moment-generating function is defined by

$$\Phi(c) = 0, \quad \Phi(x) := \int_{-\infty}^\infty dy \, \rho(y)(e^{xy} - 1), \qquad (181)$$

whose analytical characters are summarized in Appendix F. According to Appendix F, $\Phi(x) = 0$ has only two roots at

$x = 0$ and $x = c^* > 0$. This means that the general asymptotic solution is given by

$$\phi(v) \simeq C_0 e^{-(\tau c^*/\eta)v} + C_1, \qquad (182)$$

with integral constants C_0 and C_1. By imposing the natural boundary condition, C_1 must be zero (see below for the natural boundary condition). We thus have the solution

$$\phi(v) \simeq C_0 e^{-(\tau c^*/\eta)v}. \qquad (183)$$

This implies that the steady-state intensity PDF has the asymptotic form

$$P_{ss}(v) = \frac{\phi(v)}{g(v)} \propto \frac{1}{g(v)} e^{-(\tau c^*/\eta)v}, \qquad (184)$$

which implies Eq. (175) from the Jacobian relation $P_{ss}(\lambda) = P_{ss}|dv/d\lambda|$.

c. Self-consistency of the assumption. Finally, we confirm here the self-consistency of the ansatz for the solution under the assumption of FAI maps. Let us assume that the solution is given by the expansion

$$\phi(v) = \phi_0(v) + \phi_1(v) + \cdots, \quad \phi_0(v) := C^{[0]} e^{-(\tau c^*/\eta)v},$$

$$|\phi_0(v)| \gg |\phi_1(v)| \quad \text{for large } v, \qquad (185)$$

with an integral constant $C^{[0]}$. For simplicity, let us focus on the case $g(v) = v^n$ with integer $n > 2$. By assuming that the first term in Eq. (178) is subleading, we substitute this expansion into Eq. (178) to obtain

$$\int_{-\infty}^{\infty} dy\, \rho(y)\phi_1\left(v - \frac{\eta y}{\tau}\right) - \phi_1(y) \simeq -\frac{1}{\tau}\frac{\partial}{\partial v}[v^{1-n}\phi_0(v)] \implies \int_{-\infty}^{\infty} dy\, \rho(y)\phi_1\left(v - \frac{\eta y}{\tau}\right) - \phi_1(y) \simeq C^{[0]}\frac{c^*}{\eta}v^{1-n}e^{-(\tau c^*/\eta)v}. \qquad (186)$$

We make the ansatz for the solution in the form

$$\phi_1(v) \simeq C^{[1]} v^{-\kappa} e^{-(\tau c^*/\eta)v}, \quad \kappa > 0 \qquad (187)$$

to obtain the special solution with a constant $C^{[1]}$. Here the condition $\kappa > 0$ is essential; otherwise the consistency relationship $|\phi_0(v)| \gg |\phi_1(v)|$ does not hold. By direct substitution, Eq. (186) is equivalent to

$$\int_{-\infty}^{\infty} dy\, \rho(y)\phi_1\left(v - \frac{\eta y}{\tau}\right) \simeq C^{[1]} \int_{-\infty}^{\infty} dy\, \rho(y)\left(v - \frac{\eta y}{\tau}\right)^{-\kappa} e^{c^* y - (\tau c^*/\eta)v} = C^{[1]} \int_{-\infty}^{\infty} dy\, \rho(y)v^{-\kappa}\left(1 - \frac{\eta y}{\tau}v^{-1}\right)^{-\kappa} e^{c^* y} e^{-(\tau c^*/\eta)v}$$

$$= C^{[1]} \int_{-\infty}^{\infty} dy\, \rho(y)v^{-\kappa}\left(1 + \frac{\kappa \eta y}{\tau}v^{-1} + O(v^{-2})\right) e^{c^* y} e^{-(\tau c^*/\eta)v}$$

$$= C^{[1]} v^{-\kappa} e^{-(\tau c^*/\eta)v}[\Phi(c^*) + 1 + C^{[2]} v^{-1} + O(v^{-2})], \qquad (188)$$

with

$$C^{[2]} := \frac{\kappa \eta}{\tau} \int_{-\infty}^{\infty} y e^{c^* y} \rho(y)\, dy. \qquad (189)$$

By using $\Phi(c^*) = 0$, we thus obtain

$$C^{[1]} C^{[2]} v^{-\kappa - 1} e^{-(\tau c^*/\eta)v} \simeq C^{[0]}\frac{c^*}{n}v^{1-n}e^{-(\tau c^*/\eta)v}. \qquad (190)$$

This implies that the power-law exponent κ must satisfy the relationship

$$\kappa = n - 2. \qquad (191)$$

Because of the assumption $\kappa > 0$, we obtain the self-consistency condition

$$n > 2, \qquad (192)$$

which is equivalent to the assumption that the tension-intensity maps must be FAI $[g(v) \gg v^2]$.

d. Natural boundary condition. Here we impose the natural boundary condition to remove C_1. Let us use the KM expansion (33) to define the probability current as

$$\frac{\partial P_t(v)}{\partial t} = -\frac{\partial}{\partial v}J_t(v),$$

$$J_t(v) := -\frac{1}{\tau}[vP_t(v)] - \sum_{k=1}^{\infty}\frac{(-1)^k \alpha_k}{k!}\frac{\eta^k}{\tau^k}\frac{\partial^{k-1}}{\partial v^{k-1}}g(v)P_t(v). \qquad (193)$$

For the steady-state distribution, we ignore the first term in $J_t(v)$ for large v to obtain

$$J_{ss}(v) \simeq -\sum_{k=1}^{\infty}\frac{(-1)^k \alpha_k}{k!}\frac{\eta^k}{\tau^k}\frac{\partial^{k-1}}{\partial v^{k-1}}g(v)P_{ss}(v) \quad \text{for large } v. \qquad (194)$$

By direct substitution of the general solution $g(v)P_{ss}(v) = \phi(v) = C_0 + C_1 e^{-(\tau c^*/\eta)v}$, we obtain

$$J_{ss}(v) \simeq -\sum_{k=1}^{\infty}\frac{(-1)^k \alpha_k}{k!}\frac{\eta^k}{\tau^k}\frac{\partial^{k-1}}{\partial v^{k-1}}(C_1 + C_0 e^{-(\tau c^*/\eta)v})$$

$$= \frac{\eta m}{\tau}C_1 + \frac{\eta C_0}{c^* \tau}e^{-(\tau c^*/\eta)v}\sum_{k=1}^{\infty}\frac{\alpha_k}{k!}c^{*k}$$

$$= \frac{\eta m}{\tau}C_1 + \frac{\eta C_0}{c^* \tau}e^{-(\tau c^*/\eta)v}\Phi(c^*)$$

$$= \frac{\eta m}{\tau}C_1, \qquad (195)$$

where we have used $\Phi(x) = \sum_{k=1}^{\infty}(\alpha_k/k!)x^k$ and $\Phi(c^*) = 0$. Since the natural boundary condition implies $\lim_{v \to \infty} J_t(v) = 0$ for any t, we obtain $C_1 = 0$.

VII. SOLUTION 4: GENERAL MEMORY KERNEL FOR ONE-SIDED MARK DISTRIBUTION FOR THE RAMP HAWKES PROCESS

We have studied the exact solution for the NLH process assuming that (i) the memory kernel is exponential and (ii) the jump size obeys the one-sided exponential distributions. In particular, we derived the power-law tail (71) without truncation, at the critical point for the ramp Hawkes process (69). As shown in the following in this section, this exact power-law relation is robust for general ramp Hawkes processes with any memory kernel and jump-size distribution, only assuming the finiteness of

$$\langle \tau \rangle := \int_0^\infty t\, h(t)dt = \int_0^\infty x^2 \tilde{h}(x)dx < \infty, \qquad (196)$$

with $\tilde{h}(x)$ defined in (14a). Note that the critical condition is characterized by

$$\eta := \int_0^\infty h(t)dt = \int_0^\infty x \tilde{h}(x)dx = 1. \qquad (197)$$

A. Discrete sum of exponentials

Let us first consider the case of a discrete sum of exponentials. In this case, we find a power-law asymptotics at the critical point $\eta = 1$

$$h(t) = \sum_{k=1}^K \tilde{h}_k e^{-t/\tau_k},$$

$$\lambda = g(\nu) \simeq \nu - \nu_1 + o(\nu^0) \quad \text{for large } \nu,$$

$$\rho(y) = 0 \text{ for negative } y, \quad \int_0^\infty y\rho(y)dy = 1$$

$$\Longrightarrow P_{\rm ss}(\lambda) \propto \lambda^{-1-a}, \quad a := \frac{2\nu_1 \langle \tau \rangle}{\alpha_2}, \quad \langle \tau \rangle := \sum_{k=1}^K \tau_k^2\, \tilde{h}_k \qquad (198)$$

for either negative or non-negative ν_1. This relation is a true power law for positive ν_1, i.e., normalizable even without cutoff, while it is an intermediate asymptotics for nonpositive ν_1, i.e., not normalizable without cutoff. Note that the critical condition is given by

$$\eta := \sum_{k=1}^K \tau_k \tilde{h}_k = 1. \qquad (199)$$

1. Derivation

Let us first write the asymptotic form of $P_{\rm ss}(z)$ as $S(z)$,

$$P_{\rm ss}(z) = S(z) + R(z), \quad S(z) \gg R(z) \text{ for large } z, \qquad (200)$$

with a correction term $R(z)$ for small z. Here $S(z)$ is assumed to have a fat tail represented by a power law, while $R(z)$ is assumed to have a thinner tail. The ME (23) in the steady state

reduces asymptotically to

$$\sum_{k=1}^K \frac{\partial}{\partial z_k} \frac{z_k}{\tau_k} S(z) + \int_0^\infty dy\, \rho(y)\left(\sum_{k=1}^K (z_k - y\tilde{h}_k) - \nu_1\right) S(z - y\tilde{\boldsymbol{h}})$$
$$- \left(\sum_{k=1}^K z_k - \nu_1\right) S(z) \simeq 0 \quad \text{for large } \nu. \qquad (201)$$

Since the asymptotic form of this ME is the same as that for the LH process presented in Ref. [16] except that ν_1 can be either negative or non-negative, its asymptotic solution for large ν can be obtained from a calculation similar to that presented in Ref. [16]. While we refer the reader to Ref. [16] for an elementary introduction to the calculations, let us sketch the main steps of the derivation. We first define the Laplace transformations

$$\tilde{P}_{\rm ss}(\boldsymbol{s}) := \mathcal{L}_K[P_{\rm ss}(z); \boldsymbol{s}], \quad \tilde{S}(\boldsymbol{s}) := \mathcal{L}_K[S(z); \boldsymbol{s}],$$
$$\tilde{R}(\boldsymbol{s}) := \mathcal{L}_K[R(z); \boldsymbol{s}]. \qquad (202)$$

Since $P_{\rm ss(z)}$ is a PDF, the normalization implies that $\int dz\, P_{\rm ss}(z) = \tilde{P}_{\rm ss}(\boldsymbol{s} = \boldsymbol{0}) = 1$. However, $S(z)$ is just an asymptotic form of the PDF and there is no guarantee that $\tilde{S}(\boldsymbol{s} = \boldsymbol{0}) = 1$. For example, assuming that $\tilde{S}(s\boldsymbol{1}) = As^a + o(s^a)$ with any noninteger number a and indicator vector $\boldsymbol{1} := (1, 1, \ldots, 1)$, we can expand $\tilde{P}_{\rm ss}(\boldsymbol{s}) := \tilde{P}_{\rm ss}(s\boldsymbol{1})$ as

$$\tilde{P}_{\rm ss}(\boldsymbol{s}) = \tilde{S}(s\boldsymbol{1}) + \tilde{R}(s\boldsymbol{1}) \simeq As^a + \sum_{k=0}^m c_k s^k + o(s^a),$$

$$\tilde{R}(s\boldsymbol{1}) = \sum_{k=0}^m c_k s^k + o(s^a) \qquad (203)$$

for small $\boldsymbol{s}$ with $m := \max(\lfloor a \rfloor, 0)$ and the floor function $\lfloor x \rfloor = \max\{k \in \boldsymbol{Z} \mid k \leqslant x\}$ with the set of integers $\boldsymbol{Z}$. The normalization condition requires $\tilde{P}(s = 0) = c_0 = 1$. By applying the Laplace transformation to the steady-state ME (201), we obtain

$$-\sum_{k=1}^K \frac{s_k}{\tau_k} \frac{\partial \tilde{S}(\boldsymbol{s})}{\partial s_k} - [\Phi(\boldsymbol{s}) - 1]\left(\nu_1 + \sum_{k=1}^K \frac{\partial}{\partial s_k}\right)\tilde{S}(\boldsymbol{s}) \simeq 0,$$

$$\Phi(\boldsymbol{s}) := \int_0^\infty dy\, \rho(y)e^{-y\tilde{\boldsymbol{h}}\cdot\boldsymbol{s}}, \qquad (204)$$

which is valid for small $\boldsymbol{s}$. Considering $[1/\tilde{S}(\boldsymbol{s})]\partial\tilde{S}(\boldsymbol{s})/\partial s_k = (\partial/\partial s_k)\ln|\tilde{S}(\boldsymbol{s})|$, this equation can be rewritten as

$$\sum_{k=1}^K \left(1 - \Phi(\boldsymbol{s}) - \frac{s_k}{\tau_k}\right)\frac{\partial\Psi(\boldsymbol{s})}{\partial s_k} \simeq \nu_1[\Phi(\boldsymbol{s}) - 1],$$

$$\Psi(\boldsymbol{s}) := \ln|\tilde{S}(\boldsymbol{s})|. \qquad (205)$$

Since this equation belongs to the class of first-order partial differential equations, it can be solved by the method of characteristics. Let us thus consider the corresponding Lagrange-Charpit equations

$$\frac{ds_k}{dl} = 1 - \frac{s_k}{\tau_k} - \Phi(\boldsymbol{s}), \quad \frac{d\Psi}{dl} = \nu_1[\Phi(\boldsymbol{s}) - 1], \qquad (206)$$

with a parameter l describing the position on characteristic curves. By regarding l as an imaginary time of this system,
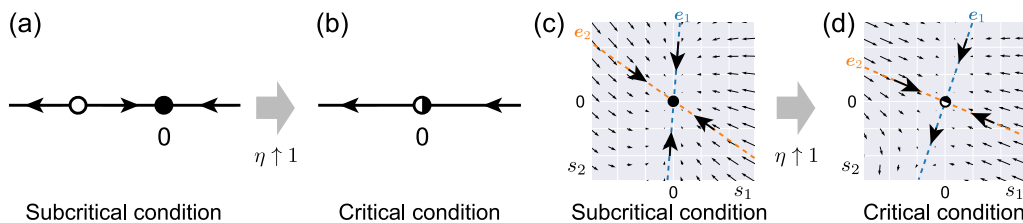
FIG. 9. Schematics of the vector field $\boldsymbol{V}(\boldsymbol{s}) := d\boldsymbol{s}/dl$ for the special cases $K = 1$ (a) below and (b) at criticality and $K = 2$ (c) below and (d) at criticality. The origin $\boldsymbol{s} = \boldsymbol{0}$ is the stable attractor below criticality $\eta < 1$. An unstable attractor merges into $\boldsymbol{s} = 0$ at criticality $\eta \uparrow 1$, which is consistent with the standard picture of the transcritical bifurcation.

we can apply the standard bifurcation theory of dynamical systems. Since we are interested in the regime of small $\boldsymbol{s}$, let us consider the long-time asymptotic limit $l \to \infty$, where $\boldsymbol{s}(l)$ relaxes to the attractor at $\boldsymbol{s} = \boldsymbol{0}$ [see the schematic in Figs. 9(a) and 9(c) for the vector field $\boldsymbol{V}(\boldsymbol{s}) := d\boldsymbol{s}/dl$ for the cases $K = 1$ and 2 below criticality $\eta < 1$, respectively] such that $\lim_{l \to \infty} \boldsymbol{s}(l) = \boldsymbol{0}$. Let us expand the equations for small $\boldsymbol{s}$,

$$\frac{ds_k}{dl} \simeq -\boldsymbol{H}\boldsymbol{s} - \frac{\alpha_2}{2}\left(\sum_{j=1}^{K} \tilde{h}_j s_j\right)^2 + o(\boldsymbol{s}^2), \quad (207a)$$

$$\frac{d\Psi}{dl} \simeq -\nu_1 \boldsymbol{K}\boldsymbol{s} + o(\boldsymbol{s}), \quad (207b)$$

with

$$\boldsymbol{H} := \begin{pmatrix} \frac{1}{\tau_1} - \tilde{h}_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K \\ -\tilde{h}_1 & \frac{1}{\tau_2} - \tilde{h}_2 & \cdots & -\tilde{h}_K \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{h}_1 & -\tilde{h}_2 & \cdots & \frac{1}{\tau_K} - \tilde{h}_K \end{pmatrix},$$

$$\boldsymbol{K} := (\tilde{h}_1, \ldots, \tilde{h}_K). \quad (208)$$

Note that the matrix $\boldsymbol{H}$ is the same as that in Refs. [16,17]. Defining its eigenvalues $\{\lambda_k\}_{k=1,\ldots,K}$ and corresponding eigenvectors $\{\boldsymbol{e}_k\}_{k=1,\ldots,K}$ by

$$\boldsymbol{H}\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k, \quad (209)$$

$\boldsymbol{H}$ has the following mathematical properties (see Ref. [17] and Appendix G for details).

(i) All the eigenvalues are real: $\lambda_k \in \boldsymbol{R}^1$. Accordingly, we assume that $\lambda_i \leqslant \lambda_j$ for $i < j$.

(ii) The determinant of $\boldsymbol{H}$ is given by

$$\det \boldsymbol{H} = \frac{1 - \sum_{k=1}^{K} \tau_k \tilde{h}_k}{\prod_{k=1}^{K} \tau_k}. \quad (210)$$

This means that the zero eigenvalue appears at criticality $\eta := \sum_{k=1}^{K} \tau_k \tilde{h}_k = 1$.

(iii) The determinant $\boldsymbol{H}$ can be diagonalized by $\boldsymbol{P}$ as follows:

$$\boldsymbol{P} := (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_K), \quad \boldsymbol{P}^{-1}\boldsymbol{H}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix}. \quad (211)$$

(iv) Let us introduce a new representation based on the eigenvectors:

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} := \boldsymbol{P}^{-1}\boldsymbol{s}, \quad \boldsymbol{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_K \end{pmatrix}, \quad \boldsymbol{P}^{-1} = \begin{pmatrix} \boldsymbol{g}_1 \\ \boldsymbol{g}_2 \\ \vdots \\ \boldsymbol{g}_K \end{pmatrix}. \quad (212)$$

At criticality $\eta = 1$, the smallest eigenvalue is zero, $\lambda_1 = 0$, and its eigenvector is given by

$$\boldsymbol{e}_1 = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_K \end{pmatrix}. \quad (213)$$

In addition, X_1 is represented by

$$X_1 = \boldsymbol{g}_1 \cdot \boldsymbol{s} = \frac{1}{\langle \tau \rangle} \sum_{k=1}^{K} \tau_k \tilde{h}_k s_k, \quad \boldsymbol{g}_1 = \left(\frac{\tau_1 \tilde{h}_1}{\langle \tau \rangle}, \ldots, \frac{\tau_K \tilde{h}_K}{\langle \tau \rangle}\right),$$

$$\langle \tau \rangle := \sum_{k=1}^{K} \tau_k^2 \tilde{h}_k. \quad (214)$$

Given these properties, let us consider the Lagrange-Charpit equation (207a) in the representation $\boldsymbol{X} := (X_1, \ldots, X_K)^T$. At criticality $\eta = 1$, the leading-order contribution in the Lagrange-Charpit equation (207a) is given by

$$\frac{dX_1}{dl} \simeq 0 - \frac{\alpha_2}{2\langle \tau \rangle}\left(\sum_{k=1}^{K} \tilde{h}_k s_k\right)^2 + o(\boldsymbol{X}^2),$$

$$\frac{dX_j}{dl} = -\lambda_j X_j + o(\boldsymbol{X}) \quad \text{for } j \geqslant 2. \quad (215)$$

Since the leading-order contribution will come from the X_1 direction, we can assume that $|X_1| \gg |X_j|$ for $j \geqslant 2$ for large l. We thus ignore contributions other than X_1 by assuming $X_j \simeq 0$ for $j \geqslant 2$:

$$\boldsymbol{s} = \boldsymbol{P}\boldsymbol{X} \simeq (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_K)\begin{pmatrix} X_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = X_1 \boldsymbol{e}_1. \quad (216)$$

We thus obtain

$$\frac{dX_1}{dl} \simeq 0 - \frac{\alpha_2}{2\langle\tau\rangle}X_1^2 + o(X^2),$$

$$\frac{dX_j}{dl} = -\lambda_j X_j + o(X) \quad \text{for } j \geqslant 2. \tag{217}$$

This is the standard normal form of the transcritical bifurcation when regarding l as a physical time [see the schematics in Figs. 9(b) and 9(d) for $K = 1$ and 2 at criticality $\eta \uparrow 1$, respectively]. The solution is given by

$$X_1(l) \simeq \frac{2\langle\tau\rangle}{\alpha_2}\frac{1}{l - l_0}, \quad X_j(l) \simeq C_j e^{-\lambda_j(l-l_0)}, \tag{218}$$

with integral constants l_0 and C_j for $j \geqslant 2$. We can assume $l_0 = 0$ as the initial point of the characteristic curve without losing generality. From expanding $\Phi(s)$, we obtain the solution

$$\Psi(l) \simeq -\nu_1 \int dl \sum_{k=1}^{K} \tilde{h}_k s_k(l) \simeq \frac{2\nu_1\langle\tau\rangle}{\alpha_2}\ln|X|_1 + O(X) + C_0, \tag{219}$$

with an integral constant C_0. According to the method of characteristics, the general solution is given by

$$\mathcal{H}(C_2, , \ldots, C_K) = C_0, \tag{220}$$

with a function $\mathcal{H}$ which needs to be determined by the initial condition. The constants C_j with $j \geqslant 2$ are related to each other such that

$$l = \frac{2\langle\tau\rangle}{\alpha_2 X_1}, \quad C_j = X_j \exp\left(\frac{2\langle\tau\rangle\lambda_j}{\alpha_2 X_1}\right). \tag{221}$$

This means that the explicit form of the general solution is given by

$$\Psi(X) \simeq \frac{2\nu_1\langle\tau\rangle}{\alpha_2}\ln|X_1| + O(X)$$
$$+ \mathcal{H}\left(X_2 \exp\left(\frac{2\langle\tau\rangle\lambda_2}{\alpha_2 X_1}\right), \ldots, X_K \exp\left(\frac{2\langle\tau\rangle\lambda_K}{\alpha_2 X_1}\right)\right). \tag{222}$$

Note the existence of the divergent term $\ln|X_1|$ resulting from neglecting the UV cutoff. Since $\Psi(X)$ must be constant for $s \to 0$, except for the artificial logarithmic divergence, we obtain

$$\lim_{X \to 0} \mathcal{H}\left(X_2 \exp\left(\frac{2\langle\tau\rangle\lambda_2}{\alpha_2 X_1}\right), \ldots, X_K \exp\left(\frac{2\langle\tau\rangle\lambda_K}{\alpha_2 X_1}\right)\right) = \text{const.} \tag{223}$$

Let us now consider the specific limit $X_1 \to 0$, by writing

$$X_j = Z_j \exp\left(-\frac{2\langle\tau\rangle\lambda_j}{\alpha_2 X_1}\right), \tag{224}$$

with any positive number Z_j for $j \geqslant 2$. This specific limit satisfies the relation

$$\lim_{X_1 \to 0} X = 0. \tag{225}$$

Since Eq. (223) should hold for any path taken to reach the limit $X \to 0$, we obtain the relation even for the specific limit

$$\lim_{X_1 \to 0} \mathcal{H}(Z_2, \ldots, Z_K) = \text{const} \tag{226}$$

for any positive $\{Z_j\}_{j=2,\ldots,K}$, implying that $\mathcal{H}$ is a constant function. We thus obtain

$$|\tilde{S}(z)| = \exp\Psi(X) \simeq \tilde{A}s^a, \quad a := \frac{2\nu_1\langle\tau\rangle}{\alpha_2}, \tag{227}$$

with some positive number $\tilde{A}$.

a. Case with negative $a < 0$. When ν_1 is negative, we have

$$\tilde{P}_{ss}(s) \simeq As^a + o(s^a) \tag{228}$$

for small s with some constant A satisfying $|A| = \tilde{A}$ and negative value $a < 0$. By applying the inverse Laplace transform (see Appendix H 1), we obtain the power-law asymptotic form (198). In this case, the sign of A is determined to be positive, i.e., $A = \tilde{A}$, for the consistency with the probability interpretation.

b. Case with $0 < a < 1$. This case is equivalent to $m := \lfloor a \rfloor = 0$. We obtain

$$\tilde{P}_{ss}(s) \simeq 1 - As^a + o(s^a), \tag{229}$$

with some constant A for small s. Assuming that $\tilde{A}$ is a positive real number, we obtain the power-law asymptotic form (198) (see Appendix H 2).

c. Case with positive noninteger a. Let us define $m := \lfloor a \rfloor$ in order to classify the solutions. Since the asymptotic series of the Laplace transformation is given by

$$\tilde{P}_{ss}(s) \simeq As^a + \sum_{k=0}^{m} c_k s^k + o(s^a), \tag{230}$$

we obtain the power-law asymptotic form (198) (see Appendix H 3), by setting A to a positive (negative) number for even (odd) m, for consistency with the probability interpretation.

d. Case with positive integer a. Technically, the positive-integer case requires a special treatment since the Γ function in the Laplace transformation formula (H14) diverges: $\Gamma(-a) = \infty$. However, since the power-law asymptotics (198) is valid for any integer a, it is straightforward to obtain the power-law asymptotics (198) for positive integer a, assuming that the power-law exponent a is a continuous function in terms of ν_1:

$$a(\nu_1) = \lim_{x \to \nu_1} a(x). \tag{231}$$

While we have numerically checked the validity of this result (198) for some specific cases (see Sec. VII C for the numerical results), a rigorous proof of the continuity assumption (231) is beyond the scope of this paper, as it requires further technical investigation while the continuity assumption (231) is physically reasonable.

In summary, we obtain the power-law asymptotics (198) for general a.

B. General memory kernel

Since the power-law asymptotics (198) holds for general discrete sums of exponentials, as a straightforward
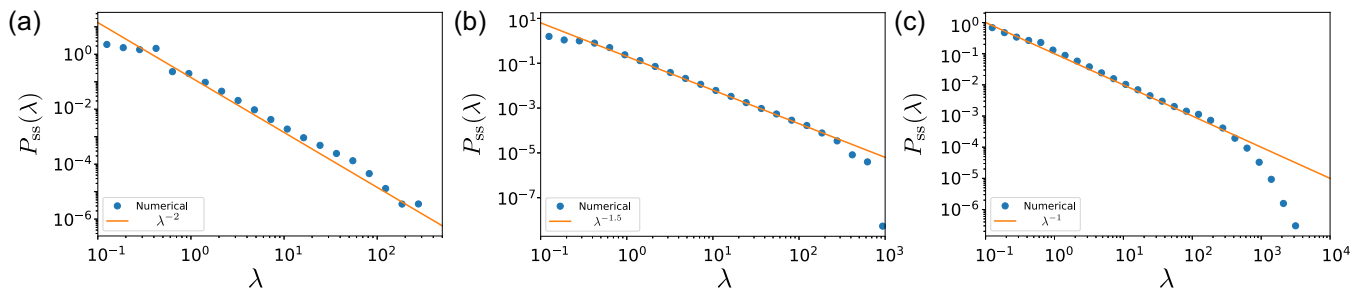
FIG. 10. Numerical confirmation of our theoretical prediction on the power-law exponents (198). (a) Simulation based on $K = 2$, $(\tau_1, \tau_2) = (1, 2)$, $(\tilde{h}_1, \tilde{h}_2) = (0.7, 0.149\,95)$, $\eta = 0.9999$, $\nu_0 = 0.01$, and $\nu_1 \simeq 0.385$, predicting $a \simeq 1.0$, i.e., a true power law (Zipf's law). (b) Simulation based on $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 2, 3)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.15, 0.1999/3)$, $\eta = 0.9999$, $\nu_0 = 0.01$, and $\nu_1 \simeq 0.147$, predicting $a \simeq 0.5$, i.e., a true power law. (c) Simulation based on $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 2, 3)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.15, 0.1999/3)$, $\eta = 0.9999$, $\nu_0 = 0.01$, and $\nu_1 = 0$, predicting $a = 0$, i.e., an intermediate power-law asymptotic.

generalization, we find a power-law asymptotics at the critical point $n = 1$ for general memory kernel $h(t)$ such that

$$h(t) = \int_0^\infty dx\, \tilde{h}(x) e^{-t/x},$$

$$\lambda = g(\nu) \simeq \nu - \nu_1 + o(\nu^0) \text{ for large } \nu,$$

$$\rho(y) = 0 \text{ for negative } y,$$

$$\int_0^\infty y\rho(y)dy = 1 \Longrightarrow P_{\text{ss}}(\lambda) \propto \lambda^{-1-a},$$

$$a := \frac{2\nu_1 \langle \tau \rangle}{\alpha_2}, \quad \langle \tau \rangle := \int_0^\infty x^2 \tilde{h}(x)dx \quad (232)$$

for either negative or non-negative ν_1. This relation is a true power law for positive ν_1, i.e., normalizable even without cutoff, while it is an intermediate asymptotics for nonpositive ν_1, i.e., not normalizable without cutoff.

C. Numerical confirmation

Figure 10 shows the numerical results based on the Monte Carlo simulation of the SDE (3) obtained for a memory function constructed as a discrete sum of exponentials, for the ramp intensity (69), and a mark distribution reducing to the Dirac function centered on $y = 1$:

$$h(t) = \sum_{k=1}^K \tilde{h}_k e^{-t/\tau_k}, \quad \lambda = g(\nu) = \max\{\nu_0, \nu - \nu_1\},$$

$$\rho(y) = \delta(y - 1). \quad (233)$$

The simulations are performed by using an adaptive time discretization scheme (see Appendix I for the detailed numerical scheme). All panels in Fig. 10 exhibit the predicted power-law tail of the intensity distribution, in excellent agreement with our theoretical prediction (198). Notably, the power-law exponent varies continuously as a function of ν_1 and the power-law

formula (198) is found to be valid even for integer exponents such as $a = 0$ and 1.

VIII. SOLUTION 5: GENERAL MEMORY KERNEL FOR FAST-ACCELERATING INTENSITY MAPS AND TWO-SIDED MARK DISTRIBUTION WITH NONPOSITIVE MEAN MARK

In Secs. V and VI we showed that a general asymptotic formula is available for the exponential memory and the two-sided mark distributions with nonpositive mean. Here we show that, by solving the corresponding MEs, the asymptotic formula is valid for a wider class of memory kernels with FAI. This analysis is a detailed version of Ref. [27].

A. Discrete sum of exponentials

We first show that the power-law tail of the PDF of intensities is robust for various memory kernels $h(t)$ for the MSA intensity function in the presence of a two-sided mark distribution with nonpositive-mean mark. Specifically, we make the assumptions, i.e., discrete sum of exponentials, MSA intensity, and two-sided mark distribution with nonpositive-mean mark,

$$h(t) = \sum_{k=1}^K \tilde{h}_k e^{-t/\tau_k}, \quad g(\nu) = \lambda_0 e^{\beta \nu},$$

$$p_+ := \int_0^\infty \rho(y)dy > 0,$$

$$p_- := \int_{-\infty}^0 \rho(y)dy > 0, \quad m := \int_{-\infty}^\infty y\rho(y)dy \leqslant 0. \quad (234)$$

Under these conditions, we obtain the power-law intensity PDF

$$P_{\text{ss}}(\lambda) \propto \lambda^{-2-\beta^{-1}u} \text{ for large } \lambda, \quad u := \frac{c^*}{h(0)}, \quad (235)$$

where c^* is the positive root of $\Phi(x)$ for $m < 0$ (or $c^* = 0$ for $m = 0$). Remarkably, we recover Zipf's law exactly for the zero-mean mark case $m = 0$.

1. Derivation

From Eq. (23) the steady-state ME is given by

$$
\sum_{k=1}^{K} \frac{1}{\tau_k \lambda_0} \frac{\partial}{\partial z_k} \left[z_k \exp\left(-\sum_{k'=1}^{K} z_{k'} \right) \phi(z) \right]
$$

$$
+ \int dy\, \rho(y) \phi(z - y\tilde{\boldsymbol{h}}) - \phi(z) = 0, \qquad (236)
$$

where we have defined $\phi(z) := G(z) P_{\mathrm{ss}}(z)$. For large z, the first term on the right-hand side is negligible due to the exponential factor, leading to

$$
\int dy\, \rho(y) \phi(z - y\tilde{\boldsymbol{h}}) - \phi(z) \simeq 0 \quad \text{for large } z. \qquad (237)
$$

We then apply the transformation from $z = (z_1, \ldots, z_K)$ to $\boldsymbol{Z} := (W, Z_2, \ldots, Z_K)$:

$$
z_1 = \tilde{h}_1 W,
$$
$$
z_2 = \tilde{h}_2 W + Z_2,
$$
$$
z_3 = \tilde{h}_3 W + Z_3,
$$
$$
\vdots
$$
$$
z_K = \tilde{h}_K W + Z_K. \qquad (238)
$$

Using this variable set, we can rewrite

$$
\psi(W - y; Z_2, \ldots, Z_K)
$$
$$
:= \phi(z - y\tilde{\boldsymbol{h}})
$$
$$
= \phi(\tilde{h}_1(W - y), \tilde{h}_2(W - y) + Z_2, \ldots, \tilde{h}_K(W - y) + Z_K). \qquad (239)
$$

The integral equation (237) is then reduced to

$$
\int dy\, \rho(y) \psi(W - y; \boldsymbol{Z}') - \psi(W; \boldsymbol{Z}') \simeq 0 \quad \text{for large } \boldsymbol{Z}, \qquad (240)
$$

with $\boldsymbol{Z}' := (Z_2, \ldots, Z_K)$. This variable transformation is useful because Eq. (237) is an effectively one-dimensional integral equation. Since the variable subset $\boldsymbol{Z}'$ is irrelevant in this integral equation, its solution is given by

$$
\psi(W; \boldsymbol{Z}') = C_0(\boldsymbol{Z}') e^{-c^* W} + C_1(\boldsymbol{Z}'), \qquad (241)
$$

with arbitrary non-negative functions $C_0(\boldsymbol{Z}')$ and $C_1(\boldsymbol{Z}')$ without the variable W (see Appendix E). In addition, by defining the moment-generating function $\Phi(x) := \int_{-\infty}^{\infty} dy\, \rho(y)(e^{xy} - 1)$, the constant c^* is the positive root of $\Phi(c^*) = 0$ for the case of negative-mean mark $m < 0$ or $c^* = 0$ for the case of zero-mean mark $m = 0$ [see Appendix F for the detailed properties of $\Phi(x)$]. Assuming the natural boundary condition, $C_1(\boldsymbol{Z}')$ must be set equal to zero, as shown later. We then derive the steady-state distribution $P_{\mathrm{ss}}(v)$ as

$$
P_{\mathrm{ss}}(v) := \int_{-\infty}^{\infty} dz\, P_{\mathrm{ss}}(z) \delta\left(v - \sum_{k=1}^{K} z_k \right) \simeq \frac{1}{\lambda_0} \int_{-\infty}^{\infty} dz \exp\left(-\beta \sum_{k=1}^{K} z_k \right) C_0(\boldsymbol{Z}') e^{-c^* W} \delta\left(v - \sum_{k=1}^{K} z_k \right)
$$

$$
= \frac{e^{-\beta v}}{\lambda_0} \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} \left(\prod_{j=2}^{K} dz_j \right) C_0\left(z_2 - \frac{\tilde{h}_2}{\tilde{h}_1} z_1, \ldots, z_K - \frac{\tilde{h}_K}{\tilde{h}_1} z_1 \right) \exp\left(-\frac{c^*}{\tilde{h}_1} z_1 \right) \delta\left(v - \sum_{k=1}^{K} z_k \right). \qquad (242)
$$

Applying the transformation

$$
z_j' := z_j - \frac{\tilde{h}_j}{\tilde{h}_1} z_1 \quad \text{for } j = 2, \ldots, K, \qquad (243)
$$

we obtain

$$
P_{\mathrm{ss}}(v) \simeq \frac{e^{-\beta v}}{\lambda_0} \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} \left(\prod_{j=2}^{K} dz_j' \right) C_0(z_2', \ldots, z_K') \exp\left(-\frac{c^*}{\tilde{h}_1} z_1 \right) \delta\left(v - r z_1 - \sum_{k=2}^{K} z_k' \right)
$$

$$
= \frac{e^{-\beta v}}{\lambda_0} \int_{-\infty}^{\infty} \left(\prod_{j=2}^{K} dz_j' \right) C_0(z_2', \ldots, z_K') \int_{-\infty}^{\infty} dz_1 \exp\left(-\frac{c^*}{\tilde{h}_1} z_1 \right) \delta\left(v - r z_1 - \sum_{k=2}^{K} z_k' \right)
$$

$$
= \frac{e^{-(\beta - u)v}}{r\lambda_0} \int_{-\infty}^{\infty} \left(\prod_{j=2}^{K} dz_j' \right) C_0(z_2', \ldots, z_K') \exp\left(\frac{c^*}{h(0)} \sum_{k=2}^{K} z_k' \right), \qquad (244)
$$

where we have used

$$
\delta\left(v - r z_1 - \sum_{k=2}^{K} z_k' \right) = \frac{1}{r} \delta\left(z_1 - \frac{v - \sum_{k=2}^{K} z_k'}{r} \right), \qquad (245)
$$

with

$$r := \frac{1}{\tilde{h}_1} \sum_{k=1}^{K} \tilde{h}_k = \frac{h(0)}{\tilde{h}_1}, \quad u := \frac{c^*}{h(0)}. \qquad (246)$$

Assuming that

$$\frac{1}{r} \int_{-\infty}^{\infty} \left(\prod_{j=2}^{K} dz_j' \right) C_0(z_2', \ldots, z_K') \exp \left(\frac{c^*}{h(0)} \sum_{k=2}^{K} z_k' \right) < \infty, \qquad (247)$$

we find that the asymptotic PDF for large ν is given by

$$P_{\mathrm{ss}}(\nu) \propto e^{-(\beta-u)\nu} \quad \text{for large } \nu. \qquad (248)$$

This asymptotic form implies the power law (235) for the intensity $\lambda := g(\nu)$.

2. Natural boundary condition

By neglecting the first term in the ME (23) and by applying the variable transformation (238), we obtain an approximate ME

$$\frac{\partial P_t(W; \mathbf{Z}')}{\partial t} \simeq \int_{-\infty}^{\infty} dy [G(W-y; \mathbf{Z}') P_t(W-y; \mathbf{Z}') - G(W; \mathbf{Z}') P_t(W; \mathbf{Z}')]. \qquad (249)$$

By applying the KM expansion, we obtain the conservation of probability

$$\frac{\partial P_t(W; \mathbf{Z}')}{\partial t} \simeq -\frac{\partial J_t(W; \mathbf{Z}')}{\partial W}, \qquad (250)$$

with the probability current

$$J_t(W; \mathbf{Z}') := \sum_{n=1}^{\infty} \frac{(-1)^{n-1} \alpha_n}{n!} \frac{\partial^{n-1}}{\partial W^{n-1}} G(W; \mathbf{Z}') P_t(W; \mathbf{Z}'). \qquad (251)$$

By substituting the solution (241), we obtain

$$J_{\mathrm{ss}}(W; \mathbf{Z}') = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} \alpha_n}{n!} \frac{\partial^{n-1}}{\partial W^{n-1}} [C_1(\mathbf{Z}') + e^{-c^*W} C_0(\mathbf{Z}')]$$

$$= m C_1(\mathbf{Z}') + \frac{C_0(\mathbf{Z}')}{c^*} \Phi(c^*) e^{-c^*W}, \qquad (252)$$

where we have used $\Phi(x) = \sum_{n=1}^{\infty} (\alpha_n/n!) x^n$. Since $\Phi(c^*) = 0$ by definition, we obtain $J_{\mathrm{ss}}(W; \mathbf{Z}') = m C_1(\mathbf{Z}')$. The natural boundary condition requires

$$\lim_{W \to \infty} J_{\mathrm{ss}}(W; \mathbf{Z}') = 0 \qquad (253)$$

for any $\mathbf{Z}'$, implying $C_1(\mathbf{Z}') = 0$.

B. General memory kernel

As done before, any memory kernel can be approximated by a sum of exponentials such that

$$h(t) = \int_0^{\infty} dx\, \tilde{h}(x) e^{-t/x} \approx \sum_{k=1}^{K} \tilde{h}_k e^{-t/\tau_k}. \qquad (254)$$

Since the power-law tail for the PDF of the intensity is found for any discrete sum of exponentials, it remains valid for general superpositions of exponential memory. Under the assumption

$$g(\nu) = \lambda_0 e^{\beta\nu}, \quad p_+ := \int_0^{\infty} \rho(y) dy > 0,$$

$$p_- := \int_{-\infty}^{0} \rho(y) dy > 0, \quad m := \int_{-\infty}^{\infty} y\rho(y) dy \leqslant 0\, \forall h(t), \qquad (255)$$

we obtain

$$P_{\mathrm{ss}}(\lambda) \propto \lambda^{-2-\beta^{-1}u} \quad \text{for large } \lambda, \quad u := \frac{c^*}{h(0)}, \qquad (256)$$

where c^* is the positive root of $\Phi(c^*) = 0$ for $m < 0$ or $c^* = 0$ for $m = 0$. For the zero-mean mark case $m = 0$, the PDF obeys Zipf's law exactly.

C. Numerical confirmation

Figure 11 shows numerical results obtained by Monte Carlo simulations of the SDE (3) for the memory function made of a discrete sum of exponential functions, for the exponential intensity MSA with finite cutoff (to ensure convergence of the numerical scheme), and a zero-mean Gaussian mark distribution

$$h(t) = \sum_{k=1}^{K} \tilde{h}_k e^{-t/\tau_k}, \quad \lambda = g(\nu) = \min\{\lambda_0 e^{\beta\nu}, \lambda_{\max}\},$$

$$\rho(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/2\sigma^2}. \qquad (257)$$

We use an adaptive time discretization scheme (see Appendix I for the detailed numerical scheme). Here $\lambda_{\max}$ is a cutoff parameter to control numerical rounding errors. Figure 11(a) exhibits the Zipf law in the intensity distribution up to the cutoff $\lambda_{\max}$, showing agreement with our theoretical prediction (235).

Figure 11(b) shows a sample trajectory of the intensity obtained for the parameter set $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 0.5, 2)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.6, 0.1)$, $\lambda_0 = 0.001$, $\beta = 10$, $\lambda_{\max} = 10^6$, and $\sigma = 0.3$. This semilogarithmic plot illustrates that the NLH model exhibits an intermittent behavior in terms of its intensity, which is qualitatively consistent with observed phenomena in various complex systems, such as seismic activity.

D. Generalization to fast-accelerating intensity maps

The above framework can be readily generalized to FAI maps defined by $g(\nu) \gg \nu^2$. Our general result can be formulated as follows. Under the assumptions

$$h(t) = \sum_{k=1}^{K} \tilde{h}_k e^{-t/\tau_k}, \quad g(\nu) \gg \nu^2 \quad \text{for large } \nu,$$

$$p_+ := \int_0^{\infty} \rho(y) dy > 0, \quad p_- := \int_{-\infty}^{0} \rho(y) dy > 0,$$

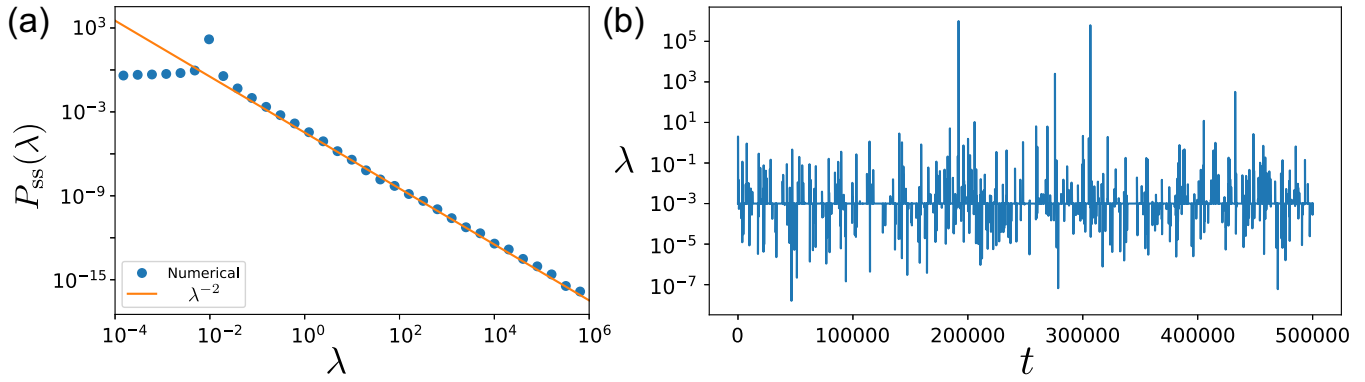$$m := \int_{-\infty}^{\infty} y\rho(y) dy \leqslant 0, \qquad (258a)$$

FIG. 11. (a) Numerical confirmation of our theoretical prediction (235) of the Zipf law for the intensity distribution. This simulation corresponds to the following set of parameters: $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 0.5, 2)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.6, 0.1)$, $\lambda_0 = 0.01$, $\beta = 6$, $\lambda_{\max} = 10^6$, and $\sigma = 0.3$. The predicted power-law exponent is given by $a = 1.0$, i.e., a true power law (the Zipf law) and is shown as the orange straight line. (b) Sample trajectory of the intensity for $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 0.5, 2)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.6, 0.1)$, $\lambda_0 = 0.001$, $\beta = 10$, $\lambda_{\max} = 10^6$, and $\sigma = 0.3$. Note that the horizontal reference line is determined by $\lambda_0 = 0.001$.

we obtain

$$P_{ss}(\lambda) \propto \lambda^{-1} \left[e^{-uv} \left(\frac{dg(v)}{dv} \right)^{-1} \right]_{v=g^{-1}(\lambda)}$$

$$\text{for large } \lambda, \quad u := \frac{c^*}{h(0)}, \tag{258b}$$

with c^* the positive root of $\Phi(c^*) = 0$ for $m < 0$ or zero for $m = 0$, where $\Phi(x) := \int_{-\infty}^{\infty} dy\, \rho(y)(e^{xy} - 1)$. The derivation of this result is essentially the same as that in Sec. VIII A, by replacing the MSA intensity map with the general FAI map.

Since any memory kernel can be approximated by a discrete sum of exponentials [see Eq. (254)], the continuous version of the statement (258) also holds for any $h(t)$,

$$g(v) \gg v^2, \quad p_+ > 0, \quad p_- > 0, \quad m \leqslant 0$$

$$\implies P_{ss}(\lambda) \propto \lambda^{-1} \left[e^{-uv} \left(\frac{dg(v)}{dv} \right)^{-1} \right]_{v=g^{-1}(\lambda)},$$

$$u := \frac{c^*}{h(0)}, \tag{259}$$

with c^* the positive root of $\Phi(c^*) = 0$ for $m < 0$ or zero for $m = 0$, where $\Phi(x) := \int_{-\infty}^{\infty} dy\, \rho(y)(e^{xy} - 1)$. Remarkably, this result implies that the power-law tail for the steady-state PDF of intensities holds robustly for superpolynomial intensity maps such as $g(v) \propto e^{\beta v}$, i.e., the MSA case, and $g(v) \propto e^{\beta v^2}$.

IX. DISCUSSION

A. Relationship to nonlinear Kesten processes

We have shown that power-law asymptotics robustly appears for the quadratic and FAI cases. Here we provide another derivation based on more heuristic arguments, by removing inessential technicalities, using the viewpoint of Kesten processes [51]. Let us focus on the case with exponential memory $h(t) = (\eta/\tau)e^{-t/\tau}$ and in the diffusive limit, i.e., for the symmetric mark distribution $\rho(y) = \rho(-y)$, described by the FPE (46) and the corresponding SDE (47).

1. Case with a quadratic intensity map: $\tilde{g}(v) = kv^2 + v_0$

Inserting $\tilde{g}(v) = kv^2 + v_0$ in the SDE (47), we obtain

$$\frac{d\hat{v}}{dt} = -\frac{\hat{v}}{\tau} + \sqrt{kv^2 + v_0}\sqrt{2D} \cdot \hat{\xi}_t^G. \tag{260}$$

Here $\hat{\xi}_t^G$ is the standard white Gaussian noise satisfying $\langle \hat{\xi}^G \rangle = 0$ and $\langle \hat{\xi}^G(t)\hat{\xi}^G(t') \rangle = \delta(t - t')$. With respect to the power-law structure of the tail of the PDF of v, by using the discretization $dv/dt = [v(t + dt) - v(t)]/dt$, this SDE can be regarded as a continuous version of the discrete-time Kesten process [51]

$$v(t + dt) \simeq a_t v(t) + b_t, \tag{261}$$

with $a_t = 1 + (2Dk)^{1/2}\hat{\xi}_t^G dt - (dt/\tau)$ and $b_t = v_0(2k)^{1/2}\hat{\xi}_t^G dt$. The first term $a_t v(t)$ controls the intermittent excursions of $v(t)$ to large values, for which v_0 can be neglected in the last term on the right-hand side of Eq. (260). The second term b_t in the Kesten map (261) is the reinjecting term, obtained when v becomes smaller than v_0. As shown in Ref. [53], the detailed shape of this reinjecting term has no impact on the existence of a power-law tail and on the value of its exponent. The only important point is that the reinjecting term exists to prevent v from being too small. Remaining no less than a stochastic variable proportional to v_0, intermittent runs of exponential growth occur when there is a succession of positive realizations of $\hat{\xi}_t^G$ for several consecutive times such that the multiplicative factor a_t is larger than 1 over this run [52,53].

The condition for the existence of a steady-state PDF for the Kesten process (261) is that $\langle \ln a_t \rangle < 0$ [51,53]. For infinitesimal dt, $\ln a_t$ can be expanded as $\ln a_t = (2Dk)^{1/2}\hat{\xi}_t^G dt - (dt/\tau)$ and its mean is then $\langle \ln a_t \rangle = -(dt/\tau)$ since $\langle \hat{\xi}_t^G \rangle = 0$ by definition. Hence, the condition for a stationarity process holds true. It is then easy to show by explicitly writing the self-consistent equation for the steady-state PDF of v that it is a power law with exponent a given as the solution of the equation

$$\langle |a_t|^a \rangle = 1. \tag{262}$$

Using the fact that $\hat{\xi}_t^G dt = dW$ is a Gaussian random variable with zero mean and variance dt, i.e., it is the infinitesimal increment of the Wiener process, the average in (262) is obtained by using the saddle-node approximation and we find that the corresponding solution recovers exactly the expression (98), namely, $a := 1/2 + c^2/2k\tau$, with $c := \tau/\eta$. This confirms that our treatment in terms of the diffusive limit gives equations in the general class of Kesten processes. Other forms of the tension-intensity map $\lambda = g(\nu)$ can thus be interpreted as continuous nonlinear extensions of the Kesten process. Therefore, from an intuitive point of view, the power-law tails of the PDFs of ν and λ can be qualitatively related to an underlying multiplicative structure together with additional ingredients to ensure the existence and stationarity of the process.

2. Case with a fast-accelerating intensity map: $\tilde{g}(\nu) \gg \nu^2$

We here consider the case of FAI maps satisfying

$$\tilde{g}(\nu) \gg \nu^2 \quad \text{for positive large } \nu,$$
$$\tilde{g}(\nu) \simeq \text{const} \quad \text{for negative large } \nu. \quad (263)$$

Let us take $D = 1/2$ to simplify notation so that the corresponding SDE is given by

$$\frac{d\hat{\nu}}{dt} = -\frac{\hat{\nu}}{\tau} + \sqrt{\tilde{g}(\hat{\nu})} \cdot \hat{\xi}^G. \quad (264)$$

Since $\tilde{g}(\nu) \simeq \text{const}$ for negative large ν, $\hat{\nu}$ cannot go to $-\infty$ due to the relaxation term $-\hat{\nu}/\tau$. On the other hand, for positive large $\hat{\nu}$, the dynamics is approximated by

$$\frac{d\hat{\nu}}{dt} \simeq \sqrt{\tilde{g}(\hat{\nu})} \cdot \hat{\xi}^G, \quad (265)$$

because $\sqrt{\tilde{g}(\hat{\nu})} \gg \hat{\nu}/\tau$ for FAI maps.

This model is thus similar to Brownian motion with a position-dependent variance or diffusion coefficient. Interestingly, such a Brownian model has a well-defined steady-state PDF for FAI maps. Let us thus consider a Brownian motion obeying the SDE

$$\frac{d\hat{\nu}}{dt} = \sqrt{\tilde{g}(\hat{\nu})} \cdot \hat{\xi}^G \quad \text{for } \nu \geqslant 0, \quad (266)$$

which is complemented by the condition of a repulsive hard wall at $\hat{\nu} = 0$, which prohibits the Brownian particle from going to $-\infty$. This condition is a simplification and ensures a similar result to that in the presence of the relaxation effect in the original model (264). The steady FPE is given by

$$\frac{d^2}{d\nu^2}[\tilde{g}(\nu)P_{ss}(\nu)] = 0 \quad \text{for } \nu > 0, \quad P_{ss}(\nu) = 0 \quad \text{for } \nu < 0. \quad (267)$$

If $\tilde{g}(\nu)$ were not a FAI map, this steady FPE might not have a normalizable steady-state solution. For example, if $\tilde{g}(\nu) = \text{const}$, the general solution of the FPE is given by a non-normalizable steady-state solution $\tilde{P}_{ss}(\nu) = c_0 + c_1\nu$, satisfying $\int_0^\infty d\nu\, P_{ss}(\nu) = \infty$. This model is therefore nonstationary. In contrast, when $\tilde{g}(\nu)$ is a FAI map, the FPE (267) has a normalizable steady-state solution for any FAI map, with

$$P_{ss}(\nu) \propto \begin{cases} \frac{1}{\tilde{g}(\nu)} & \text{for } \nu > 0 \\ 0 & \text{for } \nu < 0 \end{cases} \implies \int_{-\infty}^\infty d\nu\, P_{ss}(\nu) < \infty. \quad (268)$$
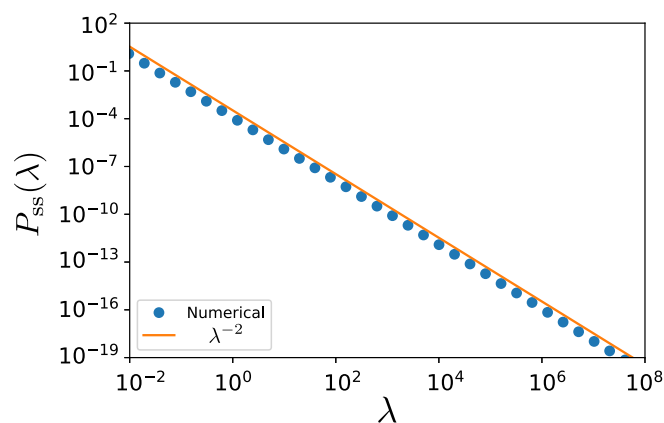

FIG. 12. Numerical simulations confirming our theoretical prediction on the Zipf law (270) for the intensity distribution. The simulated SDE describes Brownian motion with the $\hat{\nu}$-dependent diffusion constant $\tilde{g}(\hat{\nu})/2$ and a reflecting barrier at $\hat{\nu} = 0$.

We then obtain the robust expression of the steady-state PDF of the intensity $\lambda := \tilde{g}(\nu)$:

$$P_{ss}(\lambda) \propto \lambda^{-1} \left| \frac{d\tilde{g}(\nu)}{d\nu} \right|^{-1}_{\nu = \tilde{g}^{-1}(\lambda)} \quad \text{for large } \lambda. \quad (269)$$

This result readily implies that the Zipf law

$$P_{ss}(\lambda) \propto \lambda^{-2} \quad \text{for large } \lambda \quad (270)$$

is observed for a wide class of superpolynomial intensity maps, such as $\tilde{g}(\nu) = e^{\beta\nu}$ and $\tilde{g}(\nu) = e^{\beta\nu^2}$.

Let us complete this discussion by mentioning that the rigorous mathematical demonstration of the existence of steady-state solutions of the SDE (266) is obtained from the theorems presented in Ref. [55]. In particular, we refer to the theorems in Sec. 5.2 in Ref. [55].

a. Numerical simulation. Figure 12 presents the PDF of the intensity obtained from the numerical solution of the SDE (266) with an exponential intensity $\tilde{g}(\nu) = \lambda_0 e^{\beta\nu}$. The Zipf law (270) is obtained for the steady-state intensity distribution of the SDE describing Brownian motion with $\hat{\nu}$-dependent diffusion constant $\tilde{g}(\nu)/2$ in the presence of a reflecting barrier at $\hat{\nu} = 0$. The detailed numerical implementation is described in Appendix I.

b. Intuitive discussion. Why is the Brownian model with position-dependent variance stationary for FAI maps? This might be understandable from the viewpoint of step-size explosion for large $\hat{\nu}$. As $\hat{\nu}$ grows to very large values, the steps of the random walk explode even faster and thus it is very likely that a negative step occurs of huge size which brings back the Brownian particle to 0 or even pushes it to negative values if the repelling boundary is absent. In the presence of the repelling boundary, the huge negative steps bring the Brownian particle close to 0, for which the random step sizes become small, which implies that $\hat{\nu}$ remains quite a long time in the vicinity of the origin. Eventually, $\hat{\nu}$ escapes again to large values of $\hat{\nu}$, but then the huge random walk step sizes, when negative, bring it back again to the neighborhood of 0. This process occurs repeatedly and leads to a stationary PDF decaying rapidly as $1/\tilde{g}(\hat{\nu})$, due to the effect of the negative

random steps that push back $\hat{v}$ to the left and the boundary somehow traps the process in its neighborhood. In summary, this picture implies that the explosive step size leads to an effective strong trapping potential, which might be counterintuitive at the first glance.

It might be interesting to rephrase the above argument from the viewpoint of the recurrence time of one-dimensional Brownian motion. It is well known that the recurrence probability of one-dimensional Brownian motion with constant variance $\tilde{g}(v) = \text{const}$ is unity, while the average recurrence time is infinity. In other words, a Brownian particle will surely return to the origin after a long time, but this waiting time may be so large that repeated recurrence events cannot be expected within a finite observation time. In the case of Eq. (266), the variance of the step lengths depends on the distance from the origin and becomes larger and larger for large v. Thus, the time evolution of this Brownian particle becomes faster and faster for large $\hat{v}$ such that the particle can return to the origin much sooner than with a constant diffusion coefficient and repeated recurrence events can be expected in a finite time.

Finally, it is useful to stress the difference between the mechanism underlying the existence of a steady-state power-law distribution in the conventional linear Kesten process [51] and that of our FAI Hawkes model. In usual linear Kesten processes, the underlying mechanism is proportional growth or multiplicative proportional growth in order for Zipf's law and related power laws to occur, in the presence of an average contraction, i.e., the mean growth rate is negative, together with a reinjection mechanism. In contrast, our FAI Hawkes models are characterized by explosive expansions in the presence of a reflecting or bounded boundary condition. This is in stark contrast to the conventional approaches based on the Kesten-type models, proportional growth type, and preferential attachment-type mechanisms.

B. Implication to financial data analyses

From a broader perspective, our results have significant implications for financial modeling. Recall that one of the motivations for introducing the NLH family is to reproduce empirical stylized facts, in particular the power-law distribution of returns (see Sec. II C 2 for a brief review). Indeed, according to Ref. [28], one of the advantages of the quadratic Hawkes process lies in the fact that it can reproduce a power-law intensity distribution with nonuniversal exponents [see Eq. (12)], from which the power-law distribution of price changes derives.

From this point of view, our results summarized in Tables I and II show that various NLH processes can reproduce power-law intensity PDFs and not just the quadratic Hawkes processes. Even the ramp Hawkes processes with one-sided marks (which are arguably a minor modification of the LH process) can reproduce a power-law distribution with arbitrary exponent, when near criticality. If one focuses only on power-law intensity PDFs, various Hawkes models can be suitable candidate models. Moreover, by assuming symmetric marks, note that quadratic Hawkes is the model at the boundary between a class of models with universal power-law exponents and a class of models with nonuniversal exponents.

Another debatable point is whether the exponent should be universal or nonuniversal. The linear and ramp Hawkes processes with one-sided marks, the symmetric-mark quadratic Hawkes processes, and the two-sided FAI Hawkes process with nonpositive-mean mark have nonuniversal exponents for the power-law intensity PDF. In contrast, the symmetric-mark FAI Hawkes processes have universal exponents. The nonuniversal exponents are useful for flexible data calibration, while one would like to have strong justifications of why such parameters are selected in empirical data analyses. On the other hand, universal exponents are useful if the empirical exponent seems robust and universal, but this removes flexible data calibration. It might be necessary to construct a suitable framework for model selection with the goal of developing practical reverse engineering approaches. This is beyond the scope of the present paper.

C. Mathematical relation to quantum field theories

We have studied a wide variety of generalized Hawkes processes. Here we discuss their mathematical relation with quantum field theories. Let us rewrite $z(x) \to \phi(x)$ and introduce a momentum operator $\pi(x)$,

$$\pi(x) := -i\frac{\delta}{\delta\phi(x)}, \tag{271}$$

which satisfies the canonical commutative relation

$$[\phi(x), \pi(x')] = i\delta(x - x'). \tag{272}$$

By introducing the state vector

$$|P_t\rangle := \int \mathcal{D}\phi \, P_t[\phi]|\phi\rangle, \tag{273}$$

the ME (17) becomes a Schrödinger-like equation for the field $\{\phi(x)\}_x$ as

$$\frac{\partial}{\partial t}|P_t\rangle = H|P_t\rangle,$$

$$H := i\int_0^\infty \frac{dx}{x}\pi(x)\phi(x)$$

$$+ \int_{-\infty}^\infty dy\, \rho(y)(T[y\tilde{h}] - 1)G[\phi], \tag{274}$$

with the non-Hermitian Hamiltonian H and the translation operator $T[y]$, defined by

$$T[y] := \exp\left(-i\int_0^\infty dx\, y(x)\pi(x)\right), \tag{275}$$

satisfying $T[y]P_t[z] = P_t[z - y]$. This equation is nonlocal, since the Hamiltonian includes an infinite order of momentum operators.

In the diffusive limit (48), the Hamiltonian reduces to a local version as a result of the SSE

$$\frac{\partial}{\partial t}|P_t\rangle = H|P_t\rangle,$$

$$H := i \int_0^\infty \frac{dx}{x} \pi(x)\phi(x)$$

$$- \int_0^\infty dx \int_0^\infty dx' D(x, x')\pi(x)\pi(x')G[\phi],$$

$$D(x, x') := \frac{\alpha_2}{2}\tilde{h}(x)\tilde{h}(x'), \tag{276}$$

where the momentum operator appears via a quadratic form. In this sense, the SSE for the field ME can be regarded as a mathematical procedure to obtain local forms of non-Hermitian field quantum theories in an appropriate limit.

D. Future application: Field master equation for the general quadratic Hawkes processes

Our present formulation covers a part of the Zumbach Hawkes processes, but not the whole class of general quadratic Hawkes processes (10). Including the general quadratic Hawkes processes as special cases of our formalism can in principle be easily performed as follows. The linear and quadratic kernels $L(t - s)$ and $K(t - s, t - u)$ can be decomposed according to a Laplace representation

$$L(t - s) = \int_0^\infty dx\, e^{-(t-s)/x}\tilde{L}(x), \quad K(t - s, t - u)$$

$$= \int_0^\infty dx \int_0^\infty dx'\, e^{-(t-s)/x-(t-u)/x'}\tilde{K}(x, x'). \tag{277}$$

The quadratic Hawkes process then reads

$$\hat{\lambda}(t) = \lambda_0 + \int_{-\infty}^t ds\, \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s)\left(\int_0^\infty dx\, e^{-(t-s)/x}\tilde{L}(x)\right)$$

$$+ \int_{-\infty}^t ds\, \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s) \int_{-\infty}^t du\, \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(u)}(u)\left(\int_0^\infty dx \int_0^\infty dx'\, e^{-(t-u)/x'-(t-s)/x}\tilde{K}(x, x')\right)$$

$$= \lambda_0 + \int_0^\infty dx\, \tilde{L}(x)\left(\int_{-\infty}^t ds\, e^{-(t-s)/x}\hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s)\right)$$

$$+ \int_0^\infty dx \int_0^\infty dx'\, \tilde{K}(x, x')\left(\int_{-\infty}^t ds\, e^{-(t-s)/x}\hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s)\right)\left(\int_{-\infty}^t du\, e^{-(t-u)/x'}\hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(u)}(u)\right), \tag{278}$$

where we have exchanged the order of integration. By considering the Markovian SPDE

$$\frac{\partial \hat{z}(t, x)}{\partial t} = -\frac{\hat{z}(t, x)}{x} + \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(t)}(t), \tag{279a}$$

whose explicit solution is given by

$$\hat{z}(t, x) = \int_{-\infty}^t ds\, e^{-(t-s)/x}\hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s)}(s), \tag{279b}$$

we obtain

$$\hat{\lambda}(t) = \lambda_0 + \int_0^\infty dx\, \tilde{L}(x)\hat{z}(t, x)$$

$$+ \int_0^\infty dx \int_0^\infty dx'\, \tilde{K}(x, x')\hat{z}(t, x)\hat{z}(t, x'), \tag{279c}$$

which is equivalent to the original quadratic Hawkes processes (10). This means that the quadratic Hawkes dynamics has been converted into a Markovian dynamics described by the set of equations (279) in terms of the field variable $\{\hat{z}(t, x)\}_x$. Correspondingly, we obtain the field ME for the quadratic Hawkes processes as

$$\frac{\partial P_t[z]}{\partial t} = \int_0^\infty dx \frac{\delta}{\delta z(x)}\left(\frac{z(x)}{x}P_t[z]\right)$$

$$+ \int_{-\infty}^\infty dy\, \rho(y)(G[z - y\mathbf{1}]P_t[z - y\mathbf{1}] - G[y]P_t[z]), \tag{280}$$

with the indicator function $\mathbf{1}(x) = 1$ for any $x \in (0, \infty)$ and the functional intensity map $G[z]$ defined by

$$G[z] := \lambda_0 + \int_0^\infty dx\, \tilde{L}(x)z(x)$$

$$+ \int_0^\infty dx \int_0^\infty dx'\, \tilde{K}(x, x')z(x)z(x'). \tag{281}$$

Further formal generalization

Obviously, this method can be readily generalized for any functional series expansion, at least formally, such as

$$\hat{\lambda}(t) = \lambda_0 + \sum_{j=1}^J \int_{-\infty}^t ds_1 \cdots \int_{-\infty}^t ds_j K_j(t - s_1, \ldots, t - s_j)$$

$$\times \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s_1)} \cdots \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(s_j)}$$

$$= \lambda_0 + \sum_{j=1}^J \int_0^\infty ds_1 \cdots \int_0^\infty ds_j \tilde{K}_j(x_1, \ldots, x_j)$$

$$\times \hat{z}(t, x_1) \cdots \hat{z}(t, x_j). \tag{282}$$

Here we have introduced the field variable $\{\hat{z}(t, x)\}$ obeying the Markovian SPDE

$$\frac{\partial \hat{z}(t, x)}{\partial t} = -\frac{\hat{z}(t, x)}{x} + \hat{\xi}^{\mathrm{P}}_{\rho(y);\hat{\lambda}(t)} \tag{283}$$

and the Laplace decomposition

$$K_j(t - s_1, \ldots, t - s_j)$$

$$= \int_0^\infty dx_1 \cdots \int_0^\infty dx_j \exp\left(-\sum_{l=1}^{j} \frac{t - s_l}{x_l}\right) \tilde{K}_j(x_1, \ldots, x_j).$$
$$\tag{284}$$

The corresponding field ME can be derived in the same manner. This implies that our formulation has the potential to cover a wide variety of NLH families beyond the quadratic Hawkes processes. We leave to future studies the derivation of explicit analytical solutions for general quadratic Hawkes processes and beyond, based on our formulation.

X. CONCLUSION

In this article we have studied various analytical solutions to NLH processes by generalizing the field ME approach recently developed in Refs. [16,17]. We have derived the field ME for the general NLH processes and have formulated its functional KM expansion and the corresponding diffusive approximation. We then proceeded with deriving various exact solutions of the steady-state intensity distributions for an exponential-memory kernel in the absence and presence of inhibitory effects. Some of the robust asymptotic solutions have been generalized for a wide class of memory kernels, such as (i) the nonuniversal power law with an arbitrary exponent for the ramp Hawkes process in the absence of inhibitory effects, (ii) the robust Zipf law for the superexponential intensity family in the presence of symmetric inhibitory and excitatory effects, and (iii) the ubiquitous power law for the fast-acceleration intensity Hawkes models in the regime of zero- or negative-mean mark.

The summaries in Tables I and II exemplify our systematic analysis of the NLH processes. However, there are two items that are not treated in the last column of Table I for general memory kernels $h(t)$. This is because our focus has been mainly on FAI Hawkes processes and the ramp Hawkes process with nonpositive-mean mark and the quadratic Hawkes processes with symmetric mark are not FAI Hawkes processes. It is likely that different perturbative solutions are needed to solve these two cases for general memory kernels $h(t)$, which we leave for future work.

While only a few analytical solution for limited cases have been derived in the past for NLH processes due to their nonlinear and non-Markovian nature, we have significantly extended the set of solutions, obtaining exact and robust asymptotic expressions with the help of our formulation in terms of a field ME. This demonstrates the power of this approach in addressing non-Markovian stochastic processes. It would be interesting to generalize this framework for more general non-Markovian stochastic processes, such as non-Markovian point processes that have arbitrary intensities depending on the full past history. In addition, our results imply that the NLH family can accommodate various power-law relations in the intensity distribution, which could be useful for data calibration in various complex systems.

APPENDIX A: FORMAL PROPERTIES OF THE DIRAC δ FUNCTION

1. Dirac δ function

The Dirac δ function is formally defined by the relationships for real numbers x and y,

$$\delta(x - y) = \begin{cases} 0, & x \neq y \\ \infty, & x = y, \end{cases}$$

$$\int_{-\infty}^{\infty} f(x)\delta(x - y)dx = f(y),\tag{A1}$$

which is the continuous analog to the Kronecker δ, defined by

$$\delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j, \end{cases} \quad \sum_i f_i \delta_{ij} dx = f_i \tag{A2}$$

for integers i and j.

There are several formal methods to construct the Dirac δ function. In this paper we construct the Dirac δ function via a formal continuous limit from the discrete picture. Let us consider the lattice coordinate $x_i := i dx$ for an integer i with the lattice interval dx. The Dirac δ function can be formally introduced by

$$\delta(x_i - x_j) = \lim_{dx\downarrow 0} \frac{1}{dx}\delta_{i,j},\tag{A3}$$

which satisfies

$$\int_{-\infty}^{\infty} f(x)\delta(x - y)dx$$

$$:= \lim_{dx\downarrow 0} \sum_i f(x_i)\left(\frac{1}{dx}\delta_{ij}\right)dx = f(y) \quad \text{for } y = x_j. \tag{A4}$$

2. Functional derivative

The functional derivative is an analog to the partial differential such that

$$\frac{\delta}{\delta z(x)}z(y) = \delta(x - y),\tag{A5}$$

which is similar to $(\partial/\partial z_i)z_j = \delta_{ij}$. The functional derivative can be constructed via a formal continuous limit

$$\frac{\delta}{\delta z(x_k)}[\cdots] = \lim_{dx\downarrow 0} \frac{1}{dx}\frac{\partial}{\partial z_k}[\cdots].\tag{A6}$$

Indeed, this definition satisfies the relationship (A5) such that

$$\frac{\delta}{\delta z(x)}z(y) = \lim_{dx\downarrow 0} \frac{1}{dx}\frac{\partial}{\partial z_i}z_j = \lim_{dx\downarrow 0} \frac{1}{dx}\delta_{ij} = \delta(x_i - x_j) \tag{A7}$$

for $x = x_i$ and $y = x_j$.

3. Functional Taylor expansion

For a finite-dimensional vector $z := (z_1, \ldots, z_N)$, the Taylor expansion is given by

$$f(z) = \sum_{n=1}^{\infty} \frac{1}{n!} \left(\sum_{i=1}^{N} z_i \frac{\partial}{\partial x_i} \right)^n f(x) \Bigg|_{x=0}. \quad (A8)$$

As its continuous analog, the functional Taylor expansion for a function $f[z] := f(\{z(x)\}_x)$ reads

$$f[z] = \sum_{n=1}^{\infty} \frac{1}{n!} \left(\int_{-\infty}^{\infty} dx \, z(x) \frac{\delta}{\delta y(x')} \right)^n f[y] \Bigg|_{y=0}. \quad (A9)$$

APPENDIX B: ANOTHER DERIVATION OF THE FIELD MASTER EQUATION (17)

Here we provide another derivation of the field ME (17) by direct manipulation of PDFs. Let us consider the time evolution of any functional $f[\hat{z}] := f(\{\hat{z}(t, x)\}_x)$, given by

$$df(\{\hat{z}(t, x)\}_x) := f(\{\hat{z}(t + dt, x)\}_x) - f(\{\hat{z}(t, x)\}_x)$$

$$= \begin{cases} -dt \int_0^{\infty} dx \frac{\hat{z}(t,x)}{x} \frac{\delta f(\hat{z}(t,x))}{\delta \hat{z}(t,x)} & \text{(no jump during } [t, t+dt): \text{ probability} = 1 - \hat{\lambda}(t)dt) \\ f(\{\hat{z}(t, x) + \hat{y}\tilde{h}(x)\}_x) - f(\{\hat{z}(t, x)\}_x) & \text{(jump in } [t, t+dt) \text{ with } \hat{y} \in [y, y+dy): \text{ probability} = \hat{\lambda}(t)\rho(y)dtdy), \end{cases}$$

$$(B1)$$

with intensity

$$\hat{\lambda}(t) = G[\hat{z}] := g \left(\int_0^{\infty} dx \, \hat{z}(t, x) \right). \quad (B2)$$

By taking the ensemble average on both sides of the equation, we obtain

$$\int \mathcal{D}z \, f[z] \frac{\partial P_t[z]}{\partial t} dt$$

$$= \int \mathcal{D}z \left[-dt \int_0^{\infty} dx \frac{z(x)}{x} \frac{\delta f[z]}{\delta z(x)} \right.$$

$$\left. + dt \int dy \, \rho(y) G[z] (f[z + y\tilde{h}] - f[z]) \right] P_t[z]. \quad (B3)$$

By integration by parts and performing a variable transformation $z + y\tilde{h} \to z$, we obtain an identity

$$\int \mathcal{D}z \, f[z] \frac{\partial P_t[z]}{\partial t}$$

$$= \int \mathcal{D}z \, f[z] \left[\int dx \frac{\delta}{\delta z(x)} \frac{z}{x} P_t[z] \right.$$

$$\left. + \int dy \, \rho(y) G[z - y\tilde{h}] P_t[z - y\tilde{h}] - G[z] P_t[z] \right]. \quad (B4)$$

Since this identity holds for any functional $f[z]$, we obtain Eq. (17).

APPENDIX C: INTEGRAL IDENTITIES FOR EXPONENTIAL MARK DISTRIBUTIONS

Here we provide the detailed derivation of the identities (62) and (85).

1. Positive contribution

Let us consider the quantity

$$\mathcal{I}_+(\nu) := \int_0^{\infty} dy \frac{e^{-y/y^*}}{y^*} f \left(\nu - \frac{\eta y}{\tau} \right), \quad (C1)$$

with the boundary condition $\lim_{\nu \to \pm\infty} f(\nu) = 0$. Let us differentiate both sides as

$$\frac{d}{d\nu} \mathcal{I}_+(\nu) = \int_0^{\infty} dy \frac{e^{-y/y^*}}{y^*} \frac{d}{d\nu} f \left(\nu - \frac{\eta y}{\tau} \right). \quad (C2)$$

The identity

$$\frac{d}{dy} f \left(\nu - \frac{\eta y}{\tau} \right) = -\frac{\eta}{\tau} \frac{d}{d\nu} f \left(\nu - \frac{\eta y}{\tau} \right) \quad (C3)$$

leads to

$$\frac{d}{d\nu} \mathcal{I}_+(\nu) = -\frac{\tau}{\eta} \int_0^{\infty} dy \frac{e^{-y/y^*}}{y^*} \frac{d}{dy} f \left(\nu - \frac{\eta y}{\tau} \right)$$

$$= -\frac{\tau}{\eta} \left[\frac{e^{-y/y^*}}{y^*} f \left(\nu - \frac{\eta y}{\tau} \right) \right]_0^{\infty}$$

$$- \frac{\tau}{\eta y^*} \int_0^{\infty} dy \frac{e^{-y/y^*}}{y^*} f \left(\nu - \frac{\eta y}{\tau} \right), \quad (C4)$$

where we have performed an integration by parts. This means that

$$\frac{d}{d\nu} \mathcal{I}_+(\nu) = c f(\nu) - c \mathcal{I}_+(\nu), \quad c := \frac{\tau}{\eta y^*}, \quad (C5)$$

which is equivalent to

$$\left(1 + \frac{1}{c} \frac{d}{d\nu} \right) \mathcal{I}_+(\nu) = f(\nu). \quad (C6)$$

We thus have the identity (62).

2. Negative contribution

Let us consider the quantity

$$\mathcal{I}_-(\nu) := \int_{-\infty}^{0} dy \frac{e^{y/y^*}}{y^*} f \left(\nu - \frac{\eta y}{\tau} \right), \quad (C7)$$

with the boundary condition $\lim_{\nu \to \pm\infty} f(\nu) = 0$. Let us differentiate both sides as

$$
\begin{aligned}
\frac{d}{d\nu}\mathcal{I}_-(\nu) &= \int_{-\infty}^{0} dy \frac{e^{y/y^*}}{y^*} \frac{d}{d\nu} f\left(\nu - \frac{\eta y}{\tau}\right) \\
&= -\frac{\tau}{\eta} \int_{-\infty}^{0} dy \frac{e^{y/y^*}}{y^*} \frac{d}{dy} f\left(\nu - \frac{\eta y}{\tau}\right) \\
&= -\frac{\tau}{\eta}\left[\frac{e^{y/y^*}}{y^*} f\left(\nu - \frac{\eta y}{\tau}\right)\right]_{-\infty}^{0} \\
&\quad + \frac{\tau}{\eta y^*}\int_{0}^{\infty} dy \frac{e^{y/y^*}}{y^*} f\left(\nu - \frac{\eta y}{\tau}\right), \quad \text{(C8)}
\end{aligned}
$$

where we have performed an integration by parts. This means that

$$
\frac{d}{d\nu}\mathcal{I}_-(\nu) = -cf(\nu) + c\mathcal{I}_-(\nu), \quad c := \frac{\tau}{\eta y^*}, \quad \text{(C9)}
$$

which is equivalent to

$$
\left(1 - \frac{1}{c}\frac{d}{d\nu}\right)\mathcal{I}_-(\nu) = f(\nu). \quad \text{(C10)}
$$

We thus have the identity (85).

APPENDIX D: SUMMARY OF SPECIAL FUNCTIONS

Here we summarize special functions used in this paper.

1. Modified Bessel functions

The modified Bessel functions of the first and second kinds, denoted by $I_\gamma(x)$ and $K_\gamma(x)$, respectively, are defined by

$$
I_\gamma(x) := \sum_{k=0}^{\infty} \frac{1}{k!\,\Gamma(k+\gamma+1)}\left(\frac{x}{2}\right)^{2k+\gamma}, \quad \text{(D1a)}
$$

$$
K_\gamma(x) := \frac{\pi}{2}\frac{I_{-\gamma}(x) - I_\gamma(x)}{\sin(\gamma\pi)}, \quad \text{(D1b)}
$$

which satisfy the modified Bessel equation

$$
x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - (x^2 + \gamma^2)y = 0. \quad \text{(D1c)}
$$

2. Confluent hypergeometric function

The confluent hypergeometric functions of the first and second kinds are defined by

$$
{}_1F_1(a, b; x) := \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 dt\, e^{xt} t^{a-1}(1-t)^{b-a-1}, \quad \text{(D2a)}
$$

$$
{}_1U_1(a, b; x) := \frac{1}{\Gamma(a)}\int_0^\infty dt\, e^{-xt} t^{a-1}(1+t)^{b-a-1}, \quad \text{(D2b)}
$$

respectively. These functions satisfy the confluent hypergeometric differential equation

$$
x\frac{d^2y}{dx^2} + (b-x)\frac{dy}{dx} - ay = 0. \quad \text{(D3)}
$$

For positive c and β, an asymptotic formula is available for large x,

$$
{}_1F_1(c, \beta; x) \propto \frac{\Gamma(\beta)}{\Gamma(c)} e^x x^{c-\beta}, \quad {}_1U_1(c, \beta; x) \propto x^{-c}. \quad \text{(D4)}
$$

3. Generalized Laguerre function

The generalized Laguerre function $L_a^b(x)$ is defined as the solution of the differential equation

$$
x\frac{d^2}{dx^2}L_a^b(x) + (b+1-x)\frac{d}{dx}L_a^b(x) + aL_a^b(x) = 0. \quad \text{(D5)}
$$

For positive c and β, an asymptotic formula is available for large x,

$$
L_{-1-c/\beta}^{2c/\beta}(x) \simeq \frac{\Gamma(c/\beta)}{\Gamma(-c/\beta)\Gamma(1+c/\beta)} e^x x^{-c/\beta}. \quad \text{(D6)}
$$

4. Hypergeometric function

The hypergeometric function is defined as the solution of the hypergeometric differential equation

$$
z(1-z)\frac{d^2y}{dz^2} + [c - (a+b+1)z]\frac{dy}{dz} - aby = 0. \quad \text{(D7)}
$$

The explicit analytic expansion is given by

$$
{}_2F_1(a, b, c; z) := \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k}\frac{z^k}{k!} \quad \text{(D8)}
$$

for $|z| < 1$ with the Pochhammer symbol $(a)_k := \Gamma(a+k)/\Gamma(a)$ and $(a)_0 = 1$. The hypergeometric function has the integral representation

$$
{}_2F_1(a, b, c; z)
$$
$$
:= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}dt \quad \text{(D9)}
$$

for $0 < b < c$ and $|z| < 1$ with real numbers a, b, c, and z. There is an identity

$$
{}_2F_1(a, b, c; z) = (1-z)^{-a}{}_2F_1\left(a, c-b, c; \frac{z}{z-1}\right). \quad \text{(D10)}
$$

a. Useful identities

We state the following useful asymptotic formulas. For $n > 2$ and large $\nu \to +\infty$, the following asymptotic formula holds:

$$
\nu^2 {}_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -\frac{k\nu^n}{v_0}\right) = \underbrace{\frac{\frac{2\pi}{n}}{\sin\frac{2\pi}{n}}\left(\frac{v_0}{k}\right)^{2/n}}_{\text{const}} + o(\nu^0). \quad \text{(D11a)}
$$

In addition, for $n = 2$ and large ν, we obtain

$$
\nu^2 {}_2F_1\left(1, 1, 2; -\frac{k\nu^2}{v_0}\right) = \frac{v_0}{k}\ln\left(\frac{k\nu^2}{v_0}\right) + o(\nu^0), \quad \text{(D11b)}
$$

where we have used the identity ${}_2F_1(1, 1, 2; x) = -x^{-1}\ln(1-x)$.

b. Crossover between $n = 2$ and $n > 2$

The formulas (D11) are qualitatively different for $n = 2$ and $n > 2$ because of the analytical singularity of $_2F_1$ at $n = 2$. Here we consider the crossover between $n = 2$ and $n > 2$. Since the expansion holds for $n > 2$,

$$_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -x\right) = \frac{\frac{2\pi}{n}}{\sin \frac{2\pi}{n}} x^{-2/n} - \frac{\frac{2}{n}}{1 - \frac{2}{n}} x^{-1} + O(x^{-2})$$

(D12)

for large x, by the substitution $x = k\nu^n/\nu_0$, we obtain the expansion in terms of the small parameter $1/x$, valid for any $n > 2$,

$$\nu^2 {}_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -\frac{k\nu^n}{\nu_0}\right)$$

$$\simeq \nu^2 \frac{\frac{2\pi}{n}}{\sin \frac{2\pi}{n}} x^{-1} \left(x^{1-2/n} - \frac{\sin \frac{2\pi}{n}}{\pi\left(1 - \frac{2}{n}\right)}\right)$$

$$= \frac{\nu_0}{k} \nu^{-2\epsilon/(1-\epsilon)} \frac{\pi(1-\epsilon)}{\sin \pi\epsilon}\left(x^\epsilon - \frac{\sin \pi\epsilon}{\pi\epsilon}\right),$$

(D13)

with $\epsilon := 1 - 2/n > 0$. By taking the limit $\epsilon \downarrow 0$ for a large but fixed ν, we can apply

$$\frac{x^\epsilon - 1}{\epsilon} \simeq \ln x + \frac{\epsilon}{2}(\ln x)^2 + \cdots$$

(D14)

to obtain

$$\lim_{\epsilon \downarrow 0}\left\{\nu^2 {}_2F_1\left(1, \frac{2}{n}, 1 + \frac{2}{n}; -\frac{k\nu^n}{\nu_0}\right)\right\} \simeq \frac{\nu_0}{k} \ln\left(\frac{k\nu^2}{\nu_0}\right),$$

(D15)

recovering formula (D11b).

In contrast, even if $\epsilon > 0$ is small, the first-order truncation of expansion (D14) is not applicable for too large ν. Indeed, the truncation is only valid for

$$\frac{\epsilon}{2}(\ln x)^2 \ll |\ln x| \implies \frac{k}{\nu_0}\nu^n \ll e^{2/\epsilon}.$$

(D16)

We thus obtain the characteristic intensity of this crossover as $\lambda^* := \nu_0 e^{2/\epsilon}$.

APPENDIX E: SOLUTION OF AN INTEGRAL EQUATION

Here we study the solution of the integral equation with the form

$$\int_{-\infty}^\infty dy\, \rho(y)\phi(\nu - y) - \phi(\nu) = 0,$$

(E1)

which repeatedly appears in this paper, with the assumption that $\phi(\nu)$ is non-negative. We assume that the mark distribution is two sided

$$p_+ := \int_0^\infty \rho(y)dy > 0, \quad p_- := \int_{-\infty}^0 \rho(y)dy > 0$$

(E2)

and that the mean mark is nonpositive

$$m := \int_{-\infty}^\infty y\rho(y)dy \leqslant 0.$$

(E3)

1. Negative mean mark $m < 0$

Let us first consider the case with negative-mean mark $m < 0$. Let us assume that a special solution of Eq. (E1) is given by an exponential,

$$\phi(\nu) \simeq C_0 e^{-c\nu} \quad \text{for large } \nu.$$

(E4)

By substituting this solution into Eq. (E1), we obtain

$$\Phi(c) = 0, \quad \Phi(x) := \int_{-\infty}^\infty \rho(y)(e^{xy} - 1)dy.$$

(E5)

Based on this fact, we decompose the general solution of Eq. (E1) as the superposition of exponentials

$$\phi(\nu) \simeq \sum_i C_i e^{-c_i \nu} \quad \text{for large } \nu,$$

(E6)

where c_i is the ith root of $\Phi(x) = 0$. Because $\phi(\nu)$ is non-negative, oscillatory solutions (corresponding to c_i being a complex number) are excluded and thus c_i must be a real number. According to Appendix F, the solutions of $\Phi(x) = 0$ are given by $x = 0$ and $x = c^* > 0$. Thus, we obtain

$$\phi(x) \simeq C_0 + C_1 e^{-c^*\nu} \quad \text{for large } \nu.$$

(E7)

2. Zero mean mark $m = 0$

We next consider the case with zero-mean mark $m \uparrow 0$. According to Appendix F, the roots of $\Phi(x) = 0$ are given by $x = 0$ and $x = c^* > 0$.

In the zero-mean mark limit $m \uparrow 0$, the positive root c^* approaches zero as shown here. Let us assume that the mean mark is negative but very small such that $m = -\epsilon$ with a small positive parameter $\epsilon > 0$. The moment-generating function is expanded around $x = 0$ as

$$\Phi(x) = -\epsilon x + \frac{\langle y^2 \rangle x^2}{2} + \cdots, \quad \langle y^2 \rangle := \int_{-\infty}^\infty y^2 \rho(y)dy.$$

(E8)

For small ϵ, the positive root of $\Phi(x) = 0$ is given by

$$c^* = \frac{2\epsilon}{\langle y^2 \rangle} + o(\epsilon) > 0,$$

(E9)

which converges to zero for the small-ϵ limit: $\lim_{\epsilon \downarrow 0} c^* = 0$.

Assuming a small positive ϵ, let us expand the solution (E7) to obtain

$$\phi(x) \simeq C_0' + C_1' x + O(C_1 \epsilon^2),$$
$$C_0' := C_0 + C_1, \quad C_1' := -C_1 c^* = O(C_1 \epsilon).$$

(E10)

Since C_0 and C_1 are arbitrary constants, we can assume $C_1 = O(\epsilon^{-1})$, $C_0' = O(\epsilon^0)$, and $C_1' = O(\epsilon^0)$, where the divergence of C_1 is absorbed by an appropriate selection of C_0. Under this assumption, the solution is given by

$$\phi(x) \simeq C_0' + C_1' \nu + O(\epsilon^1), \quad C_0' = O(\epsilon^0), \quad C_1' = O(\epsilon^0).$$

(E11)

By taking the zero-mean limit $\epsilon \downarrow 0$, we obtain

$$\phi(x) \simeq C_0' + C_1' \nu \quad \text{for large } \nu$$

(E12)

as the general solution.

APPENDIX F: ANALYTICAL PROPERTIES OF THE MOMENT-GENERATING FUNCTION $\Phi(x)$

Here we summarize the analytical properties of the moment-generating function

$$\Phi(x) := \int_{-\infty}^{\infty} dy\, \rho(y)(e^{xy} - 1) \tag{F1}$$

in the regime where the mean mark is nonpositive

$$m := \int_{-\infty}^{\infty} y\rho(y)dy \leqslant 0 \tag{F2}$$

and the mark distribution is two sided

$$p_+ := \int_0^{\infty} \rho(y)dy > 0, \quad p_- := \int_{-\infty}^0 \rho(y)dy > 0. \tag{F3}$$

The moment-generating function $\Phi(x)$ is a strictly convex function because

$$\frac{d^2\Phi(x)}{dx^2} = \int_{-\infty}^{\infty} y^2\rho(y)e^{xy}dy > 0, \tag{F4}$$

implying that $\Phi(x)$ has no more than one minimum.

1. Negative mean mark $m < 0$

Because we have assumed $p_+ = \int_0^{\infty} dy\, \rho(y) > 0$, $\rho(y) \neq 0$ in some finite region of $y > 0$. Therefore, the relation

$$m_+ := \int_0^{\infty} y\rho(y)dy > 0 \tag{F5}$$

must hold. Based on this fact, the following three properties hold.

(i) The tangential line of the curve $\Phi(x)$ at $x = 0$ has a negative slope:

$$\left.\frac{d\Phi(x)}{dx}\right|_{x=0} = \int_{-\infty}^{\infty} y\rho(y)dy = m < 0. \tag{F6}$$

(ii) The moment-generating function is zero at $x = 0$:

$$\Phi(0) = 0. \tag{F7}$$

(iii) The moment-generating function diverges to infinity for large x

$$\lim_{x \to \infty} \Phi(x) = \infty \tag{F8}$$

because

$$\begin{aligned}
\Phi(x) &= \int_0^{\infty} dy\, \rho(y)(e^{xy} - 1) + \int_{-\infty}^0 dy\, \rho(y)(e^{xy} - 1) \\
&\geqslant \int_0^{\infty} dy\, \rho(y)xy + \int_{-\infty}^0 dy\, \rho(y)(0 - 1) \\
&= m_+ x - p_- \to \infty, \tag{F9}
\end{aligned}$$

where we have used $e^{xy} \geqslant xy + 1$ for $x \geqslant 0$ and $e^{xy} > 0$ for all x. These properties imply that the minimum of $\Phi(x)$ exists at some $x^*_{\min} > 0$ as depicted in Fig. 13(a) and that the roots of $\Phi(x) = 0$ are given by $x = 0$ and $x = c^* > 0$.
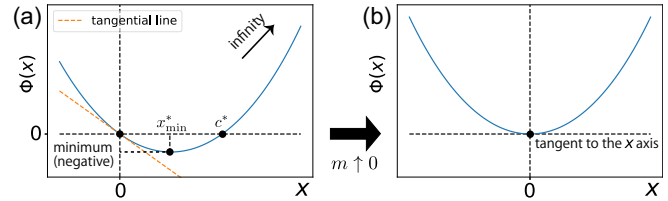

FIG. 13. Schematic of the shape of the moment-generating function $\Phi(x)$. Here $\Phi(x)$ is strictly convex and takes specific values $\Phi(0) = 0$, $\Phi(+\infty) = \infty$, and $d\Phi(0)/dx = m$. (a) For negative m, the tangential line at $x = 0$ has a negative slope $d\Phi(0)/dx = m < 0$. The minimum occurs at $x = x^*_{\min} > 0$ and the roots of $\Phi(x) = 0$ are given by $x = 0$ and $x = c^* > 0$. (b) In the zero-mean mark limit $m \uparrow 0$, c^* approaches zero and thus the only root of $\Phi(x) = 0$ is given by $x = 0$.

2. Zero mean mark $m = 0$

Let us consider the case where the mean mark is zero $m = 0$, which is realized typically for symmetric mark distributions $\rho(y) = \rho(-y)$. Under this condition, the minimum occurs at $x = 0$ because

$$\left.\frac{d\Phi(x)}{dx}\right|_{x=0} = 0. \tag{F10}$$

This implies that the solution of $\Phi(x) = 0$ is $x = 0$, which is a double root [see Fig. 13(b)]. The appearance of the double root can be understood as the zero-mean mark limit of the negative-mean mark regime. While c^* is positive for $m < 0$, it approaches zero in the zero-mean mark limit: $\lim_{m \uparrow 0} c^* = 0$ (see the discussion in Appendix E 2).

In the special case of symmetric mark distributions $\rho(y) = \rho(-y)$, the moment-generating function can be transformed into

$$\Phi(x) = \int_{-\infty}^{\infty} \rho(y)(\cosh xy - 1)dy. \tag{F11}$$

Also, considering the geometrical shape [see Fig. 13(b)], the equation

$$\Phi(x) = ax \tag{F12}$$

has a single positive root for positive a.

3. Special cases

For reference, we summarize several results for specific mark distributions.

a. Gaussian mark distribution

Let us consider the Gaussian mark distribution

$$\rho(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-m)^2/2\sigma^2}, \tag{F13}$$

with mean mark m and variance σ^2. The corresponding moment-generating function is given by

$$\Phi(x) = e^{mx + \sigma^2 x^2/2} - 1, \tag{F14}$$

which leads to the explicit formula for the root of $\Phi(c^*) = 0$ as

$$c^* = -\frac{2m}{\sigma^2}. \tag{F15}$$

b. Two-sided exponential asymmetric mark distribution

We next consider the case of the two-sided exponential asymmetric mark distribution

$$\rho(y) = \begin{cases} \frac{p_+}{y_+^*} e^{-y/y_+^*}, & y \geqslant 0 \\ \frac{p_-}{y_-^*} e^{y/y_-^*}, & y < 0, \end{cases} \tag{F16}$$

where $p_+ + p_- = 1$, $y_+^* > 0$, and $y_-^* > 0$. The mean mark is given by

$$m := y_+^* p_+ - y_-^* p_- < 0. \tag{F17}$$

We obtain

$$\Phi(x) = \frac{p_+ y_+^* x}{1 - y_+^* x} - \frac{p_- y_-^* x}{1 + y_-^* x}, \tag{F18}$$

which leads to

$$c^* = \frac{p_- y_-^* - p_+ y_+^*}{y_+^* y_-^*} = -\frac{m}{y_+^* y_-^*} > 0 \tag{F19}$$

as the unique positive root of $\Phi(c^*) = 0$.

APPENDIX G: PROOFS OF MATHEMATICAL PROPERTIES OF H [EQ. (208)]

Here we summarize the proofs of the main mathematical properties of H defined by Eq. (208).

1. Proof that eigenvalues are real

We show that all eigenvalues of H are real numbers as follows. First, H can be symmetrized as $\bar{H}$, defined by

$$\bar{H} := AHA^{-1} = \begin{pmatrix} \frac{1}{\tau_1} - \tilde{h}_1 & -\sqrt{\tilde{h}_1 \tilde{h}_2} & \cdots & -\sqrt{\tilde{h}_1 \tilde{h}_K} \\ -\sqrt{\tilde{h}_2 \tilde{h}_1} & \frac{1}{\tau_2} - \tilde{h}_2 & \cdots & -\sqrt{\tilde{h}_2 \tilde{h}_K} \\ \vdots & \vdots & \ddots & \vdots \\ -\sqrt{\tilde{h}_K \tilde{h}_1} & -\sqrt{\tilde{h}_K \tilde{h}_2} & \cdots & \frac{1}{\tau_K} - \tilde{h}_K \end{pmatrix}, \quad A := \begin{pmatrix} \sqrt{\tilde{h}_1} & 0 & \cdots & 0 \\ 0, & \sqrt{\tilde{h}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\tilde{h}_K} \end{pmatrix}. \tag{G1}$$

Indeed, by representing all the matrices by their elements $\bar{H} := (\bar{H}_{ij})$, $H := (H_{ij})$, and $A := A_{ij}$, we obtain

$$\bar{H}_{ij} = \sum_{k,l} A_{ik} H_{kl} A_{lj}^{-1} = \sum_{k,l} \sqrt{\tilde{h}_i} \delta_{ik} \left(\frac{\delta_{kl}}{\tau_k} - \tilde{h}_l \right) \sqrt{\frac{1}{\tilde{h}_j}} \delta_{lj} = \frac{\delta_{ij}}{\sqrt{\tau_i \tau_j}} - \sqrt{\tilde{h}_i \tilde{h}_j}. \tag{G2}$$

Since $\bar{H}$ is a symmetric matrix, all their eigenvalues are real. We obtain

$$H e_i = \lambda_i e_i \iff \bar{H}(A e_i) = \lambda_i (A e_i). \tag{G3}$$

This relationship implies that any of the eigenvalues of H are the same as that of $\bar{H}$. Thus, all the eigenvalues of H are likewise real.

2. Determinant

The determinant $\det H$ is derived as follows. Let us recall the invariance of determinants

$$\det H = \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_j \\ \vdots \\ a_K \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_j + c a_k \\ \vdots \\ a_K \end{pmatrix} \tag{G4}$$

for any constant c. This implies that

$$\det H = \det \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_K \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 - a_1 \\ a_3 \\ \vdots \\ a_K \end{pmatrix} = \det \begin{pmatrix} a_1 \\ a_2 - a_1 \\ a_3 - a_1 \\ \vdots \\ a_K \end{pmatrix} = \cdots = \det \begin{pmatrix} a_1 \\ a_2 - a_1 \\ a_3 - a_1 \\ \vdots \\ a_K - a_1 \end{pmatrix} := \det \begin{pmatrix} a_1' \\ a_2' \\ a_3' \\ \vdots \\ a_K' \end{pmatrix} \tag{G5}$$

and

$$
\det \boldsymbol{H} = \det \begin{pmatrix} \boldsymbol{a}'_1 \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_K \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{a}'_1 + \tau_2 \tilde{h}_2 \boldsymbol{a}'_2 \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_K \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{a}'_1 + \tau_2 \tilde{h}_2 \boldsymbol{a}'_2 + \tau_3 \tilde{h}_3 \boldsymbol{a}'_3 \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_K \end{pmatrix} = \cdots = \det \begin{pmatrix} \boldsymbol{a}'_1 + \sum_{k=2}^{K} \tau_k \tilde{h}_k \boldsymbol{a}'_k \\ \boldsymbol{a}'_2 \\ \vdots \\ \boldsymbol{a}'_K \end{pmatrix}, \tag{G6}
$$

with constants $\{\tau_k \tilde{h}_k\}_k$. Using these relations, the determinant of $\boldsymbol{H}$ is given by

$$
\begin{aligned}
\det \boldsymbol{H} &= \det \begin{pmatrix} -\tilde{h}_1 + 1/\tau_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K \\ -\tilde{h}_1 & -\tilde{h}_2 + 1/\tau_2 & \cdots & -\tilde{h}_K \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{h}_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K + 1/\tau_K \end{pmatrix} \begin{matrix} \leftarrow \boldsymbol{a}_1 \\ \leftarrow \boldsymbol{a}_2 \\ \vdots \\ \leftarrow \boldsymbol{a}_K \end{matrix} \\[2mm]
&= \det \begin{pmatrix} (1 - \tau_1 \tilde{h}_1)/\tau_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K \\ -1/\tau_1 & 1/\tau_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1/\tau_1 & 0 & \cdots & 1/\tau_K \end{pmatrix} \begin{matrix} \leftarrow \boldsymbol{a}'_1 & = \boldsymbol{a}_1 \\ \leftarrow \boldsymbol{a}'_2 & = \boldsymbol{a}_2 & - & \boldsymbol{a}_1 \\ \vdots \\ \leftarrow \boldsymbol{a}'_K & = \boldsymbol{a}_K & - & \boldsymbol{a}_1 \end{matrix} \\[2mm]
&= \det \begin{pmatrix} \left(1 - \sum_{k=1}^{K} \tau_k \tilde{h}_k\right)/\tau_1 & 0 & \cdots & 0 \\ -1/\tau_1 & 1/\tau_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1/\tau_1 & 0 & \cdots & 1/\tau_K \end{pmatrix} \begin{matrix} \leftarrow \boldsymbol{a}''_1 & = \boldsymbol{a}'_1 & + & \sum_{k=2}^{K} \tau_k \tilde{h}_k \boldsymbol{a}'_k \\ \leftarrow \boldsymbol{a}''_2 & = \boldsymbol{a}'_2 \\ \vdots \\ \leftarrow \boldsymbol{a}''_K & = \boldsymbol{a}'_K \end{matrix} \\[2mm]
&= \frac{1 - \sum_{k=1}^{K} \tau_k \tilde{h}_k}{\tau_1 \cdots \tau_K}.
\end{aligned} \tag{G7}
$$

Notably, $\det \boldsymbol{H} = 0$ at criticality $\eta = 1$. This singularity is consistent with the singularity of the inverse matrix $\boldsymbol{H}^{-1}$, as discussed in Appendix G 3.

3. Inverse matrix

The inverse matrix of $\boldsymbol{H}$ is derived from the method of row reduction

$$
\begin{aligned}
&\left(\begin{array}{cccc|cccc} -\tilde{h}_1 + 1/\tau_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K & 1 & 0 & \cdots & 0 \\ -\tilde{h}_1 & -\tilde{h}_2 + 1/\tau_2 & \cdots & -\tilde{h}_K & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -n_1/\tau_1 & -n_2/\tau_2 & \cdots & -n_K/\tau_K + 1/\tau_K & 0 & 0 & \cdots & 1 \end{array} \right) \begin{matrix} \leftarrow \boldsymbol{b}_1 \\ \leftarrow \boldsymbol{b}_2 \\ \vdots \\ \leftarrow \boldsymbol{b}_K \end{matrix} \\[2mm]
\rightarrow &\left(\begin{array}{cccc|cccc} (1 - \tau_1 \tilde{h}_1)/\tau_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K & 1 & 0 & \cdots & 0 \\ -1/\tau_1 & 1/\tau_2 & \cdots & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1/\tau_1 & 0 & \cdots & 1/\tau_K & -1 & 0 & \cdots & 1 \end{array} \right) \begin{matrix} \leftarrow \boldsymbol{b}'_1 & = \boldsymbol{b}_1 \\ \leftarrow \boldsymbol{b}'_2 & = \boldsymbol{b}_2 & - & \boldsymbol{b}_1 \\ \vdots \\ \leftarrow \boldsymbol{b}'_K & = \boldsymbol{b}_K & - & \boldsymbol{b}_1 \end{matrix} \\[2mm]
\rightarrow &\left(\begin{array}{cccc|cccc} (1-\eta)/\tau_1 & 0 & \cdots & 0 & 1 - \sum_{k=2}^{K} \tau_k \tilde{h}_k & \tau_2 \tilde{h}_2 & \cdots & \tau_K \tilde{h}_K \\ -1/\tau_1 & 1/\tau_2 & \cdots & 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -1/\tau_1 & 0 & \cdots & 1/\tau_K & -1 & 0 & \cdots & 1 \end{array} \right) \begin{matrix} \leftarrow \boldsymbol{b}''_1 & = \boldsymbol{b}'_1 & + & \sum_{k=2}^{K} \tau_k \tilde{h}_k \boldsymbol{b}'_k \\ \leftarrow \boldsymbol{b}''_2 & = \boldsymbol{b}'_2 \\ \vdots \\ \leftarrow \boldsymbol{b}''_K & = \boldsymbol{b}'_K \end{matrix} \\[2mm]
\rightarrow &\left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & \tau_1 + \tau_1 n_1/(1-\eta) & \tau_1 n_2/(1-\eta) & \cdots & \tau_1 n_K/(1-\eta) \\ -\tau_2/\tau_1 & 1 & \cdots & 0 & -\tau_2 & \tau_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\tau_K/\tau_1 & 0 & \cdots & 1 & -\tau_K & 0 & \cdots & \tau_K \end{array} \right) \begin{matrix} \leftarrow \boldsymbol{b}'''_1 & = \frac{\tau_1}{1-\eta} \boldsymbol{b}''_1 \\ \leftarrow \boldsymbol{b}'''_2 & = \tau_2 \boldsymbol{b}''_2 \\ \vdots \\ \leftarrow \boldsymbol{b}'''_K & = \tau_K \boldsymbol{b}''_K \end{matrix} \\[2mm]
\rightarrow &\left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & \tau_1 + \tau_1^2 \tilde{h}_1/(1-\eta) & \tau_1 \tau_2 \tilde{h}_2/(1-\eta) & \cdots & \tau_1 \tau_K \tilde{h}_K/(1-\eta) \\ 0 & 1 & \cdots & 0 & \tau_2 \tau_1 \tilde{h}_1/(1-\eta) & \tau_2 + \tau_2^2 \tilde{h}_2/(1-\eta) & \cdots & \tau_2 \tau_K \tilde{h}_K/(1-\eta) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \tau_K \tau_1 \tilde{h}_1/(1-\eta) & \tau_K \tau_2 \tilde{h}_2/(1-\eta) & \cdots & \tau_K + \tau_K^2 \tilde{h}_K/(1-\eta) \end{array} \right) \begin{matrix} \leftarrow \boldsymbol{b}''''_1 & = \boldsymbol{b}'''_1 \\ \leftarrow \boldsymbol{b}''''_2 & = \boldsymbol{b}'''_2 & + & \frac{\tau_2}{\tau_1} \boldsymbol{b}'''_1 \\ \vdots \\ \leftarrow \boldsymbol{b}''''_K & = \boldsymbol{b}'''_K & + & \frac{\tau_2}{\tau_1} \boldsymbol{b}'''_1 \end{matrix}
\end{aligned} \tag{G8}
$$

with the branching ratio η defined by

$$\eta := \sum_{k=1}^{K} \tau_k \tilde{h}_k. \tag{G9}$$

This relation implies

$$\boldsymbol{H}^{-1} = \begin{pmatrix} \tau_1 + \tau_1^2 \tilde{h}_1/(1-\eta) & \tau_1 \tau_2 \tilde{h}_2/(1-\eta) & \cdots & \tau_1 \tau_K \tilde{h}_K/(1-\eta) \\ \tau_2 \tau_1 \tilde{h}_1/(1-\eta) & \tau_2 + \tau_2^2 \tilde{h}_2/(1-\eta) & \cdots & \tau_2 \tau_K \tilde{h}_K/(1-\eta) \\ \vdots & \vdots & \ddots & \vdots \\ \tau_K \tau_1 \tilde{h}_1/(1-\eta) & \tau_K \tau_2 \tilde{h}_2/(1-\eta) & \cdots & \tau_K + \tau_K^2 \tilde{h}_K/(1-\eta) \end{pmatrix} \tag{G10}$$

or equivalently

$$H_{ij}^{-1} = \tau_i \delta_{ij} + \frac{\tau_i \tau_j \tilde{h}_j}{1-\eta} \tag{G11}$$

in the representation by matrix elements. The above calculation can be directly confirmed as follows:

$$\boldsymbol{H}\boldsymbol{H}^{-1} = \boldsymbol{I} \iff \sum_{j=1}^{K} H_{ij} H_{jk}^{-1} = \sum_{j=1}^{K} \left(-\tilde{h}_j + \frac{1}{\tau_j} \delta_{ij} \right) \left(\tau_j \delta_{jk} + \frac{\tau_j \tau_k \tilde{h}_k}{1-\eta} \right) = \delta_{ik}. \tag{G12}$$

The inverse matrix has a singularity at $\eta = 1$, corresponding to the criticality of the ramp Hawkes process.

4. Eigenvectors of H

Since $\boldsymbol{H}$ is directly associated with the real symmetric matric $\tilde{\boldsymbol{H}}$, $\boldsymbol{H}$ can be diagonalized such that

$$\boldsymbol{P} := (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_K), \quad \boldsymbol{P}^{-1}\boldsymbol{H}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix} \tag{G13}$$

with eigenvectors $\{\boldsymbol{e}_k\}_{k=1,\ldots,K}$ and corresponding eigenvalues $\{\lambda_k\}_{k=1,\ldots,K}$.

At criticality $\eta = 1$, the smallest eigenvalues is zero such that $\lambda_1 = 0$. In addition, the zero eigenvector $\boldsymbol{e}_1$ is explicitly given by

$$\boldsymbol{e}_1 = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_K \end{pmatrix}. \tag{G14}$$

Indeed, we can directly confirm the relationship

$$\boldsymbol{H}\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\tau_1} - \tilde{h}_1 & -\tilde{h}_2 & \cdots & -\tilde{h}_K \\ -\tilde{h}_1 & \frac{1}{\tau_2} - \tilde{h}_2 & \cdots & -\tilde{h}_K \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{h}_1 & -\tilde{h}_2 & \cdots & \frac{1}{\tau_K} - \tilde{h}_K \end{pmatrix} \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_K \end{pmatrix}$$

$$= \begin{pmatrix} 1-\eta \\ 1-\eta \\ \vdots \\ 1-\eta \end{pmatrix} = \boldsymbol{0} \quad \text{for } \eta := \sum_{k=1}^{K} \tau_k \tilde{h}_k = 1. \tag{G15}$$

We next consider the representation based on the eigenvectors

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} := \boldsymbol{P}^{-1}\boldsymbol{s}, \quad \boldsymbol{P}^{-1} = \begin{pmatrix} \boldsymbol{g}_1 \\ \boldsymbol{g}_2 \\ \vdots \\ \boldsymbol{g}_K \end{pmatrix}. \tag{G16}$$

On the basis of this representation, we obtain

$$\frac{dX_1}{dl} = 0 + O(X^2), \quad \frac{dX_k}{dl} = -\lambda_k X_k + O(X^2) \quad \text{for } k \geqslant 2. \tag{G17}$$

This implies that the leading-order contribution comes from the X_1 direction such that $|X_1| \gg |X_k|$ with $k \geqslant 2$. We can approximate

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + O(X^2) \implies \boldsymbol{s} = \boldsymbol{P}\boldsymbol{X} \simeq X_1 \boldsymbol{e}_1 + O(X^2). \tag{G18}$$

By direct substitution, we can confirm that $\boldsymbol{g}_1$ is given by

$$\boldsymbol{g}_1 = \left(\frac{\tau_1 \tilde{h}_1}{\langle \tau \rangle}, \ldots, \frac{\tau_K \tilde{h}_K}{\langle \tau \rangle} \right), \quad \langle \tau \rangle := \sum_{k=1}^{K} \tau_k^2 \tilde{h}_k. \tag{G19}$$

Indeed, this implies that X_1 is given by

$$X_1 = \boldsymbol{g}_1 \cdot \boldsymbol{s} = \frac{1}{\langle \tau \rangle} \sum_{k=1}^{K} \tau_k \tilde{h}_k s_k, \tag{G20}$$

which leads to

$$\frac{dX_1}{dl} = \frac{1}{\langle \tau \rangle} \sum_{k=1}^{K} \tau_k \tilde{h}_k \frac{ds_k}{dl} = 0 + O(X^2). \tag{G21}$$

Thus, we find that the first-order contribution is absent in Eq. (G21), confirming the correctness of the representation of Eq. (G19). In addition, this representation (G19) is consistent with the identity

$$
\boldsymbol{P}^{-1}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{g}_1 \\ \boldsymbol{g}_2 \\ \vdots \\ \boldsymbol{g}_K \end{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_K) = \begin{pmatrix} \tau_1\tilde{h}_1/\langle\tau\rangle & \tau_2\tilde{h}_2/\langle\tau\rangle & \cdots & \tau_K\tilde{h}_K/\langle\tau\rangle \\ \bigcirc & \bigcirc & \cdots & \bigcirc \\ \vdots & \vdots & \ddots & \vdots \\ \bigcirc & \bigcirc & \cdots & \bigcirc \end{pmatrix} \begin{pmatrix} \tau_1 & \bigcirc & \cdots & \bigcirc \\ \tau_2 & \bigcirc & \cdots & \bigcirc \\ \vdots & \vdots & \ddots & \vdots \\ \tau_2 & \bigcirc & \cdots & \bigcirc \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},
$$
(G22)

where $\bigcirc$ represents some unspecified value.[6] Thus, we find that Eq. (G19) is the correct and consistent representation.

APPENDIX H: SUMMARY OF ASYMPTOTIC FORMS OF THE LAPLACE TRANSFORM

Here we summarize the asymptotic forms of the Laplace transform, in particular for power-law distributions. Let us first recall the Tauberian theorem for the Laplace transform of asymptotic power-law functions [56].

Theorem. Let us consider a function $f(x)$ satisfying the asymptotic form

$$
f(x) \simeq x^{\rho-1}L(x) \quad \text{for large } x,
$$
(H1)

with $0 < \rho < \infty$ and slowly varying function $L(x)$. By definition, a slowly varying function satisfies $\lim_{x\to\infty}[L(Cx)/L(x)] = 1$ for any positive constant C. The Laplace transform of $f(x)$ has the asymptotic form

$$
\tilde{f}(s) := \mathcal{L}_1[f(x); s] \simeq \Gamma(\rho)s^{-\rho}L\left(\frac{1}{s}\right) \quad \text{for small } s. \quad (H2)
$$

Using this theorem, let us consider the Laplace transform of power-law functions $f(x) \simeq Ax^{-1-a}$ for various a and positive constant A.

1. Negative case: $a < 0$

For $a < 0$, the Tauberian theorem can be readily applied to obtain

$$
\tilde{f}(s) \simeq A's^a, \quad A' := A\Gamma(-a) > 0 \quad \text{for small } s. \quad (H3)
$$

2. Positive case: $0 < a < 1$

Let us consider the relation

$$
\begin{aligned}
\tilde{f}(s) &= \int_0^\infty f(x)e^{-sx}dx \\
&= [-F^{(0)}(x)e^{-sx}]_0^\infty - s\int_0^\infty F^{(0)}(x)e^{-sx}dx \\
&= F^{(0)}(0) - s\mathcal{L}_1[F^{(0)}(x); s], \quad F^{(0)}(x) := \int_x^\infty f(x')dx'.
\end{aligned}
$$
(H4)

Here we notice that the asymptotic tail of $F^{(0)}(x)$ satisfies the condition of the Tauberian theorem such that

$$
F^{(0)}(x) \simeq \int_x^\infty \frac{A\,dx}{x^{1+a}} = \frac{A}{a}x^{-a} \quad \text{for large } x. \quad (H5)
$$

By applying the Tauberian theorem, we obtain

$$
\tilde{f}(s) \simeq F^{(0)}(0) - A's^a + o(s^a), \quad A' := \frac{A}{a}\Gamma(1-a) > 0. \quad (H6)
$$

When $f(x)$ is a PDF, $F^{(0)}(0) = \int_0^\infty dx\, f(x) = 1$ and we obtain

$$
\tilde{f}(s) \simeq 1 - A's^a + o(s^a). \quad (H7)
$$

3. General positive case: Noninteger $0 < a$

Let us introduce the integer $m := \lfloor a \rfloor = \max\{k \in \boldsymbol{Z} \mid k \leqslant a\}$ with the set of integers $\boldsymbol{Z}$, satisfying $m \leqslant a \leqslant m+1$, and the iterated integral of $f(x)$:

$$
F^{(l)}(x) := \int_x^\infty dx_{l+1} \int_{x_{l+1}}^\infty dx_l \cdots \int_{x_2}^\infty dx_1 f(x_1). \quad (H8)
$$

We find that we can apply the Tauberian theorem to $F^{(m)}(x)$ because

$$
\begin{aligned}
F^{(m)}(x) &\simeq \frac{A(-1)^{m+1}}{(-a)(1-a)\cdots(m-a)}x^{m-a} \\
&= A(-1)^{m+1}\frac{\Gamma(-a)}{\Gamma(m-a+1)}x^{m-a}
\end{aligned}
$$
(H9)

with $-1 < m - a < 0$. Due to the identity

$$
\mathcal{L}_1[F^{(l)}(x); s] = F^{(l+1)}(0) - s\mathcal{L}_1[F^{(l+1)}(x); s], \quad (H10)
$$

we obtain an asymptotic relation for small s,

$$
\begin{aligned}
\mathcal{L}_1[f(x); s] &= \sum_{k=0}^m (-s)^k F^{(k)}(0) + (-s)^{m+1}\mathcal{L}_1[F^{(m)}(x); s] \\
&\simeq \sum_{k=0}^m (-s)^k F^{(k)}(0) + A\Gamma(-a)s^a + o(s^a).
\end{aligned}
$$
(H11)

Since the iterated integral satisfies the identity

$$
F^{(l)}(x) = \frac{1}{l!}\int_x^\infty (x'-x)^l f(x')dx', \quad (H12)
$$

---

[6]If we set $\boldsymbol{g}_1 = (\frac{\tau_1\tilde{h}_1}{c}, \ldots, \frac{\tau_K\tilde{h}_K}{c})$ with some constant $c \neq \langle\tau\rangle$, the identity $\boldsymbol{P}^{-1}\boldsymbol{P} = \boldsymbol{E}$ does not hold with the unit vector $\boldsymbol{E}$, while the relation $dX_1/dl = 0 + O(s^2)$ still holds. Therefore, c must be $\langle\tau\rangle$ for self-consistency.

we find that $F^l(0)$ is proportional to the lth-order moment M_l when $f(x)$ is a PDF:

$$M_l := \int_0^\infty x^l f(x) dx = l! F^{(l)}(0). \qquad \text{(H13)}$$

Stated differently,

$$\mathcal{L}_1[f(x); s] \simeq \sum_{k=0}^m \frac{(-1)^k M_k}{k!} s^k + A's^a + o(s^a),$$

$$A' := A\Gamma(-a). \qquad \text{(H14)}$$

Since A is positive for $f(x)$ to be a PDF, i.e., $f(x) \geqslant 0$, A' must be negative (positive) for even (odd) m.

APPENDIX I: NUMERICAL IMPLEMENTATION

This Appendix describes our numerical method for the simulation of the NLH processes. Our starting point is the Markovian SDE (18) for the discrete sum of exponentials. We introduce a discretized time series $\{t_i\}_i$ and time steps $\Delta t_i := t_{i+1} - t_i$, satisfying $0 = t_0 < t_1 < \cdots < t_{N_T} = T_{tot}$. Given that the intensity of the state-dependent Poisson process is given by $\hat{\lambda}(t) = G(\hat{z}(t))$, the discrete version of the SDE (18) is given by

$$\hat{z}_k(t_i + \Delta t_i) = \hat{z}_k(t_i) e^{-\Delta/t_i \tau_k}$$
$$+ \begin{cases} 0 & (\text{probability } 1 - G(\hat{z}(t_k))\Delta t_k) \\ \tilde{h}_k \hat{y}_i & (\text{probability } G(\hat{z}(t_k))\Delta t_k) \end{cases} \qquad \text{(I1)}$$

with the independent and identically distributed random number sequence $\{\hat{y}_i\}_i$ obeying the mark distribution $\rho(y)$. Given that $G(\hat{z}(t_k))\Delta t_k$ must be sufficiently small such that $G(\hat{z}(t_k))\Delta t_k \ll 1$ for a proper probability interpretation, we employ an adaptive scheme for the time discretization

$$\Delta t_i = \min\left\{ \Delta t_{max}^{(1)}, \frac{\Delta t_{max}^{(2)}}{G(\hat{z}(t_i))} \right\} \qquad \text{(I2)}$$

because the intensity $G(\hat{z}(t_i))$ sometimes takes extremely large values near criticality. For this setup, we obtained an empirical intensity distribution by assuming ergodicity as

$$P_{ss}(\lambda) = \lim_{t \to \infty} \langle \delta(\lambda - \hat{\lambda}(t)) \rangle$$
$$= \lim_{T_{tot} \to \infty} \frac{1}{T_{tot}} \int_0^{T_{tot}} \delta(\lambda - G(\hat{z}(t))) dt$$
$$\simeq \frac{1}{T_{tot}} \sum_{i=0}^{N_T - 1} \delta(\lambda - G(\hat{z}(t_i))) \Delta t_i. \qquad \text{(I3)}$$

For its practical implementation, we have applied a parallel computing technique for better convergence. We have obtained the empirical intensity distribution as

$$P_{ss}(\lambda) = \left\langle \lim_{T_{tot} \to \infty} \frac{1}{T_{tot}} \int_0^{T_{tot}} \delta(\lambda - G(\hat{z}(t))) dt \right\rangle$$
$$\simeq \frac{1}{N_{PC}} \sum_{j=1}^{N_{PC}} \left(\frac{1}{T_{tot}} \sum_{i=0}^{N_T^{(j)} - 1} \delta\big(\lambda - G\big(\hat{z}^{(j)}\big(t_i^{(j)}\big)\big)\big) \Delta t_i^{(j)} \right),$$
$$\text{(I4)}$$

where $\hat{z}^{(j)}(t)$ is the trajectory obtained in the jth parallel thread and N_{PC} is the number of total parallel threads.

1. Ramp intensity map without inhibitory effect (Fig. 10)

We describe the setup for Fig. 10, where the intensity function is given by the ramp function and the mark takes a single value as

$$G(\hat{z}) := \max\left\{ \sum_{k=1}^K \hat{z}_k - \nu_1, \nu_0 \right\}, \quad \rho(y) = \delta(y - 1), \quad \text{(I5)}$$

with any positive number ν_0 and any real number ν_1.

a. Figure 10(a)

The parameters are given by $K = 2$, $(\tau_1, \tau_2) = (1, 2)$, $(\tilde{h}_1, \tilde{h}_2) = (0.7, 0.149\,95)$, $\eta = 0.9999$, $\nu_0 = 0.01$, $\nu_1 \simeq 0.385$, $T_{tot} = 5 \times 10^6$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$ with the initial condition $(\hat{z}_1(0), \hat{z}_2(0)) = (1, 1)$. Since $\alpha_2 = 1$, we obtain the power-law exponent $a \simeq 1.0$ from Eq. (198). The total number of parallel threads is given by $N_{PC} = 4$.

b. Figure 10(b)

The parameters are given by $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 2, 3)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.15, 0.1999/3)$, $\eta = 0.9999$, $\nu_0 = 0.01$, $\nu_1 \simeq 0.147$, $T_{tot} = 5 \times 10^6$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$ with the initial condition $(\hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0)) = (1, 1, 1)$. Since $\alpha_2 = 1$, we obtain the power-law exponent $a \simeq 0.5$ from Eq. (198). The total number of parallel threads is given by $N_{PC} = 4$.

c. Figure 10(c)

The parameters are given by $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 2, 3)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.15, 0.1999/3)$, $\eta = 0.9999$, $\nu_0 = 0.01$, $\nu_1 = 0$, $T_{tot} = 5 \times 10^5$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$ with the initial condition $(\hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0)) = (1, 1, 1)$. Since $\alpha_2 = 1$, we obtain the power-law exponent $a = 0$ from Eq. (198). The total number of parallel threads is given by $N_{PC} = 4$.

2. MSA intensity map with inhibitory effect (Fig. 11)

We describe the setup for Fig. 11, where the intensity map and the mark distribution are given by the exponential function with finite cutoff and the normal distribution, respectively, such that

$$G(\hat{z}) = \min\left\{ \lambda_0 \exp\left(\beta \sum_{k=1}^K \hat{z}_k \right), \lambda_{max} \right\},$$
$$\rho(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/2\sigma^2}. \qquad \text{(I6)}$$

a. Figure 11(a)

The parameters are given by $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 0.5, 2)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.6, 0.1)$, $\lambda_0 = 0.01$, $\beta = 6$, $\lambda_{max} = 10^6$, $\sigma = 0.3$, $T_{tot} = 5 \times 10^6$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$ with the initial condition $(\hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0)) =$

(0, 0, 0). The total number of parallel threads is given by $N_{PC} = 8$.

b. Figure 11(b)

The parameters are given by $K = 3$, $(\tau_1, \tau_2, \tau_3) = (1, 0.5, 2)$, $(\tilde{h}_1, \tilde{h}_2, \tilde{h}_3) = (0.5, 0.6, 0.1)$, $\lambda_0 = 0.001$, $\beta = 10$, $\lambda_{max} = 10^6$, $\sigma = 0.3$, $T_{tot} = 5 \times 10^5$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$ with the initial condition $(\hat{z}_1(0), \hat{z}_2(0), \hat{z}_3(0)) = (0, 0, 0)$.

3. Brownian motion with $\hat{v}$-dependent diffusion constant (Fig. 12)

For Fig. 12 we describe the numerical method for the Brownian motion with the $\hat{v}$-dependent diffusion constant

governed by the SDE (266). The numerical simulation is based on the discrete version

$$\hat{v}(t_i + \Delta t_i) = \begin{cases} \hat{v}(t_i) + \sqrt{\tilde{g}[\hat{v}(t_i)]\Delta t_i}\hat{\xi}_i^G, & \hat{v}(t_i) \geqslant 0 \\ 0, & \hat{v}(t_i) < 0, \end{cases} \quad \text{(I7)}$$

with independent Gaussian random number $\hat{\xi}_i^G$. For Fig. 12 we employ the model

$$\tilde{g}(\hat{v}) = \lambda_0 e^{\beta\hat{v}}, \quad \Delta t_i = \min\left\{\Delta t_{max}^{(1)}, \frac{\Delta t_{max}^{(2)}}{\tilde{g}[\hat{v}(t_i)]}\right\}, \quad \text{(I8)}$$

with $\lambda_0 = 10^{-4}$, $\beta = 3$, $\Delta t_{max}^{(1)} = 0.1$, and $\Delta t_{max}^{(2)} = 0.01$. The total time of the simulation was $T = 5 \times 10^5$ and the total number of parallel threads was $N_{PC} = 4$.

---

[1] H. Scher and E. W. Montroll, Anomalous transit-time dispersion in amorphous solids, Phys. Rev. B **12**, 2455 (1975).

[2] H. Scher, H. G. Margolin, R. Metzler, J. Klafter, and B. Berkowitz, The dynamical foundation of fractal stream chemistry: The origin of extremely long retention times, Geophys. Res. Lett. **29**, 5-1 (2002).

[3] Y. Ogata, Statistical models for earthquake occurrences and residual analysis for point processes, J. Am. Stat. Assoc. **83**, 9 (1988).

[4] Y. Ogata, Seismicity analysis through point-process modeling: A review, Pure Appl. Geophys. **155**, 471 (1999).

[5] A. Helmstetter and D. Sornette, Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, J. Geophys. Res. **107**, ESE 10 (2002).

[6] S. Nandan, G. Ouillon, D. Sornette, and S. Wiemer, Forecasting the full distribution of earthquake numbers is fair, robust, and better, Seismol. Res. Lett. **90**, 1650 (2019).

[7] M. Feng, S.-M. Cai, M. Tang, and Y.-C. Lai, Equivalence and its invalidation between non-Markovian and Markovian spreading dynamics on complex networks, Nat. Commun. **10**, 3748 (2019).

[8] E. Errais, K. Giesecke, and L. R. Goldberg, Affine point processes and portfolio credit risk, SIAM J. Finan. Math. **1**, 642 (2010).

[9] A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel, Econophysics review: I. Empirical facts, Quant. Financ. **11**, 991 (2011).

[10] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, and D. Sornette, Multifractal analysis of financial markets: A review, Rep. Prog. Phys. **82**, 125901 (2019).

[11] D. Sornette, F. Deschatres, T. Gilbert, and Y. Ageon, Endogenous Versus Exogenous Shocks in Complex Networks: An Empirical Test Using Book Sale Rankings, Phys. Rev. Lett. **93**, 228701 (2004).

[12] R. Crane and D. Sornette, Robust dynamic classes revealed by measuring the response function of a social system, Proc. Natl. Acad. Sci. USA **105**, 15649 (2008).

[13] A. Hawkes, Point spectra of some mutually exciting point processes, J. R. Stat. Soc. B **33**, 438 (1971).

[14] A. Hawkes, Spectra of some self-exciting and mutually exciting point processes, Biometrika **58**, 83 (1971).

[15] A. Hawkes and D. Oakes, A cluster process representation of a self-exciting process, J. Appl. Probab. **11**, 493 (1974).

[16] K. Kanazawa and D. Sornette, Nonuniversal Power Law Distribution of Intensities of the Self-Excited Hawkes Process: A Field-Theoretical Approach, Phys. Rev. Lett. **125**, 138301 (2020).

[17] K. Kanazawa and D. Sornette, Field master equation theory of the self-excited Hawkes process, Phys. Rev. Res. **2**, 033442 (2020).

[18] G. I. Barenblatt, *Scaling, Self-Similarity, and Intermediate Asymptotics* (Cambridge University Press, Cambridge, 1996).

[19] P. Brémaud and L. Massoulié, Stability of nonlinear Hawkes processes, Ann. Probab. **24**, 1563 (1996).

[20] J.-P. Bouchaud, J. Bonart, J. Donier, and M. Gould, *Trades, Quotes and Prices* (Cambridge University Press, Cambridge, 2018).

[21] D. Sornette and G. Ouillon, Multifractal Scaling of Thermally Activated Rupture Processes, Phys. Rev. Lett. **94**, 038501 (2005).

[22] G. Ouillon and D. Sornette, Magnitude-dependent Omori law: Theory and empirical study, J. Geophys. Res. **110**, B04306 (2005).

[23] S. Nandan, G. Ouillon, J. Woessner, D. Sornette, and S. Wiemer, Systematic assessment of the static stress triggering hypothesis using interearthquake time statistics, J. Geophys. Res.: Sol. Ea. **121**, 1890 (2016).

[24] *Criticality in Neural Systems*, edited by D. Plenz and E. Niebur (Wiley, New York, 2014).

[25] I. Osorio, M. G. Frei, D. Sornette, J. Milton and Y.-C. Lai, Epileptic seizures: Quakes of the brain? Phys. Rev. E **82**, 021919 (2010).

[26] D. Sornette and I. Osorio, in *Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Physics and Engineering*, edited by I. Osorio, H. P. Zaveri, M. G. Frei, and S. Arthurs (CRC, Boca Raton, 2010), pp. 203–237.

[27] K. Kanazawa and D. Sornette, Ubiquitous Power Law Scaling in Nonlinear Self-Excited Hawkes Processes, Phys. Rev. Lett. **127**, 188301 (2021).

[28] P. Blanc, J. Donier, and J.-P. Bouchaud, Quadratic Hawkes processes for financial prices, Quant. Financ. **17**, 171 (2017).

[29] F. Gao and L. Zhu, Some asymptotic results for nonlinear Hawkes processes, Stoch. Process. Appl. **128**, 4051 (2018).

[30] S. Ciliberto, A. Guarino, and R. Scorretti, The effect of disorder on the fracture nucleation process, Physica D **158**, 83 (2001).

[31] R. Scorretti, S. Ciliberto, and A. Guarino, Disorder enhances the effects of thermal noise in the fiber bundle model, Europhys. Lett. **55**, 626 (2001).

[32] A. Saichev and D. Sornette, Andrade, Omori, and time-to-failure laws from thermal noise in material rupture, Phys. Rev. E **71**, 016608 (2005).

[33] A. Helmstetter and D. Sornette, Predictability in the epidemic-type aftershock sequence model of interacting triggered seismicity, J. Geophys. Res.: Sol. Ea. **108**, 2482 (2003).

[34] C. G. Bowsher, Modelling security market events in continuous time: Intensity based, multivariate point process models, J. Econ. **141**, 876 (2007).

[35] V. A. Filimonov and D. Sornette, Quantifying reflexivity in financial markets: Toward a prediction of flash crashes, Phys. Rev. E **85**, 056108 (2012).

[36] E. Bacry, I. Mastromatteo, and J.-F. Muzy, Hawkes processes in finance, Mark. Microstruct. Liq. **01**, 1550005 (2015).

[37] V. A. Filimonov and D. Sornette, Self-excited multifractal dynamics, Europhys. Lett. **94**, 46003 (2011).

[38] G. Soros, *The Alchemy of Finance*, 1st ed. (Simon & Schuster, New York, 1987); *ibid.*, 3rd ed. (Wiley, New York, 2015).

[39] E. Bacry, J. Delour, and J.-F. Muzy, Multifractal random walk, Phys. Rev. E **64**, 026103 (2001).

[40] A. Saichev and D. Sornette, Generic multifractality in exponentials of long memory processes, Phys. Rev. E **74**, 011111 (2006).

[41] A. I. Saichev and V. Filimonov, Numerical simulation of quasi-multifractal diffusion process, J. Exp. Theor. Phys. **107**, 324 (2008).

[42] C. W. Gardiner, *Handbook of Stochastic Methods*, 4th ed. (Springer, Berlin, 2009).

[43] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 1992).

[44] G. Ouillon and D. Sornette, Long-range static directional stress transfer in a cracked, nonlinear elastic crust, Future Gener. Comput. Syst. **22**, 500 (2006).

[45] K. Kanazawa, *Statistical Mechanics for Athermal Fluctuation: Non-Gaussian Noise in Physics* (Springer, Berlin, 2017).

[46] K. Kanazawa, T. G. Sano, T. Sagawa, and H. Hayakawa, Minimal Model of Stochastic Athermal Systems: Origin of Non-Gaussian Noise, Phys. Rev. Lett. **114**, 090601 (2015).

[47] K. Kanazawa, T. G. Sano, T. Sagawa, and H. Hayakawa, Asymptotic derivation of Langevin-like equation with non-Gaussian noise and its analytical solution, J. Stat. Phys. **160**, 1294 (2015).

[48] X. Gao and L. Zhu, Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues, Queueing Syst. **90**, 161 (2018).

[49] C. Van Den Broeck, On the relation between white shot noise, Gaussian white noise, and the dichotomic Markov process, J. Stat. Phys. **31**, 467 (1983).

[50] A. Dassios and H. Zhao, A dynamic contagion process, Adv. Appl. Probab. **43**, 814 (2011).

[51] H. Kesten, Random difference equations and renewal theory for products of random matrices, Acta Math. **131**, 207 (1973).

[52] D. Sornette, Linear stochastic dynamics with nonlinear fractal properties, Physica A **250**, 295 (1998).

[53] D. Sornette and R. Cont, Convergent multiplicative processes repelled from zero: Power laws and truncated power laws, J. Phys. (France) I **7**, 431 (1997).

[54] F. Gerhard, M. Deger, and W. Truccolo, On the stability and dynamics of stochastic spiking neuron models: Nonlinear Hawkes process and point process GLMs, PLoS Comput. Biol. **13**, e1005390 (2017).

[55] A. S. Cherny and H.-J. Engelbert, *Singular Stochastic Differential Equations*, Lecture Notes in Mathematics Vol. 1858 (Springer, Berlin, 2005), pp. 1–128.

[56] J. Krafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, 2011).