Standard form of master equations for general non-Markovian jump processes: The Laplace-space embedding framework and asymptotic solution

Kiyoshi Kanazawa
Department of Physics, Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

Didier Sornette
Institute of Risk Analysis, Prediction, and Management (Risks-X), Southern University of Science and Technology (SUSTech), Shenzhen 518055, China

(Received 31 January 2024; accepted 17 May 2024; published 12 June 2024)

We present a standard form of master equations (MEs) for general one-dimensional non-Markovian (history-dependent) jump processes, complemented by an asymptotic solution derived from an expanded system-size approach. The ME is obtained by developing a general Markovian embedding using a suitable set of auxiliary field variables. This Markovian embedding uses a Laplace-convolution operation applied to the velocity trajectory. We introduce an asymptotic method tailored for this ME standard, generalizing the system-size expansion for these jump processes. Under specific stability conditions tied to a single noise source, upon coarse graining, the generalized Langevin equation (GLE) emerges as a universal approximate model for point processes in the weak-coupling limit. This methodology offers a unified analytical tool set for general non-Markovian processes, reinforcing the universal applicability of the GLE founded in microdynamics and the principles of statistical physics.

DOI: 10.1103/PhysRevResearch.6.023270

I. INTRODUCTION

Non-Markovian stochastic processes have emerged as a powerful framework across diverse scientific disciplines, including physics [1], chemistry [2], econometrics [3], and financial modeling [4].

(1) Physics: Within statistical physics, particle motion in water is described by the generalized Langevin equation (GLE) [1]. The GLE represents a quintessential non-Markovian stochastic model, capturing the hydrodynamic memory effect.

(2) Econometrics: The autoregressive integrated moving average (ARIMA) model stands as a recognized discrete-time non-Markovian model describing many stylized structures of financial returns [3].

(3) Finance: The self-excited Hawkes process [5–7], a widely used non-Markovian point-process model, finds many applications in finance [8–11]. Here, points indicate event occurrences on the time axis.

(4) Other disciplines: The versatility of the non-Markovian self-exciting Hawkes process also extends to neuroscience [12], seismology [13–16], epidemiology [17,18], industrial and organizational psychology and sociology [19,20], criminology [21], and so on.

Central to these models is their ability to encapsulate long-memory effects inherent to various systems. This is typified by power-law decaying autocorrelation functions, transcending the conventional boundaries set by Markovian stochastic processes.

A well-established analytical toolkit has been developed for Markovian stochastic processes [22,23], which includes stochastic differential equations (SDEs), master equations (MEs), and their asymptotic solutions. For instance, the theory of standard forms has been instrumental in the systematic classification of both SDEs and MEs. Given that MEs represent linear time-evolution equations for probability density functions and functionals (PDFs), they can be solved within the framework of linear algebra, particularly through methods like the eigenfunction expansion [22,23].

There are also various asymptotic methods tailored to MEs. Prominent among these are the system-size expansion [2,24–26] and the Wentzel-Kramers-Brillouin approximation [23,27]. Notably, the system-size expansion stands as a historic cornerstone in the realm of statistical physics, especially concerning the Langevin equations. This is largely due to its role in extrapolating various Langevin equations from underlying microscopic physical dynamics. Hence, Markovian process theory offers a robust and structured foundation for statistical physics, at least in a formal sense.

In contrast to the structured theories for Markovian processes, those for non-Markovian processes remain more fragmented. A universally accepted ME theory for non-Markovian processes is absent. Current MEs pertain specifically to particular non-Markovian SDE classes, such as GLE with exponential memories [28,29], GLE with linear potential [30], and semi-Markovian point processes [31]. Without a standardized form for these MEs, corresponding asymptotic methods for general non-Markovian processes have yet to emerge. Hence, developing a systematic theory for non-Markovian processes remains a long-standing challenge in statistical physics.

Our prior studies have offered partial solutions to this challenge, specifically for linear and nonlinear Hawkes processes [32–35]. We have generalized the Markovian embedding approach, transforming a non-Markovian process into a Markovian field dynamic. Within this framework, the MEs for the Hawkes processes are conceived as time-evolution equations for the PDFs of auxiliary field variables. We refer to these equations as *field master equations* (fMEs). While we regard this methodology as a potential avenue for generating MEs for a broader range of non-Markovian processes, its scope, for now, remains confined to certain models, notably the nonlinear extensions of the Hawkes point process family.

In this paper, we focus on deriving the ME for the general class of one-dimensional non-Markovian jump processes, a subset of the broader point process family. Our approach frames the general one-dimensional non-Markovian jump process as a history-dependent jump process. As a versatile model, it can incorporate any form of historical dependency and represents the most comprehensive one-dimensional non-Markovian jump process conceivable by us. To tackle these processes, we develop a general Markovian embedding using a suitable set of auxiliary field variables. This Markovian embedding uses a Laplace-convolution operation applied to the velocity trajectory and allows us to derive the corresponding fME. Given the capability of this ME to handle all forms of one-dimensional non-Markovian properties, we suggest that it constitutes a standard ME form for general one-dimensional non-Markovian jump processes. Additionally, we introduce an asymptotic method tailored for this ME standard, generalizing the system-size expansion for this jump process. Under specific stability conditions tied to a single noise source, the GLE emerges as a universal approximate model for point processes in the weak-coupling limit. This methodology offers a unified analytical tool for general non-Markovian processes, underpinned by strong statistical physics validating the GLE's universal applicability.

This paper is organized as follows. Section II presents our mathematical notations. Section III gives a concise review of the theories of Markovian stochastic processes, of the corresponding standard form of the ME, and of the system-size expansion. Section III is prepared for readers unfamiliar with Markovian stochastic processes; expert readers can skip it to go directly to the main part from Sec. IV. Section IV introduces our model and derives the corresponding fME via the Laplace-convolution Markovian embedding. Section V describes the system-size expansion for the non-Markovian jump processes that allows us to asymptotically derive the GLE. Section VI demonstrates another application of our formalism to a financial-pricing model based on the nonlinear Hawkes processes. We conclude this paper after discussing implications and future perspectives of our work in Sec. VII. Eight appendices supplement the main text on technical issues.

II. MATHEMATICAL NOTATION

Let us describe our mathematical notation regarding stochastic variables, sets, and functionals.

A. Notation for stochastic variables

Any stochastic variable carries the hat symbol in the form $\hat{A}$ to distinguish it from the real number A. The PDF is denoted by $P_t(A) := P(\hat{A}_t = A)$, implying that the probability for $\hat{A}_t \in [A, A + dA)$ is given by $P_t(A)\, dA$. The ensemble average of any stochastic variable $\hat{A}$ is written as $\langle \hat{A}_t \rangle := \int A P_t(A)\, dA$. Using this notation, the PDF can be rewritten as $P_t(A) = \langle \delta(A - \hat{A}_t) \rangle$ with the Dirac δ function (see Appendix A).

B. Notation for sets

The set of real numbers and the set of positive integers are denoted by $\boldsymbol{R}$ and $\boldsymbol{N}$. The set of positive real numbers is denoted by $\boldsymbol{R}^+ := \{s \mid s > 0, s \in \boldsymbol{R}\}$. Here s typically represents the wave number, which should be a real positive number, and we introduce the compact notation

$$\{z(s)\}_s := \{z(s) \mid s \in \boldsymbol{R}^+\}. \tag{1}$$

Also, i and j typically represents integers, and we also introduce the corresponding compact notation

$$\{a_i\}_i := \{a_i \mid i \in \boldsymbol{N}\}. \tag{2}$$

C. Notation for functionals

If the argument of a map f is a function $\{z(s)\}_s$, f is called a *functional*. A functional is indicated by the square brackets $f[\{z(s)\}_s]$. The functional notation $f[\{z(s)\}_s]$ is sometimes abbreviated as $f[z]$ if its meaning is obvious from the context. For a stochastic field variable $\{\hat{z}_t(s)\}_s$, the corresponding PDF is written as $P_t[z] = P_t[\{z(s)\}_s]$, characterizing the probability $P_t[z]\mathcal{D}z$ that $\{\hat{z}(s)\}_s \in \prod_s [z(s), z(s) + dz(s))$, where the functional volume element is $\mathcal{D}z := \prod_s dz(s)$. The ensemble average of any functional $f[\hat{z}_t]$ is written as the path-integral representation

$$\langle f[\hat{z}_t] \rangle := \int f[z] P_t[z] \mathcal{D}z. \tag{3}$$

On the basis of this notation, the PDF is formally rewritten by $P_t[z] = \langle \delta[z - \hat{z}_t] \rangle$, where the δ functional is defined by $\delta[z - \hat{z}_t] := \prod_{s \in \boldsymbol{R}^+} \delta(z(s) - \hat{z}_t(s))$. The concept of derivative can be generalized to the functional derivative, which is denoted by $\delta f[z]/\delta z(s)$ (see Appendix A for the detail).

III. LITERATURE REVIEW: MARKOVIAN STOCHASTIC PROCESSES

This section offers a concise overview of the foundational theory of Markovian processes, serving as an introduction for readers less acquainted with Markovian processes and statistical physics. Specifically, we touch upon the standard form of the ME for these processes. Experts who are solely focused on our primary findings may bypass this section, as the main results are presented in a stand-alone, comprehensive format.

A. Review of Markovian stochastic differential equations

Let us consider a one-dimensional stochastic process characterized by the trajectory $\{\hat{v}_s\}_{s\leqslant t}$, where t is the current time. If the statistics of the infinitesimal future state $\hat{v}_{t+\mathrm{d}t}$ is completely characterized only by the current state $\hat{v}_t$, the stochastic dynamics is said to obey a *Markovian stochastic process*. However, a more general class of stochastic models can be considered that cannot be characterized only by the current state $\hat{v}_t$. For example, a stochastic model can depend on the full history $\{\hat{v}_s\}_{s\leqslant t}$. Such stochastic dynamics obey a *non-Markovian stochastic process*. This subsection reviews the theory of Markovian stochastic processes, in particular, their standard forms and the system-size expansion.

1. Review of the standard form of white noise

White noise is a noise that is independent of its history. It is formally defined as the derivative of the Lévy process $\hat{L}_t$, such that $\hat{\xi}_t^{\mathrm{W}} := \mathrm{d}\hat{L}_t/\mathrm{d}t$. According to the Lévy-Itô decomposition, any white noise can be decomposed as the sum of the white Gaussian noise and of the white Poisson noise (see Appendix B for a review), such that

$$\hat{\xi}^{\mathrm{W}} = m + \sigma\hat{\xi}^{\mathrm{G}} + \hat{\xi}^{\mathrm{P}}_{\lambda(y)}, \tag{4}$$

where m is the constant drift, σ is the standard deviation of the white noise, $\hat{\xi}^{\mathrm{G}}$ is the white noise, and $\hat{\xi}^{\mathrm{P}}_{\lambda(y)}$ is the white Poisson noise with intensity distribution function $\lambda(y)$ of the jump size y.

2. Review of the standard form of stochastic differential equations

Any one-dimensional stochastic process can be constructed from white noise by introducing the *state dependence* into the drift term a, standard deviation σ, and the intensity distribution function $\lambda(y)$, such that

$$m \to m(\hat{v}), \quad \sigma \to \sigma(\hat{v}), \quad \lambda(y) \to \lambda(y|\hat{v}). \tag{5}$$

A one-dimensional stochastic Markovian process $\{\hat{v}_s\}_{s\leqslant t}$ obeys the state-dependent SDE

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = m(\hat{v}_t) + \sigma(\hat{v}_t)\hat{\xi}^{\mathrm{G}} + \hat{\xi}^{\mathrm{P}}_{\lambda(y|\hat{v}_t)}, \tag{6}$$

with the Itô interpretation assumed. We refer to this representation as the *standard form of one-dimensional Markovian SDEs*. The Markovian property is expressed by the fact that the right-hand side of Eq. (6) depends only on the current state $\hat{v}_t$.

3. Review of the standard form of master equations

The SDE (6) describes the dynamics of stochastic systems for a single path. While the SDE are intuitive tools, they are not easy to handle because of their general nonlinear structure. For analytical calculations, the ME approach provides more systematic methods based on linear algebra. The ME is the equation governing the time evolution of the PDF $P_t(v) := \langle\delta(v-\hat{v}_t)\rangle$ as follows:

$$\frac{\partial P_t(v)}{\partial t} = \mathcal{L}P_t(v), \tag{7}$$

with a linear operator $\mathcal{L}$. The ME corresponding to the SDE (6) is given by

$$\frac{\partial P_t(v)}{\partial t} = \left[-\frac{\partial}{\partial v}m(v) + \frac{1}{2}\frac{\partial^2}{\partial v^2}\sigma^2(v)\right]P_t(v) + \int_{-\infty}^{\infty}\mathrm{d}y[\lambda(y|v-y)P_t(v-y) - \lambda(y|v)P_t(v)]. \tag{8}$$

This ME is known to cover all possible one-dimensional Markovian stochastic processes with mild assumptions [22], and thus is called the *standard form of the master equation* in this report. The ME is very useful because it is always a linear dynamical equation[1] of the PDF $P_t(v)$. In other words, a standard approach to a given Markovian process is to consider its ME and solve the corresponding eigenvalue problem with linear algebra techniques.

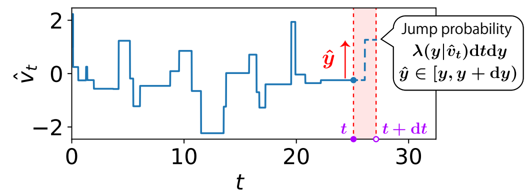

FIG. 1. Schematic of the Markovian jump process (9). The path is piecewisely continuous and occasionally has jumps. A jump occurs during $[t, t+\mathrm{d}t)$ with probability $\lambda(y|\hat{v}_t)\mathrm{d}t\mathrm{d}y$ with jump size $\hat{y} \in [y, y+\mathrm{d}y)$. The remarkable character of the Markovian jump process is that the intensity density $\lambda(y|\hat{v}_t)$ depends only on the current state $\hat{v}_t$ and does not depend on the whole history $\{\hat{v}_\tau\}_{\tau<t}$. In this sense, the Markovian jump process is a history-independent Poisson process, in contrast to the non-Markovian jump process (or the history-dependent Poisson process) defined as Eq. (16).

4. Review of the Markovian jump process (history-independent Poisson process)

Markovian jump processes constitute a large subclass of Markovian SDEs, such that

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \hat{\xi}^{\mathrm{P}}_{\lambda(y|\hat{v}_t)}, \quad \hat{\xi}^{\mathrm{P}}_{\lambda(y|\hat{v}_t)} = \sum_{k=1}^{\hat{N}(t)}\hat{y}_k\delta(t-\hat{t}_k), \tag{9}$$

where the drift term and the white Gaussian noise term are absent, and only the jump term is present (see Fig. 1). Here $\lambda(y|\hat{v}_t)$ is the conditional intensity density of the jump size y, $\hat{N}(t)$ is the total number of jumps during $[0,t)$, and $\hat{t}_k$ is the kth jump time, $\hat{y}_k$ is the kth jump size. Markovian jump processes depend only on the current state $\hat{v}_t$ and can be called *history-independent Poisson processes*, in contrast to the non-Markovian jump processes (or history-dependent Poisson processes) defined in Eq. (16) in the main section.

[1]Indeed, the formal solution of the ME (7) is given by

$$P_t(v) = \sum_i c_i e^{-\mu_i t}\phi_i(v), \quad \mathcal{L}\phi_i(v) = -\mu_i\phi_i(v).$$

with initial condition constants $\{c_i\}_i$, where μ_i and $\phi_i(v)$ are the ith eigenvalue and the corresponding eigenfunction, respectively.

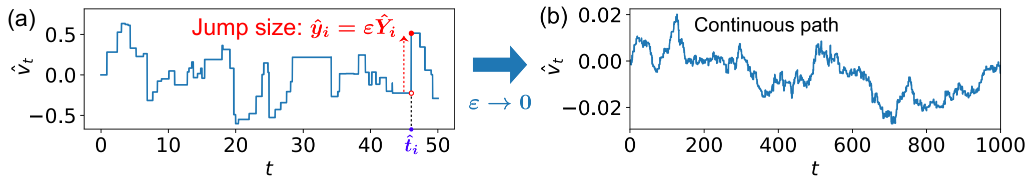

FIG. 2. Schematic of the system-size expansion. (a) A typical trajectory of a Markovian jump process (9) where the jump size is scaled with ε, such that $\hat{y}_i = \varepsilon\hat{Y}_i$ with ε-independent jump size $\hat{Y}_i$. (b) For small $\varepsilon \ll 1$, the path becomes approximately continuous due to the small jump sizes and obeys the Langevin equation (12) under an appropriate stability condition around $\hat{v}_t \simeq 0$. This picture essentially applies even to the non-Markovian jump process (16) as shown in Sec. V.

Markovian jump processes are popular models. For instance, the detailed description of physical Brownian motion is often modeled as a Markovian jump process for the velocity of the Brownian particle, where the velocity discontinuously changes due to molecular collisions.

B. Review of the system-size expansion

Solving the eigenvalue problem is, in general, difficult, in particular when the linear operator $\mathcal{L}$ leads to an integrodifferential equation of the form (8). One of the systematic methods to obtain asymptotic solutions was invented by van Kampen, which is called the *system-size expansion*. Mathematically, the system-size expansion can be regarded as a weak-noise asymptotic limit for Markovian jump processes. This assumption is very natural, particularly in the context of Brownian motions. This method provides a solid mathematical derivation of the Langevin equations from microscopic physical dynamics. In this subsection, we briefly review this methodology based on Refs. [24–26].

1. Sketch of the system-size expansion

Let us consider a Markovian jump process described by (9). In addition, let us assume that the jump size is proportional to a small positive parameter $\varepsilon > 0$ [see Fig. 2(a)], so we can write

$$\hat{y}_k = \varepsilon\hat{Y}_k. \tag{10}$$

This implies that the noise term can be rewritten as $\hat{\xi}^{\mathrm{P}}_{\lambda(y|\hat{v}_t)} = \varepsilon\hat{\xi}^{\mathrm{P}}_{W(Y|\hat{v}_t)}$, with the conditional intensity density $W(Y|\hat{v}_t)$ for the rescaled jump size Y. We thus obtain the SDE with a small jump-noise term:

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \epsilon\hat{\xi}^{\mathrm{P}}_{W(Y|\hat{v}_t)}. \tag{11}$$

This is the scaling assumption for the system-size expansion. In other words, the small parameter ϵ can be interpreted as the small constant quantifying the weak-coupling with the stochastic environment.

In the small-noise asymptotic limit $\varepsilon \to 0$ and for a broad variety of setups, assuming a stability condition around $\hat{v}_t \simeq 0$ (see Appendix C for details), the Markovian jump process reduces to the Langevin equation [see Fig. 2(b) for a schematic]

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = -\gamma\hat{v}_t + \sqrt{2\gamma T}\hat{\xi}^{\mathrm{G}}, \tag{12}$$

where γ takes the meaning of a frictional constant and T is the temperature. See Appendix C for the detailed derivation. Thus, the system-size expansion is a celebrated mathematical foundation for the derivations of the Langevin equations from microscopic dynamics.

After applying the system-size expansion, remarkably, the white Gaussian noise term emerges even though we have started our discussion with the compound Poisson noise. This result can be interpreted within the framework of the law of large numbers and the central limit theorem. In the detailed derivation (see Appendix C), the introduction of the rescaled time $\mathrm{t} := \varepsilon t$ is a technical key. Assuming $\mathrm{t} = O(1)$ [or, equivalently, $t = O(\varepsilon^{-1})$], there are sufficiently many collisions during $[0, t)$ so the law of large numbers and the central limit theorem are applicable: the deterministic drift term appears as the result of the law of large numbers, and the white Gaussian noise appears as the result of the central limit theorem. In this sense, the system-size expansion is the formal application of the law of large numbers and the central limit theorem to the Markovian jump processes.

2. Physical validity of the scaling assumption

The physical validity of the scaling (10) and (11) can be intuitively understood by considering a one-dimensional collision problem (Fig. 3). Let us prepare a small particle of mass m, velocity v and a large particle of mass M and velocity V. In a one-dimensional elastic collision, the postcollisional velocity V' of the large particle is given by

$$V' - V = \frac{2m}{m+M}(v - V). \tag{13}$$

Since the typical thermal velocities are given by $v \simeq \sqrt{T/m}$ and $V \simeq \sqrt{T/M}$, where T is the gas temperature, we have $|V/v| \propto \epsilon^{1/2} \ll 1$ with $\epsilon := m/M$. For $\epsilon \ll 1$, we obtain the velocity jump y of the large particle as

$$y := V' - V \simeq 2\epsilon v. \tag{14}$$

Thus, the velocity-jump size is proportional to ε and satisfies the system-size expansion scaling (11) exactly. This example highlights that the scaling assumption of the system-size expansion is physically reasonable[2] when the Brownian particle in a gas is much heavier than the surrounding gas particles.

[2]Note that the assumption that the dynamics is Markovian is also valid for Brownian dynamics in the dilute-gas limit.

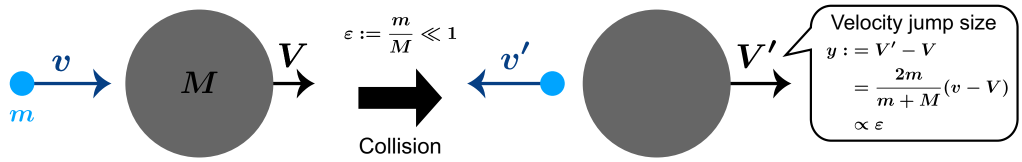

FIG. 3. The scaling assumption of the system-size expansion is natural on physical grounds for the description of a massive Brownian-particle motion. Let us denote the mass and initial velocity of the Brownian particle (respectively, of a surrounding small particle) by M (respectively, m) and $\hat{V}$ (respectively, $\hat{v}$). After an elastic collision, the velocity of the Brownian particle changes according to formula (13) deriving from the conservation of momentum. The velocity jump size y is proportional to the mass ratio $\varepsilon := m/M$, which is the small parameter of the problem in the limit of massive Brownian particle limit $m \ll M$.

3. Scaling assumption in the master equations

The scaling assumption $y = \varepsilon Y$ at a trajectory level is equivalent to the scaling assumption for the ME,

$$\lambda(y|v) = \frac{1}{\varepsilon} W\left(\frac{y}{\varepsilon}\middle|v\right), \tag{15}$$

which is derived from the conservation of probability (i.e., the Jacobian relation), such that $\lambda(y|v)\mathrm{d}y = W(Y|v)\mathrm{d}Y$.

C. Goal of this paper

On the basis of the above theory regarding the standard forms of SDEs and MEs, our goals in this paper are the following:

(1) We derive the ME for the general non-Markovian jump process analogous to the standard form of the Markovian ME (8).

(2) We asymptotically solve the ME for non-Markovian jump processes by generalizing the system-size expansion; we finally obtain the GLE via a physically reasonable coarse-graining approach.

IV. MAIN RESULT: NON-MARKOVIAN MODEL AND FORMULATION

In this section, we first present the stochastic model studied in this paper. We then introduce the Laplace-convolution Markovian embedding that converts the original low-dimensional non-Markovian dynamics onto a Markovian field dynamics. Finally, the corresponding fME is formulated.

A. Non-Markovian jump process (history-dependent compound Poisson process)

Let us consider a non-Markovian stochastic model that can encompass a large class of non-Markovian stochastic processes:

> We study the history-dependent compound Poisson process (see Fig. 4 for a schematic),
>
> $$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \hat{\xi}^{\mathrm{CP}}_{\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})}, \tag{16}$$
>
> where the intensity $\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})$ with jump size y is conditional on the full history of the system $\{\hat{v}_\tau\}_{\tau\leqslant t}$.

The non-Markovian nature of this process makes the intensity a functional of the whole history. More technically, Eq. (16) implies that

$$\mathrm{d}\hat{v}_t := \begin{cases} \hat{y} + O(\mathrm{d}t) & (\mathrm{Prob} = \mathrm{d}t\, \mathrm{d}y \lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t}) \text{ for any } \hat{y} \in [y, y+\mathrm{d}y)) \\ 0 & \left(\mathrm{Prob} = 1 - \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}y \lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})\right) \end{cases} \tag{17}$$

for any given history $\{\hat{v}_\tau\}_{\tau\leqslant t}$ with infinitesimal time evolution $\mathrm{d}\hat{v}_t := \hat{v}_{t+\mathrm{d}t} - \hat{v}_t$. Our aim is to provide the full analytical tool set for this history-dependent Poisson process by developing the corresponding fME and by analyzing its asymptotic solutions.

Throughout Sec. IV, we do not make strong assumptions and consider the general class of non-Markovian jump processes, including both stationary and nonstationary cases. On the other hand, some of the results in other sections (such as the system-size expansion in Sec. V) are only available for stationary processes.

B. Markovian embedding

In this subsection, we apply the Markovian-embedding scheme to the history-dependent Poisson process (16). We finally obtain the stochastic partial differential equation (SPDE) governing the Markovian field dynamics and derive the corresponding fME.

1. Basic idea

The idea of Markovian embedding is very simple: a low-dimensional non-Markovian dynamics can be converted onto a higher-dimensional Markovian dynamics by adding a

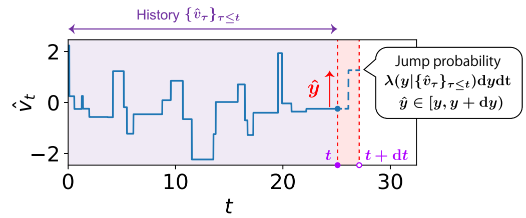

FIG. 4. Schematic of the non-Markovian jump process (or history-dependent compound Poisson process). The probability that a velocity jump occurs during $[t, t+\mathrm{d}t)$ is given by $\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})\mathrm{d}t\mathrm{d}y$ with jump size $\hat{y}\in[y, y+\mathrm{d}y)$. Here the intensity density $\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})$ explicitly depends on the whole history $\{\hat{v}_\tau\}_{\tau\leqslant t}$ and the system is thus truly non-Markovian.

sufficient number of auxiliary variables. In the statistical-physics context, this approach dates back to Mori [36] around the mid-1960s.[3] Also, the theory of the Kac-Zwanzig model [28,29,37–39] can be regarded as a theory of Markovian embedding between the generalized Langevin equation and the Hamiltonian-particle model with harmonic interaction. For example, the generalized Langevin equation with the sum of K-exponential memories can be thought of as a K-dimensional Markovian dynamics [28,29,33]. This idea can even be applied to the Hawkes processes [4,40,41] for memory kernels expressed as a sum of exponential functions. Remarkably, this idea of Markovian embedding has been also applied to non-Markovian stochastic processes in quantum systems [42–48] in the context of the pseudomode approach around the mid-1990s.

The dimension needed for the Markovian embedding depends on the model but can be infinite in general. In this case, the dynamics can be regarded as a Markovian field dynamics. For instance, the GLE and the Hawkes processes have been converted onto Markovian field dynamics [32–35], which can be analyzed by the fME (which is a functional-differential equation for the probability density functional).

Markovian embedding is nontrivial and technically tricky for continuous-time stochastic processes, while Markovian embedding is rather straightforward for discrete-time stochastic processes (see Appendix E for brief clarification). This paper aims at formulating a general embedding theory of the non-Markovian jump process (16), even though it is based on continuous time.

2. Variable set

Before proceeding with the derivation of the Markovian field dynamics, let us introduce a complete set of system variables useful for Markovian embedding. In the previous section, we used $\{\hat{v}_\tau\}_{\tau\leqslant t}$ as a naive complete set of system variables. This set is equivalent in information content to another set $(\hat{v}_t; \{\hat{a}_\tau\}_{\tau<t})$, with acceleration $\hat{a}_t := \mathrm{d}\hat{v}_t/\mathrm{d}t$. Note that the acceleration can include the impulses described by the Dirac δ functions associated with the jumps. Let us now introduce the *Laplace-convolution Markovian-embedding representation* of the velocity trajectory as

$$\hat{z}_t(s) := \int_0^\infty e^{-s\tau}\hat{a}_{t-\tau}\,\mathrm{d}\tau, \quad \hat{a}_t := \frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t}, \tag{18}$$

which is defined for $s > 0$ (see Fig. 5 for a schematic). We then adopt the variable set

$$(\hat{v}_t, \{\hat{z}_t(s)\}_{s>0}) \tag{19}$$

as a useful complete variable set.

The introduction of the auxiliary field variable $\{\hat{z}_t(s)\}_{s>0}$ is the technical but crucial trick to convert the general nonlinear non-Markovian model onto a Markovian field model. In the following, the wave number s is always considered strictly positive ($s > 0$), and the set $\{f(s)\}_{s>0}$ for any function $f(s)$ is sometimes abbreviated by $\{f(s)\}_s$ if its meaning is clear from the context.

[3] He proposed a systematic expansion of the relaxation memory kernel by the continued-fraction expansion. Truncating the expansion leads to an approximation based on the sum of several exponential memories.

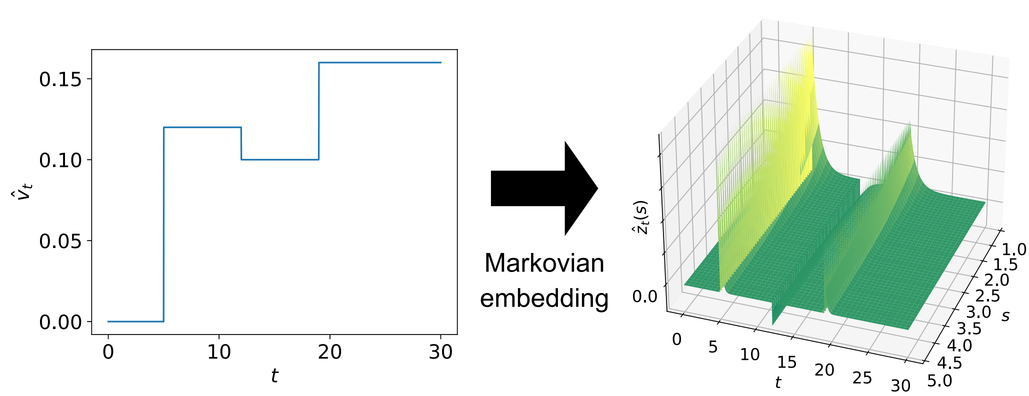

FIG. 5. Schematic of the Markovian embedding of the original one-dimensional non-Markovian jump process (16) onto the Markovian field dynamics $\{\hat{z}_t(s)\}_s$. The auxiliary field variables $\{\hat{z}_t(s)\}_s$ are defined by Eq. (18) on the wave-number axis $s\in(0,\infty)$ (i.e., one-dimensional field) and obey the first-order Markovian SPDE (23a).

3. Phase space

Let us introduce the state variables $\hat{\Gamma}_t$ as points in the phase space $\boldsymbol{S}$, such that

$$\hat{\Gamma}_t := (\hat{v}_t, \{\hat{z}_t(s)\}_s) \in \boldsymbol{S}, \quad \boldsymbol{S} := \{(v, \{z(s)\}_s) : v \in \boldsymbol{R}, z(s) \in \boldsymbol{F}\}, \tag{20}$$

where $\boldsymbol{R}$ is the space of real numbers and $\boldsymbol{F}$ is the function space. In the following, we simplify the notation of functionals such as the intensity as a functional in terms of the history in the following way:

$$\lambda[y|\hat{\Gamma}_t] := \lambda[y|\hat{v}_t; \hat{z}_t] := \lambda(y|\hat{v}_t, \{\hat{a}_\tau\}_{\tau\leqslant t}) := \lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t}). \tag{21}$$

In our notation, the functional argument (e.g., $\hat{\Gamma}_t$ and $\{\hat{z}_t(s)\}_s$) follows other variables (e.g., y and $\hat{v}_t$) after the separation by the semicolon.

Note that the ordinary Markov compound Poisson process corresponds to the case where the intensity does not depend on the historical velocities, such that

$$\lambda[y|\hat{v}_t; \hat{a}_t] = \lambda(y|\hat{v}_t), \tag{22}$$

with a nonnegative function $\lambda(y|\hat{v}_t)$.

C. Markovian field dynamics

The history-dependent compound Poisson process (16) characterizing the original variables $\{\hat{v}_\tau\}_{\tau\leqslant t}$ is equivalent to the set of the following SDE and SPDEs characterizing the new variables $\hat{\Gamma}_t := (\hat{v}_t, \{\hat{z}_t(s)\}_s)$:

$$\frac{d\hat{v}_t}{dt} = \hat{\xi}^{\mathrm{CP}}_{\lambda[y|\hat{\Gamma}_t]}, \quad \frac{\partial \hat{z}_t(s)}{\partial t} = -s\hat{z}_t(s) + \hat{\xi}^{\mathrm{CP}}_{\lambda[y|\hat{\Gamma}_t]}, \tag{23a}$$

where the jump term $\hat{\xi}^{\mathrm{CP}}_{\lambda[y|\hat{\Gamma}_t]}$ simultaneously acts on both $\hat{v}_t$ and $\hat{z}_t(s)$ for all $s > 0$ (see Fig. 5 for a typical configuration of the auxiliary field). The initial condition is given by

$$\hat{z}_0(s) = \int_0^\infty e^{-s\tau}\hat{a}_{-\tau}\,d\tau, \quad \hat{a}_\tau := \frac{d\hat{v}_\tau}{d\tau}. \tag{23b}$$

This set of SDEs characterizes the complete dynamics of the phase point $\hat{\Gamma}_t = (\hat{v}_t, \{\hat{z}_t(s)\}_s)$ in a closed form.

This is the first main result of this paper, stating that the original non-Markovian SDE is converted onto a Markovian field dynamics for the auxiliary variables. In other words, the non-Markovian memory effect has been removed in the extended space by considering all the components in the Laplace space.

The essence of the trick is to use the Laplace-convolution transform, which encodes the whole history of $\hat{v}_t$ (or, equivalently, its acceleration $\hat{a}_t$) into a function of t (now, thus Markovian) and of an additional variable s. The function $\hat{z}_t(s)$ dependent on s serves as the key device to render the system Markovian, utilizing an infinite series of equations for all $\hat{z}_t(s)$. It is remarkable that this system is Markovian in the extended phase space $\hat{\Gamma}_t \in \boldsymbol{S}$, while the original one-dimensional process is non-Markovian. This means that we have successfully transformed the original non-Markovian dynamics into a Markovian dynamics by adding a sufficient number of variables. Since the resulting dynamics is Markovian, we can derive the corresponding ME for the PDF for the phase point $\hat{\Gamma}_t$ in the extended phase space.

1. Derivation

By directly solving Eq. (23a), we obtain

$$\begin{aligned}
\hat{z}_t(s) &= \hat{z}_0(s)e^{-st} + \int_0^t e^{-s(t-t')}\hat{\xi}^{\mathrm{CP}}_{t',\lambda[y|\hat{\Gamma}_t]}\,dt' \\
&= \int_0^\infty e^{-s(\tau+t)}\hat{a}_{-\tau}\,d\tau + \int_0^t e^{-s(t-t')}\hat{a}_{t'}\,dt' \\
&= \int_t^\infty e^{-s\tau'}\hat{a}_{t-\tau'}\,d\tau' + \int_0^t e^{-st''}\hat{a}_{t-t''}\,dt'' \\
&= \int_0^\infty e^{-s\tau}\hat{a}_{t-\tau}\,d\tau,
\end{aligned} \tag{24}$$

with the dummy-variable transformation $\tau' := t + \tau$ and $t'' := t - t'$ from the second to the third line. Thus, the set of the SDEs (23a) is consistent with the Markovian-embedding representation (18).

2. Interpretation of the Laplace embedding

Here, we offer an intuitive interpretation of the Laplace embedding from multiple perspectives. First, the Laplace embedding bears resemblance to the exponential moving average (EMA)

$$\hat{a}_t^{(\mathrm{EMA})}(T) := \frac{1}{T}\int_0^\infty e^{-\tau/T}\hat{a}_{t-\tau}\,d\tau = s\hat{z}_t(s) \tag{25}$$

with the characteristic timescale $T := 1/s$. In finance, the EMA finds widespread use in technical analysis, particularly in detecting financial price trends. The auxiliary variable $\hat{z}_t(s)$ encapsulates the same information as the EMA, with a characteristic timescale of $1/s$. In essence, the Laplace embedding implies that complete memory effects are accounted for by incorporating the EMA with diverse timescales.

Second, the effectiveness of the Laplace embedding stems from its ability to transform the original non-Markovian dynamics into simultaneous first-order SPDEs. There is thus a perfect equivalence between the first-order ordinary-differential equation (ODE) and its general integral solution:

$$\frac{dx_t}{dt} = -sx_t + f_t \quad \Longleftrightarrow \quad x_t = x_0 + \int_0^t e^{-s(t-\tau)}f_s\,d\tau. \tag{26}$$

This formula is essentially equivalent to the Laplace-convolution transformation by assuming consistent initial values. Inversely, the Laplace-convolution transformation can be rewritten as an ODE. We have utilized this fact to convert the original non-Markovian dynamics onto a first-order SPDE.

D. Field master equation for the history-dependent compound Poisson process

The functional ME of the field corresponding to the SPDEs (23a) is given by

$$\frac{\partial P_t[\Gamma]}{\partial t} = \mathcal{L}P_t[\Gamma] := \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma]) + \int_{-\infty}^\infty \mathrm{d}y\{\lambda[y|\Gamma - \Delta\Gamma_y]P_t[\Gamma - \Delta\Gamma_y] - \lambda[y|\Gamma]P_t[\Gamma]\}, \tag{27}$$

with jump size vector

$$\Delta\Gamma_y := (y, \{y\mathbf{1}(s)\}_s), \tag{28}$$

with indicator function $\mathbf{1}(s) = 1$ for any s.

This is the second main result of this paper. The fME (27) determines the PDF of the auxiliary field dynamics, generally applicable to non-Markovian jump processes.

1. Derivation

We derive the fME as follows: For any functional $f[\hat{\Gamma}_t] := f[\hat{v}_t; \hat{z}_t] = f(\hat{v}_t, \{\hat{z}_t(s)\}_s)$, from Eq. (23a), its path-level differential $\mathrm{d}f[\hat{\Gamma}_t] := f[\hat{\Gamma}_{t+dt}] - f[\hat{\Gamma}_t]$ is given by

$$\mathrm{d}f[\hat{\Gamma}_t] = \begin{cases} f[\hat{\Gamma}_t + \Delta\Gamma_{\hat{y}}] - f[\hat{\Gamma}_t] + O(\mathrm{d}t) & (\text{Prob} = \mathrm{d}t\, \mathrm{d}y\lambda[y|\hat{\Gamma}_t] \text{ for any } \hat{y} \in [y, y+\mathrm{d}y)) \\ -\,\mathrm{d}t \int_0^\infty s\hat{z}_t(s)\frac{\delta f[\hat{\Gamma}_t]}{\delta \hat{z}_t(s)}\mathrm{d}s + O(\mathrm{d}t^2) & \left(\text{Prob} = 1 - \mathrm{d}t \int_{-\infty}^\infty \mathrm{d}y\lambda[y|\hat{\Gamma}_t]\right) \end{cases} \tag{29}$$

at leading order.[4] By taking the ensemble average of both sides, up to the order of $\mathrm{d}t$, we obtain

$$\langle \mathrm{d}f[\hat{\Gamma}_t]\rangle = \mathrm{d}t \int \mathrm{d}\Gamma P_t[\Gamma]\left[\int_{-\infty}^\infty \mathrm{d}y\lambda[y|\Gamma](f[\Gamma + \Delta\Gamma_y] - f[\Gamma]) - \int_0^\infty \mathrm{d}s\left\{sz(s)\frac{\delta}{\delta z(s)}f[\Gamma]\right\}\right] + O(\mathrm{d}t^2) \tag{30}$$

with integral volume element $\mathrm{d}\Gamma := \mathrm{d}v\mathcal{D}z$. By applying a variable transformation $\Gamma + \Delta\Gamma_y = \Gamma'$, the first term on the right-hand side is given by

$$\int \mathrm{d}\Gamma P_t[\Gamma]\lambda[y|\Gamma]f[\Gamma + \Delta\Gamma_y] = \int \mathrm{d}\Gamma' P_t[\Gamma' - \Delta\Gamma_y]\lambda[y|\Gamma' - \Delta\Gamma_y]f[\Gamma'] = \int \mathrm{d}\Gamma P_t[\Gamma - \Delta\Gamma_y]\lambda[y|\Gamma - \Delta\Gamma_y]f[\Gamma], \tag{31}$$

where the dummy variable Γ' is finally replaced with Γ. By applying the functional partial integration (A17), the third term on the right-hand side of Eq. (30) is given by

$$\int sz(s)P_t[\Gamma]\frac{\delta f[\Gamma]}{\delta z(s)}\mathrm{d}\Gamma = -\int f[\Gamma]\frac{\delta}{\delta z(s)}\{sz(s)P_t[\Gamma]\}\,\mathrm{d}\Gamma. \tag{32}$$

By considering that the left-hand side of Eq. (30) is given by

$$\langle \mathrm{d}f[\hat{\Gamma}_t]\rangle = \langle f[\hat{\Gamma}_{t+\mathrm{d}t}]\rangle - \langle f[\hat{\Gamma}_t]\rangle = \int \mathrm{d}\Gamma(P_{t+\mathrm{d}t}[\Gamma] - P_t[\Gamma])f[\Gamma] = \mathrm{d}t\int \mathrm{d}\Gamma f[\Gamma]\frac{\partial P_t[\Gamma]}{\partial t} + O(\mathrm{d}t^2), \tag{33}$$

we finally obtain the following integral identity regarding any functional $f[\Gamma]$:

$$\int \mathrm{d}\Gamma f[\Gamma]\frac{\partial P_t[\Gamma]}{\partial t} = \int \mathrm{d}\Gamma f[\Gamma]\left[\int_{-\infty}^\infty \mathrm{d}y\{\lambda[y|\Gamma - \Delta\Gamma_y]P_t[\Gamma - \Delta\Gamma_y] - \lambda[y|\Gamma]P_t[\Gamma]\} + \int_0^\infty \mathrm{d}s\frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma])\right] \tag{34}$$

in the limit $\mathrm{d}t \to 0$. Since this relation holds for an arbitrary $f[\Gamma]$, we obtain the fME (27).

E. Functional Kramers-Moyal expansion

By applying the identity [see Eq. (A11) for the functional Taylor expansion],

$$\int_{-\infty}^\infty \mathrm{d}y\lambda[y|\Gamma - \Delta\Gamma_y]P_t[\Gamma - \Delta\Gamma_y] = \sum_{n=0}^\infty \frac{1}{n!}\int_{-\infty}^\infty \mathrm{d}yy^n\left(-\frac{\partial}{\partial v} - \int_0^\infty \mathrm{d}s\frac{\delta}{\delta z(s)}\right)^n \lambda[y|\Gamma]P_t[\Gamma], \tag{35}$$

[4]Notably, while the jump probability during $[t, t+\mathrm{d}t)$ is of order $\mathrm{d}t$, the leading-order contribution $f[\hat{\Gamma}_t + \Delta\Gamma_{\hat{y}}] - f[\hat{\Gamma}_t]$ is of order 1. In addition, while the no-jump probability during $[t, t+\mathrm{d}t)$ is of order 1, the leading-order contribution $-\,\mathrm{d}t\int_0^\infty s\hat{z}_t(s)\{\delta f[\hat{\Gamma}_t]/\delta\hat{z}_t(s)\}\,\mathrm{d}s$ is of order $\mathrm{d}t$. Therefore, the contributions of their averages are balanced at the order $\mathrm{d}t$.

to the fME (27), we obtain the functional Kramers-Moyal (KM) expansion:

$$\frac{\partial P_t[\Gamma]}{\partial t} = \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma]) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!}\left(\frac{\partial}{\partial v} + \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}\right)^n \alpha_n[\Gamma]P_t[\Gamma] \tag{36a}$$

with the KM coefficient

$$\alpha_n[\Gamma] := \int_{-\infty}^{\infty} y^n \lambda[y|\Gamma]\,\mathrm{d}y. \tag{36b}$$

F. Remark: Systematic calculations based on linear algebra

Since the fME (27) is linear, it can be analyzed with the tools of linear algebra. Starting from the standard form (7) $\frac{\partial P_t[\Gamma]}{\partial t} = \mathcal{L}P_t[\Gamma]$, let us consider the eigenvalue problem

$$\mathcal{L}\phi_i[\Gamma] = -\mu_i\phi_i[\Gamma], \tag{37}$$

where μ_i is the ith eigenvalue and $\phi_i[\Gamma]$ is the corresponding eigenfunction.[5] The time-dependent solution is then given by superposition of the eigenfunctions

$$P_t[\Gamma] = \sum_i c_i e^{-\mu_i t}\phi_i[\Gamma], \tag{38}$$

where the coefficients $\{c_i\}_i$ are determined by the initial condition. The steady-state PDF corresponds to the zeroth eigenfunction $\phi_i[\Gamma]$ with $\mu_0 = 0$:

$$P_{\mathrm{ss}}[\Gamma] \propto \phi_0[\Gamma]. \tag{39}$$

Various physical quantities can be systematically calculated by the time-dependent (38) or steady-state solutions (39). For example, the correlation function is formally given by the path integral

$$\langle \hat{A}_t \hat{B}_{t'} \rangle = \int A_t[\Gamma]B_{t'}[\Gamma]P_{\mathrm{ss}}[\Gamma]\,\mathrm{d}\Gamma, \tag{40}$$

where A_t and $B_{t'}$ are expressed as functionals of Γ.

V. APPLICATION 1: SYSTEM-SIZE EXPANSION AND GENERALIZED LANGEVIN EQUATION

In this section, we illustrate the utilization of our fME framework in relation to an asymptotic theory, drawing upon the system-size expansion. By enforcing a stability condition upon the system-size expansion, we ultimately infer the GLE as a plausible coarse-graining process rooted in physical reasoning.

A. Assumptions

(1) Small noise assumption: Let us consider the non-Markovian process with small jumps,

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \varepsilon\hat{\xi}^{\mathrm{P}}_{W[Y|\{\hat{v}_\tau\}_{\tau\leqslant t}]}, \tag{41}$$

where W is the ε-independent conditional intensity of the jump size Y (see Fig. 2 for the schematic of the small-jump assumption). This assumption is equivalent to the following scaling relation of the conditional intensity density:

$$\lambda[y|\Gamma] = \frac{1}{\varepsilon}W\left[\frac{y}{\varepsilon}\middle|\Gamma\right]. \tag{42}$$

This assumption leads to the scaling relation for the KM coefficients

$$\alpha_n[\Gamma] = \varepsilon^n\mathcal{A}_n[\Gamma], \quad \mathcal{A}_n[\Gamma] := \int_{-\infty}^{\infty} Y^n W[Y|\Gamma]\,\mathrm{d}Y, \tag{43}$$

with the ε-independent KM coefficient $\mathcal{A}_n[\Gamma]$.

(2) Linear stability: Let us additionally assume that the first-order KM coefficient $\mathcal{A}_1[\Gamma]$ has a single stable point:

$$\mathcal{A}_1[\Gamma = \mathbf{0}] = 0, \quad \mathbf{0} := (0, \{0\}_s) \tag{44}$$

We also assume that $\mathcal{A}_1[\Gamma]$ is linearly stable around $\Gamma = \mathbf{0}$, such that

$$\gamma = -\frac{\partial}{\partial v}\mathcal{A}_1[\Gamma]\bigg|_{\Gamma=\mathbf{0}} > 0, \quad \Upsilon(u) = -\frac{\delta}{\delta z(s)}\mathcal{A}_1[\Gamma]\bigg|_{\Gamma=\mathbf{0}} > 0 \ \text{for any } u, \tag{45}$$

where the rescaled wave number u is defined by $u := \frac{s}{\varepsilon}$.

(3) Existence of the noise term: The noise term is assumed to be present even for $\varepsilon \to 0$ and thus the variance term is nonzero:

$$\sigma^2 = \mathcal{A}_2[\Gamma = \mathbf{0}] > 0. \tag{46}$$

The assumption of linear stability around $\mathbf{\Gamma} = \mathbf{0}$ implies that the velocity $\hat{v}_t$ typically relaxes to zero. In other words, we are focusing on stationary processes; nonstationary processes, where the state variable can go to infinity (such as for the Brownian motion), are not considered in Sec. V.

In addition, as a technical assumption, we assume that all the considered integrals converge. This assumption implies

[5] While the eigenvalue spectral may be continuous technically, we formally write the eigenvalues with discrete notation.

that physically singular processes, such as with long time tail with decaying speed slower than t^{-1}, are out of the scope of this paper. Note that the above stability assumptions parallel the conventional stability assumption for the Markovian jump process (see Appendix C for their comparison). Also, we note that all even-order KM coefficients are positive under assumption 3, such that $\mathcal{A}_k[\mathbf{0}] > 0$ for all even k due to Pawula's theorem [23].

B. Asymptotic derivation of the functional Fokker-Planck equation

Given the following rescaled variables:

$$\mathrm{t} := \varepsilon t, \quad u := \frac{s}{\varepsilon}, \quad V := \frac{v}{\sqrt{\varepsilon}}, \quad Z(u) := \frac{z(s)}{\sqrt{\varepsilon}}, \quad G := (V, \{Z(u)\}_u), \tag{47}$$

we obtain the following functional Fokker-Planck equation for $\varepsilon \to 0$:

$$\frac{\partial P_{\mathrm{t}}[G]}{\partial \mathrm{t}} = \Bigg[\frac{\partial}{\partial V}\left\{\gamma V + \int_0^\infty du_2 \Upsilon(u_2) Z(u_2)\right\} + \int_0^\infty du_1 \frac{\delta}{\delta Z(u_1)}\left\{u Z(u_1) + \gamma V + \int_0^\infty du_2 \Upsilon(u_2) Z(u_2)\right\}$$
$$+ \frac{\sigma^2}{2}\left\{\frac{\partial}{\partial V} + \int_0^\infty du \frac{\delta}{\delta Z(u)}\right\}^2\Bigg] P_{\mathrm{t}}[G]. \tag{48}$$

See Appendix D 1 for the detailed derivation. This result is an analog to the system-size expansion for Markovian jump processes (see Appendix C for comparison).

The functional Fokker-Planck (48) is equivalent to the stochastic dynamics described by

$$\frac{d\hat{V}_{\mathrm{t}}}{d\mathrm{t}} = -\gamma \hat{V}_{\mathrm{t}} - \int_0^\infty du \Upsilon(u) \hat{Z}_{\mathrm{t}}(u) + \sigma \hat{\xi}^{\mathrm{G}}_{\mathrm{t}}, \tag{49a}$$

$$\frac{\partial \hat{Z}_{\mathrm{t}}(u)}{\partial \mathrm{t}} = -u \hat{Z}_{\mathrm{t}}(u) - \gamma \hat{V}_{\mathrm{t}} - \int_0^\infty du \Upsilon(u) \hat{Z}_{\mathrm{t}}(u) + \sigma \hat{\xi}^{\mathrm{G}}_{\mathrm{t}}, \tag{49b}$$

with standard white Gaussian noise $\hat{\xi}^{\mathrm{G}}_{\mathrm{t}}$ that is common to the stochastic dynamics of $\hat{V}_{\mathrm{t}}$ and $\hat{Z}_{\mathrm{t}}(u)$.

C. Asymptotic derivation of the generalized Langevin equation

The stochastic dynamics (49) is equivalent to the GLE,

$$\frac{d\hat{V}_{\mathrm{t}}}{d\mathrm{t}} = -\gamma \hat{V}_{\mathrm{t}} - \int_{-\infty}^{\mathrm{t}} d\mathrm{t}' \mathcal{M}(\mathrm{t}-\mathrm{t}') \hat{V}_{\mathrm{t}'} + \hat{\eta}_{\mathrm{t}}, \tag{50a}$$

with memory kernel $\mathcal{M}(\mathrm{t})$ given by expression (54) and colored Gaussian noise:

$$\hat{\eta}_{\mathrm{t}} := \sigma \hat{\xi}^{\mathrm{G}}_{\mathrm{t}} + \sigma \int_{-\infty}^{\mathrm{t}} d\mathrm{t}' \mathcal{M}(\mathrm{t}-\mathrm{t}') \hat{\xi}^{\mathrm{G}}_{\mathrm{t}'}. \tag{50b}$$

This is the third main result, implying that the GLE is a minimal model for the coarse-grained description of general non-Markovian jump processes in the weak coupling limit $\varepsilon \to 0$ under the stability condition (Sec. V A). The memory effect naturally manifests itself in both the friction and noise terms, stemming from the historical dependence inherent in the original non-Markovian jump process. See Appendix D 2 for the detailed derivation.

The memory kernel $\mathcal{M}(\mathrm{t})$ and the noise statistics can be explicitly derived as follows. Let us define the matrix $K(u, u')$,

$$K(u, u') := u\delta(u - u') + \Upsilon(u'), \tag{51}$$

and the corresponding eigenvalue μ and eigenfunctions $\{e(\mu; u)\}_\mu$, satisfying

$$\int_0^\infty du' K(u, u') e(\mu; u') = \mu e(\mu; u). \tag{52}$$

The matrix $K(u, u')$ has the following properties (see Appendix F): (i) All of its eigenvalues are real and positive $\mu > 0$. (ii) $K(u, u')$ is a positive symmetric matrix and thus has an inverse matrix $K^{-1}(u, u')$. We assume that the eigenfunctions $\{e(\mu; u)\}_\mu$ constitute a complete set, and have inverse matrices $e^{-1}(u; \mu')$ such that[6]

$$\int_0^\infty du\, e(\mu; u) e^{-1}(u; \mu') = \delta(\mu - \mu'),$$
$$\int_0^\infty d\mu\, e^{-1}(u; \mu) e(\mu; u') = \delta(u - u'). \tag{53}$$

[6]In N-dimensional linear algebra, the set of all eigenvectors $\{\boldsymbol{e}_i\}_i$ with $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{iN})$ of a symmetric matrix constitute a complete set. In addition, the matrix $A := (e_{ij})$ has the inverse matrix $A^{-1} = (e^{-1}_{ij})$, such that $\sum_j e_{ij} e^{-1}_{jk} = \delta_{ik}$ and $\sum_j e^{-1}_{ij} e_{jk} = \delta_{ik}$. This property is a straightforward generalization from finite-dimensional to infinite-dimensional linear algebra.

With these notations, the memory kernel and the noise statistical properties are, respectively, given by

$$\mathcal{M}(\mathrm{t}) := \int_0^\infty \nu(\mu)e^{-\mu\mathrm{t}}\,\mathrm{d}\mu, \quad \nu(\mu) := \int_0^\infty \kappa(\mu)\Upsilon(u)e(\mu;u)\,\mathrm{d}u, \quad \kappa(\mu) := \int_0^\infty e^{-1}(u;\mu)\,\mathrm{d}u, \tag{54}$$

$$\langle\hat{\eta}_\mathrm{t}\rangle = 0, \quad \langle\hat{\eta}_{\mathrm{t}_1}\hat{\eta}_{\mathrm{t}_2}\rangle = \sigma^2\left[\delta(\mathrm{t}_1-\mathrm{t}_2) + \mathcal{M}(|\mathrm{t}_1-\mathrm{t}_2|) + \int_0^\infty \mathrm{d}\mathrm{t}'\mathcal{M}(\mathrm{t}')\mathcal{M}(\mathrm{t}'+|\mathrm{t}_1-\mathrm{t}_2|)\right]. \tag{55}$$

D. Physical interpretation

It is interesting that the colored Gaussian noise emerges from the system-size expansion, while our starting point is the non-Markovian jump process. This approximate result can be interpreted within the law of large numbers and the central limit theorem. In formulating the system-size expansion, the rescaled time $\mathrm{t} := \varepsilon t$ is assumed to be order 1. This assumption implies that the original time $t = \mathrm{t}/\varepsilon$ is very large; the typical number of collisions during $[0, t)$ is very large likewise. Therefore, both the law of large numbers and the central limit apply. The time-delayed friction term results from the law of large numbers, while the colored Gaussian noise term results from the central limit theorem. In other words, we have developed a formal application of the law of large numbers and of the central limit theorem for general non-Markovian jump processes.

VI. APPLICATION 2: PRICE DYNAMICS BASED ON NONLINEAR HAWKES PROCESSES

In this section, we illustrate another application of our formalism. We focus on modeling financial price dynamics based on a nonlinear Hawkes process. Linear Hawkes processes have become popular in econophysics as well as in econometrics of market microstructure.

A. Model

Let us consider a stochastic financial model based on the nonlinear Hawkes processes, which has recently become popular to describe the price dynamics of financial assets [49,50]. Let us denote $\hat{v}_t$ the logarithm of the price of some stock at time t. The log-price dynamics is given by

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \sum_{k=1}^{\hat{N}_t}\hat{y}_k\delta(t-\hat{t}_k), \tag{56a}$$

where $\hat{y}_k$ is kth jump size of the log price occurring at time $\hat{t}_k$. The amplitude of the jumps are independently and identically distributed with mark distribution $\rho(y)$. The sequence of jumps defines the jump size series $\{\hat{y}_k\}_k$ and the jump time series $\{\hat{t}_k\}$. We denote by $\hat{N}_t$ the total number of jumps during $[0, t)$. We assume that both excitatory and inhibitory effects are balanced, which is realized when the mark distribution is symmetric:

$$\rho(y) = \rho(-y). \tag{56b}$$

The intensity $\hat{\lambda}_t$ of the jumps is assumed to obey the nonlinear Hawkes process

$$\hat{\lambda}_t = g\left(\sum_{k=1}^{\hat{N}_t}\hat{y}_k h(t-\hat{t}_k)\right), \tag{56c}$$

with non-negative intensity function $g > 0$ and memory kernel $h(t)$. Recall that the intensity $\hat{\lambda}_t$ gives the probability per unit time for the next jump to occur: $\hat{\lambda}_t\,\mathrm{d}t$ is the probability for the next jump to occur during $[t, t+\mathrm{d}t)$. See Fig. 6 for the schematic paths of this model regarding the intensity $\hat{\lambda}_t$ and the log-price $\hat{v}_t$.

This model is an example of a history-dependent Poisson processes. Indeed, the following specific history-dependent Poisson process:

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = \hat{\xi}^{\mathrm{CP}}_{\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t})},$$
$$\lambda(y|\{\hat{v}_\tau\}_{\tau\leqslant t}) := \rho(y)g\left(\int_0^\infty h(\tau)\hat{a}_{t-\tau}\,\mathrm{d}\tau\right), \tag{57}$$
$$\hat{a}_t := \frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t}$$

is equivalent to the nonlinear Hawkes price model (56) [34,35].

B. Markovian embedding

Our Laplace-convolution Markovian-embedding scheme (23) fully converts the nonlinear non-Markovian Hawkes process (56) into a Markovian field process. Indeed, by decomposing the memory kernel as the sum of exponentials

$$h(t) := \int_0^\infty e^{-st}\tilde{h}(s)\,\mathrm{d}s, \tag{58}$$

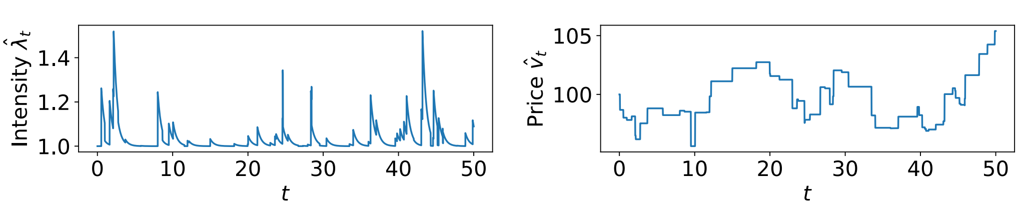

FIG. 6. Schematic paths of the intensity $\hat{\lambda}_t$ and the price $\hat{v}_t$ described by the nonlinear Hawkes model (56) for the financial price dynamics.

the conditional intensity can be rewritten as

$$\lambda[y|\Gamma] = \rho(y)g\left(\int_0^\infty \tilde{h}(s)\hat{z}_t(s)\,\mathrm{d}s\right). \tag{59}$$

This is equivalent to the Markov-embedding representation introduced in our previous works [32–35].

C. Field master equation

The fME for the nonlinear Hawkes price model (56) is

$$\frac{\partial P_t[\Gamma]}{\partial t} = \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma]) + \int_{-\infty}^{\infty} \mathrm{d}y\rho(y)[G[z-y\mathbf{1}]P_t[\Gamma-\Delta\Gamma_y] - G[z]P_t[\Gamma]], \tag{60a}$$

with

$$G[z] := g\left(\int_0^\infty \tilde{h}(s)\hat{z}_t(s)\,\mathrm{d}s\right) \quad \text{and} \quad \hat{\Gamma}_t = (\hat{v}_t, \{\hat{z}_t(s)\}_s). \tag{60b}$$

By integrating out both sides over $\hat{v}_t$, this fME reduces to the fME for a marginal PDF $P_t[z] := \int P_t[\Gamma]\,\mathrm{d}x$ that was introduced in our previous works [32–35] (see Appendix G for the explicit derivation). In addition, the reduced fME has been analytically solved in Refs. [32–35] for the asymptotic intensity PDF in the steady state.

D. Diffusive approximation

Let us apply the diffusive approximation by using the KM series (36a) for the fME (60) and truncating it at second order. This leads to the following approximate Fokker-Planck equation:

$$\frac{\partial P_t[\Gamma]}{\partial t} \simeq \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma]) + \frac{1}{2}\left(\frac{\partial}{\partial v} + \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}\right)^2 \alpha_2[\Gamma]P_t[\Gamma] \tag{61}$$

with

$$\alpha_2[\Gamma] := \sigma^2 g\left(\int_0^\infty \tilde{h}(s)z(s)\,\mathrm{d}s\right), \quad \sigma^2 := \int_{-\infty}^{\infty} y^2\rho(y)\,\mathrm{d}y. \tag{62}$$

This field Fokker-Planck equation is equivalent to

$$\frac{\mathrm{d}\hat{v}_t}{\mathrm{d}t} = D_t\hat{\xi}_t^{\mathrm{G}}, \quad D_t := \sigma^2\hat{\lambda}_t. \tag{63}$$

This recovers the standard Geometric Brownian Motion model of price dynamics for constant $\hat{\lambda}_t$. For nonconstant $\hat{\lambda}_t$, Eq. (63) recovers the general class of stochastic volatility models [51]. Here, we derived that the volatility is proportional to the intensity $\hat{\lambda}_t$ of the underlying point process. In other words, our nonlinear Hawkes (56) combined with our Markovian embedding and the diffusive approximation provide an interpretation of the source of stochastic volatility, which is here interpreted as resulting from the underlying jump intensity and its nonlinear memory structure.

VII. DISCUSSION AND CONCLUSION

This section delves into the ramifications of our research and outlines our perspective on several outstanding technical challenges yet to be addressed. We finally conclude this paper with some remarks.

A. Comparison with the projection-operator formalism

Our formulation bears similarities to the projection-operator formalism, as both theories pertain to the derivations of the GLEs. In this subsection, we juxtapose the two approaches, evaluating their respective advantages and disadvantages.

The projection-operator formalism originated in the 1950s and 1960s, crafted by pioneers like Nakajima, Mori, Zwanzig, and Kawasaki [28,52–56]. Particularly, Mori's approach focuses on establishing a microscopic foundation for the GLEs. In the projection-operator formalism, the selection of several slow variables is necessitated, guided by physical intuitions or empirical findings, as these variables cannot be determined theoretically. Subsequently, a projection operator is defined to dissect the phase-space dynamics between the function space, exclusive to slow variables, and the remainder.

Through the application of integral identities associated with projection operators, GLEs are derived. A notable merit of this approach is the formal derivation of GLEs from microscopic dynamics, providing a rigorous connection to underlying physical processes. However, a significant drawback lies in the inherent ambiguity of the approximation involved. While all calculations are theoretically exact, eliciting nontrivial predictions mandates the approximate computation of noise statistics and friction coefficients. This level of approximation is notably more intricate compared to conventional statistical-physics theories. In fact, the determination of theoretical key perturbation and control parameters for the conclusive deduction of the GLEs from microscopic dynamics remains unambiguous, making this process elusive.

Within the foundational framework of statistical physics pertaining to the GLEs, a drawback of our theory is the requisite assumption of the one-dimensional non-Markovian jump process (16) as an initial standpoint. This assumption is fundamentally heuristic, primarily rooted in phenomenological considerations. Conversely, a significant advantage of our approach is the explicit definition of the key perturbation parameter. Specifically, the small-jump scaling parameter, ϵ—generally anticipated to represent the mass ratio between the Brownian particle and surrounding entities—serves a crucial and explicit role in our asymptotic computations. This is particularly coherent for modeling dynamics of massive Brownian particles. In this sense, we have successfully established the GLEs through a physically plausible coarse-graining process, pinpointing the essential control parameter for mathematical derivation, a contrast to the methodologies embedded in the projection-operator formalism.

B. Future issue 1: Physical validation of the non-Markovian jump model

Our theory is premised on the non-Markovian jump model (16). While intrinsic to models in seismic activity, finance, and

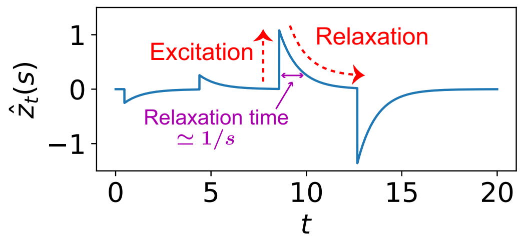

FIG. 7. Absence of time-reversal symmetry for the auxiliary field variables $\{\hat{z}_t(s)\}$. According to the SPDE (23a), the path of $\hat{z}_t(s)$ responds to the excitations due to jumps and then relaxes toward zero. The relaxation is time-irreversible, and the auxiliary field variables have thus no time-reversal symmetry.

social science—for instance, the Hawkes process is a subset of this model—its applicability in physics remains indeterminate. Addressing this uncertainty will necessitate further theoretical or data-driven analysis in the future.

From a theoretical standpoint, the Markovian ME formalism (8) has been substantiated in the dynamics of Brownian particles amid dilute gases [57–59]. Indeed, the linearized Boltzmann equation, derivable from Newtonian microscopic dynamics through the Bogoliubov-Born-Green-Kirkwood-Yvon hierarchy [57,58] in the low-density limit, is an instance of a Markovian jump process, thus allowing systematic theoretical validation of the Markovian ME formalism (8) via kinetic theory.

Conversely, theoretical validation for the non-Markovian jump model (16) is yet to be achieved. Formulating a statistical physics theory analogous to the Markovian kinetic theory to derive the non-Markovian jump process (16) from microscopic Hamiltonian dynamics is imperative.

C. Future issue 2: Time-reversal symmetry of our field master equation

Exploring time-reversal symmetry is crucial when examining stochastic dynamics influenced by equilibrium fluctuations. Regrettably, this symmetry is not upheld for the fME (27). The indispensable condition for general master equations, fully detailed in Gardiner's textbook [22] and Appendix H, is invariably breached in our fME (27).

The absence of time-reversal symmetry in our Laplace-type embedding representation can be intuitively understood by considering a typical path of $\{\hat{z}_t(s)\}_s$. Indeed, the SPDE (23a) states that the dynamics of $\hat{z}_t(s)$ is composed of the excitation due to the Poisson jump $\hat{\xi}^{\rm CP}_{\lambda[y|\hat{\Gamma}_t]}$ and the relaxation due to the term $-s\hat{z}_t(s)$ with the characteristic timescale $\simeq 1/s$. Since the relaxation dynamics is time irreversible, the dynamics of $\{\hat{z}_t(s)\}_s$ has no time-reversal symmetry by construction (see Fig. 7). Also, the Laplace-embedding can be interpreted as an exponential-moving average of the history, which is time-irreversible, intuitively.

The maintenance of time-reversal symmetry in a ME largely depends on the selected state variables [60]. To highlight this mathematical fact, let us consider a trivial example based on Hamiltonian dynamics. The canonical equation for a harmonic oscillator is given by

$$\boldsymbol{x} := (p, q), \quad \frac{{\rm d}p}{{\rm d}t} = -q, \quad \frac{{\rm d}q}{{\rm d}t} = p. \tag{64}$$

The corresponding master equation is called the Liouville equation,

$$\frac{\partial P_t(\boldsymbol{x})}{\partial t} = \left[-\frac{\partial}{\partial p}A_1(\boldsymbol{x}) - \frac{\partial}{\partial q}A_2(\boldsymbol{x})\right]P_t(\boldsymbol{x}), \tag{65}$$

with $A_1(\boldsymbol{x}) := -q$ and $A_2(\boldsymbol{x}) := p$. This Liouville equation satisfies the detailed-balance condition (H2) by defining the time-reversal operators $(\epsilon_1, \epsilon_2) = (-1, 1)$ and $B_{kl}(\boldsymbol{x}) = 0$ for any k and l. Thus, this process represented by $\boldsymbol{x} := (p, q)$ has the time-reversal symmetry.

As another mathematically equivalent representation, let us consider the following state variables:

$$\boldsymbol{y} := (r, q) := (p + q, q), \quad \frac{{\rm d}r}{{\rm d}t} = r - 2q, \quad \frac{{\rm d}q}{{\rm d}t} = r - q. \tag{66}$$

The corresponding Liouville equation is given by

$$\frac{\partial P_t(\boldsymbol{y})}{\partial t} = \left[-\frac{\partial}{\partial r}A'_1(\boldsymbol{y}) - \frac{\partial}{\partial q}A'_2(\boldsymbol{y})\right]P_t(\boldsymbol{y}) \tag{67}$$

with $A'_1(\boldsymbol{x}) := r - 2q$ and $A'_2(\boldsymbol{x}) := r - q$. For any definition $(\epsilon'_1, \epsilon'_2)$, the detailed-balance condition (H2a) is always violated in the representation $\boldsymbol{y}$. Indeed, we obtain

$$\begin{aligned}&\epsilon'_1A'_1(\boldsymbol{\epsilon}'\boldsymbol{y})P_{\rm ss}(\boldsymbol{y}) + A'_1(\boldsymbol{y})P_{\rm ss}(\boldsymbol{y})\\ &\quad = 2[r - (1 + \epsilon'_1\epsilon'_2)q]P_{\rm ss}(\boldsymbol{y}) \neq 0\end{aligned} \tag{68}$$

for the steady-state PDF $P_{\rm ss}(\boldsymbol{y})$, by assuming $(\epsilon'_1)^2 = 1$ and $B'_{kl}(\boldsymbol{y}) = 0$ for any k and l. Thus, the detailed-balance condition (H2) does not hold for the new representation $\boldsymbol{y}$, even though it holds for the original representation $\boldsymbol{x}$. This example clearly shows that the time-reversal symmetry of the obtained ME depends on the selection of the state variables.

This key mathematical insight reveals that multiple Markovian-embedding approaches can be applied to the same ME, even when the stochastic dynamics are uniquely defined. Consequently, time-reversal symmetry may be effectively captured in certain Markovian-embedding models. As such, an alternative Markovian-embedding approach might be more appropriate for expressing time-reversal symmetry when it holds true at the microscopic level. Addressing this aspect is crucial for advancing the development of stochastic thermodynamics and energetics [61,62], particularly in the context of non-Markovian jump processes. Further exploration and resolution of this matter are planned for future research endeavors.

D. Future issue 3: The fluctuation-dissipation relation of the second kind

When the environment is in thermal equilibrium, the thermal fluctuation of the GLE must satisfies the fluctuation-dissipation relation of the second kind:

$$\langle \hat{\eta}_{t_1}\hat{\eta}_{t_2}\rangle = 2T\{\gamma\delta(t_1 - t_2) + \mathcal{M}(|t_1 - t_2|)\}, \tag{69}$$

where the left-hand side is the cross-correlation between $\hat{\eta}_{t_1}$ and $\hat{\eta}_{t_2}$, T is the temperature and we have taken units where the Boltzmann constant is unity. This fluctuation-dissipation relation of the second kind is equivalent to the time-reversal symmetry of the GLE. This was one of the most important issues for the statistical-physics foundation of the GLE particularly within the context of linear response theory [1] and the projection-operator formalism [28].

Since our Markovian-embedding formulation does not yet convert the time-reversal symmetry of the fME, the necessary and sufficient condition for this fluctuation-dissipation relation of the second kind is not yet identified. The identification of these conditions is also an important future challenge.

E. Future issue 4: Formal relations to quantum field theory

Our fME is formally related to quantum field theory. Indeed, the field Fokker-Planck equation for the GLE with time-reversal symmetry is equivalent to a non-Hermitian quantum field theory with the Hermitian part of its Hamiltonian describing a field of harmonic oscillators [33]. Also, a similar renormalisation issue appears regarding the infinite zero-point energy of the field harmonic oscillators. Since the first-order contribution of the system-size expansion for the non-Markovian jump process leads to the GLE, its next-order perturbation theory might require the use of methods developed in quantum field theory, such as the Feynman-diagram expansion. Establishing such field-theoretical techniques will be an interesting future topic.

F. Future issue 5: The backward field master equation

Technically, the fME (27) is a *forward master equation*, for which the time-evolution goes from the past to the future with imposed initial conditions. On the other hand, there is another ME called the *backward master equation*, for which the time evolution goes from the future to the past with imposed final conditions. Generally, the backward MEs can be essentially derived as the self-adjoint equation of the forward MEs. Therefore, it is straightforward to derive the backward fME corresponding to the forward fME (27). Studying the properties of the backward fME will be an interesting avenue for future research.

G. Conclusion

In conclusion, we have introduced a comprehensive stochastic framework through a field master equation, encompassing all one-dimensional non-Markovian jump processes. Utilizing the Laplace-convolution embedding representation, we have demonstrated the transformation of any non-Markovian jump process into Markovian-field dynamics. We subsequently derived the corresponding field master equation and procured an asymptotic solution using a generalized system-size expansion. In essence, this framework can be applied to any jump processes, assuming one-dimensional dynamics are driven by collisions. We posit that this model's flexibility makes it adept at accommodating a wide array of point-process data, proving invaluable for data analyzation.

ACKNOWLEDGMENTS

K.K. was supported by JST PRESTO (Grants No. JPMJPR20M2) and No. JSPS KAKENHI (Grants No. 22H04830, No. 22H01141, No. 23H00467, and No. 24H00833). D.S. was partially supported by the National Natural Science Foundation of China (Grant No. T2350710802 and No. U2039202), Shenzhen Science and Technology Innovation Commission Project (Grants No. GJHZ20210705141805017 and No. K23405006), and the Center for Computational Science and Engineering at Southern University of Science and Technology. We thank M. Itami for his careful reading of our manuscripts with useful feedback. We also thank A. Dechant for his crucial comment that the condition of time-reversal symmetry explicitly depends on the selection of Markovian-embedding representations. In addition, K.K. had a fruitful discussion with A. Teretenkov in the international conference Statphys 28 regarding Markovian embedding techniques in quantum stochastic processes. Furthermore, the discussions with H. Tajima and S. Ito were quite inspiring for future applications to stochastic thermodynamics.

APPENDIX A: DIRAC'S δ FUNCTION AND FUNCTIONAL DERIVATIVE

In this Appendix, we formally define the Dirac δ function and the functional derivatives. While our formulation is systematic enough at the theoretical-physics level, presenting mathematically rigorous formulations is out of scope in this paper.

1. Formal definition

a. Dirac's δ function

Dirac's δ function is formally defined by

$$\delta(s-s') = \begin{cases} 0 & (s \neq s') \\ \infty & (s = s') \end{cases}, \quad \int_{-\infty}^{\infty} \mathrm{d}s f(s)\delta(s-s') = f(s'), \tag{A1}$$

with any real numbers s and s'. The δ function is the continuous analog of the Kronecker δ for discrete variables, which is defined by

$$\delta_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}, \quad \sum_k \delta_{ik} f_k = f_i \tag{A2}$$

for any integers i and j.

Dirac's δ function can be formally constructed via a lattice model. Let us discretize the real number line $(0,\infty)$, such that $s_k = k\,\mathrm{d}s$ with the lattice constant $\mathrm{d}s > 0$ and any integer $k > 0$. The Dirac δ function is formally defined by

$$\delta(s-s') := \lim_{\mathrm{d}s \searrow 0} \frac{1}{\mathrm{d}s}\delta_{ij}, \tag{A3}$$

where $s \in [s_i, s_{i+1})$ and $s' \in [s_j, s_{j+1})$. Indeed, with this definition, we obtain the consistent relationship

$$\int_{-\infty}^{\infty} \mathrm{d}s f(s)\delta(s-s') = \lim_{\mathrm{d}s \searrow 0} \sum_i \mathrm{d}s_i f(s_i)\frac{1}{\mathrm{d}s}\delta_{ij} = f(s'). \tag{A4}$$

b. Functional derivative

Let us define the functional derivatives as a formal limit from the finite-dimensional vector function (i.e., a lattice model). Let us consider the K-dimensional vector $z := (z_1, \dots, z_K)$ and an arbitrary function $f(z)$. The partial derivative of $f(z)$ is written as $(\partial f(z))/(\partial z_k)$ for an integer k.

We then consider a formal continuous limit from such a finite-dimensional models. Let us introduce $s_k := k(\mathrm{d}s)$ with the lattice constant $\mathrm{d}s > 0$ for integer $k > 0$, and take the continuous limit $\mathrm{d}s \searrow 0$ and $K \to \infty$. The functional derivative is defined by

$$\frac{\delta f[z]}{\delta z(s)} := \lim_{\substack{\mathrm{d}s\searrow 0\\ K\to\infty}} \frac{1}{\mathrm{d}s}\frac{\partial f(z)}{\partial z(s_k)}, \tag{A5}$$

where $s_k = k\,\mathrm{d}s$ and $s \in [s_k, s_{k+1})$ with an integer k.

2. Useful identities

a. First-order functional Taylor expansion

For a finite-dimensional vector function $f(z)$, the first-order Taylor expansion is given by

$$\mathrm{d}f(z) = \sum_{k=1}^{K} \frac{\partial f(z)}{\partial z_k}\mathrm{d}z_k, \tag{A6}$$

with $\mathrm{d}f(z) := f(z + \mathrm{d}z) - f(z)$ for infinitesimal $\mathrm{d}z := (\mathrm{d}z_1, \dots, \mathrm{d}z_K)$. In the continuous limit, we apply the replacement

$$\sum_{k=1}^{K}\mathrm{d}s[\dots] \to \int_0^\infty \mathrm{d}s[\dots], \quad \frac{\partial f(z)}{\partial z(s_k)} \to \mathrm{d}s\frac{\delta f[z]}{\delta z(s)},$$
$$\mathrm{d}z_k \to \delta z(s) \tag{A7}$$

to obtain the first-order functional Taylor expansion

$$\delta f[z] = \int \mathrm{d}s \frac{\delta f[z]}{\delta z(s)}\delta z(s) + O(\delta z^2), \tag{A8}$$

with $\delta f[z] := f[z + \delta z] - f[z]$ with infinitesimal δz.

b. Full-order functional Taylor expansion

For a finite-dimensional vector function $f(z)$, the full-order Taylor expansion is given by

$$f(z + \Delta z) - f(z) = \sum_{n=1}^{\infty}\frac{1}{n!}\left(\sum_{k=1}^{K}\Delta z_k\frac{\partial}{\partial z_k}\right)^n f(z), \tag{A9}$$

with $\Delta z := (\Delta z_1, \dots, \Delta z_K)$. In the continuous limit based on the formal replacement (A7), we obtain the full-order functional Taylor expansion:

$$f[z + \Delta z] - f[z] = \sum_{n=1}^{\infty}\frac{1}{n!}\left(\int \mathrm{d}s\Delta z(s)\frac{\delta}{\delta z(s)}\right)^n f[z]. \tag{A10}$$

We note that this calculation can be readily generalized for a two-argument functional $f[v; z]$ with a real value v and a function $\{z(s)\}_s$, such that

$$f[v + \Delta v; z + \Delta z] - f[v; z] = \sum_{n=1}^{\infty}\frac{1}{n!}\left(\Delta v\frac{\partial}{\partial v} + \int \mathrm{d}s\Delta z(s)\frac{\delta}{\delta z(s)}\right)^n f[v; z], \tag{A11}$$

with small Δv and $\{\Delta z(s)\}_s$. Particularly, the Maclaurin series is given by

$$f[v; z] = \sum_{n=0}^{\infty}\frac{1}{n!}\left(v\frac{\partial}{\partial\chi} + \int \mathrm{d}s z(s)\frac{\delta}{\delta\zeta(s)}\right)^n f[\chi; \zeta]\Big|_{(\chi,\{\zeta(s)\}_s)=\mathbf{0}}, \tag{A12}$$

where the dummy argument variables χ and ζ are introduced to distinguish the arguments involved in the derivatives from the arguments v and z of the function $f[v; z]$.

c. Variable transformation formula

Let us consider a simple variable transformation

$$\tilde{s} = as, \tag{A13}$$

with a positive constant a. Considering the definition (A5), we obtain

$$\frac{\delta}{\delta z(s)} := \lim_{\substack{\mathrm{d}s\searrow 0\\ K\to\infty}} \frac{a}{(a\,\mathrm{d}s)}\frac{\partial}{\partial z(s_k)} = a \lim_{\substack{\mathrm{d}\tilde{s}\searrow 0\\ K\to\infty}} \frac{1}{\mathrm{d}\tilde{s}}\frac{\partial}{\partial z(\tilde{s}_k)} = a\frac{\delta}{\delta z(\tilde{s})}, \tag{A14}$$

which leads to the invariant integral relationship:

$$\int \mathrm{d}s\frac{\delta}{\delta z(s)} = \int \mathrm{d}\tilde{s}\frac{\delta}{\delta z(\tilde{s})}. \tag{A15}$$

d. Partial integration

For a finite-dimensional vector z, the partial integration is given by

$$\int_{-\infty}^{\infty} P(z)\frac{\partial f(z)}{\partial z_k}\mathrm{d}z = -\int_{-\infty}^{\infty} f(z)\frac{\partial P(z)}{\partial z_k}\mathrm{d}z \tag{A16}$$

by assuming vanishing boundary conditions $\lim_{|z|\to\infty} P(z) = 0$. As a straightforward generalization, by considering the formal definition (A5), the partial integration of a functional $f[z]$ is given by

$$\int P[z]\frac{\partial f[z]}{\partial z(s)}\mathcal{D}z = -\int f[z]\frac{\delta P[z]}{\delta z(s)}\mathcal{D}z \tag{A17}$$

by also assuming vanishing boundary conditions.

APPENDIX B: BRIEF REVIEW OF THE WHITE GAUSSIAN AND POISSON NOISES

1. White Gaussian noise

Let us consider the following SDE with finite time step $\mathrm{d}t$:

$$\hat{W}_{t+dt} = \hat{W}_t + \sqrt{\mathrm{d}t}\hat{\eta}_t^{\mathrm{G}}, \tag{B1}$$

with the standard normal random variable $\hat{\eta}_t^{\mathrm{G}}$ that is independent and identically distributed: $\langle\hat{\eta}_t^{\mathrm{G}}\hat{\eta}_{t'}^{\mathrm{G}}\rangle = 1$ for $t = t'$ and $\langle\hat{\eta}_t^{\mathrm{G}}\hat{\eta}_{t'}^{\mathrm{G}}\rangle = 0$ for $t \neq t'$.

We then consider the stochastic dynamics for the infinitesimal time step limit $\mathrm{d}t \to 0$ to define the *Wiener process* $\hat{W}_t$. The formal derivative of the Wiener process is called the *white Gaussian noise*,

$$\hat{\xi}_t^{\mathrm{G}} := \frac{\mathrm{d}\hat{W}_t}{\mathrm{d}t}, \tag{B2}$$

which satisfies the relationship of the white noise:

$$\left\langle\hat{\xi}_t^{\mathrm{G}}\hat{\xi}_{t'}^{\mathrm{G}}\right\rangle = \delta(t - t'). \tag{B3}$$

2. White Poisson noise

The white Poisson noise is composed of the sum of δ functions, such that

$$\hat{\xi}^{\rm P}_{t,\lambda(y)} := \sum_{i=1}^{\hat{N}(t)} \hat{y}_i \delta(t-\hat{t}_i), \tag{B4}$$

which is characterized by the intensity density function $\lambda(y)$. $\{\hat{t}_i\}_i$ is the time sequence of jump events, $\{\hat{y}_i\}_i$ is the sequence of jump sizes (called *mark* in the context of point processes), and $\hat{N}(t)$ is the total number of jump events during the interval $[0,t)$. The probability that an event with jump size $\hat{y}_i \in [y, y+{\rm d}y)$ occurs during $[t, t+{\rm d}t)$ is given by

$$\lambda(y)\,{\rm d}y\,{\rm d}t. \tag{B5}$$

When the total intensity $\lambda_{\rm tot} := \int_{-\infty}^{\infty} \lambda(y)\,{\rm d}y$ is finite ($\lambda_{\rm tot} < \infty$), an event occurs during $[t, t+{\rm d}t)$ with the probability

$$\lambda_{\rm tot}\,{\rm d}t \tag{B6}$$

and the jump size distribution is given by

$$\rho(y) = \frac{\lambda(y)}{\int_{-\infty}^{\infty} \lambda(y)\,{\rm d}y}. \tag{B7}$$

3. White noise

The white noise $\hat{\xi}^{\rm W}(t)$ is the time-homogeneous noise without time correlation and is defined as the formal time-derivative of the Lévy process. The Lévy process $\hat{L}_t$ is defined as the stochastic process satisfying the following properties: (i) $L_0 = 0$. (ii) For any $0 \leqslant t_1 < t_2 < \cdots < t_n$, $L_{t_2} - L_{t_1}, L_{t_3} - L_{t_2}, \ldots, L_{t_n} - L_{t_{n-1}}$ are independent of each other. (iii) For any $s < t$, the PDF of $L_t - L_s$ is equal to that of L_{t-s}. With mean $m' := \langle \hat{\xi}^{\rm W} \rangle$, the white noise has no correlation, such that

$$\left\langle \left(\hat{\xi}^{\rm W}_t - m'\right)\left(\hat{\xi}^{\rm W}_{t'} - m'\right) \right\rangle = \delta(t-t'). \tag{B8}$$

According to the Lévy-Itô decomposition, any white noise is decomposed of the sum of a constant drift, the white Gaussian noise, and the white Poisson noise as given by Eq. (4). Thus, the white Gaussian and Poisson noises are the fundamental components of the Markovian noise sources.

APPENDIX C: REVIEW OF THE SYSTEM-SIZE EXPANSION FOR THE MARKOVIAN JUMP PROCESS

Let us briefly explain the system-size expansion for the Markovian jump process (9). With the scaling assumption (11), the master equation (8) can be rewritten as

$$\begin{aligned} \frac{\partial P_t(v)}{\partial t} &= \frac{1}{\varepsilon}\int_{-\infty}^{\infty} {\rm d}y \left[W\left(\frac{y}{\varepsilon}\middle| v-y\right) P_t(v-y) - W\left(\frac{y}{\varepsilon}\middle| v\right) P_t(v) \right] \\ &= \sum_{n=1}^{\infty} \frac{(-\varepsilon)^n}{n!} \frac{\partial^n}{\partial v^n} [\mathcal{A}_n(v) P_t(v)], \end{aligned} \tag{C1}$$

with the transformation $y = \varepsilon Y$ and the ε-independent KM coefficient defined by

$$\mathcal{A}_n(v) := \int_{-\infty}^{\infty} Y^n W(Y|v)\,{\rm d}Y. \tag{C2}$$

We assume the following stability conditions around $v = 0$:

(1) Linear stability: The first-order KM coefficient has a single stable point, such that

$$\mathcal{A}_1(0) = 0, \quad \gamma := -\frac{\partial}{\partial v}\mathcal{A}_1(v)\bigg|_{v=0} = -\mathcal{A}_1^{(1)}(0) > 0, \tag{C3}$$

with $\mathcal{A}_n^{(k)}(v) := \partial^k \mathcal{A}_n(v)/\partial v^k$.

(2) Existence of the noise term: The Gaussian noise term is assumed to be present even for $\varepsilon \to 0$, such that

$$\sigma^2 := \mathcal{A}_2(0) > 0. \tag{C4}$$

(3) Scaled variables: Furthermore, we apply the transformation of variables:

$$\mathsf{t} := \varepsilon t, \quad V := \frac{v}{\sqrt{\varepsilon}}. \tag{C5}$$

These scaled variables are introduced to focus on the long-time limit [i.e., $\mathsf{t} = O(1) \iff t = O(\varepsilon^{-1}) \gg 1$] and to enlarge the peak of the velocity PDF (i.e., $V = O(1) \iff v = O(\varepsilon^{1/2}) \ll 1$) in the small-noise limit.

With these assumptions, the KM series (C1) can be rewritten as

$$\begin{aligned} \frac{\partial P_t(V)}{\partial \mathsf{t}} &= \sum_{n=1}^{\infty}\sum_{k=0}^{\infty} \frac{\partial^n}{\partial V^n}\left[\sum_{k=1}^{\infty} V^k \frac{(-1)^n \varepsilon^{\frac{k+n}{2}-1}}{n!} \frac{\mathcal{A}_n^{(k)}(0)}{k!} P_t(V)\right] \\ &= \gamma \frac{\partial}{\partial V}[V P_t(V)] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial V^2} P_t(V) + o(\varepsilon^{1/2}), \end{aligned} \tag{C6}$$

where we applied the Taylor expansion of the nth-order KM coefficient:

$$\mathcal{A}_n(v) = \sum_{k=0}^{\infty} \frac{\varepsilon^{k/2} V^k}{k!} \mathcal{A}_n^{(k)}(0). \tag{C7}$$

In the small-noise limit $\varepsilon \to 0$, we obtain the Fokker-Planck (FP) equation,

$$\frac{\partial P_t(V)}{\partial \mathsf{t}} = \gamma \frac{\partial}{\partial V}[V P_t(V)] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial V^2} P_t(V), \tag{C8}$$

which is equivalent to the Langevin equation

$$\frac{{\rm d}\hat{V}}{{\rm d}\mathsf{t}} = -\gamma \hat{V} + \sigma \hat{\xi}^{\rm G}. \tag{C9}$$

APPENDIX D: DETAILED CALCULATIONS FOR THE SYSTEM-SIZE EXPANSION APPLIED TO THE NON-MARKOVIAN JUMP PROCESSES

This Appendix offers the detailed calculations for the system-size expansion applied to the non-Markovian jump processes. In particular, we derive Eqs. (48) and (50).

1. Derivation of Eq. (48)

With the assumptions in Sec. V A, let us formulate the system-size expansion for this model and derive the result (48). We have

$$\int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}(sz(s)P_t[\Gamma]) = \varepsilon \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}(uZ(u)P_t[\Gamma]) \tag{D1}$$

and

$$\begin{aligned}\left(\frac{\partial}{\partial v} + \int_0^\infty \mathrm{d}s \frac{\delta}{\delta z(s)}\right)^n \alpha_n[\Gamma]P_t[\Gamma] &= \left(\varepsilon^{-1/2}\frac{\partial}{\partial V} + \varepsilon^{-1/2}\int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}\right)^n \varepsilon^n \mathcal{A}_n[\Gamma]P_t[\Gamma] \\ &= \varepsilon^{n/2}\left(\frac{\partial}{\partial V} + \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}\right)^n \mathcal{A}_n[\Gamma]P_t[\Gamma].\end{aligned} \tag{D2}$$

We also consider the functional Maclaurin series (A12) for the KM coefficient around $\Gamma = \mathbf{0}$,

$$\begin{aligned}\mathcal{A}_n[\Gamma] &= \mathcal{A}_n[\mathbf{0}] + \sum_{k=1}^\infty \frac{1}{k!}\left(v\frac{\partial}{\partial \chi} + \int_0^\infty \mathrm{d}s z(s)\frac{\delta}{\delta\zeta(s)}\right)^k \mathcal{A}_n(\chi;\{\zeta(s)\}_s)\Big|_{(\chi,\{\zeta(s)\}_s)=\mathbf{0}} \\ &= \mathcal{A}_n[\mathbf{0}] + \sum_{k=1}^\infty \frac{\varepsilon^{k/2}}{k!}\left(V\frac{\partial}{\partial \chi} + \int_0^\infty \mathrm{d}u Z(u)\frac{\delta}{\delta\zeta(u)}\right)^k \mathcal{A}_n(\chi;\{\zeta(u)\}_u)\Big|_{(\chi,\{\zeta(u)\}_u)=\mathbf{0}},\end{aligned} \tag{D3}$$

with the dummy-variable arguments χ and ζ. This relation implies that

$$\begin{aligned}\mathcal{A}_1[\Gamma] &= \mathcal{A}_1[\mathbf{0}] + \sum_{k=1}^\infty \frac{\varepsilon^{k/2}}{k!}\left(V\frac{\partial}{\partial \chi} + \int_0^\infty \mathrm{d}u Z(u)\frac{\delta}{\delta\zeta(u)}\right)^k \mathcal{A}_1(\chi;\{\zeta(u)\}_u)\Big|_{(\chi,\{\zeta(u)\}_u)=\mathbf{0}} \\ &= 0 - \varepsilon^{1/2}\left(\gamma V + \int_0^\infty \mathrm{d}u \Upsilon(u)Z(u)\right) + O(\varepsilon)\end{aligned} \tag{D4}$$

and

$$\begin{aligned}\mathcal{A}_n[\Gamma] &= \mathcal{A}_n[\mathbf{0}] + \sum_{k=1}^\infty \frac{\varepsilon^{k/2}}{k!}\left(V\frac{\partial}{\partial \chi} + \int_0^\infty \mathrm{d}u Z(u)\frac{\delta}{\delta\zeta(u)}\right)^k \mathcal{A}_n(\chi;\{\zeta(u)\}_u)\Big|_{(\chi,\{\zeta(u)\}_u)=\mathbf{0}} \\ &= \mathcal{A}_n[\mathbf{0}] + O(\varepsilon^{1/2}) \quad \text{for } n \geqslant 2.\end{aligned} \tag{D5}$$

From the KM expansion (36a), by introducing $\mathrm{t} := \varepsilon t$, we obtain

$$\begin{aligned}\frac{\partial P_\mathrm{t}[G]}{\partial \mathrm{t}} &= \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}(uZ(u)P_\mathrm{t}[G]) + \sum_{n=1}^\infty \frac{(-1)^n\varepsilon^{n/2-1}}{n!}\left(\frac{\partial}{\partial V} + \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}\right)^n \mathcal{A}_n[G]P_\mathrm{t}[G] \\ &= \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}(uZ(u)P_\mathrm{t}[G]) + \left(\frac{\partial}{\partial V} + \int_0^\infty \mathrm{d}u_1 \frac{\delta}{\delta Z(u_1)}\right)\left(\gamma V + \int_0^\infty \mathrm{d}u_2 \Upsilon(u_2)Z(u_2)\right)P_\mathrm{t}[G] \\ &\quad + \frac{\sigma^2}{2}\left(\frac{\partial}{\partial V} + \int_0^\infty \mathrm{d}u \frac{\delta}{\delta Z(u)}\right)^2 P_\mathrm{t}[G] + O(\varepsilon^{1/2}).\end{aligned} \tag{D6}$$

We thus obtain the functional Fokker-Planck equation (48) in the weak coupling limit $\varepsilon \to 0$.

2. Derivation of Eq. (50)

Let us derive the GLE (50) as the leading-order approximation of the system-size expansion. We rewrite expression (49b) as

$$\frac{\partial \hat{Z}_\mathrm{t}(u)}{\partial \mathrm{t}} = -\int_0^\infty K(u,u')\hat{Z}_\mathrm{t}(u')\,\mathrm{d}u' - \gamma\hat{V}_\mathrm{t} + \sigma\hat{\xi}_\mathrm{t}^\mathrm{G}, \quad K(u,u') = u\delta(u-u') + \Upsilon(u'). \tag{D7}$$

By introducing the representation based on the eigenvectors $\{e(\mu;u)\}_\mu$ of $K(u,u')$,

$$\hat{w}_{\mathrm{t}}(\mu) := \int_0^\infty e^{-1}(u;\mu)\hat{Z}_{\mathrm{t}}(u)\,\mathrm{d}u, \tag{D8}$$

we formally obtain the explicit representation of $\hat{Z}_{\mathrm{t}}(u)$,

$$\hat{Z}_{\mathrm{t}}(u) = \int \hat{w}_{\mathrm{t}}(\mu)e(\mu;u)\,\mathrm{d}\mu, \tag{D9}$$

where we used Eq. (53). We thus obtain

$$\frac{\partial \hat{w}_{\mathrm{t}}(\mu)}{\partial \mathrm{t}} = -\mu\hat{w}_{\mathrm{t}}(\mu) + \kappa(\mu)\big(-\gamma\hat{V}_{\mathrm{t}} + \sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}}\big), \quad \kappa(\mu) := \int_0^\infty e^{-1}(u;\mu)\,\mathrm{d}u, \tag{D10}$$

whose solution is given by

$$\hat{w}_{\mathrm{t}}(\mu) = \kappa(\mu)\int_{-\infty}^{\mathrm{t}} e^{-\mu(\mathrm{t}-\mathrm{t}')}\big(\sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}'} - \gamma\hat{V}_{\mathrm{t}'}\big)\,\mathrm{d}\mathrm{t}', \tag{D11}$$

which leads to the explicit form of $\hat{Z}_{\mathrm{t}}$ as

$$\hat{Z}_{\mathrm{t}}(u) = \int \mathrm{d}\mu e(\mu;u)\kappa(\mu)\int_{-\infty}^{\mathrm{t}} \mathrm{d}\mathrm{t}' e^{-\mu(\mathrm{t}-\mathrm{t}')}\big(\sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}'} - \gamma\hat{V}_{\mathrm{t}'}\big) \tag{D12}$$

from Eq. (D9). From Eqs. (49a) and (D12), we obtain

$$\frac{\mathrm{d}\hat{V}_{\mathrm{t}}}{\mathrm{d}\mathrm{t}} = \sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}} - \gamma\hat{V}_{\mathrm{t}} - \int_{-\infty}^{\mathrm{t}}\mathrm{d}\mathrm{t}'\left[\int_0^\infty \mathrm{d}u\int_0^\infty \mathrm{d}\mu\kappa(\mu)\Upsilon(u)e(\mu;u)e^{-\mu(\mathrm{t}-\mathrm{t}')}\right]\big(\sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}'} - \gamma\hat{V}_{\mathrm{t}'}\big). \tag{D13}$$

This equation can be written as

$$\frac{\mathrm{d}\hat{V}_{\mathrm{t}}}{\mathrm{d}\mathrm{t}} = -\gamma\hat{V}_{\mathrm{t}} - \int_{-\infty}^{\mathrm{t}}\mathcal{M}(\mathrm{t}-\mathrm{t}')\hat{V}_{\mathrm{t}'}\,\mathrm{d}\mathrm{t}' + \hat{\eta}_{\mathrm{t}} \tag{D14}$$

with the memory kernel

$$\mathcal{M}(\mathrm{t}) := \int_0^\infty \nu(\mu)e^{-\mu\mathrm{t}}\,\mathrm{d}\mu, \quad \nu(\mu) := \int_0^\infty \kappa(\mu)\Upsilon(u)e(\mu;u)\,\mathrm{d}u \tag{D15}$$

and the colored Gaussian noise:[7]

$$\hat{\eta}_{\mathrm{t}} := \sigma\hat{\xi}^{\mathrm{G}}_{\mathrm{t}} + \sigma\int_{-\infty}^{\mathrm{t}}\mathrm{d}\mathrm{t}'\mathcal{M}(\mathrm{t}-\mathrm{t}')\hat{\xi}^{\mathrm{G}}_{\mathrm{t}'}. \tag{D16}$$

APPENDIX E: TRIVIAL MARKOVIAN EMBEDDING FOR DISCRETE-TIME STOCHASTIC PROCESSES

Here we show a trivial approach of Markovian embedding available only for discrete-time stochastic processes.

1. Discrete-time stochastic process and Markovian embedding

Let us consider a discrete-time stochastic difference equation,

$$\hat{x}_{t+1} = f(\hat{x}_t, \hat{x}_{t-1}, \dots, \hat{x}_{t-K}), \tag{E1}$$

with a positive integer $K > 0$. We assume f includes noise terms, in general, and can be stochastic, such as the ARIMA model.

This model can be trivially converted onto Markovian dynamics by introducing the phase-space vector

$$\hat{\boldsymbol{\Gamma}}_t := (\hat{x}_t, \hat{x}_{t-1}, \dots, \hat{x}_{t-K})^{\mathrm{T}}, \tag{E2}$$

with the superscript T signifying the transpose operator. Indeed, we obtain a first-order stochastic difference equation

$$\hat{\boldsymbol{\Gamma}}_{t+1} = S\hat{\boldsymbol{\Gamma}}_t + \boldsymbol{f}(\hat{\boldsymbol{\Gamma}}_t), \quad S := \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

$$\boldsymbol{f}(\hat{\boldsymbol{\Gamma}}_t) := \begin{pmatrix} f(\hat{\Gamma}_t) \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{E3}$$

[7] Any noise composed of a sum of Gaussian random numbers obeys the Gaussian statistics [1].

Here S is the finite-dimensional shifting operator, such that

$$\begin{pmatrix} 0 \\ \hat{x}_t \\ \hat{x}_{t-1} \\ \vdots \\ \hat{x}_{t-K-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{x}_t \\ \hat{x}_{t-1} \\ \hat{x}_{t-2} \\ \vdots \\ \hat{x}_{t-K} \end{pmatrix}. \tag{E4}$$

Similar ideas are used in econometrics [3] regarding the lag operator. If the time is discrete, this formulation can be straigtforwardly generalized even for $K \to \infty$, where the embedding dimension is infinite and thus the dynamics is truly non-Markovian.

2. Technical contribution of the Laplace-convolution representation

This fact implies that Markovian embedding is trivial for discrete-time stochastic processes. However, a straightforward generalization of this specific embedding is difficult for continuous-time stochastic processes. Indeed, it is challenging to generalize the shifting operator S for continuous-time representations, even at a formal level.

Let us attempt to write the formal continuous representation from the naive discrete-time embedding equation (E3). By considering the continuous limit with $K \to \infty$ and $dt \to 0$ for the time interval, let us write the phase-space vector as $\{\hat{\Gamma}_t(s)\}_{s\geqslant 0}$ parametrized with $s \geqslant 0$ defined by

$$\hat{\Gamma}_t(s) := \hat{x}_{t-s}. \tag{E5}$$

Equation (E3) can be formally written as

$$\hat{\Gamma}_{t+dt}(s) = \int_0^\infty ds' \delta(s - dt - s')\hat{\Gamma}_t(s') + f[\hat{\Gamma}_t]\delta_{s,0}, \tag{E6}$$

or, equivalently,

$$\frac{d\hat{\Gamma}_t(s)}{dt} = \int_0^\infty ds' K(s,s')\hat{\Gamma}_t(s') + f[\hat{\Gamma}_t]\delta_{s,0},$$
$$K(s,s') := \frac{d}{ds'}\delta(s-s'). \tag{E7}$$

This equation does not make sense even at the theoretical physics level due to the apparent singularity of the δ function and its derivative. Thus, the naive embedding (E1) for the discrete-time processes cannot be straightforwardly generalized toward the continuous-time processes, even at the formal level.

The Laplace-convolution representation technically solves this problem. The shifting operator S has an analytically tractable representation in the Laplace-convolution space, and thus the original non-Markovian dynamics is mapped onto a first-order Markovian SPDE.

APPENDIX F: EIGENVALUES AND EIGENFUNCTIONS OF THE MATRIX $K(u, u')$

Let us prove that all the eigenvalues of $K(u, u')$ defined by Eqs. (51) are real and positive. We define

$$\mathcal{K}(u,u') := \sqrt{\frac{\Upsilon(u)}{\Upsilon(u')}}K(u,u') = \sqrt{uu'}\delta(u-u') + \sqrt{\Upsilon(u)\Upsilon(u')}, \tag{F1}$$

where $\Upsilon(u)$ is defined by Eqs. (45). Since $\mathcal{K}$ is symmetric $[\mathcal{K}(u,u') = \mathcal{K}(u',u)]$, its eigenvalues $\tilde{\mu}$ are real, such that

$$\int_0^\infty du' \mathcal{K}(u,u')\tilde{e}(\tilde{\mu};u') = \tilde{\mu}\tilde{e}(\tilde{\mu};u), \quad \tilde{\mu} \in \boldsymbol{R}, \tag{F2}$$

with the eigenfunctions $\{\tilde{e}(\tilde{\mu};u)\}_{\tilde{\mu}}$. In addition, we find a positive-definite inequality for any function $f(u)$, such that

$$\int_0^\infty f(u_1)\mathcal{K}(u_1,u_2)f(u_2)\,du_1\,du_2$$
$$= \int_0^\infty uf^2(u)\,du + \left(\int_0^\infty \sqrt{\Upsilon(u)}f(u)\right)^2 du > 0, \tag{F3}$$

except for the trivial case $f(u) = 0$ for all u. This implies that the symmetric matrix $\mathcal{K}(u,u')$ is positive definite, and thus has only real eigenvalues. In addition, since all the eigenvalues are positive for the symmetric matrix $\mathcal{K}(u,u')$, it has an inverse matrix.

Finally, from the definition (F1), we find that

$$\int_0^\infty du' K(u,u')\frac{\tilde{e}(\tilde{\mu};u')}{\sqrt{\Upsilon(u')}} = \tilde{\mu}\frac{\tilde{e}(\tilde{\mu};u)}{\sqrt{\Upsilon(u)}}, \tag{F4}$$

implying that all the eigenvalues of $K(u,u')$ correspond to those of $\mathcal{K}(u,u')$, such that

$$\mu = \tilde{\mu}, \quad e(\mu;u) = \frac{\tilde{e}(\tilde{\mu};u)}{\sqrt{\Upsilon(u)}}. \tag{F5}$$

This means that all the eigenvalues of $K(u,u')$ are real and positive. Furthermore, $K(u,u')$ has an inverse matrix.

APPENDIX G: EXPLICIT RELATION WITH THE FIELD MASTER EQUATION FOR NONLINEAR HAWKES PROCESSES PREVIOUSLY DERIVED IN REFS. [34,35]

The fME (60) can be easily transformed. Let us define the following quantities:

$$s := \frac{1}{x}, \quad z'(x) := \tilde{h}'(x)z(s), \quad \tilde{h}'(x) := \frac{\tilde{h}(s)}{x^2}, \tag{G1}$$

satisfying $h(t) = \int_0^\infty \tilde{h}'(x)e^{-t/x}\,dx$ and $G'[z'] = g(\int_0^\infty \tilde{z}'(x)\,dx)$. By integrating $P_t[\Gamma]$ over $\hat{v}$ to define the marginal PDF

$$P_t[z'] := \int P_t[v;z]\,dv, \tag{G2}$$

we obtain from Eq. (60)

$$\frac{\partial P_t[z']}{\partial t} = \int \mathrm{d}x \frac{\delta}{\delta z'(s)}\left(\frac{z'(s)}{x}P_t[z']\right) + \int_{-\infty}^{\infty} \mathrm{d}y\rho(y)G'[z' - y\tilde{h}']P_t[z' - y\tilde{h}'] - G'[z']P_t[z']. \tag{G3}$$

This equation is equivalent to the fME in Refs. [34,35].

APPENDIX H: TIME-REVERSAL SYMMETRY OF THE MASTER EQUATION

We review the necessary and sufficient condition of the validity of time-reversal symmetry according to Ref. [22]. For a finite-dimensional Markovian stochastic processe $\boldsymbol{x} := (x_1, \ldots, x_K)^{\mathrm{T}}$, the general master equation is given by

$$\frac{\partial P_t(\boldsymbol{x})}{\partial t} = \left[-\sum_k \frac{\partial}{\partial x_k} A_k(\boldsymbol{x}) + \frac{1}{2}\sum_{k,l} \frac{\partial^2}{\partial x_k \partial_l} B_{kl}(\boldsymbol{x})\right] P_t(\boldsymbol{x}) + \int \mathrm{d}\boldsymbol{y}[\lambda(\boldsymbol{x}|\boldsymbol{y})P_t(\boldsymbol{y}) - \lambda(\boldsymbol{y}|\boldsymbol{x})P_t(\boldsymbol{x})], \tag{H1}$$

where A_k is the drift term, $B_{kl} \geqslant 0$ is the diffusion term, and $\lambda(\boldsymbol{y}|\boldsymbol{x}) \geqslant 0$ is the jump-intensity density for jumps from $\boldsymbol{x}$ to $\boldsymbol{y}$.

Let us define the time-reversal operator ϵ_k such that $\epsilon_k = 1$ if x_k is an even variable, and $\epsilon_k = -1$ if x_k is an odd variable. Typically, the velocity (position) is an odd (even) variable because it has odd (even) parity under time reversal. The necessary and sufficient condition for time-reversal symmetry to hold is given by

$$\lambda(\boldsymbol{y}|\boldsymbol{x})P_{\mathrm{ss}}(\boldsymbol{x}) = \lambda(\boldsymbol{\epsilon x}|\boldsymbol{\epsilon y})P_{\mathrm{ss}}(\boldsymbol{y}), \tag{H2a}$$

$$\epsilon_k A_k(\boldsymbol{\epsilon x})P_{\mathrm{ss}}(\boldsymbol{x}) = -A_k(\boldsymbol{x})P_{\mathrm{ss}}(\boldsymbol{x}) + \sum_l \frac{\partial}{\partial x_l}[B_{kl}P_{\mathrm{ss}}(\boldsymbol{x})], \tag{H2b}$$

$$\epsilon_k \epsilon_l B_{kl}(\boldsymbol{\epsilon x}) = B_{kl}(\boldsymbol{x}), \tag{H2c}$$

with

$$\boldsymbol{\epsilon} := \begin{pmatrix} \epsilon_1 & 0 & \ldots & 0 \\ 0 & \epsilon_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \epsilon_K \end{pmatrix}. \tag{H2d}$$

[1] R. Kubo, M. Toda, and N. Hashitsume, *Statsitical Physics II*, 2nd ed. (Springer-Verlag, Berlin, 1991).
[2] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 1992).
[3] J. D. Hamilton, *Time Series Analysis* (Princeton University Press, Princeton, NJ, 1994).
[4] J.-P. Bouchaud, J. Bonart, J. Donier, and M. Gould, *Trades, Quotes and Prices* (Cambridge University Press, Cambridge, 2018).
[5] A. Hawkes, Point spectra of some mutually exciting point processes, J. Royal Stat. Soc. Ser. B (Methodological) **33**, 438 (1971).
[6] A. Hawkes, Spectra of some self-exciting and mutually exciting point processes, Biometrika **58**, 83 (1971).
[7] A. Hawkes and D. Oakes, A cluster process representation of a self-exciting process, J. Appl. Probab. **11**, 493 (1974).
[8] A. G. Hawkes, Hawkes processes and their applications to finance: A review, Quant. Finance **18**, 193 (2018).
[9] C. G. Bowsher, Modelling security market events in continuous time: Intensity based, multivariate point process models, J. Econ. **141**, 876 (2007).
[10] V. A. Filimonov and D. Sornette, Quantifying reflexivity in financial markets: Toward a prediction of flash crashes, Phys. Rev. E **85**, 056108 (2012).
[11] E. Bacry, I. Mastromatteo, and J.-F. Muzy, Hawkes processes in finance, Mark. Microstruct. Liq. **01**, 1550005 (2015).
[12] 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).
[13] Y. Ogata, Statistical models for earthquake occurrences and residual analysis for point processes, J. Am. Stat. Assoc. **83**, 9 (1988).
[14] Y. Ogata, Seismicity analysis through point-process modeling: A review, Pure Appl. Geophys. **155**, 471 (1999).
[15] A. Helmstetter and D. Sornette, Subcritical and supercritical regimes in epidemic models of earthquake aftershocks, J. Geophys. Res. **107**, 2237 (2002).
[16] 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).
[17] 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).
[18] F. P. Schoenberg, M. Hoffmann, and R. J. Harrigan, A recursive point process model for infectious diseases, Ann. Inst. Stat. Math. **71**, 1271 (2019).
[19] S. K. Ram, S. Nandan, S. Boulebnane, and D. Sornette, Synchronized bursts of productivity and success in individual careers, Sci. Rep. **12**, 7637 (2022).
[20] S. K. Ram, S. Nandan, and D. Sornette, Significant hot hand effect in international cricket, Sci. Rep. **12**, 11663 (2022).
[21] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, Self-exciting point process modeling of crime, J. Am. Stat. Assoc. **106**, 100 (2011).
[22] C. W. Gardiner, *Handbook of Stochastic Methods*, 4th ed. (Springer, Berlin, 2009).
[23] H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1984).

[24] K. Kanazawa, *Statistical Mechanics for Athermal Fluctuation: Non-Gaussian Noise in Physics* (Springer, Berlin, 2017).

[25] 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).

[26] 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).

[27] R. Graham and T. Tél, On the weak-noise limit of Fokker-Planck models, J. Stat. Phys. **35**, 729 (1984).

[28] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).

[29] R. Kupferman, Fractional kinetics in Kac-Zwanzig heat bath models, J. Stat. Phys. **114**, 291 (2004).

[30] P. Siegle, I. Goychuk, P. Talkner, and P. Hänggi, Markovian embedding of non-Markovian superdiffusion, Phys. Rev. E **81**, 011136 (2010).

[31] J. Krafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, UK, 2011).

[32] K. Kanazawa and D. Sornette, Non-universal power law distribution of intensities of the self-excited Hawkes process: A field-theoretical approach, Phys. Rev. Lett. **125**, 138301 (2020).

[33] K. Kanazawa and D. Sornette, Field master equation theory of the self-excited Hawkes process, Phys. Rev. Res. **2**, 033442 (2020).

[34] K. Kanazawa and D. Sornette, Ubiquitous power law scaling in nonlinear self-excited Hawkes processes, Phys. Rev. Lett. **127**, 188301 (2021).

[35] K. Kanazawa and D. Sornette, Asymptotic solutions to nonlinear Hawkes processes: A systematic classification of the steady-state solutions, Phys. Rev. Res. **5**, 013067 (2023).

[36] H. Mori, A continued-fraction representation of the time-correlation functions, Prog. Theor. Phys. **34**, 399 (1965).

[37] G. Ford, M. Kac, and P. Mazur, Statistical mechanics of assemblies of coupled oscillators, J. Math. Phys. **6**, 504 (1965).

[38] G. Ford and M. Kac, On the quantum Langevin equation, J. Stat. Phys. **46**, 803 (1987).

[39] R. Zwanzig, Problems in nonlinear transport theory, in *Systems far from Equilibrium*, (Springer, Berlin, 1980).

[40] A. Dassios and H. Zhao, Exact simulation of Hawkes process with exponentially decaying intensity, Electron. Commun. Probab. **18**, 1 (2013).

[41] D. Hainaut, *Continuous Time Processes for Finance*, Bocconi Series in Mathematics, Statistics, Finance and Economics (Springer, Cham, Switzerland, 2022).

[42] A. Imamoglu, Stochastic wave-function approach to non-Markovian systems, Phys. Rev. A **50**, 3650 (1994).

[43] B. M. Garraway and P. L. Knight, Cavity modified quantum beats, Phys. Rev. A **54**, 3592 (1996).

[44] B. M. Garraway, Nonperturbative decay of an atomic system in a cavity, Phys. Rev. A **55**, 2290 (1997).

[45] D. Tamascelli, A. Smirne, S. F. Huelga, and M. B. Plenio, Nonperturbative treatment of non-Markovian dynamics of open quantum systems, Phys. Rev. Lett. **120**, 030402 (2018).

[46] D. Tamascelli, A. Smirne, J. Lim, S. F. Huelga, and M. B. Plenio, Efficient simulation of finite-temperature open quantum systems, Phys. Rev. Lett. **123**, 090402 (2019).

[47] A. E. Teretenkov, Non-Markovian evolution of multi-level system interacting with several reservoirs. exact and approximate, Lobachevskii J. Math. **40**, 1587 (2019).

[48] G. Pleasance, B. M. Garraway, and F. Petruccione, Generalized theory of pseudomodes for exact descriptions of non-Markovian quantum processes, Phys. Rev. Res. **2**, 043058 (2020).

[49] P. Blanc, J. Donier, and J.-P. Bouchaud, Quadratic Hawkes processes for financial prices, Quant. Financ. **17**, 171 (2017).

[50] C. Aubrun, M. Benzaquen, and J.-P. Bouchaud, Multivariate quadratic Hawkes processes–part I: Theoretical analysis, Quant. Financ. **23**, 741 (2023).

[51] L. Bergomi, *Stochastic Volatility Modeling*, 1st ed., Chapman and Hall/CRC Financial Mathematics Series (Chapman and Hall/CRC, Routledge, Boca Raton, London, New York, 2016).

[52] S. Nakajima, On quantum theory of transport phenomena: Steady diffusion, Prog. Theor. Phys. **20**, 948 (1958).

[53] H. Mori, Transport, collective motion, and Brownian motion, Prog. Theor. Phys. **33**, 423 (1965).

[54] R. Zwanzig, Ensemble method in the theory of irreversibility, J. Chem. Phys. **33**, 1338 (1960).

[55] R. Zwanzig, Memory effects in irreversible thermodynamics, Phys. Rev. **124**, 983 (1961).

[56] D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, 2nd ed. (Cambridge University Press, Cambridge, 2008).

[57] P. Résibois and M. de Leener, *Classical Kinetic Theory of Fluids* (Wiley, New York, 1977).

[58] J.-P. Hansen and I. McDonald, *Theory of Simple Liquids*, 3rd ed. (Academic Press, Amsterdam, 2006).

[59] H. Spohn, Kinetic equations from Hamiltonian dynamics: Markovian limits, Rev. Mod. Phys. **52**, 569 (1980).

[60] A. Dechant, private communication (2023).

[61] N. Shiraishi, An Introduction to Stochastic Thermodynamics (Springer-Nature, Singapore, 2023).

[62] K. Sekimoto, *Stochastic Energetics* (Springer, Heidelberg, 2010).