Thermodynamic length in stochastic thermodynamics of far-from-equilibrium systems: Unification of fluctuation relation and thermodynamic uncertainty relation

Atul Tanaji Mohite [*] and Heiko Rieger
Department of Theoretical Physics and Center for Biophysics, Saarland University, Saarbrücken, Germany

(Received 7 November 2025; accepted 11 June 2026; published 1 July 2026)

The Boltzmann distribution for an equilibrium system constrains the statistics of the system by the energetics. Despite the nonequilibrium generalization of the Boltzmann distribution being studied extensively, a unified framework valid for far-from-equilibrium discrete-state systems is lacking. Here, we derive an exact path-integral representation for discrete-state processes and represent it using the exponential of the action for stochastic transition dynamics. Solving the variational problem, the effective action is shown to be equal to the inferred entropy production rate (a thermodynamic quantity) and a nonquadratic dissipation function of the thermodynamic length (TL) defined for microscopic stochastic currents (a dynamic quantity). This formulates a far-from-equilibrium analog of the Boltzmann distribution, namely, the minimum action principle. The nonquadratic dissipation function is physically attributed to incorporating non-Gaussian fluctuations or far-from-equilibrium nonconservative driving. Further, an exact large deviation dynamical rate functional is derived. The equivalence of the variational formulation with the information geometric formulation is proved. The nonquadratic TL recovers the nonquadratic thermodynamic-kinetic uncertainty relation (TKUR) and the speed limits, which are tighter than the close-to-equilibrium quadratic formulations. Moreover, if the transition affinities are known, then the nonquadratic TL recovers the fluctuation relation (FR). The minimum action principle manifests the nonquadratic TKUR and FR as two faces corresponding to the thermodynamic inference and partial control descriptions, respectively. In addition, the validity of these results is extended to coarse-grained observable currents, strengthening the experimental/numerical applicability of them.

DOI: 10.1103/fmlt-px9y

I. INTRODUCTION

The Boltzmann distribution is the most fundamental principle in statistical physics. It formulates an equivalence between thermodynamics and statistics for equilibrium systems, valid in the thermodynamic limit [1]. Finite-size/particle systems prone to nonequilibrium fluctuations are ubiquitous and violate the assumption of the thermodynamic limit. By relaxing the assumption of the thermodynamic limit, the framework of stochastic thermodynamics (ST) enables to define thermodynamic quantities for the stochastic transition of a microscopic system [2–4]. In ST, the thermodynamic dissipation cost to sustain nonequilibrium fluctuations and/or driving is quantified by the entropy production rate (EPR). Recently, ST has been extended to "nonreciprocal" systems that violate "actio-reactio" symmetry, and to coarse-grained macroscopic systems [5,6] due to an exact thermodynamically consistent coarse-graining of microscopic systems [6]. This has cemented the applicability of ST to experimentally/practically relevant real-world systems, such as nonideal particles and coarse-grained mesoscopic/macroscopic degrees of freedom.

The fluctuation relation (FR) is a fundamental seminal law in ST, which connects the time-reversal asymmetry of dynamics to the stochastic thermodynamic cost [2–5,7–26]. The first-order mean-field approximation of FR recovers the second law of thermodynamics (an approximate law). Recently, the thermodynamic-kinetic uncertainty relation (TKUR) has revealed a lower bound on thermodynamic dissipation (a thermodynamic quantity) using the current precision (a dynamic quantity) [27–34]. TKUR obtains a tighter lower bound on the thermodynamic dissipation required to sustain a nonequilibrium process than the second law of thermodynamics. TKUR's relation to speed limits (SL) has been explored [35–39]. FR and TKUR are often treated as distinct fundamental laws in ST, and their connection is missing. TKUR has been derived using FR [40–43], but the lower bound obtained on dissipation was loose [40–43]. Although TKUR has a practical advantage for thermodynamic inference, in contrast to FR, the fundamental/seminal origin of TKUR is debatable, which we will prove.

Nonequilibrium generalizations of the Boltzmann distribution have been explored extensively [26,44–69], and its applications to biological systems are studied [70–73]. However, the nonequilibrium generalization of the Boltzmann distribution has three major drawbacks. First, a Gaussian approximation for fluctuations/driving, which is identified by a quadratic relation between EPR and driving forces/fluctuations [74–80].

[*]Contact author: atul.mohite@uni-saarland.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

The Gaussian approximation for fluctuations was originally derived to study close-to-equilibrium (cEQ) systems [81,82] and extended to path-integral formulism around the mean-field description [83–86]: a *top-down approach* toward state-space fluctuations. However, non-Gaussian fluctuations are important for far-from-equilibrium (fEQ) or finitely small size systems. Notable non-Gaussian state-space formulations assume a specific form of the non-Gaussian noise rather than its microscopic derivation; again, a *top-down approach* to state-space fluctuations [87,88]. Second, except for a few notable exceptions, existing works focus on the state-space representation of fEQ processes. However, stochasticity in physical systems originates from microscopic transitions, which are the most fundamental level of description. Moreover, state-space formulations fail to distinguish between multiple transition channels connecting the same states, for example, those arising from coupling to different thermodynamic reservoirs or from distinct effective driving mechanisms. Therefore, a transition-space representation of fEQ systems is necessary, since, a state-space formulation is a cEQ methodology that does not fully capture the physical aspects of transitions in fEQ systems. Third, a coherent and unified description of fEQ systems grounded in a single underlying principle is missing, due to contradictions between different formulations, which is again attributed to a phenomenological *top-down approach* toward fEQ principles.

In this work, we derive the minimum action principle for the EPR of discrete-state processes [89]. To this end, we use the second quantization method, namely, the Doi-Peliti field theory (DPFT), which preserves non-Gaussian transition fluctuations due to its *bottom-up construction* [6,90–94]. We derive an exact transition probability measure for discrete-state processes, which is equal to the exponential of the action. Hence, a variational formulation for discrete-state processes is formulated, namely, the "minimum action principle" (MinAP). We prove that the effective action Lagrangian is equivalent to inferred EPR (a thermodynamic quantity). The Lagrangian is shown to be a nonquadratic function of the cumulants of the microscopic stochastic transition currents (a dynamic quantity) and quantifies the thermodynamic length of stochastic currents. The threefold equality between the transition probability measure, inferred EPR, and current cumulants formulates a far-from-equilibrium analog of the Boltzmann distribution.

Using the thermodynamic length (TL) [39,95–101], we demonstrate that the variational formulation yields a nonquadratic TKUR, which provides a tighter bound than the quadratic TKUR. If the transition affinities are known (a partial control description), then the Lagrangian reduces to the bilinear form of EPR, which recovers FR for microscopic stochastic currents. Using TL, the nonquadratic TKUR and FR are unified within MinAP; they correspond to the thermodynamic inference and partial control descriptions, respectively. Further, we derive the exact large deviation functional for discrete-state processes and discuss its importance compared to the Gaussian and Hessian approximations of the large deviation functional. We extend the validity of MinAP for coarse-grained observable stochastic currents. To this end, we solve the variational problem for the Lagrangian under the constraint imposed by the observable currents. We prove that MinAP is extended to coarse-grained observable stochastic currents with the same underlying framework/structure as the one derived for microscopic transition currents.

Moreover, the variational formulation broadens the numerical applicability of MinAP [102], where an exact analytical solution is not feasible, which is usually the case beyond a few exactly solvable prototypical models. For example, variational formulations have been utilized to study the nonequilibrium dynamics: phase transitions, first-passage times, metastability in stochastic systems [103–111], applications in machine learning [112–114], and for thermodynamic inference in ST [115–119]. Hence, the variational formulation of far-from-equilibrium discrete state systems allows us to explore its applications in the future and broadens the practical applicability of ST. As an application of MinAP, assuming "full control" of transition affinities and mobilities, we develop *Generalized finite-time optimal control framework* in Ref. [120].

II. MINIMUM ACTION PRINCIPLE

A. Setup

1. Thermodynamically consistent discrete-state processes and graphs

The discrete-state systems are represented by a graph, such that the state and transition form the nodes and directional edges of the graph, respectively [7]. The set of all states is denoted by $\{i\}$. The state probability/density and transition between the states are denoted by ρ_i and γ, respectively. For each forward unidirectional transition γ between reactant state ρ_{r_γ} to product state ρ_{p_γ}, there exists a backward unidirectional transition $-\gamma$ [121]. The state probability and density on the graph correspond to the physical descriptions provided by Markov jump processes and linear chemical reaction networks, respectively. The intermediate coarse-graining step between them and the resulting hydrodynamic description is elucidated in Ref. [6]. The transition currents for the forward and backward reactions are denoted by $j_\gamma = \rho_{r_\gamma} k_\gamma$ and $j_{-\gamma} = \rho_{p_\gamma} k_{-\gamma}$, respectively, where k_γ and $k_{-\gamma}$ are the forward and backward transition rates. The set of all unidirectional and bidirectional transitions of a graph is denoted by $\{\gamma^{\rightarrow}\}$ and $\{\gamma^{\rightleftharpoons}\}$, respectively.

The thermodynamic consistency condition for the transition currents implies that they satisfy the local detailed balance (LDB) condition, $A_\gamma = \ln(j_\gamma / j_{-\gamma}) = F_\gamma - \Delta_\gamma E + \Delta_\gamma S^{\text{state}}$. Here, A_γ is the transition affinity that quantifies the time-reversal asymmetry or the directionality of the transition currents. It is decomposed into three terms: an external nonconservative nonequilibrium driving F_γ supported by a thermodynamic reservoir, change in the equilibrium energy functional $\Delta_\gamma E$ due to transition γ, and the change in the state entropy $S_i^{\text{state}} = -\ln(\rho_i)$ [2]. Here, the conservative force is represented as a change of chemical potential (μ) between the reactant and product state $-\Delta_\gamma E = \mu_{r_\gamma} - \mu_{p_\gamma}$ [5]. The energy E is controlled by the set of parameters $\{\lambda\}$. The equilibrium Boltzmann distribution is given by $\rho_i^E = e^{-E_i + \psi_E}$, with the corresponding equilibrium free energy $\psi_E = -\ln(\sum_{\{i\}} e^{-E_i})$. The formulation of the LDB condition in ST relies on timescale separation and weak

coupling between the system degrees of freedom and the environment. Nevertheless, LDB allows one to exactly quantify the thermodynamic cost associated with any stochastic change of the system's state. This is of paramount importance for establishing the physical connection between the statistical and thermodynamic properties of fEQ systems, much like the Boltzmann distribution does for equilibrium systems. However, the derivation of MinAP is valid irrespective of the LDB condition.

The total bidirectional current and traffic for the bidirectional transition $\gamma^{\rightleftharpoons}$ are defined as $J_\gamma = j_\gamma - j_{-\gamma}$ and $T_\gamma = j_\gamma + j_{-\gamma}$. They correspond to the decomposition of bidirectional currents into linearly independent time-antisymmetric and time-symmetric parts, respectively. The scaled traffic physically quantifies the variance of the current, with a large deviation scaling parameter that plays a role similar to the inverse temperature for equilibrium systems [5,6,122]. For example, the observation time τ for dynamical systems [26,122,123], and the system volume $\mathcal{V}$ for fluctuating hydrodynamic description [5,6,26,79] are the relevant large deviation scaling parameters. Defining the mobility $D_\gamma = \sqrt{j_\gamma j_{-\gamma}} = \sqrt{\rho_{p_\gamma}\rho_{r_\gamma} e^{\mu_{p_\gamma}+\mu_{r_\gamma}}}$ for $\gamma^{\rightleftharpoons}$, mean currents and traffics satisfy relations with affinities and mobilities, $J_\gamma = 2D_\gamma \sinh(A_\gamma/2)$ and $T_\gamma = 2D_\gamma \cosh(A_\gamma/2)$, which justifies the nomenclature of mobility.

2. Dynamics

The continuity equation constrains the temporal state-space flow generated due to transitions; it reads

$$\partial_t \vec{\rho} = \mathbb{S}\vec{J}, \tag{1}$$

where $\mathbb{S}$ is the stoichiometry matrix for the graph that dictates the contraction from the transition space to the state space [7]. The states and transitions are represented in the column vector notation $\vec{\rho} = (.., \rho_i, ..)^T$ and $\vec{J} = (.., J_\gamma, ..)^T$. The state space and transition space have dimensions $|\{i\}|$ and $|\{\gamma^{\rightleftharpoons}\}|$, respectively. Hence, $\mathbb{S}$ is a matrix of dimension $|\{i\}| \times |\{\gamma^{\rightleftharpoons}\}|$. Here, $||$ denotes the dimension of the set. For a transition $\gamma^{\rightleftharpoons}$, $\mathbb{S}_\gamma$ is the row vector corresponding to the transition $\gamma^{\rightleftharpoons}$. The negative and positive entries of $\mathbb{S}_\gamma$ correspond to the reactant and product species vectors, respectively. For a graph, $\vec{r}_\gamma$ and $\vec{p}_\gamma$ have only one nonvanishing entry denoted by index r_γ and p_γ for the reactant and product, respectively. The reactant and product indices r_γ and p_γ for each transition γ dictates the contraction from the transition-space ($\gamma \in \{\gamma^{\rightleftharpoons}\}$) to the state space ($r_\gamma, p_\gamma \in \{i\}$).

To highlight the importance of the transition-space representation of fEQ systems discussed in the introduction, let us consider two conformations of a protein "1" and "2," with transitions occurring between them via an equilibrium thermodynamic reservoir and an enzymatic thermodynamic reservoir that favors the product formation through an effective nonequilibrium driving F_{enz}. For this prototypical system, $\{i\} = \{1, 2\}$ and $\{\gamma^{\rightleftharpoons}\} = \{\text{eq}, \text{enz}\}$. Both transition channels connect the same pair of states, such that $r_{\text{eq}} = 1$, $p_{\text{eq}} = 2$ and $r_{\text{enz}} = 1$, $p_{\text{enz}} = 2$. But the thermodynamic cost associated with the transition depends on the channel taken: $A_{\text{eq}} = \mu_1 - \mu_2 + \ln(\rho_1) - \ln(\rho_2)$ and $A_{\text{enz}} = F_{\text{enz}} + \mu_1 - \mu_2 + \ln(\rho_1) - \ln(\rho_2)$. Therefore, the transition space representation of the microscopic dynamics preserves the physical (both dynamical and thermodynamic) distinction between different transition channels connecting the same states. Hence, throughout this work, we employ the transition space representation. If a state-space description is required, then one may invoke the continuity equation to contract the dynamics from transition space to state space.

3. Orthogonal decomposition of thermodynamic dissipation for graphs

The mean EPR $\langle\dot{\Sigma}\rangle$ for the graph has a bilinear form, namely, force times current [7],

$$\langle\dot{\Sigma}\rangle = \sum_{\{\gamma^{\rightleftharpoons}\}} J_\gamma A_\gamma, \tag{2}$$

$\langle\dot{\Sigma}\rangle$ is further decomposed into its three linearly independent contributions,

$$\begin{aligned} -\dot{\psi}_E &= -\dot{\lambda}\,\partial_\lambda \psi_E, \\ \langle\dot{\Sigma}_E^{\text{ex}}\rangle &= -\sum_{\{i\}} d_t\rho_i \ln\left(\frac{\rho_i}{\rho_i^E}\right) = -d_t D_E^{\text{KL}}, \\ \langle\dot{\Sigma}^{\text{hk}}\rangle &= \sum_{\{\gamma^{\rightleftharpoons}\}} T_\gamma^{\perp} F_\gamma \sinh\left(\frac{F_\gamma}{2}\right), \end{aligned} \tag{3}$$

the quasistatic driving work rate ($-\dot{\psi}_E$), the excess EPR ($\langle\dot{\Sigma}^{\text{ex}}\rangle$), and the housekeeping EPR ($\langle\dot{\Sigma}^{\text{hk}}\rangle$), respectively. They physically correspond to the instantaneous change in free energy implemented by changing control parameters, the statistical distance of instantaneous state distribution $\{\rho_i\}$ from the Boltzmann distribution for states $\{\rho_i^E\}$, and the thermodynamic cost of maintaining nonconservative driving $\{F_\gamma\}$, respectively. They are defined in $\{\lambda\}$, $\{i\}$, and $\{\gamma^{\rightleftharpoons}\}$ space, respectively. $-\dot{\psi}_E$ and $\langle\dot{\Sigma}_E^{\text{ex}}\rangle$ do not require information on the transition topology. Thus, they correspond to dissipation terms that are integrated using Eq. 1 to obtain boundary terms for EPR in the control parameter space and state space, respectively.

The excess-housekeeping decomposition of the EPR is formulated by identifying the conservative and nonconservative decompositions of the transition affinity, which satisfy time-reversal symmetry and antisymmetry, respectively. The decomposition of the affinity is $A_\gamma = F_\gamma + \Delta_\gamma S_i^{\text{state}/E} = A_\gamma^{\text{nc}} + A_\gamma^{\text{rel}}$ with the relative state entropy defined with respect to the Boltzmann distribution $S_i^{\text{state}/E} = -\ln(\rho_i/\rho_i^E)$ for fixed values of $\{\lambda\}$. A_γ^{rel} and A_γ^{nc} are transition affinities that generate instantaneous (short-time) and steady-state dynamics given by Eq. (1). Physically, this symmetry is a manifestation of the short-time relaxation and long-time steady-state symmetry of currents. This symmetry generates an orthogonal decomposition of EPR, and $\langle\dot{\Sigma}^{\text{hk}}\rangle$ is simplified to, $\langle\dot{\Sigma}^{\text{hk}}\rangle = \sum_{\{\gamma^{\rightleftharpoons}\}} F_\gamma J_\gamma^a$, where J_γ^a is the antisymmetric part of J_γ under the orthogonal transformation, $F_\gamma \to -F_\gamma$ [5,6]. This resolves to $\langle\dot{\Sigma}^{\text{hk}}\rangle = \sum_{\{\gamma^{\rightleftharpoons}\}} T_\gamma^{\perp} F_\gamma \sinh(F_\gamma/2)$ in Eq. (3), where $T_\gamma^{\perp} = T_\gamma|_{F_\gamma=0}$ is the equilibrium traffic obtained by plugging in $F_\gamma = 0$ (the direction orthogonal to nonconservative driving). This implies that the mean housekeeping EPR is equal to the equilibrium traffic multiplied by a nonquadratic function of the noncon-

servative driving force F_γ [5,6]. Putting in the steady-state distribution as the reference Boltzmann distribution $E = \text{ss}$, one obtains the usual adiabatic-nonadiabatic decomposition [2], a subcase of the orthogonal decomposition.

B. Variational formulation

1. Derivation using DPFT

We derive the exact path-integral representation of discrete-state processes using an exact second-quantized approach that preserves non-Gaussian transition statistics, namely, DPFT [6,90–94].

Microscopic transitions represented as a second-quantized Hamiltonian. In DPFT, the second-quantized representation for the creation and annihilation of state i is given by creation and annihilation operators $\hat{\eta}_i^\dagger$ and $\hat{\eta}_i$, respectively. Hence, $\hat{\eta}_{p_\gamma}^\dagger$ and $\hat{\eta}_{r_\gamma}$ correspond to the creation of the product state and the annihilation of the reactant state for the transition $\gamma^{\rightharpoonup}$. The corresponding inverse transition $\gamma^{\leftharpoondown}$ is represented in the second-quantized form by $\hat{\eta}_{r_\gamma}^\dagger$ and $\hat{\eta}_{p_\gamma}$ for the creation of the reactant state and the annihilation of the product state. Using Doi-Peliti field theory, the second-quantized Hamiltonian operator for the bidirectional transition $\gamma^{\rightleftharpoons}$ reads [6,90–94]

$$\hat{H}_\gamma\big[\hat{\eta}_{\mathbf{r}_\gamma}^\dagger, \hat{\eta}_{\mathbf{p}_\gamma}^\dagger, \hat{\eta}_{\mathbf{r}_\gamma}, \hat{\eta}_{\mathbf{p}_\gamma}\big] = \big(\hat{\eta}_{\mathbf{r}_\gamma}^\dagger - \hat{\eta}_{\mathbf{p}_\gamma}^\dagger\big)\big(\hat{\eta}_{\mathbf{r}_\gamma} k_\gamma - \hat{\eta}_{\mathbf{p}_\gamma} k_{-\gamma}\big). \quad (4)$$

The transition rates k_γ and $k_{-\gamma}$ are assumed not to depend on ρ_i. A state dependence of the transition rates $\{k_\gamma\}$ would create a technical sophistication in implementing DPFT, discussed later [6]. It can be addressed by a careful implementation of DPFT for the state-dependent $\{k_\gamma\}$ and represents a technical intricacy rather than a conceptual one [6]. The second-quantized Hamiltonian for all linearly independent microscopic transitions reads [6,90–94]

$$\hat{H}[\{\hat{\eta}_i^\dagger, \hat{\eta}_i\}] = \sum_{\{\gamma^{\rightleftharpoons}\}} \hat{H}_\gamma\big[\hat{\eta}_{\mathbf{r}_\gamma}^\dagger, \hat{\eta}_{\mathbf{p}_\gamma}^\dagger, \hat{\eta}_{\mathbf{r}_\gamma}, \hat{\eta}_{\mathbf{p}_\gamma}\big]. \quad (5)$$

Coherent state and master equation. The coherent state $|\phi_i\rangle$ and its corresponding dual state $\langle\phi_i^*|$ are defined as [6,90–94]

$$|\phi_i\rangle = \sum_l \frac{(\phi_i)^l (\hat{\eta}_i^\dagger)^l}{l!}|0\rangle_i, \quad \langle\phi_i^*| = {}_i\langle 0| \sum_l \frac{(\phi_i^*)^l (\hat{\eta}_i)^l}{l!}. \quad (6)$$

Here, ϕ_i is the eigenvalue of the coherent state with complex conjugate ϕ_i^* and satisfies $\hat{\eta}_i|\phi_i\rangle = \phi_i|\phi_i\rangle$ and $\langle\phi_i|\hat{\eta}_i^\dagger = \langle\phi_i|\phi_i^*$. The physical interpretation of the coherent state Eq. (6) is visualized using its alternative representation, $|\phi_i\rangle = \sum_l P_i(l)|l_i\rangle$, which gives the sum over a Poissonian probability distribution for the state occupancy l for state i ρ_i with $P_i(l) = (\phi_i)^l/l!$. Then, the master Eq. (1) for the state probability flow using the coherent state has the second-quantized form $\partial_t|\{\phi_i\}\rangle = -\hat{H}[\{\hat{\eta}_i^\dagger, \hat{\eta}_i\}]|\{\phi_i\}\rangle$ [6,90–94].

Therefore, the inner product of an operator $\hat{O}$ using the coherent state is equivalent to computing the expectation value over a Poissonian probability measure for the state occupancy ($\mathcal{O}[\{\phi_i^*, \phi_i\}] = \langle\phi_i^*|\hat{O}[\{\hat{\eta}_i^\dagger, \hat{\eta}_i\}]|\phi_i\rangle/\langle\phi_i^*|\phi_i\rangle$). The coherent state and its dual are eigenvectors of $\hat{\eta}_i|\phi_i\rangle = \phi_i|\phi_i\rangle$ and $\langle\phi_i|\hat{\eta}_i^\dagger = \langle\phi_i|\phi_i^*$. Thus, computing the expectation value of the operator over Poissonian occupancy of the states is equivalent to replacing $\hat{\eta}_i \to \phi_i$ and $\hat{\eta}_i^\dagger \to \phi_i^*$ if $\hat{O}[\{\hat{\eta}_i^\dagger, \hat{\eta}_i\}]$ is normal ordered [6]. This highlights the reason for choosing transition rates $\{k_\gamma\}$ that are independent of the state occupancies $\{\rho_i\}$, since it adds another sophistication to obtain a normal-ordered expression for $\{k_\gamma\}$ in the transition Hamiltonian Eq. (4) [6]. Using the tensor product, the coherent state $|\{\phi_i\}\rangle = \prod \otimes|\phi_i\rangle$ and its conjugate dual state $\langle\{\phi_i^*\}| = \prod \otimes\langle\phi_i|$ for the set of all microstates are defined.

The mesoscopic Doi-Peliti Action, Lagrangian and Hamiltonian in the creation-annihilation picture. Using DPFT [6,90–94], the exact "stochastic" path integral formulation for the transition probability measure for the discrete-state process written using the eigenvalues of the coherent state reads [124],

$$\mathcal{P}[\{\phi_i^*, \phi_i\}] = e^{-\mathcal{S}_{\text{DP}}[\{\phi_i^*, \phi_i\}]}. \quad (7)$$

Using ϕ_i, ϕ_i^* and Eqs. (4) and (5), $\mathcal{S}_{\text{DP}}[\{\phi_i^*, \phi_i\}]$ derived for the transition dynamics reads [6,90–94]

$$\mathcal{S}_{\text{DP}}[\{\phi_i^*, \phi_i\}] = \int_{t_i}^{t_f} dt\, \mathcal{L}[\{\phi_i^*, \phi_i\}], \quad (8)$$

with the mesoscopic Lagrangian $\mathcal{L}[\{\phi_i^*, \phi_i\}]$ and the mesoscopic Hamiltonian $\mathcal{H}[\{\phi_i^*, \phi_i\}]$ are

$$\mathcal{L}[\{\phi_i^*, \phi_i\}] = -\sum_{\{i\}} \phi_i \partial_t \phi_i^* - \mathcal{H}[\{\phi_i^*, \phi_i\}],$$
$$\mathcal{H}[\{\phi_i^*, \phi_i\}] = -\sum_{\{\gamma^{\rightleftharpoons}\}} \big(\phi_{r_\gamma}^* - \phi_{p_\gamma}^*\big)\big(\phi_{r_\gamma} k_\gamma - \phi_{p_\gamma} k_{-\gamma}\big). \quad (9)$$

In the definition of $\mathcal{L}[\{\phi_i, \phi_i^*\}]$, the first and second terms lie in the state space and transition space, respectively. Their physical origin is attributed to the left and right sides of the Master Eq. (1). Excluding the boundary terms, the first term of $\mathcal{L}[\{\phi_i^*, \phi_i\}]$ is reorganized and equal to $\sum_{\{i\}} \phi_i^* \partial_t \phi_i$.

Cole-Hopf transform and mesoscopic Doi-Peliti Lagrangian in the density-affinity picture. Equations (7)–(9) formulate the "stochastic" path integral representation of the discrete-state process in the second-quantized coherent-state picture. However, the coherant state eigenvalues ϕ_i and ϕ_i^* do not have intuitive physical meaning for classical stochastic systems [6,125]. The Cole-Hopf transform transforms the complex eigenvalue fields ϕ_i and ϕ_i^* to classical fields: the state probability/density (ρ_i) and the corresponding conjugate field (χ_i) [6,125]. The conjugate field χ_i has been referred to by multiple names: response field, noise field, or bias field. However, we will stick to its mathematical nomenclature convention, namely, the conjugate field that generates state-cumulants.

We define the Cole-Hopf transform as $\phi_i = \rho_i e^{\chi_i}$ and $\phi_i^* = e^{-\chi_i}$. Compared to the usual definition of Cole-Hopf transform [125], we have defined the Cole-Hopf transform with a negative sign in χ_i. This convention allows χ_i to have the same sign as the transition affinity A_γ or the chemical potential μ_i of the state ρ_i. Hence, it gives a physical and thermodynamic interpretation to conjugate fields $\{\chi_i\}$, elaborated below. The Cole-Hopf transform changes the Lagrangian $\mathcal{L}[\{\phi_i^*, \phi_i\}]$ to $\mathcal{L}[\{\rho_i; \chi_i\}] = \mathcal{L}[\{\phi_i^*, \phi_i\}]|_{CF}$, in a physically intuitive and

legible form,

$$\mathcal{L}[\{\rho_i;\chi_i\}] = \vec{\chi}\cdot\partial_t\vec{\rho} - \mathcal{H}[\{\rho_i;\chi_i\}],$$
$$\mathcal{H}[\{\rho_i;\chi_i\}] = \sum_{\{\gamma^{\rightleftharpoons}\}}\left[j_\gamma\left(e^{\chi_{r_\gamma}-\chi_{p_\gamma}}-1\right)+j_{-\gamma}\left(e^{\chi_{p_\gamma}-\chi_{r_\gamma}}-1\right)\right]. \tag{10}$$

2. Doi-Peliti Lagrangian in the current-affinity picture

The first and second terms of $\mathcal{L}[\{\rho_i;\chi_i\}]$ lie in the state space and the transition space, respectively, similar to Eq. (8). Hence, our objective is to represent the Lagrangian $\mathcal{L}[\{\rho_i;\chi_i\}]$ in the transition space, since the microscopic transitions are the most fundamental physical origin of the system's stochasticity. To this end, the first term of $\mathcal{L}[\{\rho_i;\chi_i\}]$ is rewritten using the continuity Eq. (1), which equals replacing $\partial_t\rho_i$ by j_γ for all transitions $\gamma^{\rightleftharpoons}$ that contribute to the temporal change of state ρ_i due to the continuity equation. State-space conjugate fields are also represented in the transition space by defining $\chi_\gamma = \chi_{p_\gamma} - \chi_{r_\gamma}$ and $\chi_{-\gamma} = \chi_{r_\gamma} - \chi_{p_\gamma}$, hence, by definition, χ_γ satisfies the topological constraint $\chi_{-\gamma} = -\chi_\gamma$.

This reduces the state-space representation of the Lagrangian $\mathcal{L}[\{\rho_i;\chi_i\}]$ to the transition-space representation $\mathcal{L}[\{j_\gamma;\chi_\gamma\}]$ for each unidirectional transition $\gamma^{\rightarrow}$:

$$\mathcal{L}[\{j_\gamma;\chi_\gamma\}] = \sum_{\{\gamma^{\rightarrow}\}} j_\gamma\left(\chi_\gamma + 1 - e^{-\chi_\gamma}\right). \tag{11}$$

Furthermore, recalling definitions $\chi_{-\gamma} = -\chi_\gamma$ and $J_\gamma = j_\gamma - j_{-\gamma}$ and $T_\gamma = j_\gamma + j_{-\gamma}$, Eq. (11) is simplified to

$$\mathcal{L}[\{J_\gamma, T_\gamma;\chi_\gamma\}] = \sum_{\{\gamma^{\rightleftharpoons}\}}[J_\gamma(\chi_\gamma + \sinh(\chi_\gamma)) + T_\gamma(1-\cosh(\chi_\gamma))], \tag{12}$$

defined for the set of all bidirectional transitions. Equation (12) incorporates time-reversal symmetry/asymmetry due to the topological constraint imposed on χ_γ. It reveals that the symmetric and antisymmetric parts of transition currents are coupled to a single conjugate field χ_γ. Moreover, the time antisymmetric term J_γ and the time-symmetric term T_γ are coupled to odd and even powers of χ_γ, respectively. This structure of Eq. (12) reveals that the statistical properties of all higher-order transition fluctuations are effectively encapsulated by the first and second cumulants (J_γ and T_γ), due to the nonlinear coupling to the conjugate field χ_γ.

Combining Eqs. (7) and (8) with Eq. (12), the "stochastic" path integral representation of the discrete-state processes in the current-affinity picture is given by the transition probability measure,

$$\mathcal{P}[\{J_\gamma, T_\gamma;\chi_\gamma\}] = e^{-\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma;\chi_\gamma\}]}, \tag{13}$$

and the corresponding Doi-Peliti action and Lagrangian:

$$\mathcal{S}_{\rm DP}[\{J_\gamma, T_\gamma;\chi_\gamma\}] = \int_{t_i}^{t_f} dt\,\mathcal{L}[\{J_\gamma, T_\gamma;\chi_\gamma\}]. \tag{14}$$

The coherent-state sums over all Poissonian realization of the state, which is encapsulated in its eigenvalues, but the stochasticity associated with the microscopic stochastic transition between states is preserved through the conjugate fields $\{\chi_\gamma\}$, which couples nonlinearly to both J_γ and T_γ, and the transition probability measure is quantified by Eq. (13). The Stratonovich-Hubbard transform is an example of such path-integral formulism, where a conjugate field χ_o couples linearly to an observable $\mathcal{O}$, and quantifies it's gaussian fluctuations (stochasticity of the observable $\mathcal{O}$) with a quadratic coupling to χ_o, the validity of this approach relies on cEQ Gaussian approximation: *a top-down approach* toward the fluctuation of the relevant coarse-grained observable [81,82], detailed discussion in Sec. II C 2. However, the nonlinear coupling of χ_γ to J_γ and T_γ in Eqs. (11) and (12) is attributed to incorporating Poissonian transitions between states, due to the exact *bottom-up approach* toward the transition fluctuations. The set of Eqs. (12) to (14) concludes the derivation of the exact "stochastic" path integral formulism for discrete-state processes [94], which incorporates poissonian transition fluctuations and is the central result of this work.

The results obtained in Sec. II B 1 fall within the existing DPFT framework, which is formulated in state space. In contrast, the transition-space representation developed in Sec. II B 2 constitutes the main technical advance of this work. Since χ_i is the generator associated with the state ρ_i, its state-space nature prevents it from tracking the specific transition channel through which a state change occurs. By comparison, the transition-space representation in Eq. (10) assigns a generator χ_γ to each linearly independent transition. To illustrate this distinction, consider again the example of two protein conformations undergoing transitions via two channels: an equilibrium channel and an enzymatically driven one. Using $\chi_{r_{\rm eq}} = \chi_1$, $\chi_{p_{\rm eq}} = \chi_2$, $\chi_{r_{\rm enz}} = \chi_1$, and $\chi_{p_{\rm enz}} = \chi_2$, Eq. (10) reduces to a function of $\chi_1 - \chi_2$, i.e., a state-space representation that cannot distinguish which transition channel is taken. In contrast, Eq. (12), formulated in transition space, depends on $\chi_{\rm eq}$ and $\chi_{\rm enz}$, and therefore explicitly resolves the contribution of each transition channel. This also highlights the physical equivalence of the conjugate fields χ_i and χ_γ, which can be interpreted, respectively, as a stochastic chemical potential associated with the state and a stochastic affinity associated with the transition, discussed below.

3. Most likelihood path and the inferred EPR

The transition probability measure Eq. (13) is dominated by the saddle-point of Eq. (12). Therefore, the most-likelihood path gives the optimization problem for the Lagrangian with respect to the conjugate affinity χ_γ:

$$\mathcal{L}^*[\{J_\gamma, T_\gamma\}] = \sup_{\{\chi_\gamma\}}\mathcal{L}[\{J_\gamma, T_\gamma;\chi_\gamma\}]. \tag{15}$$

The extremization $\delta\mathcal{L}[\{J_\gamma, T_\gamma;\chi_\gamma\}]/\delta\chi_\gamma = 0$ gives the optimal 'effective' affinity for the stochastic current $\chi^*_\gamma = 2\tanh^{-1}(J_\gamma/T_\gamma)$ and the corresponding "effective" Lagrangian $\mathcal{L}^*[\{J_\gamma, T_\gamma\}]$, where, for any given stochastic realization of J_γ and T_γ, χ^*_γ and $\mathcal{L}^*$ are unique. Physically, it corresponds to the most likelihood transition affinity that generates the given instantaneous stochastic current and traffic; see Fig. 1(a). It will play a key role throughout this work.

The saddle-point approximation parameter is 1 for the extremization with respect to χ_γ. This physically corresponds

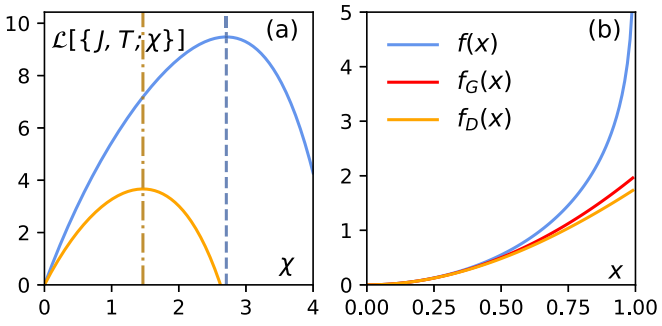


FIG. 1. (a) Lagrangian $\mathcal{L}[\{J, T; \chi\}]$ for fixed $J_\gamma = 3.5, T_\gamma = 4$ (cyan) and $J_\gamma = 2.5, T_\gamma = 4$ (orange). The corresponding most-likely transition affinity $\chi^* = 2\tanh^{-1}(J_\gamma/T_\gamma)$ is shown as a vertical dotted line. (b) Comparison between the exact large deviation rate functional $I = 2x\tanh^{-1}(x)$, the dynamical rate functional $I_D = 2x\sinh^{-1}(x)$, and the close-to-equilibrium quadratic (Gaussian) approximated rate functional $I_G = 2x^2$, where $x = J_\gamma/T_\gamma$ is the current precision.

to incorporating Poisssonian transition fluctuations for mesoscopic systems, with a quantitatively equal weight to all higher-order current cumulants. The extremization procedure in Eq. (15) is mathematically equivalent to the inverse of the Stratonovich-Hubbard transform [126,127], with a nonlinear coupling to the conjugate field. Due to which, it leads to a nonquadratic dependence of $\mathcal{L}^*$ on J_γ and T_γ, because of the quantification of non-Gaussian fluctuations, due to the *bottom-up construction* toward the transition fluctuations developed here. In contrast, the usual quadratic formulation corresponds to Gaussian fluctuations and yields a quadratic Onsager-Machlup functional [81,82]. This assumes the validity of the mean-field approximation of the transition currents and incorporates the gaussian fluctuations around it by implementing the Stratonovich-Hubbard transform that couples the conjugate field to the current and traffic linearly and quadratically, respectively; see Eq. (20): a *top-down approach* toward incorporating the transition fluctuations [83–86].

χ^*_γ depends nonlinearly on the precision of the current defined as the ratio J_γ/T_γ. Physically, this is a consequence of exactly incorporating the non-Gaussian transition fluctuations. Therefore, taking into account all higher-order cumulants renormalizes the effective affinity to give a nonlinear relation between the current precision and the effective affinity. Remarkably, the first two cumulants completely determine the exact transition statistics, a physical manifestation of the $\mathbb{Z}_2$ symmetry of the forward and backward transitions, a consequence of time-reversal symmetry imposed by the transition topology. Then, the "effective" Lagrangian $\mathcal{L}^*[\{J_\gamma, T_\gamma\}]$ reads

$$\dot{\Sigma} = \mathcal{L}^*[\{J_\gamma, T_\gamma\}] = \sum_{\{\gamma^{\rightleftharpoons}\}} 2J_\gamma \tanh^{-1}\left(\frac{J_\gamma}{T_\gamma}\right), \tag{16}$$

If the transition affinities have been known, then Eq. (16) is equal to the mean transition EPR. This follows trivially using analytical expressions $J_\gamma = 2D_\gamma \sinh(A_\gamma/2)$ and $T_\gamma = 2D_\gamma \cosh(A_\gamma/2)$ which implies $\chi^*_\gamma = A_\gamma$. However, Eq. (16) defines EPR for a given stochastic realization of J_γ and T_γ, without knowing the transition affinities, which is equivalently seen by the bilinear form of Eq. (16), $\mathcal{L}^*[\{J_\gamma, T_\gamma\}] = \sum_{\{\gamma^{\rightleftharpoons}\}} J_\gamma \chi^*_\gamma$ and gives a thermodynamic meaning to $\mathcal{L}^*[\{J_\gamma, T_\gamma\}]$: the "inferred" mean EPR in the absence of knowledge about the transition affinities.

$\chi^*_\gamma = A_\gamma$ does not necessarily hold. To highlight this point, consider that an observed value of current and traffic are $J_\gamma = 2.5$ and $T_\gamma = 4$, which results in an effective inferred affinity χ^*_γ, vertical dark orange dashed line in Fig. 1(a), the corresponding underlying effective Lagrangian that generates this most likelihood path is given by the solid orange curvy line in Fig. 1(a), this case corresponds to $\chi^*_\gamma = A_\gamma$. However, this realization of observed stochastic current and traffic could have been a lesser likelihood path of a different underlying model, for example, the intersection point of the cyan solid line (other model) and the vertical dark orange dashed line, this case corresponds to $\chi^*_\gamma \neq A_\gamma$, since A_γ is given by the dotted vertical dark cyan line in Fig. 1(a). This allows defining the inferred mean EPR for any given lesser likelihood stochastic realization of the current and traffic, without knowing the underlying affinities. The inferred mean EPR is equal to the deterministic mean EPR of the underlying true model, only if the sampled realization of the stochastic current and traffic is the most likelihood path, otherwise it is a lesser likelihood fluctuation of EPR.

4. Thermodynamic length and entropy production

Defining a time-integrated stochastic microscopic current $\tau\tilde{J}_\gamma = \int_0^\tau \hat{J}_\gamma$ and traffic $\tau\tilde{T}_\gamma = \int_0^\tau \hat{T}_\gamma$. Here, $\int_0^\tau$ is defined using the counting observable such that $\tau\tilde{J}_\gamma$ and $\tau\tilde{T}_\gamma$ count the total directional current and bidirectional transitions for $\gamma^{\rightleftharpoons}$, respectively. Physically, $\tau\tilde{J}_\gamma$ and $\tau\tilde{T}_\gamma$ quantify the thermodynamic length and dynamical activity of the transition $\gamma^{\rightleftharpoons}$ over the given observation time τ. Here, $\tilde{J}_\gamma$ and $\tilde{T}_\gamma$ are scaled time-integrated stochastic current and traffic, with the observation time τ being the scaling parameter that assumes the dissipative scaling of the transition currents. We integrate Eq. (16) from the initial time $t = 0$ to the final time τ and obtain the following nonquadratic relation between the thermodynamic length and the EP,

$$\tau\tilde{\Sigma} = \Sigma = \mathcal{S}^*_{\mathrm{DP}} = \int_0^\tau \mathcal{L}^* dt \geqslant \sum_{\{\gamma^{\rightleftharpoons}\}} 2\tau\tilde{J}_\gamma \tanh^{-1}\left(\frac{\tilde{J}_\gamma}{\tilde{T}_\gamma}\right), \tag{17}$$

where we have used Jensen's inequality to integrate Eq. (16), which turns the equality between $\dot{\Sigma}$ and J_γ, T_γ in Eq. (16) to the inequality (lower bound) between Σ and $\tilde{J}_\gamma, \tilde{T}_\gamma$ in Eq. (17), such that the equality is recovered in $\tau \to 0$.

Equations (16) and (17) relate the short-time and finite-time nonquadratic thermodynamic lengths to the thermodynamic quantities $\dot{\Sigma}$ and Σ. It reveals that the dynamic quantities ($\tilde{J}_\gamma/J_\gamma$ and $\tilde{T}_\gamma/T_\gamma$ here) of the fEQ systems are fundamentally constrained by the thermodynamic EP/EPR, and delineates a trade-off between current mean, current fluctuations, and dissipation. The exact formulation of the nonquadratic thermodynamic length gives the tightest bound compared to the known phenomenological quadratic and nonquadratic counterparts. In contrast, equilibrium or cEQ systems exhibit a quadratic relation of thermodynamic length and entropy production and are valid for the free energy or

excess EP, respectively [39,95–98], which represents only a part of the total EP, reducing its applicability to fEQ systems due to massive underestimation of EPR.

5. Min-Max principle for the EPR

The set of Eqs. (11)–(16) formulates a min-max variational problem for the action, in particular, $\Sigma = \inf_{\{J_\gamma,T_\gamma\}}\sup_{\{\chi_\gamma\}}\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma;\chi_\gamma\}]$. It is a minimum action principle valid for fEQ systems, which allows a unified description of stochastic discrete-state systems [89]. The min-max principle is a variational consequence of Eq. (12) being concave in χ_γ [Fig. 1(a)], and Eq. (16) being convex in J_γ/T_γ [blue line in Fig. 1(b)] combined with the saddle-point approximation that aims to minimize the action. The min-max formulation here is more subtle than the existing cEQ formulations. It requires a first maximization over the conjugate affinity and then a subsequent minimization over the transition currents. The maximization over the conjugate affinity is more important for systems exhibiting prominent stochastic effects. To elaborate on this point, we highlight two major fundamental regimes for the applicability of min-max principle that lead to different physical principles.

The first regime corresponds to the constraining of the effective transition affinity χ_γ, defined as, "force constraint" systems. In this regime, min-max formulation is effectively realized as the "minimization of inferred EPR," since $\inf_{\{J_\gamma,T_\gamma\}}\sup_{\{\chi_\gamma\}}\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma;\chi_\gamma\}] = \inf_{\{J_\gamma,T_\gamma\}}\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma\}]$. The "force constraint" system is physically realized in two important scenarios: when χ_γ is constant and/or small. First, when the noise effects (stochasticity) are not prominent, the effective transition affinity is a constant, for example, a deterministic limit of chemical reaction networks, and the study of nonequilibrium steady-state analysis using irreversible thermodynamics [46]. These systems have been paradigms for justifying the "minimum entropy production principle" [46]. Second, when χ_γ is small, it quantifies cEQ gaussian fluctuations using the quadratic Onsager-Machlup functional [81,82]. The quadratic Onsager-Machlup functional has been used to study fluctuations of fEQ systems, despite it being obtained relying on the cEQ Gaussian approximation.

The second regime corresponds to constraining the transition current J_γ, namely, "current constraint" systems. In this case, min-max formulation is effectively realized as the maximization of inferred EPR, namely, MaxEP, because $\inf_{\{J_\gamma,T_\gamma\}}\sup_{\{\chi_\gamma\}}\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma;\chi_\gamma\}] = \sup_{\{\chi_\gamma\}}\mathcal{S}_{\rm DP}[\{J_\gamma,T_\gamma;\chi_\gamma\}]$. Hence, physically, this implies that if the transition currents are fixed using an external thermodynamic reservoir, then the system maximizes the corresponding effective affinity, which effectively minimize the corresponding current fluctuations. Equivalently, a similar mechanism applies to systems prone to non-Gaussian stochastic effects. Here, maximization over the conjugate noise field becomes important, and hence the system effectively realizes "MaxEP" [69]. Physically, "MinEP" and "MaxEP" are valid approximations of the Min-Max principle.

However, more practical systems lie in an intermediate regime, where partially "current constraint" and "force constraint" on the system can coexist. Such sophisticated setups are expected to exhibit a more complex interplay between the "minimization" and "maximization" of entropy production, with the transition between them dictated by the control parameters of the specific model studied [69]. Nevertheless, the MinAP is expected to hold for any control parameter regime. In this respect, similar to equilibrium systems, MinAP can incorporate experimental observable constraints via a Lagrange multiplier.

Defining conjugate affinities $\chi_{J_\gamma} = \partial_{J_\gamma}\mathcal{L}^*$ and $\chi_{T_\gamma} = \partial_{T_\gamma}\mathcal{L}^*$ for currents and traffics, $\mathcal{L}^*[\{J_\gamma,T_\gamma\}]$ satisfies the zero-cost flow constraint $\sup_{\{J_\gamma,T_\gamma\}}(J_\gamma\chi_{J_\gamma} + T_\gamma\chi_{T_\gamma} - \mathcal{L}^*) = 0$, establishing the connection between the mean inferred EPR and frenetic activity [123]. Mathematically, this dependence was rather clear from Eq. (12), where the conjugate field χ_γ couples to both J_γ and T_γ, but we still explicitly highlight it.

6. Mapping to information geometry

From an information geometric viewpoint, the right-side of Eq. (15) is the variational representation of the KL divergence defined between the forward and backwards path probability measures, which is defined as the EPR in Refs. [128–130] and other relevant results on variational formulation [115–117]. This is prominently visible using the representation of Lagrangian in unidirectional transition currents Eq. (11), which is the Donsker-Varadhan representation of KL divergence used in Information geometry, and is equal to $\sup_{\{\chi_\gamma\}}\mathcal{L}[\{j_\gamma;\chi_\gamma\}]$ [128–130]. Hence, our formulation proves the equivalence of the EPR obtained using the information geometric and statistical mechanical formulation. Importantly, in contrast to Refs. [128–130], we have defined EPR using the microscopic observables of the model itself and does not require identification of the backward process [131], which is sensitive to correct/incorrect identification of the backward process, due to the resolution of the given trajectory/path. This exact mapping allows us to study ST using information geometry as a mathematical tool with a physical interpretation given by the 'Minimum action principle'.

7. Beyond the local detailed balance condition

χ_γ is a parameter that is associated with the statistical properties of the system. In comparison, A_γ is associated with the thermodynamically consistent model definition imposed through the LDB condition required in stochastic thermodynamics. However, the derivation of MinAP does not utilize LDB condition. Therefore, the "stochastic" path integral formulation and "MinAP" derived in Sec. II B and its physical formulation as a fEQ analog of the Boltzmann distribution holds irrespective of the LDB condition for a broad class of fEQ discrete-state systems beyond validity regimes of Stochastic Thermodynamics. When the LDB condition is satisfied, the thermodynamic consistency of transition currents endows the effective Lagrangian with a clear physical interpretation as the mean EPR, as discussed in Sec. II B 3. Even in the absence of LDB, the information-geometric mathematical interpretation of the Lagrangian remains valid. Importantly, even if the LDB condition is violated, or if the "observer" is practically agnostic about its validity (even when it holds microscopically), the "physical" inferred mean EPR, which is constructed from the time-symmetric and time-antisymmetric components of physical observables (such as currents and

traffics), remains well-defined. Owing to this distinction, the information-geometric interpretation presented here is directly grounded in physically observable quantities.

χ_γ^* being time-homogeneous (independent of observation time) is equivalent to the system being memory-less (if valid $\forall\, \gamma^{\rightleftharpoons} \in \{\gamma^{\rightleftharpoons}\}$); therefore, in this regime, due to the time-scale separation, the initial-value condition for the states and the correlations between the system and the thermodynamic reservoir become irrelevant, and the "*bottom-up*" constructed validity of the LDB condition emerges naturally [50,51]. Since the derivation of MinAP does not rely on the LDB condition, it allows the study of transient effects that violate LDB and goes beyond the assumption of timescale separation in ST. In ST, the weak coupling between the system and the reservoir allows for the formulation of the LDB condition, which in itself is interpreted as a form of cEQ assumption (in the sense of weak coupling between the system and environment). This highlights that the MinAP extends the dual thermodynamic structure between affinities and the currents and traffics to fEQ systems, beyond the weak coupling regime in ST.

C. Large deviation principle

The Boltzmann distribution establishes an equivalence between the statistical properties of the physical quantities and the energetics of the systems [1]. Using the variational formulation and the large deviation theory, our aim is to investigate a similar principle for fEQ systems [26]. The large deviation theory studies fluctuations of dynamical observables in nonequilibrium systems [26]. The probability distribution $\mathcal{P}(\mathcal{O})$ for nonequilibrium physical observables ($\mathcal{O}$) is said to satisfy the large deviation principle (LDP) with a large deviation parameter Ω and a rate functional $I(\mathcal{O})$, if it satisfies $\mathcal{P}(\mathcal{O}) \asymp e^{-\Omega I(\mathcal{O})}$. The scaling parameter Ω dictates the convergence of the probability distribution to the minimum of the rate functional (the most likelihood value of the observable) and also quantifies the fluctuations around the most likelihood value. Here, the observable is an intensive variable and satisfies the scaling for the mean $\langle\mathcal{O}\rangle \propto O(1)$ and variance $\mathrm{Var}(\mathcal{O}) \propto 1/\Omega$. The Boltzmann distribution for equilibrium systems ($\mathcal{P}^{\mathrm{eq}} \asymp e^{-\beta E}$) is an example of LDP. Where the inverse temperature β is the LDP scaling parameter, and the equilibrium energy functional E is the corresponding rate functional that quantifies the energy of the system. Fluctuations vanish in the zero-temperature limit, and the system's energy converges to the minimum of E.

However, the LDP formulation is prone to two assumptions/approximations: (1) the existence and particular analytical form of such a rate functional, for instance, a Gaussian approximation for the observable statistics; (2) choosing a coarse-grained macroscopic observable that discards information about other relevant microscopic observables. Here, we aim to address both issues and derive an exact rate functional using a systematic "*bottom-up approach*" for all microscopically relevant physical quantities. To highlight the importance of the exact LDP, we will compare it to two phenomenological cases that correspond to the Gaussian approximation of noise and dynamical rate functionals, which correspond to the thermodynamic uncertainty relation and the nonequilibrium fluctuation-response relation, respectively.

1. Exact result

We combine Eqs. (17) and (13) and obtain the exact LDP for the scaled-time-integrated current and traffic, $\tilde{J}_\gamma$ and $\tilde{T}_\gamma$, it reads

$$\mathcal{P}[\{\tilde{J}_\gamma, \tilde{T}_\gamma\}] \asymp e^{-\tau I[\{\tilde{J}_\gamma, \tilde{T}_\gamma\}]}, \tag{18}$$

where $I[\tilde{J}_\gamma, \tilde{T}_\gamma] = 2\tilde{J}_\gamma \tanh^{-1}(\tilde{J}_\gamma/\tilde{T}_\gamma)$ is the exact dynamical rate functional for $\tilde{J}_\gamma$ and $\tilde{T}_\gamma$, with the observation time τ is the corresponding LDP scaling parameter for the dynamical canonical ensemble [26,123]. Due to the second saddle-point approximation (with respect to τ), the transition probability measure for $\tilde{J}_\gamma$ and $\tilde{T}_\gamma$ peaks at the minimum of the rate functional [26,123].

We define the current precision for the finite- and short-time processes $\tilde{x}_\gamma = \tilde{J}_\gamma/\tilde{T}_\gamma$ and $x_\gamma = J_\gamma/T_\gamma$, respectively. $I[\tilde{J}_\gamma, \tilde{T}_\gamma]$ is rewritten using the current precision and $f(x) = 2x\tanh^{-1}(x)$ such that $I[\tilde{J}_\gamma, \tilde{T}_\gamma] = \tilde{T}_\gamma f(\tilde{x}_\gamma)$. This scaling implies that $\tilde{T}_\gamma$ defines the timescale corresponding to $\gamma^{\rightleftharpoons}$, and $\tilde{x}_\gamma$ is the relevant observable with rate funtional $f(\tilde{x})$. We can absorb $\tilde{T}_\gamma$ by redefining time $t' = t\tilde{T}_\gamma$. However, different transitions have different timescales, so $\tilde{T}_\gamma$ is also a relevant parameter. Importantly, because the scaling parameter τ characterizes the convergence of the transition probability measure to the minimum of the rate functional (most likelihood value), with a relaxation timescale $\tau_\gamma^{\mathrm{rel}} = 1/\tilde{T}_\gamma$. Hence, for a given fixed observation time τ, higher values of $\tilde{T}_\gamma$ accelerate the convergence of $\tilde{x}_\gamma$ to it's most likelihood value. Physically, a faster current dynamics (degree of freedom) is approximated with a constant value, and the stochasticity of slower dynamics could be studied. This physical mechanism usually comes under different names, for example, time-scale separation, adiabatic approximation of fast dynamics.

$f(x)$ also relates the precision of time-integrated currents to the EP Σ_γ associated with transition $\gamma^{\rightleftharpoons}$. Due to Eq. (17), we obtain $\Sigma_\gamma = \tau\tilde{T}_\gamma f(\tilde{x}_\gamma)$. Using the inverse function $f^{-1}(x)$, we obtain the nonquadratic upper bound on the current precision for the given EP and traffic, which reads

$$\tilde{J}_\gamma \leqslant \tilde{T}_\gamma f^{-1}\left(\frac{\Sigma_\gamma}{\tau\tilde{T}_\gamma}\right). \tag{19}$$

Importantly, the scaling form of $\Sigma_\gamma/\tau\tilde{T}_\gamma$ in Eq. (19) signifies that the EP for a transition Σ_γ should be measured in its natural timescale $\tau\tilde{T}_\gamma$. Equation (19) is the fundamental universal scaling relation between the precision of the time-integrated current and EP, that bounds the precision of the transition current with the scaled EP (in the natural timescale of transition). Equations (16)–(18) formulate the fundamental foundation of this work, namely, an exact canonical ensemble analog for the statistical properties of dynamical physical quantities (current and traffic) and their connection to thermodynamic dissipation.

2. Gaussian approximation and thermodynamic uncertainty relation

The Gaussian approximation of the transition fluctuations is equivalent to the second-order Taylor-series expansion of

Eq. (12) in χ_γ,

$$\mathcal{L}_G[\{J_\gamma, T_\gamma; \chi_\gamma\}] = \sum_{\{\gamma \rightleftharpoons\}} 2J_\gamma \chi_\gamma - \frac{1}{2} T_\gamma \chi_\gamma^2, \tag{20}$$

which ignores all higher-order current cumulants. Having solved the variational problem for Eq. (20), $\delta \mathcal{L}_G / \delta \chi_\gamma = 0$, the effective affinity for the transition $\gamma^{\rightleftharpoons}$ is $\chi_\gamma^* = 2J_\gamma / T_\gamma$ corresponds to the most-likelihood path. The effective Gaussian Lagrangian is

$$\dot{\Sigma}_G = \mathcal{L}_G^*[\{J_\gamma, T_\gamma\}] = \sum_{\{\gamma \rightleftharpoons\}} \frac{2J_\gamma^2}{T_\gamma}. \tag{21}$$

Equation (21) is the quadratic dissipation function originally defined for Gaussian fluctuations around the equilibrium state [81,82] but is generalized here for Gaussian fluctuations around any NESS. This also clarifies the nomenclature convention of the inverse Stratonovich-Hubbard transform as the transform from $\mathcal{L}$ to $\mathcal{L}^*$ [126,127]. The inferred EPR $\dot{\Sigma}_G$ using the Gaussian approximation is also known as pseudo-EPR and obtains a lower bound on $\dot{\Sigma}$ [36]. Using Eq. (21), the transition probability measure satisfies LDP:

$$\mathcal{P}[\{\tilde{J}_\gamma, \tilde{T}_\gamma\}] \asymp e^{-\tau I_G[\{\tilde{J}_\gamma, \tilde{T}_\gamma\}]}, \tag{22}$$

with scaling relation, $I_G[\tilde{J}_\gamma, \tilde{T}_\gamma] = \tilde{T}_\gamma f_G(\tilde{x}_\gamma)$, where $f_G(\tilde{x}) = 2\tilde{x}^2$. Compared to the exact results, χ_γ^* is a linear function of current precision x_γ. This led to a quadratic dissipation function in Eqs. (21) and (22) due to the Gaussian approximation of the transition noise.

3. Dynamical rate functional and nonequilibrium fluctuation-response relation

The nonequilibrium fluctuation response relation (NFRR) between the instantaneous current and the instantaneous traffic is [5,132–135]

$$\frac{\partial J_\gamma}{\partial \chi_\gamma} = \frac{T_\gamma}{2}. \tag{23}$$

We have used the conjugate field χ_γ instead of A_γ, allowing NFRR to be parametrically evaluated around any effective affinity χ_γ^*. Choosing $\chi_\gamma = A_\gamma$ recovers NFRR for the given physical model. Using the nonlinear dependence of $T_\gamma = \sqrt{J_\gamma^2 + 4D_\gamma^2}$, the nonlinear force-current relation is

$$J_\gamma = 2D_\gamma \sinh\left(\frac{\chi_\gamma}{2}\right). \tag{24}$$

By plugging $\chi_\gamma = A_\gamma$ into Eq. (24), the nonlinear relation between the mean transition current and the affinity is recovered.

We define the dual dissipation functions ψ_χ and ψ_J, to further unveil the Legendre dual structure between the current and the force. To this end, we use the definition of the Legendre transform, $J_\gamma = \partial_{\chi_\gamma} \psi_\chi(\{\chi_\gamma\})$ and $\chi_\gamma = \partial_{J_\gamma} \psi_J(\{J_\gamma\})$, combined with Eq. (24), we obtain

$$\begin{aligned} \psi_\chi(\{\chi_\gamma\}) &= \sum_{\{\gamma \rightleftharpoons\}} 4D_\gamma \left(\cosh\left(\frac{\chi_\gamma}{2}\right) - 1\right), \\ \psi_J(\{J_\gamma\}) &= \sum_{\{\gamma \rightleftharpoons\}} 2J_\gamma \sinh^{-1}\left(\frac{J_\gamma}{2D_\gamma}\right) \\ &\quad - 2\left(\sqrt{J_\gamma^2 + 4D_\gamma^2} - 2D_\gamma\right), \end{aligned} \tag{25}$$

where we have imposed the constraint $\psi_F(0) = 0$ and $\psi_J(0) = 0$, physically corresponding to the vanishing EPR for vanishing forces and currents. The dual dissipation functions characterize the EPR, the variational Lagrangian corresponding to them reads [123]

$$\mathcal{L}_D[\{J_\gamma; \chi_\gamma\}] = \sum_{\{\gamma \rightleftharpoons\}} [\psi_\chi + \psi_J]. \tag{26}$$

Using Eq. (24) and simplifying Eqs. (26) and (25), the effective Lagrangian $\mathcal{L}_D^* = \sup_{\{\chi_\gamma\}} \mathcal{L}_D[\{J_\gamma; \chi_\gamma\}]$ is obtained. Where the second term of $\psi_J(\{J_\gamma^*\})$ cancels with $\psi_\chi(\{\chi_\gamma^*\})$ in Eq. (25) and leads to

$$\dot{\Sigma}_D = \mathcal{L}_D^*[\{J_\gamma, D_\gamma\}] = \sum_{\{\gamma \rightleftharpoons\}} 2J_\gamma \sinh^{-1}\left(\frac{J_\gamma}{2D_\gamma}\right). \tag{27}$$

Equation (27) is the previously computed dynamical large deviation rate functional for the transition currents [136,137] and recently utilized to formulate and study the "Hessian" structure for discrete-state processes in Refs. [123,136–152]. If LDP were to be formulated using Eq. (27), then the transition probability measure reads

$$\mathcal{P}[\{\tilde{J}_\gamma, \tilde{D}_\gamma\}] \asymp e^{-\tau I_D[\{\tilde{J}_\gamma, \tilde{D}_\gamma\}]}, \tag{28}$$

with time-integrated mobility ($\tau \tilde{D}_\gamma = \int_0^\tau \hat{D}_\gamma$) and rate functional $I_D[\tilde{J}_\gamma, \tilde{D}_\gamma] = 2\tilde{J}_\gamma \sinh^{-1}(\tilde{J}_\gamma / 2\tilde{D}_\gamma)$ satisfying the scaling relation $I_D[\tilde{J}_\gamma, \tilde{D}_\gamma] = 2\tilde{D}_\gamma f_D(\tilde{x}_\gamma)$ with $f_D(\tilde{x}) = 2\tilde{x} \sinh^{-1}(\tilde{x})$.

4. Physical implications of the exact LDP

The analytical form of $f(x)$ plays a key role in the relationship between the current precision and the EPR. We plot $f(x)$, $f_G(x)$ and $f_D(x)$ in Fig. 1(b), which exhibit the hierarchy of inequality,

$$f(x) \geqslant f_G(x) \geqslant f_D(x). \tag{29}$$

This hierarchy of inequality, combined with Eq. (17) implies that the exact rate functional computes the best bound on Σ_γ for a given value of current precision. Inverting Eq. (29) and using Eq. (19), we obtain the hierarchy of bounds on the current precision for the given EP corresponding to the transition,

$$\tilde{J}_\gamma \leqslant \tilde{T}_\gamma f^{-1}\left(\frac{\Sigma_\gamma}{\tau \tilde{T}_\gamma}\right) \leqslant \tilde{T}_\gamma \sqrt{\frac{\Sigma_\gamma}{2\tau \tilde{T}_\gamma}} \leqslant 2\tilde{D}_\gamma f_D^{-1}\left(\frac{\Sigma_\gamma}{2\tau \tilde{D}_\gamma}\right). \tag{30}$$

Equation (30) holds for an inverse problem, when EP corresponding to a transition is known, and the objective is to obtain the tightest bound on the corresponding current precision. The inequality between the rate functional hierarchy becomes prominent for fEQ systems, that exhibit more precise currents.

The Gaussian approximation of the rate functional [74,75,79,80] has been extensively studied to obtain bounds on EP using quadratic TKUR [29–35]. Here, the exact rate functional addresses the issue of massive underestimation of EPR associated with quadratic TKUR. This mismatch is particularly pronounced for fEQ systems or those exhibiting non-Gaussian fluctuations. Similarly, I_D has been utilized to study the "Hessian" dual structure between force and current [123,136–152]. I avoids an underestimation of $\dot{\Sigma}$ and obtains the tightest bound on $\dot{\Sigma}$. The proof follows using, $I[\tilde{J}_\gamma, \tilde{T}_\gamma] > I_D[\tilde{J}_\gamma, \tilde{D}_\gamma]$, since $\tilde{T}_\gamma > 2\tilde{D}_\gamma$ combined with $f(\tilde{x}) > f_D(\tilde{x})$, which implies $\dot{\Sigma} \geqslant \dot{\Sigma}_D$. Physically, the tightness of the bounds obtained using $I(\tilde{J}_\gamma, \tilde{T}_\gamma)$ and $f(\tilde{x})$ results from the incorporation of exact nonequilibrium current fluctuations characterized by T_γ instead of $2D_\gamma$. Where $2D_\gamma$ computes the lower bound on the equilibrium current fluctuations, as $T_\gamma^{\mathrm{eq}} = j_\gamma^{\mathrm{eq}} + j_{-\gamma}^{\mathrm{eq}} = (j_\gamma + j_{-\gamma})|_{F_\gamma=0}$ due to the inequality $T_\gamma^{\mathrm{eq}} \geqslant 2D_\gamma = 2\sqrt{j_\gamma j_{-\gamma}}$. Using T_γ instead of $2D_\gamma$, the renormalization of the variance of the nonequilibrium currents is taken into account. In contrast, Eq. (25) assumes a constant static equilibrium diffusivity D_γ, which is the transition mobility. Using $2D_\gamma$ instead of T_γ reveals a violation of NFRR for fEQ systems, attributed to an underestimation of variance (fluctuations) due to $2D_\gamma$. However, as formulated above, the NFRR between the transition current response and traffic holds for fEQ systems. Due to the equivalence between the exact variational formulation and the information geometric formulation, we have a mathematical proof an open problem quoted in [128,129], information geometric methods obtains tighter bounds on the EPR in Refs. [128,129], in comparison to the Hessian structure in Refs. [123,136–152]. Importantly, our analysis reveals that the shortcomings of the quadratic TKUR and Hessian structure are remedied by the exact formulation.

III. COARSE-GRAINED OBSERVABLE CURRENTS

Microscopic transition currents and traffics assume complete information of the system. However, experimental constraints or ignorance of the microscopic transitions restrict our access to the complete information. Experimentally, coarse-grained observable (macroscopic) currents are easily accessible. This creates a necessity to examine the possibility of the variational formulation for observable currents. In this section, we extend the applicability of the variational formulation to coarse-grained observable currents. The main results are summarized in Sec. III A. Their derivation is detailed in Sec. III B for completeness. The reader is encouraged to read through the derivation Sec. III B for a qualitative understanding of the structural impact of the coarse-graining procedure, although the technical details may be skipped.

A. Setup

We define a set of observable (macroscopic) $\{o\}$ time-antisymmetric currents $\{J_o\}$ and the corresponding time-symmetric traffics $\{T_o\}$, with the many-to-one coarse-graining mapping $\mathcal{CG} : \{\gamma^{\rightleftharpoons}\} \to \{o\}$, thus $J_{o'} = \sum_{\{\gamma^{\rightleftharpoons}\}} \mathbb{O}_{o'\gamma} J_\gamma$ and $T_{o'} = \sum_{\{\gamma^{\rightleftharpoons}\}} \mathbb{O}_{o'\gamma} T_\gamma$, $\forall o' \in \{o\}$, where the matrix elements $\mathbb{O}_{o'\gamma} \in \{0, 1\}$ of $\mathbb{O}$ defines the coarse-graining mapping $\mathcal{CG} : \{\gamma^{\rightleftharpoons}\} \to \{o\}$, which is represented mathematically as $\mathbb{O}_{o'\gamma'} \neq 0 \Rightarrow \mathbb{O}_{o'\gamma'} = 0, \forall o' \in \{o\} - o'', \forall \gamma' \in \{\gamma^{\rightleftharpoons}\}$. The support of observable o' ($\mathrm{supp}(o') = \{\gamma' | \mathbb{O}_{o'\gamma'} \neq 0\}$) defines the set of microscopic transitions, and its dimension quantifies the number of microscopic transitions $|N_{o'}| = |\mathrm{supp}(o')|$ that contribute to o'. If observable currents account for all microscopic transitions, then $|\{\gamma^{\rightleftharpoons}\}| = \sum_{o' \in \{o\}} |N_{o'}|$ holds. Introducing the vector notation $\vec{J}_o = \mathbb{O}\vec{J}_\gamma$ and $\vec{T}_o = \mathbb{O}\vec{T}_\gamma$. Physically, the binary-ness of $\mathbb{O}_{o'\gamma} \in \{0, 1\}$ implies that a microscopic transition is either observable or not observable and imposes a scaling constraint on the observable EPR, ensuring that each microscopic transition is counted once in the observable currents. The many-to-one mapping constraint also ensures the linear independence of observable currents.

1. Main results

Using the contraction principle [26], we derive the observable Lagrangian/EPR $\mathcal{L}^*_{\{o\}}[\{J_o, T_o\}]/\dot{\Sigma}_{\{o\}}$ corresponding to observable currents and traffics in Sec. III B 1, it reads

$$\dot{\Sigma}_{\{o\}} = \mathcal{L}^*_{\{o\}}[\{J_o, T_o\}] = \sum_{\{o\}} 2J_o \tanh^{-1}\left(\frac{J_o}{T_o}\right). \tag{31}$$

Using the bilinear form, $\dot{\Sigma}_{\{o\}} = \sum_{\{o\}} J_o \chi^*_o$, we obtain the inferred affinity $\chi^*_o = 2\tanh^{-1}(J_o/T_o)$ using J_o and T_o. $\mathcal{L}^*[\{J_\gamma, T_\gamma\}] \geqslant \mathcal{L}^*_{\{o\}}[\{J_o, T_o\}]$ holds due to the generalized log-normal inequality combined with the definitions of J_o and T_o [153]. Physically, it corresponds to the observable currents and traffics being able to capture a part of the microscopic EPR. Hence, in addition to using the exact rate functional, selecting all linearly independent microscopic observable currents is the second important criterion to obtain exact bounds on $\dot{\Sigma}$ using observable currents and traffics. Choosing $\{o\} = \{\gamma^{\rightleftharpoons}\}$ saturates the bound between $\dot{\Sigma}$ and $\dot{\Sigma}_{\{o\}}$.

Defining the scaled time-integrated stochastic observable current and the corresponding traffic, $\tilde{J}_o = \frac{1}{\tau}\int_0^\tau \hat{J}_o$ and $\tilde{T}_o = \frac{1}{\tau}\int_0^\tau \hat{T}_o$, respectively. Integrating Eq. (31) from the initial $t = 0$ to the final time $t = \tau$, the relation between the inferred EP (Σ_o) and $\tilde{J}_o, \tilde{T}_o$ is

$$\tau\tilde{\Sigma}_{\{o\}} = \Sigma_{\{o\}} = \mathcal{S}^*_{\{o\}} = \int_0^\tau \mathcal{L}^*_{\{o\}} dt \geqslant \sum_{\{o\}} 2\tau\tilde{J}_o \tanh^{-1}\left(\frac{\tilde{J}_o}{\tilde{T}_o}\right). \tag{32}$$

Equations (31) and (32) formulate the short- and finite-time nonquadratic thermodynamic lengths for the observable currents and traffics, analogous to Eqs. (16) and (17), respectively. They relate the dynamical quantities: $\{J_o, T_o\}/\{\tilde{J}_o, \tilde{T}_o\}$ for short-time/finite-time to $(\dot{\Sigma}_{\{o\}})/(\Sigma_{\{o\}})$.

Using Eq. (32), the transition probability measure for the observable current and traffic reads

$$\mathcal{P}[\{\tilde{J}_o, \tilde{T}_o\}] \asymp e^{-\tau I[\{\tilde{J}_o, \tilde{T}_o\}]}, \tag{33}$$

Equation (33) formulates the canonical ensemble using $\{\tilde{J}_o, \tilde{T}_o\}$, analogously to Eq. (18). Defining the precision of time-integrated observable current $\tilde{x}_o = \tilde{J}_o/\tilde{T}_o$), the rate functional of the observable ($I[\tilde{J}_o, \tilde{T}_o]$) satisfies the scaling relation $I[\tilde{J}_o, \tilde{T}_o] = \tilde{T}_o f(\tilde{x}_o)$, implying that, for a fixed observation time τ, the convergence of $\mathcal{P}[\{\tilde{J}_o, \tilde{T}_o\}]$ to the minimum of ($I[\tilde{J}_o, \tilde{T}_o]$) is accelerated due to larger values of $\tilde{T}_o$ compared to their microscopic counterparts. Thus, the relaxation timescale of $\tilde{T}_o$ is $\tau_o^{\mathrm{rel}} = 1/\tilde{T}_o$, and is smaller than the microscopic counterparts.

The derivation of variational formulations for observable currents relied on two key constraints. *Constraint 1*: a binary notion of a microscopic current being observable or not observable, $\mathbb{O}_{o\gamma} \in \{0, 1\}$. *Constraint 2*: the many-to-one mapping ($\mathcal{CG} : \{\gamma^{\rightleftharpoons}\} \to \{o\}$) is used to obtain the coarse-grained currents. The validity of Eqs. (31)–(33) holds even if these constraints are relaxed; see Sec. III B 2.

2. Condition for the saturation of the bound

The condition for the saturation of equality between $\dot{\Sigma}$ and $\dot{\Sigma}_{\{o\}}$ is obtained using the log-normal inequality [153]. In particular, $\mathcal{L}^*[\{J_\gamma, T_\gamma\}] = \mathcal{L}^*_{\{o\}}[\{J_o, T_o\}]$, if $J_\gamma/T_\gamma = J_o/T_o$. Subsequently, this condition implies $\chi^*_\gamma = \chi^*_o, \forall\gamma \in \mathrm{supp}(o)$. This can be easily verified using the bilinear form of the EPR, $\dot{\Sigma} = \sum_{\{\gamma^{\rightleftharpoons}\}} J_\gamma \chi^*_\gamma = \sum_o J_o \chi^*_o = \dot{\Sigma}_o$, if $\chi^*_\gamma = \chi^*_o, \forall\gamma \in \mathrm{supp}(o)$. The bound $\dot{\Sigma} \geqslant \dot{\Sigma}_o$ is saturated if microscopic transition currents with equal effective affinities are counted together as a single observable current. Therefore, this bound is saturated for unicyclic graphs in steady state, where a single effective nonconservative affinity quantifies all nonconservative currents. However, for multicyclic systems, the affinities associated with all linearly independent cyclic currents need to be known, unless all of them are equal. We briefly mention the relevance of this result in the context of Ref. [154]. A "Third Kirchhoff's Law" for Stochastic Transport [Eq. (13) in Ref. [154]], which is a trivial subcase of the condition $\chi^*_\gamma = \chi^*_o, \forall\gamma \in \mathrm{supp}(o)$ obtained here for any generic multicyclic graph.

B. Derivation

1. Contraction principle approach

In large deviation theory, the contraction principle deals with deriving the rate functional for an observable with knowledge of another rate functional [26]. Here, we aim to apply the contraction principle to obtain the LDP for observable currents, given that the exact LDP for microscopic currents has been derived. The contraction of Eq. (16) under the constraint of the observable currents and traffics is formulated as the following constrained Lagrangian optimization problem:

$$\mathcal{L}^*[\{J_o, T_o\}] = \inf_{\{J_\gamma\},\{T_\gamma\},\{\lambda_{J_o}\},\{\lambda_{T_o}\}} [\mathcal{L}^*[\{J_\gamma, T_\gamma\}] + \vec{\lambda}_{J_o} \cdot (\vec{J}_o - \mathbb{O}\vec{J}_\gamma) + \vec{\lambda}_{T_o} \cdot (\vec{T}_o - \mathbb{O}\vec{J}_\gamma)], \tag{34}$$

where $\{\lambda_{J_o}\}$ and $\{\lambda_{T_o}\}$ are the Lagrange multipliers corresponding to the constraints imposed by $\{J_o\}$ and $\{T_o\}$, respectively. The extremization of Eq. (34) with respect to λ_{J_o} and λ_{T_o} leads to the trivial constraint equations for $\{J_o\}$ and $\{T_o\}$. The optimization problem in Eq. (34) for the set of observables $\{o\}$ is decoupled into independent optimization problems for J_o and T_o, due to the many-to-one coarse-graining mapping.

Solving the decoupled optimization problem requires computing the Euler-Lagrange equations, $\forall\gamma \in \{\gamma^{\rightleftharpoons}\}$, given by

$$\frac{\delta\mathcal{L}^*[\{J_\gamma, T_\gamma\}]}{\delta J_\gamma} = 2\tanh^{-1}\left(\frac{J_\gamma}{T_\gamma}\right) + \frac{2J_\gamma T_\gamma}{T_\gamma^2 - J_\gamma^2} - \lambda_{J_o}\mathbb{O}_{o\gamma}, \tag{35a}$$

$$\frac{\delta\mathcal{L}^*[\{J_\gamma, T_\gamma\}]}{\delta T_\gamma} = -\frac{2J_\gamma^2}{T_\gamma^2 - J_\gamma^2} - \lambda_{T_o}\mathbb{O}_{o\gamma}, \tag{35b}$$

where Eq. (35) holds $\forall\gamma \in \{\gamma^{\rightleftharpoons}\}$ and maps γ to a unique $o \in \{o\}$. Solving the optimization problem trivially implies $\delta\mathcal{L}^*[\{J_\gamma, T_\gamma\}]/\delta J_\gamma = 0$ and $\delta\mathcal{L}^*[\{J_\gamma, T_\gamma\}]/\delta T_\gamma = 0$,

$$2\tanh^{-1}\left(\frac{J_\gamma}{T_\gamma}\right) + \frac{2J_\gamma T_\gamma}{T_\gamma^2 - J_\gamma^2} = \lambda_{J_o}\mathbb{O}_{o\gamma}, \tag{36a}$$

$$-\frac{2J_\gamma^2}{T_\gamma^2 - J_\gamma^2} = \lambda_{T_o}\mathbb{O}_{o\gamma}. \tag{36b}$$

Computing $J_\gamma\times$ Eq. (36a) + $T_\gamma\times$ Eq. (36b) leads to $2J_\gamma \tanh^{-1}(J_\gamma/T_\gamma) = \mathbb{O}_{o\gamma}(\lambda_{J_o}J_\gamma + \lambda_{T_o}T_\gamma)$. This simplifies $\inf_{\{J_\gamma,T_\gamma\}} \mathcal{L}^*[\{J_\gamma, T_\gamma\}] = \sum_{\{\gamma^{\rightleftharpoons}\}} O_\gamma(J_\gamma\lambda_{J_o} + T_\gamma\lambda_{T_o})$ to $\mathcal{L}^*[\{J_o, T_o\}] = \sum_{\{o\}}(J_o\lambda_{J_o} + T_o\lambda_{T_o}) = \vec{\lambda}_{J_o}^T\vec{J}_o + \vec{\lambda}_{T_o}^T\vec{T}_o$ in its bilinear form.

For $\gamma, \gamma' \in \mathrm{supp}(o)$, the right-hand side of Eq. (36a) or Eq. (36b) is equal to λ_{J_o}. Hence, it imposes equality on the left-hand side of Eqs. (36a) and (36b) for $\gamma, \gamma' \in \mathrm{supp}(o)$. However, the left-hand sides are monotonic functions of $x_\gamma = J_\gamma/T_\gamma$, defined as $b_1(x) = 2\tanh^{-1}(x) + 2x/(1 - x^2)$ and $b_2(x) = 2x^2/(1 - x^2)$. This implies $\forall\gamma, \gamma' \in \mathrm{supp}(o)$, the solution of Eq. (34) satisfies $x_\gamma = x_{\gamma'} = x_o = J_o/T_o$. Plugging it into Eq. (34) leads to

$$\mathcal{L}^*[J_o, T_o] = 2J_o\tanh^{-1}\left(\frac{J_o}{T_o}\right), \ \forall o \in \{o\}, \tag{37}$$

hence, $\mathcal{L}^*[\{J_o, T_o\}] = \sum_{\{o\}} 2J_o\tanh^{-1}(J_o/T_o)$. Therefore, $\chi^*_o = 2\tanh^{-1}(J_o/T_o)$ corresponds to the effective transition affinity of the observable current, which can be equivalently inferred using J_o and T_o. Comparing χ^*_o to the bilinear form derived previously $\mathcal{L}^*[\{J_o, T_o\}] = \vec{\lambda}_{J_o}^T\vec{J}_o + \vec{\lambda}_{T_o}^T\vec{T}_o$, it is equal to an effective Lagrange multiplier for J_o with $\chi^*_o = \lambda_{J_o} + \lambda_{T_o}T_o/J_o$.

2. Linear algebra approach

For partial (scaled) observation of transition currents, such that $\mathbb{O}_{o\gamma} \in \mathbb{R}^+$ is any real number and not necessarily $\mathbb{O}_{o\gamma} \in \{0, 1\}$, we note that to respect the scale invariance of the EPR $\dot{\Sigma}$, $\mathcal{L}[\{J_\gamma, T_\gamma\}]$ has to be invariant under the scaling transformation of currents and traffics $J_\gamma \to \bar{J}_\gamma = cJ_\gamma$ and

$T_\gamma \to \bar{T}_\gamma = cT_\gamma$ [155], implying scaling transformation $x_\gamma \to \bar{x}_\gamma = x_\gamma$. Hence, our naive calculation implies that the effective affinity is also invariant under the scaling transformation, $\chi^*_\gamma \to \bar{\chi}^*_\gamma = \chi^*_\gamma$. Using the bilinear form of the EPR, $\dot{\Sigma} = \sum_{\gamma=} J_\gamma \chi^*_\gamma$, we reach the scaling transformation of the EPR, $\dot{\Sigma} \to \dot{\bar{\Sigma}} = c\dot{\Sigma}$. However, this contradicts the invariance of the EPR under the scaling transformation of currents.

To address this problem associated with scaling, one notices that a simultaneous scaling of time, $t \to \bar{t} = t/k$, restores the scaling invariance of the EPR, also seen equivalently by the scaling of traffic $T_\gamma \to \bar{T}_\gamma = kT_\gamma$ that defines the inverse timescale for $\gamma^=$. This key physical insight implies that Eq. (19) is invariant under scaling transformation and holds for scaled microscopic currents, therefore, the proof derived in Sec. III B 1 is extended by relaxing *Constraint 1*. This amounts to applying the contraction principle to $\bar{J}_\gamma = \mathbb{O}_{o\gamma} J_\gamma$. This symmetry has previously been realized on the Lagrange multipliers λ_{J_o} and λ_{T_o} in the right-hand side of Eq. (36), where χ^*_o depends linearly on λ_{J_o} and λ_{T_o} with a constant multiplied by $\mathbb{O}_{o\gamma}$.

Similarly, we exploit the scale invariance of EPR to extend the variational formulation for observable currents by relaxing *Constraint 2*. For this purpose, the linear algebraic formulation implies, to the preserve the invariance of $\dot{\Sigma}_{\{o\}}$ $J_\gamma \to \bar{J}_\gamma = kJ_\gamma$ should be compensated by $\chi^*_\gamma \to \bar{\chi}^*_\gamma = \chi^*_\gamma/k$ defined for microscopic currents. To this end, we utilize the bilinear form of the EPR, $\dot{\Sigma} = \sum_{\{\gamma=\}} J_\gamma \chi^*_\gamma = (\vec{\chi}^*_\gamma)^T \vec{J}_\gamma$. Thus, the problem is reduced to a norm-preserving basis transformation in linear algebra. How do effective observable affinities transform under the dual transformation that preserves the correct scaling of $\dot{\Sigma}$ using $\dot{\Sigma}_{\{o\}}$?, given that microscopic currents transform as: $\vec{J}_o = \mathbb{O}\vec{J}_\gamma$ with $\mathbb{O}_{o'\gamma} \in \mathbb{R}^+$. By not assuming any constraint on the structure of the matrix $\mathbb{O}$, the linear algebraic formulation of the problem inherently violates *Constraint 2*. Therefore, the effective affinity of the observable current should transform as $(\vec{\chi}^*_o)^T = (\vec{\chi}^*_\gamma)^T \mathbb{O}^\dagger$, where $\mathbb{O}^\dagger = \mathbb{O}^T(\mathbb{O}\mathbb{O}^T)^{-1}$ is the right pseudoinverse that dictates the transformation of $(\vec{\chi}^*_\gamma)^T$ in the dual conjugate space to the current space. This leads to $\dot{\Sigma}_{\{o\}} = \mathcal{L}^*_{\{o\}}[\{J_o, T_o\}] = \sum_{\{o\}} J_o \chi^*_o = (\vec{\chi}^*_o)^T \vec{J}_o$.

Using the linear algebraic solution, we highlight the consequences of the two constraints on the structure of effective observable affinities. Due to *Constraint 2*: the observable conjugate affinities simplify to $\chi^*_o = \sum_{\gamma \in \mathrm{supp}(o)} \chi^*_\gamma \mathbb{O}_{o\gamma}/||O_o||$, where $||O_o|| = \sum_\gamma \mathbb{O}^2_{o\gamma}$, since $\mathbb{O}\mathbb{O}^T$ is decomposed into a block-diagonal form, resulting in decoupling of χ_o for linearly independent observable currents. Due to *Constraint 1*: $\chi^*_o = \sum_{\gamma^= \in \mathrm{supp}(o)} \chi^*_\gamma/|N_o|$. Violation of *Constraint 2*: in this case, $\{J_o\}$ are not linearly independent due to $\mathbb{O}_{o^1\gamma'} \neq 0 \not\Rightarrow \mathbb{O}_{o^2\gamma'} = 0$. Although χ^*_o is obtained using $\mathbb{O}^\dagger$, a simplified closed-form expression for χ^*_o is not available compared to previous cases. Because, due to cross-coupling terms, a microscopic transition current contributes to multiple observable currents, which leads to nonvanishing nondiagonal terms in $(\mathbb{O}\mathbb{O}^T)^{-1}$. This cross-coupling is transferred from the transformation $\{\chi^*_\gamma\} \to \{\chi^*_o\}$, making it rather difficult to visualize the transformation. However, a system-specific brute-force or numerical computation of $\mathbb{O}^\dagger$ is always feasible.

IV. MANIFESTATIONS OF MINIMUM ACTION PRINCIPLE

A. Single observable currents: Stochastic EPR is the most precise current

For a single observable current, $\mathbb{O}$ is reduced a vector $\vec{O}$, which leads to $(\vec{O}\vec{O}^T) = \sum_{\{\gamma=\}}(O_\gamma)^2$, therefore, $\chi^*_o = [\sum_{\{\gamma=\}} \chi^*_\gamma O_\gamma]/[\sum_{\{\gamma=\}}(O_\gamma)^2]$. Choosing $\vec{O} = \vec{\chi}^*_\gamma$ corresponds to the stochastic EPR as an observable current; which leads to $\chi^*_o = 1$. Imposing the normalization of the observable current, $\sum_{\{\gamma=\}}(O_\gamma)^2 = 1$, we investigate the condition to observe the most precise observable current. This amounts to solving the variational optimization problem: $\chi^{\max}_o = \sup_{\vec{O}}(\chi_o)$ under the normalization constraint. Therefore, $\delta\chi^*_o/\delta O_\gamma = 0, \forall\gamma \in \{\gamma^=\}$, which implies $O_\gamma/\chi^*_\gamma = O_{\gamma'}/\chi^*_{\gamma'} = \text{constant}, \forall\gamma, \gamma' \in \{\gamma^=\}$. Using the normalization constraint, its unique optimal solution is $O_\gamma = \chi^*_\gamma/||\chi^*||$ or $\vec{O}_\gamma = \vec{\chi}^*_\gamma/||\chi^*||$ and $\chi^{\max}_o = ||\chi^*||$, where $||\chi^*|| = \sqrt{\sum_{\{\gamma=\}}(\chi^*_\gamma)^2}$ denotes the absolute value or the length.

Physically, this implies that, among all normalized single observable currents J_o, the maximum observable affinity corresponds to choosing a current along (parallel to) the stochastic EPR, $J_o = \dot{\Sigma}_{\mathrm{st}}/||\chi^*||$ with an effective affinity $\chi_{\dot{\Sigma}_{\mathrm{st}}} = \chi^{\max}_o/||\chi^*|| = 1$. Analogously formulated, for the thermodynamic inference using a single observable current, the stochastic EPR is the most precise observable current that maximizes the quantification of nonequilibrium-ness, it is a 'special' current that exactly quantifies the microscopic thermodynamic dissipation on the macroscale, which is otherwise lost due to the suboptimal choice J_o. The least optimal observable current for thermodynamic inference is orthogonal to the stochastic EPR. Since $J_o \perp \vec{\chi}^*_\gamma$, in this case, $\chi^*_o = 0$, because $\sum_{\{\gamma=\}} O_\gamma \chi^*_\gamma = 0$. These results hold independently of the dimension of $\{\gamma^=\}$ and the steady-state assumption for any observation time τ, and apply to any multicyclic system.

B. Partial control description

If the transition affinities are known, then $\chi^*_\gamma = A_\gamma$, which corresponds to the partial control description of MinAP. Then $\mathcal{L}^*_{\{o\}}$ is reduced to its bilinear form. Here, the partial control description refers to the control of the transition affinities, which are fixed and known, but the transition currents are still the "uncontrollable" stochastic observables.

1. Fluctuation relation

One observes that the normalization condition for the probability distribution Eq. (33) trivially implies the integrated FR $\langle e^{-\tau(\vec{\chi}^*_o)^T \vec{J}_o}\rangle = 1$. Defining the scaled cumulant generating function (SCGF) for observable currents, $\mathcal{K}_{\vec{J}_o}(\chi_o) = \lim_{\tau\to\infty} \frac{1}{\tau} \ln \langle e^{\tau \vec{\chi}_o \cdot \vec{J}_o}\rangle$ [26], where SCGF is defined with respect to the large deviation parameter. SCGF defines the nonequilibrium analog of the system's free energy. Using $\mathcal{P}[\vec{J}_o] \asymp e^{-\tau \vec{\chi}^*_o \cdot \vec{J}_o}$, it is known SCGF satisfies the Gallavotti-

Cohen FR symmetry: $\mathcal{K}_{\vec{\tilde{J}}_o}(\vec{\chi}_o) = \mathcal{K}_{\vec{\tilde{J}}_o}(-\vec{\chi}_o^* - \vec{\chi}_o)$ [18,24–26] and reveals the asymmetry/symmetry of the detailed FR. Since χ_o^* corresponds to the nontrivial root of SCGF, thereby quantifying the detailed FR symmetry for the time-integrated observable stochastic currents $\vec{\tilde{J}}_o$ [24–26]:

$$\ln\left(\frac{\mathcal{P}[\vec{\tilde{J}}_o = \langle\vec{\tilde{J}}_o\rangle]}{\mathcal{P}[\vec{\tilde{J}}_o = -\langle\vec{\tilde{J}}_o\rangle]}\right) = \tau\vec{\chi}_o^* \cdot \langle\vec{\tilde{J}}_o\rangle = \mathcal{S}_{\{o\}}^*, \tag{38}$$

where $\mathcal{P}[\vec{\tilde{J}}_o = \langle\vec{\tilde{J}}_o\rangle]$ is the shorthand notation for the probability density of observing a value $\langle\vec{J}_o\rangle$ for $\vec{\tilde{J}}_o$.

We consider the notable case of observable currents. First, the stochastic EPR as an observable current, $J_o = \dot{\Sigma}_{\rm st} = \sum_{\{\gamma^{\rightleftharpoons}\}} \chi_\gamma^*$, which results in $\chi_o^* = 1$, derived in Sec. IV A, leading to the Lebowitz-Spohn symmetry: $\mathcal{K}_{\tilde{\Sigma}_{\rm st}}(\chi_{\tilde{\Sigma}_{\rm st}}) = \mathcal{K}_{\tilde{\Sigma}_{\rm st}}(-1 - \chi_{\tilde{\Sigma}_{\rm st}})$, and the detailed FR for $\dot{\Sigma}_{\rm st}$,

$$\ln\left(\frac{\mathcal{P}[\tilde{\Sigma}_{\rm st} = \langle\tilde{\Sigma}_{\rm st}\rangle]}{\mathcal{P}[\tilde{\Sigma}_{\rm st} = -\langle\tilde{\Sigma}_{\rm st}\rangle]}\right) = \tau\langle\tilde{\Sigma}_{\rm st}\rangle, \tag{39}$$

and the integrated FR for $\dot{\Sigma}_{\rm st}$ is $\langle e^{-\tau\tilde{\Sigma}_{\rm st}}\rangle = 1$ [19,26]. Second, if we consider the most fundamental case of all microscopic transition currents, $\{o\} = \{\gamma^{\rightleftharpoons}\}$, with known transition affinities, $\{\chi_\gamma^*\} = \{A_\gamma\}$. The detailed FR for $\{J_\gamma\}$ is

$$\ln\left(\frac{\mathcal{P}[\vec{\tilde{J}}_\gamma = \langle\vec{\tilde{J}}_\gamma\rangle]}{\mathcal{P}[\vec{\tilde{J}}_\gamma = -\langle\vec{\tilde{J}}_\gamma\rangle]}\right) = \tau\vec{A}_\gamma \cdot \langle\vec{\tilde{J}}_\gamma\rangle, \tag{40}$$

and the integrated FR for J_γ is $\langle e^{-\tau A_\gamma \tilde{J}_\gamma}\rangle = 1$.

2. *Effective affinity and martingale property*

Effective affinity χ_γ^* plays a key role in quantifying FR symmetry. Here, we highlight its underlying mathematical structure, namely, the martingale property [156,156–164], which has important implications for thermodynamic inference in the absence of observable currents [165–167], using first-passage time statistics or waiting-time statistics [158,163,164], and applications to thermodynamic inference [168].

TL satisfies the additive property $\tau\tilde{J}_o[\tau] = \tau'\tilde{J}_o[\tau'] + (\tau - \tau')\tilde{J}_o[\tau - \tau']$, with the initial value condition $\tilde{J}_o[\tau] = 0$. If χ_o^* is time-homogeneous, that is, $\chi_o^*[\tau] = \chi_o^*[\tau'] = \chi_o^*[\tau - \tau']$, then the action satisfies the additive property $\mathcal{S}_{\{o\}}^*[\tau] = \mathcal{S}_{\{o\}}^*[\tau'] + \mathcal{S}_{\{o\}}^*[\tau - \tau']$, with the initial value condition $\mathcal{S}_{\{o\}}^*[\tau] = 0$. The integrated FR $\langle e^{-\mathcal{S}_{\{o\}}^*}\rangle = 1$ is also satisfied. These are sufficient conditions for $e^{-\mathcal{S}_{\{o\}}^*}$(or equivalently $e^{-\Sigma_{\{o\}}}$), to be a martingale [156]. The time-homogeneity of χ_o^* required for the martingale property of J_o is equivalent to the observable being memory-less; a regime in which the LDB condition effectively emerges from MinAP as discussed in Sec. II B 7. This results in the linear scaling of $\mathcal{S}_{\{o\}}^*[\tau]$ in τ, ensuring the time-additive property, and violations of it due to transient effects need to be studied [50,51]. From the most fundamental perspective, if we choose all microscopic currents $\{o\} = \{\gamma^{\rightleftharpoons}\}$, then the exponentiated negative of J_γ is a Martingale with effective affinity χ_γ^* that characterizes the directional asymmetry between observation of the positive and negative amplitudes of its value. This physical property has been realized earlier through the detailed FR symmetry Eq. (40) for microscopic currents. Therefore, the martingale property of the microscopic transition currents is the most fundamental thermodynamic symmetry, the stochastic EP is a 'special' case, whose martingale property has been rigorously studied [156–161].

In the measure-theoretic formalism, the Radon-Nikodym derivative (RND) is defined here as the ratio of the transition probability measure between the forward and backwards processes, which by definition is the exponential of left side of Eqs. (38)–(40). The equivalence between the stochastic EP (a physical property) and the logarithm of RND (a mathematical property) assigns a thermodynamic meaning to it [169–172]. It unveils the FR symmetry and martingale property of stochastic EP [169–172]. Here, we extend the measure-theoretical formalism to the most fundamental microscopic transition currents of discrete-state processes Eq. (40), so that the existing measure-theoretical understanding of stochastic EP is recovered by the contraction of microscopic currents to stochastic EP [Eq. (40) to Eq. (39)] [169–172]. We have briefly outlined the connection to martingale structure and measure-theoretical formulation; however, a more systematic and rigorous mathematical analysis, as well as stronger implications of martingale theory [156] or measure theory [169] for other physical observables and practical applications such as thermodynamic inference [168] remain to be explored.

3. *Orthogonal decomposition of EPR*

Time-integrated relaxation and dissipative currents satisfy the scalings $\int_0^\tau dt J_\gamma^{\rm rel} \propto O(1)$ and $\int_0^\tau dt J_\gamma^{\rm ss} \propto \tau$, respectively. Physically, this results in short-time and long-time symmetries of currents for the relaxation and dissipative currents, respectively. The origin of the orthogonality of relaxation and steady state is attributed to the decomposition of $A_\gamma = A_\gamma^{\rm rel} + A_\gamma^{\rm ss}$ into its boundary and bulk terms, respectively, which results in the excess and housekeeping EPR scalings, $\Sigma^{\rm ex} \propto O(1)$ and $\Sigma^{\rm hk} \propto O(\tau)$. However, we have assumed a dissipative scaling of $\{J_\gamma\}$ to obtain the LDP from the variational formulation. Thus, the LDP must be modified to accurately account for the different scalings of the relaxation and dissipative currents, thereby restoring the short-time symmetry of the stochastic currents.

Consider the orthogonal decomposition of the transition affinities $\chi_\gamma^* = \chi_\gamma^{\rm ss} + \chi_\gamma^{\rm ex}$, such that, $\chi_\gamma^{\rm ss} = F_\gamma$ and $\chi_\gamma^{\rm ex} = \Delta_\gamma \chi_i^{\rm ex}$ with $\chi_i^{\rm ex} = -\ln(\rho_i/\rho_i^E) = S_i^{{\rm state}|E}$, where $\chi_\gamma^{\rm ex}$ characterizes the distance from the Boltzmann distribution $\rho_i^E = e^{-E_i + \psi_E}$. Using the linearly independent decomposition of the transition affinities into relaxation and dissipative components, we define, $\vec{J}_o = (-\partial_t\psi_E, -\sum_{\{i\}}\ln(\rho_i/\rho_i^E), \vec{F}_\gamma)$, with the respective scaling vector $\vec{\Omega} = (1, 1, \vec{\tau})$. Furthermore, we consider the total dissipative transition affinity as a single dissipative current. Therefore, $\vec{J}_o = (-\partial_t\psi_E, -\sum_{\{i\}}\ln(\rho_i/\rho_i^E), \sum_{\{\gamma^{\rightleftharpoons}\}} F_\gamma)$, with the scaling vector $\vec{\Omega} = (1, 1, \tau)$. The third term here implies, using the observable transition current F_γ, when a transition $\gamma^{\rightleftharpoons}$ takes place. Hence, the choice of $\vec{J}_o$ is equivalent to the orthogonal decomposition of the stochastic EPR $\vec{\dot{\Sigma}}_{\rm st} = (\dot{W}_{\rm qs}, \sum_i \dot{S}_i^{{\rm state}|E}, \dot{\Sigma}^{\rm hk})$ with $\vec{\chi}_{\rm st}^* = (1, 1, 1)$. Here, the stochas-

tic EPR $\vec{\tilde{\Sigma}}_{\rm st}$ is decomposed into three linearly independent contributions. First, the quasistatic driving work rate $\dot{W}_{\rm qs}$, a boundary term in the control parameter space $\{\lambda\}$ of $E(\{\lambda\})$, which depends on the explicit time-dependent driving of $\{\lambda\}$. Second, the relaxation/excess EPR, a boundary term in the probability state space $\{\rho_i\}$, which is the statistical distance between the initial and final states relative to the reference Boltzmann distribution. Third, the housekeeping EPR, a bulk term that scales with τ and is supported by a dissipative bath: a nonvanishing dissipative EP contribution in the steady state. Thus, a detailed FR symmetry for the orthogonal decomposition of the EP reads

$$\ln\left(\frac{\mathcal{P}[\vec{\tilde{\Sigma}}_{\rm st}=\langle\vec{\tilde{\Sigma}}_{\rm st}\rangle]}{\mathcal{P}[\vec{\tilde{\Sigma}}_{\rm st}=-\langle\vec{\tilde{\Sigma}}_{\rm st}\rangle]}\right)=\vec{\chi}^{*}_{\rm st}\cdot\vec{\Omega}\odot\langle\vec{\tilde{\Sigma}}_{\rm st}\rangle, \tag{41}$$

where $\odot$ denotes a component-wise Hadamard product defined between $\vec{\Omega}$ and $\vec{\tilde{\Sigma}}_{\rm st}$, simplified to obtain the total time-integrated EP, $\chi^{*}_{\rm st}\cdot\vec{\Omega}\odot\langle\vec{\tilde{\Sigma}}_{\rm st}\rangle=\langle-\Delta^{\tau}_{0}\psi_E+\Delta^{\tau}_{0}S^{E}_{\rm state}+\tau\tilde{\Sigma}_{\rm hk}\rangle$.

C. Thermodynamic inference description

We show applications of the MinAP to thermodynamic inference, and discuss three cases using state-space observables, which are experimentally easily assessable [173], compared to TKUR which requires current statistics.

1. Nonquadratic speed limit

If observable currents $\{o\}=\{J_i\}$ into the state $\{\rho_i\}$ are chosen, this choice of observable currents corresponds to the contraction from the transition space to the state space through the continuity equation, $\partial_t\rho_i=J_i$. The traffic defined for the state ρ_i is $T_i=\sum_{i\in\{\gamma=\}}T_\gamma$ and quantifies the total scaled variance of ρ_i due to microscopic transitions (where ρ_i is involved). The time-integrated continuity equation implies $\Delta^{\tau}_{0}\rho_i=\rho_i(\tau)-\rho_i(0)=\tau\tilde{J}_i$. This reduces Eq. (32) to

$$\Sigma_{\rm SL}=2\Delta^{\tau}_{0}\rho_i\tanh^{-1}\left(\frac{\Delta^{\tau}_{0}\rho_i}{\tau\tilde{T}_i}\right). \tag{42}$$

Equation (42) is a nonquadratic speed limit and generalizes the quadratic speed limit from Ref. [99]. Here, $\tilde{T}_i$ is the scaled diffusion constant for ρ_i. The approximation $\tanh^{-1}(x)\approx x$ gives the quadratic speed limit in its more familiar form [99]. However, the mismatch increases for fEQ systems, as discussed earlier.

2. Nonquadratic Onsager-Machlup functional and fluctuations around steady state

We aim to use the MinAP to study the fluctuations around the steady state. For this purpose, we choose the relaxation currents of states $\{i\}$, $\{J_o\}=\{J^{\rm rel}_i\}$. Using the orthogonal decomposition $\{J_i\}$ into dissipative and relaxation currents, the continuity equation $\partial_t\rho_i=J^{\rm rel}_i+J^{\rm ss}_i$, and the short-time nonquadratic TKUR Eq. (31), we obtain the excess Lagrangian $\mathcal{L}^{*}_{\rm rel}$ for fluctuations around the steady state,

$$\mathcal{L}^{*}_{\rm rel}=\sum_{\{i\}}2(\partial_t\rho_i-J^{\rm ss}_i)\tanh^{-1}\left(\frac{\partial_t\rho_i-J^{\rm ss}_i}{\tau\tilde{T}_i}\right), \tag{43}$$

with a nonquadratic Onsager-Machlup functional for the probability distribution of the fluctuations around the steady state,

$$\mathcal{P}[\{\rho_i,J^{\rm ss}_i\}]\asymp e^{-\int_0^\tau\mathcal{L}^{*}_{\rm rel}dt}. \tag{44}$$

The Gaussian approximation $\tanh^{-1}(x)\approx x$ of Eqs. (43) and (44) leads to the quadratic Onsager-Machlup functional [81,82], derived originally for Gaussian fluctuations around the equilibrium steady state but generalized here with Eqs. (43) and (44) for any nonequilibrium steady state $\{J^{\rm ss}_i\}$ and incorporating non-Gaussian fluctuations.

Importantly, if the relaxation-fluctuation symmetry is satisfied, then $\mathcal{L}^{*}_{\rm rel}=-d_tD^{\rm KL}_{\rm ss}$. However, this is not generally the case, since Eq. (43) assigns a nonquadratic thermodynamic EPR cost to fluctuations around the steady state. In contrast, $-d_tD^{\rm KL}_{\rm ss}$ is the thermodynamic EPR cost associated with the relaxation process (the gradient descent) toward the steady state, with $-D^{\rm KL}_{\rm ss}$ being the corresponding Lyapunov functional for the relaxation process. Since the relaxation-fluctuation symmetry is not necessarily satisfied in fEQ systems, we clarify these differences between the fluctuations around the steady state and the relaxation toward the steady state, governed by $\mathcal{L}^{*}_{\rm rel}$ in Eq. (43) and $-d_tD^{\rm KL}_{\rm ss}$, respectively.

3. Nonquadratic state-space TKUR

A novel class of systems that breaks the "actio=reactio" symmetry are called "nonreciprocal systems" and manifests the formation of vorticity currents, defined between two states is defined as $\omega_{ij}=\rho_j\partial_t\rho_i-\rho_i\partial_t\rho_j$ [5]. Importantly, ω_{ij} is analog of an nonequilibrium current, which is defined in state space and does not require knowledge of the underlying topology of transitions (graph). This makes thermodynamic inference using ω_{ij} suitable and appealing for experimental purposes [173–177].

We choose ρ_jJ_i as the unidirectional observable current, which leads to a bidirectional current $J_o=\rho_iJ_j-\rho_jJ_i$. Using the continuity equation, $\omega_{ij}=\rho_i\partial_t\rho_j-\rho_j\partial_t\rho_i=J_o$. The corresponding observable traffic is defined as $\omega^{s}_{ij}=T_o=\rho_i\partial_t\rho_j+\rho_j\partial_t\rho_i$. Choosing the set of all possible combinations of state pairs $n(n-1)/2$ leads to all linearly independent vorticity currents as observable currents in state space, $\{J_o\}=\{\omega_{ij}\}$. Defining the temporal state correlations, $C_{ij}(\tau)=\rho_i(\tau)\rho_j(0)$, between ρ_i and ρ_j over time τ [174] and decomposing into its state-symmetric and state-antisymmetric components, $C^{s}_{ij}(\tau)=\rho_i(\tau)\rho_j(0)+\rho_j(\tau)\rho_i(0)$ and $C^{a}_{ij}(\tau)=\rho_j(\tau)\rho_i(0)-\rho_i(\tau)\rho_j(0)$, respectively. The time-integrated vorticity $\tau\tilde{\omega}_{ij}=\int_0^\tau\omega_{ij}dt=\Delta^{\tau}_{0}C^{a}_{ij}(\tau)$ and time-integrated traffic $\tau\tilde{\omega}^{s}_{ij}=\int_0^\tau\omega^{s}_{ij}dt=\Delta^{\tau}_{0}C^{s}_{ij}(\tau)$ are simplified, where $\Delta^{\tau}_{0}C^{a}_{ij}(\tau)=C^{a}_{ij}(\tau)-C^{a}_{ij}(0)$ and $\Delta^{\tau}_{0}C^{s}_{ij}(\tau)=C^{s}_{ij}(\tau)-C^{s}_{ij}(0)$ quantify the change over observation time τ.

Hence, using Eq. (32) for $\{J_o\}=\{\omega_{ij}\}$ leads to

$$\Sigma_{\{\omega\}}=\int_0^\tau\mathcal{L}^{*}_{\{\omega\}}dt\geqslant\sum_{\{ij\}}2\Delta^{\tau}_{0}C^{a}_{ij}\tanh^{-1}\left(\frac{\Delta^{\tau}_{0}C^{a}_{ij}}{\Delta^{\tau}_{0}C^{s}_{ij}}\right). \tag{45}$$

Equation (45) is the nonquadratic state-space TKUR quoted in Ref. [5] for nonreciprocal systems. It obtains a bound on Σ using using state-space temporal correlations $C^{a}_{ij}(\tau)$ and $C^{s}_{ij}(\tau)$ (instead of the usual current-space formulation). The choice $\{J_o\}=\{\omega_{ij}\},\forall i,j\in\{i\}$ obtains the tightest bound on Σ using

all linearly independent microscopic vorticity currents. By implementing a state-space contraction, the results derived in this section hold for any choice of coarse-grained vorticity current defined between two observable state-like quantities: "effectively" nonreciprocal systems, and are closely related to the results obtained in Refs. [176–179].

V. CONCLUSION AND OUTLOOK

We have presented a unified framework of the MinAP for the EPR of discrete-state systems. By deriving an exact stochastic path integral representation of discrete-state transition dynamics, which is equal to exponentiated action and incorporates non-Gaussian transition fluctuations/effective drivings, which results in an exact nonquadratic dissipation function. This formulation provides a physical interpretation of the action Lagrangian as mean inferred EPR, analogous to the role of the energy functional in the equilibrium Boltzmann distribution. This generalization allows us to formulate a far-from-equilibrium analog of the canonical ensemble that relates EPR to transition-space mean currents and its variances and defines the TL of microscopic currents, which are linked through the exact nonquadratic dissipation function. Using this, we derive an exact nonquadratic large deviation rate functional, which tightens the bounds on EP/EPR compared to previous close-to-equilibrium Gaussian (quadratic) and far-from-equilibrium Hessian formulations, which physically correspond to quadratic Thermodynamic-kinetic uncertainty relation and nonequilibrium linear-response. We show that the variational formulation derived here is equivalent to the Information geometric formulation, extending the applicability of Information geometric methodologies to Stochastic thermodynamics, provided thermodynamic consistency is ensured.

Using TL, we show that the nonquadratic TKUR and the FR are manifestations of the MinAP as thermodynamic inference and partial control descriptions, respectively. This unifies FR and nonquadratic TKUR within a single framework. Moreover, we extend the applicability of MinAP to coarse-grained observable currents, making it applicable to practically accessible experimental setups/systems. The variational formulation is also particularly helpful for implementing numerical optimization in cases where an analytical solution cannot be obtained. Although our framework is developed for discrete-state systems modeled by graphs [7], it is easily extended to hypergraphs that model other physical systems, for example, nonlinear chemical reaction networks [144]. This work lays the foundation for practical applications of the minimum action principle in stochastic thermodynamics of far-from-equilibrium systems. For example, the generalized finite-time optimal control framework for discrete-state systems is developed in Ref. [120].

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

[1] L. Boltzmann, *Lectures on Gas Theory* (University of California Press, Berkeley, CA, 1964).

[2] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).

[3] K. Sekimoto, *Stochastic Energetics* (Courier Corporation, North Chelmsford, MA, 2010).

[4] N. Shiraishi, *An Introduction to Stochastic Thermodynamics: From Basic to Advanced* (Springer Nature, Singapore, 2023), pp. 31–47.

[5] A. T. Mohite and H. Rieger, Stochastic thermodynamics of nonreciprocally interacting particles and fields, Phys. Rev. E **113**, 064136 (2026).

[6] A. T. Mohite and H. Rieger, Thermodynamically consistent coarse-graining: From interacting particles to fields via second quantization, arXiv:2508.11430.

[7] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Rev. Mod. Phys. **48**, 571 (1976).

[8] G. N. Bochkov and I. E. Kuzovlev, General theory of thermal fluctuations in nonlinear systems, Zh. Eksp. Teor. Fiz. **72**, 238 (1977) [Sov. Phys. JETP **45**, 125 (1977)].

[9] G. N. Bochkov and Y. E. Kuzovlev, Fluctuation-dissipation relations for nonequilibrium processes in open systems, Zh. Eksp. Teor. Fiz. **76**, 1071 (1979) [Sov. Phys. JETP **49**, 543 (1979)].

[10] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Probability of second law violations in shearing steady states, Phys. Rev. Lett. **71**, 2401 (1993).

[11] D. J. Evans and D. J. Searles, Equilibrium microstates which generate second law violating steady states, Phys. Rev. E **50**, 1645 (1994).

[12] C. Jarzynski, Nonequilibrium equality for free energy differences, Phys. Rev. Lett. **78**, 2690 (1997).

[13] C. Jarzynski, Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach, Phys. Rev. E **56**, 5018 (1997).

[14] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[15] H. Tasaki, Jarzynski relations for quantum systems and some applications, arXiv:cond-mat/0009244.

[16] G. E. Crooks, Path-ensemble averages in systems driven far from equilibrium, Phys. Rev. E **61**, 2361 (2000).

[17] C. Maes and K. Netočný, Time-reversal and entropy, J. Stat. Phys. **110**, 269 (2003).

[18] G. Gallavotti and E. G. D. Cohen, Dynamical ensembles in nonequilibrium statistical mechanics, Phys. Rev. Lett. **74**, 2694 (1995).

[19] J. L. Lebowitz and H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).

[20] J. Kurchan, Fluctuation theorem for stochastic dynamics, J. Phys. A: Math. Gen. **31**, 3719 (1998).

[21] K. Sekimoto, Kinetic characterization of heat bath and the energetics of thermal ratchet models, J. Phys. Soc. Jpn. **66**, 1234 (1997).

[22] K. Sekimoto, Langevin equation and thermodynamics, Prog. Theor. Phys. Suppl. **130**, 17 (1998).
[23] U. Seifert, Entropy production along a stochastic trajectory and an integral fluctuation theorem, Phys. Rev. Lett. **95**, 040602 (2005).
[24] D. Andrieux and P. Gaspard, Fluctuation theorem for currents and Schnakenberg network theory, J. Stat. Phys. **127**, 107 (2007).
[25] D. Andrieux and P. Gaspard, Network and thermodynamic conditions for a single macroscopic current fluctuation theorem, Comptes Rendus. Physique **8**, 579 (2007).
[26] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).
[27] R. Gilmore, Uncertainty relations of statistical mechanics, Phys. Rev. A **31**, 3237 (1985).
[28] J. Uffink and J. van Lith, Thermodynamic uncertainty relations, Found. Phys. **29**, 655 (1999).
[29] J. M. Horowitz and T. R. Gingrich, Thermodynamic uncertainty relations constrain non-equilibrium fluctuations, Nat. Phys. **16**, 15 (2020).
[30] A. C. Barato and U. Seifert, Thermodynamic uncertainty relation for biomolecular processes, Phys. Rev. Lett. **114**, 158101 (2015).
[31] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Dissipation bounds all steady-state current fluctuations, Phys. Rev. Lett. **116**, 120601 (2016).
[32] J. M. Horowitz and T. R. Gingrich, Proof of the finite-time thermodynamic uncertainty relation for steady-state currents, Phys. Rev. E **96**, 020103(R) (2017).
[33] I. Di Terlizzi and M. Baiesi, Kinetic uncertainty relation, J. Phys. A: Math. Theor. **52**, 02LT03 (2019).
[34] E. Kwon, J.-M. Park, J. S. Lee, and Y. Baek, Unified hierarchical relationship between thermodynamic tradeoff relations, Phys. Rev. E **110**, 044131 (2024).
[35] T. Van Vu and K. Saito, Thermodynamic unification of optimal transport: Thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits, Phys. Rev. X **13**, 011013 (2023).
[36] V. T. Vo, T. V. Vu, and Y. Hasegawa, Unified thermodynamic–kinetic uncertainty relation, J. Phys. A: Math. Theor. **55**, 405004 (2022).
[37] V. T. Vo, T. Van Vu, and Y. Hasegawa, Unified approach to classical speed limit and thermodynamic uncertainty relation, Phys. Rev. E **102**, 062132 (2020).
[38] J. S. Lee, S. Lee, H. Kwon, and H. Park, Speed limit for a highly irreversible process and tight finite-time Landauer's bound, Phys. Rev. Lett. **129**, 120603 (2022).
[39] S. Ito, Stochastic thermodynamic interpretation of information geometry, Phys. Rev. Lett. **121**, 030605 (2018).
[40] N. Merhav and Y. Kafri, Statistical properties of entropy production derived from fluctuation theorems, J. Stat. Mech.: Theory Exp. (2010) P12022.
[41] Y. Hasegawa and T. Van Vu, Fluctuation theorem uncertainty relation, Phys. Rev. Lett. **123**, 110602 (2019).
[42] A. M. Timpanaro, G. Guarnieri, J. Goold, and G. T. Landi, Thermodynamic uncertainty relations from exchange fluctuation theorems, Phys. Rev. Lett. **123**, 090604 (2019).
[43] G. Francica, Fluctuation theorems and thermodynamic uncertainty relations, Phys. Rev. E **105**, 014129 (2022).
[44] M. J. Klein and P. H. Meijer, Principle of minimum entropy production, Phys. Rev. **96**, 250 (1954).
[45] H. B. Callen, Principle of minimum entropy production, Phys. Rev. **105**, 360 (1957).
[46] P. Glansdorff, G. Nicolis, and I. Prigogine, The thermodynamic stability theory of non-equilibrium states, Proc. Natl. Acad. Sci. USA **71**, 197 (1974).
[47] E. T. Jaynes, The minimum entropy production principle, Annu. Rev. Phys. Chem. **31**, 579 (1980).
[48] H. Struchtrup and W. Weiss, Maximum of the local entropy production becomes minimal in stationary processes, Phys. Rev. Lett. **80**, 5048 (1998).
[49] H. Qian, Entropy production and excess entropy in a nonequilibrium steady-state of single macromolecules, Phys. Rev. E **65**, 021111 (2002).
[50] R. M. L. Evans, Rules for transition rates in nonequilibrium steady states, Phys. Rev. Lett. **92**, 150601 (2004).
[51] R. M. L. Evans, Detailed balance has a counterpart in nonequilibrium steady states, J. Phys. A: Math. Gen. **38**, 293 (2005).
[52] V. Lecomte, C. Appert-Rolland, and F. van Wijland, Chaotic properties of systems with Markov dynamics, Phys. Rev. Lett. **95**, 010601 (2005).
[53] L. Martyushev and V. Seleznev, Maximum entropy production principle in physics, chemistry and biology, Phys. Rep. **426**, 1 (2006).
[54] Q. Wang, Maximum entropy change and least action principle for nonequilibrium systems, Astrophys. Space Sci. **305**, 273 (2006).
[55] S. Bruers, C. Maes, and K. Netočný, On the validity of entropy production principles for linear electrical circuits, J. Stat. Phys. **129**, 725 (2007).
[56] S. Bruers, A discussion on maximum entropy production and information theory, J. Phys. A: Math. Theor. **40**, 7441 (2007).
[57] V. Lecomte, C. Appert-Rolland, and F. van Wijland, Thermodynamic formalism for systems with Markov dynamics, J. Stat. Phys. **127**, 51 (2007).
[58] Z. Yoshida and S. M. Mahajan, "Maximum" entropy production in self-organized plasma boundary layer: A thermodynamic discussion about turbulent heat transport, Phys. Plasmas **15**, 032307 (2008).
[59] A. Baule and R. M. L. Evans, Invariant quantities in shear flow, Phys. Rev. Lett. **101**, 240601 (2008).
[60] A. Baule and R. M. L. Evans, Nonequilibrium statistical mechanics of shear flow: Invariant quantities and current relations, J. Stat. Mech.: Theory Exp. (2010) P03030.
[61] L. M. Martyushev, The maximum entropy production principle: Two basic questions, Phil. Trans. R. Soc. B **365**, 1333 (2010).
[62] R. K. Niven, Simultaneous extrema in the entropy production for steady-state fluid flow in parallel pipes, J. Non-Equilib. Thermodyn. **35**, 347 (2010).
[63] Y. Kawazura and Z. Yoshida, Entropy production rate in a flux-driven self-organizing system, Phys. Rev. E **82**, 066403 (2010).
[64] M. Doi, Onsager's variational principle in soft matter, J. Phys.: Condens. Matter **23**, 284118 (2011).
[65] C. Monthus, Non-equilibrium steady states: Maximization of the Shannon entropy associated with the distribution of dy-

namical trajectories in the presence of constraints, J. Stat. Mech.: Theory Exp. (2011) P03008.
[66] Y. Kawazura and Z. Yoshida, Comparison of entropy production rates in two different types of self-organized flows: Benard convection and zonal flow, Phys. Plasmas **19**, 012305 (2012).
[67] R. Chetrite and H. Touchette, Nonequilibrium microcanonical and canonical ensembles and their equivalence, Phys. Rev. Lett. **111**, 120601 (2013).
[68] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, Principles of maximum entropy and maximum caliber in statistical physics, Rev. Mod. Phys. **85**, 1115 (2013).
[69] R. G. Endres, Entropy production selects nonequilibrium states in multistable systems, Sci. Rep. **7**, 14437 (2017).
[70] W. Bialek, Stability and noise in biochemical switches, arXiv:cond-mat/0005235.
[71] M. Sasai and P. G. Wolynes, Stochastic gene expression as a many-body problem, Proc. Natl. Acad. Sci. USA **100**, 2374 (2003).
[72] Y. Lan, P. G. Wolynes, and G. A. Papoian, A variational approach to the stochastic aspects of cellular signal transduction, J. Chem. Phys. **125**, 124106 (2006).
[73] M. Vellela and H. Qian, Stochastic dynamics and nonequilibrium thermodynamics of a bistable chemical system: The Schlögl model revisited, J. R. Soc. Interface. **6**, 925 (2009).
[74] T. Bodineau and B. Derrida, Current fluctuations in nonequilibrium diffusive systems: An additivity principle, Phys. Rev. Lett. **92**, 180601 (2004).
[75] B. Derrida, Non-equilibrium steady states: Fluctuations and large deviations of the density and of the current, J. Stat. Mech.: Theory Exp. (2007) P07023.
[76] A. Prados, A. Lasanta, and P. I. Hurtado, Large fluctuations in driven dissipative media, Phys. Rev. Lett. **107**, 140601 (2011).
[77] P. I. Hurtado and P. L. Garrido, Spontaneous symmetry breaking at the fluctuating level, Phys. Rev. Lett. **107**, 180601 (2011).
[78] C. P. Espigares, P. L. Garrido, and P. I. Hurtado, Dynamical phase transition for current statistics in a simple driven diffusive system, Phys. Rev. E **87**, 032115 (2013).
[79] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, Macroscopic fluctuation theory, Rev. Mod. Phys. **87**, 593 (2015).
[80] H. Qian, Y.-C. Cheng, and Y.-J. Yang, Kinematic basis of emergent energetics of complex dynamics, Europhys. Lett. **131**, 50002 (2020).
[81] L. Onsager and S. Machlup, Fluctuations and irreversible processes, Phys. Rev. **91**, 1505 (1953).
[82] S. Machlup and L. Onsager, Fluctuations and irreversible process. II. Systems with kinetic energy, Phys. Rev. **91**, 1512 (1953).
[83] P. C. Martin, E. D. Siggia, and H. A. Rose, Statistical dynamics of classical systems, Phys. Rev. A **8**, 423 (1973).
[84] H.-K. Janssen, On a Lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties, Z Physik B **23**, 377 (1976).
[85] R. Bausch, H. K. Janssen, and H. Wagner, Renormalized field theory of critical dynamics, Z. Phys. B: Condens. Matter **24**, 113 (1976).
[86] C. De dominicis, Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques, J. Phys. Colloques **37**, C1-247 (1976).
[87] A. Baule and P. Sollich, Exponential increase of transition rates in metastable systems driven by non-Gaussian noise, Sci. Rep. **13**, 3853 (2023).
[88] Y. Huang, X. Zhou, and J. Duan, Probability flow approach to the onsager–machlup functional for jump-diffusion processes, SIAM J. Appl. Math. **85**, 524 (2025).
[89] A. T. Mohite and H. Rieger, Minimum action principle for entropy production rate of far-from-equilibrium systems, arXiv:2511.00967.
[90] M. Doi, Second quantization representation for classical many-particle system, J. Phys. A: Math. Gen. **9**, 1465 (1976).
[91] M. Doi, Stochastic theory of diffusion-controlled reaction, J. Phys. A: Math. Gen. **9**, 1479 (1976).
[92] L. Peliti, Path integral approach to birth-death processes on a lattice, J. Phys. France **46**, 1469 (1985).
[93] L. Peliti, Renormalisation of fluctuation effects in the A+A to A reaction, J. Phys. A: Math. Gen. **19**, L365 (1986).
[94] M. F. Weber and E. Frey, Master equations and the theory of stochastic path integrals, Rep. Prog. Phys. **80**, 046601 (2017).
[95] P. Salamon and R. S. Berry, Thermodynamic length and dissipated availability, Phys. Rev. Lett. **51**, 1127 (1983).
[96] P. Salamon, J. D. Nulton, and R. S. Berry, Length in statistical thermodynamics, J. Chem. Phys. **82**, 2433 (1985).
[97] F. Schlögl, Thermodynamic metric and stochastic measures, Z. Phys. B: Condens. Matter **59**, 449 (1985).
[98] G. E. Crooks, Measuring thermodynamic length, Phys. Rev. Lett. **99**, 100602 (2007).
[99] K. Yoshimura and S. Ito, Thermodynamic uncertainty relation and thermodynamic speed limit in deterministic chemical reaction networks, Phys. Rev. Lett. **127**, 160601 (2021).
[100] K. Yoshimura and S. Ito, Information geometric inequalities of chemical thermodynamics, Phys. Rev. Res. **3**, 013175 (2021).
[101] D. Loutchko, Y. Sughiyama, and T. J. Kobayashi, Riemannian geometry of optimal driving and thermodynamic length and its application to chemical reaction networks, Phys. Rev. Res. **4**, 043049 (2022).
[102] G. L. Eyink, Action principle in nonequilibrium statistical dynamics, Phys. Rev. E **54**, 3419 (1996).
[103] W. E, W. Ren, and E. Vanden-Eijnden, Minimum action method for the study of rare events, Commun. Pure Appl. Math. **57**, 637 (2004).
[104] M. Heymann and E. Vanden-Eijnden, Pathways of maximum likelihood for rare events in nonequilibrium systems: Application to nucleation in the presence of shear, Phys. Rev. Lett. **100**, 140601 (2008).
[105] E. Vanden-Eijnden and M. Heymann, The geometric minimum action method for computing minimum energy paths, J. Chem. Phys. **128**, 061103 (2008).
[106] P. C. Bressloff, Path-integral methods for analyzing the effects of fluctuations in stochastic hybrid neural networks, J. Math. Neurosci. **5**, 4 (2015).
[107] T. Grafke, T. Schäfer, and E. Vanden-Eijnden, Long term effects of small random perturbations on dynamical systems: Theoretical and computational tools, in *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science*, edited by R. Melnik, R. Makarov,

and J. Belair (Springer New York, New York, NY, 2017) pp. 17–55.

[108] T. Grafke and E. Vanden-Eijnden, Numerical computation of rare events via large deviation theory, Chaos **29**, 063118 (2019).

[109] P. Gagrani and E. Smith, Action functional gradient descent algorithm for estimating escape paths in stochastic chemical reaction networks, Phys. Rev. E **107**, 034305 (2023).

[110] R. Zakine and E. Vanden-Eijnden, Minimum-action method for nonequilibrium phase transitions, Phys. Rev. X **13**, 041044 (2023).

[111] E. Smith, H. B. Smith, and J. L. Andersen, Rules, hypergraphs, and probabilities: The three-level analysis of chemical reaction systems and other stochastic stoichiometric population processes, PLOS Complex Systems **1**, e0000022 (2024).

[112] A. Cherukuri, B. Gharesifard, and J. Cortés, Saddle-point dynamics: Conditions for asymptotic stability of saddle points, SIAM J. Control Optim. **55**, 486 (2017).

[113] T. Lin, C. Jin, and M. Jordan, On gradient descent ascent for nonconvex-concave minimax problems, in *Proceedings of the 37th International Conference on Machine Learning*, edited by H. D. III and A. Singh, Proceedings of Machine Learning Research, Vol. 119 (PMLR, 2020), pp. 6083–6093.

[114] J. Yan, H. Touchette, and G. M. Rotskoff, Learning nonequilibrium control forces to characterize dynamical phase transitions, Phys. Rev. E **105**, 024115 (2022).

[115] D.-K. Kim, Y. Bae, S. Lee, and H. Jeong, Learning entropy production via neural networks, Phys. Rev. Lett. **125**, 140604 (2020).

[116] S. Otsubo, S. Ito, A. Dechant, and T. Sagawa, Estimating entropy production by machine learning of short-time fluctuating currents, Phys. Rev. E **101**, 062106 (2020).

[117] S. Otsubo, S. K. Manikandan, T. Sagawa, and S. Krishnamurthy, Estimating time-dependent entropy production from non-equilibrium trajectories, Commun. Phys. **5**, 11 (2022).

[118] S. A. Horiguchi and T. J. Kobayashi, Optimal control of stochastic reaction networks with entropic control cost and emergence of mode-switching strategies, PRX Life **3**, 033027 (2025).

[119] T. Tottori and T. J. Kobayashi, Theory for optimal estimation and control under resource limitations and its applications to biological information processing and decision-making, Phys. Rev. Res. **7**, 043048 (2025).

[120] A. T. Mohite and H. Rieger, Generalized finite-time optimal control framework in stochastic thermodynamics, arXiv:2511.00974.

[121] We consider systems with all bidirectional transitions with positive transition rates, this is to ensure the absolute continuousness of the forward path probability measure with respect to the conjugate path probability measure.

[122] C. Maes, Frenesy: Time-symmetric dynamical activity in nonequilibria, Phys. Rep. **850**, 1 (2020).

[123] C. Maes, Frenetic bounds on the entropy production, Phys. Rev. Lett. **119**, 160601 (2017).

[124] The equality in Eq. (7) for the transition probability measure is trivially assumed by incorporating the probability normalization factor, $\int \mathbb{D}\{\phi_i^*\}\,\mathbb{D}\{\phi_i\}\mathcal{P}[\{\phi_i^*,\phi_i\}] = 1$. Here, $\mathbb{D}\{\phi_i^*\}$ and $\mathbb{D}\{\phi_i\}$ denote the path integral over all coherent-state eigenvalues. Henceforth, we omit the normalization constant for the transition probability measure; however, its inclusion for probability conservation is presumed.

[125] K. Itakura, J. Ohkubo, and S. ichi Sasa, Two Langevin equations in the Doi–Peliti formalism, J. Phys. A: Math. Theor. **43**, 125001 (2010).

[126] R. L. Stratonovich, On a method of calculating quantum distribution functions, Sov. Phys. Dokl. **2**, 416 (1957).

[127] J. Hubbard, Calculation of partition functions, Phys. Rev. Lett. **3**, 77 (1959).

[128] A. Kolchinsky, A. Dechant, K. Yoshimura, and S. Ito, Information geometry of excess and housekeeping entropy production, arXiv:2206.14599.

[129] A. Kolchinsky, A. Dechant, K. Yoshimura, and S. Ito, Generalized free energy and excess entropy production for active systems, Phys. Rev. Res. **8**, 023025 (2026).

[130] S.-I. Amari, *Information Geometry and Its Applications*, Applied Mathematical Sciences (Springer, Tokyo, 2016).

[131] Due to our *bottom-up approach*, we had identified the backward process on a microscopic level by decomposing the unidirectional transitional currents into its time-antisymmetric (J_γ) and time-symmetric (T_γ) components.

[132] M. O. Vlad and J. Ross, Thermodynamic approach to nonequilibrium chemical fluctuations, J. Chem. Phys. **100**, 7295 (1994).

[133] M. O. Vlad and J. Ross, Random paths and fluctuation–dissipation dynamics for one-variable chemical systems far from equilibrium, J. Chem. Phys. **100**, 7279 (1994).

[134] M. O. Vlad and J. Ross, Fluctuation–dissipation relations for chemical systems far from equilibrium, J. Chem. Phys. **100**, 7268 (1994).

[135] J. Ross and M. O. Vlad, Nonlinear kinetics and new approaches to complex reaction mechanisms, Annu. Rev. Phys. Chem. **50**, 51 (1999).

[136] C. Maes and K. Netočný, Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states, Europhys. Lett. **82**, 30003 (2008).

[137] C. Maes, K. Netočný, and B. Wynants, Dynamical fluctuations for semi-Markov processes, J. Phys. A: Math. Theor. **42**, 365002 (2009).

[138] A. Mielke, M. A. Peletier, and D. R. M. Renger, On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion, Potential Anal. **41**, 1293 (2014).

[139] A. Mielke, R. I. A. Patterson, M. A. Peletier, and D. R. Michiel Renger, Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics, SIAM J. Appl. Math. **77**, 1562 (2017).

[140] M. Kaiser, R. L. Jack, and J. Zimmer, Canonical structure and orthogonality of forces and currents in irreversible Markov chains, J. Stat. Phys. **170**, 1019 (2018).

[141] D. R. M. Renger and J. Zimmer, Orthogonality of fluxes in general nonlinear reaction networks, Discrete Contin. Dyn. Syst. **14**, 205 (2021).

[142] M. A. Peletier, R. Rossi, G. Savaré, and O. Tse, Jump processes as generalized gradient flows, Calc. Var. Partial Differ. Equation **61**, 33 (2022).

[143] T. J. Kobayashi, D. Loutchko, A. Kamimura, and Y. Sughiyama, Hessian geometry of nonequilibrium chemical reaction networks and entropy production decompositions, Phys. Rev. Res. **4**, 033208 (2022).

[144] T. J. Kobayashi, D. Loutchko, A. Kamimura, S. A. Horiguchi, and Y. Sughiyama, Information geometry of dynamics on graphs and hypergraphs, Inf. Geom. **7**, 97 (2024).
[145] M. A. Peletier and A. Schlichting, Cosh gradient systems and tilting, Nonlinear Analysis **231**, 113094 (2023), Variational Models for Discrete Systems.
[146] R. I. A. Patterson, D. R. M. Renger, and U. Sharma, Variational structures beyond gradient flows: A macroscopic fluctuation-theory perspective, J. Stat. Phys. **191**, 18 (2024).
[147] D. R. M. Renger, Macroscopic fluctuation theory versus large-deviation-induced generic, arXiv:2402.04092.
[148] T. J. Kobayashi, D. Loutchko, A. Kamimura, and Y. Sughiyama, Kinetic derivation of the Hessian geometric structure in chemical reaction networks, Phys. Rev. Res. **4**, 033066 (2022).
[149] Y. Sughiyama, D. Loutchko, A. Kamimura, and T. J. Kobayashi, Hessian geometric structure of chemical thermodynamic systems with stoichiometric constraints, Phys. Rev. Res. **4**, 033065 (2022).
[150] D. R. Michiel Renger and U. Sharma, Untangling dissipative and Hamiltonian effects in bulk and boundary-driven systems, Phys. Rev. E **108**, 054123 (2023).
[151] M. H. Duong and J. Zimmer, On decompositions of non-reversible processes, J. Phys.: Conf. Ser. **2514**, 012007 (2023).
[152] T. Mizohata, T. J. Kobayashi, L.-S. Bouchard, and H. Miyahara, Information geometric bound on general chemical reaction networks, Phys. Rev. E **109**, 044308 (2024).
[153] F. M. Dannan, P. Neff, and C. Thiel, On the sum of squared logarithms inequality and related inequalities, arXiv:1411.1290.
[154] Y.-J. Yang and K. A. Dill, A principled basis for nonequilibrium network flows, Nat. Commun. **17**, 4980 (2026).
[155] We use an overbar to denote a scaling transformation, avoiding confusion with scaled time-integrated physical quantities denoted using a tilde.
[156] E. Roldan, I. Neri, R. Chetrite, S. Gupta, S. Pigolotti, F. Juelicher, and K. Sekimoto, Martingales for physicists: A treatise on stochastic thermodynamics and beyond, Adv. Phys. **72**, 1 (2023).
[157] R. Chetrite and S. Gupta, Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale, J. Stat. Phys. **143**, 543 (2011).
[158] I. Neri, E. Roldan, and F. Julicher, Statistics of infima and stopping times of entropy production and applications to active molecular processes, Phys. Rev. X **7**, 011019 (2017).
[159] S. Pigolotti, I. Neri, E. Roldan, and F. Julicher, Generic properties of stochastic entropy production, Phys. Rev. Lett. **119**, 140604 (2017).
[160] R. Chetrite, S. Gupta, I. Neri, and E. Roldan, Martingale theory for housekeeping heat, Europhys. Lett. **124**, 60006 (2019).
[161] H.-M. Chun and J. D. Noh, Universal property of the housekeeping entropy production, Phys. Rev. E **99**, 012136 (2019).
[162] I. Neri, Universal tradeoff relation between speed, uncertainty, and dissipation in nonequilibrium stationary states, SciPost Phys. **12**, 139 (2022).
[163] A. Raghu and I. Neri, Effective affinity for generic currents in Markov processes, J. Stat. Phys. **192**, 50 (2025).
[164] I. Neri, Martingale approach for first-passage problems of time-additive observables in Markov processes, J. Phys. A: Math. Theor. **58**, 145002 (2025).
[165] I. A. Martínez, G. Bisker, J. M. Horowitz, and J. M. R. Parrondo, Inferring broken detailed balance in the absence of observable currents, Nat. Commun. **10**, 3542 (2019).
[166] D. J. Skinner and J. Dunkel, Estimating entropy production from waiting time distributions, Phys. Rev. Lett. **127**, 198101 (2021).
[167] D. J. Skinner and J. Dunkel, Improved bounds on entropy production in living systems, Proc. Natl. Acad. Sci. USA **118**, e2024300118 (2021).
[168] A. Nadal-Rosa and G. Manzano, Thermodynamic inference in molecular motors: A martingale approach Phys. Rev. E **112**, 034114 (2025).
[169] Y.-J. Yang and H. Qian, Unified formalism for entropy production and fluctuation relations, Phys. Rev. E **101**, 022129 (2020).
[170] L. Hong, H. Qian, and L. F. Thompson, Representations and divergences in the space of probability measures and stochastic thermodynamics, J. Comput. Appl. Math. **376**, 112842 (2020).
[171] H. Qian, Y.-C. Cheng, and L. F. Thompson, Ternary representation of stochastic change and the origin of entropy and its fluctuations, arXiv:1902.09536.
[172] D.-Q. Jiang, M. Qian, and F.-X. Zhang, Entropy production fluctuations of finite Markov chains, J. Math. Phys. **44**, 4176 (2003).
[173] I. D. Terlizzi, M. Gironella, D. Herraez-Aguilar, T. Betz, F. Monroy, M. Baiesi, and F. Ritort, Variance sum rule for entropy production, Science **383**, 971 (2024).
[174] K. Tomita and H. Tomita, Irreversible circulation of fluctuation, Prog. Theor. Phys. **51**, 1731 (1974).
[175] I. Di Terlizzi, M. Baiesi, and F. Ritort, Variance sum rule: Proofs and solvable models, New J. Phys. **26**, 063013 (2024).
[176] N. Ohga, S. Ito, and A. Kolchinsky, Thermodynamic bound on the asymmetry of cross-correlations, Phys. Rev. Lett. **131**, 077101 (2023).
[177] N. Shiraishi, Entropy production limits all fluctuation oscillations, Phys. Rev. E **108**, L042103 (2023).
[178] A. Kolchinsky, N. Ohga, and S. Ito, Thermodynamic bound on spectral perturbations, with applications to oscillations and relaxation dynamics, Phys. Rev. Res. **6**, 013082 (2024).
[179] S. Liang and S. Pigolotti, Thermodynamic bounds on time-reversal asymmetry, Phys. Rev. E **108**, L062101 (2023).