Universal trade-off between irreversibility and intrinsic timescale in thermal relaxation with applications to thermodynamic inference

Ruicheng Bao,[1,2,*] Chaoqun Du,[1] Zhiyu Cao,[1] and Zhonghuai Hou[1,†]

[1]*Department of Chemical Physics & Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*

[2]*Department of Physics, Graduate School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received 5 August 2024; revised 25 August 2025; accepted 2 October 2025; published 20 October 2025)

We establish a general lower bound for the entropy production rate (EPR) based on the Kullback-Leibler divergence and the logarithmic-Sobolev constant that characterizes the timescale of relaxation. This bound can be considered as an enhanced second law of thermodynamics. When applied to thermal relaxation, it reveals a universal trade-off relation between the dissipation rate and the intrinsic relaxation timescale. From this relation, a thermodynamic upper bound on the relaxation time between two given states emerges, acting as an inverse speed limit over the entire time region. We also obtain a quantum version of this upper bound, which is always tighter than its classical counterpart, incorporating an additional term due to decoherence. Remarkably, we further demonstrate that the trade-off relation remains valid for any generally non-Markovian coarse-grained relaxation dynamics, highlighting its significant applications in thermodynamic inference. This trade-off relation is a new tool in inferring EPRs in molecular dynamics simulations and practical experiments.

DOI: 10.1103/fmsz-rdbj

I. INTRODUCTION

The past 20 years have seen extraordinary progress in nonequilibrium statistical physics of small systems with non-negligible fluctuations. Significant advances include the celebrated fluctuation theorems [1–12] containing all information of the stochastic entropy production (EP), the speed limit in quantum and classical systems [13–20], some refined versions of the second law of thermodynamics [21–28] and the recently proposed thermodynamic uncertainty relations [29–39]. Thermodynamic irreversibility, typically quantified by EP, is key to most of the important theorems and relations mentioned above. As a central concept in modern thermodynamics, it plays a pivotal role across various fields, including the optimization of heat engines, the design and operation of nanomachines, and the understanding of biological system functions. The thermodynamic irreversibility in transient processes, which are common in nature and inherently out-of-equilibrium, has not been studied as extensively as that in stationary processes [40].

Our main focus here is on a crucial and nontrivial class of transient processes known as thermal relaxation. This fundamental class of physical processes is ubiquitous in the real world and has numerous applications across various fields [41]. Interestingly, thermal relaxation phenomena are complex and varied even under Markov approximations. Typical examples are dynamical phase transitions [41,42], anomalous relaxation like the Mpemba effect [43] and asymmetric relaxation from different directions [40,44,45]. One of the central quantities in thermal relaxation is its timescale of convergence, which has been intensively studied. A well-developed theory on that is the spectral gap theory, which says that the relaxation timescale is typically characterized by the spectral gap of the generator of dynamics in the large time regime. Around the spectral gap, some important frameworks on metastability [46–50] and the Mpemba effect [43,44,51–59] have been established, which mainly focus on the large time limit. However, there are much fewer works concentrating on the entire time region of relaxation processes [60]. In particular, general principles that constrain the behaviors of instantaneous irreversibility [entropy production rate (EPR)] and the relaxation timescale, applicable at any time during relaxation processes, remain to be investigated.

In this study, we propose a general lower bound for irreversibility based on Kullback-Leibler (KL) divergence [61] and logarithmic-Sobolev (LS) constant [62], which is strengthened compared to the standard second law of thermodynamics. The general bound is then applied to thermal relaxation, revealing a trade-off relation between the intrinsic timescale and EPR that is valid throughout the entire relaxation process, not just in the large-time region. A thermodynamic upper bound on the transformation time between any pair of given states during thermal relaxation follows from the trade-off, which we term the inverse speed limit. More importantly, we theoretically and numerically show that our trade-off relation holds even for generally non-Markovian coarse-grained dynamics, significantly broadening the applicability of the relation. It can aid in the design of real-world rapid relaxation processes, which are desirable in numerous situations [56].

A key distinction between our findings and previous results is that we provide an experimentally feasible lower bound for the instantaneous EPR. This contrasts with prior related results, which mainly focus on lower bounds for the EP over a

*Contact author: ruicheng@g.ecc.u-tokyo.ac.jp

†Contact author: hzhlj@ustc.edu.cn

time interval during relaxation [13,19,25], thus representing a different aspect of nonequilibrium phenomena [63]. Additionally, the intrinsic timescale considered here is a property of the underlying dynamics, characterized by the spectral properties of the dynamical generator and independent of the initial and final distributions of the system. In comparison, timescales incorporated in previous findings, such as speed limits, depend on those distributions.

The trade-off relation for coarse-grained dynamics is fundamentally new and serves as a valuable tool for thermodynamic inference, a crucial task in nonequilibrium statistical physics [64]. We apply this coarse-grained trade-off to infer the EPR in relaxation processes of complex systems, where only coarse-grained observations are feasible, typical in experimental settings. This lower bound complements previous thermodynamic inference results, which are primarily limited to stationary processes [65–73]. Although some studies have focused on nonstationary dynamics [74–78], these methods often require substantial trajectory data, complex procedures, and assume Markovianity, making them impractical for non-Markovian coarse-grained dynamics. In contrast, our methodology is applicable to highly coarse-grained dynamics with very few coarse-grained states, without requiring trajectory data, relying solely on the statistics of coarse-grained states, which are easier to obtain in practice. We demonstrate our method with coarse-grained data from molecular-dynamics simulations, where transitions or currents are undetectable and heat dissipation cannot be measured directly. This demonstration highlights the potential applicability of the trade-off relation for inferring EPR in real experiments.

II. A GENERAL LOWER BOUND FOR EPR

We are considering a system with N states coupled to a heat bath with inverse temperature $\beta = 1/(k_B T)$, though the generalization of our results to multiple heat baths is straightforward. The dynamics of the probability of the system being in state i at time t, $p_i(t)$, is described by a master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i(t) = \sum_{j=1}^{N}[k_{ij}(t)p_j(t) - k_{ji}(t)p_i(t)], \tag{1}$$

where $k_{ij}(t)$ denotes the transition rate from state j to state i at time t. The master equation can be rewritten in a more compact matrix form as $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) = \mathcal{L}(t)\boldsymbol{p}(t)$, where $\boldsymbol{p}(t) = [p_1(t), p_2(t), \ldots, p_N(t)]^{\mathrm{T}}$ and $\mathcal{L}_{ij}(t) = k_{ij}(t) - \delta_{ij}\sum_l k_{li}(t)$ is the stochastic matrix (strictly speaking, $\mathcal{L}$ is an operator) at time t. The stochastic matrix changes over time due to external protocols. In this work, we focus on both cases where the detailed balance condition $k_{ij}(t)\pi_j^t = k_{ji}(t)\pi_i^t$ holds for all pairs of i, j at any time t, and when it does not, where π_i^t is the (instantaneous) stationary distribution at time t for state i. We denote $\boldsymbol{\pi}_t = [\pi_1^t, \ldots, \pi_N^t]^{\mathrm{T}}$, in which $\boldsymbol{\pi}_t$ is defined as the stationary state, and it will be reached if the stochastic matrix is frozen at time t. When the detailed balance condition holds, $\boldsymbol{\pi}_t$ will be an (instantaneous) equilibrium state $\boldsymbol{p}_t^{\mathrm{eq}}$ whose entries are $p_{t,i}^{\mathrm{eq}} = e^{-\beta E_i(t)}/Z$, with $E_i(t)$ being the instantaneous energy of state i at time t, and Z is the normalization constant. The KL divergence, which quantifies the difference between two probability distributions, is defined as $D[\boldsymbol{p}^a||\boldsymbol{p}^b] \equiv \sum_i p_i^a \ln(p_i^a/p_i^b)$. For any continuous-time Markov processes obeying the master equation (1) with an instantaneous equilibrium distribution $\boldsymbol{p}_t^{\mathrm{eq}}$ at time t, we demonstrate that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}D\big[\boldsymbol{p}(\tau)||\boldsymbol{p}_t^{\mathrm{eq}}\big]|_{\tau=t} \leqslant -4\lambda_{\mathrm{LS}}(t)D\big[\boldsymbol{p}(t)||\boldsymbol{p}_t^{\mathrm{eq}}\big], \tag{2}$$

where $\lambda_{\mathrm{LS}}(t)$ is a positive real number determined by $\mathcal{L}(t)$. Further, without a detailed balance condition, we still have a similar inequality $\frac{\mathrm{d}}{\mathrm{d}\tau}D[\boldsymbol{p}(\tau)||\boldsymbol{\pi}_t]|_{\tau=t} \leqslant -2\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{\pi}_t]$, where a factor $1/2$ is multiplied on the right-hand side. Before proceeding, we denote $\langle f, g\rangle_\pi \equiv \sum_i f g_i^\dagger \pi_i$ the inner product induced by the stationary distribution $\boldsymbol{\pi}$ (may be instantaneous).

The positive real number $\lambda_{\mathrm{LS}}(t)$ in Eq. (2) is the LS constant [62] corresponding to the stochastic matrix $\mathcal{L}(t)$, whose definition is

$$\lambda_{\mathrm{LS}} \equiv \inf_{\mathrm{Ent}(f)\neq 0} \frac{\mathrm{Re}\langle -\mathcal{L}f, f\rangle_\pi}{\mathrm{Ent}(f)}, \tag{3}$$

where $\mathrm{Ent}(f)$ is an entropy-like quantity defined as $\mathrm{Ent}(f) = \sum_{i=1}^{N}|f_i|^2 \ln(\frac{|f_i|^2}{\langle f,f\rangle_\pi})\pi_i$ and f is any function in the state space of the system.

According to the stochastic thermodynamics, the average EP rate $\dot{\sigma}(t)$ at time t in this system is (k_B is set to be 1) [79]

$$\dot{\sigma}(t) = \sum_{i,j} k_{ij}(t)p_j(t)\ln\frac{k_{ij}(t)p_j(t)}{k_{ji}(t)p_i(t)}. \tag{4}$$

If the stochastic matrix satisfies the detailed balance condition (i.e., the Markov process in focus is reversible), $\dot{\sigma}(t)$ is related to the KL divergence between the current distribution and the instantaneous equilibrium distribution $\boldsymbol{p}_t^{\mathrm{eq}}$ as $\dot{\sigma}(t) = -\partial_\tau D[\boldsymbol{p}(\tau)||\boldsymbol{p}_t^{\mathrm{eq}}]|_{\tau=t}$ [80,81]. Combining this with Eq. (2) leads to

$$\dot{\sigma}(t) \geqslant 4\lambda_{\mathrm{LS}}(t)D\big[\boldsymbol{p}(t)\big|\big|\boldsymbol{p}_t^{\mathrm{eq}}\big]. \tag{5}$$

This general lower bound for the EPR at any given time is our first main result. The bound will always be positive unless the system is in an equilibrium state, since $\lambda_{\mathrm{LS}}(t)$ is always positive [62], which makes it generally stronger than the conventional second law. It also shows that the possible EPR increases as the system deviates further from the instantaneous equilibrium state.

In the absence of the detailed balance condition, a similar lower bound for the nonadiabatic EPR (also named as Hatano-Sasa EP) can be obtained as

$$\dot{\sigma}^{\mathrm{na}}(t) \geqslant 2\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{\pi}_t], \tag{6}$$

where the definition of $\dot{\sigma}^{\mathrm{na}}(t)$ is given by $\dot{\sigma}^{\mathrm{na}}(t) = -\sum_i \dot{p}_i(t)\ln\frac{p_i(t)}{\pi_i^t}$ and the relation $\dot{\sigma}^{\mathrm{na}}(t) = -\partial_\tau D[\boldsymbol{p}(\tau)||\boldsymbol{\pi}_\tau]|_{\tau=t}$ has been used [82]. Since the total EP rate satisfies $\dot{\sigma}(t) \geqslant \dot{\sigma}^{\mathrm{na}}(t)$, the bound can still serve as a stronger second law, i.e., $\dot{\sigma}(t) \geqslant \dot{\sigma}^{\mathrm{na}}(t) \geqslant 2\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{\pi}_t] \geqslant 0$.

In what follows, we focus on an important application of our lower bounds in thermal relaxation processes, where the stochastic matrix $\mathcal{L}$ becomes time-independent and λ_{LS} is a constant uniquely determined by $\mathcal{L}$. Nonetheless, we

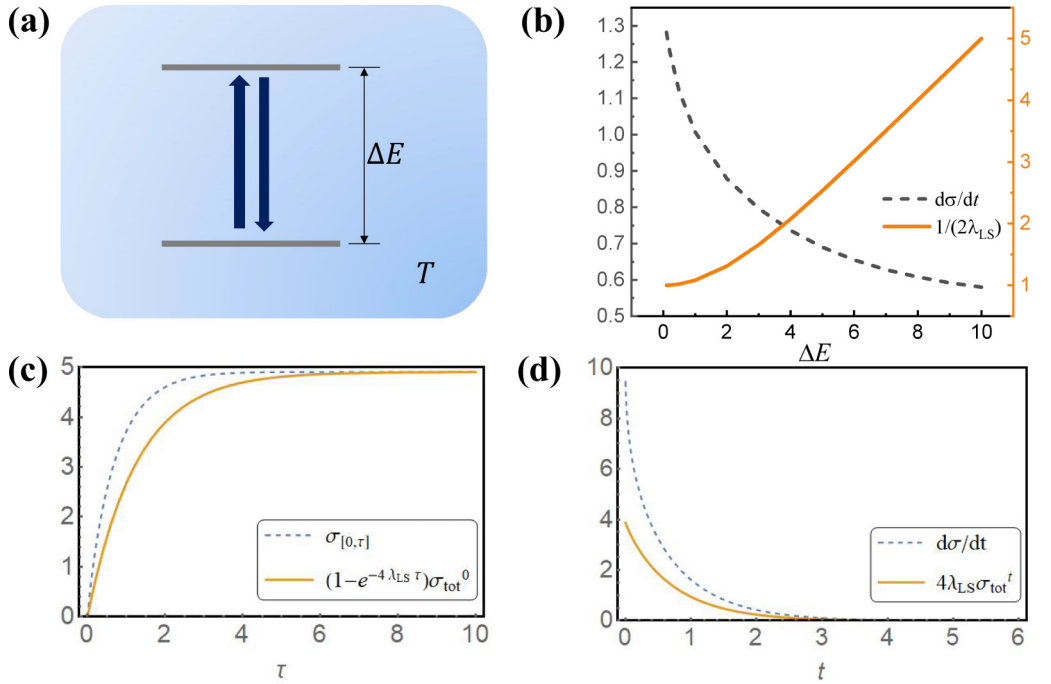

FIG. 1. Illustration of the trade-off relation and the lower bound for entropy production in a two-state mode. Here, we set $\beta = 1/(k_B T) = 1$. (a) The two-state model coupled to a heat reservoir with temperature T. (b) The trade-off relation between the EPR $\dot{\sigma}(t)$ at $t = 0$ and the relaxation timescale $1/(2\lambda_{\rm LS})$ for relaxation processes with different ΔE, where the distance to equilibrium $\sigma^t_{\rm tot} \equiv D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]$ is fixed to be 0.5. (c) and (d) We demonstrate Eqs. (8) and (9) for this model, in which $\Delta E = 5.0$ and the initial distribution is chosen to be $(p_u, p_d) = (0.99, 0.01)$.

also present another application of Eq. (5) in a system with time-dependent dynamics, where the stochastic matrix is periodically switching [83,84]. In that example, we show that our lower bound can help in recovering part of the "hidden" EPR [85] of an effective equilibrium state.

III. TRADE-OFF RELATION FOR THERMAL RELAXATION

In thermal relaxation, $\lambda_{\rm LS}$ is related to the intrinsic timescale $\tau^0_{\rm rel} \equiv \inf_\tau\{\sup_i \|\frac{\boldsymbol{p}(\tau)|p_j(0)=\delta_{ij})}{\boldsymbol{\pi}} - \mathbb{I}\|_2 \leqslant 1/e\}$, which characterizes the slowest (dominant) mode of relaxation and is independent of the initial and final distributions. $\tau^0_{\rm rel}$ satisfies $\tau^0_{\rm rel} \geqslant 1/(2\lambda_{\rm LS})$.

Operationally, $\lambda_{\rm LS}$ can be measured via another expression:

$$\frac{1}{2\lambda_{\rm LS}} \equiv \tau^1_{\rm rel} \overset{t\gg 1}{\simeq} \frac{-t}{2\ln D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]}. \tag{7}$$

Here, $\tau^1_{\rm rel}$ is the measurable relaxation timescale. This may require more statistical data for precise measurement but is more adaptable to non-Markovian coarse-grained dynamics. For both definitions, Eq. (5) leads to

$$\sigma_{[0,\tau]} \geqslant \left(1 - e^{-2\tau/\tau^{0,1}_{\rm rel}}\right) D[\boldsymbol{p}(0)||\boldsymbol{p}^{\rm eq}], \quad \tau \geqslant 0, \tag{8}$$

and

$$\dot{\sigma}(t)\tau^{0,1}_{\rm rel} \geqslant 2D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]. \tag{9}$$

Here, $\boldsymbol{p}(t)$ is the distribution at time t, $\boldsymbol{p}^{\rm eq}$ is the equilibrium distribution of $\mathcal{L}$ and the entropy production reads $\sigma_{[0,\tau]} = \int_0^\tau \dot{\sigma}(t)dt = D[\boldsymbol{p}(0)||\boldsymbol{p}^{\rm eq}] - D[\boldsymbol{p}(\tau)||\boldsymbol{p}^{\rm eq}]$. Equations (8) and (9) reveal close connections between EP (rate) and intrinsic relaxation timescale, which is our second main result. These two inequalities hold for any $t > 0$ ($\tau > 0$), and they are saturated at the large time limit t, $\tau \to \infty$. Equation (8) also saturates at the small time limit $\tau \to 0$ when both sides equal zero. Equation (9) rigorously shows that the minimal possible dissipation rate in thermal relaxation increases as the distance from equilibrium grows, a fact that was not explicitly known before. Previously, it was only proven that the accumulated EP from time 0 to t in an irreversible process can be lower bounded by the KL divergence from the state at t to the equilibrium state [63,86,87].

To illustrate the results, we take a two-state model, which may be used to model a single spin or a qubit, as an example. As shown in Fig. 1(a), the model system is comprised of an up state u with energy E_u and an down state d with energy E_d, and it is coupled to a heat bath with temperature T. The energy difference between two states is $\Delta E = E_u - E_d > 0$. The transition rates from u to d and from d to u are given by $k_{u\to d} = e^{\beta\Delta E}/(1 + e^{\beta\Delta E})$ and $k_{d\to u} = 1/(1 + e^{\beta\Delta E})$, respectively. Under this setting, the stationary distribution will be an equilibrium one $[p_u, p_d]^{\rm T} = [1/(1 + e^{\beta\Delta E}), e^{\beta\Delta E}/(1 + e^{\beta\Delta E})]^{\rm T}$. The LS constant in this case can be exactly computed as [62] $\lambda_{\rm LS} = \frac{\tanh(\frac{\beta\Delta E}{2})}{\beta\Delta E}$. In Fig. 1(b), a trade-off relation between $\tau_{\rm rel} = 1/(2\lambda_{\rm LS})$ and $\dot{\sigma}(t)$ is demonstrated when $D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]$ is fixed. Figures 1(c) and 1(d) shows that two relations (8) and (9) are valid for any time.Additionally, we obtain another bound related to the spectral gap λ_g of $\mathcal{L}$ as

$$\dot{\sigma}(t) \geqslant 4C\lambda_g D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}], \tag{10}$$

where $C = (1-2\pi_\star)/\ln(1-\pi_\star/\pi_\star)$. The relation between EP and spectral gap is of broad interest [88–91]. Equation (C6) is a complementary relation to previous results on stationary EP.

We further show generalizations of the trade-off relation to discrete-time Markov processes and continuous-space Markov processes.

IV. INVERSE SPEED LIMIT

A corollary of Eq. (9) is an inverse speed limit:

$$\tau \geqslant \frac{1}{4\lambda_{\rm LS}} \ln\left\{\frac{D[\boldsymbol{p}(0)||\boldsymbol{p}^{\rm eq}]}{D[\boldsymbol{p}(0)||\boldsymbol{p}^{\rm eq}] - \sigma_{[0,\tau]}}\right\}, \tag{11}$$

which gives the upper bound for the time τ of the relaxation from an initial distribution $\boldsymbol{p}(0)$ to a target distribution $\boldsymbol{p}(\tau)$. Here, the EP $\sigma_{[0,\tau]}$ for the state transformation from $\boldsymbol{p}(0)$ to $\boldsymbol{p}(\tau)$ can be interpreted as the distance between these two states, as it is monotonic in time during relaxation. Thus, the maximal time that the system takes to relax through such a distance $\sigma_{[0,\tau]}$ is given by the inverse speed limit (11). The system should initially be farther from equilibrium so that the transformation time τ can be shorter. The upper bound still holds when the detailed balance condition is not satisfied. The only difference is that the EP should be replaced with the nonadiabatic EP $\sigma^{\rm na}$ and a factor $1/2$ should be multiplied on the right.

The above relation (11) can be generalized to open quantum systems described by the Lindblad master equation. The population in open quantum system is defined as $P_n(t) \equiv \langle n|\rho_t|n\rangle$, where ρ_t is the density matrix and $|n\rangle$ is the nth energy eigenstate. Then, one can show that an upper bound on the transformation time τ that is tighter than its classical counterpart is given by

$$\tau \leqslant \frac{1}{4\lambda_{\rm LS}} \ln\left\{\frac{D[\boldsymbol{P}(0)||\boldsymbol{P}_{\boldsymbol{\beta}}]}{D[\boldsymbol{P}(0)||\boldsymbol{P}_{\boldsymbol{\beta}}] - (\sigma_{[0,\tau]} - \Delta A)}\right\}, \tag{12}$$

with $\boldsymbol{P}(0)$ being the population vector at initial time, P_β being the population vector for Gibbs state, $\Delta A \equiv A(0) - A(\tau) \geqslant 0$, and $A(t)$ being the asymmetry defined as $A(t) := D(\rho_t||\rho_t^{\rm d})$. Here, $\rho_t^{\rm d}$ is the fully decohered version of ρ_t. This result implies that relaxation may be accelerated by quantum coherence.

V. TRADE-OFF RELATION FOR ARBITRARY COARSE-GRAINED DYNAMICS AND THERMODYNAMIC INFERENCE

Measuring the EPR at the coarse-grained level is challenging due to experimental resolution limitations and the large amount of data needed for convergence. Remarkably, our trade-off relation can be generalized to arbitrarily coarse-grained relaxation dynamics, allowing it to be further applied to estimate EP when only coarse-grained observations are feasible. The trade-off relation for coarse-grained dynamics reads

$$\dot\sigma(t) \geqslant \frac{2}{\tau_{\rm rel}^{1,\rm CG}} D[\mathcal{P}(t)||\mathcal{P}^{\rm eq}], \tag{13}$$

where $\mathcal{P}(t)$ and $\mathcal{P}^{\rm eq}$ are the probability distributions for coarse-grained states, and $\tau_{\rm rel}^{1,\rm CG} \overset{t\gg 1}{\simeq} \frac{-t}{2\ln D[\mathcal{P}(t)||\mathcal{P}^{\rm eq}]}$ is the measurable relaxation timescale measured at the coarse-grained level (another definition $\tau_{\rm rel}^{0,\rm CG} \equiv \inf_\tau\{\sup_i \|\frac{\mathcal{P}[|\mathcal{P}_j(0)=\delta_{ij}]}{\mathcal{P}^{\rm eq}} - \mathbb{I}\|_2 \leqslant 1/e\}$ works when there is timescale separation). The lower bound (13) is our third main result. Physically, we have that $\tau_{\rm rel}^{\rm CG} \leqslant \tau_{\rm rel}$ because equilibrium of microscopic dynamics implies the convergence of macroscopic coarse-grained dynamics, but not vice versa. These inequalities must hold when $\tau_{\rm rel}^{\rm CG} \sim \tau_{\rm rel}$, which is a criterion of good coarse-graining [92]. The coarse-grained version of Eq. (6) similarly holds. Note that even when there is timescale separation and $\tau_{\rm rel}^{\rm CG} \sim \tau_{\rm rel}$, the resulting coarse-grained dynamics can still be non-Markovian [93]. We also provide theoretical justification for the validity of the bound in general cases, where $\tau_{\rm rel}^{\rm CG}$ can be much smaller than $\tau_{\rm rel}$. We prove a stronger lower bound without assuming a timescale separation and argue that it gives the desired experimentally feasible bound. To our knowledge, there are no similar results like (13) that can infer the EPR in arbitrarily coarse-grained relaxation dynamics without knowing the model details.

We use a system consisting of many interacting Brownian particles under an external harmonic field (which may be produced using an optical trap) as an example to illustrate the power of our bound [Fig. 2(a)]. This example illustrates its applicability in molecular-dynamics simulations and highlights its potential for practical experimental applications. Even if we know the details of the dynamics, we still need much more data to measure the EPR (initial positions of each particle) without our method. In contrast, our method only requires very coarse-grained data to provide an estimation, with no prior knowledge of the model details. We define a coarse-grained state as the state where a randomly picked particle from the system is in a given spatial region. Then the coarse-grained distribution $\mathcal{P}(t)$ becomes the spatial distribution of particle number density in different regions of the space, which is experimentally feasible. For instance, if the total space is divided into two regions A and B, $\mathcal{P}(t) = (\langle n_A\rangle/n, \langle n_B\rangle/n)$, where $\langle n_i\rangle$ $(i = A, B)$ is the average particle number in region i, and n is the total particle number. We remark that the novel coarse-grained mapping employed here differs from the conventional many-to-one mappings in previous literature. We refer to this experimentally beneficial mapping as random coarse-graining (see [94] for more details). The true EPR and the lower bound are shown in Fig. 2(b). Our bound could reproduce over 20% of the real EPR, which is significant considering that the data we use to obtain the bound are very coarse-grained (four coarse-grained states, compared to the state space spanned by 100 Brownian particles in continuous space). As the degree of coarse-graining decreases, the lower bound will be closer to the real value, according to the inverse scaling law by Yu and Tu [95]. Note that this example lacks measurable transitions, currents, or trajectories, which precludes the application of any existing thermodynamic inference methods.

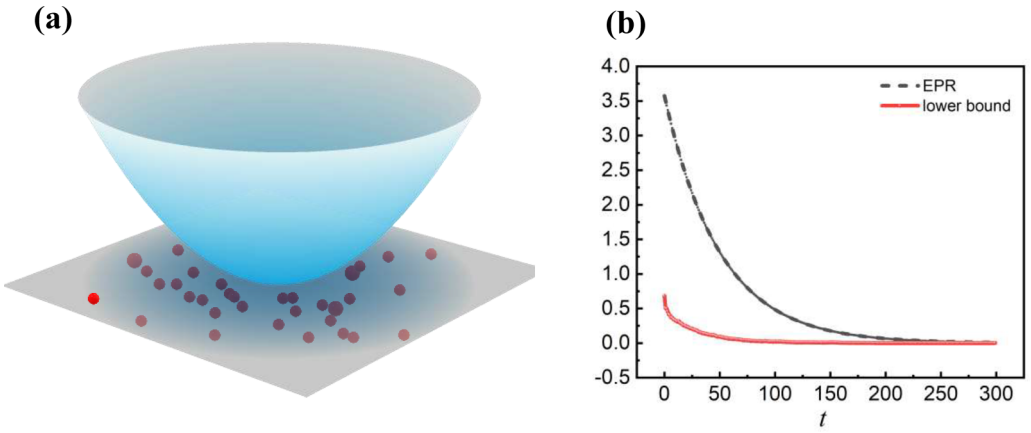

FIG. 2. Application of the coarse-grained trade-off relation to molecular-dynamics simulation. (a) The two-dimensional interacting Brownian particles model. The total particle number is chosen to be 100. The interacting potential is the spring potential, with a strength of $\kappa = 0.01$. The stiffness of the external harmonic potential field is $k = 0.1$. The initial distribution is such that every particle is in one of the four spatial regions divided artificially. (b) The comparison between the true EPR (EPR, the dashed black line) obtained from an approximate analytical expression, and our lower bound (the red line).

VI. DISCUSSION

In this work, we propose a general lower bound for EP related to the LS constant and KL divergence. We utilize it to identify a trade-off between intrinsic timescale and EPR in thermal relaxation. A consequence of this trade-off is an inverse speed limit for transforming states during thermal relaxation, providing a thermodynamic upper bound for the transformation time. It indicates that a system is unlikely to remain in a metastable state beyond a certain time threshold. The inverse speed limit can be extended to open quantum systems, where an additional contribution from decoherence emerges.

Our trade-off relation strikingly holds for arbitrary coarse-grained dynamics, allowing it to be applied to estimate entropy production irrespective of model details in molecular dynamics and real experiments. It remains effective even when there are only two coarse-grained states in relaxation dynamics and no observable current, which stands in sharp contrast to previous results. Our current study opens up several avenues for future research. Interestingly, similar trade-off relations exist for steady state dissipation rate, a counterintuitive finding that warrants further exploration in a forthcoming paper. Another promising direction would be to investigate these trade-off relations in nonlinear chemical reaction networks. Additionally, while our focus has been on the dominant timescale, it would be valuable to analyze the role of other timescales, aside from the dominant one.

ACKNOWLEDGMENTS

R.B. is grateful to Guangyi Zou, Tan Van Vu, and Shiling Liang for useful discussions. Part of this work was finished by R.B. based on the discussion during his participation in the long-term workshop "Frontiers in Non-equilibrium Physics 2024 Workshop" (YITP-T-24-01). R.B. would like to acknowledge the warm hospitality during his stay in the YITP. This work is supported by MOST (2022YFA1303100), NSFC (32090044).

R.B. conceived the research, performed the mathematical derivations, designed the numerical simulations, and wrote the paper. C.D. performed the numerical simulations. Z.C. contributed to the numerical simulations. Z.H. supervised the research.

APPENDIX A: DETAILED DERIVATION OF EQ. (2) IN THE MAIN TEXT

Without loss of generality, we assume transition rates k_{ij} satisfy the normalization condition $\sum_i k_{ij} = 1$ throughout this section (here, we temporarily let k_{ii} be the escape rate from i, in contrast with the main text). Releasing the constraint, the only difference is a multiplicative factor $\sum_i k_{ij}$, which will not affect the derivations here. Under the condition, the stochastic matrix $\mathcal{L}(t)$ in the main text can be written as $K(t) - \mathbb{I}$, where $K_{ij}(t) = k_{ij}(t)$ and $\mathbb{I}$ is the identity matrix. Then for any function f, the operator K satisfies $(Kf)_i = \sum_j k_{ij} f_j$. Recall that the inner product induced by the stationary distribution is defined as

$$\langle f, g\rangle_\pi = \sum_i f_i g_i^\dagger \pi_i. \tag{A1}$$

Based on this inner product, one can further define an adjoint operator $\mathcal{L}^\star$ of $\mathcal{L}$ as $\langle f, \mathcal{L}g\rangle_\pi = \langle \mathcal{L}^\star f, g\rangle_\pi$ for any function f and g. Likewise, another adjoint operator $K^\star$ of K is given by $\langle f, Kg\rangle_\pi = \langle K^\star f, g\rangle_\pi$. Consequently, one can readily check that the operator $K^\star$ satisfies $(K^\star f)_j = \sum_i k_{ij} f_i$. These relations will be useful in the derivations below. For more details, see Ref. [96]. Moreover, one can define the LS constant with respect to $\mathcal{L}_s \equiv (\mathcal{L} + \mathcal{L}^\star)/2$, the symmetrized version of $\mathcal{L}$, as

$$\begin{aligned}\lambda_{\mathrm{LS}} &= \inf_{\mathrm{Ent}(f)\neq 0} \frac{\mathrm{Re}\langle -\mathcal{L}f, f\rangle_\pi}{\mathrm{Ent}(f)} \\ &= \inf_{\mathrm{Ent}(f)\neq 0} \frac{\langle -\mathcal{L}_s f, f\rangle_\pi}{\mathrm{Ent}(f)}.\end{aligned} \tag{A2}$$

Note that $\mathcal{L}^\star = \mathcal{L}$ so that $\mathcal{L}_s = \mathcal{L}$ when the detailed balance condition holds. We drop t dependence in the following for notation's brevity.

Lemma A1.

$$\mathrm{Re}\langle -\mathcal{L}^\star f, f\rangle_\pi = \frac{1}{2}\sum_{i,j}|f_i - f_j|^2 k_{ij}\pi_j, \tag{A3}$$

where $|A| \equiv \sqrt{AA^\dagger}$.

Proof. Notice that $\mathrm{Re}\langle -\mathcal{L}^\star f, f\rangle_\pi = \mathrm{Re}\langle(\mathbb{I} - K^\star)f, f\rangle_\pi = \langle f, f\rangle_\pi - \mathrm{Re}\langle K^\star f, f\rangle_\pi$, and the right-hand side of Eq. (A3) can be rewritten as

$$\begin{aligned}
&\frac{1}{2}\sum_{i,j}|f_i - f_j|^2 k_{ij}\pi_j \\
&= \frac{1}{2}\sum_{i,j}[|f_i|^2 + |f_j|^2 - 2\mathrm{Re}(f_i f_j^\dagger)]k_{ij}\pi_j &\text{(A4)}\\
&= \frac{\sum_i |f_i|^2 \sum_j k_{ij}\pi_j + \sum_j |f_j|^2\pi_j \sum_i k_{ij}}{2} \\
&\quad -\mathrm{Re}\sum_{i,j} k_{ij} f_i f_j^\dagger \pi_j &\text{(A5)}\\
&= \frac{\sum_i |f_i|^2\pi_i + \sum_j |f_j|^2\pi_j}{2} - \mathrm{Re}\sum_j (K^\star f)_j f_j^\dagger \pi_j &\text{(A6)}\\
&= \langle f, f\rangle_\pi - \mathrm{Re}\langle K^\star f, f\rangle_\pi, &\text{(A7)}
\end{aligned}$$

where in the third line, the identities $\sum_j k_{ij}\pi_j = \pi_i$ and $\sum_i k_{ij} = 1$ have been used. Therefore, $\mathrm{Re}\langle -\mathcal{L}^\star f, f\rangle_\pi = \langle f, f\rangle_\pi - \mathrm{Re}\langle K^\star f, f\rangle_\pi = \frac{1}{2}\sum_{i,j}|f_i - f_j|^2 k_{ij}\pi_j$.

Further, with the detailed balance condition $k_{ij}\pi_j = k_{ji}\pi_i$ holding, one can similarly show that

$$\mathrm{Re}\langle -\mathcal{L} f, g\rangle_\pi = \frac{1}{2}\sum_{i,j}(f_i - f_j)(g_i - g_j)k_{ij}\pi_j. \tag{A8}$$

Note that f can be a complex function and $f_i^\dagger$ denote the complex conjugate of f_i.

Lemma A2. For a system with detailed balance condition ($\boldsymbol{\pi} = \boldsymbol{p}^{\mathrm{eq}}$), any function f in the state space of the system satisfies

$$\langle -\mathcal{L}^\star \ln f, f\rangle_\pi \geqslant 4\langle -\mathcal{L}^\star\sqrt{f}, \sqrt{f}\rangle_\pi. \tag{A9}$$

Additionally, in the absence of detailed balance condition, a weaker inequality

$$\langle -\mathcal{L}^\star \ln f, f\rangle_\pi \geqslant 2\langle -\mathcal{L}^\star\sqrt{f}, \sqrt{f}\rangle_\pi \tag{A10}$$

holds.

Proof: For any $a, b > 0$,

$$\begin{aligned}
\left(\frac{\sqrt{a} - \sqrt{b}}{a - b}\right)^2 &= \left[\frac{1}{2(a-b)}\int_b^a x^{-\frac{1}{2}}\mathrm{d}x\right]^2 \\
&\leqslant \left[\int_b^a \frac{1}{4(a-b)^2}\mathrm{d}x\right]\left(\int_b^a x^{-1}\mathrm{d}x\right) \\
&= \frac{1}{4}\frac{\ln a - \ln b}{a - b} &\text{(A11)}
\end{aligned}$$

thus the inequality below is fulfilled:

$$(a - b)[\ln a - \ln b] \geqslant 4(\sqrt{a} - \sqrt{b})^2. \tag{A12}$$

Then using Lemma A1 and Eq. (A8), the inequality $\langle -\mathcal{L}^\star \ln f, f\rangle_\pi \geqslant 4\langle -\mathcal{L}^\star\sqrt{f}, \sqrt{f}\rangle_\pi$ is immediately derived. For any $a, b > 0$, there is another inequality

$$\ln a^2 - \ln b^2 \leqslant \frac{2(a - b)}{b} \tag{A13}$$

due to the concavity of the function $\ln x^2$. Multiplying both sides by b^2 leads to

$$b^2(\ln a^2 - \ln b^2) \leqslant 2b(a - b). \tag{A14}$$

Then letting $f_i = b^2$ and $(K \ln f)_i = \ln a^2$, one obtains

$$\begin{aligned}
f_i[(K^\star - \mathbb{I})\ln f]_i &\leqslant 2\sqrt{f_i}(\sqrt{e^{K^\star \ln f}} - \sqrt{f})_i \\
&\leqslant 2\sqrt{f_i}[(K^\star - \mathbb{I})\sqrt{f}]_i, &\text{(A15)}
\end{aligned}$$

where the inequality

$$(\sqrt{e^{K^\star \ln f}})_i = \sqrt{e^{\sum_j k_{ji}(\ln f)_j}} \leqslant \sum_j k_{ji}(\sqrt{e^{\ln f_j}}) = K^\star\sqrt{f}$$

has been used (the inequality is from the convexity of the function $\sqrt{e^x}$ and the Jensen inequality). Notice that $-\mathcal{L}^\star = \mathbb{I} - K^\star$, thus Eq. (A15) is equal to

$$[(-\mathcal{L}^\star)\ln f]_i f_i\pi_i \geqslant 2[(-\mathcal{L}^\star)\sqrt{f}]_i\sqrt{f_i}\pi_i, \tag{A16}$$

which directly yields $\langle -\mathcal{L}^\star \ln f, f\rangle_\pi \geqslant 2\langle -\mathcal{L}^\star\sqrt{f}, \sqrt{f}\rangle_\pi$.

Proof of the Eq. (2) in the main text. Equipped with Lemmas A1 and A2, we can prove the inequality (3) (with detailed balance) as follows:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}D[\boldsymbol{p}(t)||\boldsymbol{\pi}_\tau]|_{\tau=t} \\
&= \sum_i \left[1 + \ln\frac{p_i(t)}{\pi_i^t}\right]\dot{p}_i(t) &\text{(A17)}\\
&= \sum_i \ln\frac{p_i(t)}{\pi_i^t}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{p_i(t)}{\pi_i^t}\right]\pi_i &\text{(A18)}\\
&= -\sum_i \ln\frac{p_i(t)}{\pi_i^t}\left[-\mathcal{L}\frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}\right]_i\pi_i^t &\text{(A19)}\\
&= -\left\langle \ln\frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}, -\mathcal{L}\frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}\right\rangle_\pi &\text{(A20)}\\
&\equiv -\langle \ln f(t), -\mathcal{L}f(t)\rangle_\pi &\text{(A21)}\\
&= -\langle -\mathcal{L}^\star \ln f(t), f(t)\rangle_\pi &\text{(A22)}\\
&\leqslant -4\langle -\mathcal{L}^\star\sqrt{f(t)}, \sqrt{f(t)}\rangle_\pi &\text{(A23)}\\
&= -4\mathrm{Re}\langle -\mathcal{L}\sqrt{f(t)}, \sqrt{f(t)}\rangle_\pi \leqslant -4\lambda_{\mathrm{LS}}(t)\mathrm{Ent}(\sqrt{f(t)}) &\text{(A24)}\\
&= -4\lambda_{\mathrm{LS}}(t)\sum_{i=1}^N \left|\sqrt{\frac{p_i(t)}{\pi_i^t}}\right|^2 \ln\left(\frac{\left|\sqrt{\frac{p_i(t)}{\pi_i}}\right|^2}{\left\langle\sqrt{\frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}}, \sqrt{\frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}}\right\rangle_\pi}\right)\pi_i^t &\text{(A25)}\\
&= -4\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{\pi}_t]. &\text{(A26)}
\end{aligned}$$

It should be noted that, since $f(t) = \frac{\boldsymbol{p}(t)}{\boldsymbol{\pi}_t}$ is a real function and $\langle -\mathcal{L}^\star f, f\rangle_\pi = \langle f, -\mathcal{L}f\rangle_\pi = \langle -\mathcal{L}f, f\rangle_\pi^\dagger$, we get

that $\langle -\mathcal{L}^\star f, f\rangle_\pi = \mathrm{Re}\langle -\mathcal{L}^\star f, f\rangle_\pi = \mathrm{Re}\langle -\mathcal{L} f, f\rangle_\pi$, which has been used in the third to last line. Without detailed balance, Eq. (A23) should be substituted with $-2\langle -\mathcal{L}^\star\sqrt{f(t)}, \sqrt{f(t)}\rangle_\pi$, where the only difference is a multiplicative constant $1/2$.

The discussions above can be naturally generalized to the system coupled to multiple heat baths, in which the stochastic matrix consists of contributions from each independent baths as $\mathcal{L}(t) = \sum_\nu \mathcal{L}^\nu(t)$, with $\mathcal{L}^\nu(t)$ being the stochastic matrix related to the νth bath. With multiple heat baths, the nonadiabatic entropy production rate can still be associated with the KL divergence as

$$\dot\sigma^{\mathrm{na}}(t) = -\sum_i \dot p_i(t)\ln\frac{p_i(t)}{\pi_i^t} = -\frac{\mathrm{d}}{\mathrm{d}t}D[\boldsymbol{p}(t)||\boldsymbol{\pi}_\tau]|_{\tau=t}. \tag{A27}$$

Notably, $\dot\sigma^{\mathrm{na}}(t)$ is only a function of the coarse-grained transition rates $k_{ij}(t) = \sum_\nu k_{ij}^\nu(t)$, which is not pertinent to the individual contribution from the νth heat bath. Therefore, the general bound (6) in the main text can be directly generalized to the system coupled to multiple heat baths as

$$\dot\sigma(t) \geqslant \dot\sigma^{\mathrm{na}}(t) \geqslant 2\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{\pi}_t]. \tag{A28}$$

As mentioned in the main text, the total EPR can be decomposed into two parts: one part is the nonadiabatic entropy production, and another part (housekeeping or adiabatic EP rate) reads

$$\dot\sigma^{\mathrm{hs}}(t) = \sum_\nu\sum_{i,j} k_{ij}^\nu(t)p_j(t)\ln\frac{k_{ij}^\nu(t)\pi_j^t}{k_{ji}^\nu(t)\pi_i^t}. \tag{A29}$$

It can be seen from this expression that only if the transitions induced by every heat bath all satisfy the detailed balance condition, i.e., $k_{ij}^\nu(t)\pi_j^t = k_{ji}^\nu(t)\pi_i^t$ for any ν, will the housekeeping part vanish (so that $\boldsymbol{\pi}_t = \boldsymbol{p}_t^{\mathrm{eq}}$). In this case, $\dot\sigma(t) = \dot\sigma^{\mathrm{na}}(t) \geqslant 4\lambda_{\mathrm{LS}}(t)D[\boldsymbol{p}(t)||\boldsymbol{p}_t^{\mathrm{eq}}]$.

APPENDIX B: AN APPLICATION OF EQ. (3) TO TIME-DEPENDENT DYNAMICS

In this Appendix, we are interested in an example considered in Refs. [83,84], where the transition matrix is under periodic oscillations. This setting has actual applications in biological and chemical systems. As a result, the dynamics is governed by a time-dependent stochastic matrix: $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) = \mathcal{L}(t)\boldsymbol{p}(t)$, where

$$\mathcal{L}(t) = \begin{cases}\mathcal{L}_1 & t\in[2n\tau,(2n+1)\tau]\\ \mathcal{L}_2 & t\in[(2n+1)\tau,(2n+2)\tau]\end{cases},\quad n\in\mathbb{N}. \tag{B1}$$

Here, $\mathcal{L}_1$ and $\mathcal{L}_2$ are time-independent stochastic matrices satisfying the detailed balance condition, and τ is the period of oscillation. The LS constants of $\mathcal{L}_1$ and $\mathcal{L}_2$ are denoted as $\lambda_{\mathrm{LS},1}$ and $\lambda_{\mathrm{LS},2}$, respectively. The system under this setting will finally converge to a periodic stationary state in which $\boldsymbol{p}(t) = \boldsymbol{p}(t+\tau)$.

In the fast oscillation limit $\tau\to 0$, it has been demonstrated that the periodic stationary state reduces to an effective equilibrium state $\boldsymbol{p}^{\mathrm{eff}}$ corresponding to an effective stochastic matrix

$$\mathcal{L}^{\mathrm{eff}} \equiv \frac{\mathcal{L}_1+\mathcal{L}_2}{2}, \tag{B2}$$

i.e.,

$$\mathcal{L}^{\mathrm{eff}}\boldsymbol{p}^{\mathrm{eff}} = 0. \tag{B3}$$

The effective LS constant corresponding to the effective stochastic matrix is given by

$$\begin{aligned}\lambda_{\mathrm{LS}}^{\mathrm{eff}} &= \inf_{\mathrm{Ent}(f)\neq 0}\frac{\mathrm{Re}\langle -\mathcal{L}^{\mathrm{eff}}f, f\rangle_\pi}{\mathrm{Ent}(f)}\\ &= \frac{1}{2}\inf_{\mathrm{Ent}(f)\neq 0}\frac{\mathrm{Re}\langle -(\mathcal{L}_1+\mathcal{L}_2)f, f\rangle_\pi}{\mathrm{Ent}(f)}\\ &= \frac{\inf_{\mathrm{Ent}(f)\neq 0}\frac{\mathrm{Re}\langle -\mathcal{L}_1 f, f\rangle_\pi}{\mathrm{Ent}(f)} + \inf_{\mathrm{Ent}(f)\neq 0}\frac{\mathrm{Re}\langle -\mathcal{L}_2 f, f\rangle_\pi}{\mathrm{Ent}(f)}}{2}\\ &= \frac{\lambda_{\mathrm{LS},1}+\lambda_{\mathrm{LS},2}}{2}.\end{aligned} \tag{B4}$$

Then, applying our first main result Eq. (7) to the situation when the system has reached the effective equilibrium yields that

$$\dot\sigma(t) \geqslant 2(\lambda_{\mathrm{LS},1}+\lambda_{\mathrm{LS},2})D\big[\boldsymbol{p}^{\mathrm{eff}}\big|\big|\boldsymbol{p}_t^{\mathrm{eq}}\big], \tag{B5}$$

since the system is in the effective equilibrium state and the effective LS constant is given by $(\lambda_{\mathrm{LS},1}+\lambda_{\mathrm{LS},2})/2$. Note that the instantaneous equilibrium distribution $\boldsymbol{p}_t^{\mathrm{eq}}$ at time t will be one of the equilibrium distributions $\boldsymbol{p}_1^{\mathrm{eq}}$ or $\boldsymbol{p}_2^{\mathrm{eq}}$ corresponding to $\mathcal{L}_1$ or $\mathcal{L}_2$, which are not matched with the effective equilibrium state. As a consequence, $D[\boldsymbol{p}^{\mathrm{eff}}||\boldsymbol{p}_t^{\mathrm{eq}}] > 0$ so that the lower bound given by Eq. (B5) is positive. For a very large time $t\gg\tau$, one can further bound the entropy production $\sigma_{[0,t]}$ during the interval $[0,t]$ asymptotically from below as

$$\sigma_{[0,t]} \gtrsim \big(2\lambda_{\mathrm{LS},1}D\big[\boldsymbol{p}^{\mathrm{eff}}\big|\big|\boldsymbol{p}_1^{\mathrm{eq}}\big] + 2\lambda_{\mathrm{LS},2}D\big[\boldsymbol{p}^{\mathrm{eff}}\big|\big|\boldsymbol{p}_2^{\mathrm{eq}}\big]\big)t. \tag{B6}$$

This is interesting because the system in an effective equilibrium state may not be distinguished from a real equilibrium state in a coarse grained level, e.g., in the experimental observations level, which may lead to the wrong conclusion that there is no EP. However, our positive bound can recover at least part of the "hidden" EP (rate), which is notably stronger than the conventional second law of thermodynamics. We should emphasize that even when the system is not periodically switching, but randomly switching between two configurations $\mathcal{L}_1$ and $\mathcal{L}_2$ at a constant Poisson rate r, the above result can still apply in the fast switching limit $r\to\infty$, when the system is still in an effective equilibrium state corresponding to $\mathcal{L}^{\mathrm{eff}}$.

Our bound is not limited to the fast oscillation limit, when the period τ is finite, our lower bound can still be applied and probably give a positive value. For example, assuming $t > \tau$ and $\mathcal{L}(0) = \mathcal{L}_1$, one can utilize Eq. (6) to obtain a positive lower bound for $\sigma_{[0,t]}$ in this case:

$$\begin{aligned}\sigma_{[0,t]} \geqslant{}& [1-e^{-4\lambda_{\mathrm{LS},1}\tau}]D\big[\boldsymbol{p}(0)\big|\big|\boldsymbol{p}_{\mathcal{L}_1}^{\mathrm{eq}}\big]\\ &+ [1-e^{-4\lambda_{\mathrm{LS},2}(t-\tau)}]D\big[\boldsymbol{p}(\tau)\big|\big|\boldsymbol{p}_{\mathcal{L}_2}^{\mathrm{eq}}\big],\end{aligned} \tag{B7}$$

where $\boldsymbol{p}(\tau) = \mathcal{T}e^{\int_0^\tau \mathcal{L}_1\mathrm{d}t}\boldsymbol{p}(0)$, $\mathcal{T}$ is the time-ordering.

APPENDIX C: PROPERTIES OF THE LS CONSTANT AND ITS RELATION TO THE SPECTRAL GAP

The LS constant characterizes the intrinsic relaxation timescale (mixing time) as [96]

$$\frac{1}{2\lambda_{\rm LS}} \leqslant \tau_{\rm rel} \leqslant \frac{4+\log\log[1/\pi_\star]}{2\lambda_{\rm LS}}, \tag{C1}$$

where $\pi_\star \equiv \min_i \pi_i$ and the relaxation time $\tau_{\rm rel}$ is defined as

$$\tau_{\rm rel} = \inf\left\{t>0 : \sup_i \left\| \frac{\boldsymbol{p}(t|p_j(0)=\delta_{ij})}{\boldsymbol{\pi}} - \mathbb{I}\right\|_2 \leqslant \frac{1}{e}\right\}, \tag{C2}$$

with the L_2 norm being defined as $||f||_2 \equiv \sqrt{\langle f,f\rangle_\pi} = \sqrt{\sum_i |f_i|^2 \pi_i}$ and $\mathbb{I}$ being the unit vector. Another definition of the relaxation time is given by

$$\tau_{\rm rel} \equiv \frac{1}{2\lambda_{\rm LS}} = -\lim_{t\to\infty} \frac{t}{2\ln D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]}, \tag{C3}$$

which seems require more statistical data to produce a precise value. The advantage of the above definition of $\tau_{\rm rel}$ is that it can be applied to any non-Markovian coarse-grained dynamics without assuming timescale separation. By contrast, the initial definition $\tau_{\rm rel} = \inf\{t>0 : \sup_i \|\frac{\boldsymbol{p}(t|p_j(0)=\delta_{ij})}{\boldsymbol{\pi}} - \mathbb{I}\|_2 \leqslant \frac{1}{e}\}$ can only be effective for fine-grained Markov dynamics and the coarse-grained relaxation dynamics whose KL divergence from the equilibrium state decays exponentially in the large time limit. To assure exponential decay in the large time limit, a clear timescale separation is required. Without timescale separation, the maximization over the initial distribution on coarse-grained level cannot uniquely determine the relaxation timescale $\tau_{\rm rel}^{\rm CG}$, because different microscopic distributions within coarse-grained states will affect $\tau_{\rm rel}^{\rm CG}$.

When the unique stationary distribution $\boldsymbol{\pi}$ is an equilibrium distribution $\boldsymbol{p}^{\rm eq}$, the upper bound of $\tau_{\rm rel}$ can be enhanced by a factor $1/2$. There are similar inequalities for $\tau_{\rm rel}$ using the spectral gap λ_g when detailed balance condition holds, i.e.,

$$\frac{1}{\lambda_g} \leqslant \tau_{\rm rel} \leqslant \frac{2+\log[1/\pi_\star]}{2\lambda_g}.$$

Further, there is a hierarchical relation between the spectral gap λ_g and LS constant $\lambda_{\rm LS}$ [62]:

$$\begin{aligned} \frac{\lambda_g}{2} \geqslant \lambda_{\rm LS} &\geqslant \frac{1-2\pi_\star}{\ln[(1-\pi_\star)/\pi_\star]} \\ \lambda_g &\geqslant \max\left\{\frac{1-2\pi_i}{\ln[(1-\pi_i)/\pi_i]}\lambda_g, 0\right\}. \end{aligned} \tag{C4}$$

The spectral gap λ_g is the second largest eigenvalue of $-\mathcal{L}_s$, and it has a similar definition to $\lambda_{\rm LS}$ as

$$\begin{aligned} \lambda_g &= \inf_{\langle f,f\rangle_\pi \neq 0} \frac{{\rm Re}\langle -\mathcal{L}f, f\rangle_\pi}{\langle f,f\rangle_\pi} \\ &= \inf_{\langle f,f\rangle_\pi \neq 0} \frac{\langle -\mathcal{L}_s f, f\rangle_\pi}{\langle f,f\rangle_\pi}. \end{aligned} \tag{C5}$$

Note that when the detailed balance condition holds, $\mathcal{L}_s = \mathcal{L}$ so that λ_g becomes the second largest eigenvalue of $-\mathcal{L}$ in this case.

Consequently, $\lambda_{\rm LS}$ may characterize the relaxation timescale better compared with the spectral gap λ_g due to the hierarchical relation above (the inequality from $\lambda_{\rm LS}$ is tighter than the inequality from λ_g).

Due to the close connection between $\lambda_{\rm LS}$ and the spectral gap λ_g which is usually easier to determined, one can obtain another useful bound related to λ_g as

$$\dot\sigma(t) \geqslant 4C\lambda_g \sigma_{\rm tot}^t, \tag{C6}$$

where $C = (1-2\pi_\star)/\ln(1-\pi_\star/\pi_\star)$. This bound uncovers a connection between the thermodynamic irreversibility and the spectrum of the dynamical generator in thermal relaxation.

APPENDIX D: THEORETICAL JUSTIFICATIONS OF THE COARSE-GRAINED RESULTS (11)

1. The case when there is timescale separation and the coarse-graining is appropriate

In this case, we can rigorously prove that the trade-off relation for coarse-grained dynamics, i.e., Eq. (11) in the main text is valid by using the data-processing inequality (or log-sum inequality). The data-processing inequality reads

$$D[p(x)||q(x)] \geqslant D[p(y)||q(y)], \tag{D1}$$

where $y = f(x)$ is an arbitrary function of x. This follows from the chain rule of KL divergence, i.e.,

$$\begin{aligned} D[p(x,y)||q(x,y)] &= D[p(y|x)||q(y|x)] + D[p(x)||q(x)] \\ &= D[p(x|y)||q(x|y)] + D[p(y)||q(y)]. \end{aligned} \tag{D2}$$

The term $D[p(y|x)||q(y|x)] = 0$ because $p(y|x) = q(y|x) = 1$ only when $y = f(x)$ and $p(y|x) = q(y|x) = 0$ otherwise. $D[p(x|y)||q(x|y)] \geqslant 0$ and the equality holds when $y = f(x)$ is a one-to-one mapping. Thus, $D[p(x)||q(x)] \geqslant D[p(y)||q(y)]$. Choosing $y = \sum_i c_i \mathbb{I}_{A_i}$, where the universal set $\chi = A_1 + A_2 + \cdots + A_n$, leads to

$$D[p(x)||q(x)] \geqslant \sum_i p(x\in A_i)\ln\frac{p(x\in A_i)}{q(x\in A_i)} \equiv D[\mathcal{P}||\mathcal{Q}], \tag{D3}$$

where $\mathcal{P}_i \equiv p(x\in A_i)$ is the probability that the system is in coarse-grained state i. Due to the arbitrariness of the set A_i, this prove that the KL divergence between always decreases under any conceivable state coarse-graining. Therefore, $D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}] \geqslant D[\mathcal{P}(t)||\mathcal{P}^{\rm eq}]$ and

$$\dot\sigma(t) \geqslant \frac{2}{\tau_{\rm rel}} D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}] \geqslant \frac{2}{\tau_{\rm rel}^{CG}} D[\mathcal{P}(t)||\mathcal{P}^{\rm eq}] \tag{D4}$$

whenever $\tau_{\rm rel} \sim \tau_{\rm rel}^{CG}$ (when $\frac{\tau_{\rm rel}^{CG}}{\tau_{\rm rel}} \geqslant \frac{D[\mathcal{P}(t)||\mathcal{P}^{\rm eq}]}{D[\boldsymbol{p}(t)||\boldsymbol{p}^{\rm eq}]}$), which is a criterion of good coarse-graining mapping. In other words, the breakdown of the above inequality is a strong witness of inappropriate coarse-graining procedure (but we argue that this equality holds for general coarse-graining).

2. General coarse-grained dynamics

In general, we only know that the probability distribution of microscopic dynamics evolves according to the master equation $\frac{d}{dt}p_i(t) = \sum_{ij} k_{ij}(t)p_j - k_{ji}(t)p_i$ and each coarse-grained state consists of many microscopic states. In this

setting, the coarse-grained dynamics can be described by an effective master equation,

$$\frac{d}{dt}\mathcal{P}_m(t) = \sum_n k_{mn}^{\mathrm{CG}}(t)\mathcal{P}_n - k_{nm}^{\mathrm{CG}}(t)\mathcal{P}_m, \tag{D5}$$

where the time-dependent coarse-grained transition rate $k_{mn}^{\mathrm{CG}}(t)$ reads

$$k_{mn}^{\mathrm{CG}}(t) = \sum_{i\in m}\sum_{j\in n} k_{ij}\, p_t(j|n), \tag{D6}$$

where $p_t(j|n) = \frac{p_j(t)}{\sum_{x\in n} p_x(t)}$ is the conditional probability of the system being in microscopic state j at time t given that it is in the coarse-grained state n. The dynamics under this effective master equation is in general non-Markovian because transition rates here are dependent on the probability distribution of microscopic states.

If the detailed balance condition $k_{ij}(t)p_j^{\mathrm{eq}} = k_{ji}(t)p_i^{\mathrm{eq}}$ holds in the fine-grained level, in the coarse-grained level we still have that $k_{mn}^{\mathrm{CG}}(t)\mathcal{P}_n^{\mathrm{eq}} = k_{nm}^{\mathrm{CG}}(t)\mathcal{P}_m^{\mathrm{eq}}$ for any pair of coarse-grained states m, n. If this is not the case, the coarse-grained dynamics can be nonequilibrium even when the original fine-grained dynamics is in equilibrium, which is not physical. With detailed balance condition, one can write that

$$\begin{aligned}
-\frac{\mathrm{d}}{\mathrm{d}t}D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}]|_{\tau=t} &= \sum_i \dot{\mathcal{P}}_i(t)\ln\frac{\mathcal{P}_i^{\mathrm{eq}}}{\mathcal{P}_i(t)} \\
&= \sum_{i,j} k_{ij}^{\mathrm{CG}}(t)\mathcal{P}_j(t)\ln\frac{k_{ij}^{\mathrm{CG}}(t)\mathcal{P}_j(t)}{k_{ji}^{\mathrm{CG}}(t)\mathcal{P}_i(t)} \\
&= \mathcal{K}_{\mathrm{CG}}\sum_{l_{\mathrm{CG}}} p(l_{\mathrm{CG}})\ln\frac{p(l_{\mathrm{CG}})}{p(\tilde{l}_{\mathrm{CG}})} \\
&\leqslant \mathcal{K}\sum_l p(l)\ln\frac{p(l)}{p(\tilde{l})} \\
&= \sum_{i,j} k_{ij}(t)p_j(t)\ln\frac{k_{ij}(t)p_j(t)}{k_{ji}(t)p_i(t)} = \dot{\sigma}(t)
\end{aligned} \tag{D7}$$

Here, $\mathcal{K}_{\mathrm{CG}}$ and $\mathcal{K}$ are dynamical activities in the coarse-grained and fine-grained level, quantifying the mean time between two consecutive jumps in different level. l, $\tilde{l}$ denote microscopic transitions and their time-reversal transitions, and l_{CG}, $\tilde{l}_{\mathrm{CG}}$ are their coarse-grained counterparts. In the second line, the detailed balance condition has been used. The last inequality is due to data-processing inequality. $\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)$ for the coarse-grained dynamics can be defined using the time-dependent generator $\mathcal{L}^{\mathrm{CG}}$. In addition, the inequality

$$-\frac{\mathrm{d}}{\mathrm{d}t}D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}]|_{\tau=t} \geqslant 4\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}] \tag{D8}$$

is valid, because the proof of the inequality only relies on the mathematical form of the master equation. $\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)$ still quantifies the largest relaxation timescale at time t. Combining the above two inequalities, we deduce that the inequality

$$\dot{\sigma}(t) \geqslant 4\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}] \tag{D9}$$

holds for coarse-grained relaxation dynamics. Equation (D9) is a stronger lower bound than the Eq. (12) in the main text. Defining $\lambda_{\mathrm{LS}}^{\mathrm{CG}} \equiv \min_t\{\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)\}$, we then conclude that $\tau_{\mathrm{rel}}^{\mathrm{CG}} \equiv \inf_\tau\{\sup_i \|\frac{\mathcal{P}[\tau|\mathcal{P}_j(0)=\delta_{ij}]}{\mathcal{P}^{\mathrm{eq}}} - \mathbb{I}\|_2 \leqslant 1/e\}$ (the definition works only when there is a timescale separation in fine-grained dynamics) is lower bounded as

$$\tau_{\mathrm{rel}}^{\mathrm{CG}} \geqslant \frac{1}{2\lambda_{\mathrm{LS}}^{\mathrm{CG}}}, \tag{D10}$$

so that we can replace $\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)$ with $\frac{2}{\tau_{\mathrm{rel}}^{\mathrm{CG}}}$ for every t in Eq. (D9).

For another definition of the coarse-grained relaxation timescale, $\tau_{\mathrm{rel}}^{\mathrm{CG}} \equiv \lim_{t\to\infty}\frac{-t}{2\ln D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}]}$, we argue that it satisfies $\frac{1}{2\lambda_{LS}^{CG}(t)} \leqslant \tau_{\mathrm{rel}}^{CG} \leqslant \tau_{\mathrm{rel}}$ for any t. By definition, $\tau_{\mathrm{rel}}^{CG} = \lim_{t\to\infty}\frac{-t}{2\ln D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}]} \leqslant \tau_{\mathrm{rel}} \equiv \lim_{t\to\infty}\frac{-t}{2\ln D[\boldsymbol{p}(t)||\boldsymbol{p}^{\mathrm{eq}}]}$, due to the data-processing inequality $D[\mathcal{P}(t)||\mathcal{P}^{\mathrm{eq}}] \leqslant D[\boldsymbol{p}(t)||\boldsymbol{p}^{\mathrm{eq}}]$ as shown in Appendix A. The first inequality also holds with physical assumptions that

(1) $\lim_{t\to\infty}\frac{1}{2\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)}$ exists and converges to $\tau_{\mathrm{rel}}^{\mathrm{CG}}$.

(2) $\min_t\{\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)\} = \lim_{t\to\infty}\lambda_{\mathrm{LS}}^{\mathrm{CG}}(t)$.

Although the assumptions cannot be rigorously proven, the physics that all faster relaxation modes vanish in the large time limit strongly supports their validity. Consequently, we obtain the desired lower bound, Eq. (12) in the main text, which is experimentally feasible. Note that when the detailed balance condition is broken, i.e., when the stationary state is not an equilibrium state, the justification above does not hold because the data-processing inequality cannot be applied. However, if the coarse-graining mapping is appropriate (as shown in Appendix A), the coarse-grained lower bound for the nonadiabatic EPR still holds.

Additionally, it should be noted that the Markov jump process can be obtained by faithfully discretizing the continuous Fokker-Planck equations [97], which assures that our trade-off relation holds for Langevin dynamics by taking the continuum limits. Therefore, our results can be applied to molecular dynamics simulations, where the underlying dynamics can be described by Langevin equations.

APPENDIX E: ANALYTICAL CALCULATION OF THE EP RATE FOR THE INTERACTING BROWNIAN PARTICLES SYSTEM

Here, we derive an analytical expression for the EPR of a single Brownian particle in a harmonic potential field. We then use this single-particle expression to approximate the EPR of the interacting Brownian particle system discussed in the main text.

Considering a single Brownian particle under a harmonic field $U(\boldsymbol{x}) = k||\boldsymbol{x}-\boldsymbol{a}||^2/2$ in 2-dimensional space, where k is the stiffness and $\boldsymbol{a} = (a_x, a_y)$ is the center of the field. The corresponding Langevin function reads $\dot{\boldsymbol{x}} = -\mu\nabla U(\boldsymbol{x}) + \sqrt{2D}\boldsymbol{\xi}(t)$. The probability distribution of the position $\boldsymbol{x} = (x, y)$ of the particle evolves according to the Fokker-Planck equation

$$\begin{aligned}
\frac{\partial p_t(x,y)}{\partial t} &= \mu[\partial_x(k(x-a_x) - T\partial_x) \\
&\quad + \partial_y(k(y-a_y) - T\partial_y)]p_t(x,y),
\end{aligned} \tag{E1}$$

with μ being the mobility and satisfying $\mu = D/k_BT$. Multiplying both sides with x (or y) and then integrating both sides

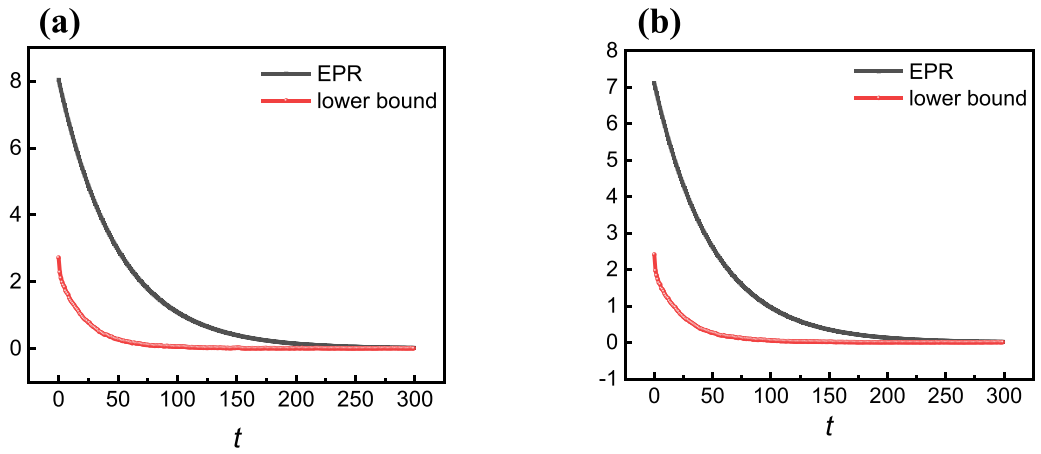

FIG. 3. The coarse-grained trade-off relations for the interacting Brownian particles system with more coarse-grained states: (a) The lower bound is calculated by dividing the space uniformly into 16 regions (16 states). (b) The lower bound is calculated by dividing the space uniformly into 25 regions (25 states).

with respect to y (or x), leading to two decoupled equations for the first moments as

$$d_t\langle x\rangle_t = -\mu k(\langle x\rangle_t - a_x)$$
$$d_t\langle y\rangle_t = -\mu k\big(\langle y\rangle_t - a_y\big), \tag{E2}$$

where the conservation of probability has been used. We let $a_x = a_y = 0$, i.e., the harmonic field is posed at the center of the space. Then the moments at time t are solved as

$$\langle x\rangle_t = \langle x\rangle_0 e^{-\mu kt},\ \langle y\rangle_t = \langle y\rangle_0 e^{-\mu kt}. \tag{E3}$$

If the initial distribution is Gaussian, the probability $p_t(x, y)$ will keep Gaussian at any time, i.e.,

$$p_t(x, y) = \frac{k}{2\pi T}\exp\left(-\frac{k||\boldsymbol{x} - \langle\boldsymbol{x}\rangle_t||^2}{2T}\right).$$

The EPR is

$$\dot\sigma(t) = \frac{1}{\mu T}\int d\boldsymbol{x}||\boldsymbol{v}_t(\boldsymbol{x})||^2 p_t(\boldsymbol{x}), \tag{E4}$$

where the local mean velocity is defined as

$$\boldsymbol{v}_t(\boldsymbol{x}) \equiv \mu(-\nabla U(\boldsymbol{x}) - T\nabla \ln p_t(\boldsymbol{x})). \tag{E5}$$

Using the expressions for $p_t(\boldsymbol{x}) = p_t(x, y)$ and $\langle x\rangle_t$, $\langle y\rangle_t$, we obtain that the EPR for a single Brownian particle at time t is

$$\dot\sigma(t) = \frac{Dk^2}{k_BT^2}\big(\langle x\rangle_0^2 + \langle y\rangle_0^2\big)e^{-\frac{2Dk}{k_BT}t}. \tag{E6}$$

Since the interacting potential in our case is also harmonic, and the interaction strength $\kappa \ll k$, we conclude that the total EPR for N such Brownian particles at time t is simply

$$\dot\sigma_{\rm tot}(t) \approx \frac{NDk^2}{k_BT^2}\big(\overline{\langle x\rangle_0^2} + \overline{\langle y\rangle_0^2}\big)e^{-\frac{2Dk}{k_BT}t}, \tag{E7}$$

where $\overline{\langle x\rangle_0^2} \equiv \frac{1}{N}\sum_k\langle x_k\rangle_0^2$ and $\overline{\langle y\rangle_0^2} \equiv \frac{1}{N}\sum_k\langle y_k\rangle_0^2$. This expression may slightly underestimate the true EP rate, which will not affect the validity of our trade-off relation. It is clear that the information of microscopic states is still needed to calculate the EPR using this expression, even if the model details are known in prior.

APPENDIX F: SIMULATION DETAILS AND FURTHER NUMERICAL RESULTS

1. Calculating lower bounds for interacting brownian particles with different levels of coarse-graining

The interaction potential is the spring potential, which reads

$$U_{in}(r_{ij}) = \begin{cases}\frac{1}{2}\kappa(r_{ij} - r_c)^2 & r_{ij} < r_c\\ 0 & r_{ij} \geqslant r_c\end{cases}. \tag{F1}$$

Here, the $r_c = 1.0$ is the cutoff distance. $\kappa = 0.01 \ll k = 0.1$, where k is the stiffness of the external field $U(\boldsymbol{x}) = k||\boldsymbol{x} - \boldsymbol{a}||^2/2$. The diffusion constant $D = k_BT = 1.0$. The simulation box has a size of 50×50 with periodic boundary conditions. In Fig. 3, we divide the two-dimensional box uniformly into 16 and 25 regions, in contrast with the 4 regions in the main text. When the coarse-grained level increases (the number of coarse-grained states increases), the lower bound becomes closer to the real EPR.

2. Another example: Interacting active Brownian particles system

We consider a more complex scenario involving interacting active Brownian particles (see Fig. 4). The dynamics of this system is described by the following overdamped Langevin equations

$$\dot{\boldsymbol{r}} = D\beta[\boldsymbol{F}_i + \boldsymbol{n}_i v] + \sqrt{2D}\boldsymbol{\xi}_i(t)$$
$$\dot\theta_i = \sqrt{2D_r}\eta_i(t), \tag{F2}$$

where $\boldsymbol{F}_i = -\sum_{i\neq j}\nabla U(r_{ij})$ and v denotes the strength of active force whose orientation is described by the unit vector $\boldsymbol{n}_i = (\cos\theta_i, \sin\theta_i)$. $\beta = \frac{1}{k_BT} = 10.0$. The strength of active force is chosen to be $v = 20.0$. More simulation details can be found in Ref. [98]. We define $\boldsymbol{F}_{\rm tot} \equiv \boldsymbol{F}_i + \boldsymbol{n}_i v$. With detailed knowledged of the model, the (non-adiabatic) EPR at time t is calculated approximately by numerically integrating $\dot\sigma(t) = \frac{1}{\delta t}\int_t^{t+\delta t}\boldsymbol{F}_{\rm tot}\circ\dot{\boldsymbol{r}}dt$, where δt is a small time step.

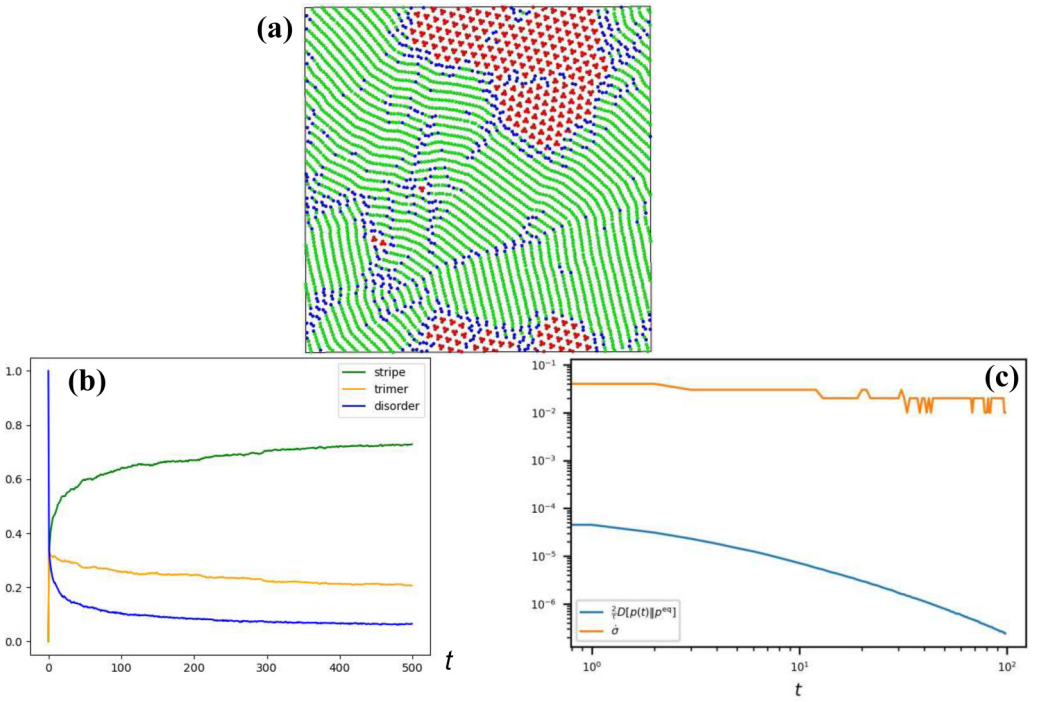

FIG. 4. Demonstration of the coarse-grained trade-off relation in an interacting active particles system. There are 4096 active particles. (a) A snapshot of the simulation, showing the system self-assembling into three coarse-grained states: the stripe state (green part), the trimer state (red part), and the disorder state (blue part). (b) The evolution of the probability distributions of the three coarse-grained states during relaxation. (c) The nonadiabatic EPR calculated from detailed knowledge (orange curve) and the coarse-grained lower bound (blue curve) using only the statistics of the coarse-grained states.

This example demonstrates the limitation of our inference strategy, specifically that it can only provide a very loose bound when the accessible dynamics are highly coarse-grained. This remains an open question within the stochastic thermodynamics community, as the reduction of EP due to coarse-graining (information loss) is inevitable. If we apply our lower bound to systems with more coarse-grained states, we anticipate that the bound will more closely approximate the true EPR, similar to the first example.

APPENDIX G: GENERALIZATION TO OPEN QUANTUM SYSTEMS AND CONTINUOUS-SPACE MARKOV PROCESSES

1. Details of generalization to Markovian open quantum systems

Here we show that, our main results can be generalized to quantum Markov processes described by the Lindblad master equations. In this setting, the dynamics of the density operator $\rho_t \equiv \rho(t)$ of the system at time t is given by $\dot{\rho}_t = \mathcal{L}(\rho_t)$, where

$$\mathcal{L}_t(\rho) \equiv -i[H_t, \rho] + \sum_i \left[J_i \rho J_i^\dagger - \frac{1}{2}\{J_i^\dagger J_i, \rho\} \right] \tag{G1}$$

is the Lindbladian. Here, H_t is the Hamiltonian (can possibly be time-dependent) in a $d-$dimensional Hilbert space $\mathcal{H}^d$ and J_i is the jth jump operator describing dissipation effect due to the environment. To proceed, we assume that the Lindbladian satisfies the quantum detailed balance condition [99], in which case the density operator will finally converge to a Gibbs state $\rho_\beta = \lim_{t\to\infty} e^{\mathcal{L}t}\rho_0 = e^{-\beta H}/\mathrm{Tr}(e^{-\beta H})$ when H is time-independent. The inner product should be redefined as $\langle A, B\rangle_\pi \equiv \mathrm{Tr}(A^\dagger B \rho_\beta)$, the average over the Gibbs state reads $\langle A\rangle_\pi \equiv \mathrm{Tr}(A\rho_\beta)$ and the quantum KL divergence is given by $D(\rho||\rho') = \mathrm{Tr}(\rho \ln \rho - \rho \ln \rho')$. The EPR $\dot{\sigma}_t$ at time t in the open quantum systems can be separated to the change rate of system entropy $\dot{s}_t = \mathrm{Tr}(\dot{\rho}_t \ln \rho_t)$ and the heat flow $\beta\dot{q} = \beta\mathrm{Tr}(\dot{\rho}_t H_t)$ as $\dot{\sigma}_t = \dot{s}_t - \beta\dot{q}$. Like in the classical case, $\dot{\sigma}_t$ has a direct connection with KL divergence that $\dot{\sigma}_t = -\partial_t D(\rho_t||\rho_\beta)$ [100]. When H is time-dependent, $\dot{\sigma}_t = -\partial_t D(\rho_t||\rho_{\beta,\tau})|_{t=\tau}$, where $\rho_{\beta,\tau}$ is the instantaneous Gibbs state defined as $\mathcal{L}_\tau(\rho_{\beta,\tau}) = 0$. Recently, it has been proved that there always exist a positive constant α assuring that the quantum LS inequality

$$-\langle \mathcal{L}(f), \ln f\rangle_\pi \geqslant \alpha \langle f \ln f\rangle_\pi \tag{G2}$$

holds [101] for any postive operator $f \in \mathcal{H}^d$ satisfying $\langle f\rangle_\pi = 1$, once the quantum detailed balance condition is satisfied. Consequently, a straightforward calculation shows that

$$\begin{aligned}
\dot{\sigma}_t =& -\partial_t D(\rho_t||\rho_{\beta,\tau})|_{t=\tau} \\
=& -\left\langle \mathcal{L}\left(\frac{\rho_t}{\rho_{\beta,t}}\right), \ln\frac{\rho_t}{\rho_{\beta,t}} \right\rangle_\pi (t) \\
\geqslant& \lambda_{\mathrm{QLS}}(t) \left\langle \frac{\rho_t}{\rho_{\beta,t}} \ln \frac{\rho_t}{\rho_{\beta,t}} \right\rangle_\pi, \\
=& \lambda_{\mathrm{QLS}}(t) D(\rho_t||\rho_{\beta,t}),
\end{aligned} \tag{G3}$$

where

$$\lambda_{\rm QLS} \equiv \inf_{f>0,\ \langle f\rangle_\pi=1} \frac{-\langle \mathcal{L}(f), \ln f\rangle_\pi}{\langle f \ln f\rangle_\pi} > 0 \tag{G4}$$

is the quantum LS constant. Integrating both parts of Eq. (G3) from 0 to t results in $D(\rho_t||\rho_\beta) \leqslant D(\rho_0||\rho_\beta)e^{-\lambda_{\rm QLS}t}$. Then, an inverse quantum speed limit can still be directly obtained as in the classical case when H is time-independent:

$$\tau \leqslant \frac{1}{\lambda_{\rm QLS}} \ln\left\{\frac{\sigma^0_{\rm tot}}{\sigma^0_{\rm tot}-\sigma_{[0,\tau]}}\right\}, \tag{G5}$$

where $\sigma_{[0,\tau]} = \int_0^\tau \dot\sigma_t {\rm d}t = D(\rho_0||\rho_\beta) - D(\rho_t||\rho_\beta)$ and $\sigma^0_{\rm tot} = D(\rho_0||\rho_\beta)$. The quantum LS constant can also be connected with the relaxation time scale by using the Pinsker inequality $({\rm Tr}|\rho-\rho'|)^2 \leqslant 2D(\rho||\rho')$. In open quantum systems, the distance between two density operator is commonly described by the trace distance defined as $D_{\rm Tr}(\rho||\rho') \equiv \frac{1}{2}{\rm Tr}|\rho-\rho'|$. Thus the relaxation time scale to the Gibbs state is naturally characterized by the convergence rate of $D_{\rm Tr}(\rho_t||\rho_\beta)$. Here, we have that

$$\begin{aligned} D_{\rm Tr}(\rho_t||\rho_\beta) &\leqslant \sqrt{D(\rho_t||\rho_\beta)/2} \\ &\leqslant \sqrt{D(\rho_0||\rho_\beta)/2}e^{-\lambda_{\rm QLS}t/2}, \end{aligned} \tag{G6}$$

which implies that $\lambda_{\rm QLS}$ is a characterization of relaxation timescale in open quantum systems. In summary, the upper bound of the transformation time τ in relaxation of open quantum systems depends both on the relaxation time scale of the whole process and the initial energetic cost, similar to the classical case.

In addition, we provide an alternative way to generalize our results to open quantum systems when H is time-independent, which is the generalization shown in the Eq. (10) of main text. As is known, the Lindblad master equation can lead to an equation of motion for the populations

$$P_n(t) \equiv \langle n|\rho_t|n\rangle \tag{G7}$$

of the eigenstates $|n\rangle$ of the system Hamiltonian $H_s = \sum_n \epsilon_n|n\rangle\langle n|$ (assuming that H_s is nondegenerate) [100]. The equation of motion is given by

$$\frac{{\rm d}P_n(t)}{{\rm d}t} = \sum_m [W_{nm}P_m(t) - W_{mn}P_n(t)], \tag{G8}$$

which is often referred to as the Pauli master equation. In this equation, time-independent transition rates W_{nm} from the energy level ϵ_m to ϵ_n are given by

$$W_{nm} = \sum_i (\epsilon_m - \epsilon_n)\langle n|J_i|m\rangle^2.$$

If the LS constant associated with the Markov generator $\tilde{\boldsymbol{W}}$ (whose entries are $\tilde W_{nm} = W_{nm} - \delta_{nm}\sum_l W_{ln}$) is $\lambda_{\rm LS}$, one has that

$$-\frac{\rm d}{{\rm d}t}D[\boldsymbol{P}(t)||\boldsymbol{P_\beta}] \geqslant 4\lambda_{\rm LS}D[\boldsymbol{P}(t)||\boldsymbol{P_\beta}] \tag{G9}$$

$$D[\boldsymbol{P}(\tau)||\boldsymbol{P_\beta}] \leqslant e^{-4\lambda_{\rm LS}\tau}D[\boldsymbol{P}(\tau)||\boldsymbol{P_\beta}], \tag{G10}$$

where $D[\cdot||\cdot]$ is the classical KL divergence and entries of $\boldsymbol{P_\beta}$ are given by $P_{\beta,n} = \langle n|\rho_\beta|n\rangle$. The quantum EP can be decomposed as

$$\dot\sigma_t = -\partial_t D(\rho_t||\rho_\beta) = -\frac{\rm d}{{\rm d}t}D[\boldsymbol{P}(t)||\boldsymbol{P_\beta}] - \frac{{\rm d}A(t)}{{\rm d}t}, \tag{G11}$$

where $A(t)$ is the asymmetry defined as $A(t) := D(\rho_t||\rho_t^{\rm d})$. Here, $\rho_t^{\rm d}$ is the fully decohered version of ρ_t. It has been shown that $-\frac{{\rm d}A(t)}{{\rm d}t} \geqslant 0$, which quantifies the EPR from destroying quantum coherence. Consequently, we have that

$$\dot\sigma_t \geqslant 4\lambda_{\rm LS}D[\boldsymbol{P}(t)||\boldsymbol{P_\beta}] - \frac{{\rm d}A(t)}{{\rm d}t} \geqslant 4\lambda_{\rm LS}D[\boldsymbol{P}(t)||\boldsymbol{P_\beta}], \tag{G12}$$

and the EP

$$\sigma_{[0,\tau]} = D[\boldsymbol{P}(0)||\boldsymbol{P_\beta}] - D[\boldsymbol{P}(\tau)||\boldsymbol{P_\beta}] + A(0) - A(\tau) \tag{G13}$$

during any interval $[0,\tau]$ is also bounded from below as

$$\sigma_{[0,\tau]} \geqslant (1-e^{-4\lambda_{\rm LS}\tau})D[\boldsymbol{P}(0)||\boldsymbol{P_\beta}] + \Delta A, \tag{G14}$$

with $\Delta A \equiv A(0) - A(\tau) \geqslant 0$. Thus, a quantum inverse speed limit tighter than its classical counterpart is obtained as

$$\tau \leqslant \frac{1}{4\lambda_{\rm LS}} \ln\left\{\frac{D[\boldsymbol{P}(0)||\boldsymbol{P_\beta}]}{D[\boldsymbol{P}(0)||\boldsymbol{P_\beta}] - (\sigma_{[0,\tau]} - \Delta A)}\right\}. \tag{G15}$$

Whether the above relations hold true in the nonequilibrium open quantum systems remains an interesting open problem.

2. Generalization to continuous-space Markov processes

The Markov processes in continuous-space can be described by the Fokker-Planck equation. Here, we would like to discuss a system within a time-dependent conservative force field $U(x,t)$, where $x \in \mathbb{R}^n$ is a n-dimensional vector. The dynamics of the system is described by a Langevin equation (we have set the mobility $\mu = 1$)

$$\dot x(t) = F(x,t) + \sqrt{2D}\xi(t), \tag{G16}$$

where the force $F(x,t) \equiv -\nabla U(x,t)$. The corresponding Fokker-Planck equation reads

$$\frac{\partial p(x,t)}{\partial t} = -\nabla j(x,t), \tag{G17}$$

where the current $j(x,t) \equiv F(x,t)p(x,t) - D\nabla p(x,t)$. The Fokker-Planck equation has an instantaneous stationary solution $p^{\rm st}_t(x) \propto e^{-\beta U(x,t)}$ with Boltzmann form at any time t.

According to Ref. [102], the LS inequality

$$\frac{\lambda}{2}\int u \ln u p^{\rm st}(x){\rm d}x \leqslant \int |\nabla\sqrt{u}|^2 p^{\rm st}(x){\rm d}x \tag{G18}$$

holds for any positive function $u = u(x,t)$ satisfying $\int u(x,t)p^{\rm st}(x){\rm d}x = 1$ and any stationary distribution $p^{\rm st}(x) \propto e^{-\beta U(x)}$, if the following condition is fulfilled for the positive constant λ:

$$\nabla^2 U(x) \geqslant \lambda \mathbb{I}_n,$$

where $\mathbb{I}_n$ is the n-dimensional identity matrix. For instance, consider a one-dimensional Brownian particle confined in a harmonic potential $U(x) = \frac{1}{2}kx^2$. In this case, $U''(x) = k$, such that the positive constant λ is exactly equal to the stiffness k of the potential.

The EPR $\dot{\sigma}(t)$ at time t obtained from Eq. (G17) can be expressed as the time-derivative of the KL divergence [103], i.e.,

$$\dot{\sigma}(t) = -\frac{\mathrm{d}}{\mathrm{d}t} D\big[p(x,s)\big|\big|p_t^{st}(x)\big]_{|s=t} = \int \left|\nabla \ln\left[\frac{p(x,t)}{p_t^{st}(x)}\right]\right|^2 p(x,t)\mathrm{d}x \geqslant 0. \tag{G19}$$

Notice that $\dot{\sigma}(t)$ can be rewritten as

$$\dot{\sigma}(t) = 4\int \left|\nabla\sqrt{\frac{p(x,t)}{p_t^{st}(x)}}\right|^2 p_t^{st}(x)\mathrm{d}x, \tag{G20}$$

then applying the LS inequality (G18) to it yields

$$\dot{\sigma}(t) \geqslant 2\lambda \int \frac{p(x,t)}{p_t^{st}(x)} \ln\frac{p(x,t)}{p_t^{st}(x)} p^{st}(x)\mathrm{d}x = 2\lambda D\big[p(x,t)\big|\big|p_t^{st}(x)\big]. \tag{G21}$$

Therefore, there is still a general lower bound $\dot{\sigma}(t) \geqslant 2\lambda D[p(x,t)||p_t^{st}(x)]$ for the continuous-space Markov process described by Eq. (G17), serving as a stronger second law of thermodynamics.

Whether it is possible to find a general lower bound for the local EPR using local version of the KL divergence or other distance function is another interesting open question.

3. Generalization to discrete-time Markov processes

The nonadiabatic EP for such process during relaxation can also be expressed by KL divergence as

$$\sigma_{[0,\tau]} = D[\boldsymbol{p}_0||\boldsymbol{\pi}] - D[\boldsymbol{p}_\tau||\boldsymbol{\pi}]. \tag{G22}$$

It can be proved that

$$D[\boldsymbol{p}_\tau||\boldsymbol{\pi}] \leqslant (1-\alpha_d)^\tau D[\boldsymbol{p}_0||\boldsymbol{\pi}], \tag{G23}$$

leading to the result

$$\sigma_{[0,\tau]} \geqslant [1-(1-\alpha_d)^\tau] D[\boldsymbol{p}_0||\boldsymbol{\pi}], \tag{G24}$$

where the LS constant α_d for the discrete-time case is defined as

$$\alpha_d \equiv \min\left\{\frac{\langle (1-KK^\star)f, f\rangle}{\mathrm{Ent}(f)}\right\}.$$

[1] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[2] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[3] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).
[4] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[5] C. Jarzynski and D. K. Wójcik, Phys. Rev. Lett. **92**, 230602 (2004).
[6] M. Esposito, U. Harbola, and S. Mukamel, Rev. Mod. Phys. **81**, 1665 (2009).
[7] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
[8] G. Verley and D. Lacoste, Phys. Rev. E **86**, 051127 (2012).
[9] G. Verley, R. Chétrite, and D. Lacoste, Phys. Rev. Lett. **108**, 120601 (2012).
[10] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Nature (London) **437**, 231 (2005).
[11] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002).
[12] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, Phys. Rev. Lett. **113**, 030601 (2014).
[13] N. Shiraishi, K. Funo, and K. Saito, Phys. Rev. Lett. **121**, 070601 (2018).
[14] Z. Cao, R. Bao, J. Zheng, and Z. Hou, J. Phys. Chem. Lett. **14**, 66 (2023).
[15] G. Falasco and M. Esposito, Phys. Rev. Lett. **125**, 120604 (2020).
[16] L. P. García-Pintos, S. B. Nicholson, J. R. Green, A. del Campo, and A. V. Gorshkov, Phys. Rev. X **12**, 011038 (2022).
[17] R. Hamazaki, PRX Quantum **3**, 020319 (2022).
[18] T. Van Vu and K. Saito, Phys. Rev. Lett. **130**, 010402 (2023).
[19] J. S. Lee, S. Lee, H. Kwon, and H. Park, Phys. Rev. Lett. **129**, 120603 (2022).
[20] T. Van Vu and K. Saito, Phys. Rev. X **13**, 011013 (2023).
[21] E. Aurell, K. Gawędzki, C. Mejía-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, J. Stat. Phys. **147**, 487 (2012).
[22] E. Roldán, I. Neri, M. Dörpinghaus, H. Meyr, and F. Jülicher, Phys. Rev. Lett. **115**, 250602 (2015).
[23] S. Pigolotti, I. Neri, E. Roldán, and F. Jülicher, Phys. Rev. Lett. **119**, 140604 (2017).
[24] I. Neri, E. Roldán, and F. Jülicher, Phys. Rev. X **7**, 011019 (2017).
[25] N. Shiraishi and K. Saito, Phys. Rev. Lett. **123**, 110603 (2019).
[26] I. Neri, Phys. Rev. Lett. **124**, 040601 (2020).
[27] T. Van Vu and Y. Hasegawa, Phys. Rev. Lett. **127**, 190601 (2021).
[28] T. Van Vu and Y. Hasegawa, Phys. Rev. Lett. **126**, 010601 (2021).
[29] A. C. Barato and U. Seifert, Phys. Rev. Lett. **114**, 158101 (2015).
[30] T. R. Gingrich, J. M. Horowitz, N. Perunov, and J. L. England, Phys. Rev. Lett. **116**, 120601 (2016).
[31] J. M. Horowitz and T. R. Gingrich, Phys. Rev. E **96**, 020103(R) (2017).
[32] Z. Cao, H. Jiang, and Z. Hou, Phys. Rev. Res. **2**, 043331 (2020).
[33] A. Dechant and S.-I. Sasa, Phys. Rev. Res. **3**, L042012 (2021).
[34] A. Dechant and S.-I. Sasa, Phys. Rev. X **11**, 041061 (2021).
[35] B. Das, S. K. Manikandan, and A. Banerjee, Phys. Rev. Res. **4**, 043080 (2022).
[36] R. Bao and Z. Hou, Phys. Rev. E **107**, 024112 (2023).
[37] S. K. Manikandan, D. Gupta, and S. Krishnamurthy, Phys. Rev. Lett. **124**, 120603 (2020).
[38] T. R. Gingrich and J. M. Horowitz, Phys. Rev. Lett. **119**, 170601 (2017).
[39] Z. Cao, J. Su, H. Jiang, and Z. Hou, Phys. Fluids **34**, 053310 (2022).

[40] A. Lapolla and A. C. V. Godec, Phys. Rev. Lett. **125**, 110602 (2020).
[41] J. Meibohm and M. Esposito, New J. Phys. **25**, 023034 (2023).
[42] J. Meibohm and M. Esposito, Phys. Rev. Lett. **128**, 110603 (2022).
[43] Z. Lu and O. Raz, Proc. Natl. Acad. Sci. USA **114**, 5083 (2017).
[44] S. K. Manikandan, Phys. Rev. Res. **3**, 043108 (2021).
[45] T. Van Vu and Y. Hasegawa, Phys. Rev. Res. **3**, 043160 (2021).
[46] G. Biroli and J. Kurchan, Phys. Rev. E **64**, 016101 (2001).
[47] D. C. Rose, K. Macieszczak, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. E **94**, 052132 (2016).
[48] K. Macieszczak, M. Guţă, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. Lett. **116**, 240404 (2016).
[49] T. Mori, Phys. Rev. Res. **3**, 043137 (2021).
[50] K. Macieszczak, D. C. Rose, I. Lesanovsky, and J. P. Garrahan, Phys. Rev. Res. **3**, 033047 (2021).
[51] I. Klich, O. Raz, O. Hirschberg, and M. Vucelja, Phys. Rev. X **9**, 021060 (2019).
[52] A. Kumar and J. Bechhoefer, Nature (London) **584**, 64 (2020).
[53] A. Santos and A. Prados, Phys. Fluids **32**, 072010 (2020).
[54] A. Gal and O. Raz, Phys. Rev. Lett. **124**, 060602 (2020).
[55] M. Baity-Jesi, E. Calore, A. Cruz, L. A. Fernandez, J. M. Gil-Narvión, A. Gordillo-Guerrero, D. Iñiguez, A. Lasanta, A. Maiorano, E. Marinari *et al.*, Proc. Natl. Acad. Sci. USA **116**, 15350 (2019).
[56] F. Carollo, A. Lasanta, and I. Lesanovsky, Phys. Rev. Lett. **127**, 060401 (2021).
[57] D. M. Busiello, D. Gupta, and A. Maritan, New J. Phys. **23**, 103012 (2021).
[58] R. Bao, Z. Cao, J. Zheng, and Z. Hou, Phys. Rev. Res. **5**, 043066 (2023).
[59] R. Bao and Z. Hou, Phys. Rev. Lett. **135**, 150403 (2025).
[60] T. Mori and T. Shirai, Phys. Rev. Lett. **130**, 230404 (2023).
[61] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley & Sons, New York, 1991).
[62] E. Giné, G. R. Grimmett, and L. Saloff-Coste, *Lectures on Probability Theory and Statistics: Ecole d'eté de probabilités de Saint-Flour XXVI-1996* (Springer, 2006).
[63] S. Deffner and E. Lutz, Phys. Rev. Lett. **105**, 170402 (2010).
[64] U. Seifert, Annu. Rev. Condens. Matter Phys. **10**, 171 (2019).
[65] E. Roldán and J. M. R. Parrondo, Phys. Rev. Lett. **105**, 150607 (2010).
[66] I. A. Martínez, G. Bisker, J. M. Horowitz, and J. M. R. Parrondo, Nat. Commun. **10**, 3542 (2019).
[67] D. J. Skinner and J. Dunkel, Phys. Rev. Lett. **127**, 198101 (2021).
[68] S. Ro, B. Guo, A. Shih, T. V. Phan, R. H. Austin, D. Levine, P. M. Chaikin, and S. Martiniani, Phys. Rev. Lett. **129**, 220601 (2022).
[69] J. van der Meer, B. Ertel, and U. Seifert, Phys. Rev. X **12**, 031025 (2022).
[70] P. E. Harunari, A. Dutta, M. Polettini, and E. Roldán, Phys. Rev. X **12**, 041026 (2022).
[71] J. van der Meer, J. Degünther, and U. Seifert, Phys. Rev. Lett. **130**, 257101 (2023).
[72] K. Blom, K. Song, E. Vouga, A. Godec, and D. E. Makarov, Proc. Natl. Acad. Sci. USA **121**, e2318333121 (2024).
[73] I. Di Terlizzi, M. Gironella, D. Herraez-Aguilar, T. Betz, F. Monroy, M. Baiesi, and F. Ritort, Science **383**, 971 (2024).
[74] N. Shiraishi and T. Sagawa, Phys. Rev. E **91**, 012130 (2015).
[75] S. Lee, D.-K. Kim, J.-M. Park, W. K. Kim, H. Park, and J. S. Lee, Phys. Rev. Res. **5**, 013194 (2023).
[76] J. Kappler and R. Adhikari, Phys. Rev. E **105**, 044107 (2022).
[77] S. Otsubo, S. K. Manikandan, T. Sagawa, and S. Krishnamurthy, Commun. Phys. **5**, 11 (2022).
[78] N. Ohga and S. Ito, Phys. Rev. Res. **6**, 013315 (2024).
[79] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[80] M. Esposito and C. Van den Broeck, Phys. Rev. E **82**, 011143 (2010).
[81] Y.-Z. Zhen, D. Egloff, K. Modi, and O. Dahlsten, Phys. Rev. Lett. **127**, 190602 (2021).
[82] K. Yoshimura, A. Kolchinsky, A. Dechant, and S. Ito, Phys. Rev. Res. **5**, 013017 (2023).
[83] D. M. Busiello, S. Liang, F. Piazza, and P. De Los Rios, Commun. Chem. **4**, 16 (2021).
[84] Z. Zhang, V. Du, and Z. Lu, Phys. Rev. E **107**, L012102 (2023).
[85] S.-W. Wang, K. Kawaguchi, S.-I. Sasa, and L.-H. Tang, Phys. Rev. Lett. **117**, 070601 (2016).
[86] S. Vaikuntanathan and C. Jarzynski, Europhys. Lett. **87**, 60005 (2009).
[87] M. Esposito and C. Van Den Broeck, Europhys. Lett. **95**, 40004 (2011).
[88] L. Oberreiter, U. Seifert, and A. C. Barato, Phys. Rev. E **106**, 014106 (2022).
[89] N. Ohga, S. Ito, and A. Kolchinsky, Phys. Rev. Lett. **131**, 077101 (2023).
[90] A. Kolchinsky, N. Ohga, and S. Ito, Phys. Rev. Res. **6**, 013082 (2024).
[91] T. Van Vu, V. T. Vo, and K. Saito, Phys. Rev. Res. **6**, 013273 (2024).
[92] W. Yang, C. Templeton, D. Rosenberger, A. Bittracher, F. Nüske, F. Noé, and C. Clementi, ACS Central Sci. **9**, 186 (2023).
[93] X. Zhao, D. Hartich, and A. C. V. Godec, Phys. Rev. Lett. **132**, 147101 (2024).
[94] R. Bao, N. Ohga, and S. Ito, arXiv:2508.11586.
[95] Q. Yu, D. Zhang, and Y. Tu, Phys. Rev. Lett. **126**, 080601 (2021).
[96] P. Diaconis and L. Saloff-Coste, Ann. Appl. Prob. **6**, 695 (1996).
[97] N. Shiraishi, K. Saito, and H. Tasaki, Phys. Rev. Lett. **117**, 190601 (2016).
[98] Y. Du, H. Jiang, and Z. Hou, J. Chem. Phys. **151**, 154904 (2019).
[99] R. Firanko, M. Goldstein, and I. Arad, J. Math. Phys. **65**, 051901 (2024).
[100] H.-P. Breuer, F. Petruccione *et al.*, *The Theory of Open Quantum Systems* (Oxford University Press on Demand, 2002).
[101] L. Gao and C. Rouzé, Arch. Ration. Mech. Anal. **245**, 183 (2022).
[102] P. A. Markowich and C. Villani, Mat. Contemp **19**, 1 (2000).
[103] A. Dechant, S.-I. Sasa, and S. Ito, Phys. Rev. E **106**, 024125 (2022).