

Consistent Time Reversal and Reliable and Accurate Inference in the Presence of Memory

Tassilo Schwarz^{1,2}, Anatoly B. Kolomeisky^{3,4} and Aljaž Godec^{5,1,*}

¹*Mathematical bioPhysics group, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany*

²*Mathematical Institute, University of Oxford, Oxford, OX2 6GG, United Kingdom*

³*Department of Chemistry, Rice University, Houston, Texas 77005, USA*

⁴*Center for Theoretical Biological Physics, Department of Chemical and Biomolecular Engineering and Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA*

⁵*Mathematical Physics and Stochastic Dynamics, Faculty of Mathematics and Physics, University of Freiburg*



(Received 14 October 2024; revised 15 December 2025; accepted 25 September 2026; published 29 October 2026)

Thermodynamic inference from coarse observations remains a key challenge. Memory, in particular correlations between consecutively observed mesostates, blur signatures of irreversibility and must be accounted for in defining physical time reversal, which remains an open problem. We derive an experimentally accessible k th order estimator for the entropy production rate. Using novel measure-theoretic techniques we prove necessary and sufficient conditions for *guaranteed* lower bounds on the dissipation even in the strongly non-Markovian setting. The proof reveals that estimators saturated in the order unravel the duration of memory that needs to be considered in defining physically consistent time reversal. We show that Markovian estimators in the absence of a timescale separation lead to artifacts that convey no physical meaning. Similarly, estimators *not* saturated in the order may overestimate the dissipation. The necessity of *correctly* accounting for memory in thermodynamic inference from strongly non-Markovian observations underscores the still underappreciated challenges and intricacies in defining and understanding irreversibility in presence of memory. Our results will hopefully stimulate experiments systematically considering thermodynamic inference on multiple scales *consistently* accounting for memory.

DOI: 10.1103/gwf7-6kvr

Measurements in complex systems usually have finite resolution [1–4] or do *not* resolve all relevant degrees of freedom—they just probe some low-dimensional projections of the dynamics, where many microscopic states become “lumped” onto the same observable “state.” Examples are macroscopic observables in condensed matter systems, such as magnetization [5], dielectric response [6], and diverse order parameters [7–9], as well as other observables (e.g., molecular extensions or Förster resonance energy transfer efficiencies and lifetimes) in neural activity [10], single-molecule [11–20], and particle-tracking [21,22] experiments.

It is well known that such coarse graining introduces memory in the dynamics of the observable quantities [23–26] (see Ref. [3] for the effect of a finite resolution). Much less is understood about out-of-equilibrium systems that display memory. In particular, a consistent formulation [26–36] and inference [35,37–41] of dissipation in the

presence of slow dissipative hidden degrees of freedom is often a daunting task. It is very difficult to reliably infer violations of time reversibility (i.e., the detailed balance) from experiments on individual molecules [36,38,42].

Violations of detailed balance on the level of individual stochastic trajectories $\Gamma_t \equiv (x_\tau)_{0 \leq \tau \leq t}$ may be quantified via the steady-state dissipation (entropy production) rate $\dot{S}$. Given a *physically consistent* time-reversal operation θ accounting for the highly nontrivial (anti)persistence due to memory [36], $\dot{S}$ is defined as the relative entropy between the probability measure of a path $\mathbb{P}[\Gamma_t]$ and its time reverse $\mathbb{P}[\theta\Gamma_t]$ per unit time [43–45], i.e.,

$$\dot{S}[\Gamma_t] = k_B \lim_{t \rightarrow \infty} \frac{1}{t} \left\langle \ln \frac{d\mathbb{P} \Gamma_t}{d\mathbb{P} [\theta\Gamma_t]} \right\rangle, \quad (1)$$

where $\langle \cdot \rangle$ denotes the average over $d\mathbb{P}[\Gamma_t]$ and k_B is Boltzmann’s constant. When x_t is an overdamped Markov dynamics, we simply have $\theta\Gamma_t = (x_{t-\tau})_{0 \leq \tau \leq t}$. Specifically, for an ergodic Markov process on a discrete state space V with transition rates L_{ji} from state i to j and stationary probabilities P^s , Eq. (1) becomes [43–45]

$$\dot{S}_M[\Gamma_t] = k_B \sum_{i < j} (L_{ji} P_i^s - L_{ij} P_j^s) \ln \frac{L_{ji}}{L_{ij}}. \quad (2)$$

*Contact author: agodec@physik.uni-freiburg.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. Open access publication funded by the Max Planck Society.

For thermodynamic consistency we must always have $L_{ji} > 0 \Rightarrow L_{ij} > 0$ for all $i, j \neq i \in V$ [45]. We henceforth express all estimates of $\dot{S}$ in units of k_B .

When a coarse-grained version $\hat{X}_t$ of the process is observed, where several microscopic states are lumped into the same observable state, provided a physically consistent time reversal operation θ [27,33,40,46], the definition in Eq. (1) with lumped paths $\hat{\Gamma}_t = (\hat{X}_\tau)_{0 \leq \tau \leq t}$ may still hold, whereas Eq. (2) does *not*. The reason is that $\hat{X}_t$ displays memory due to coarse graining, i.e., correlations between consecutively observed mesostates, which must be taken explicitly into account in the time reversal [33–35,40,47,48]. Determining the precise manner in which memory must enter the consistent time reversal remains elusive. If one simply replaces the rate L_{ji} between microscopic states in the Markovian result (2) with the effective rates between (mesoscopic) lumped states averaged over P^s , one finds [29] $\dot{S}_M[\hat{\Gamma}_t] \leq \dot{S}_M[\Gamma_t]$. However, unless one assumes the existence of a (potentially unphysical [34]) infinite timescale separation between dynamics within and between lumped states, $\dot{S}_M[\hat{\Gamma}_t]$ is *not* the physical dissipation. In particular, this inequality does *not* imply that a dependence of $\dot{S}_M[\hat{\Gamma}_t]$ on the properties of the lumped state space carries any physical meaning since $\dot{S}_M[\hat{\Gamma}_t]$ is *not* the dissipation rate of the coarse-grained process, $\dot{S}_M[\hat{\Gamma}_t] \neq \dot{S}[\hat{\Gamma}_t]$. In absence of an infinite timescale separation thermodynamic interpretations of $\dot{S}_M[\hat{\Gamma}_t]$ (see, e.g., Refs. [49,50]) may lead to misconceptions and guaranteed higher-order estimators $\dot{S}_k^{\text{est}}$ are required, as we explicitly show below. A timescale separation on some scale does *not* imply a separation on coarser scales. Under timescale separation on *all* scales, $\dot{S}_M$ does *not* display any scale dependence; see Appendix C.

Energy dissipation is a critically important quantity since it measures how the system deviates from the equilibrium. It also sets the thermodynamic cost of biological functions [51,52] and transport phenomena [53–59], such as molecular-motor mediated transport of cargo [60,61], biochemical oscillations [62,63], error correction in biochemical reactions [64–66], or cellular computations [67]. These functions typically involve many processes coupled over multiple spatial and temporal scales [49,51,52]. It is thus of great interest to unravel how much energy is dissipated on different length scales.

Physically inconsistent time reversal ignoring memory may lead to misinterpretations [49]. More importantly, memory implies correlations between consecutively visited mesostates but the extent and manner in which these need to be accounted for when defining time-reversal symmetry, and in turn when estimating the entropy production rate, remain a grand challenge.

We prove here with measure-theoretic techniques that the *only* relevant transition probabilities that enter the estimation of $\dot{S}$ are those accounting for the *relevant extent of memory*

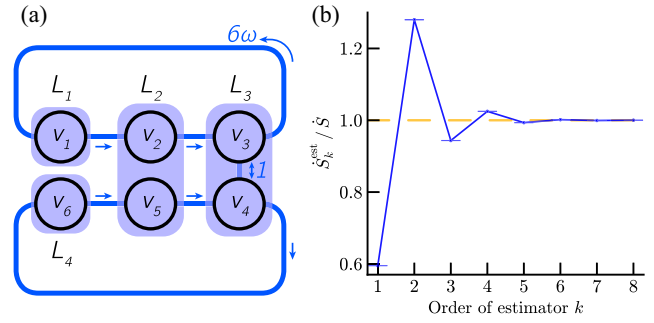


FIG. 1. (a) Six-state Markov system lumped to a process of order $n > 3$. The order is at least 3 because a transition out of L_3 depends on whether we arrived from L_1 or L_4 two state changes before. Since multiple jumps between $L_2 \leftrightarrow L_3$ may have occurred before, the order is in fact higher, but the influence of back-and-forth jumps vanishes exponentially. (b) Entropy production rate (with error bars from 30 independent identically distributed simulations) estimated with $\dot{S}_k^{\text{est}}$ from trajectories as a function of k ; saturation is observed from $k = 5$ onward. Each depicted $\dot{S}_k^{\text{est}}$ was evaluated from path data with 10^9 transitions with our code optimized for GPUs (available at [68]) within seconds (i.e., 1.2 s for order $k = 1$, and 2.4 s for $k = 8$ on an H100 GPU). For details see Supplemental Material [69].

(but not those accounting for shorter or longer memory). This means that saturation in the order is an inherent quantifier for the extent of memory that enters the definition of physical time reversal. We construct an experimentally accessible estimator of $\dot{S}$ for strongly non-Markovian observations (see examples in Figs. 1 and 2) and prove that it yields *guaranteed* sharp lower bounds. Further examples, with strong memory and hidden cycles, demonstrate the power and necessity of the framework, by highlighting pitfalls in using inconsistent Markovian estimators.

In particular, we present fractal networks dissipating (i) only on a single (i.e., largest) length scale and (ii) networks dissipating mostly on the smallest length scale and less on larger scales (i.e., networks with an “inverse energy dissipation cascade” [49]) whereby both display the same scaling of $\dot{S}_M$. A self-similar network structure is shown to *not* imply a power-law scaling of $\dot{S}_M$, therefore there is

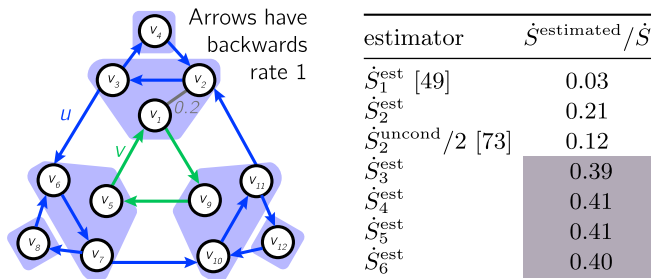


FIG. 2. Coarse graining (left) and estimators (right) for $u = v = 2$: $\dot{S}_k^{\text{est}}$ allows for good inference of $\dot{S}$ already for $k = 2$ and saturates in the order for $k \geq 3$ up to statistical fluctuations, indicated in gray.

generally *no* relation between the power-law scaling of $\dot{S}_M$ and the energy dissipated on a given scale. The apparent power-law dependence of $\dot{S}_M$ has *no* relation to the actual energy dissipated on a given scale and may simply be an artifact of an inconsistent time reversal, i.e., of neglecting correlations between consecutively visited mesostates due to memory. Most importantly, when no dissipative cycles become hidden by the coarse graining, any scale dependence vanishes entirely once the memory is accounted for exactly when estimating $\dot{S}$. Furthermore, with a measure-theoretic framework, our higher-order estimates are proved to be reliable indicators of hidden dissipative cycles and, if saturated in the order that ensures consistent time reversal, provide *guaranteed* sharper lower bounds. These results emphasize the importance of properly accounting for memory in analyzing measurements from coarse observations.

Setup—Consider a continuous-time Markov process x_t on a finite discrete state space, with L_{ji} denoting transition rates $i \rightarrow j$. We are interested in the steady-state dissipation rate $\dot{S}$ in Eq. (1) in situations where one might only observe some coarse (i.e., lumped) version of the process that we denote as $\hat{X}_t$. For simplicity, and to align with the assumptions in [49], we assume the absence of kinetic hysteresis [33,34,40], i.e., that the coarse graining and time reversal commute. The observed process is hence a non-Markovian jump process on a partitioned “lumped space,” whereby the memory duration is finite but may become arbitrarily long depending on the coarse graining. In the following examples, the lumped process will correspond to a k th order semi-Markov process with the instantaneous state $\hat{X}_t$ depending on the previous k states [78]. Thus, $k = 1$ corresponds to a renewal process and, if the waiting time is exponentially distributed, to a Markov process.

Guaranteed higher-order estimators—We will estimate $\dot{S}[\hat{\Gamma}_t]$ within the simplest Markov approximation as in [49] as well as in terms of k th order estimators [40] that account for memory. For simplicity and without any loss of generality we will neglect any potential waiting-time contributions (for further details see, e.g., Ref. [40]). If we let $\hat{\gamma}_k^{(k_i:k_j)}$ denote a subsequence $k_i, k_{i+1}, \dots, k_j$ of a particular tuple of $k + 1$ (with $k \geq 1$) consecutive observed states $\hat{\gamma}_k$ with final state $\hat{\gamma}_k^{(f)}$, and $\mathbb{P}(\hat{\gamma}_k^{(k_i:k_j)})$ the probability to observe this subsequence along a (formally infinitely long) trajectory, then the k th order estimator of $\dot{S}$ is given by

$$\dot{S}_k^{\text{est}} \equiv \frac{1}{T} \sum_{\hat{\gamma}_k} \mathbb{P}[\hat{\gamma}_k^{(1:k+1)}] \ln \frac{\mathbb{P}[\hat{\gamma}_k^{(f)} | \hat{\gamma}_k^{(1:k)}]}{\mathbb{P}[(\theta \hat{\gamma}_k)^{(f)} | (\theta \hat{\gamma}_k)^{(1:k)}]}, \quad (3)$$

where T is the average waiting time per state in the observed trajectories, $\theta \hat{\gamma}_k$ is the corresponding reversed tuple of $k + 1$ states, and the upper indices refer to the discrete state sequence [79]. In [69], we prove that

$\dot{S}_k^{\text{uncond}} = \sum_{i=1}^k \dot{S}_i^{\text{est}}$. If $\hat{\Gamma}_t$ is an n th order semi-Markov process then $\dot{S}_k^{\text{est}} = \dot{S}_n^{\text{est}}$, $\forall k \geq n$ (see proof in [69]), and if $\hat{\Gamma}_t = \Gamma_t$ is a Markov process we have $\dot{S}_1^{\text{est}} = \dot{S}$. Notably, if the observed semi-Markov process $\hat{\Gamma}$ is of order n , then $\dot{S}_k^{\text{est}}$ for $k \geq n$ provides a lower bound on the microscopic $\dot{S}$,

$$\dot{S}(\Gamma) \geq \dot{S}_k^{\text{est}}(\hat{\Gamma}) \quad \forall k \geq n. \quad (4)$$

We further prove in [69] that the estimator will reach a plateau upon increasing k once reaching the correct order n ; at this order (and *not* before) consistent time reversal is obtained, correctly accounting for correlations between consecutive mesostates. Furthermore, $\dot{S}_n^{\text{est}}$ (but *not* $\dot{S}_{k < n}^{\text{est}}$) always provides a lower bound on $\dot{S}(\Gamma)$ (see our proof in [69]; this has been stated, but not proven, for *discrete* time in [80]). Continuous-time mesoscopic paths require more care since they involve long-time limits of unbounded hidden-jump sums (see Appendix D for a trivial argument), creating the need for a novel measure-theoretic technique that we develop here. Our proof includes the discrete time result. Crucially, *no* statement can be made about the monotonicity of $\dot{S}_k^{\text{est}}$ as a function of k for $k < n$ as the inequality (4) may be violated (see Fig. 1 and Appendix A). Inferring the exact $\dot{S}$ from projected observations is generally *not* possible because of dissipative cycles hidden in lumps.

We provide an instructive demonstration of the consequences of our framework in Fig. 1. The six-state example shows that estimating with insufficient order may lead to underestimation (e.g., if a Markovian estimator is used) and even overestimation. Importantly, at sufficiently high order the estimator plateaus. In the present example, the lower bound [Eq. (4)] even becomes sharp.

Sharpening inference at strong memory—The system in Fig. 2 illustrates the benefits of accounting for long-range memory in thermodynamic inference: We observe that accounting for progressively longer memory k yields a better estimate of $\dot{S}$ via $\dot{S}_k^{\text{est}}$, until we reach saturation for $k \geq 3$. Therefore, physically consistent time reversal requires accounting for $k = 3$ past states. Note the substantial improvement over the Markovian estimate $\dot{S}_M = \dot{S}_1^{\text{est}}$ [49]. Further, while $\dot{S}_2^{\text{uncond}}/2$ provides a lower bound [81] (this estimator was originally developed for the broader setting of blurred transitions [82]), we recover with $\dot{S}_3^{\text{est}}$ in Fig. 2 substantially higher fraction of the microscopic dissipation, and get saturation for $k \geq 3$. In contrast to Fig. 1, full recovery is not feasible due to indistinguishable hidden cycles. Further parameters are considered in Appendix E. An efficient GPU implementation of the estimator is provided in [68]. We now turn to the necessity of accounting for memory in time reversal, when considering sequential (scale-dependent) lumping.

On hierarchical lumping and $\dot{S}$ —We first consider Markovian dynamics on a (trivially self-similar) simple

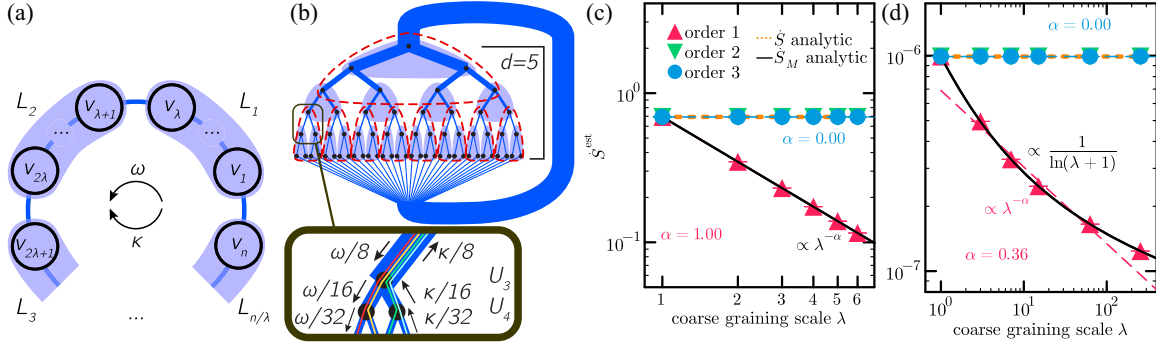


FIG. 3. Model systems with sequential lumping and their entropy production rate estimation. (a) Cycle graph with transition rates ω and κ in counterclockwise (+) and clockwise (−) direction, respectively. Lumps $L_1, L_2, \dots, L_{n/\lambda}$ are highlighted in blue. (b) Tree of depth $d = 5$ with lumps of depth $l = 1$ (shaded in blue) and depth $l = 2$ (circled in dashed red). Transition rates down and up (cyclic with respect to the thick edge) are set $\propto \omega$ and $\propto \kappa$, respectively, chosen such that the cumulative transition rate between each level is $\omega - \kappa$. Self-similarity becomes obvious upon noticing that each tree consists of multiple smaller trees. Inset: four fundamental cycles passing through the highlighted subgraph on levels U_3, U_4 , and corresponding transition rates. (c) $\dot{S}_k^{\text{est}}$ for the ring graph with $n = 60$ states. Higher-order estimators ($k \geq 2$) allow for correct inference of $\dot{S}$ (green, blue), while omitting memory leads to an artefactual power law (analytic black line overlays red dashed fit to order-1-estimates). Analytical results are fully corroborated by simulations. (d) $\dot{S}_k^{\text{est}}$ for the tree of depth $d = 23$, i.e., $n = 16'777'215$. Higher-order estimators, in combination with exploiting the system's symmetry, allow for correct inference of $\dot{S}$ (green, blue), whereas ignoring memory leads to a spurious inverse energy-dissipation cascade (black line), which may be easily mistaken for a power law (red dashed fit). In both cases, we considered 100 independent stationary trajectories, each visiting 10^9 microscopic states.

cycle graph with $n \geq 4$ states with forward and backward transition rates ω and $\kappa \neq \omega$, respectively [see Fig. 3(a)]. The system has a single dissipative scale with an exact total dissipation rate,

$$\dot{S} = (\omega - \kappa) \ln\left(\frac{\omega}{\kappa}\right). \quad (5)$$

As in [49], we decimate the state space in n/λ observable mesostates (we assume $n = 0 \bmod \lambda$). The coarse process is non-Markovian; to be precise, it is a second-order semi-Markov process (see also Ref. [34]). This means that the waiting time within, and probability for a forward or backward transition from, a lump depends on the previous state. The exact $\dot{S}$ may be determined in terms of J and the two-step affinity A_2 [34,40,47],

$$\dot{S}_2^{\text{est}} = JA_2 = J \ln\left(\frac{\Phi_{+|+}}{\Phi_{-|-}}\right) = \dot{S}, \quad (6)$$

where $\Phi_{\pm|\pm}$ denotes the conditional stationary probability for a step in the $\pm$ direction given that the previous step was also in the $\pm$ direction. The proof of the last equality is given in Supplemental Material [69] and verified by simulations in Fig. 3(c). Equality, as opposed to a lower bound, emerges because no dissipative cycles become hidden, and the waiting-time contribution vanishes (see Supplemental Material [69]).

Assuming instead memoryless coarse dynamics with “Markovian” forward and backward rates defined as in [49] given by $\Omega(\lambda) = \omega/\lambda$ and $K(\lambda) = \kappa/\lambda$, respectively, we

obtain a single dissipative cycle with homogeneous steady-state current $J = (\lambda/n)[\Omega(\lambda) - K(\lambda)] = (\omega - \kappa)/n$ and affinity $A = (n/\lambda) \ln(\omega/\kappa)$, yielding the Markov estimate

$$\dot{S}_M = JA = \frac{1}{\lambda}(\omega - \kappa) \ln\left(\frac{\omega}{\kappa}\right) \propto \lambda^{-1}, \quad (7)$$

indeed an inverse power law according to [49]. However, the microscopic system has a single dissipative scale, the entire cycle. Thus, the scaling has *no* implications for the energy dissipation on distinct length scales; it is merely an artifact of inconsistent time reversal in the presence of memory. Consistent time reversal and hence thermodynamics thus improve the inference and remove the artefactual scaling behavior.

We now consider a perfectly self-similar network with a single dissipative *scale* (i.e., the system size) but multiple (equivalent) dissipative *cycles* each with a stationary current J_c [see Fig. 3(b)]. In this binary tree of depth d , every edge is bidirectional with rate $\propto \omega$ in the “down” (+) and $\propto \kappa$ in the “up” (−) direction, respectively, such that the cumulative transition rate between each level is $\omega - \kappa$. This way, the rates along incident edges in consecutive levels in both directions $\pm$ are halved [see inset in Fig. 3(b)]. The detailed mathematical construction is described in [69]. The exact entropy production of the system in the cycle basis $\mathcal{C}$ with 2^d basis cycles reads (see proof in [69])

$$\dot{S} = \sum_{c \in \mathcal{C}} J_c A_c = \frac{d+1}{2^{d+1}-1} (\omega - \kappa) \ln\left(\frac{\omega}{\kappa}\right), \quad (8)$$

where the basis-cycle current (see Ref. [83] for details on cycle bases) is $J_c = (\omega - \kappa)/2^d(2^{d+1} - 1)$.

We lump the tree into mesoscopic states each having a depth $l < d$ [see Fig. 3(b)]. The lumping yields a second-order semi-Markov process and hides *no* dissipative cycles. By accounting for memory via $\dot{S}_{k \geq 2}^{\text{est}}$ [34,40,47] we recover the *exact* microscopic result (see derivation in [69]),

$$\dot{S}_2^{\text{est}} = \frac{\omega - \kappa}{2^{d+1} - 1} \frac{d+1}{l+1} \ln \left(\frac{\Phi_{++}}{\Phi_{--}} \right) = \dot{S}_{k \geq 2}^{\text{est}} = \dot{S}, \quad (9)$$

where $\Phi_{\pm|\pm}$ is the conditional probability for a transition between two lumps (at given coarseness) in the $\pm$ direction given that the preceding step occurred in the $\pm$ direction. The exact result (9) is confirmed by simulations in Fig. 3(d) and generalized to a tree-diamond graph in [69]. It is *independent* of λ since the microscopic system has a single dissipative scale.

The Markovian estimate $\dot{S}_M$ for such a lumped system instead reads (see derivation in [69])

$$\dot{S}_M = \frac{d+1}{2^{d+1} - 1} \frac{\ln(2)}{\ln(\lambda + 1)} (\omega - \kappa) \ln \left(\frac{\omega}{\kappa} \right) \propto \frac{1}{\ln(\lambda + 1)}, \quad (10)$$

where $\lambda = n/n_s = 2^{l+1} - 1$. Now there is in fact strictly *no* power-law scaling of $\dot{S}_M$ in λ despite a perfectly self-similar network structure. However, one may confuse the scaling for a power law $\sim \lambda^{-0.36}$ [see Fig. 3(d)]. An observed power law thus has *no* relation with energy dissipation on distinct scales and may not even be a power law. It is an artifact of inconsistent time reversal in the presence of memory. Consistent time reversal removes the artefactual scaling behavior.

Conclusion—A novel, *experimentally accessible* dissipation estimator for coarse-grained observations in continuous-time was introduced and proved to yield a guaranteed lower bound. The estimator is applicable to strongly non-Markovian observations (i.e., beyond weakly Markovian [84,85] and correlation-time expansions [86,87]). For the first time, the estimator provides insight into the extent of memory that needs to be considered when defining physically consistent time reversal.

The necessity of correctly accounting for memory in time reversal was demonstrated by invalidating a recently proposed relation between an apparent power-law dependence of Markovian estimates for the energy dissipation rate $\dot{S}_M$ from coarse-grained observations, and the actual dissipation on different length scales. We constructed fractal networks with a single dissipative scale and self-similar networks (see Appendix B) with an “inverse energy dissipation cascade” [49], where most energy is dissipated on the smallest and less energy on larger scales. Both systems exhibit the same scaling behavior of the energy

dissipation at different scales. The apparent scaling exponent gradually reduces upon accounting for memory in the dissipation estimates and only if saturated in the order, provides information on hidden dissipative cycles (see Supplemental Material [69]); scaling laws ignoring memory reflect a mixture of memory and hidden cycles. With a hypothesized power-law scaling of entropy production rate [49] being already mentioned but not validated in experimental work [88], we hope that our novel framework stimulates experiments systematically considering multiple scales. When no dissipative cycles become hidden by the coarse graining, any scale dependence vanishes as soon as the memory is accounted for exactly in the time-reversal operation and we can infer the microscopic dissipation. Therefore, an inverse power-law dependence of the dissipation rate generally has *no* implications for scale-dependent energy dissipation and may simply be an artifact of inconsistent time reversal in presence of memory.

However, it may still be possible to infer meaningful information about the microscopic dynamics, e.g., the presence of hidden dissipative cycles, via higher-order dissipation estimators established here. These incorporate a consistent time reversal in the presence of memory and, if (and only if) saturated in the order, yield *experimentally accessible* [89,90], *feasible* [91,92], *sharper*, and *guaranteed* lower bounds on dissipation, further highlighting the importance and benefits of accounting for memory in analyzing data from coarse observations.

Acknowledgments—Financial support from the European Research Council (ERC) under the European Union’s Horizon Europe research and innovation program (Grant Agreement No. 101086182 to A.G.) and from the Studienstiftung des Deutschen Volkes, the EPSRC Centre for Doctoral Training in Mathematics of Random Systems: Analysis, Modelling and Simulation (EPSRC Grant No. EP/S023925/1), and the Rhodes Trust (to T.S.) are acknowledged. A.B.K. also acknowledges the support from the Welch Foundation (Grant No. C-1559) and the Center for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

- [1] J. Gladrow, N. Fakhri, F. C. MacKintosh, C. F. Schmidt, and C. P. Broedersz, Broken detailed balance of filament dynamics in active networks, *Phys. Rev. Lett.* **116**, 248301 (2016).
- [2] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. MacKintosh, Broken detailed balance at mesoscopic scales in active biological systems, *Science* **352**, 604 (2016).
- [3] C. Dieball and A. Godec, Mathematical, thermodynamical, and experimental necessity for coarse graining empirical densities and currents in continuous space, *Phys. Rev. Lett.* **129**, 140601 (2022).

- [4] C. Dieball and A. Godec, Coarse graining empirical densities and currents in continuous-space steady states, *Phys. Rev. Res.* **4**, 033243 (2022).
- [5] D. Hérisson and M. Ocio, Fluctuation-dissipation ratio of a spin glass in the aging regime, *Phys. Rev. Lett.* **88**, 257202 (2002).
- [6] H. Oukris and N. E. Israeloff, Nanoscale non-equilibrium dynamics and the fluctuation–dissipation relation in an ageing polymer glass, *Nat. Phys.* **6**, 135 (2009).
- [7] A. J. Bray, S. N. Majumdar, and G. Schehr, Persistence and first-passage properties in nonequilibrium systems, *Adv. Phys.* **62**, 225 (2013).
- [8] S. Burov, R. Metzler, and E. Barkai, Aging and non-ergodicity beyond the Khinchin theorem, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13228 (2010).
- [9] J. H. P. Schulz, E. Barkai, and R. Metzler, Aging renewal theory and application to random walks, *Phys. Rev. X* **4**, 011028 (2014).
- [10] T. GrandPre, G. Teza, and W. Bialek, Direct estimates of irreversibility from time series, [arXiv:2412.19772](https://arxiv.org/abs/2412.19772).
- [11] A. K. Sangha and T. Keyes, Proteins fold by subdiffusion of the order parameter, *J. Phys. Chem. B* **113**, 15886 (2009).
- [12] S. S. Plotkin and P. G. Wolynes, Non-Markovian configurational diffusion and reaction coordinates for protein folding, *Phys. Rev. Lett.* **80**, 5015 (1998).
- [13] W. Min, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, Observation of a power-law memory kernel for fluctuations within a single protein molecule, *Phys. Rev. Lett.* **94**, 198302 (2005).
- [14] D. E. Makarov, Interplay of non-Markov and internal friction effects in the barrier crossing kinetics of biopolymers: Insights from an analytically solvable model, *J. Chem. Phys.* **138**, 014102 (2013).
- [15] J. Brujić, R. I. Hermans Z., K. A. Walther, and J. M. Fernandez, Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin, *Nat. Phys.* **2**, 282 (2006).
- [16] I. Grossman-Haham, G. Rosenblum, T. Namani, and H. Hofmann, Slow domain reconfiguration causes power-law kinetics in a two-state enzyme, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 513 (2018).
- [17] I. M. Sokolov, Cyclization of a polymer: First-passage problem for a non-Markovian process, *Phys. Rev. Lett.* **90**, 080601 (2003).
- [18] T. Guérin, O. Bénichou, and R. Voituriez, Non-Markovian polymer reaction kinetics, *Nat. Chem.* **4**, 568 (2012).
- [19] A. M. Berezhkovskii and D. E. Makarov, Single-molecule test for Markovianity of the dynamics along a reaction coordinate, *J. Phys. Chem. Lett.* **9**, 2190 (2018).
- [20] A. Lapolla and A. Godec, Toolbox for quantifying memory in dynamics along reaction coordinates, *Phys. Rev. Res.* **3**, L022018 (2021).
- [21] F. Ginot, J. Caspers, M. Krüger, and C. Bechinger, Barrier crossing in a viscoelastic bath, *Phys. Rev. Lett.* **128**, 028001 (2022).
- [22] N. Narinder, C. Bechinger, and J. R. Gomez-Solano, Memory-induced transition from a persistent random walk to circular motion for achiral microswimmers, *Phys. Rev. Lett.* **121**, 078003 (2018).
- [23] S. Nordholm and R. Zwanzig, A systematic derivation of exact generalized Brownian motion theory, *J. Stat. Phys.* **13**, 347 (1975).
- [24] H. Grabert, P. Talkner, P. Hänggi, and H. Thomas, Microdynamics and time-evolution of macroscopic non-Markovian systems. II, *Z. Phys. B* **29**, 273 (1978).
- [25] M. Ferrario and P. Grigolini, The non-Markovian relaxation process as a “contraction” of a multidimensional one of Markovian type, *J. Math. Phys. (N.Y.)* **20**, 2567 (1979).
- [26] S. Rahav and C. Jarzynski, Fluctuation relations and coarse-graining, *J. Stat. Mech.* (2007) P09012.
- [27] H. Wang and H. Qian, On detailed balance and reversibility of semi-Markov processes and single-molecule enzyme kinetics, *J. Math. Phys. (N.Y.)* **48**, 013303 (2007).
- [28] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Role of hidden slow degrees of freedom in the fluctuation theorem, *Phys. Rev. Lett.* **108**, 220601 (2012).
- [29] M. Esposito, Stochastic thermodynamics under coarse graining, *Phys. Rev. E* **85**, 041125 (2012).
- [30] A. Puglisi, S. Pigolotti, L. Rondoni, and A. Vulpiani, Entropy production and coarse graining in Markov processes, *J. Stat. Mech.* (2010) P05015.
- [31] G. Teza and A. L. Stella, Exact coarse graining preserves entropy production out of equilibrium, *Phys. Rev. Lett.* **125**, 110601 (2020).
- [32] P. Talkner and P. Hänggi, Colloquium: Statistical mechanics and thermodynamics at strong coupling: Quantum and classical, *Rev. Mod. Phys.* **92**, 041002(C) (2020).
- [33] D. Hartich and A. Godec, Emergent memory and kinetic hysteresis in strongly driven networks, *Phys. Rev. X* **11**, 041047 (2021).
- [34] D. Hartich and A. Godec, Violation of local detailed balance upon lumping despite a clear timescale separation, *Phys. Rev. Res.* **5**, L032017 (2023).
- [35] J. van der Meer, B. Ertel, and U. Seifert, Thermodynamic inference in partially accessible Markov networks: A unifying perspective from transition-based waiting time distributions, *Phys. Rev. X* **12**, 031025 (2022).
- [36] C. Dieball and A. Godec, Perspective: Time irreversibility in systems observed at coarse resolution, *J. Chem. Phys.* **162**, 090901 (2025).
- [37] P. E. Harunari, A. Dutta, M. Polettini, and E. Roldán, What to learn from a few visible transitions’ statistics?, *Phys. Rev. X* **12**, 041026 (2022).
- [38] A. Godec and D. E. Makarov, Challenges in inferring the directionality of active molecular processes from single-molecule fluorescence resonance energy transfer trajectories, *J. Phys. Chem. Lett.* **14**, 49 (2023).
- [39] J. van der Meer, J. Degünther, and U. Seifert, Time-resolved statistics of snippets as general framework for model-free entropy estimators, *Phys. Rev. Lett.* **130**, 257101 (2023).
- [40] K. Blom, K. Song, E. Vouga, A. Godec, and D. E. Makarov, Milestoning estimators of dissipation in systems observed at a coarse resolution, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2318333121 (2024).
- [41] J. Degünther, J. van der Meer, and U. Seifert, General theory for localizing the where and when of entropy production meets single-molecule experiments, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2405371121 (2024).

- [42] C. Ratzke, B. Hellenkamp, and T. Hugel, Four-colour FRET reveals directionality in the Hsp90 multicomponent machinery, *Nat. Commun.* **5**, 4192 (2014).
- [43] 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).
- [44] P. Gaspard, Fluctuation theorem for nonequilibrium reactions, *J. Chem. Phys.* **120**, 8898 (2004).
- [45] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [46] D. Hartich and A. Godec, Comment on “Inferring broken detailed balance in the absence of observable currents”, *Nat. Commun.* **15**, 8678 (2024).
- [47] 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).
- [48] X. Zhao, D. Hartich, and A. Godec, Emergence of memory in equilibrium versus nonequilibrium systems, *Phys. Rev. Lett.* **132**, 147101 (2024).
- [49] Q. Yu, D. Zhang, and Y. Tu, Inverse power law scaling of energy dissipation rate in nonequilibrium reaction networks, *Phys. Rev. Lett.* **126**, 080601 (2021).
- [50] Q. Yu and Y. Tu, State-space renormalization group theory of nonequilibrium reaction networks: Exact solutions for hypercubic lattices in arbitrary dimensions, *Phys. Rev. E* **105**, 044140 (2022).
- [51] H. Qian, Phosphorylation energy hypothesis: Open chemical systems and their biological functions, *Annu. Rev. Phys. Chem.* **58**, 113 (2007).
- [52] R. D. Astumian, Stochastic conformational pumping: A mechanism for free-energy transduction by molecules, *Annu. Rev. Biophys.* **40**, 289 (2011).
- [53] N. Shiraishi, K. Funo, and K. Saito, Speed limit for classical stochastic processes, *Phys. Rev. Lett.* **121**, 070601 (2018).
- [54] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Optimal protocols and optimal transport in stochastic thermodynamics, *Phys. Rev. Lett.* **106**, 250601 (2011).
- [55] S. Ito and A. Dechant, Stochastic time evolution, information geometry, and the Cramér-Rao bound, *Phys. Rev. X* **10**, 021056 (2020).
- [56] N. Shiraishi and K. Saito, Information-theoretical bound of the irreversibility in thermal relaxation processes, *Phys. Rev. Lett.* **123**, 110603 (2019).
- [57] G. Falasco and M. Esposito, Dissipation-time uncertainty relation, *Phys. Rev. Lett.* **125**, 120604 (2020).
- [58] 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).
- [59] C. Dieball and A. Godec, Thermodynamic bounds on generalized transport: From single-molecule to bulk observables, *Phys. Rev. Lett.* **133**, 067101 (2024).
- [60] A. B. Kolomeisky and M. E. Fisher, Molecular motors: A theorist’s perspective, *Annu. Rev. Phys. Chem.* **58**, 675 (2007).
- [61] U. Seifert, Efficiency of autonomous soft nanomachines at maximum power, *Phys. Rev. Lett.* **106**, 020601 (2011).
- [62] L. Oberreiter, U. Seifert, and A. C. Barato, Universal minimal cost of coherent biochemical oscillations, *Phys. Rev. E* **106**, 014106 (2022).
- [63] Y. Cao, H. Wang, Q. Ouyang, and Y. Tu, The free-energy cost of accurate biochemical oscillations, *Nat. Phys.* **11**, 772 (2015).
- [64] J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4135 (1974).
- [65] J. Ninio, Kinetic amplification of enzyme discrimination, *Biochimie* **57**, 587 (1975).
- [66] J. D. Mallory, O. A. Igoshin, and A. B. Kolomeisky, Do we understand the mechanisms used by biological systems to correct their errors?, *J. Phys. Chem. B* **124**, 9289 (2020).
- [67] P. Mehta and D. J. Schwab, Energetic costs of cellular computation, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17978 (2012).
- [68] See Code at SchwarzTass/consistent-time-reversal for an efficient implementation of the estimator. <https://github.com/schwarzTass/Consistent-Time-Reversal>.
- [69] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/gwf7-6kvr> for details on the models and the underlying calculations and mathematical proofs of the claims, which includes Refs. [70–75].
- [70] C. Godrèche and J. M. Luck, Statistics of the occupation time of renewal processes, *J. Stat. Phys.* **104**, 489 (2001).
- [71] K. J. Falconer, *Fractals: A Very Short Introduction*, 2 (Oxford University Press, Oxford, 2013), 10.1093/ac-trade/9780199675982.003.0002.
- [72] C. Song, S. Havlin, and H. A. Makse, Self-similarity of complex networks, *Nature* **433**, 392 (2005).
- [73] B. Krön, Growth of self-similar graphs, *J. Graph Theory* **45**, 224 (2004).
- [74] L. Malozemov and A. Teplyaev, Pure point spectrum of the Laplacians on fractal graphs, *J. Funct. Anal.* **129**, 390 (1995).
- [75] L. Malozemov and A. Teplyaev, Self-similarity, operators and dynamics, *Math. Phys. Anal. Geom.* **6**, 201 (2003).
- [76] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. (John Wiley & Sons, Inc, New York, 1968).
- [77] L. Cocconi, G. Salbreux, and G. Pruessner, Scaling of entropy production under coarse graining in active disordered media, *Phys. Rev. E* **105**, L042601 (2022).
- [78] X. Zhao, D. E. Makarov, and A. Godec, Towards Markov-state holography, *New J. Phys.* **27**, 125002 (2025).
- [79] That is, the sequence of states of the embedded Markov chain.
- [80] É. Roldán and J. M. R. Parrondo, Estimating dissipation from single stationary trajectories, *Phys. Rev. Lett.* **105**, 150607 (2010).
- [81] Q. Yu and P. E. Harunari, Dissipation at limited resolutions: Power law and detection of hidden dissipative scales, *J. Stat. Mech.* **2024**, 103201 (2024).
- [82] B. Ertel and U. Seifert, Estimator of entropy production for partially accessible Markov networks based on the observation of blurred transitions, *Phys. Rev. E* **109**, 054109 (2024).

- [83] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, *Rev. Mod. Phys.* **48**, 571 (1976).
- [84] K. Brandner, Dynamics of microscale and nanoscale systems in the weak-memory regime: A mathematical framework beyond the Markov approximation, *Phys. Rev. E* **111**, 014137 (2025).
- [85] K. Brandner, Dynamics of microscale and nanoscale systems in the weak-memory regime, *Phys. Rev. Lett.* **134**, 037101 (2025).
- [86] R. F. Fox, Correlation time expansion for non-Markovian, Gaussian, stochastic processes, *Phys. Lett.* **94A**, 281 (1983).
- [87] R. F. Fox, Functional-calculus approach to stochastic differential equations, *Phys. Rev. A* **33**, 467 (1986).
- [88] P. J. Foster, J. Bae, B. Lemma, J. Zheng, W. Ireland, P. Chandrakar, R. Boros, Z. Dogic, D. J. Needleman, and J. J. Vlassak, Dissipation and energy propagation across scales in an active cytoskeletal material, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2207662120 (2023).
- [89] L. Vollmar, R. Bebon, J. Schimpf, B. Flietel, S. Celiksoy, C. Sönnichsen, A. Godec, and T. Hugel, Model-free inference of memory in conformational dynamics of a multi-domain protein, *J. Phys. A* **57**, 365001 (2024).
- [90] M. Schörner, S. R. Beyer, J. Southall, R. J. Cogdell, and J. Köhler, Conformational memory of a protein revealed by single-molecule spectroscopy, *J. Phys. Chem. B* **119**, 13964 (2015).
- [91] K. Song, R. Park, A. Das, D. E. Makarov, and E. Vouga, Non-Markov models of single-molecule dynamics from information-theoretical analysis of trajectories, *J. Chem. Phys.* **159**, 064104 (2023).
- [92] V. Salnikov, M. T. Schaub, and R. Lambiotte, Using higher-order Markov models to reveal flow-based communities in networks, *Sci. Rep.* **6**, 23194 (2016).
- [93] G. Nicolis and I. Prigogine, *Self-Organization in Non-equilibrium Systems*, 1st ed. (Wiley, New York, 1977).
- [94] H. Qian, S. Saffarian, and E. L. Elson, Concentration fluctuations in a mesoscopic oscillating chemical reaction system, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10376 (2002).
- [95] J. H. Fritz, B. Nguyen, and U. Seifert, Stochastic thermodynamics of chemical reactions coupled to finite reservoirs: A case study for the Brusselator, *J. Chem. Phys.* **152**, 235101 (2020).
- [96] That is, W is a lumping that refines the partition U .

End Matter

Appendix A: Overestimation of $\dot{S}$ with insufficient-order estimators—While we proved that the saturated-order estimator always provides a lower bound on the microscopic entropy production rate, this was *not* possible for estimators of insufficient order. It is thus natural to ask whether a too-low-order estimator may lead to an over estimation. Indeed, an overestimation occurs in the process shown in Fig. 1(a): the entropy production rate of the $n > 3$ order lumped process is clearly overestimated with $\dot{S}_2^{\text{est}}$. In particular, we see from Fig. 4 that the overestimation persists when increasing the number of

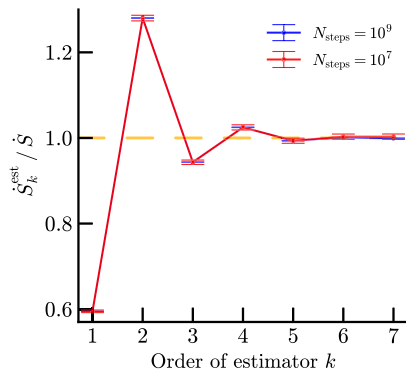


FIG. 4. Entropy production rate for the system in Fig. 1(a) estimated with $\dot{S}_k^{\text{est}}$ up to order $k = 7$ from lumping of 10^7 (red) and 10^9 (blue) microscopic steps. We see that the overestimation at $k = 2$ is *not* an undersampling artifact, showing the necessity of sufficiently high estimator order as proved in the present Letter. Each estimator was determined 30 times; we show the mean and reported empirical standard deviations (error bars).

steps (from $N_{\text{steps}} = 10^7$ steps to $N_{\text{steps}} = 10^9$). While we prove that $\dot{S}_k^{\text{est,uncond}}$ (see Supplemental Material [69]) is monotone in k , the estimator $\dot{S}_k^{\text{est}}$ is in particular *not*, as this example shows. However, it is necessary to use—as defined in Eq. (3)—the splittings rather than the joint probabilities, so that we have robustness to overestimation (see Supplemental Material [69]).

Appendix B: Problematic extrapolation to microscopic scales—Assuming that the power-law scaling of dissipation rate truly reflects dissipation on distinct length scales, one may attempt to extrapolate $\dot{S}_M \propto \lambda^{-\alpha}$ and $\dot{S}_{k \geq 2}^{\text{est}} \propto \lambda^{-\alpha}$ to microscopic scales, that is, to vanishing coarse graining $\lambda \rightarrow 1$ (see Ref. [49]). In the case of the cycle graph, which has a single (i.e., macroscopic) dissipative length scale, this would coincidentally (and conceptually somewhat paradoxically) yield the exact microscopic result [compare Eqs. (5) and (7) in the limit $\lambda \rightarrow 1$]. In all other examples, extrapolating $\dot{S}_M$ to $\lambda = 1$ would yield a lower bound. This may be useful for bounding $\dot{S}$ from below (the bound $\dot{S}_M \leq \dot{S}$ was obtained in [29]). However, there are strictly self-similar networks (e.g., the tree and tree-diamond graph) that do *not* display a power-law scaling of energy dissipation rate but where one may nevertheless (erroneously) identify a power law. It is conceivable that such examples may lead to an overestimation of $\dot{S}$. Extrapolated values should thus be interpreted with great care, especially when the underlying microscopic topology of the network is not known. Similarly, our examples in Figs. 3 and 5 show

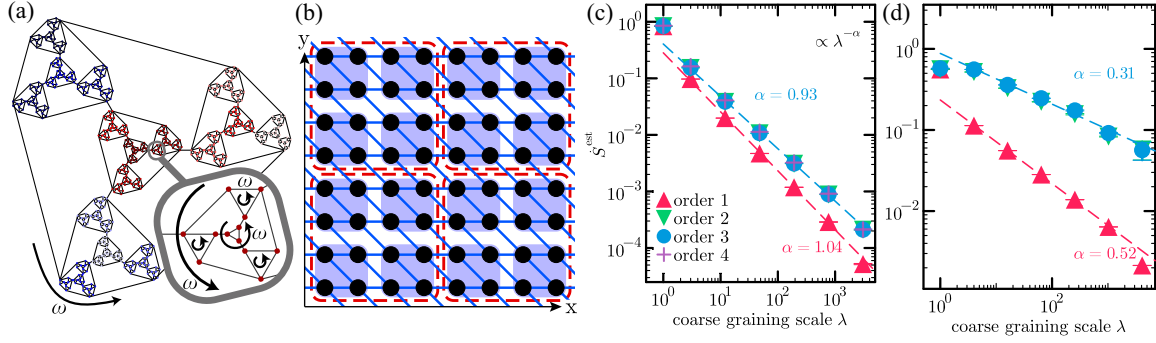


FIG. 5. Model systems and their entropy production rate estimation. (a) Self-similar Sierpinski-type graph with lumps of size $\lambda = 12$ indicated by the respective vertex coloring. A stationary current ω runs through the outer side of each polygon, whereby the size of said polygons depends on the respective recursive depth. The depicted graph has recursion depth 5 and $n = 768$ vertices. Inset: element of recursion-depth 2 with all irreversible currents indicated by arrows. (b) Brusselator depicted as a grid graph as in [49]. A portion of the total $n = 450 \times 450$ states is shown with lumps of size $\lambda = 2 \times 2$ (shaded in blue) and $\lambda = 4 \times 4$ (circled in dashed red). (c) $\dot{S}^{\text{est}}$ for Sierpinski-type graph for recursion depth 7 ($n = 12\,288$ vertices). We find $\dot{S}_M \propto \lambda^{-1}$ while higher-order estimators $\dot{S}_{k \geq 2}^{\text{est}}$ yield a power law with $\alpha = 0.93$, corroborating numerically that the lumped process is a second-order semi-Markov process. (d) $\dot{S}^{\text{est}}$ for the Brusselator with $n = 202\,500$ vertices. The virtual power law for the Markov estimate $\dot{S}_M^{\text{est}} \propto \lambda^{-0.52}$ is faster decaying than $\dot{S}_{k \geq 2}^{\text{est}} \propto \lambda^{-0.31}$ obtained by accounting for memory. For details, see Supplemental Material [69].

that one must also *not* extrapolate apparent power-law scalings of higher-order estimates $\dot{S}_{k \geq 2}^{\text{est}}$ as these may lead to overestimating $\dot{S}$.

We conducted numerical experiments to highlight the issue of extrapolating an apparent power law to determine the entropy production rate.

We consider a self-similar graph composed of nested polygons in the spirit of the Sierpinski fractal [see Fig. 5(a)]. The rates are set such that each nested polygon is driven with ω . Hence a polygon with recursive depth l has affinity $A_l = 3/\ln \omega$ (see derivation in [69]). The depth $l = 1$ corresponds to innermost triangles. The detailed construction is described in [69]. In contrast to the previous examples, this network is designed to indeed display a scale-dependent energy dissipation rate $\dot{S}$, whereby most energy is dissipated on the smallest and less energy on larger scales (for quantitative statements see Supplemental Material [69]). For networks with a recursive depth larger than 3, we are not able to determine P_i^s (and hence $\dot{S}$) analytically and therefore resort to numerical methods to compute $\dot{S}$ (see Supplemental Material [69] for details). The system with recursive depth 7 (12 288 microscopic states) was coarse-grained into lumps with recursive depth l , such that the lumped graph perfectly preserves the original fractal structure [see Fig. 5(c)]. However, in stark contrast to the previous examples, the coarse graining here hides dissipative cycles. We determine $\dot{S}_M$ via Eqs. (2) and (3) (with $k = 1$) and $\dot{S}_{k \geq 2}^{\text{est}}$ using Eq. (3) from 108 simulated steady-state trajectories each with a total duration 5×10^8 steps.

Despite having properties that are “orthogonal” to the previous examples—featuring a scale-dependent energy dissipation rate and dissipative cycles becoming hidden by the lumping—it *resembles* the same power-law scaling of the

Markovian estimate $\dot{S}_M \propto \lambda^{-\alpha}$ [red line Fig. 5(c)] in fact with the same exponent $\alpha \approx 1$ as in Eq. (7). This shows once more that a power-law scaling of $\dot{S}_M$, and in particular also the exponent, does *not* give any indication about dissipation on different length scales. Moreover, the exponent α changes to ≈ 0.9 upon accounting for memory via $\dot{S}_{k \geq 2}^{\text{est}}$ [blue line Fig. 5(c)], and the lumping yields a second-order semi-Markov process, since adjacent lumps are connected by exactly one microscopic transition (see proof in [69]). Accordingly, the numerical results show $\dot{S}_1^{\text{est}} \neq \dot{S}_2^{\text{est}} = \dot{S}_3^{\text{est}} = \dot{S}_4^{\text{est}}$. The fact that $\dot{S}_{k \geq 2}^{\text{est}} < \dot{S}$ (i.e., equality not reached) is because dissipative cycles are hidden by the coarse graining.

Finally, let us consider the Brusselator model [93–95] [see Fig. 5(b)] as in [49], here with 202 500 microscopic states. We perform the coarse graining as was done in [49], two levels of coarse graining are highlighted in Fig. 5(b). We estimate $\dot{S}$, $\dot{S}_M$, and $\dot{S}_{k \geq 2}^{\text{est}}$ from 25 simulated stationary trajectories each having 10^9 steps (skipping 10^9 initial steps). Using Eq. (3) with $k = 1$ we infer $\dot{S}_M$ that agrees with the results in [49] [red triangles in Fig. 5(d)]. In particular, we reproduce the apparent power-law scaling $\dot{S}_M \propto \lambda^{-\alpha}$ with the same exponent as in [49], i.e., $\alpha \approx 0.52$. Coincidentally, it is of the same order as the *virtual* power law of the tree and tree-diamond with $\alpha \approx 0.3$ (see Supplemental Material [69]), which in reality is *not* a power law at all [see Eq. (10)].

We now account for memory in the time reversal via $\dot{S}_{k \geq 2}^{\text{est}}$. The lumped process is apparently second- or third-order semi-Markov as $\dot{S}_{2,3}^{\text{est}} = \dot{S}_4^{\text{est}}$. The fact that $\dot{S}_{k \geq 2}^{\text{est}} < \dot{S}$ is a consequence of dissipative cycles that become hidden by the coarse graining. Higher-order estimates also display an apparent power-law scaling albeit with a smaller exponent

$\dot{S}_{\geq 2}^{\text{est}} \propto \lambda^{-0.31}$ for considered ω, κ . However, according to previous examples, we must conclude that these apparent power laws in fact have *no* implications about energy dissipation on different length scales (see Sec. IX in Supplemental Material [69]).

Appendix C: Nonexistence of scaling of $\dot{S}_M$ under timescale separation—We already showed that a power-law scaling of $\dot{S}_M$ has no physical implications in absence of a timescale separation due to inconsistent time reversal. In the presence of a timescale separation the time reversal becomes consistent [29], but the power-law scaling disappears, as we now show.

Suppose there is a timescale separation, i.e., suppose the dynamics within each element (lump) of a lumping U relaxes fast enough for the transitions *between* elements of U to proceed as in Ref. [29] and (consistently) apply a Markovian estimator $\dot{S}_1^{\text{est}}$. We now argue that timescale separation in U implies timescale separation in each refined scale. Consider W to be a refinement of U [96]. Because of timescale separation in U any transition between and within lumps of W happens at least as fast as in U . Hence, we also have timescale separation in W .

Hence, in presence of timescale separation, the Markovian time reversal is consistent, and the Markovian estimate gives the exact result [29] for scale U and any refined scale. This naturally implies the absence of (power-law) scaling of $\dot{S}_1^{\text{est}}$.

To illustrate this, consider the cycle graph C_{36} with rates chosen to have timescale separation (the detailed parameters are given in Fig. S8 in [69]). At a lumping of size 6 (first lump are vertices 1, ..., 6, second lump vertices 7, ..., 12, etc.), we observe a timescale separation. Thus, we may use a Markovian estimator for a lumping of size 6. To get the correct entropy production rate at a lumping of

size 12, however, we must use an estimator $\dot{S}_k^{\text{est}}$ of order $k \geq 2$ (as argued in the Letter). As we see in Fig. 6, any refinement of the lumping of size 6 will be Markovian *and result in the same estimate*. Thus, there is no power-law scaling under timescale separation.

Appendix D: Uncountability of microscopic paths—We finally show that the number of microscopic paths projecting onto a mesoscopic path in any time window $[0, t]$ (with $t > 0$) is generally *not* countable. To that end, suppose that a mesostate M contains at least two neighboring microscopic states $\{0, 1\}$. Take some real number $r \in [0, t)$, and consider its binary encoding, say $\text{bin}(r)$. Take the corresponding microscopic trajectory of jumping from 0 to 1 at time r . This is a unique mapping to a microscopic path, but all those paths correspond to the mesoscopic path of staying in M . By uncountability of $\mathbb{R}$, the claim follows.

Appendix E: Other parameter choices—We present in Table I how other choices of u, v in Fig. 2 affect the estimators' quality: Any $k \geq 2$ is better than the Markovian estimate $\dot{S}_M = \dot{S}_1^{\text{est}}$ [49]. We always observe saturation for $k \geq 3$ (up to statistical fluctuations) and our estimate is already at $k = 2$ better than $\frac{1}{2} \dot{S}_2^{\text{uncond}} / \dot{S}$ [81]. Note how saturation happens for $u = 1$ already at $k \geq 2$, indicating that at $u = 1$ we have an order-2 process: While $\dot{S}_2^{\text{est}}$ allows for 100% recovery, using in that setting $\frac{1}{2} \dot{S}_2^{\text{uncond}} / \dot{S}$ [81] would allow only for recovery of 82% and 83% respectively, with the remaining recovery only possible through the statistically very costly (binning of a continuous-time random variable for integral computation) inference from the waiting-time distribution asymmetry. Note that σ_2^{ℓ} in Ref. [81] is—in contrast to σ_2^t —statistically equally feasible as $\dot{S}_2^{\text{est}}$.

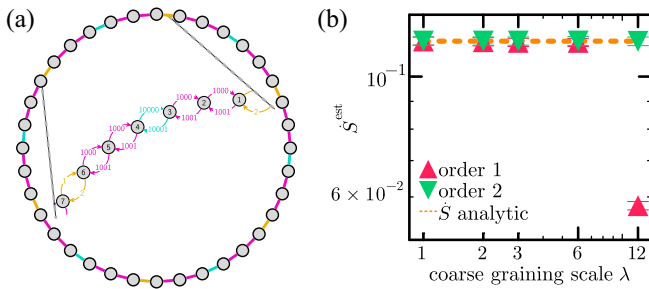


FIG. 6. (a) Cycle graph C_{36} with transition rates chosen as indicated in the inset and detailed in Supplemental Material [69] (Sec. XIII) such that the dynamics exhibit timescale separation. (b) Entropy-production rate estimates for C_{36} with timescale separation between lumps of size 6: the timescale separation assumption implies that any refined scale is Markovian ($\dot{S}_1^{\text{est}} = \dot{S}_2^{\text{est}}$). This also implies, however, that there is *no* scaling of the entropy production rate (observation).

TABLE I. Ability of different estimators under more parameter choices for the system of Fig. 2: we observe a consistently better approximation of $\dot{S}$ through $\dot{S}_k^{\text{est}}$ for every $k \geq 2$ compared to $\dot{S}_M = \dot{S}_1^{\text{est}}$ [49] and compared to $\dot{S}_2^{\text{uncond}}/2 = \frac{1}{2}(\dot{S}_1^{\text{est}} + \dot{S}_2^{\text{est}}) = \sigma_2^{\ell}$ in [81].

Parameter u	2.0	1.0	2.0	1.0	1.5
Parameter v	2.0	2.0	1.0	3.0	2.0
$\dot{S}_1^{\text{est}}/\dot{S}$ [49]	0.03	0.58	0.06	0.63	0.14
$\dot{S}_2^{\text{est}}/\dot{S}$	0.21	1.05	0.09	1.03	0.45
$\frac{1}{2} \dot{S}_2^{\text{uncond}}/\dot{S}$ [81]	0.12	0.82	0.07	0.83	0.30
$\dot{S}_3^{\text{est}}/\dot{S}$	0.39	1.02	0.20	1.02	0.61
$\dot{S}_4^{\text{est}}/\dot{S}$	0.41	1.02	0.22	1.02	0.63
$\dot{S}_5^{\text{est}}/\dot{S}$	0.41	1.00	0.21	1.01	0.62
$\dot{S}_6^{\text{est}}/\dot{S}$	0.4	0.99	0.21	1.00	0.61