PCCP

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2026, **28**, 9840

Identification and quantification of irreversibility in stochastic systems

Aishani Ghosal [ab] and Gili Bisker *[cdefgh]

Advances in single-molecule measurements, active-matter control, and nonequilibrium statistical mechanics are transforming our understanding of thermodynamics in small, strongly fluctuating systems. Biological molecular motors, driven chemical-reaction networks (*e.g.*, gene regulation), artificial active matter, autonomous engines, and synthetic nanomachines all operate *via* inherently irreversible, dissipative processes in noisy environments, while producing entropy. Quantifying this entropy production (EP) has therefore become central to understanding both the physical limits and design principles of nanoscale systems. This review surveys principled routes to characterize and quantify EP from time-series and trajectory data. Because experimental observables are often coarse-grained and only partially resolve the underlying dynamics, we discuss how dissipation can be inferred from incomplete information, and how coarse-graining systematically biases EP estimates. This overview maps the current toolkit for estimating EP and outlines open challenges in unifying inference approaches to obtain reliable and tight bounds on EP in living and engineered nanoscale systems.

Received 4th December 2025,
Accepted 23rd March 2026

DOI: 10.1039/d5cp04712a

rsc.li/pccp

1 Introduction and overview

The concept of irreversibility, a cornerstone of modern thermodynamics, was not present in classical thermodynamics, which emerged in the 19th century through the foundational work of Carnot, Clausius, Joule, Helmholtz, Kelvin, and Gibbs. Classical thermodynamics primarily addressed equilibrium states and quasi-static processes, focusing on energy exchange without accounting for time-dependent or directional (arrow of time) processes. In contrast, the formal theory of irreversible processes was developed in the 20th century by Lars Onsager,[1,2] Théophile De Donder,[3] Ilya Prigogine,[4–6] and others, laying the foundation for modern thermodynamics.[4,5,7] This framework introduced linear irreversible thermodynamics to describe systems close to equilibrium and was later extended to capture the behavior of systems operating far from equilibrium.[8]

Modern thermodynamics provides a basis for understanding how highly ordered structures can arise and persist in nature—phenomena that classical frameworks could not explain. Irreversible processes, such as those involved in biological evolution, are fundamental to the emergence and maintenance of such structures. These processes can both generate[9] and degrade order,[10] with sustained order requiring continuous dissipation of energy. In this sense, irreversibility is intimately linked to the flow of energy and the arrow of time. To further generalize the thermodynamics of small-scale, highly fluctuating systems, the field of stochastic thermodynamics has emerged as a sub-discipline of nonequilibrium statistical physics at the end of the twentieth century.[11] It extends the principles of thermodynamics, which were originally formulated for ensemble-averaged systems near equilibrium, to individual trajectories of systems operating far-from-equilibrium.[12,13] Early developments focused on steady-state systems[14,15] and ensemble averages,[16] leading to the discovery of key results such as fluctuation theorems.[17–23]

A significant advancement in this field is the formulation of thermodynamics at the trajectory level. In the framework of stochastic energetics, introduced by Sekimoto,[11] thermodynamic quantities such as work, heat, and entropy production are defined along single trajectories. Within this framework, even the first and second laws of thermodynamics have been recast in terms of path-dependent quantities.[24,25] These theoretical predictions have been experimentally verified through

[a] *School of Chemical Sciences, National Institute of Science Education and Research, Khurdha, Jatni Rd, Bhubaneswar 752050, India. E-mail: aishanig@niser.ac.in*
[b] *Department of Physics, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO, USA*
[c] *School of Biomedical Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: bisker@tauex.tau.ac.il*
[d] *Center for Physics and Chemistry of Living Systems, Tel Aviv University, Tel Aviv 6997801, Israel*
[e] *Center for Nanoscience and Nanotechnology, Tel Aviv University, Tel Aviv 6997801, Israel*
[f] *Center for Light-Matter Interaction, Tel Aviv University, Tel Aviv 6997801, Israel*
[g] *Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 6997801, Israel*
[h] *Sagol School for Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel*

advances in single-molecule manipulation and detection techniques.[26–30] More recently, thermodynamic uncertainty relations (TUR)[31–34] have emerged, offering fundamental bounds on the trade-offs between precision and dissipation in fluctuating systems. These bounds have been extended to a wide range of systems, including those with multiple time-integrated observables,[35] underdamped dynamics,[36] broken time reversal symmetry,[37] and non-Markovian (memory-retaining) behavior.[38] Additional relations, such as the dissipation-time uncertainty relation[39] and energy-efficiency trade-offs,[40] have proven particularly relevant for analyzing and constraining the performance of biological and artificial nanoscale systems.

The framework of stochastic thermodynamics has been successfully applied across a wide range of systems to quantify dissipation. In biological systems, dissipation has been investigated in contexts such as Darwinian dynamics,[41] biochemical reaction networks,[42] self-replication,[43,44] active cytoskeleton material,[45] population dynamics,[46] ecological systems, sensory adaption,[47] and even neural[48] and brain function.[49] Beyond biology, dissipation estimates have been crucial in understanding the thermodynamic costs of information processing,[47,50–54] including contexts involving Shannon information,[55] information geometry,[56,57] and information ratchets.[58–61] Studies have also extended to the thermodynamic cost of computation and cognition, including analyses of Turing machines.[62–65] Additionally, stochastic thermodynamics has been applied to study dissipation in stochastic transport and optimal transport theory,[66,67] uncovering profound connections between physical cost and the geometry of probability flows. Collectively, these applications highlight the versatility of stochastic thermodynamics in quantifying dissipation across a wide range of systems. Together, these diverse applications reveal the versatility of stochastic thermodynamics across biological, informational, and physical domains.

Both natural and synthetic nanoscale systems face the same fundamental energetic constraints: dissipation sets limits on precision and performance. This unifying perspective makes dissipation not just a cost, but a design principle. Small-scale living systems, such as biological molecular machines, function in highly fluctuating environments and rely on precise energy management to maintain robust performance. The thermodynamic cost of these processes—often quantified by dissipation—plays a critical role in determining how such systems can perform reliably under thermal noise,[68] maintain a target nonequilibrium state,[69] and other stochastic influences. Understanding dissipation is therefore essential for uncovering the design principles of molecular machines in biology and for guiding the development of efficient artificial Brownian motor[70] or nanoscale devices. Recent advances in nanotechnology have intensified interest in dissipation, as researchers strive to engineer synthetic molecular machines capable of performing specific tasks, such as molecule synthesis, cargo transport, and vesicle fusion.[71] These artificial systems, much like their biological counterparts, must operate under tight energetic constraints, making dissipation not only a measure of inefficiency but also a design constraint for functionality and control. As such, dissipation has become a key quantity in the design, optimization, and theoretical understanding of both natural and engineered molecular systems.

A significant challenge in analyzing small-scale systems lies in the limited spatiotemporal resolution of experimental techniques, particularly in single-molecule measurement.[72] This limitation makes it difficult to access all relevant, driven degrees of freedom, resulting in only partial observations of the system's true dynamics. Consequently, the observed dynamics are inherently coarse-grained, representing a projection of the full microscopic evolution onto a lower-dimensional space.[73,74]

In molecular simulations, the term coarse-graining typically refers to a modeling strategy in which several atoms are grouped into a single effective particle. This reduction eliminates fine-scale degrees of freedom and interactions that are deemed irrelevant at the scale of interest. By smoothing the configurational landscape and reducing the dimensionality of the phase space, such coarse-grained models significantly accelerate molecular dynamics simulations of complex systems. The coarser the representation, the greater the computational gain. This form of coarse-graining is widely used to enable simulations over longer time and length scales. However, the notion of coarse-graining considered in this review does not arise from a deliberate modeling simplification, but from limited observational access to the underlying dynamics. Many nonequilibrium systems, particularly biological ones, operate across multiple timescales, and experimental techniques often cannot resolve all relevant microstates or transitions. As a result, certain degrees of freedom remain hidden, and the observed dynamics represent only a projection of the full microscopic process. Thus, coarse-graining reflects partial information due to finite spatiotemporal resolution or intrinsically unobservable states, rather than a reduction performed for computational efficiency.

In this context, coarse-graining reflects partial observational access to the system's dynamics: hidden degrees of freedom, unresolved transitions, or finite spatiotemporal resolution effectively project the full microscopic evolution onto a reduced description. Such observational coarse-graining introduces fundamental discrepancies between the observed dynamics and the underlying microscopic process, with direct consequences for the thermodynamic inference of irreversibility. In particular, dissipation estimated from coarse-grained data often underestimates the true entropy production, due to the loss of information about hidden or unobserved degrees of freedom.[75] Understanding how coarse-graining affects dissipation has thus become a critical area of investigation in stochastic thermodynamics. Various methods have been developed to model coarse-grained systems, often differing in the quantities preserved during the coarse-graining procedure—such as total probability densities,[76] fluxes,[74] or transition rates.[77] Recent efforts have focused on formulating coarse-graining frameworks that preserve thermodynamic consistency[78,79] and enable reliable inference of dissipation from partial data.[80] This has led to ongoing theoretical advancements aimed at extending stochastic thermodynamics to partially observed systems, bridging the gap between experimental accessibility and microscopic thermodynamic laws.

In this review, we discuss recent advances in the detection and quantification of irreversibility in small-scale systems, with a particular focus on the effects of coarse-graining on dissipation within the framework of stochastic thermodynamics. The review proceeds as follows: Section 2 outlines criteria for irreversibility; Section 3 examines dissipation estimation for fully resolved systems near equilibrium; Section 4 addresses the impact of coarse-graining and partial observations; Section 5 surveys theoretical and experimental advances in dissipation quantification; and Section 6 concludes.

2 Qualitative and quantitative criteria for nonequilibrium

Systems governed by time-independent conservative potentials relax to thermal equilibrium, characterized by a time-independent probability distribution and vanishing probability currents between all pairs of states, *i.e.*, detailed balance.[81] Time-independent non-conservative driving breaks detailed balance and, in the long-time limit, leads to a non-equilibrium steady state (NESS) characterized by a stationary probability distribution with non-zero steady-state probability currents. Under explicitly time-dependent forces, however, the system generally does not reach a steady state, and the probability distribution remains time-dependent. In this section, we focus on criteria for identifying and characterizing NESS.

2.1 Broken detailed balance and time reversal symmetry

The principle of detailed balance (DB), which serves as a microscopic basis for thermodynamics, was identified by Ludwig Boltzmann.[82] Systems at thermodynamic equilibrium are considered to follow DB, or the Kolmogorov criterion.[83] DB refers to the reversible and pairwise balance of transition rates between any two discrete microstates, therefore, the ratio of products of the transition rates along the clockwise direction to counterclockwise direction equals unity. The latter results in a vanishing net transition flux or current. At NESS, in contrast, the probability distribution becomes time-independent, yet detailed balance is violated and persistent probability currents circulate through state space.

If all microstates are experimentally accessible, distinguishing between equilibrium and NESS becomes trivial, provided sufficiently long trajectories are available. At equilibrium, the following DB equality holds at steady state:

$$P_i^{ss} k_{ji} = P_j^{ss} k_{ij} \quad \forall i, j \neq i \tag{1}$$

where the stationary probabilities at state i are denoted P_i^{ss}, which can be obtained from the fraction of time the system spends at a particular state, given a long trajectory, and the transition rate from state i to j is denoted as k_{ji}. Under this condition, all pairwise probability currents vanish. By contrast, at NESS, detailed balance is broken and closed loops of probability flux persist.[84] The stationary probability density is maintained by constant probability currents, reflecting sustained entropy production despite time-independent statistics. Fig. 1(a) illustrates the detailed balance criterion (top) and the broken detailed balance leading to net transition flux (bottom).

The presence of a time-dependent or external non-conservative force or field drives a system out of equilibrium, thus, detailed balance or time-reversal symmetry is broken. Examples of violations of detailed balance during active processes in biological systems have been reported.[49,85,90–96] For broken time reversal symmetry, the probability distribution of having a time forward series does not equal the probability distribution of having the time-reversed series. Mathematically, if $P[\{x_\tau\}_{0\le\tau\le t}]$ represents the probability density of having a time-series for a duration of time t of a conformational or chemical degrees of freedom (x) at time τ, then $\tilde{P}[\{\theta x_\tau\}_{0\le\tau\le t}]$ would be the time-reversed probability distribution function, where $\theta = \pm 1$ depending on whether the variable is even or odd under time reversal. For an even variable, like position, $P[\{\theta x_\tau\}_{0\le\tau\le t}] = \tilde{P}[\{x_{t-\tau}\}_{0\le\tau\le t}]$. Broken time reversal symmetry can also be measured by an asymmetry factor, $A_{i,j}^{\tau}$,[97] defined by

$$A_{i,j}^{\tau} = \left| \ln\left(\frac{P(i,\tau|j) P_{ss}(j)}{P(j,\tau|i) P_{ss}(i)} \right) \right|, \tag{2}$$

where $P_{ss}(i)$ is the steady state probability density function at state i and $P(i,\tau|j)$ is the conditional probability density that the system is at state i at time τ given that initially the system was at steady state j.

In partially observed systems, some of the net transition fluxes may be obscured, unobserved, or inaccessible. For example, protein copy number inside a living cell as a function of time can be accessed experimentally using live-cell imaging, but whether the corresponding gene is in an active or an inactive state cannot be accessed experimentally.[98] Although gene activity influences protein copy number, other sources of stochastic noise from transcription and translation also affect copy frequency. As a result, protein copy number alone cannot be used to determine whether a gene is active or inactive. That scenario may lead to an illusion that a nonequilibrium system is at equilibrium, and one needs to rely on the broken time-reversal symmetry in the available, coarse-grained, data to infer irreversibility.

If the coarse-graining commutes with time-reversal symmetry, broken time-reversal symmetry can be quantified *via* the mean-dwell time asymmetry factor (MDAF) for a partially observed system even with a vanishing net current. These systems exhibit second-order semi-Markov statistics.[86] † In this

† A stochastic process is Markovian if transitions to future states depend only on the present state. In other words, if the system is currently in state i, the probability of transitioning to another state j depends solely on being at i, and not on previously visited states. In contrast, coarse-graining a Markov process generally induces non-Markovian dynamics, in which the future evolution of the observed system may depend not only on its current state but also on its past trajectory. Such memory-dependent dynamics can often be represented as semi-Markov processes of finite order. In an n-th order semi-Markov process, the probability of the next transition may depend on the current state, the time already spent in the current state, and all the previous $n-1$ visited states. As the order increases, more historical information is required to determine future evolution. For example, in a second-order semi-Markov process, if the system is currently in state j and previously visited state i, the present state is represented by the ordered pair $[i,j]$. The transition from j to a subsequent state k then depends both on the residence time in state j and on the identity of the preceding state i.

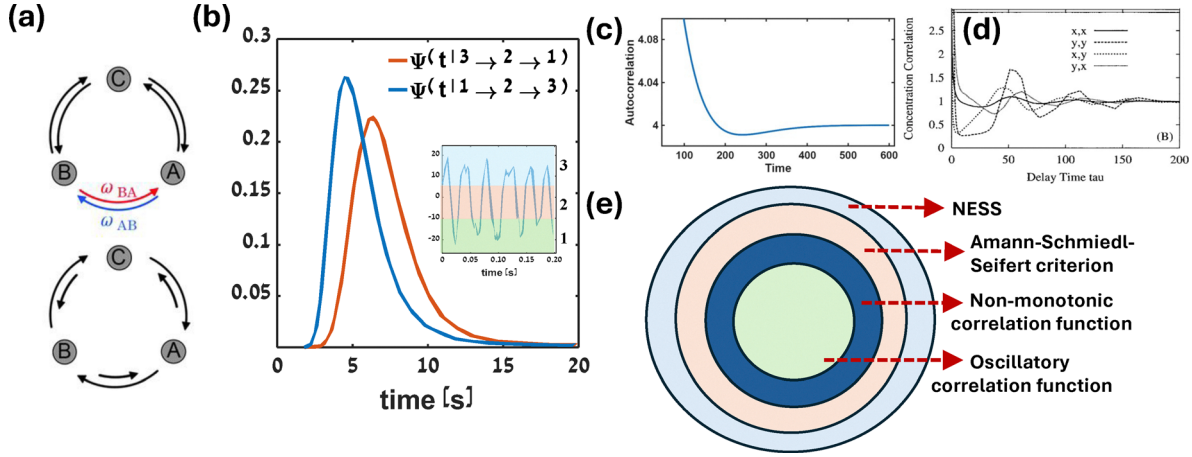


Fig. 1 Criteria for irreversibility detection: (a) detailed balance with zero pairwise flux (top) and broken detailed balance with a non-zero net flux (bottom). This figure has been reproduced from ref. 85 with permission from the American Association for the Advancement of Science, copyright 2026. (b) An example of broken time-reversal symmetry, where a continuous variable is discretized into three states (as shown in the inset with different shaded colors, which are labeled by 1, 2, and 3) following a commonly used spatial coarse-graining of a positional degree of freedom x_1, plotted as a function of time. The underlying process in the discretized space is modeled as a second-order semi-Markov process, and the asymmetry in the distribution of waiting times at state 2 for an upward transition (1 → 2 → 3) *versus* a downward transition (3 → 2 → 1) reflects the broken time-reversal symmetry. This figure has been reproduced from ref. 86 with permission from The Royal Society of Chemistry, copyright 2026. (c) A non-monotonic time autocorrelation function for a three-state kinetic cycle, indicative of a non-equilibrium steady state. This figure has been adapted from ref. 87 with permission from the American Physical Society, copyright 2026. (d) An oscillatory correlation function exhibited by an open chemical or biochemical reaction network. This figure has been reproduced from ref. 88 with permission from the National Academy of Sciences, copyright 2026. (e) A schematic view of the expansion of the parameter space defining a non-equilibrium steady state: the subset of system parameters for which the system exhibits an oscillatory correlation function lies within the parameter regime where the time autocorrelation function of system variables shows non-monotonic behavior. Furthermore, this regime is itself a subset of the parameter range identified by the Amann–Schmiedl–Seifert criteria or their generalizations. This figure has been adapted from ref. 89 with permission from the IOP Publishing Ltd copyright 2026.

framework, the state of the system is represented by an ordered doublet $[i,j]$, indicating that the system is currently in state j and previously visited state i. The MDAF is defined as the ratio of the time associated with transitions between doublets of states (*e.g.*, from $[i,j]$ to $[j,k]$ for the transition $i \to j \to k$) to the time associated with the reverse transition $k \to j \to i$.[99–101] Fig. 1(b) illustrates time-reversal symmetry breaking in a system coarse-grained into three discrete states (see inset): 1 (green shaded area), 2 (orange shaded area), and 3 (blue shaded area). The probability distribution of the waiting time at state 2 before transitioning to state 3 in the forward process (1 → 2 → 3) does not coincide with the distribution of the waiting time at state 2 in the reverse process (3 → 2 → 1).

2.2 Violation of fluctuation dissipation theorem

Systems at equilibrium follow a fluctuation dissipation relation (FDR) as derived by Kubo,[102] which links the response of a system to an infinitesimal perturbation to the statistics of its spontaneous equilibrium fluctuations. When a system with an unperturbed Hamiltonian H_0 is perturbed by a force F, the perturbed Hamiltonian is given by $H(\mathbf{x},t) = H_0(\mathbf{x}) - F(t)A(\mathbf{x})$, where A is the conjugate variable corresponding to the force. The response or the average variation of another generic observable B that is coupled to the perturbed observable A is given by

$$R_{BA}(t) = -\beta \frac{\mathrm{d}C_{BA}(t)}{\mathrm{d}t}, \tag{3}$$

where $R_{BA}(t)$ is the response function and C_{BA} denotes the correlation function between two observables, A and B, and $\beta = \frac{1}{k_B T}$ in the inverse temperature. Eqn (3) is valid for a system at equilibrium. However, a violation of eqn (3) is an indication for an out-of-equilibrium system.

Originally developed within the framework of Hamiltonian statistical mechanics, the FDR has since been generalized to broad conditions, independent of whether the system is Hamiltonian or in equilibrium. There are several other works on the generalization of the fluctuation–dissipation relation for out-of-equilibrium systems.[103–106] The fluctuation–dissipation theorem (FDT) states that the linear response of a system to an external perturbation can be expressed in terms of the system's equilibrium fluctuations. For systems governed by Langevin dynamics, Kubo showed that the FDT relates the frequency-dependent friction force to the (non-white) random force.[106] A general linear response formula for a system under time-dependent perturbation was formulated at a NESS, which remains valid far from equilibrium.[105] A recent review has also discussed FDR in the context of granular systems, nanosystems, and biological systems.[104]

Notably, violation of FDT has been found in a range of systems and processes.[107–112] Harada and Sasa proposed an equality between the extent of violation of the FDR in the NESS and the rate of energy dissipation into the environment.[113]

Therefore, the violation of FDR had been used to infer the energy dissipation.[113–115]

2.3 Non-monotonic time correlation function

Time correlation function of a quantity Ω is defined as follows

$$C(\tau) = \langle \Omega(t+\tau)\Omega(t) \rangle \quad (4)$$

The angular brackets refer to averages with respect to time dependent probability density function, *i.e.* $\langle \Omega(t+\tau)\Omega(t) \rangle = \int d\Gamma \Omega(\Gamma; t+\tau)\Omega(\Gamma; t) f(\Gamma, t)$, where $f(\Gamma,t)$ represents the time dependent probability density function of variable Γ at time t. The decay of a correlation function provides information on whether a system is at equilibrium or far from equilibrium.[87] The presence of oscillation or non-monotonic decay of correlation function[87] is an indication of a far-from-equilibrium process, whereas monotonic decay of correlation functions without any oscillations denotes equilibrium. In terms of the eigenvalues of the transition rate matrix,‡ real eigenvalues signify equilibrium, and complex eigenvalues denote otherwise.[116,117] Qian *et al.* showed that for a pumped reaction, the eigenvalues are complex.[87] Similarly, oscillations in the correlation function of the two-state trajectory signify an underlying nonequilibrium system with several microstates lumped into two coarse-grained states.[87]

Let us consider a variable c at time t. For systems that reach a non-equilibrium steady state (NESS), the time-correlation function $\langle c(t)c(t+\tau) \rangle$ exhibits non-monotonic or oscillatory behavior. The angular brackets denote averages over the probability distribution function of the variable at time t. A non-monotonic Fig. 1(c) and oscillatory Fig. 1(d) correlation function is an indication for underlying non-equilibrium behaviour. Fig. 1(d) illustrates how concentration fluctuations of two chemical species (namely x and y) are correlated at time t. The normalized correlation function is defined as $G_{ij}(\tau) = \frac{\langle c_i(t)c_j(t+\tau) \rangle}{\langle c_i \rangle \langle c_j \rangle}$, where $i,j = x,y,z$.

2.4 Amann–Schmiedl–Seifert criteria and its generalization for NESS

Amann, Schmiedl, and Seifert[118] identified an alternative criterion for NESS in a three-state system where only one state is observed and the remaining two states are merged into a single mesostate. They proposed a criterion based on observables such as the waiting-time distributions in the observed state and the merged state, as well as the time-dependent probability of being in the observed state. All these quantities can be expressed in terms of the transition rates among the microstates of a given network topology. After rescaling these quantities, conditions leading to negative transition rates are discarded, and violations of the detailed-balance condition are identified as signatures of a NESS. The Amann–Schmiedl–Seifert criterion suggests that if the steady-state probability of occupying state exceeds a certain threshold, the system resides in a NESS. In other words, if the sum of the weights of all spanning trees directed into the 'on' state—or equivalently, the time spent in the 'on' state—exceeds this bound, then the coarse-grained two-state Markov system will be in a NESS. Therefore, the criterion sets a lower bound for when a three-state system, coarse-grained into two states, can be in a NESS. Notably, the region of parameter space corresponding to a NESS under this criterion is larger than that obtained from the oscillatory correlation function. Subsequently, Jia and Chen[89] proposed another criterion for NESS based on an observed coarse-grained two-state trajectory arising from an underlying three-state Markov system. They provided a mathematical derivation and a probabilistic interpretation of the Amann–Schmiedl–Seifert criterion, and demonstrated that it also captures non-monotonic or oscillatory correlation functions. Wu *et al.* formulated theoretical framework that detects a nonequilibrium criteria based on coarse-grained observation for a general Markov network with an arbitrary number of microstates and an arbitrary coarse-grained partitioning, given long enough statistics of the trajectory.[119] The Amann–Schmiedl–Seifert criterion detects a larger NESS region in terms of system parameter values compared to the oscillatory correlation function.[118]

Fig. 1(e), adapted from ref. 89, schematically illustrates the NESS criteria, which shows that the parameter space for a system to be at NESS in terms of the transition rates among the microstates of the full netowrk increases for Amann–Schmiedl–Seifert criterion, and a subset of this parameter space is applicable for the non-monotonic correlation function, and even smaller parameter space is applicable for the oscillatory correlation function. Although, the Amann–Schmiedl–Seifert criterion identifies NESS regimes that cannot be accessed using criteria based on oscillatory or non-monotonic correlation functions, it can also capture the parameter space for NESS which is traced from oscillatory or non-monotonic correlation function.

2.5 Other methods

Mori *et al.* showed that if a stochastic process defined by the variable x as a function of τ, $x(\tau)$, over a time interval $[0,T]$ $(0 \leq t_m \leq T)$, is at equilibrium, then the probability distribution of $P(t_m|T)$ would be symmetric with respect to the time t_m at which x reaches its global maximum value.[120] The area enclosing rate (AER) is another quantity to detect the nonequilibrium nature of a process from two or more particles in a system.[121] The advantage of this quantity is that it can be computed using only two degrees of freedom.[121] A three-time-point-position correlation, known as mean back relaxation (MBR), have been recently used to detect broken time reversal symmetry in confinement.[122] MBR correlates the displacement of a particle between two given time points with its displacement in a prior time period with a finite distance cut-off. For microscopic densities as stochastic observables, the deviation of long time value of MBR from 1/2 is an indication of the nonequilibrium nature of the process.[123]

‡ The off-diagonal elements of the transition rate matrix, q_{ij}, are nonzero and denote the rate of transitioning from state j to state i per unit time. Each column of the transition rate matrix sums to zero.

3 Entropy production in fully observed systems

3.1 Systems close to equilibrium

Modern thermodynamics provides a unified framework that connects entropy, a central thermodynamic quantity, to irreversible processes occurring in nonequilibrium systems.[12,24,124,125] In this framework, the total entropy change over an infinitesimal time interval, $\mathrm{d}S$, is expressed as the sum of two contributions: one due to exchange with the environment and another due to internal irreversible processes. The exchange term, often written as $\mathrm{d}_\mathrm{e}S$, accounts for entropy flow resulting from the transfer of heat or matter, *e.g.*, $\mathrm{d}_\mathrm{e}S = \mathrm{d}Q/T$ for heat exchange. The second term, $\mathrm{d}_\mathrm{i}S$, represents the irreversible entropy production (EP) within the system and is strictly non-negative, in accordance with the second law of thermodynamics.

Irreversible processes such as heat conduction, diffusion, and chemical reactions generate entropy and are characterized as thermodynamic flows driven by corresponding thermodynamic forces. The entropy production rate (EPR), $\mathrm{d}_\mathrm{i}S/\mathrm{d}t$, can thus be formulated in terms of the product of these flows and forces, providing a quantitative description of nonequilibrium dynamics.[81]

3.2 EP for systems governed by master equations

In systems described by a continuous-time Markov chain over a discrete set of states, the dynamics are governed by a master equation (ME):[126]

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \sum_j \left[w_{ij}p_j(t) - w_{ji}p_i(t)\right], \tag{5}$$

where $p_i(t)$ is the probability of being in state i at time t, and w_{ij} is the transition rate from state j to i.

For a stochastic trajectory, $\gamma = \{i_0 \to i_1 \to \cdots \to i_N\}$, occurring over the time interval $[0,\tau]$, the total entropy production ΔS_tot (see Table 1 for a complete list of EP-related symbols) along the trajectory can be written as:

$$\Delta S_\mathrm{tot}[\gamma] = \ln\frac{P[\gamma]}{\tilde{P}[\tilde{\gamma}]}, \tag{6}$$

where $P[\gamma]$ is the path probability of the trajectory and $\tilde{P}[\tilde{\gamma}]$ is that of the time-reversed trajectory ($\tilde{\gamma}$).

Expanding the path probabilities yields:

$$\Delta S_\mathrm{tot} = \ln\frac{p_{i_0}(0)}{p_{i_n}(\tau)} + \sum_{k=1}^{N}\ln\frac{w_{i_k i_{k-1}}}{w_{i_{k-1}i_k}}. \tag{7}$$

The first term represents the change in system entropy:

$$\Delta S_\mathrm{sys} = \ln\frac{p_{i_0}(0)}{p_{i_n}(\tau)}, \tag{8}$$

while the second term is interpreted as the entropy flow into the environment:

$$\Delta S_\mathrm{m} = \sum_{k=1}^{N}\ln\frac{w_{i_k i_{k-1}}}{w_{i_{k-1}i_k}}. \tag{9}$$

This decomposition implies:

$$\Delta S_\mathrm{tot} = \Delta S_\mathrm{sys} + \Delta S_\mathrm{m}. \tag{10}$$

As in the continuous case, this total entropy production obeys the integral fluctuation theorem:[127]

$$\langle \mathrm{e}^{-\Delta S_\mathrm{tot}}\rangle = 1, \tag{11}$$

which provides a strong statistical constraint on fluctuations in entropy production for finite-time, nonequilibrium trajectories. It also ensures consistency with the second law: $\langle \Delta S_\mathrm{tot}\rangle \geq 0$.

Below, we derive the calculation of the EPR rate for a system following Markovian statistics.[128] We consider a trajectory γ with a sequence of N states $\{i_0,\ldots,i_N\}$ and the corresponding waiting times $\{t_0,\ldots,t_N\}$ for a total observation time T, $\sum_i t_i = T$, given by $\gamma = \{(i_0,t_0),\ldots,(i_N,t_N)\}$. The probability of observing such

Table 1 List of entropy symbols and their definitions

Symbol	Description
ΔS_tot	Total entropy production
ΔS_sys	System entropy production
ΔS_m	Medium entropy production
σ	Entropy production rate (EPR)
σ_KLD	EPR estimator based on KLD
σ_aff	EPR estimator based on affinity
σ_WTD	EPR estimator based on waiting times
σ^TM	EPR estimator for full network
σ^AM	EPR estimator for approximated master equation
σ^L	EPR estimator from merging states
σ^SCGF	EPR estimator using scaled cumulant generating function
σ_pp	Passive partial entropy production estimator
σ_ip or σ^IPEP	Informed partial entropy production
σ_plug	Plug-in EPR estimator
σ_zk	Zero-knowledge EPR estimator
σ_RNEEP	Recurrent neural network estimator for entropy production
σ^n_tot	EPR estimator based on an optimization problem considering nth moments of waiting times

a trajectory with initial probability distribution π_{i_0} is given by

$$P[\gamma] = \mathrm{e}^{-t_N\lambda_{i_N}} \prod_{n=0}^{N-1} \left[w_{i_{n+1},i_n}\mathrm{e}^{-t_n\lambda_{i_n}}\right]\pi_{i_0} \tag{12}$$

where λ_i is the escape rate from state i, $\lambda_j = \sum_{i\neq j} w_{ij}$.

Similarly, one can define the probability $\tilde{P}[\tilde{\gamma}]$ for the time-reversed trajectory $\tilde{\gamma} = \{(i_N,t_N),\ldots,(i_0,t_0)\}$, and express the total EP along the trajectory, Δ, using their ratio:[129]

$$\begin{aligned}\Delta S_{\mathrm{tot}} &= \ln\frac{P[\gamma]}{\tilde{P}[\tilde{\gamma}]} = \ln\frac{\pi_{i_0}}{\pi_{i_N}} + \sum_{n=0}^{N-1}\ln\frac{w_{i_{n+1},i_n}}{w_{i_n,i_{n+1}}} \\ &= \ln\frac{\pi_{i_0}}{\pi_{i_N}} + \sum_{i<j}\phi_{ij}\ln\frac{w_{ij}}{w_{ji}}\end{aligned} \tag{13}$$

where ϕ_{ij} is the net number of transitions from state j to state i.

The EPR can be calculated from the evolution of the EP governed by the Master equation. Following Teza and Stella,[78] we denote the probability $P_i(S,t)$ that the system is in state i at time t, having produced entropy S from all possible trajectories up to the time t. Since each transition between states i and j adds $\ln(w_{ji}/w_{ij})$ to the total EP, the Master equation for $P_i(S,t)$ reads:

$$\partial_t P_i(S,t) = -\lambda_i P_i(S,t) + \sum_{j\neq i} w_{ij}P_j\left(S - \ln\frac{w_{ij}}{w_{ji}}, t\right). \tag{14}$$

Let us denote

$$G_i(\Lambda,t) = \sum_S \mathrm{e}^{\Lambda S}P_i(S,t) \tag{15}$$

as the discrete Laplace transform of $P_i(S,t)$ with respect to S. Eqn (14) then becomes:

$$\partial_t G_i(\Lambda,t) = -\lambda_i G_i(\Lambda,t) + \sum_{j\neq i} w_{ij}\mathrm{e}^{\Lambda\ln\left(w_{ij}/w_{ji}\right)}G_j(\Lambda,t) \tag{16}$$

In matrix form,

$$\partial_t G(\Lambda,t) = \tilde{W}G(\Lambda,t) \tag{17}$$

where $G(\Lambda,t)$ is a column vector with $G_i(\Lambda,t)$ as its i-th entry, and $\tilde{W}$ is the tilted transition matrix:[78,130]

$$\tilde{W}_{ji} = \begin{cases} w_{ji}\mathrm{e}^{\Lambda\ln\left(w_{ji}/w_{ij}\right)}, & i\neq j \\ -\lambda_i, & i = j\end{cases} \tag{18}$$

The dominant eigenvalue Ω^{TM} of $\tilde{W}$ is the scaled cumulant generating function (SCGF) of the entropy production. The mean EPR is given by:[78,130,131]

$$\sigma^{\mathrm{TM}} = \left.\frac{\partial\Omega^{\mathrm{TM}}(\Lambda)}{\partial\Lambda}\right|_{\Lambda\to 0} \tag{19}$$

Here, TM refers to "Total Mean", indicating that all state dynamics are considered.

3.3 EP along a stochastic trajectory

For systems described by overdamped Langevin dynamics, the evolution of the probability density ($p(x,t)$) is governed by a Fokker–Planck equation (FPE). Within the framework of stochastic thermodynamics, the entropy production[136] associated with a single trajectory can be decomposed into two components: the change in system entropy and the entropy flow into the surrounding medium.[81,124] The system entropy at time t is defined as[23,124]

$$S_{\mathrm{sys}}(t) = -\ln p(x(t),t), \tag{20}$$

where $x(t)$ is the system's state along a stochastic trajectory, and $p(x(t),t)$ is the instantaneous probability density at time t.

The entropy change of the medium, often associated with the heat dissipation into the environment (assumed to be at constant temperature), is given by

$$\Delta S_{\mathrm{m}} = \frac{1}{T}\int_0^t \dot{x}(\tau)\circ F(x(\tau),\tau)\mathrm{d}\tau, \tag{21}$$

where $F(x,t)$ is the total time-dependent force, which can result from a combination of a conservative potential and a non-conservative force, $\dot{x}$ is the velocity along the trajectory, and $\circ$ is the Stratonovich product.[124] The total EP along a single trajectory as defined by eqn (10), satisfies the integral fluctuation theorem, eqn (11),[127] which holds for arbitrary initial conditions and time-dependent driving. This result generalizes the second law of thermodynamics to the level of individual stochastic trajectories and highlights the inherent irreversibility of nonequilibrium processes, even in the presence of thermal fluctuations.

3.4 Fully observed systems

The methods described in Sections 3.2 and 3.3 can provide the total EP if all dissipative degrees of freedom are known. Full EP has been calculated in various systems, including particles exhibiting overdamped Brownian dynamics,[137] active Brownian particles (ABPs) under velocity-dependent active forces,[138] coupled harmonic oscillators connected to heat baths at two different temperatures,[132,139] Brownian duets,[140] Brownian particles in moving traps,[141] particles undergoing underdamped Langevin dynamics with active forces, both with and without confinement,[142] harmonically dragged Brownian particles with time-varying stiffness,[143,144] harmonically bound particles under time-dependent forces,[145] polymers in hydrodynamic flow fields,[146] run-and-tumble particles diffusing in harmonic potentials,[134] and electric circuits.[147–149] Fig. 2 shows examples of systems with various driven degrees of freedom.

4 EP from partial information

4.1 Partially observed system

As discussed in Sections 3.2 and 3.3, inferring dissipation in terms of EP requires all relevant mesostates to be experimentally accessible with thermally equilibrated fast hidden dynamics on the microstates. However, due to finite spatiotemporal resolution, all relevant mesostates might not be accessible, and therefore, we obtain a coarser level description of the system. In practice, finite spatiotemporal resolution yields a coarse-grained description in the absence of timescale separation between the microstates. The unresolved microscopic/

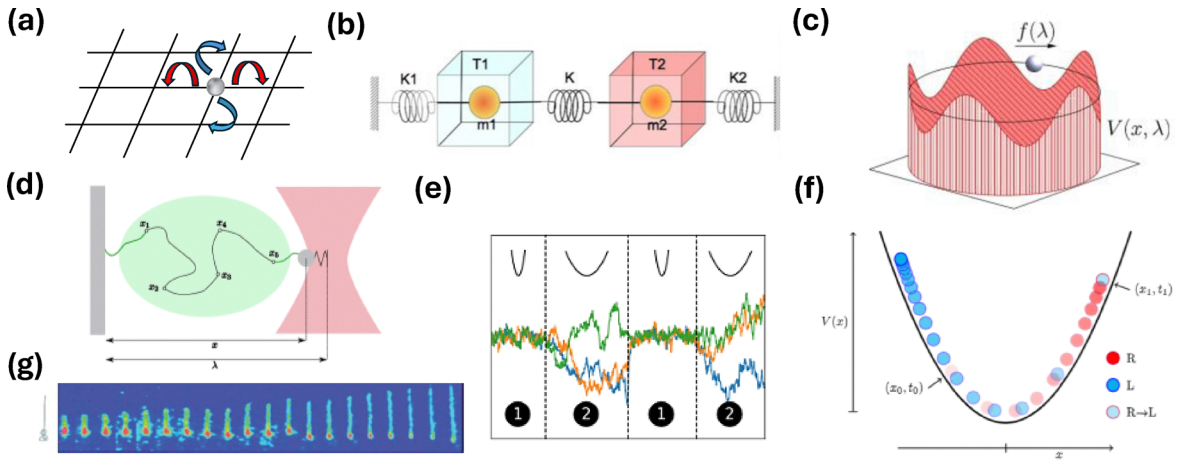


Fig. 2 Schematic of systems with full accessibility of the dissipative degrees of freedom. (a) Random walk on a two-dimensional lattice. (b) Two Brownian oscillators of masses m_1 and m_2 are bound by harmonic potentials of strengths k_1 and k_2, respectively. The oscillators are coupled by a harmonic potential of strength k, and each is in contact with a heat bath, one at temperature T_1 and the other at temperature T_2.[132] This figure has been reproduced from ref. 132 with permission from the IOP Publishing, copyright 2026. (c) The state of a colloidal particle is defined by its position x; it is driven by a periodic potential $V(x,\lambda)$ and a non-conservative force $f(\lambda)$, where λ is a control parameter. This figure has been reproduced from ref. 23 with permission from the IOP Publishing, copyright 2026. (d) Schematic of polymer pulling using an optical trap: one end of the macromolecule is fixed to a surface, while the other end is attached to a colloidal particle pulled by an optical trap. The distance between the bead and the fixed surface (λ) is varied. This figure has been reproduced from ref. 23 with permission from the IOP Publishing, copyright 2026. (e) Trajectories of three diffusing particles driven by harmonic potentials of varying stiffness, shown in different colors. This figure has been reproduced from ref. 133 with permission from the American Physical Society, copyright 2026. (f) Run-and-tumble particle. This figure has been reproduced from ref. 134 with permission from the IOP Publishing, copyright 2026. (g) Experimental images of polymer extension in a hydrodynamic flow field. This figure has been reproduced from ref. 135 with permission from the Royal Society of Chemistry, copyright 2026.

mesoscopic degrees of freedom render the observed dynamics non-Markovian. Therefore, residence times in the observable states for partially observed systems are no longer exponentially distributed.

A model example of a partially observed system is a fully-connected four-state system, in which two of the states cannot be resolved and are therefore coarse-grained into a single hidden state (Fig. 3(a)).[129,151] Specifically, two states remain as observable (Markovian) states, while the other two are combined into one macrostate. An external control parameter is introduced to tune the transition rates across the observed link, thereby influencing the observed dynamics. The transition rates between state 1 and 2: $w_{12}(F) = w_{12}e^{-\beta FL}$ and $w_{21}(F) = w_{21}e^{\beta FL}$, where $\beta = T^{-1}$ is the inverse temperature (with $k_B = 1$), F is the applied force and L is a characteristic length scale.

Fig. 3(b) schematically shows another model system that serves as an example of a partially observed process. It depicts a molecular motor that moves along a one-dimensional track in discrete steps, either "up" or "down." At each spatial position, the motor can exist in either an active or a passive internal state. As illustrated in the figure, the system can undergo spatial transitions between neighboring positions ($i \leftrightarrow i+1$) or switch between internal passive and active states ($i \leftrightarrow i'$). In the active state, upward transitions are favored by a chemical potential difference $\Delta\mu$, while an external force (F) acts downward, opposing this preferred direction. We assume that an external observer cannot distinguish between the active and passive internal states and can only record the motor's position. As a result, the observed dynamics reduce to a second-order semi-Markov process, which can be represented as a three-state cyclic network, where each state corresponds to a physical location that combines the active and passive substates. The transition rates satisfy local detailed balance, with ($\Delta\mu$) affecting only the active-state transitions and (F) influencing all spatial transitions, such that $\frac{w_{d_1}}{w_{u_1}} = e^{\beta FL}$ and $\frac{w_{d_2}}{w_{u_2}} = e^{\beta(FL-\mu)}$.

Theoretical developments in stochastic thermodynamics employ different coarse-graining methods to consider partial information that mimics experimental spatiotemporal limitations. We now introduce two types of coarse-graining (CG) approaches in ref. 152, as shown in Fig. 4. In the full-CG approach, a subset of microstates is lumped into a single observed state, giving rise to a second-order semi-Markov process, since the waiting time in the hidden state depends on the previously visited state. In this example, states 1 and 2 are observed, whereas states 3 and 4 cannot be distinguished and are recorded as a single state H. The waiting time in H is then the sum of the corresponding waiting times in microstates 3 and 4 before transitioning to one of the observed states. In the semi-CG scheme, we assume that an observer can record intra-transitions within the hidden states (Fig. 4(c)). For example, a sequence $1 \to 4 \to 3 \to 2$ is recorded as $1 \to H \to H \to 2$, with the corresponding waiting times (*i.e.*, the time spent during the first and second visits to H) recorded separately. In this case, although the initial and final microstates are not distinguished because they are both lumped into same macrostate, the

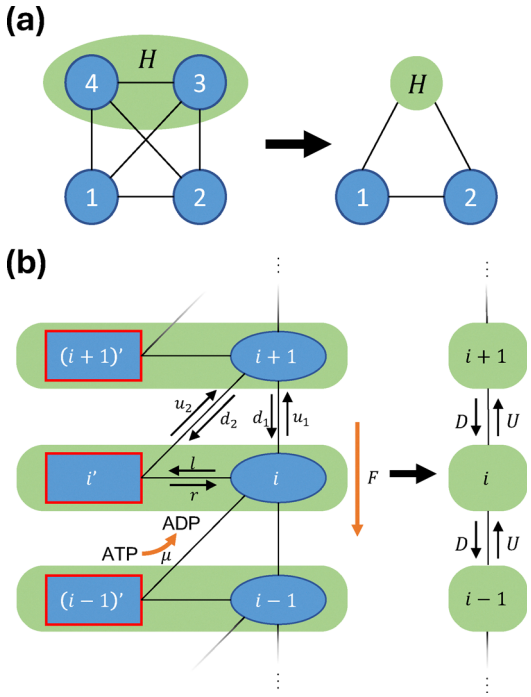


Fig. 3 Fully resolved (left) and coarse-grained (right) representations of two model systems: (a) schematic shows a 4-state fully connected network where the discrete states are denoted by circles with numbers 1, 2, 3, and 4, and transitions between them are represented by solid lines. States 1 and 2 are observed and states 3 and 4 are merged together into a lumped state H; (b) schematic of a fully resolved (left) and a coarse-grained (right) model of a molecular motor. The discrete states representing the position of the molecular motor are represented by the numbers $i-1$, i, and $i+1$, where at each position, the motor can be active (red rectangles) or passive (ellipses). Only the position of the motor is accessible to an observer, while whether it is active or passive cannot be resolved. As such, it is a partially observed system. This figure has been reproduced from ref. 150 with permission from the American Physical Society, copyright 2026.

additional information can still be exploited to improve the lower bound on the total EPR.

Van der Meer *et al.* developed a method to infer entropy production in partially observable Markov systems by analyzing observable transitions and the waiting times between them.[153] Using ratios of waiting time distributions, they built an estimator that either recovers the full entropy production (when no hidden cycles exist) or provides an improved lower bound when hidden cycles are present. They also showed that these waiting-time statistics reveal hidden network features such as cycle presence, length, and affinity. By formulating the problem in an equivalent semi-Markov framework, they unified earlier entropy estimators under a fluctuation-theorem perspective and clarified the role of the correct time-reversal operation.[153]

Around the same time, Harunari *et al.* developed a framework for inferring dynamical and thermodynamic properties from systems where only a subset of transitions is observable.[154] They derived analytic expressions for the probabilities and timing of successive visible transitions and used these to build a lower bound on entropy production that remains informative even when no net currents are observed. They showed that repeated and alternated transition statistics carry distinct physical information, allowing one to detect irreversibility or hidden disorder in the underlying network. Their approach was validated with numerical simulations and demonstrated its applicability to experimentally motivated models such as molecular motors, highlighting the broader power of transition statistics for probing partially accessible Markov processes.[154]

4.2 EPR from partial information

Evaluating the total EP requires full knowledge of all the underlying system dynamics, which is not always accessible. However, partial EPRs can be obtained from the observed information. For these trajectory observables, the auxiliary generator is obtained by modifying the hidden transitions.

The passive partial EP is defined by modified transition rates as follows:[129,155]

$$\bar{W}_{ij}^{\mathrm{pp}} = \begin{cases} w_{ij} & ij \in \mathcal{O} \\ w_{ji}\dfrac{\pi_i}{\pi_j} & i \neq j;\ ij \in \mathcal{O} \\ -\bar{\lambda}_i & i = j \end{cases} \tag{22}$$

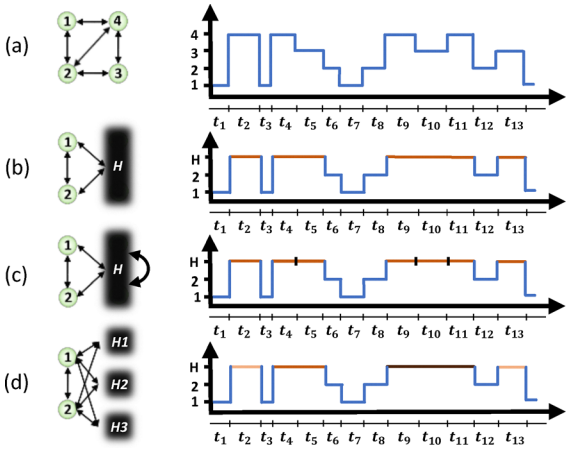
where pp in $\bar{W}_{ij}^{\mathrm{pp}}$ stands for "passive partial". The observed substates form a set denoted by $\mathcal{O}$. w_{ji} and π_j denote the transition rate from i to j and the steady state probability for the state j, respectively. The steady state probabilities are obtained from $\mathbf{W}\pi = 0$, and also $\pi_j = \lim_{T\to\infty} \frac{T_j}{T}$ (the fraction of time spent at each state at long time limit). $\bar{\lambda}_i$ denotes exit rate from state i and ensures probability conservation (see ref. 129 for the analytical expression of $\bar{\lambda}_i$).

The other partial EP is known as the informed partial entropy production, and it is defined by the following:

$$\bar{W}_{ij}^{\mathrm{ip}} = \begin{cases} w_{ij} & ij \in \mathcal{O} \\ w_{ji}\dfrac{\pi_i^{\mathrm{st}}}{\pi_j^{\mathrm{st}}} & i \neq j;\ ij \in \mathcal{O} \\ -\bar{\lambda}_i & i = j \end{cases} \tag{23}$$

where ip in $\bar{W}_{ij}^{\mathrm{ip}}$ stands for "informed partial". π_j^{st} refers to the stall probability distribution for which the flux between the observed substates vanishes. The stall probability distribution satisfies $w_{12}(x^{\mathrm{st}})\pi_2(x^{\mathrm{st}}) = w_{21}(x^{\mathrm{st}})\pi_1(x^{\mathrm{st}})$, where x^{st} is the value of the control parameter that stalls the observed current. Note that in both cases of the passive and informed partial entropy production, the entries of the auxiliary rate matrices, $\bar{W}^{\mathrm{pp}}$ and $\bar{W}^{\mathrm{ip}}$, are identical to the ones of original rate matrix W for the observed transitions $ij \in \mathcal{O}$, where the hidden transitions are multiplied by ratios of either the state-state probabilities or the stalling probabilities, respectively. This formulation gives rise to the development of fluctuation theorems for both partial entropy productions from a unifying perspective.[129]

The passive partial entropy production rate (PPEP), σ_{pp}, is an estimator of the total EPR calculated from the transitions between two observed states, which provides a lower bound on the full entropy production.[129] Suppose we observe only two states i and j, that form Markovian subsystems and record the



Full Trajectory	State	1	4	1	4	3	2	1	2	4	3	4	2	3
	WT	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Full-CG	State	1	H	1	H		2	1	2	H			2	H
	WT	t_1	t_2	t_3	t_4+t_5		t_6	t_7	t_8	$t_9+t_{10}+t_{11}$			t_{12}	t_{13}
Semi-CG	State	1	H	1	H	H	2	1	2	H	H	H	2	H
	WT	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Transformed Semi-CG	State	1	H1	1	H2		2	1	2	H3			2	H1
	WT	t_1	t_2	t_3	t_4+t_5		t_6	t_7	t_8	$t_9+t_{10}+t_{11}$			t_{12}	t_{13}

Fig. 4 Illustration of partial-information frameworks for a four-state system. (a) Fully observed system, with trajectories specified by microstates labeled 1, 2, 3, and 4, and the waiting times (WTs) spent in each state before transitioning to the next state. (b) Full coarse-graining (full-CG): states 3 and 4 are merged into a single mesostate labeled *H*. (c) Semi-coarse-graining (semi-CG): states 3 and 4 remain spatially indistinguishable, but transitions between them can be recorded, with each WT measured for the hidden microstates. (d) Transformed semi-CG: every *n* consecutive visits to the hidden states are grouped into H_n, with the WT given by the sum of the WTs over the *n* visits. This figure has been reproduced from ref. 152 with permission from the American Physical Society, copyright 2026.

transitions between them. In this case, we can compute the transition fluxes or transition rate times the steady state probability density $n_{ij} = w_{ij}\pi_j$ and $n_{ji} = w_{ji}\pi_i$, and use them to evaluate the EP inferred from these transitions:

$$\sigma_{\mathrm{pp}} = \left(n_{ij} - n_{ji}\right) \ln\left(\frac{n_{ij}}{n_{ji}}\right). \tag{24}$$

The average informed partial entropy production rate (IPEP), σ_{ip}, is given by

$$\sigma_{\mathrm{ip}} = \left(n_{ij} - n_{ji}\right) \ln\frac{w_{12}\pi_2^{\mathrm{st}}}{w_{21}\pi_1^{\mathrm{st}}}, \tag{25}$$

where only states 1 and 2 are observed. Like passive partial EP, informed partial entropy production provides alower bound on the total entropy production. PPEP or IPEP and their corresponding complementary parts sum up to the total entropy production rate. Contribution to the complementary parts of the EP arises from the hidden part of the network. Both PPEP and IPEP satisfies the detailed fluctuation theorem.

Ehrich obtained a tight lower bound on the entropy production from dynamics of a hidden Markov model that is compatible with the observed data.[156] Shiraishi *et al.* showed that the partial entropy production satisfies the integral fluctuation theorem for both autonomous and non-autonomous nanomachines.[155]

4.3 Coarse-grained (space) EPR

The amount of information lost describing a partially observed physical system would depend on the coarse-graining procedure. Therefore, recovering the total dissipation from imperfection of partial information and role of finite spatial resolution and finite statistics on the dissipation estimation has attracted a lot of attention in recent times.

Busiello *et al.*[157,158] investigated EP in nonequilibrium systems by comparing stochastic dynamics described by a Master equation (ME) with their coarse-grained representation *via* FPE. It showed that coarse-grained EP contains contributions from microscopic probability currents that are absent in the FPE description and therefore cannot be recovered from it. By comparing ME and Fokker–Planck representations for discrete- and continuous-state systems, the authors derived analytical corrections to the entropy production and identified conditions under which no information is lost in the coarse-graining when moving from an ME to an FPE. The findings, which were experimentally testable, offer a method to infer hidden microscopic processes from coarse-grained measurements.[157,158]

Generally, coarse-graining reduces the complexity of a system either by merging states or by eliminating them. State-merging approaches, including lumping and milestoning, combine multiple microstates into a single mesostate. In contrast, state-elimination methods, such as trimming and decimation, remove selected states from the original network. State-merging methods may reflect experimental systems observed with finite spatiotemporal resolution or systems exhibiting timescale separation, where states connected by fast transitions are grouped into an effective state. State-elimination methods, on the other hand, may correspond to experimental situations in which certain states are unobserved. Whether trajectory functionals of the coarse-grained system remain identical to those of the full system depends on the specific coarse-graining procedure employed.[153]

State lumping is a widely used approach for coarse-graining, in which several microstates are coarse-grained into mesoscopic states as shown in Fig. 5(a).[160] Recently, Igoshin, Kolomeisky, and Makarov have proposed a state eliminating coarse-graining method called trimming[79] (Fig. 5(b)) that preserves the mean EPR in the absence of a dissipative cycle. This

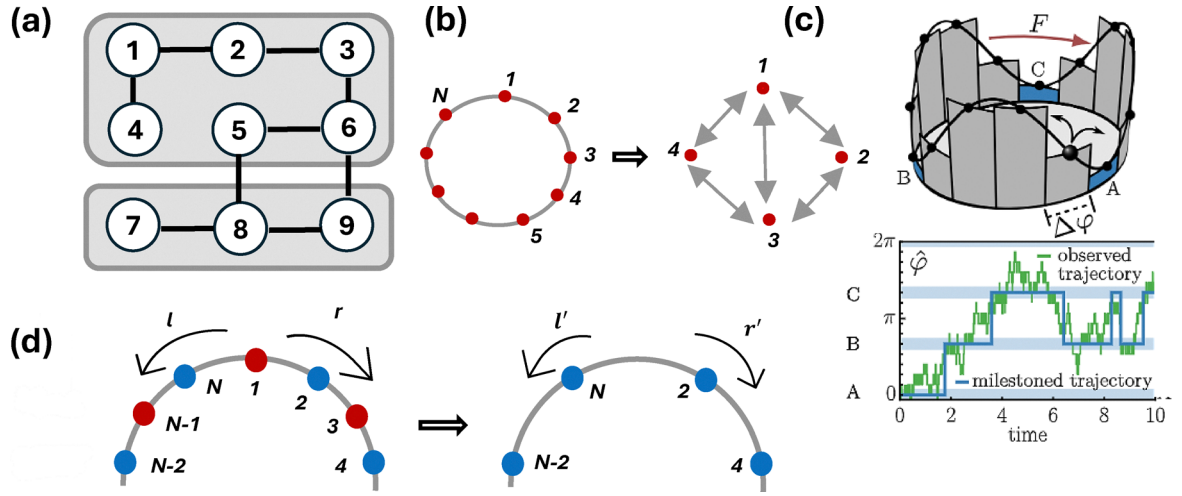


Fig. 5 Schematic of different coarse-graining methods. Lumping and milestoning are state-merging coarse-graining methods, whereas decimation and trimming are state-eliminating coarse-graining methods. (a) Lumping of a discrete Markov chain, in which multiple substates are merged into a single mesostate. (b) Trimming, where states are progressively eliminated from the network while keeping the transition flux among the remaining states unchanged. This figure has been adapted from ref. 79 with permission from the American Chemical Society, copyright 2026. (c) Milestoning – another state-merging coarse-graining method – in which the system remains in a coarse-grained (''milestoned'') state until its system trajectory reaches another milestoned state. This figure has been reproduced from ref. 159 with permission from the National Academy of Sciences, copyright 2026. (d) Decimation, a state-eliminating coarse-graining approach in which selected states are removed from the network. In this schematic, odd-numbered states are eliminated. This figure has been adapted from ref. 78 with permission from the American Physical Society, copyright 2026.

coarse-graining method eliminates states and reassigns the transition rates determined by splitting probabilities and mean first passage times. The observed dynamics of the remaining microstates that do not undergo decimation remain unchanged after the CG method. However, this method might not preserve the network topology for eliminating any state with more than two connections. In this method, the dynamics of the coarse-grained system follow a semi-Markov process, with nonexponential distributions of waiting times between jumps, and it does not require any knowledge about the true microscopic dynamics.

Milestoning[159,161] is another coarse-graining framework (Fig. 5(c)) in which the continuous state space of a stochastic dynamical system is partitioned into a set of milestones, *i.e.*, carefully placed, low-dimensional hypersurfaces that delineate metastable regions or ''cores''. Instead of tracking the full microscopic trajectory, milestoning records only the sequence of milestone crossings and the statistics of transition times between them.

By stitching together these short trajectory segments, one can compute rate constants and estimate thermodynamic quantities. If the equilibration time in each milestoned state is short compared to the excursion times (time spent in between visiting the milestoned states), the system may be approximated as Markovian.[161]

Blom *et al.*[159] clarify the fundamental difference between coarse-graining by lumping and by milestoning. Lumping maps many microscopic configurations onto broad observable states, which typically introduces memory and results in non-Markovian dynamics even when the underlying system is Markov. Milestoning applies an additional postprocessing step. Instead of tracking the system whenever it occupies a lumped state, it records only first passages between selected milestones. This filtering eliminates repeated visits within the same region and often yields a low-order semi-Markov process that, under some conditions, may bring the observable dynamics closer to the underlying irreversibility.

Decimating states (for example, odd numbered states are decimated in Fig. 5(d)), as mentioned by Pigolotti and Vulpiani[162,163] approximate the dynamics of the coarse-grained system by a Markovian master equation. This method eliminates short-lived states by assuming that these are in a quasi-steady state. The probabilities of being in short-lived states can then be expressed as those for the remaining states, resulting in an effective approximate dynamic description that does not explicitly consider the short-lived states. Subsequently, it was demonstrated that in the presence of time-scale separation, the decimation of fast states does not affect entropy production estimates. In trimming, jumps between observed states are counted, whereas jumps between an unobserved state and an observed state are not counted. This distinguishes trimming from the decimation method.

Teza *et al.*[78] showed that memoryless jump processes can be coarse-grained (Fig. 5(d)) while exactly preserving the stationary average and fluctuations of EP. This method assumes complete knowledge of the underlying microscopic dynamics. It eliminates a subset of states from the ME for the probability of the system being in state i at time t, and expresses the entropy S in terms of non-normalized probabilities of the remaining states, which are sufficient to exactly recover the long-time average EPR. This CG method assumes the knowledge of microscopic dynamics before coarse-graining for feasibility of ME formulation.

Bilotto *et al.* studied how coarse-graining affects on entropy production by considering a model system of a

one-dimensional colloidal particle in contact with a thermal bath under a sinusoidal potential and driven out of equilibrium by a small constant force.[164] The authors found that at large friction coefficients, both underdamped and overdamped dynamics yield identical entropy production. In contrast to that, at smaller friction, the overdamped approximation overestimates entropy production compared to the underdamped case. By approximating the continuous dynamics of a colloidal particle under periodic potential as a Markov chain model *via* spatio-temporal discretization, the entropy production at underdamped and overdamped limits for varying friction coefficients was also studied. The authors found that at a small friction limit, inertia plays a significant role in the loss of entropy production in small friction coefficient regime and gain of entropy production in a regime of intermediate frictions.[164] Another instance where the actual EP turns out to be lower than the CG entropy production estimated in ref. 165, where the authors studied a kinetic network with fast-pumping dynamics and slow-network dynamics.

In a coarse-grained description, we often integrate out bath degrees of freedom. However, whether this CG would affect the entropy production rate of the probe particle depends on whether the integrated-out degrees of freedom are dissipative or non-dissipative. Busiello *et al.* employed a generalized Langevin equation (GLE) to model the dynamics of a probe particle immersed in an active bath.[166] The authors found that if the active bath degrees of freedom interact with the probe particle non-reciprocally, then eliminating those entropic degrees of freedom would result in a change in the entropy production. However, for elimination of non-entropic degrees of freedom, there would by no effect on the EPR estimates.[166] Yu *et al.*[167] showed that CG EPR estimates followed a power law relation with the spatial resolution and highlighted the importance of accounting for the correlation between flux at lower resolutions.

4.4 Coarse-grained (time) EPR

Yu *et al.*[167] found that CG EPR is non-monotonic in temporal resolution, with its peak position revealing characteristic timescale corresponding to the underlying dissipative process. This observation is also found in the actomyosin cortex of starfish oocytes.[168] This method provides characteristic dissipative scales from dissipation measurements.[169] In a recent study, Fritz *et al.* investigated the effect of EPR estimators based on waiting times and TUR on finite temporal resolution, assuming that all observed transitions are registered.[170] They found that the waiting-time estimator based on resolved transitions performs best at finite temporal resolution, given perfect measurement statistics.

5 Quantifying degree of irreversibility

In this section, we discuss theoretical tools for estimating dissipation in non-invasive ways that can be directly applied to experimental data. Quantifying dissipation is crucial because a non-zero steady entropy production rate not only characterizes a nonequilibrium process in biological and artificial systems,[12,23,124,171] but also determines the "biological quality"[172] or efficiency of energy transduction in microscopic machines.[173–177] Moreover, dissipation constrains precision in sensory adaptation,[178,179] the regularity of biological clocks,[172] and several other fundamental processes.

As we will see in this section, estimating the probability distribution or the probability current is essential for quantifying entropy production. However, obtaining the probability distribution depends on the underlying system dynamics, and the probability current is not always experimentally accessible or theoretically known without the dynamical equations. In such cases, it is necessary to estimate entropy production using model-free approaches. These approaches will also be discussed in the following subsections.

5.1 EPR calculation from KLD

Kullback–Leibler divergence (KLD) quantifies the distinguishability between two probability distributions,

$$D[p(x) \parallel q(x)] = \int \mathrm{d}x p(x) \ln \frac{p(x)}{q(x)} \tag{26}$$

KLD is a positive quantity that vanishes only when the probability densities are equal, $p(x) = q(x)$ for all x. If an experiment measures a single random variable at a finite frequency, one obtains a discrete trajectory (say length n) as an output which can be written as $x = \{x_1,x_2,x_3,x_4,\ldots,x_n\}$. However, if we do not know the true distribution (say Q) from which x is sampled, and incorrectly guess it as P, then the KLD quantifies the amount of information lost when the true probability distribution Q is approximated by another distribution P.[180]

If $P(x)$ and $Q(x)$ represent probability distributions for time-forward and time-backward series or phase space densities, then KLD measures the irreversibility or relative entropy.[181] In that scenario, the expression for average entropy production rate, $\langle \dot{S} \rangle$, at NESS ($\tau \to \infty$) becomes[182]

$$\langle \dot{S} \rangle = \lim_{\tau\to\infty} \frac{k_B}{\tau} D\left[\{P(z(t)\}_{t=0}^{\tau}) \middle\| \tilde{P}(\{z(\tau - t)\}_{t=0}^{\tau})\right]. \tag{27}$$

Eqn (27) applies for variables with even parity. KLD between phase space densities of the forward (ρ) and backward ($\tilde{\rho}$) transitions for systems transitioning between two equilibrium states is related to the average dissipated work,[183]

$$\langle W_{\mathrm{diss}} \rangle = k_B T D(\rho \| \tilde{\rho}). \tag{28}$$

k_B is the Boltzmann constant and T is the temperature of the heat bath where the system relaxes to a steady state. The average dissipated work is the difference between the average work, $\langle W \rangle$, and the equilibrium free energy change, ΔF, *i.e.*, $\langle W_{\mathrm{diss}} \rangle = \langle W \rangle - \Delta F$. $D(\rho \| \tilde{\rho})$ in eqn (28) is called the relative entropy.[184] Total entropy production was found to be related to the distinguishability of a time forward to its time-reversed process, quantified by the relative entropy between forward and

backward states.[181] In other words, the total entropy production is the dissipated work divided by temperature,[181]

$$S = \langle W_{\mathrm{Diss}} \rangle / T = k_{\mathrm{B}} D(\rho \| \tilde{\rho}) \tag{29}$$

Eqn (29) applies to different initial states and both classical and quantum systems. The dissipated work can also be calculated from the forward and reverse work distribution (P_{F} and P_{B}, respectively) where the protocol is reversed for obtaining the reverse work distribution, $\langle W_{\mathrm{Diss}} \rangle = k_{\mathrm{B}} D(P_{\mathrm{F}}(W) \| P_{\mathrm{B}}(W))$.[182]

The zero-knowledge estimator, σ_{zk}, does not assume any dynamical model, and calculates the EPR from the observed dynamical activity and observed probability distribution of the observed transitions using the formula:[185]

$$\sigma_{\mathrm{zk}} = \frac{1}{2} \sum_{x \in \mathscr{X}} j_x \ln \frac{P(x)}{P(\bar{x})}. \tag{30}$$

The zero-knowledge EPR estimator σ_{zk} requires the observed transition flux, $j_x = \mathscr{K}_{\mathrm{obs}}(P(x) - P(\bar{x}))$ ($\mathscr{K}_{\mathrm{obs}}$ being the dynamical activity or total number of events per unit time), the unconditional probability distribution of the observed event $P(x)$ and its time-reversed counterpart, $P(\bar{x})$. Ref. 185 studies multifilar events and showed σ_{zk} provides a lower bound on the total dissipation.

KLD is a versatile tool for calculating several thermodynamic quantities,[182,186] and several KLD estimators exist, such as the plug-in method[187] and estimators based on compression algorithms.[188] The plug-in estimator, σ_{plug}, was introduced to approximate the Kullback–Leibler divergence (KLD) between forward and reverse sequences of discrete stationary time series by counting data sequences and estimating their probabilities.[182,189] The mth-order approximation of the KLD between sequences of length m is:

$$D_m^x = \sum_{x_1, x_2, \ldots, x_m} p(x_{1 \to m}) \ln \left(\frac{p(x_{1 \to m})}{p(x_{m \to 1})} \right), \tag{31}$$

where $p(x_{1 \to m})$ and $p(x_{m \to 1})$ denote the probabilities of observing the forward sequence $x_{1 \to m} = (x_1, \ldots, x_m)$ and the corresponding backward sequence $x_{m \to 1} = (x_m, \ldots, x_1)$, respectively. These probabilities can be estimated from the frequency of each sequence in a sufficiently long trajectory. The incremental slope of D_m^x as a function of m,[182]

$$\hat{d}_m^x = D_m^x - D_{m-1}^x, \tag{32}$$

converges to the entropy production per step in the limit of large m. Non-Markov processes can be represented as a semi-Markov process of any order, however, calculating $\hat{d}_m^x$ becomes difficult for large m. To address this, the following ansatz[182] has been proposed:

$$\hat{d}_m^x \simeq \hat{d}_\infty^x - c \frac{\ln(m)}{m^\gamma}, \tag{33}$$

where $\hat{d}_\infty^x$, c, and γ are fit parameters for $\hat{d}_m^x$ as a function of m.

The plug-in estimator for the entropy production rate per unit time is then given by

$$\sigma_{\mathrm{plug}} = \frac{\hat{d}_\infty^x}{\tau}, \tag{34}$$

with τ denoting the mean waiting time per step. We applied this estimator to a partially observed four-state system, as well as to molecular motor and flashing ratchet models; the results are discussed in Section 5.7.[152]

Eqn (28) and (29) hold only when the full phase-space density or all driven degrees of freedom are accessible. For coarse-grained descriptions, eqn (29) reduces to[190]

$$\langle W_{\mathrm{diss}} \rangle \geq k_{\mathrm{B}} T D(\rho \| \tilde{\rho}), \tag{35}$$

With partial information or finite-length trajectories, KLD provides only a lower bound on entropy production. In this case, eqn (29) for a trajectory of length t becomes[191]

$$\dot{S} \geq \dot{\sigma}_{\mathrm{KLD}} \equiv \lim_{t \to \infty} \frac{k_{\mathrm{B}}}{t} D[P(\gamma_t) \| P(\tilde{\gamma}_t)]. \tag{36}$$

A hierarchy of inequalities can then be established for k-variable irreversibility:[192]

$$0 \leq S_1 \leq S_2 \leq \cdots \leq S_k \leq \cdots \leq S_{\mathrm{tot}}. \tag{37}$$

where the subscript k of S_k denotes the number of driven variables considered when evaluating the KLD.

Partially observed systems generally lose their Markovian character but can be modeled as semi-Markov processes. In a second-order semi-Markov process, the observed states are reformulated as doublets, $[ij]$, where the first index denotes the previous state and the second the current state. For such processes, the KLD-based estimator of entropy production, σ_{KLD}, can be decomposed into two contributions,[99]

$$\sigma_{\mathrm{KLD}} = \sigma_{\mathrm{aff}} + \sigma_{\mathrm{WTD}}, \tag{38}$$

where σ_{aff} corresponds to entropy production from state affinities, and σ_{WTD} arises from asymmetries in the waiting-time distributions (WTDs). The affinity contribution is given by[99]

$$\sigma_{\mathrm{aff}} = \frac{1}{\tau} \sum_{i,j,k} p(ijk) \ln \left(\frac{p([ij] \to [jk])}{p([kj] \to [ji])} \right), \tag{39}$$

where $p(ijk)$ is the probability of observing the sequence $i \to j \to k$, expressed as[99]

$$p(ijk) = p([ij] \to [jk]) R[ij], \tag{40}$$

with $p([ij] \to [jk])$ the probability of a transition $j \to k$ given $i \to j$, and $R[ij]$ the fraction of visits to the state pair $[i,j]$. Thus, σ_{aff} is determined by the ratio of forward to backward transition probabilities.

The WTD contribution stems from the KLD between forward and backward waiting-time distributions:[99]

$$\sigma_{\mathrm{WTD}} = \frac{1}{\tau} \sum_{i,j,k} p(ijk) D[\psi(t|[ij] \to [jk]) \| \psi(t|[kj] \to [ji])], \tag{41}$$

where $\psi(t|[ij] \to [jk])$ is the WTD in state j, conditioned on the

previous state i and subsequent state k, and τ is the mean waiting time per state.

Eqn (41) is valid for second-order semi-Markov processes when coarse-graining is local in time and commutes with time reversal.[100,101] One such example was demonstrated in ref. 86, which we discuss below. The entropy production was inferred for an experimental partially observed system of hair-cell bundle oscillations, driven by molecular motors.[86,193,194] For this system, the tip position of the hair bundle is accessible, whereas the positions of the molecular motors remain hidden. We estimated the irreversibility of such partially observed oscillatory dynamics, governed by coupled overdamped Langevin equations. However, estimating entropy production rates in continuous-variable systems was challenging due to finite spatiotemporal resolution and the limited accessibility of all driven variables. To address this, we coarse-grained the observed variable of the nonequilibrium system into a few discrete states and estimated a lower bound on the total entropy production rate (EPR) from the Kullback–Leibler divergence (KLD) between waiting-time distributions (WTDs), exploiting the underlying time irreversibility.[195] We further proposed the mean dwell-time asymmetry factor, defined as the ratio between the mean dwell times along the forward and backward directions, as an estimator of EPR. This quantity provided a qualitative measure of broken time-reversal symmetry and increased with finer spatial resolution. We applied this estimator to the above-mentioned example of a second-order semi-Markov process.

EPR inference for partially observed systems was done using the coarse-graining approach as discussed in Section 4. Additional approaches rely on observed transitions and waiting time distributions.[153,154,196,197] Dominic *et al.* introduced a bound on the EPR from the waiting time statistics of hidden Markov processes. They further used their estimator to quantify the irreversibility in various biological processes like in gene regulatory networks, mammalian behavioral dynamics, and others.[196] Van der Meer *et al.* formulated an entropy estimator using the ratios of the forward and backward probability distributions of two consecutive transitions to quantify irreversibility.[153] The authors also determined criteria whether the EPR estimator recovers total EPR or just a lower bound and this estimator also works with network topology with hidden cycles.[153,154] In another study, Van der Meer *et al.* developed a framework to lower bound EP from measuring the time-resolved statistics of events.[197]

5.2 EPR calculation from fluctuation of currents

If the amount of heat delivered to the reservoirs or the gradients of forces applied on a nonequilibrium system is unknown, EP can be obtained from fluctuations of an observable, which is known as the variational characterization of the entropy production rate. Fluctuation–dissipation relation connects observable fluctuations and dissipation for physical systems operating at thermodynamic equilibrium. A new class of inequalities emerged that state dissipation constraints current fluctuations in steady states arbitrarily far from equilibrium, and is known as TUR.[198] These relations capture the trade-off between the precision of a process and its thermodynamic cost. We will briefly discuss extensions and applications of the TUR, along with related results such as the kinetic uncertainty relation (KUR).

It has been shown that at the steady state, the dispersion or uncertainty of the time-integrated quantities like number of reactants and products or the number of steps taken by molecular motors are constrained by the thermodynamic cost associated with the process.[32] If the thermodynamic cost of a chemical process consuming a total number of reactants X during a time interval t, the product of squared relative uncertainty and the total dissipation Σ is constant:[32,199,200]

$$Q = 2\Sigma \frac{\mathrm{var}(X)}{\mathrm{mean}(X)^2} \tag{42}$$

where Q is a dimensionless quantity, and var(X) and mean(X) represent the variance and mean of the variable X.

In 2016, Gingrich *et al.*[33] derived a linear-response like bound on the large deviation of the current for Markov jump process and for small fluctuation limit uncertainty bound on the current fluctuation is shown. The authors also found nonequilibrium current fluctuations are more likely compared to the one obtained from linear-response relation for nonequilibrium systems. In 2018, Katarzyna *et al.* also used linear response theorem to obtain the relative uncertainty of the time integrated quantities, and their approach is applicable for systems with broken time-reversal symmetry.[201] Andreas *et al.* rederived the thermodynamic uncertainty relation in 2020.[202] Gianmaria *et al.* derived a generalization of the so-called TUR without needing the large deviation theory or information-theoretic techniques.[34] In 2016, Patrick *et al.* derived a universal parabolic bound on the generating function of an arbitrary current which depends solely on the average entropy production.[203] They also obtained a power-efficiency trade-off of heat engines working under the temperature gradient between two heat baths. For broken time reversal symmetry, the bounds on the relative uncertainty are controlled both by dissipation and by a parameter encoding the asymmetry of the Onsager matrix.[204]

TUR has been extended to finite-times,[205] discrete times,[206] multidimensions,[35] systems following generalized Langevin equation with memory,[38] underdamped Langevin dynamics,[38] unidirectional processes,[207] optimal TUR for Markov process,[208] TUR including measurement and feedback,[209] and for arbitrary initial states.[210] A recently derived thermodynamic uncertainty relation shows that the minimum scaled variance of a charge, as well as the charge's variance, is governed not only by the mean entropy production but also by the higher moments of its probability distribution.[200]

Ivan *et al.* derived a new inequality called kinetic uncertainty relation (KUR) that states how observable fluctuations in discrete stochastic systems are bounded by the mean number of jumps among discrete states of stochastic systems, applicable for all times.[130] There have been unified thermodynamic and kinetic uncertainty relation.[211] A tighter bound on the current

fluctuation is achieved from both thermodynamic and kinetic uncertainty relation, and it leads to a stronger classical speed limit. This framework can also be extended to first passage time observables and is applicable for systems with unidirectional transitions.[212]

Gingrich *et al.* derived an inequality relating the dissipation rate to current fluctuations in jump processes and found a lower bound on the total dissipation rate for driven diffusive process from observed coarse-grained currents.[213] Li *et al.* estimated entropy production rate from probability currents.[214] Using the finite-time generalization of the thermodynamic uncertainty relation, the mean and fluctuations of the entropy production were obtained, and the framework requires short experimental time-series data.[215] It was further used to infer a lower bound to entropy production rate from flickering data generated by interference reflection microscopy of HeLa cells.[216]

So far, we have discussed variants of the TUR based on transition currents. TUR, however, has also been extended to account for fluctuations in first-passage times.[217,218] In this formulation, the TUR is expressed as

$$\frac{\mathrm{var}(T)}{\langle T\rangle^2} \geq \frac{2}{T\sigma}, \tag{43}$$

where $\langle T\rangle$ and $\mathrm{var}(T)$ denote the mean and variance of the first-passage time, respectively, and σ is the entropy production rate. More recently, the TUR framework has been extended to bound the characteristic timescale $t^\star$ over which a system can exhibit anomalous diffusion, where the mean-squared displacement follows a power law in time, $\langle x^2(t)\rangle \simeq D_\alpha t^\alpha$, with generalized diffusion coefficient D_α (units: $\mathrm{m}^2\ \mathrm{s}^{-\alpha}$) and anomalous exponent $\alpha \neq 1$:[219]

$$t^\star \simeq \frac{\dot{W}_{\mathrm{ss}} D_\alpha}{2k_{\mathrm{B}}Tv^2}, \tag{44}$$

where $\dot{W}_{\mathrm{ss}}$ is the steady-state dissipation rate, and v is the mean particle velocity, defined through the linear growth of displacement, $\langle x(t)\rangle = vt$.

5.3 EPR from optimization

For discrete-time models, Ehrich[156] introduced σ_{fit}, which searches over possible hidden dynamics consistent with the observed statistics, given knowledge of the number of hidden states. The optimization problem was formulated for a four-state system, with two states observed and the other two coarse-grained into one. This approach can be extended to similar settings, provided that the number of observed states matches the number of hidden states and the jump probability matrices are invertible.

In continuous-time models, Skinner and Dunkel[220] derived a lower bound on the total EPR, σ_2, by minimizing it over a canonical representation of the system that preserves both first- and second-order transition statistics. The canonical form is obtained through a sequence of transformations that do not increase the EPR while maintaining the mass rate statistics up to second order. They further showed that for any triplet of coarse-grained states I, J, K, the number of intermediate states within J connected to I and K can be reduced to at most four without affecting the minimum EPR. Because the canonical form has a simple structure, one can readily identify the constraints ensuring that the observed statistics are preserved.

In another study, Skinner and Dunkel[196] proposed an optimization method to estimate the EPR (σ_{T}) in systems with two observed states using waiting-time statistics. By rescaling the rates and steady-state probabilities of the underlying system, they obtained an estimator expressed as a factor $\frac{2k_{\mathrm{B}}}{t_{\mathrm{A}}+t_{\mathrm{B}}}$ multiplied by a function Λ, which depends solely on the ratio $\frac{\mathrm{var}(t_{\mathrm{A}})}{{t_{\mathrm{A}}}^2}$, where t_{A} and t_{B} are the waiting times in the observed states. The function Λ can be computed numerically by solving an optimization problem for each value of $\frac{\mathrm{var}(t_{\mathrm{A}})}{{t_{\mathrm{A}}}^2}$.

In another study,[150] we introduced a new method that incorporates additional information about the underlying system topology, which was not accounted for in earlier approaches and found EPR for a partially observed system using an optimization techniques based on the statistics of the transition rates and the cumulants of the waiting times (Fig. 6). Given a coarse-grained system and a model of the full Markovian network topology, we formulated an optimization problem to obtain a tight bound on the total EPR. The observables considered include the coarse-grained steady-state probabilities, π_I, representing the probability of being in state I; the first-order mass rates (steady state probability density times the transition rates), n_{IJ}, corresponding to transitions $I \to J$ (Fig. 6(a)); the second-order mass rates, n_{IJK}, associated with transitions $I \to J \to K$, and the conditional waiting-time distributions $\Psi_{IJK}(t)$, describing the distribution of waiting times in state J before transitioning to K, conditioned on a prior transition $I \to J$. The optimization problem was defined over all possible underlying systems with the same topology as the assumed Markovian model that reproduce the observed statistics, while minimizing the EPR. By construction, the EPR of the coarse-grained system is bounded from below by the minimal entropy production of the underlying Markovian system consistent with the observed statistics.

We first derived analytical results for the Laplace transforms of conditional probabilities of two- and three-state transitions, which enabled us to formulate the optimization problem.[150] Let S denote the true underlying Markovian system and R a general system with the same topology as S, that is, the same states and allowed transitions, but arbitrary mass rates and steady-state probabilities. Considering all systems R with steady-state probabilities $\pi_I^R = \pi_I^S$, first-order mass rates $n_{IJ}^R = n_{IJ}^S$, second-order mass rates $n_{IJK}^R = n_{IJK}^S$, and conditional waiting-time distributions $\Psi_{IJK}^R(t) = \Psi_{IJK}^S(t)$, the following inequality holds between the EPR of S and R:

$$\sigma_{\mathrm{tot}}(S) \geq \min_{\mathscr{R}}\{\sigma_{\mathrm{tot}}(\mathscr{R}) | \forall_{I,J,K}: \pi_I^{\mathscr{R}} = \pi_I^S, n_{IJ}^{\mathscr{R}} = n_{IJ}^S, \; n_{IJK}^{\mathscr{R}} = n_{IJK}^S, \psi_{IJK}^{\mathscr{R}}(t) = \psi_{IJK}^S(t)\} \equiv \sigma_{\mathrm{opt}}^{(\infty)}. \tag{45}$$

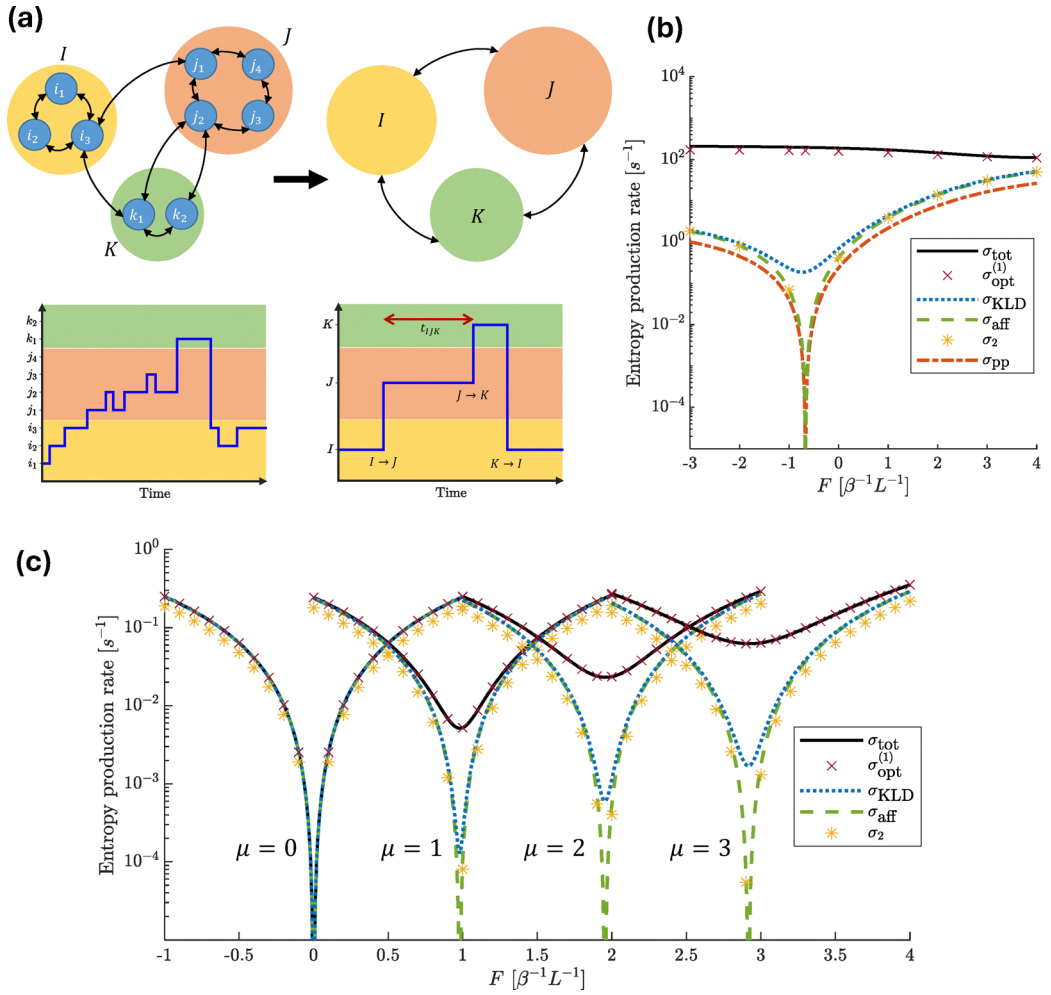


Fig. 6 (a) The full Markovian system consists of the substates (i_n, j_n, and k_n) (left), while the coarse-grained system consists of the mesostates (I, J, and K) (right). An example of a full trajectory (left), illustrating transitions between substates, and the corresponding coarse-grained trajectory (right), showing transitions only among the observed states. (b) Comparison of different EPR estimates for the four-state model shown in Fig. 3(a): the total EPR, σ_{tot} (solid black line); our bound, σ_1^{opt} (brown cross), calculated using eqn (46); the KLD estimator, σ_{KLD} (dotted blue line); the affinity estimator, σ_{aff} (dashed green line); the two-step estimator, σ_2 (yellow asterisk); and the passive partial entropy production, σ_{pp} (dash-dotted orange line). (c) EPR estimates for the molecular motor model shown in Fig. 3(b): the total EPR, (σ_{tot}) (solid black line); our bound, (σ_1^{opt}) (brown cross); the KLD estimator, (σ_{KLD}) (dotted blue line); the affinity estimator, σ_{aff} (dashed green line); and the two-step estimator, σ_2 (yellow asterisk). The details of the parameters used can be found in ref. 150. This figure has been reproduced from ref. 150 with permission from the American Physical Society, copyright 2026.

Unlike constraints on steady-state probabilities and mass rates, enforcing equality of the full waiting-time distributions $\Psi_{IJK}(t)$ requires matching continuous functions, which cannot be exactly reconstructed from finite trajectory data. Moreover, directly imposing functional constraints with non-trivial dependence on optimization variables is computationally demanding. To address this, we reformulated the optimization problem using the moments of the waiting-time distributions:

$$\sigma^{(n)}_{\mathrm{opt}}(S) \equiv \min_{\mathscr{R}}\left\{\sigma_{\mathrm{tot}}(\mathscr{R}) \middle| \forall_{I,J,K}: \pi^{\mathscr{R}}_I = \pi^S_I, n^{\mathscr{R}}_{IJ} = n^S_{IJ}, \right.$$
$$\left. n^{\mathscr{R}}_{IJK} = n^S_{IJK},\ \forall_{k\in\{1,2,\ldots,n\}}: \left\langle t^k_{IJK}\right\rangle^{\mathscr{R}} = \left\langle t^k_{IJK}\right\rangle^S \right\}, \tag{46}$$

where $\langle t^k_{IJK}\rangle$ denotes the k-th moment of $\Psi_{IJK}(t)$.

The analytical expressions for the moments can be obtained from the Laplace transform as $\left\langle t^k_{IJK}\right\rangle = (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}s^k}\Psi_{IJK}(s)\Big|_{s\to 0}$.

This reformulation simplifies the dependence of the observables on the optimization variables, making the problem more tractable. After evaluating the observables, the optimization is solved using a global non-linear search algorithm.[221]

Fig. 6 shows different EPR bounds on a four-state system (Fig. 6(b)) and a molecular motor model (Fig. 6(c)).[150] Fig. 6(b) shows comparison of existing estimators like σ_2, σ_{KLD} ($\dot{S}_{KLD}$ in eqn (38)), σ_{aff} ($\dot{S}_{aff}$ in eqn (39)), and σ_{pp} (eqn (24)) with our estimator $\sigma^{(1)}_{opt}$ (eqn (46) with n = 1). We conclude that, at the stalling force (condition for no current in the visible link) we obtain $\sigma_{pp} = \sigma_{aff} = \sigma_2 = 0$, which corresponds to the trivial bound. In contrast, σ_{KLD} and our estimator $\sigma^{(1)}_{opt}$ provide non-trivial bounds. Moreover, $\sigma^{(1)}_{opt}$ significantly surpasses σ_{KLD} and yields a tight bound. For this 4 state system (Fig. 3(a)), using higher moments to calculate $\sigma^{(2)}_{opt}$ (eqn (46) with n = 2) did not lead to any improvement compared to $\sigma^{(1)}_{opt}$. While σ_2 can be

larger than σ_{KLD} in some cases,[220] for the rate values considered here, we find $\sigma_2 < \sigma_{\mathrm{KLD}}$. In fact, although σ_2 and σ_{aff} appear to be similar in Fig. 6(a), we confirmed that $\sigma_2 < \sigma_{\mathrm{aff}}$ for all values of F used in the transition rates. We found the following hierarchy of the different EPR estimators for the molecular motor given the transition rates considered: $\sigma_{\mathrm{opt}}^{(1)} \geq \sigma_{\mathrm{KLD}} \geq \sigma_{\mathrm{aff}} \geq \sigma_2$. At the stalling force for each value of μ, we found $\sigma_{\mathrm{aff}} = \sigma_2 = 0$, corresponding to the trivial bound. In contrast, similar to the four-state system. In Fig. 6(b), $\sigma_{\mathrm{opt}}^{(1)}$ significantly surpasses σ_{KLD} and provides a tight bound.

Other EPR estimators that use observed transitions only. For example, Dominic *et al.*[220] introduced EPR estimators from observed transition statistics using an optimization framework, which would produce non-vanishing values for processes that appear to maintain detailed balance and time-symmetric. They used their framework for bacterial flagella motors, calcium oscillations within human embryonic kidney cells, and growing microtubules.

5.4 EPR from thermodynamic speed limit

EPR has also been estimated from speed limits.[66,222–224] The fastest rate (τ_ζ^{-1}) at which any physical observable can measurably change is limited by the rate (τ^{-1}) at which the probability distribution can change.[225] The thermodynamic speed limit can be mathematically expressed as

$$\tau^{-1} \geq \tau_\zeta^{-1}. \tag{47}$$

The intrinsic speed τ^{-1} is related to the time parameterized Fisher information[226] ($I_{\mathrm{F}}(t)$) *via* $\sqrt{I_{\mathrm{F}}(t)} = \Delta\dot{I} = \tau^{-1}$. The thermodynamic speed (τ_ζ^{-1}), defined by the ratio of the function's rate of change to one standard deviation of the associated variable, $\tau_\zeta^{-1} = \left|\dot{\zeta}\right| \Big/ \Delta X$, where $\left|\dot{\zeta}\right| = \left|\mathrm{cov}(X, -\dot{I})\right|$ is the covariance between the surprisal rate and observable or variable X.

In another study,[227] Nicholson *et al.* proposed a theoretical framework to derive the thermodynamic speed of observables (X) in models where the full probability distribution is known, based on linear regression of the surprisal rate ($\dot{I} = -d_t \ln P(X,t)$):

$$\dot{I} = bX + a. \tag{48}$$

They showed that the optimal intercepts (a_{opt}) and optimal slopes (b_{opt}) of the linear regression model are related to free energies and observable speed, respectively.[227] Consider the rate of change of (log) probability, $\hat{r}$ to be a function of the driven variables, X, where $X^{\mathrm{T}} = (X_1, X_2, \ldots, X_N)$. For the function

$$\hat{r} = a + b^{\mathrm{T}}X + \mathscr{E}, \tag{49}$$

we can find the optimal parameters by minimizing the mean-squared error:

$$\langle \mathscr{E}^2 \rangle = \langle (\hat{r} - (a + X^{\mathrm{T}}b))^{\mathrm{T}} (\hat{r} - (a + X^{\mathrm{T}}b)) \rangle. \tag{50}$$

Setting $\left\langle \frac{\partial \mathscr{E}^2}{\partial a} \right\rangle = \left\langle \frac{\partial \mathscr{E}^2}{\partial b} \right\rangle = 0$, gives

$$a = \langle \hat{r} \rangle - \langle X^{\mathrm{T}} \rangle b \tag{51}$$

and

$$\langle XX^{\mathrm{T}} \rangle b = \langle X(\hat{r} - a) \rangle. \tag{52}$$

Solving eqn (52), we find the optimal slope

$$b_{\mathrm{opt}} = \frac{\mathrm{cov}(X, \hat{r})}{(\langle XX^{\mathrm{T}} \rangle) - \langle X \rangle \langle X^{\mathrm{T}} \rangle)}. \tag{53}$$

The covariance in the numerator corresponds to physical observables (*e.g.*, heat rate, rate of dissipated work, entropy production rate). Substituting b_{opt} into eqn (51), we find the optimal intercept, a_{opt}. For example, if the observable is the system's energy, the observable speed corresponds to the heat dissipation rate (assuming no work is performed in the process). However, evaluating the time-parameterized Fisher information still requires access to the full probability distribution. The theory was then applied to discrete state models.[227]

We recently showed that the need for the full probability distribution can be bypassed by considering the coordinate transformation of the Fisher information which would relate the coordinate-transformed Fisher information to the variance of an experimental variable.[195] This method enables us to use the statistical moments of the experimental observables to obtain the intrinsic speed. This coordinate-transformed Fisher information is further used to calculate the dissipation rate using the variance of observable responsible for dissipation using:

$$\left|\dot{Q}\right| \equiv \left|\mathrm{cov}(\varepsilon, -\dot{I})\right| = \frac{|d_t X|}{\Delta X} \Delta\varepsilon. \tag{54}$$

X could be the position of a particle exhibiting a nonequilibrium process, concentration, number of molecules of a chemical species, *etc.* The key advantage of our method is to experimentally exploit the known thermodynamic speed limit to calculate the dissipation rate. This method is not only limited to entropy production estimation, moreover, ζ can represent different thermodynamic variables depending on the observable X. We illustrated our theoretical approach to estimate dissipation rates for two systems: (1) harmonically bound Brownian particle dragged by a constant velocity (Fig. 7(a) and (b)) and (2) active gel composed of microtubule and kinesin molecule that are driven out of equilibrium by ATP hydrolysis (Fig. 7(c)).

For harmonically dragged Brownian particle, we predicted the heat dissipated using only the measured uncertainty in position Δx, trap speed v, and the temperature of the surrounding fluid T:

$$\left|\tilde{\dot{Q}}\right| \approx \frac{|v|}{\Delta\hat{x}} \Delta\hat{\varepsilon} = |v| \sqrt{k_{\mathrm{f}}} \sqrt{k_{\mathrm{B}} T} \tag{55}$$

The first expression of the heat rate can use data directly. We estimated $\Delta\hat{x} \approx \sigma_x = \sqrt{k_{\mathrm{B}} T / k_{\mathrm{f}}} = 78.32$ nm (errors can be of order ± 2 nm) and $\Delta\hat{\varepsilon} \approx k_{\mathrm{B}}T/2$. In this case, our prediction used quantities that are experimentally controlled or can be experimentally measured. As validation of the numerical prediction, we used the dissipation rate by numerically simulating

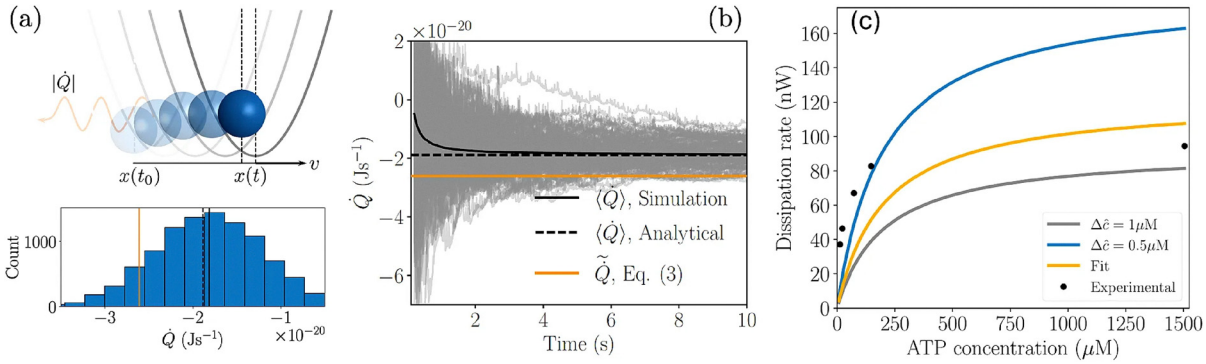


Fig. 7 (a) A Brownian particle releases energy in the form of heat when it is dragged through a viscous medium. At time $t = t_0$, the particle is located at the center of the trapping potential, $x(t_0)$. As the trap is translated, the particle's average position shifts upward along the potential landscape. Throughout this process, heat is dissipated into the surroundings at a rate ($|\dot{Q}|$). A histogram of the heat transferred to the environment, obtained from simulations of overdamped Langevin dynamics with 10 000 noise realizations at $t = 4$ s is shown. The housekeeping heat rate predicted by eqn (55) (eqn (3) in the legend) (orange) matches both the sample mean from the simulations (black solid line) and the analytical mean (black dotted line). (c) Comparison of the heat rates predicted using eqn (56) for different ATP concentration uncertainties (gray and blue solid lines), experimentally measured average dissipation rates (shown by black filled circles), and a fit based on a chemical kinetics model is shown as a function of ATP concentration for a microtubule active gel. This figure has been reproduced from ref. 195 with permission from the American Physical Society, copyright 2026.

Langevin dynamics. Our predicted value $\left|\tilde{\dot{Q}}\right| = |v|\sqrt{k_f/\beta} = 2.61 \times 10^{-20}$ W agrees well with the value from our numerical simulations 1.85×10^{-20} W. Fig. 7(b) shows our prediction also agreed well with an analytical formula for the heat rate $|\langle\dot{Q}\rangle| = v^2/\gamma = 1.885 \times 10^{-20}$ W.[228]

For the active gel system, the predicted heat rate from our formalism based on coordinate-transformed thermodynamic speed (Fig. 7(c)) limit[195]

$$\left|\tilde{\dot{Q}}\right| \approx \frac{|d_t\hat{c}|}{\Delta\hat{c}}\Delta\hat{\varepsilon}. \tag{56}$$

In this case, we did use a chemical kinetics model to estimate the dissipation rate because the time derivative of concentration, $d_t\hat{c}$, has not yet been reported. However, when available, $d_t\hat{c}$ can be directly used to estimate the heat rate. As part of the expanded discussion, we describe in more detail how we estimated the energy fluctuations to be 10^{-8} J or 10 nW in the active gel. We first estimated this value based on Foster *et al.*'s experimental measurements of the heat release rate.[45] Over 7 minute intervals for a total of 180 minutes, their measurements show fluctuations on the order of 10–15 nW when the gel was prepared with and without the pyruvate kinase-based ATP regeneration system. As a point of comparison, these fluctuations are smaller than the mean heat release rate they measured: the microtubule active gel dissipates energy at a rate on the order of 100 nW. The experimental uncertainty is 0.2 nW energy resolution of the picocalorimeter.[45] Since 10 nW is an order of magnitude smaller than the measured heat release and two orders of magnitude larger than the experimental error, we used the 10 nW deviations as an estimate of the energy fluctuations.

Our formalism can also be extended into higher dimensions. One of such example is shown here. Suppose we consider a moving trap along the x-coordinate, while the particle remains equilibrated and static in the y-coordinate. If the y-coordinate is equilibrated and does not contribute to dissipation, then using y alone to estimate the heat rate would yield a poor estimate unless it is coupled to the driven x-coordinate.

When the equilibrated y-coordinate is coupled to the driven x-coordinate, the covariance cov(x,y) captures this coupling. In this two-dimensional case, one can perform a linear regression for the surprisal rate, $\hat{r}(x,y)$, of the form

$$\hat{r} = ax + by + c. \tag{57}$$

The variance of $\hat{r}$,

$$\tilde{\tau}^{-2} \equiv \tilde{I}_F \equiv \Delta\hat{r}^2, \tag{58}$$

represents the coordinate-transformed Fisher information. Expressed in terms of position uncertainties, this Fisher information becomes

$$\tilde{\tau}^{-2} = \left(\frac{\partial\hat{r}}{\partial x}\right)^2\Delta x^2 + \left(\frac{\partial\hat{r}}{\partial y}\right)^2\Delta y^2 + 2\left(\frac{\partial\hat{r}}{\partial x}\right)\left(\frac{\partial\hat{r}}{\partial y}\right)\mathrm{cov}(x,y). \tag{59}$$

In this scenario, if the driven x-coordinate is inaccessible, the terms involving $(\partial\hat{r}/\partial x)$ vanish. The optimal parameters for the linear statistical model then give $(\partial\hat{r}/\partial y) = \partial_t y/\sigma_y{}^2$. Since y is equilibrated, $\partial_t y$ vanishes, leading to a poor estimate of the heat rate when only the y-coordinate is used in the coordinate transformation.

5.5 EPR from machine learning approaches

Kim *et al.*[229] proposed a neural estimator for entropy production (NEEP), that estimates EP from trajectories of relevant variables without detailed knowledge of system dynamics, applied their approach to bead spring and discrete flashing ratchet models. Their method can estimate coarse-grained EP even for higher-dimensional data. This approach is independent of the system dynamics, and can be useful for other contexts.[230]

Kim *et al.*[231] also introduced a machine-learning framework to estimate EP in systems with odd-parity variables using neural networks trained on trajectory data and parity information. They further demonstrated their approach with an underdamped bead-spring model and an odd -parity Markov jump process. Odd-parity systems require additional estimators to account for asymmetry (ΔS_{as}) and waiting-time contributions (ΔS_{WTD}). The bead-spring model and the Markov jump process used two and three such estimators, respectively. Related studies[99,220] found out that WTD fluctuations can reveal EP in hidden Markov or semi-Markov processes.

We also studied the RNEEP estimator, $\sigma_{\mathrm{RNEEP},m}$, which approaches the entropy production rate as an optimization task solved by stochastic gradient descent.[229,232,233] Its input is a collection of sequences of length m taken from a long trajectory, and the output is the coarse-grained entropy production per step. Similar to the Plug-in estimator mentioned in Section 5.1, RNEEP relies only on discrete state sequences and does not require explicit knowledge of waiting-time distributions. A recurrent neural network can be used to compare forward and time-reversed sequences, thereby quantifying trajectory irreversibility. Although its estimates are typically close to those from the plug-in or affinity-based approaches, the RNEEP framework offers a flexible, machine-learning-based method for refining lower bounds on entropy production. The EP estimates from RNEEP are further discussed in Section 5.7. Moreover, there are other studies where machine learning approaches have been used to estimate dissipation.[234]

5.6 EPR from error propagation

Di Terlizzi *et al.* introduced a variance sum rule (VSR) for displacement and force variances that allows direct EP estimation in steady.[235,236] They demonstrated their approach to estimate EP for active Brownian particles in optical traps and red blood cell flickering that agree with calorimetric measurements. The authors further generalized the VSR[235] relating variances of displacement and force for overdamped Langevin systems at NESS. They found out deviation of mean-squared displacement from normal diffusion is a result of non-equilibrium condition. The VSR highlights that deviations from normal diffusion are caused by nonequilibrium effects. They derived entropy production rate which depends on second-order time derivatives of position correlations, and applied their framework to exactly solvable models. This approach can infer the hidden nonequilibrium behavior.[235] Our method[195] as presented in Section 5.4 is a generalized version of VSR to estimate EP.

5.7 EPR comparison

This section compares inferred dissipation bounds from spatiotemporally coarse-grained data. Passive partial EP σ_{pp} (eqn (24)) provides a lower bound for the total entropy production. The total entropy production is given by

$$\sigma_{\mathrm{tot}} = \sum_{i<j} \left(n_{ij} - n_{ji}\right) \ln\left(\frac{n_{ij}}{n_{ji}}\right). \tag{60}$$

Restricting the sum to a pair of observed states i and j yields

$$\begin{aligned}\sigma_{\mathrm{tot}} &= \left(n_{ij} - n_{ji}\right) \ln\left(\frac{n_{ij}}{n_{ji}}\right) + \cdots \geq \left(n_{ij} - n_{ji}\right) \ln\left(\frac{n_{ij}}{n_{ji}}\right) \\ &= \sigma_{\mathrm{pp}}.\end{aligned} \tag{61}$$

Ref. 129 showed

$$\sigma_{\mathrm{tot}} \geq \sigma_{\mathrm{ip}} \geq \sigma_{\mathrm{pp}} \geq 0, \tag{62}$$

where using a molecular motor model system, Martínez *et al.* showed[99]

$$\sigma_{\mathrm{KLD}} \geq \sigma_{\mathrm{aff}} \geq \sigma_{\mathrm{pp}}. \tag{63}$$

Mean EPR rates can be estimated using the following formula for different CG methods[78,130,131]

$$\sigma = \left.\frac{\partial \Omega(\Lambda)}{\partial \Lambda}\right|_{\Lambda \to 0}. \tag{64}$$

where Ω represents the dominant eigenvalue of a coarse-grained or tilted transition rate matrix as expressed below

$$\left[\hat{R}\right]_{ji} = \begin{cases} [R]_{ji} \mathrm{e}^{\Lambda \ln\left(\frac{[R]_{ji}}{[R]_{ij}}\right)}, & i \neq j, \\ -\sum\limits_{j \neq i} [R]_{ji}, & i = j. \end{cases} \tag{65}$$

where $[\hat{R}]_{ji}$ would vary depending on the coarse-graining method employed, and different forms of the matrix are summarized in ref. 128. We investigated various notions of partial information in driven systems and the corresponding mean EPRs.[128] TM refers to total mean EPR considering full knowledge about system dynamics.

We discussed mainly two approaches. One is lumping (Fig. 5(a)) and the other one is decimation (Fig. 5(d)). We call EPRs obtained from different CG methods as SCGF, AM, L, and compare them with partial information-based inference like IPEP as discussed in Section 4.2. SCGF refers to a decimation-based CG approach which employs scaled cumulant generating function (SCGF). Whereas, AM refers to a decimated or coarse-grained system dynamics that follows an approximated Master equation with effective transition rates. L refers to a lumping method[74] where the probabilities of the merged states are summed up, while the steady states probabilities of the rest of the states remain unchanged, and the coarse-grained system dynamics follow an approximated Master equation. Another lumping procedure (HS) was developed by Hummer and Szabo.[160] Their method ensures that the time-dependent occupancy number correlation functions in the coarse-grained system are equal to the ones of the original system. The same reduced matrix (rate matrix for the coarse-grained system) can be obtained from the projection operator technique. In the approximated Markovian limit, the reduced transition rate matrix was calculated analytically.[160]

Fig. 8(a) shows the different network topologies used to compare EPR estimators in Fig. 8(b), in which TM, SCGF, AM, L, and IPEP correspond to σ^{TM}, σ^{SCGF}, σ^{AM}, σ^{L}, and σ^{IPEP}, respectively. TM provides the mean EPR for fully observed systems, so the closer the EPR bounds are to TM, the better the estimators. Coarse-graining *via* scaled cumulant generating functions (SCGF), which involves decimation and redistribution of steady-state probabilities and transition rates, accurately reproduces the total mean EPR as expected. SCGF approach was meant to preserve the total mean EPR (σ^{TM} or TM).[78] We compared AM, L, HS, and IPEP with the total mean EPR (TM) across all network topologies shown in Fig. 8(b). The approximated Master equation (AM) yields a lower bound on TM because decimation produces a non-Markovian coarse-grained system, preventing AM from capturing the full entropy production.

When the hidden substates are disconnected (network topology III and VII in Fig. 8(a)), AM provides values closer to the total mean EPR (σ^{TM}) as no cycles carrying transition currents are lost during the coarse-graining method. For topology VII of Fig. 8(a), the available observations are sufficient to determine the total entropy production exactly, because the observed cycle is the sole entropy-producing fundamental cycle in the network; consequently, AM and IPEP equals the total EPR (TM). For specific topologies, reducing connectivity among hidden states leads the AM method to converge to the TM EPR, since the entropy contribution from loss cycles with transition flux diminishes.

By comparing state-based coarse-graining methods (lumping L and HS, and decimation-based SCGF and AM) that rely on knowledge of the network topology, with the mean EPR inferred solely from observed states and transitions (IPEP) in Fig. 8(b), we showed that the impact of partial information on inferred EPR depends strongly on network structure. Notably, IPEP yields entropic information comparable to lumping (L), which preserves steady-state properties, but their difference increases

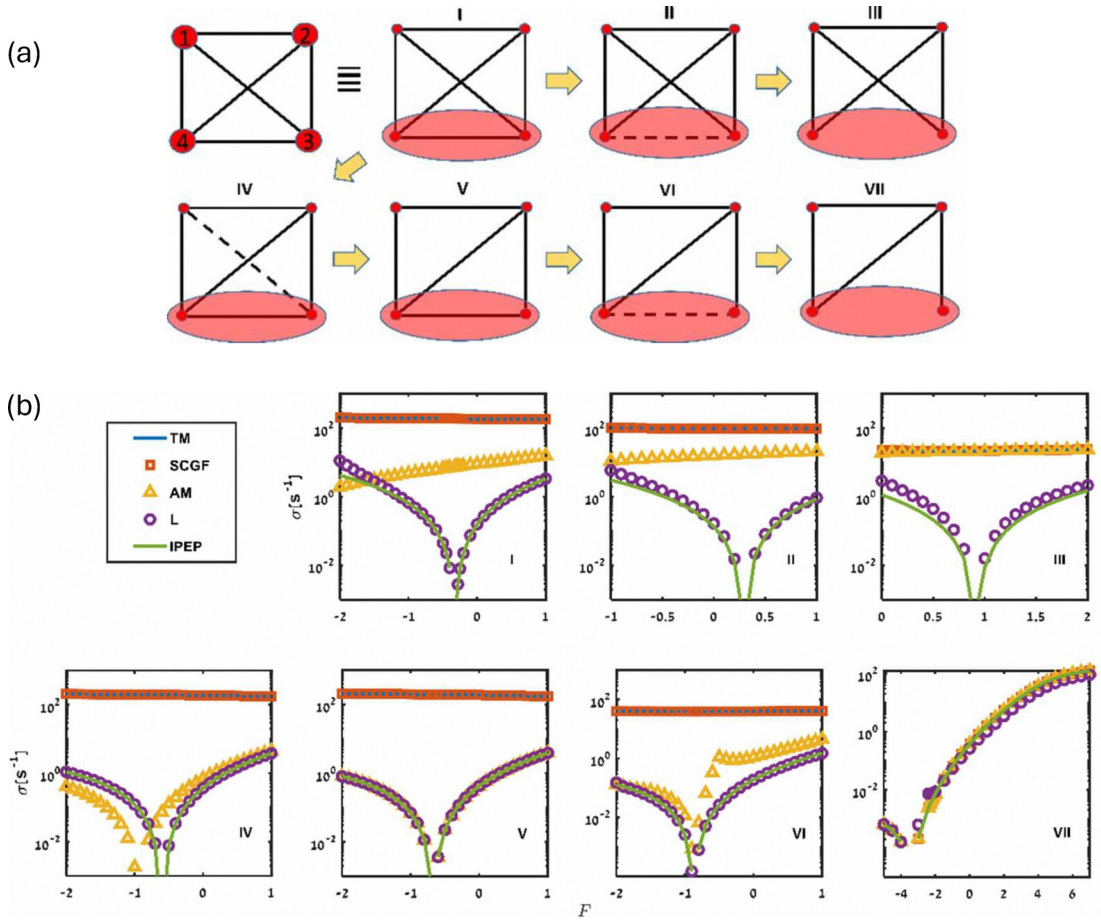


Fig. 8 Different network topologies in which states 1 and 2 are observed, while states 3 and 4 are indistinguishable, are shown. States 3 and 4 are shaded in red and are collectively referred to as the ''hidden state.'' The weights of the 3–4 and 1–3 links are varied, resulting in the different network topologies shown. (b) The total mean (TM) is calculated using a fully resolved network, and the coarse-grained mean EPRs estimated using eqn (65) are compared for each topology. The EPR from the scaled cumulant generating function (SCGF) is obtained after decimating states and applying a Markovian approximation. The EPR obtained from a coarse-grained network by applying lumping – while preserving transition fluxes and steady-state probabilities during the coarse-graining process – is denoted by L. The EPR obtained from the approximated master equation is called AM, and informed partial entropy production (IPEP) is based on transition rates among the observed states and the stalling probability distribution. This figure has been reproduced from ref. 128 with permission from the IOP Publishing, copyright 2026.

when hidden substates are disconnected, reflecting loss of hidden state information. The discrepancy between HS and L varies with stalling force and timescale differences: HS is sensitive to timescale separation between merged states, while L is not. These findings highlighted the interplay between topology, hidden connectivity, and timescale separation in entropy production inference. The framework can be extended to more complex systems with multiple hidden microstates or hidden cycle currents.

In another work, we compared three estimators, the plug-in estimator (σ_{plug} (eqn (34) in Section 5.1), RNEEP estimator, and KLD estimator for different model systems described below. First, we discuss 4-state system, with three observed states among which 2 states are Markovian Fig. 4. In the full-CG case, the KLD estimator σ_{KLD} provides the tightest lower bound on the entropy production rate, while $\sigma_{\text{RNEEP},m}$ converges to the affinity σ_{aff} with increasing sequence length, and the plug-in estimator σ_{plug} shows bias near the stall force. For the semi-CG data, $\sigma_{\text{RNEEP},m}$ becomes comparable to σ_{KLD} for $m \approx 20$, both outperforming σ_{plug}. Thus, KLD dominates in the full-CG setting, whereas in the semi-CG setting, both KLD and RNEEP provide similarly tight bounds given sufficient sequence length. The results have been reproduced with permission and presented in Fig. 9(a).

Next, we discuss EP values comparison from three EPR estimators (σ_{KLD}, $\sigma_{\text{RNEEP},m}$, σ_{plug}) for the molecular motor system (Section 4) and analyzed in Fig. 9(b) under both the full-CG and semi-CG schemes for three values of $\Delta\mu$ and range of F values including the stall force. As we can infer from Fig. 9(b), the semi-CG scheme consistently provides tighter bounds on the entropy production rate (EPR) than the full-CG case, since it exploits additional information from intra-

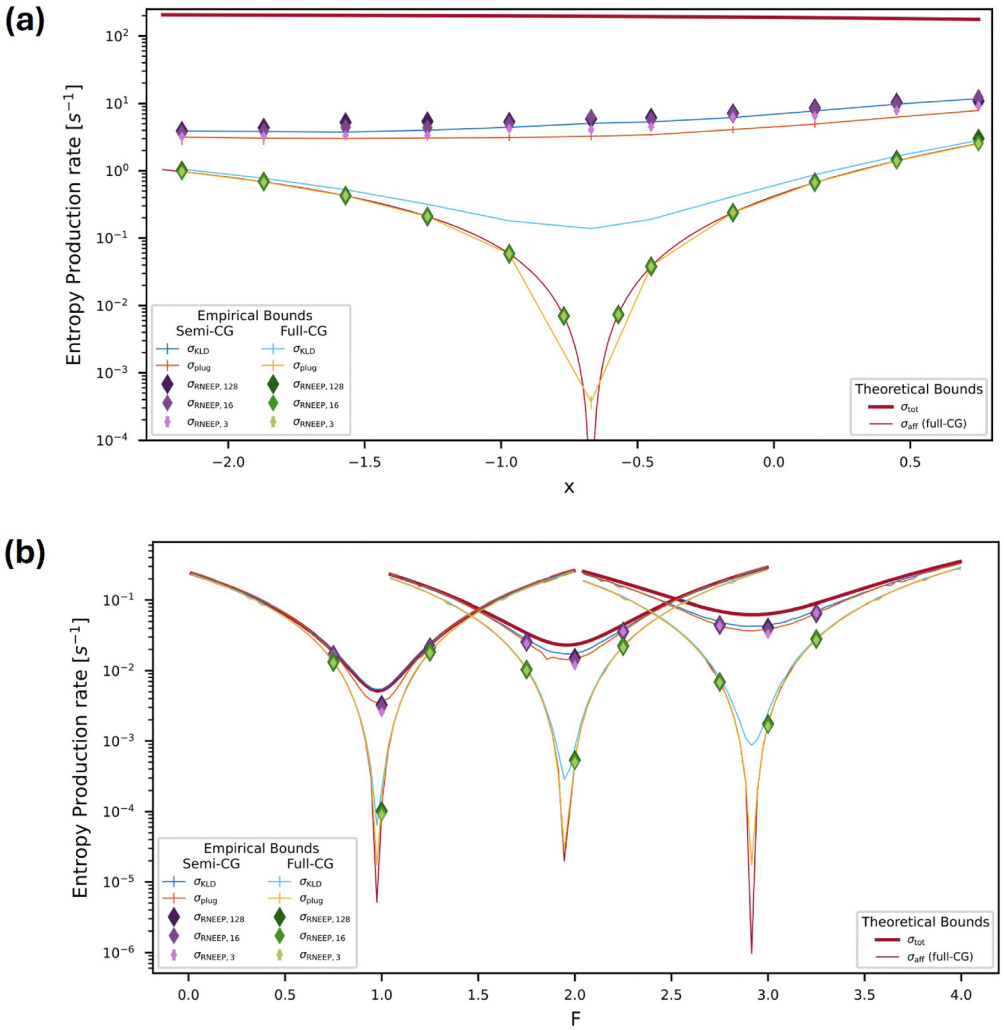


Fig. 9 (a) Entropy production rates for the 4-state system. Total EPR, σ_{tot} (dark red); KLD estimator, σ_{KLD}, for semi-CG (dark blue) and full-CG (light blue) data; plug-in estimator, σ_{plug}, for semi-CG (dark orange) and full-CG (light orange) data; RNEEP estimator, $\sigma_{\text{RNEEP},m}$, for semi-CG (light to dark purple with increasing sequence length m) and for full-CG data (light to dark green with increasing sequence length m); and the affinity contribution, σ_{aff}, for full-CG data (red). (b) Entropy production rates for the molecular motor model: the total EPR, σ_{tot} (dark red); the KLD estimator, σ_{KLD} (dark blue); the plug-in estimator, σ_{plug} (dark orange); and the RNEEP estimator, $\sigma_{\text{RNEEP},m}$ (light to dark purple for increasing sequence length (m)). This figure has been reproduced from ref. 152 with permission from the American Physical Society, copyright 2026.

transitions within macrostates. In the full-CG trajectories, the KLD, plug-in, and RNEEP estimators yield similar results, except near the stalling force, where σ_{KLD} remains the tightest bound. Under the semi-CG scheme, however, the transformed KLD estimator outperforms both the plug-in and RNEEP, producing the sharpest lower bound on EPR. The RNEEP estimator still improves with sequence length m, converging toward the KLD results, while the plug-in estimator eventually falls within the error of $\sigma_{\mathrm{RNEEP},m}$. This difference between the semi-CG and full-CG frameworks highlights the role of hidden transitions in encoding irreversibility, especially evident near the stalling force.

6 Conclusions

Substantial progress over the past two decades has greatly advanced our understanding of thermodynamic irreversibility in systems operating far from equilibrium, clarifying how dissipation, entropy production, and broken time-reversal symmetry are connected.[237] Foundational questions have been explored, including the relationship between information-theoretic entropy production (ITEP) and thermodynamic EPR, how irreversibility can be estimated from partial observations, how coarse-graining modifies entropy-production estimates, and how the parity of coarse-grained observables under time reversal affects the definition of broken time-reversal symmetry. Optimization principles—such as minimal entropy production and optimal transport—have provided unifying structure, and significant effort has been devoted to identifying measurable signatures of non-equilibrium behavior.

Despite this progress, determining whether a system operates at equilibrium or out of equilibrium remains difficult in many experimental contexts. In several scenarios – including one-dimensional driven variables with non-Markovian statistics,[238] Gaussian observables, or multidimensional linear Markovian systems with multiple timescales[239] – equilibrium *versus* nonequilibrium behavior can only be distinguished through invasive perturbations. This highlights an urgent need for reliable, noninvasive approaches to detect time irreversibility directly from experimental time-series data.

Although entropy production and broken time-reversal symmetry remain the primary markers of irreversibility, other indicators such as effective temperature[240] are often invoked. Recent experiments on active fluctuations of an AFM cantilever tip embedded in the mitotic cell cortex reveal a striking decoupling between effective temperature and entropy-production rate,[241] underscoring the limitations of effective temperature as a universal metric. Additional measures, such as response functions and transfer entropy, provide complementary perspectives on temporal asymmetry.[242]

As stochastic thermodynamics interfaces with fields such as active matter, soft materials, machine learning, and field theory, new classes of questions have emerged that go beyond traditional thermodynamic settings. In active matter systems,[243–245] quantifying dissipation in the presence of persistent non-equilibrium fluxes remains a central challenge. In machine learning, thermodynamic inference is increasingly used to analyze learning dynamics, model complexity, and information bottlenecks. Meanwhile, field-theoretic frameworks are being developed to describe fluctuations, symmetry constraints, and coarse-grained entropy production in spatially extended systems. These cross-disciplinary links are reshaping conceptual foundations and prompting new questions about universality, inference, and scalability in complex driven systems.

Parallel theoretical and experimental advances, especially in single-molecule tracking and manipulation, have paved the way for engineering synthetic materials capable of performing functional tasks such as catalysis,[246] directed motion,[247] and autonomous propulsion.[248] While stochastic thermodynamics was originally built on explicit knowledge of underlying system dynamics, recent model-free approaches[249] aim to infer thermodynamic quantities directly from observed trajectories. Such methods promise to connect experiments more directly with theory, especially for synthetic living systems characterized by strongly coupled degrees of freedom, memory effects, and finite-time dynamics. Machine learning-based inference frameworks may play a prominent role in this direction.

Looking forward, major open challenges include establishing model-free theoretical frameworks with predictive power for multivariate driven systems, and exploiting biased-ensemble methods for reverse engineering active materials.[237] Continued integration of theoretical, computational, and experimental perspectives will be essential for uncovering universal principles of non-equilibrium physics and enabling the design of next-generation synthetic systems.

Author contributions

All authors contributed equally to this work.

Conflicts of interest

There are no conflicts to declare.

Data availability

No primary research results, software or code have been included, and no new data were generated or analyzed as part of this review.

Acknowledgements

We regret that, due to space limitations, we are unable to cite all valuable contributions. G. Bisker acknowledges support from the Zuckerman STEM Leadership Program, the European Research Council (ERC) under the NanoNonEq project (Grant No. 101039127), the Air Force Office of Scientific Research (AFOSR) under Award No. FA9550-20-1-0426, the Army Research Office (ARO) under Grant No. W911NF-21-1-0101, the Tel Aviv

University Center for AI and Data Science (TAD), the Israel Science Foundation (Grant No. 196/22), the Ministry of Science, Technology, and Space, Israel (Grant No. 1001818370), the Zimin Institute for Engineering Solutions Advancing Better Lives, the Marian Gertner Institute for Medical Nanosystems at Tel Aviv University, and the Naomi Prawer Kadar Foundation.

References

1 L. Onsager, *Phys. Rev.*, 1931, **37**, 405–426.
2 L. Onsager, *Phys. Rev.*, 1931, **38**, 2265–2279.
3 T. de Donder, F. van den Dungen and G. van Lerberghe, *Leçons de thermodynamique et de chimie physique*, Gauthier-Villars et cie., 1920.
4 I. Prigogine and P. Van Rysselberghe, *J. Electrochem. Soc.*, 1963, **110**, 97C.
5 I. Prigogine, *Science*, 1978, **201**, 777–785.
6 I. Prigogine, *Capillarity Today*, Berlin, Heidelberg, 1991, pp. 1–13.
7 S. De Groot and P. Mazur, *Non-Equilibrium Thermodynamics*, Dover Publications, 2013.
8 D. Reguera, J. Rubi and J. Vilar, *The mesoscopic dynamics of thermodynamic systems*, 2005.
9 M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.*, 1993, **65**, 851.
10 W. A. Bickmore, *Annu. Rev. Genomics Hum. Genet.*, 2013, **14**, 67–84.
11 K. Sekimoto, *Prog. Theor. Phys. Suppl.*, 1998, **130**, 17–27.
12 L. Peliti and S. Pigolotti, *Stochastic Thermodynamics: An Introduction*, Princeton University Press, 2021.
13 R. F. Streater, *Statistical dynamics: a stochastic approach to nonequilibrium thermodynamics*, World Scientific Publishing Company, 2009.
14 Y. Oono and M. Paniconi, *Prog. Theor. Phys. Suppl.*, 1998, **130**, 29–44.
15 S.-i Sasa and H. Tasaki, *J. Stat. Phys.*, 2006, **125**, 125–224.
16 T. Hatano and S.-i Sasa, *Phys. Rev. Lett.*, 2001, **86**, 3463–3466.
17 J. L. Lebowitz and H. Spohn, *J. Stat. Phys.*, 1999, **95**, 333–365.
18 G. E. Crooks, *Phys. Rev. E:Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1999, **60**, 2721.
19 D. J. Evans, D. J. Searles and E. Mittag, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2001, **63**, 051105.
20 D. J. Evans and D. J. Searles, *Adv. Phys.*, 2002, **51**, 1529–1585.
21 E. M. Sevick, R. Prabhakar, S. R. Williams and D. J. Searles, *Annu. Rev. Phys. Chem.*, 2008, **59**, 603–633.
22 C. Jarzynski, *Time: Poincaré Seminar*, 2010, vol. 2012, pp. 145–172.
23 U. Seifert, *Rep. Prog. Phys.*, 2012, **75**, 126001.
24 C. Van den Broeck, *Physics of Complex Colloids*, IOS Press, 2013, pp. 155–193.
25 E. Aurell, K. Gawdzki, C. Mejía-Monasterio, R. Mohayaee and P. Muratore-Ginanneschi, *J. Stat. Phys.*, 2012, **147**, 487–505.
26 S. Ciliberto, *Phys. Rev. X*, 2017, **7**, 021051.
27 F. Ritort, *Nonequilibrium Fluctuations in Small Systems:From Physics to Biology*, John Wiley Sons, Ltd, 2007, ch. 2, pp. 31–123.
28 J. Bechhoefer, S. Ciliberto, S. Pigolotti and E. Roldán, *J. Stat. Mech.:Theory Exp.*, 2020, **2020**, 064001.
29 F. Ritort, *C. R. Phys.*, 2007, **8**, 528–539.
30 F. Ritort, in *Fluctuation Theorems, Brownian Motors and Thermodynamics of Small Systems*, ed. B. Chakraborty, Springer US, New York, NY, 2022, pp. 477–494.
31 J. M. Horowitz and T. R. Gingrich, *Nat. Phys.*, 2020, **16**, 15–20.
32 A. C. Barato and U. Seifert, *Phys. Rev. Lett.*, 2015, **114**, 158101.
33 T. R. Gingrich, J. M. Horowitz, N. Perunov and J. L. England, *Phys. Rev. Lett.*, 2016, **116**, 120601.
34 G. Falasco, M. Esposito and J.-C. Delvenne, *New J. Phys.*, 2020, **22**, 053046.
35 A. Dechant, *J. Phys. A: Math. Theor.*, 2018, **52**, 035001.
36 T. Van Vu and Y. Hasegawa, *Phys. Rev. E*, 2019, **100**, 032130.
37 S. Saryal, S. Mohanta and B. K. Agarwalla, *Phys. Rev. E*, 2022, **105**, 024129.
38 I. Di Terlizzi and M. Baiesi, *J. Phys. A: Math. Theor.*, 2020, **53**, 474002.
39 G. Falasco and M. Esposito, *Phys. Rev. Lett.*, 2020, **125**, 120604.
40 R. Rao and L. Peliti, *J. Stat. Mech.:Theory Exp.*, 2015, **2015**, P06001.
41 P. Ao, *Commun. Theor. Phys.*, 2008, **49**, 1073.
42 H. Qian and D. A. Beard, *Biophys. Chem.*, 2005, **114**, 213–220.
43 A. Kolchinsky, *J. Chem. Phys.*, 2024, **161**, 124101.
44 J. L. England, *J. Chem. Phys.*, 2013, **139**, 121923.
45 P. J. Foster, J. Bae, B. Lemma, J. Zheng, W. Ireland, P. Chandrakar, R. Boros, Z. Dogic, D. J. Needleman and J. J. Vlassak, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, e2207662120.
46 H. Qian, *The dynamics of biological systems*, Springer, 2019, pp. 149–188.
47 P. Sartori, L. Granger, C. F. Lee and J. M. Horowitz, *PLoS Comput. Biol.*, 2014, **10**, e1003974.
48 D. Sekizawa, S. Ito and M. Oizumi, *Phys. Rev. X*, 2024, **14**, 041003.
49 C. W. Lynn, E. J. Cornblath, L. Papadopoulos, M. A. Bertolero and D. S. Bassett, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2109889118.
50 J. M. Parrondo, J. M. Horowitz and T. Sagawa, *Nat. Phys.*, 2015, **11**, 131–139.
51 T. Sagawa, *Prog. Theor. Phys.*, 2012, **127**, 1–56.
52 A. Barato and U. Seifert, *Phys. Rev. Lett.*, 2014, **112**, 090601.
53 G. Micali and R. G. Endres, *Curr. Opin. Microbiol.*, 2016, **30**, 8–15.
54 A. B. Boyd, A. Patra, C. Jarzynski and J. P. Crutchfield, *J. Stat. Phys.*, 2022, **187**, 17.
55 C. E. Shannon and W. Weaver, *The mathematical theory of communication*, University of Illinois Press, 1998.

56 S. Ito, *Phys. Rev. Lett.*, 2018, **121**, 030605.
57 A.-J. Guel-Cortez and E.-J. Kim, *J. Stat. Mech.:Theory Exp.*, 2023, **2023**, 033204.
58 V. Serreli, C.-F. Lee, E. R. Kay and D. A. Leigh, *Nature*, 2007, **445**, 523–527.
59 M. Alvarez-Pérez, S. M. Goldup, D. A. Leigh and A. M. Slawin, *J. Am. Chem. Soc.*, 2008, **130**, 1836–1838.
60 T. E. Ouldridge and R. Mukherjee, *Nat. Nanotechnol.*, 2025, 1–2.
61 L. Binks, S. Borsley, T. R. Gingrich, D. A. Leigh, E. Penocchio and B. M. Roberts, *Chem*, 2023, **9**, 2902–2917.
62 D. H. Wolpert, J. Korbel, C. W. Lynn, F. Tasnim, J. A. Grochow, G. Kardes, J. B. Aimone, V. Balasubramanian, E. De Giuli and D. Doty, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2321112121.
63 G. Manzano, G. Kardes, É. Roldán and D. H. Wolpert, *Phys. Rev. X*, 2024, **14**, 021026.
64 P. Strasberg, J. Cerrillo, G. Schaller and T. Brandes, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2015, **92**, 042104.
65 A. Kolchinsky and D. H. Wolpert, *Phys. Rev. Res.*, 2020, **2**, 033312.
66 E. Aurell, C. Mejía-Monasterio and P. Muratore-Ginanneschi, *Phys. Rev. Lett.*, 2011, **106**, 250601.
67 S. Blaber and D. A. Sivak, *J. Phys. Commun.*, 2023, **7**, 033001.
68 P. Reimann, *Phys. Rep.*, 2002, **361**, 57–265.
69 G. Bisker and J. L. England, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E10531–E10538.
70 P. Hänggi and F. Marchesoni, *Rev. Mod. Phys.*, 2009, **81**, 387–442.
71 S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan and A. L. Nussbaumer, *Chem. Rev.*, 2015, **115**, 10081–10206.
72 D. Keller, D. Swigon and C. Bustamante, *Biophys. J.*, 2003, **84**, 733–738.
73 M. Esposito, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 041125.
74 D. Seiferth, P. Sollich and S. Klumpp, *Phys. Rev. E*, 2020, **102**, 062149.
75 R. García-García, S. Lahiri and D. Lacoste, *Phys. Rev. E*, 2016, **93**, 032103.
76 G. Teza, Out of equilibrium dynamics: from an entropy of the growth to the growth of entropy production, PhD thesis, University of Padova, 2020.
77 B. Altaner and J. Vollmer, *Phys. Rev. Lett.*, 2012, **108**, 228101.
78 G. Teza and A. L. Stella, *Phys. Rev. Lett.*, 2020, **125**, 110601.
79 O. A. Igoshin, A. B. Kolomeisky and D. E. Makarov, *J. Phys. Chem. Lett.*, 2025, **16**, 1229–1237.
80 D. Gupta and S. Sabhapandit, *J. Stat. Mech.:Theory Exp.*, 2020, **2020**, 013204.
81 U. Seifert, *Stochastic Thermodynamics*, Cambridge University Press, 2025.
82 L. Boltzmann, *Sitzungsberichte Akad. Wiss., Vienna, II, 66: 275–370; English transl.: Brush, SG (1966) Kinetic Theory: Vol. 2 Irreversible Processes*, 1872.
83 F. P. Kelly, *Reversibility and stochastic networks*, Cambridge University Press, 2011.
84 R. K. Zia and B. Schmittmann, *J. Stat. Mech.:Theory Exp.*, 2007, **2007**, P07012.
85 C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt and F. C. MacKintosh, *Science*, 2016, **352**, 604–607.
86 A. Ghosal and G. Bisker, *Phys. Chem. Chem. Phys.*, 2022, **24**, 24021–24031.
87 H. Qian and M. Qian, *Phys. Rev. Lett.*, 2000, **84**, 2271.
88 H. Qian, S. Saffarian and E. L. Elson, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 10376–10381.
89 C. Jia and Y. Chen, *J. Phys. A: Math. Theor.*, 2015, **48**, 205001.
90 H. Turlier, D. A. Fedosov, B. Audoly, T. Auth, N. S. Gov, C. Sykes, J.-F. Joanny, G. Gompper and T. Betz, *Nat. Phys.*, 2016, **12**, 513–519.
91 P. Martin, A. Hudspeth and F. Jülicher, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 14380–14385.
92 J. Gladrow, N. Fakhri, F. C. MacKintosh, C. Schmidt and C. P. Broedersz, *Phys. Rev. Lett.*, 2016, **116**, 248301.
93 S. Ma, R. Zhang and J. Yuan, *Biophys. J.*, 2022, **121**, 2345–2352.
94 J. Holehouse, *arXiv*, 2024, preprint, arXiv:2405.12897, DOI: 10.48550/arXiv.2405.12897.
95 F. S. Gnesotto, F. Mura, J. Gladrow and C. P. Broedersz, *Rep. Prog. Phys.*, 2018, **81**, 066601.
96 C. F. Schmidt, J. Gladrow, N. Fakhri, F. C. Mackintosh and C. Broedersz, *APS March Meeting Abstracts*, 2017, pp. A4-007.
97 S. Liang and S. Pigolotti, *Phys. Rev. E*, 2023, **108**, L062101.
98 I. Golding, J. Paulsson, S. M. Zawilski and E. C. Cox, *Cell*, 2005, **123**, 1025–1036.
99 I. A. Martínez, G. Bisker, J. M. Horowitz and J. M. Parrondo, *Nat. Commun.*, 2019, **10**, 3542.
100 D. Hartich and A. Godec, *Nat. Commun.*, 2024, **15**, 8678.
101 G. Bisker, I. A. Martínez, J. M. Horowitz and J. M. Parrondo, *Nat. Commun.*, 2024, **15**, 8679.
102 R. Kubo, *J. Phys. Soc. Jpn.*, 1957, **12**, 570–586.
103 J. Prost, J.-F. Joanny and J. M. Parrondo, *Phys. Rev. Lett.*, 2009, **103**, 090601.
104 U. M. B. Marconi, A. Puglisi, L. Rondoni and A. Vulpiani, *Phys. Rep.*, 2008, **461**, 111–195.
105 D. Ruelle, *Phys. Lett. A*, 1998, **245**, 220–224.
106 R. Kubo, *Rep. Prog. Phys.*, 1966, **29**, 255.
107 A. Crisanti and F. Ritort, *J. Phys. A: Math. Gen.*, 2003, **36**, R181.
108 T. S. Grigera and N. Israeloff, *Phys. Rev. Lett.*, 1999, **83**, 5038.
109 K. Hayashi and M. Takano, *Biophys. J.*, 2007, **93**, 895–901.
110 A. Pérez-Madrid, D. Reguera and J. Rub, *Phys. A*, 2003, **329**, 357–364.
111 D. V. Averin and J. P. Pekola, *Phys. Rev. Lett.*, 2010, **104**, 220601.
112 R. Mauri and D. Leporini, *Europhys. Lett.*, 2006, **76**, 1022.
113 T. Harada and S.-i Sasa, *Phys. Rev. Lett.*, 2005, **95**, 130602.
114 T. Harada and S.-i Sasa, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2006, **73**, 026131.

115 S.-W. Wang, *Phys. Rev. E*, 2018, **97**, 052125.
116 Y.-D. Chen, *Adv. Chem. Phys.*, 1978, **37**, 67–97.
117 J. Wei and C. D. Prater, *Advances in catalysis*, Elsevier, 1962, vol. 13, pp. 203–392.
118 C. P. Amann, T. Schmiedl and U. Seifert, *J. Chem. Phys.*, 2010, **132**, 041102.
119 B. Wu and C. Jia, *Phys. Rev. Lett.*, 2025, **134**, 087103.
120 F. Mori, S. N. Majumdar and G. Schehr, *Europhys. Lett.*, 2021, **135**, 30003.
121 S. Thapa, D. Zaretzky, R. Vatash, G. Gradziuk, C. Broedersz, Y. Shokef and Y. Roichman, *SciPost Phys.*, 2024, **17**, 096.
122 T. M. Muenker, G. Knotz, M. Krüger and T. Betz, *Nat. Mater.*, 2024, **23**, 1283–1291.
123 G. Knotz and M. Krüger, *Phys. Rev. E*, 2024, **110**, 044137.
124 N. Shiraishi, *Fundamental Theories of Physics*, Springer, Singapore, 2023.
125 D. T. Limmer, *Statistical mechanics and stochastic thermodynamics: A textbook on modern approaches in and out of equilibrium*, Oxford University Press, 2024.
126 J. Schnakenberg, *Rev. Mod. Phys.*, 1976, **48**, 571.
127 U. Seifert, *Phys. Rev. Lett.*, 2005, **95**, 040602.
128 A. Ghosal and G. Bisker, *J. Phys. D: Appl. Phys.*, 2023, **56**, 254001.
129 G. Bisker, M. Polettini, T. R. Gingrich and J. M. Horowitz, *J. Stat. Mech.:Theory Exp.*, 2017, **2017**, 093210.
130 I. Di Terlizzi and M. Baiesi, *J. Phys. A:Math. Theor.*, 2018, **52**, 02LT03.
131 A. Ghosal, *J. Phys. Chem. B*, 2021, **125**, 1760–1767.
132 S. Ciliberto, A. Imparato, A. Naert and M. Tanase, *J. Stat. Mech.:Theory Exp.*, 2013, **2013**, P12014.
133 M. Biroli, M. Kulkarni, S. N. Majumdar and G. Schehr, *Phys. Rev. E*, 2024, **109**, L032106.
134 R. Garcia-Millan and G. Pruessner, *J. Stat. Mech.:Theory Exp.*, 2021, **2021**, 063203.
135 D. J. Mai, C. Brockman and C. M. Schroeder, *Soft Matter*, 2012, **8**, 10560–10572.
136 T. Tomé and M. J. de Oliveira, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2010, **82**, 021120.
137 G. Nicolis and Y. De Decker, *Entropy*, 2017, **19**, 434.
138 D. Chaudhuri, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2014, **90**, 022131.
139 A. Ghosal and B. Cherayil, *J. Stat. Mech.:Theory Exp.*, 2016, **2016**, 043201.
140 K. Proesmans, Y. Dreher, M. Gavrilov, J. Bechhoefer and C. Van den Broeck, *Phys. Rev. X*, 2016, **6**, 041010.
141 G. Wang, J. C. Reid, D. Carberry, D. Williams, E. M. Sevick and D. J. Evans, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 046142.
142 D. Frydel, *Phys. Rev. E*, 2023, **107**, 014604.
143 D. Carberry, J. C. Reid, G. Wang, E. M. Sevick, D. J. Searles and D. J. Evans, *Phys. Rev. Lett.*, 2004, **92**, 140601.
144 J. Gomez-Solano, A. Petrosyan and S. Ciliberto, *Phys. Rev. Lett.*, 2011, **106**, 200602.
145 A. Saha, S. Lahiri and A. Jayannavar, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2009, **80**, 011117.
146 A. Ghosal and B. J. Cherayil, *Eur. Phys. J. B*, 2019, **92**, 243.
147 R. Van Zon, S. Ciliberto and E. Cohen, *Phys. Rev. Lett.*, 2004, **92**, 130601.
148 N. Garnier and S. Ciliberto, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 060101.
149 S. Joubaud, N. Garnier and S. Ciliberto, *Europhys. Lett.*, 2008, **82**, 30007.
150 E. Nitzan, A. Ghosal and G. Bisker, *Phys. Rev. Res.*, 2023, **5**, 043251.
151 M. Polettini and M. Esposito, *Phys. Rev. Lett.*, 2017, **119**, 240601.
152 U. Kapustin, A. Ghosal and G. Bisker, *Phys. Rev. Res.*, 2024, **6**, 023039.
153 J. Van der Meer, B. Ertel and U. Seifert, *Phys. Rev. X*, 2022, **12**, 031025.
154 P. E. Harunari, A. Dutta, M. Polettini and É. Roldán, *Phys. Rev. X*, 2022, **12**, 041026.
155 N. Shiraishi and T. Sagawa, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2015, **91**, 012130.
156 J. Ehrich, *J. Stat. Mech.:Theory Exp.*, 2021, **2021**, 083214.
157 D. M. Busiello, J. Hidalgo and A. Maritan, *New J. Phys.*, 2019, **21**, 073004.
158 D. M. Busiello and A. Maritan, *J. Stat. Mech.:Theory Exp.*, 2019, **2019**, 104013.
159 K. Blom, K. Song, E. Vouga, A. Godec and D. E. Makarov, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2318333121.
160 G. Hummer and A. Szabo, *J. Phys. Chem. B*, 2015, **119**, 9029–9037.
161 D. Hartich and A. Godec, *Phys. Rev. Res.*, 2023, **5**, L032017.
162 S. Pigolotti and A. Vulpiani, *J. Chem. Phys.*, 2008, **128**, 154114.
163 A. Puglisi, S. Pigolotti, L. Rondoni and A. Vulpiani, *J. Stat. Mech.:Theory Exp.*, 2010, **2010**, P05015.
164 P. Bilotto, L. Caprini and A. Vulpiani, *Phys. Rev. E*, 2021, **104**, 024140.
165 M. Esposito and J. M. Parrondo, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2015, **91**, 052114.
166 D. M. Busiello, M. Ciarchi and I. Di Terlizzi, *Phys. Rev. Res.*, 2024, **6**, 013190.
167 Q. Yu and P. E. Harunari, *J. Stat. Mech.:Theory Exp.*, 2024, **2024**, 103201.
168 T. H. Tan, G. A. Watson, Y.-C. Chao, J. Li, T. R. Gingrich, J. M. Horowitz and N. Fakhri, *arXiv*, 2021, preprint, arXiv:2107.05701, DOI: 10.48550/arXiv.2107.05701.
169 F. A. Cisneros, N. Fakhri and J. M. Horowitz, *J. Stat. Mech.:Theory Exp.*, 2023, **2023**, 073201.
170 J. H. Fritz, B. Ertel and U. Seifert, *Phys. Rev. E*, 2025, **111**, 044106.
171 C. Maes, *Prog. Math. Phys.*, 2004, **38**, 145.
172 M. Baiesi and C. Maes, *J. Phys. Commun.*, 2018, **2**, 045017.
173 G. Verley, M. Esposito, T. Willaert and C. Van den Broeck, *Nat. Commun.*, 2014, **5**, 4721.
174 G. T. Landi and M. Paternostro, *Rev. Mod. Phys.*, 2021, **93**, 035008.
175 I. A. Martínez, É. Roldán, L. Dinis, D. Petrov, J. M. Parrondo and R. A. Rica, *Nat. Phys.*, 2016, **12**, 67–70.

176 S. K. Manikandan, L. Dabelow, R. Eichhorn and S. Krishnamurthy, *Phys. Rev. Lett.*, 2019, **122**, 140601.
177 A. Parmeggiani, F. Jülicher, A. Ajdari and J. Prost, *Phys. Rev. E:Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1999, **60**, 2127.
178 G. De Palo and R. G. Endres, *PLoS Comput. Biol.*, 2013, **9**, e1003300.
179 W. Buijsman and M. Sheinman, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2014, **89**, 022712.
180 T. M. Cover, *Elements of information theory*, John Wiley & Sons, 1999.
181 J. M. Parrondo, C. Van den Broeck and R. Kawai, *New J. Phys.*, 2009, **11**, 073008.
182 É. Roldán and J. M. Parrondo, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 031129.
183 R. Kawai, J. M. Parrondo and C. V. den Broeck, *Phys. Rev. Lett.*, 2007, **98**, 080602.
184 B. Gaveau, L. Granger, M. Moreau and L. Schulman, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2014, **89**, 032107.
185 P. E. Harunari, *Phys. Rev. E*, 2024, **110**, 024122.
186 É. Roldán and É. Roldán, *Irreversibility and Dissipation in Microscopic Systems*, 2014, pp. 37–59.
187 Q. Wang, S. R. Kulkarni and S. Verdú, *IEEE Trans. Inf. Theory*, 2005, **51**, 3064–3074.
188 J. Ziv and N. Merhav, *IEEE Trans. Inf. Theory*, 1993, **39**, 1270–1279.
189 É. Roldán and J. M. Parrondo, *Phys. Rev. Lett.*, 2010, **105**, 150607.
190 J. Horowitz and C. Jarzynski, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2009, **79**, 021106.
191 C. Maes and K. Netočný, *J. Stat. Phys.*, 2003, **110**, 269–310.
192 É. Roldán, J. Barral, P. Martin, J. M. Parrondo and F. Jülicher, *New J. Phys.*, 2021, **23**, 083013.
193 P. Martin, D. Bozovic, Y. Choe and A. Hudspeth, *J. Neurosci.*, 2003, **23**, 4533–4548.
194 J.-Y. Tinevez, F. Jülicher and P. Martin, *Biophys. J.*, 2007, **93**, 4053–4067.
195 A. Ghosal and J. R. Green, *Phys. Rev. Res.*, 2025, **7**, L012078.
196 D. J. Skinner and J. Dunkel, *Phys. Rev. Lett.*, 2021, **127**, 198101.
197 J. van der Meer, J. Degünther and U. Seifert, *Phys. Rev. Lett.*, 2023, **130**, 257101.
198 A. Puglisi, *Europhys. Lett.*, 2025, **150**, 67001.
199 K. Proesmans, *Commun. Phys.*, 2023, **6**, 226.
200 K. J. Ray, A. B. Boyd, G. Guarnieri and J. P. Crutchfield, *Phys. Rev. E*, 2023, **108**, 054126.
201 K. Macieszczak, K. Brandner and J. P. Garrahan, *Phys. Rev. Lett.*, 2018, **121**, 130601.
202 A. Dechant and S.-i Sasa, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 6430–6436.
203 P. Pietzonka, A. C. Barato and U. Seifert, *Phys. Rev. E*, 2016, **93**, 052145.
204 P. Pietzonka and U. Seifert, *Phys. Rev. Lett.*, 2018, **120**, 190602.
205 J. M. Horowitz and T. R. Gingrich, *Phys. Rev. E*, 2017, **96**, 020103.
206 K. Proesmans and C. Van den Broeck, *Europhys. Lett.*, 2017, **119**, 20001.
207 A. Pal, S. Reuveni and S. Rahav, *Phys. Rev. Res.*, 2021, **3**, 013273.
208 N. Shiraishi, *J. Stat. Phys.*, 2021, **185**, 19.
209 P. P. Potts and P. Samuelsson, *Phys. Rev. E*, 2019, **100**, 052137.
210 K. Liu, Z. Gong and M. Ueda, *Phys. Rev. Lett.*, 2020, **125**, 140602.
211 V. T. Vo, T. Van Vu and Y. Hasegawa, *J. Phys. A: Math. Theor.*, 2022, **55**, 405004.
212 T. Van Vu and Y. Hasegawa, *et al.*, *J. Phys. A: Math. Theor.*, 2022, **55**, 405004.
213 T. R. Gingrich, G. M. Rotskoff and J. M. Horowitz, *J. Phys. A: Math. Theor.*, 2017, **50**, 184004.
214 J. Li, J. M. Horowitz, T. R. Gingrich and N. Fakhri, *Nat. Commun.*, 2019, **10**, 1666.
215 S. K. Manikandan, D. Gupta and S. Krishnamurthy, *Phys. Rev. Lett.*, 2020, **124**, 120603.
216 S. K. Manikandan, T. Ghosh, T. Mandal, A. Biswas, B. Sinha and D. Mitra, *Phys. Rev. Res.*, 2024, **6**, 023310.
217 T. R. Gingrich and J. M. Horowitz, *Phys. Rev. Lett.*, 2017, **119**, 170601.
218 T. Wampler and A. C. Barato, *J. Phys. A: Math. Theor.*, 2021, **55**, 014002.
219 D. Hartich and A. Godec, *Phys. Rev. Lett.*, 2021, **127**, 080601.
220 D. J. Skinner and J. Dunkel, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2024300118.
221 Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly and R. Mart, *INFORMS J. Comput.*, 2007, **19**, 328–340.
222 T. Van Vu and K. Saito, *Phys. Rev. X*, 2023, **13**, 011013.
223 S. Ito and A. Dechant, *Phys. Rev. X*, 2020, **10**, 021056.
224 N. Shiraishi, K. Funo and K. Saito, *Phys. Rev. Lett.*, 2018, **121**, 070601.
225 S. B. Nicholson, L. P. García-Pintos, A. del Campo and J. R. Green, *Nat. Phys.*, 2020, **16**, 1211–1215.
226 B. R. Frieden, *Science from Fisher information: a unification*, Cambridge University Press, 2004.
227 S. B. Nicholson and J. R. Green, *arXiv*, 2021, preprint, arXiv:2105.01588, DOI: 10.48550/arXiv.2105.01588.
228 A. Imparato, L. Peliti, G. Pesce, G. Rusciano and A. Sasso, *Phys. Rev. E:Stat., Nonlinear, Soft Matter Phys.*, 2007, **76**, 050101.
229 D.-K. Kim, Y. Bae, S. Lee and H. Jeong, *Phys. Rev. Lett.*, 2020, **125**, 140604.
230 F. Mura, G. Gradziuk and C. P. Broedersz, *Phys. Rev. Lett.*, 2018, **121**, 038002.
231 D.-K. Kim, S. Lee and H. Jeong, *Phys. Rev. Res.*, 2022, **4**, 023051.
232 I. Goodfellow, Y. Bengio, A. Courville and Y. Bengio, *Deep learning*, MIT Press, Cambridge, 2016, vol. 1.
233 Y. LeCun, Y. Bengio and G. Hinton, *Nature*, 2015, **521**, 436–444.
234 A. Nir, E. Sela, R. Beck and Y. Bar-Sinai, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 30234–30240.
235 I. Di Terlizzi, M. Gironella, D. Herraez-Aguilar, T. Betz, F. Monroy, M. Baiesi and F. Ritort, *Science*, 2024, **383**, 971–976.

236 I. Di Terlizzi, M. Baiesi and F. Ritort, *New J. Phys.*, 2024, **26**, 063013.
237 É. Fodor, R. L. Jack and M. E. Cates, *Annu. Rev. Condens. Matter Phys.*, 2022, **13**, 215–238.
238 R. R. Netz, *Phys. Rev. E*, 2025, **112**, 014132.
239 D. Lucente, A. Baldassarri, A. Puglisi, A. Vulpiani and M. Viale, *Phys. Rev. Res.*, 2022, **4**, 043103.
240 L. F. Cugliandolo, *J. Phys. A: Math. Theor.*, 2011, **44**, 483001.
241 N. Narinder and E. Fischer-Friedrich, *arXiv*, 2025, preprint, arXiv:2503.17016, DOI: 10.48550/arXiv.2503.17016.
242 D. Lucente, M. Baldovin, F. Cecconi, M. Cencini, N. Cocciaglia, A. Puglisi, M. Viale and A. Vulpiani, *New J. Phys.*, 2025, **27**, 041201.
243 S. Ramaswamy, *J. Stat. Mech.:Theory Exp.*, 2017, **2017**, 054002.
244 L. Dabelow, S. Bo and R. Eichhorn, *Phys. Rev. X*, 2019, **9**, 021009.
245 P. Pietzonka, É. Fodor, C. Lohrmann, M. E. Cates and U. Seifert, *Phys. Rev. X*, 2019, **9**, 041032.
246 G. De Bo, M. A. Gall, S. Kuschel, J. De Winter, P. Gerbaux and D. A. Leigh, *Nat. Nanotechnol.*, 2018, **13**, 381–385.
247 I. Y. Loh, J. Cheng, S. R. Tee, A. Efremov and Z. Wang, *ACS Nano*, 2014, **8**, 10293–10304.
248 I. Ortiz-Rivera, M. Mathesh and D. A. Wilson, *Acc. Chem. Res.*, 2018, **51**, 1891–1900.
249 T. GrandPre, G. Teza and W. Bialek, *arXiv*, 2024, preprint, arXiv:2412.19772, DOI: 10.48550/arXiv.2412.19772.

